

Validation documentation
Hepart raw materials

HiperScan GmbH

March 22, 2021

Contents

Introduction	7
Context of this document	7
Criteria for the inclusion of substances	8
Validation concept	9
Model creation procedure and validation runs	9
Particularities of individual substance classes	12
Significance of testing with <i>Apo-Ident</i>	13
Conclusion	13
Explanation of terminology	14
Summary	19
Validation samples	19
Validation results	19
Validation reports	21
Acacia gum (encapsia) / acacia gum (R001/R0141)	21
a-lipoic acid (R0003)	25
Apple powder (R0005)	29
Artichoke extract (R0208)	33
Ascorbic acid (R0007)	39
Beta carotene 10% (R0237)	43
Beta carotene 20% (R0008)	47
beta-glucan 28% (oat fibre) (R0334)	51
Betaine HCl (R0211)	55
Biotin (R0010)	59
Black cumin extract (R0193)	63
Calcium caseinate (Lactonate RCE) (R0239)	67
Calcium D panthothenate (R0012)	71
Calcium lactate pentahydrate (R0013)	75
Calmegh extract (R0201)	79
Chaga extract (R0016)	83
Chili extract (R0198)	87
Cholecalciferol (vitamin D3) (R0018) / Menaquinone-7 0.2% (R0134)	91
Choline chloride (R0019)	97
Citric acid monohydrate (R0143)	103
Citric acid powder (R0258)	107
Citrus bioflavonoids 60% (R0021)	111
Coenzyme Q10 (R0022)	115
Coriolus extract (R0207)	119
Creatine monohydrate (R0048)	123
Curcumin 95% extract (R0023)	127
Cyanocobalamin (R0024) / Methylcobalamin (R0190)	131
D-alpha tocopheryl acetate 50% dry CWS (R0081)	135
D-galactose (R0135)	139
DL-methionine (R0209)	143
D-ribose (R0137)	147
Galactomannans (R0035/R0094) / Potato starch, pregelatinized (R0047)	153
Garcinia Cambogia (R0028)	159
Ginger extract (R0199)	165
Ginkgo extract (R0093)	169
Ginseng extract (R0029) / Lutein 20% (R0052) / Lycopene 5% extract (R0053)	173
Glucosamine sulfate 2 KCl (R0030)	179
Glutathione (reduced form) (R0032)	183
Glycine (R0033)	187
Green tea extract (R0034)	191

Guarana extract (R0289)	195
Hydrangea extract (R0203)	199
Hyssop extract (R0205)	203
Inositol (myo-) (R0039)	207
Inositol hexanicotinate (R0040)	213
Instant yeast (dry yeast) (R0112)	219
Inulin (R0041)	223
Isoflavones 10% (R0042)	229
Katuka herb extract (R0210)	233
Konjac glucomannan 80% (R0293)	237
L-5-hydroxytryptophan (R0050)	241
L-arginine base (R0006)	245
L-carnitine-L-tartrate (R0014)	249
L-carnosine (R0015)	253
L-glutamine (R0031)	257
L-histidine base (R0037)	261
L-isoleucine (R0043)	265
L-leucine (R0051)	271
L-lysine HCl (R0054)	277
L-ornithine HCl (R0066)	281
L-phenylalanine (R0067)	285
L-proline (R0070)	289
L-threonine (R0080)	293
L-tryptophan (R0085)	297
L-tyrosine (R0086)	303
Magnesium citrate nonahydrate (R0056)	307
Magnesium glycerophosphate (R0114)	311
Magnesium stearate (R0140)	315
Marshmallow root extract (R0204) / Gentian extract (R0200) / Juniper extract (R0192)	319
Methionine (R0059)	323
Methylsulfonylmethane (R0060)	329
N-acetyl-L-cysteine (R0002)	333
Nicotinamide (R0064)	337
Olive leaf extract (R0202)	343
OPC (grape seed) extract (R0065)	347
Oregano extract (R0206)	351
Osha root extract (R0220)	355
p-aminobenzoic acid (R0004)	359
Pea protein (R0092)	363
Piperine 95% extract (R0069)	367
Progesterone, micronized (R0108)	371
Pumpkin seed protein (R0306)	375
Pyridoxal-5-phosphate (R0195)	379
Pyridoxine hydrochloride (vitamin B6) (R0071)	383
Resveratrol (R0339)	387
Retinyl palmitate (vitamin A palmitate) (R0073)	391
Rhodiola rosea extract 3% (R0159)	395
Riboflavin (vit. B2) (R0074)	399
Sodium borate (R0191)	403
Sodium chondroitin sulfate (R0061)	409
Sunflower oil powder 25 (R0319)	413
Taurine (R0078)	417
Thiamine HCl (Vitamin B1) (R0079)	421
Thiamine mononitrate (R0212)	425
Tripotassium citrate monohydrate (R0117)	429
Valine (R0087)	433
Vanilla flavoring (natural flavor) (R0240)	437

Walnut leaf extract (R0197)	441
Yeast extract A (R0292)	445
Zinc citrate trihydrate (R0089)	449
Appendix	453
Additional calibration samples (<i>Type A</i>)	453
Additional validation samples (<i>Type B</i>)	455
Additional validation samples (<i>Type C</i>)	457
Requirements of validation	459
Conformity of <i>Apo-Ident</i> with the <i>European pharmacopeia</i>	460
References	461
Index	463

Introduction

The unequivocal proof of the identity of pharmaceutical raw materials based on a monograph or traditional alternative methods is work-intensive, time-intensive and economically often no longer makes sense. Near-infrared spectroscopy (NIR) provides a new means here. It enables relatively easy, fast and nonetheless safe identity testing via the preparation and evaluation of spectra.

The analysis system *Apo-Ident* was developed specially for pharmacy use. Pharmacists have the duty to check the identity of all raw materials for extemporaneous products in their pharmacies. This normally takes place based on the monographs for the respective substances in the European Pharmacopoeia. But also NIR spectroscopy is described in the European Pharmacopoeia as an identification method which, as opposed to the methods incorporated in the respective monographs, is approved for testing [1] (quotation translated)

with the prerequisite that the same results (“namely the determination of the identity” [2]), (quotation translated) are achieved as with the described methods and instruments.

The *Apo-Ident* analysis system serves to identify raw materials for prescriptions in the defined manner according to *ApBetrO* [Pharmacies Rules and Regulations] §§ 6 and 11 at pharmacies (NIR spectroscopy as an alternative testing method). *Apo-Ident* consists of three components:

- An *NIR spectrometer*, which records the spectra of non-preprocessed raw materials in a measuring glass in diffuse reflection or transfection.
- The *QuickStep* spectroscopy software controls the instrument and records the spectra and user inputs via a pharmacy-specific software plug-in. It also generates the test protocol for documentation of the testing and storage of the printout to be signed in the pharmacy.
- The software module *IdentModule* incorporates *reference databases*. The spectra from the *QuickStep* software are presented to it for evaluation.

NIR spectroscopy is a very powerful analytical method. It is also able to establish the identity of several chemical compounds and mixes in as far as an appropriate database (technically correct: a [chemometric model](#)) was created. Identity testing with *Apo-Ident* is a very safe, very fast and easy to operate analytical method for testing a large number of raw materials.

Context of this document

The suitability of the instrument, method and database is proven as follows:

- *NIR spectroscopy as a method for identity testing*: The *Ph. Eur.* [3] describes NIR spectroscopy in *Section 2.2.40* as an analytical method which is also suitable for the identification of raw materials. Therefore, validation of the method as such is not necessary.
- *Performance of the instrument*: The *Ph. Eur.* [3] furthermore describes the apparatus and the testing of its performance in *Section 2.2.40*. The document *Erfüllung von 2.2.40 Ph. Eur. durch Apo-Ident* [4] compares the implementation by *Apo-Ident* with this monograph in order to prove that *Apo-Ident* meets the specifications of the Pharmacopoeia. Each individual instrument delivered to a pharmacy is qualified in accordance with the tests described in “*Control of Instrument Performance*”. In this test, the unit consisting of analysis instrument hardware and the *QuickStep* spectroscopy software is assessed. The result is documented in a test protocol which is kept at the pharmacy.
- *Validation of the database* is documented separately for each substance class. The report at hand documents the substance class *Hepart raw materials*.

The *Arbeitsgemeinschaft der Pharmazieräte Deutschlands (APD)* [Working Group of German Pharmacy Inspectors] has clarified the following in its resolution dated October 16, 2013 ([5], quotation translated):

NIR is a testing method incorporated in the Pharmacopoeia. The testing quality depends on the quality of the database stored. The APD views the use of NIR instruments in case of ensured validation of the databases used in conjunction with it as one of several options for identity testing.

The APD defined more precisely ([6], quotation translated) on October 1, 2014:

The use of near-infrared is a recognised testing method according to Ph. Eur. 8. For the use of NIR instruments in pharmacies for testing the identity of raw materials, sufficient and verifiable validation of the instrument used is required. The quality of the database stored by the instrument manufacturer is decisive for quality. Batch-specific differences with the same original substances must be taken into account if present.

So NIR is basically suitable. The validity of the reference database is proven with the existing validation documentation.

Criteria for the inclusion of substances

This validation documentation describes the results of the validation of the reference database for the substance class *Hepat raw materials*. Validation documentation is created for each published version of the reference database for all substance classes incorporated.

The reference database is incorporated in the software module *IdentModule*. During identification testing with *Apo-Ident*, spectra which are used for evaluation purposes are presented to it by the *QuickStep* software. In the same manner, the *IdentModule* is presented all validation spectra successively during the validation runs for evaluation purposes. The *IdentModule* responds respectively (without taking the initial assumption into account) with the identified substance or rejects it as unknown. The correctness of this response is checked for each possible initial assumption and counted.

The results are summarised for each substance and reproduced in this document. The core statement of this validation report is that the following criteria must be fulfilled for each database entry, so that *Apo-Ident* offers verification of identity by means of NIR for the relevant substance/substance group:

- The database is exclusively generated from spectra which have been recorded by *HiperScan GmbH* on traceable samples in pharmaceutical quality.
 - The samples are procured via typical pharmacy sources (*DAC III.2.: Bezugsquellennachweis für Rezepturbestandteile [reference source for prescription components]* [7]).
 - A valid manufacturer's certificate exists (content, purity and identity of the batch).
 - The identity was confirmed by a certified test laboratory or *HiperScan GmbH*.
- Each version of the reference database (every update) is validated in-full.
 - Calibration spectra (*Type A*), other spectra recorded under the control of *HiperScan GmbH* (*Type B*), and spectra from the field (*Type C*) are presented to the *IdentModule* for evaluation in three separately evaluated validation runs.
 - Here, no single *false positive* result may arise.
 - Here, the various substance classes are also tested for reciprocal rejection, where this is objectively justified (see *Summary* section).
- In the validation with spectra recorded under the control of *HiperScan GmbH*, spectra of at least one independent sample must be considered, i.e. spectra from a batch whose spectra have not been used for the generation of the database. In addition, the set of *Type A* and *Type B* spectra must originate from at least three different batches.
- Spectra of additional substances may be used for the generation of the database even though they will not be offered for identification by this database. The purpose is the reliable distinction from these substances.

- Any positive result of *Apo-Ident* confirms the identity of the substance/substance group and distinguishes it from all other substances of the database. In the case of substance groups, the result is ambiguous: Distinction from all substances not belonging to the group is proven. The substance is identified as a member of this group. However, within the substance group, it is not possible to reliably classify which substance has been tested.
- The criteria for clear identifiability are a [specificity](#) of 100 % ([true negative rate](#)) and a minimum distance in the distance matrix. See 2. d) under [Model creation procedure and validation runs](#).

Validation concept

[Chemometrics](#) is a statistical technique for the extraction of relevant chemical information from spectra. In mathematics, this method is described as *multivariate data analysis*. Chemometrics proceeds here as follows:

1. Collection of spectra for the *calibration sample*. The results (identities) of the calibration sample must be known. The calibration samples must be representative for the samples which are to be evaluated later. Therefore, they must take the various possible (physical) compositions into account. (Therefore, sourcing calibration samples for NIR from the specialist trade is superior to the use of CRS reference substances.)
2. The first mathematical step is *calibration*. Here, the [chemometric model](#) is calculated from the *calibration sample spectra* ([reference spectra](#)) and limits as well as some parameters are stipulated. The chemometric model is used later to calculate the analysis result (*prediction*).
3. Collection of further spectra for the *validation sample* which should be independent of the *calibration sample*. The results (identities) of the *validation sample* must also be known. The textbook suggests a random sample with a normal scope of 25 % to 50 % of the *calibration sample* [8].
4. The second technical data step is *validation*. Here, the [chemometric model](#) created is evaluated based on the spectra of the *validation samples*. As validation parameters for the identification, the *Ph. Eur. Section 2.2.40* [3] specifies the [specificity](#) and [robustness](#).

The validation step according to the textbook has the target of estimating the performance capability of the model created based on a random sample. In order to achieve the best possible precision, attention is paid to the calibration sample. In the field of pharmaceuticals, the safety of the method has priority. In order to be able to [validate](#) the model within the regulatory scope, the validation step must include probative force. For this purpose, the validation sample must be *representative and complete* in order to enable the testing of all cases.

A *sufficient number of batches* must be secured for validation because validation finally proves whether the number of batches in calibration suffices.

Each substance is validated individually. The validation results are documented per substance in this document. Moreover, the documents show how many and which batches have been used for creating the model or model validation.

At least one certificate is taken in for each substance from an accredited test laboratory for the independent testing of identity of the sample. The identification number of the corresponding test certificate is listed in the report, enabling traceability of a substance tested according to the monographs in the Pharmacopoeia.

Model creation procedure and validation runs

The safety of the [chemometric models](#) is guaranteed by several measures during model creation, of which the validation step is the final one. Normally, the procedure is as follows. It is in particular valid for the active pharmaceutical ingredients (APIs) *solid API excipients, liquid/semi-solid API excipients (with a test certificate), narcotics - solid medicinal substances and drugs*. If, for individual substance classes, variations are required, they are depicted in the section [Particularities of individual substance classes](#).

1. Collecting the reference spectra (calibration sample)

- a) Procurement of the samples from the same sources from which pharmacies source their raw material for compounding (Caelo, Fagron, Euro-OTC, ..., see also *DAC III.2. Bezugsquellen-nachweis für Rezepturbestandteile* [Sources of supply for compounding] [7]).
- b) Testing the suitability according to *ApBetrO* [Pharmacies Rules and Regulations] §§ 6, 11, that is to say the availability of a valid manufacturer certificates via identity, purity and contents of the batch.
- c) Recording standard 40 spectra of the sample in different positions, as a standard on four instruments. Here, handling and presentation of the samples as later in the pharmacy.
- d) Visual checking for anomalies in the spectra. In case of indications of measurement errors, measurement must be repeated. If a signature is missing in the spectrum, the substance may be excluded from the start as not promising (the spectra are nonetheless entered in the database validation as independent *Type B* spectra).
- e) Testing identity. For each substance, a certificate of correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the respective following substance page of this validation documentation the *Mahalanobis distance* to this reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance. Such samples underline the statistical spread of the original reference substance, but cannot add any new characteristics of the substance.

HiperScan GmbH cooperates with some suppliers as follows: the raw materials supplier takes a sufficiently large sample in his incoming goods area so that part of it can be used for recording the NIR spectra. The remainder of the sample goes to analytics for market approval. The manufacturer's batch certificate emerges from these identity, contents and purity tests, which consequently also prove the correct identity of the NIR reference sample. Therefore, the NIR spectra are suitable for structuring the database (*Type A*) and can also optionally be used for validation (*Type B*). The samples which this applies to are marked in the validation report with a footnote.

- f) If the identity of the new sample is proven, it is declared as a reference sample and the spectra are approved for structuring the database.

2. Generating the chemometric models (calibration)

- a) Determination of the transformation matrix from the reference spectra using variance maximisation [8, 9]. (All reference spectra are always included, even if only a few spectra are added for an update.) All reference spectra receive the same data pre-treatment, which is also later applied to all measurement spectra in the field (in the pharmacies).
- b) Checking that the number of principal components used is still sufficient.
- c) Calculating the limits for each substance from the spread of the reference spectra. The calculation regulation is identical for each substance in a substance class.
- d) Überprüfen der Abstände zwischen den Grenzen der trennbaren Substanzen: Die Distanzmatrix enthält die *Mahalanobis-Abstände* von jeder Substanz zu jeder anderen. Die Werte hin und zurück sind jeweils unterschiedlich, weil die Streuung der Ausgangssubstanz eingeht. Ist eine Distanz kleiner als der Mindestabstand, so gelten die Substanzen als nicht sicher trennbar. Der Mindestabstand ist auf 9 festgelegt. Der Entwickler des Modells darf einen größeren Mindestabstand festlegen (ein Wert für das gesamte *chemometrische Modell*), um die Trennschärfe zu erhöhen.
- e) Überprüfung des Modells anhand der Referenzspektren. Es sind keine *falsch-positiven* Ergebnisse erlaubt.

f) Wird eines der Kriterien verletzt (d) *Unterschnittener Mindestabstand zwischen zwei Substanzen* oder (e) *Eine Substanz wird als eine andere identifiziert*, entscheidet der Entwickler der Datenbank, welche der folgenden Optionen er anwendet:

- Er nimmt beide Substanzen aus der Datenbank. (Die Spektren bleiben in der Validierung und dürfen auch in den Aufbau eingehen. Sie werden aber nicht zur Prüfung angeboten.)
- Er bildet eine Substanzgruppe mehrerer nicht sicher trennbarer Substanzen. Dann ist das Ergebnis mehrdeutig: Das chemometrische Modell stellt fest, dass es sich bei der Probe um eine der Substanzen aus der Gruppe handelt und dass es sich um keine andere Substanz handelt. Es kann aber nicht sagen, um welche der Substanzen es sich handelt. Um die eindeutige Identität festzustellen, muss der Anwender eine geeignete ergänzende Prüfung durchführen.
- Er erstellt ein weiteres *chemometrisches Modell* mit geringerem Umfang, in das mindestens alle Substanzen der nicht sicher trennbaren Substanzgruppe eingehen (Zweite-Stufe-Modell). Zweite-Stufe-Modelle werden nur aufgerufen, wenn die erste Stufe festgestellt hat, dass es sich nur um eine der Substanzen handeln kann, die in den Aufbau der Zweiten Stufe eingegangen ist.

3. Set of validation spectra (validation samples)

The following is provided for validation:

- a) *Type A*: The reference spectra = calibration spectra from which the database was generated. These also include spectra from substances which the *chemometric model* should not identify, but were also recorded during generation in order to increase selectivity. (As a result, the model “learns” to differentiate from other substances which are actually unknown to it.)
- b) *Type B*: Spectra recorded under the control of *HiperScan GmbH* and not used for the generation of the database. These also include reference spectra of other substance classes, and spectra that are not used as reference spectra. Samples are considered to be independent, if they originate from a batch, of which no spectra have been used for the generation of the database. (Up to *IdentModule 2018-01*, samples were still considered independent if the sampling was done independently, i.e. if they originate from another sales container.)
- c) *Type C*: Spectra from the field, which have not been recorded under the control of *HiperScan GmbH* and have not been used for the generation of the database. The spectra include not only substances of the substance class to be tested, but also substances from other classes.

All manufacturers’ batches from which spectra are used for the validation are listed by substance in this document: for substances included in the substance class *Hepart raw materials* in the respective validation reports; otherwise in attachments *A*, *B* and *C*.

Furthermore remains valid: validation spectra may only be removed if a error in the spectrum can be proven. Here, the spectra are not deleted, but instead placed on a *blacklist* incorporating the reason, date and initials in the commentary.

The section *Particularities of individual substance classes* treats the other substance classes from which *Type B* and *Type C* spectra are cited for validation purposes.

4. Validation runs and approval

- a) Validation spectra are transferred holistically to the *IdentModule* for evaluation in the same way as the spectroscopy software *QuickStep* transfers measured spectra.
- b) Following the provision of each spectrum, the *IdentModule* responds as to whether it has recognised a substance and which substance was recognised.
- c) The correctness of this response is checked for each possible initial assumption (each measurable substance with the substance class) and counted according to *true negative*, *false negative*, *true positive* and *false positive*. These figures are provided for each substance and additionally in the section *Summary*, separated according to types *A*, *B* and *C*.

- d) No *false positive* results whatsoever are permissible.
- e) If the criterion is also met for all substance classes, the *IdentModule* is approved.

Particularities of individual substance classes

Basically, *HiperScan GmbH* procures and tests the manufacturer's certificate for the batch, commissions external testing of the identity of the sample or carries it out independently and stores the certificates. As described, this process is established for the Pharmacopoeia substances, that is to say for substance classes **APIs & excipients, solid**, **APIs & excipients, liquid/semi-solid (with analysis certificate)**, **Narcotic substances, liquid/semi-solid** and **Drugs**. Therefore, *HiperScan GmbH* is able to furnish proof of the identity of the reference samples. In case of manufacturer-specific substance classes and others, individual steps are organised differently in-part:

The substance class **APIs & excipients, liquid/semi-solid (other)** (often described as cosmetics) incorporates substances for which no specification of the requirements of the pharmaceutical quality is determined, neither in a Pharmacopoeia monograph, a DAC/NRF monograph nor via a manufacturer's specification. Consequently, neither the identity nor contents can be tested independently. No certificates whatsoever exist for the reference samples. So here, merely the matching of the sample with former samples of this product is established and confusion with the other substances is ruled-out. (If the manufacturer of such a substance prepares a specification, determines testing methods and provides manufacturer's certificates in accordance with *ApBetrO* [Pharmacies Rules and Regulations] §§ 6, 11, *HiperScan GmbH* can assign the substance to the substance class *APIs & excipients, liquid/semi-solid (with analysis certificate)* again in the future).

Substance class **HCK – nutritional supplements (Hepart)** contains the HCK micro-nutrients from the Swiss company *Hepart AG*. *HiperScan GmbH* receives the reference samples directly from the manufacturer. For each reference sample, *HiperScan GmbH* also receives manufacturer's certificates and keeps these. New checking of the identity of the reference sample is not carried out by *HiperScan GmbH*. The identity of the reference samples is therefore documented by *Hepart AG*. The spectra of all batches provided by *Hepart AG* are recorded by *HiperScan GmbH* and entered in the database.

All the manufacturer's batches are used for the generation and validation of the substance class **HCK**. The expected variation is also represented in the generation and validation if there are less than three batches.

Also, for the substance class **PhytoComm** (TCM-Granulated herbal extracts of the manufacturer *PhytoComm*) spectra for all useable batches are recorded by *HiperScan GmbH* and entered in the database. The supplier organises the respective tests themselves and keeps the test certificates.

A new evaluation option was created for the class *PhytoComm* with the update 2016-01. As the risks are considerably fewer than those from chemical agents, the pharmacist can specify a reasonable criterion for the *specificity* in accordance with internal risk estimation. The database for this is created without taking safety distances into account and no criterion is determined in advance for the *specificity*. Instead, the *specificity* for testing the identity with this concrete substance is calculated in the validation for each substance and provided with the measurement result. The pharmacist then judges himself whether this safety is reasonable with regard to the risk of the substance.

Additionally, a statistical forecast is provided for the *specificity* which is determined according to the *Rule of Three* [10, 11]. For this forecast, it is assumed that there would have been three wrong results more and is provided with a lower limit for *specificity*. This value has a special meaning if a *specificity* of 100 % is achieved for a substance during validation. In this case, the lower limit allows conclusions regarding the scale of existing safety for which with an endless number of validation spectra a value of less than 100 % is to be assumed.

If, for example 14 000 spectra not belonging to the substance are presented and no *false positive* classification is made, a hypothetical number of three *false positive* results is assumed (*Rule of Three* [10, 11]) and the *specificity* is defined with 100.0000 % (> 99.9786 %). Here, it applies that the higher the number of validation spectra which form the statistical basis, the better the *specificity* calculated via the lower *specificity* limit will be approximated.

The positive result of the identity test using *Apo-Ident* establishes that the sample spectrum is in accordance with a batch of the specified granulate from the supplier *PhytoComm*, whereby all useable batches from the supplier are known.

The *PhytoComm* class can only confirm the identity of batches that have been used for the generation of the database. As a consequence there cannot be any validation spectra of other batches. Therefore, the criterion reads that two samples (from different sales containers) from each batch must exist, one for the structure of the database (*Type A*) and one for the validation (*Type B*).

Significance of testing with *Apo-Ident*

The analysis result is determined using sophisticated statistical methods according to state-of-the-art science and technology. Chemical and pharmaceutical knowledge is applied for the selection of the samples from which the calibration spectra and validation spectra are recorded. Otherwise it does not influence the further steps of model creation.

Verbally, the statement of the analysis result can be expressed as follows. Here “*the spectra match*” means that the criteria *Mahalanobis distance*, *outlier analysis* and *correlation* are met as shown in *Erfüllung von 2.2.40 Ph. Eur. durch Apo-Ident* [4]. “The spectra do not match”, on the other hand, means that at least the criterion *Mahalanobis distance* is not met.

The positive analysis result “*was identified as ...*” is very meaningful because both the quantity of substances to be taken into account and the number of underlying samples is very comprehensive.

1. The spectrum of the sample measured matches spectra of the defined substance.
2. The spectrum of the sample measured does not match any spectrum of any other substance in this substance class. Therefore, all other substances can be clearly ruled-out.
3. As the spectra from other substance classes were used for validation, it is proven that no spectrum of one of these other substances matches the defined substance. (All substance classes with which a spectrum comparison is possible and makes sense are used for validation. This is documented for each substance class in the section *Summary*.)
4. If the defined substance belongs to a *substance group* which in itself is not clearly separable with *Apo-Ident*, matching with the spectra of one or several substances in this group is confirmed. Which of these substances it actually is cannot be determined clearly. All other substances are excluded analogous to 2 and 3.

On the other hand, a negative analysis result “*was not identified as ...*” means:

1. The substance offered could not be recognised based on the spectrum of this sample.
2. The identity of this sample is not confirmed.
3. Testing must be repeated in accordance with the specifications of the Pharmacopoeia.

Conclusion

NIR spectroscopy is a testing method incorporated in the Pharmacopoeia. In case of successful database validation, it is a possible method for identity testing [5]. *Apo-Ident* meets the criteria of the *European Pharmacopoeia* as a near-infrared spectrometer and proves the validity of the reference database with the existing validation documentation. This means that *Apo-Ident* can be used as an alternative testing method for testing raw materials at pharmacies.

Explanation of terminology

The following section serves to explain or define specialist terminology which is required in order to understand this document. If necessary, definitions for the analysis system *Apo-Ident* are defined more precisely.

The term database is used in this document exactly as in the *Ph. Eur. Section 2.2.40* [3] synonymous with [chemometric model](#). In order to differentiate the databases which are relatively independent of each other, *HiperScan GmbH* frequently also uses the term [substance class](#) (primarily in the plural). On the other hand, the spectra used to structure the database are termed spectrum collection and not database.

Substance classes are units of the organisational structure of the *IdentModule*. The substance classes are substance [databases](#) which are also broadly independently subscribable. On the one hand, the liquid and semi-solid substances are separated from the solid powders because they are measured against different references and therefore the spectra cannot be compared. On the other hand, for example the Pharmacopoeia substances are kept separated from the manufacturer-specific database *PhytoComm* for TCM (traditional Chinese medicine) raw materials.

The individual substance classes need only be limited against each other in-part. Often, no risk of confusion exists because they can only be procured from different sources. On the other hand, in several cases we handle substances which need not be distinguished. For example, em Huang Qi granulate from the company *PhytoComm* neither needs to be delimited from *Huang Qi* granulate from the company *HerbaSinica* nor is matching required. Respectively one single [chemometric model](#) is behind a substance class. (Even if several reciprocally secured chemometric models would be permissible.) The terms *substance class*, *chemometric model* and *databases* are mostly used here as synonyms.

A substance group respectively summarises all the substances within a [substance class](#) which cannot safely be distinguished from one another based on their NIR spectra. However, all the other substances in the database can be excluded.

The formation of subgroups is mentioned in the *Ph. Eur. Section 2.2.40* [3]. In this manner, technical restrictions in case of extensive databases can be avoided and it is possible to prepare individual subgroups with different spectrum pre-treatment. Validation of the subgroups against each other is required. *HiperScan GmbH* has solved these technical restrictions and doesn't use any subgroups within a substance class any longer.

Principal component analysis (PCA) [8, 9] is a multivariate statistics process or multivariate data analysis. It serves to structure, simplify and illustrate comprehensive data records by describing a large number of statistical variables by describing a lower number of linear combinations (the *principal components*) which are as significant as possible. In the *Apo-Ident IdentModule*, *PCA* is used to evaluate the recorded spectrum data (corresponding with *Ph. Eur. 2.2.40* [3]).

The term validation is defined in both relevant contexts here with different (even if related) meanings.

Within the sense of the expert discipline of *chemometrics*, validation is a process step when creating a [chemometric model](#): after a transformation matrix, limits and various parameters have been calculated or determined from a set of reference spectra during the course of the calibration step [8, 9], the validation step determines the performance capability of the model (selectivity, precision, ...) based on the validation spectra. Normally, random sampling is planned here. In order for the validation to gain strength of proof, the validation spectrum set must be selected with an appropriately wide scope (*representative* and *complete*). The terms *validation run* and *validation step* always actually mean the process step in this sense.

In the regulatory sense (of pharmaceutical production), validation is the documented proof that a process or system meets the previously specified requirements reproducibly when applied practi-

cally. In this sense, the *Apo-Ident* databases only become validated databases with the validation documentation, which this document is part of.

The *European Pharmacopeia* uses the term validation in *Section 2.2.40* within the sense of the specialist discipline of *chemometrics* [3].

The robustness of a process is the property of only being influenced by environmental fluctuations (e.g. temperature or humidity) a little. A method is robust if the environmental conditions do not or hardly falsify the final result.

The specificity of a classification (of a [chemometric model](#)) is the [true negative rate](#).

The recognition rate (also sensitivity) is the [true positive rate](#). It defines in how many percent of cases a correctly set up substance is actually confirmed.

The true negative rate describes the share of spectra correctly classified as non-identity during validation. This is equivalent to correct classification. It means that a substance *A* within identity checking as substance *B* is judged as “*not identified*”. The *true negative rate* is equivalent to the conditional frequency

$$h(\text{rejected}|\text{genuinely no identity}) = \frac{r_n}{r_n + f_p}$$

with r_n as the total number of *true negative* classifications and f_p as the total number of *false positive* classifications. For successful validation of an *IdentModule*, all spectra presented belonging to this category must be classified as *not in accordance*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of this number. The weight of each spectrum of a substance/substance group *i* therefore results as

$$w_i = \frac{1}{n_i}$$

with n_i number of spectra in this substance/substance group. This weighting ensures that the overall result cannot be enhanced by adding especially large numbers of spectra from easily separable substances.

The true positive rate describes the share of spectra correctly classified as identity during validation. This is equivalent to correct classification. It means that a substance *A* within identity checking as substance *A* is judged as “*identified*”. The *true positive rate* is equivalent to the conditional frequency

$$h(\text{identified}|\text{genuine identity}) = \frac{r_p}{r_p + f_n}$$

with r_p as the total number of *true positive* classifications and f_n as the total number of *false negative* classifications. The *true positive rate* is a measure for the recognition rate of the validated *Apo-Ident IdentModule*.

In order to ensure that each substance is received with the same weight, the spectra are weighted as described for the [true negative rate](#).

The true negative result describes a spectrum correctly classified as non-identity during validation. It is equivalent to correct classification. It means that a substance *A* within identity checking as substance *B* is judged as “*not identified*”.

The false positive result describes a spectrum falsely classified as non-identity during validation. This is the most critical type of possible false classification. It means that a substance *A* within identity checking as substance *B* is judged as “*identified*”. For successful validation of an *IdentModule*, a number of false positive events of zero are demanded for all spectra entering the validation. The exception to this restriction is the class of TCM granulates from the company *PhytoComm* as described under [Particularities of individual substance classes](#).

The true positive result describes a spectrum correctly classified as identity during validation. It is equivalent to correct classification. It means that a substance *A* within identity checking as substance *A* is judged as “*identified*”.

The false negative result describes a spectrum falsely classified as non-identity during validation. It is equivalent to false classification. It means that a substance *A* within identity checking as substance *A* is judged as “*not identified*”.

The ‘Rule of Three’ says that with a probability of 95 % the next random sample of the same size no more than three false results are to be expected if no false result existed in the existing random sample [10, 11].

The *specificity* and *recognition rate* are determined both globally and from the validation runs for all substances. The information is supplemented with the hypothetical value if there had been three false results more. The percent information is provided in parentheses with the “greater than” symbol ‘>’, e.g. *specificity* 100.000 % (>99.983 %) if 17 567 false spectra have been presented without one single *false positive* result. The larger the statistical basis, the lower the influence of the hypothetical false results.

The Mahalanobis distance is a distance measure between two points in *n*-dimensional vector space. Here, the respective direction component of the distance to *standard deviation* [12] of an *n*-dimensional distribution is standardised. In case of the *principal component analysis* [8, 9] this standardisation relates to the distribution of the respective calibration data set for a classification (substance/substance group) in the *principal component space* [8]. The *Mahalanobis distance* of a point (mapping of a spectrum) $\vec{y}$ in the *n*-dimensional principal component space to the expected value of an *n*-dimensional distribution $\mathbf{X}$ then results as

$$d(\mathbf{X}, \vec{y}) = \sqrt{(\bar{\mathbf{X}} - \vec{y})^T \mathbf{S}^{-1} (\bar{\mathbf{X}} - \vec{y})} \quad \text{with} \quad \mathbf{X} \in \mathbb{R}^{m \times n}, \vec{y} \in \mathbb{R}^m$$

[13]. Here, *m* is equivalent to the number of principal components used (dimension of the principal component space) and *n* the number of measurements existing in the calibration data set (spectra). $\bar{\mathbf{X}}$ is the expected value of the resulting distribution for the calibration data set (the average value of *n* measurements received). $\mathbf{S}^{-1}$ is the inverse covariance matrix [12] for distribution $\mathbf{X}$.

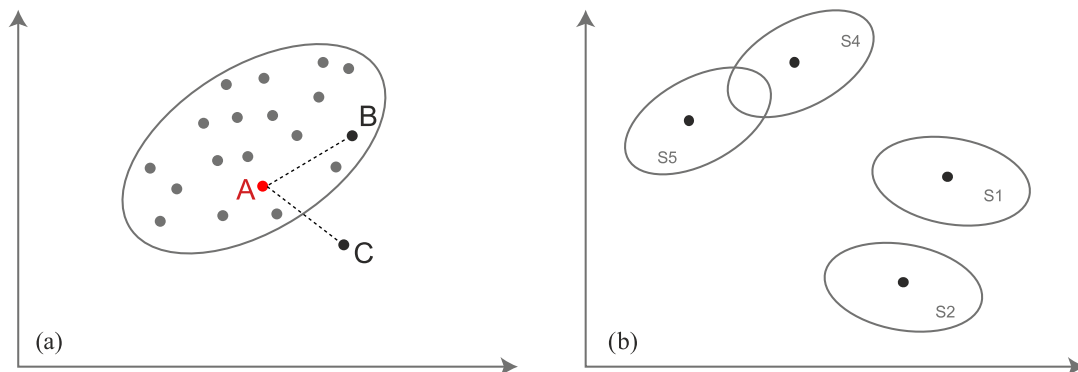


Figure 1: (a) The *Mahalanobis distance* from *A* to *B* is less than from *A* to *C*. However, the *Euclidean distances* are identical. (b) The *Mahalanobis distance* between the two measurement series *S4* and *S5* is smaller than between *S1* and *S2*. However, the *Euclidean distances* are identical.

The *Mahalanobis distance* offers advantages compared to the *Euclidean distance*: For the calculation of the distance it takes the statistical properties of a data point distribution (measurement series), i.e. average value, variance and covariance of the data points [14] into account. The *Mahalanobis distance* is applied while creating the reference database for evaluating the spectra from different samples of a substance.

A chemometric model is a classifier based on statistical methods [8, 9]. Through the respective algorithm used (e.g. *Principal Component Analysis*, *Cluster Analysis*), a maximum of chemical information is extracted from measurement data. Here, systematical or physical disturbances are eliminated using appropriate data pre-processing [15, 16].

At several places in this document, in order to simplify understanding, the term **database** is used instead of *chemometric model* – in the same manner as in the *Ph. Eur. Section 2.2.40* [3].

A sample (with its own sample ID) refers to substance in a sales container. Repeated sampling from the same sales container is listed under the same sample ID. (The suffix “SI” is not part of the sample ID.) Several samples may originate from the same batch. Samples are called “independent”, if they originate from a batch, of which no spectra have been used for the generation of the database. (Up to *IdentModule 2018-01*, samples from different sales packages were considered to be independent.) The information above the list of validation spectra now includes also the number of batches that deliver independent samples for the validation (for both *Type B* and *Type C*).

In case a supplier takes a sample for testing from its incoming goods and splits it to multiple laboratory containers, the substance in all laboratory containers will still be ascribed to the same sample. *HiperScan GmbH* only uses one of the subsamples.

Reference samples are used to structure the database. The *reference spectra* originate from these samples. In chemometric technical jargon you would normally say: For *calibration*, a **chemometric model** is generated from the *calibration spectra* recorded from the *calibration samples*, whose quality is subsequently assessed in **validation**.

Reference samples are procured via typical pharmacy sources. Their identity is tested. The *reference spectra* are recorded by *HiperScan GmbH*. The documentation also includes the manufacturer’s name and batch number.

Reference samples are clearly identified by a sample ID. Samples without sample ID may not be used as *reference samples*.

Summary

A total of 23 502 spectra from 511 different batches for a total of 138 substances were used to validate the substance class *Hepart raw materials*.

Validation samples

The validation samples can be categorised as follows:

Typ A Calibration spectra. These are the spectra used to generate the chemometric model. They were recorded by *HiperScan GmbH*. Detailed information regarding the batches or samples can be found in the following validation reports under *calibration samples* and under *Type A*. Further information is listed in [Appendix A](#).

Substance class	Substances	Batches	Spectra
Hepart raw materials	108	274	13 410

From category *A* a total of 13 410 spectra from 274 batches for a total of 108 substances were taken into account for validation.

Typ B Spectra from independent samples which are not included in database generation. These spectra were recorded by *HiperScan GmbH*. Detailed information regarding the batches or samples can be found in the following validation reports in the section *Type B* or in [Appendix B](#).

Substance class	Substances	Batches	Spectra
Hepart raw materials	106	237	10 089

From category *B* a total of 10 089 spectra from 237 batches for a total of 106 substances were taken into account for validation.

Typ C Spectra from independent samples which are not included in database generation. *Apo-Ident* customers carried out the measurements. Detailed information regarding the batches or samples can be found in the following validation reports in the section *Type C* or in [Appendix C](#).

Substance class	Substances	Batches	Spectra
Hepart raw materials	1	3	3

From category *C* a total of 3 spectra from 3 batches for a total of 1 substances were taken into account for validation.

Validation results

The validation runs checked whether all substances/substance groups in the substance class *Hepart raw materials* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, the matching of all relevant spectra of various substances with the substances/substance groups in the substance class *Hepart raw materials* was checked and the correctness of the results was evaluated. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	13 360	10	1 327 630
Type B	0	7914	765	1 000 171
Type C	0	3	0	297

All substances/substance groups in the substance class *Hepart raw materials* can be clearly distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Typ A	100.000 00 % (> 99.999 31 %)	99.950 00 % (> 99.915 83 %)
Typ B	100.000 00 % (> 99.999 01 %)	83.616 57 % (> 83.558 53 %)
Typ C	100.000 00 % (> 97.979 80 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Acacia gum (encapsia) / acacia gum (R001/R0141)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85084-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Acacia gum (encapsia) / acacia gum (R001/R0141)

Special notes

When selecting the *Acacia gum (encapsia) / acacia gum (R001/R0141)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Acacia gum (encapsia) / acacia gum (R001/R0141)	5	5	0

Second-stage model

For differentiation of the substance/substance group *Acacia gum (encapsia) / acacia gum (R001/R0141)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Acacia gum (encapsia) / acacia gum (R001/R0141)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Acacia gum (enca...	1212002	85084	50	not required
Hepart	Acacia gum (enca...	14001701	85232	60	not required
Hepart	Acacia gum (enca...	15000744	85285	60	not required
Hepart	Acacia gum (enca...	15001443	85323	60	not required
Hepart	Acacia gum (enca...	15001837	85339	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 290 spectra of 5 reference samples from the substance/substance group *Acacia gum (encapsia) / acacia gum (R001/R0141)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 5 different batches.
- 13 120 spectra from a total of 268 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 240 spectra of 5 reference samples from the substance/substance group *Acacia gum (encapsia) / acacia gum (R001/R0141)*.
- Among them are spectra of independent samples from 5 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Acacia gum (encapsia) / acac...	15001444	85322	60
Hepart	Acacia gum (encapsia) / acac...	16000477	85354	60
Hepart	Acacia gum (encapsia) / acac...	16001008	85430	40
Hepart	Acacia gum (encapsia) / acac...	19001212/0	85595	40
Hepart	Acacia gum (encapsia) / acac...	19001212/1	85596	40

- 9849 spectra from a total of 232 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Acacia gum (encapsia) / acacia gum (R001/R0141)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Acacia gum (encapsia) / acacia gum (R001/R0141)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Acacia gum (encapsia) / acacia gum (R001/R0141)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	290	0	13 120
Type B	0	240	0	9799
Type C	0	0	0	3

The substance/substance group *Acacia gum (encapsia) / acacia gum (R001/R0141)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9304 %)	100.0000 % (> 97.9310 %)
Type B	100.0000 % (> 99.9014 %)	100.0000 % (> 97.5000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Acacia gum (encapsia) / acacia gum (R001/R0141)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Potato starch, pregelatinized (R0047)	17.64	—
Juniper extract (R0192)	22.67	—
Galactomannans (R0035/R0094)	23.17	—
Gentian extract (R0200)	26.09	—
Marshmallow root extract (R0204)	32.88	—
Menaquinone-7 0.2% (R0134)	35.43	—
Konjac glucomannan 80% (R0293)	37.88	—
beta-glucan 28% (oat fibre) (R0334)	42.92	—
Ginseng extract (R0029)	46.85	—
Guarana extract (R0289)	49.77	—
Calcium lactate pentahydrate (R0013)	50.56	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Acacia gum (encapsia) / acacia gum (R001/R0141)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85285	85285	0.00	20.30
85084	85084	0.00	20.28
85232	85232	0.00	21.96
85323	85323	0.00	19.91
85339	85339	0.00	17.64

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **a-lipoic acid (R0003)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85227-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

a-lipoic acid (R0003)

Special notes

When selecting the *a-lipoic acid (R0003)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
a-lipoic acid (R0003)	2	1	0

Second-stage model

For differentiation of the substance/substance group *a-lipoic acid (R0003)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *a-lipoic acid (R0003)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	a-lipoic acid (R...	1011008	85227	60	not required
Hepart	a-lipoic acid (R...	16001690	85466	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 100 spectra of 2 reference samples from the substance/substance group *a-lipoic acid (R0003)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 310 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *a-lipoic acid (R0003)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	a-lipoic acid (R0003)	19002348/0	85626	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *a-lipoic acid (R0003)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *a-lipoic acid (R0003)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *a-lipoic acid (R0003)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	100	0	13 310
Type B	0	40	0	10 049
Type C	0	0	0	3

The substance/substance group *a-lipoic acid (R0003)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9308 %)	100.0000 % (> 94.0000 %)
Type B	100.0000 % (> 99.9016 %)	100.0000 % (> 85.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *a-lipoic acid (R0003)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Magnesium stearate (R0140)	83.92	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *a-lipoic acid (R0003)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85227	85227	0.00	84.84
85466	85466	0.00	83.92

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Apple powder (R0005)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85055-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Apple powder (R0005)

Special notes

When selecting the *Apple powder (R0005)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Apple powder (R0005)	1	0	0

Second-stage model

For differentiation of the substance/substance group *Apple powder (R0005)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Apple powder (R0005)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Apple powder (R0...	20001157	85866	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 40 spectra of 1 reference samples from the substance/substance group *Apple powder (R0005)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 370 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Apple powder (R0005)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Apple powder (R0005)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Apple powder (R0005)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Apple powder (R0005)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	40	0	13 370
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Apple powder (R0005)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9317 %)	100.0000 % (> 85.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Apple powder (R0005)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Marshmallow root extract (R0204)	24.70	—
Acacia gum (encapsia) / acacia gum (R0...	26.65	—
Guarana extract (R0289)	27.83	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gentian extract (R0200)	30.01	—
Konjac glucomannan 80% (R0293)	31.76	—
Inulin (R0041)	34.72	—
Citrus bioflavonoids 60% (R0021)	42.75	—
Galactomannans (R0035/R0094)	42.89	—
Vanilla flavoring (natural flavor) (R0...	47.34	—
Juniper extract (R0192)	48.18	—
D-ribose (R0137)	51.29	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Apple powder (R0005)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85866	85866	0.00	27.83

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group	Artichoke extract (R0208)
Substance class	Hepart raw materials
Report date	22/03/2021
Report number	85289-2021-03-22
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Artichoke extract (R0208)

Special notes

When selecting the *Artichoke extract (R0208)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Artichoke extract (R0208)	4	1	0

Second-stage model

For differentiation of the substance/substance group *Artichoke extract (R0208)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Artichoke extract (R0208)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Artichoke extrac...	15000797	85289	60	not required
Hepart	Artichoke extrac...	16000723	85371	40	not required
Hepart	Artichoke extrac...	20001465	85854	40	not required
Hepart	Artichoke extrac...	20001147	85855	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 180 spectra of 4 reference samples from the substance/substance group *Artichoke extract (R0208)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 4 different batches.
- 13 230 spectra from a total of 269 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *Artichoke extract (R0208)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Artichoke extract (R0208)	18001106	85545	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Artichoke extract (R0208)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Artichoke extract (R0208)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Artichoke extract (R0208)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	180	0	13 230
Type B	0	39	1	10 049
Type C	0	0	0	3

The substance/substance group *Artichoke extract (R0208)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9305 %)	100.0000 % (> 96.6667 %)
Type B	100.0000 % (> 99.9016 %)	97.5000 % (> 90.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Artichoke extract (R0208)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Katuka herb extract (R0210)	11.45	—
Hyssop extract (R0205)	15.08	—
Juniper extract (R0192)	19.72	—
Marshmallow root extract (R0204)	24.94	—
Ginseng extract (R0029)	25.75	—
Garcinia Cambogia (R0028)	27.02	—
Oregano extract (R0206)	27.05	—
Coriolus extract (R0207)	30.21	—
Black cumin extract (R0193)	31.84	—
Gentian extract (R0200)	32.03	—
Magnesium glycerophosphate (R0114)	32.07	—
Hydrangea extract (R0203)	32.14	—
Citrus bioflavonoids 60% (R0021)	32.62	—
Chili extract (R0198)	33.94	—
Acacia gum (encapsia) / acacia gum (R0...	34.82	—
Osha root extract (R0220)	35.20	—
Chaga extract (R0016)	37.43	—
Walnut leaf extract (R0197)	41.48	—
Calcium lactate pentahydrate (R0013)	51.94	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Artichoke extract (R0208)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85371	85371	0.00	19.35
85289	85289	0.00	11.45
85854	85854	0.00	25.15
85855	85855	0.00	28.69

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Ascorbic acid (R0007)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85005-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Ascorbic acid (R0007)

Special notes

When selecting the *Ascorbic acid (R0007)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Ascorbic acid (R0007)	7	5	0

Second-stage model

For differentiation of the substance/substance group *Ascorbic acid (R0007)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Ascorbic acid (R0007)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Ascorbic acid (R...	14001127	85197	40	not required
Hepart	Ascorbic acid (R...	15000274	85253	60	not required
Hepart	Ascorbic acid (R...	15000642	85270	60	not required
Hepart	Ascorbic acid (R...	15000642/1	85277	60	not required
Hepart	Ascorbic acid (R...	16000189	85352	60	not required
Hepart	Ascorbic acid (R...	16001693	85463	40	not required
Hepart	Ascorbic acid (R...	20000822	85841	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 360 spectra of 7 reference samples from the substance/substance group *Ascorbic acid (R0007)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 7 different batches.
- 13 050 spectra from a total of 266 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 200 spectra of 5 reference samples from the substance/substance group *Ascorbic acid (R0007)*.
- Among them are spectra of independent samples from 5 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Ascorbic acid (R0007)	1109001	85171	40
Hepart	Ascorbic acid (R0007)	16000865	85408	40
Hepart	Ascorbic acid (R0007)	18001129	85554	40
Hepart	Ascorbic acid (R0007)	19000252	85584	40
Hepart	Ascorbic acid (R0007)	19002129	85828	40

- 9889 spectra from a total of 232 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Ascorbic acid (R0007)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Ascorbic acid (R0007)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ascorbic acid (R0007)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	360	0	13 050
Type B	0	159	41	9889
Type C	0	0	0	3

The substance/substance group *Ascorbic acid (R0007)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9303 %)	100.0000 % (> 98.3333 %)
Type B	100.0000 % (> 99.9004 %)	79.5000 % (> 78.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ascorbic acid (R0007)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Glutathione (reduced form) (R0032)	90.12	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ascorbic acid (R0007)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85253	85253	0.00	90.48
85463	85463	0.00	91.92
85270	85270	0.00	91.14
85277	85277	0.00	92.15
85841	85841	0.00	119.33
85197	85197	0.00	90.12
85352	85352	0.00	90.66

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Beta carotene 10% (R0237)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85436-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Beta carotene 10% (R0237)

Special notes

When selecting the *Beta carotene 10% (R0237)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Beta carotene 10% (R0237)	1	1	0

Second-stage model

For differentiation of the substance/substance group *Beta carotene 10% (R0237)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Beta carotene 10% (R0237)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Beta carotene 10...	16000748	85436	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 40 spectra of 1 reference samples from the substance/substance group *Beta carotene 10% (R0237)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 370 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *Beta carotene 10% (R0237)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Beta carotene 10% (R0237)	16001028	85437	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Beta carotene 10% (R0237)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Beta carotene 10% (R0237)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Beta carotene 10% (R0237)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	40	0	13 370
Type B	0	17	23	10 049
Type C	0	0	0	3

The substance/substance group *Beta carotene 10% (R0237)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9317 %)	100.0000 % (> 85.0000 %)
Type B	100.0000 % (> 99.9016 %)	42.5000 % (> 35.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Beta carotene 10% (R0237)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Lutein 20% (R0052)	19.67	—
Ginseng extract (R0029)	24.60	—
Retinyl palmitate (vitamin A palmitate...	32.89	—
Ginger extract (R0199)	37.55	—
Sunflower oil powder 25 (R0319)	39.13	—
Menaquinone-7 0.2% (R0134)	43.47	—
Instant yeast (dry yeast) (R0112)	56.66	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Beta carotene 10% (R0237)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85436	85436	0.00	19.67

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Beta carotene 20% (R0008)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85216-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Beta carotene 20% (R0008)

Special notes

When selecting the *Beta carotene 20% (R0008)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Beta carotene 20% (R0008)	2	1	0

Second-stage model

For differentiation of the substance/substance group *Beta carotene 20% (R0008)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Beta carotene 20% (R0008)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Beta carotene 20...	14001398	85216	60	not required
Hepart	Beta carotene 20...	15001163	85312	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 120 spectra of 2 reference samples from the substance/substance group *Beta carotene 20% (R0008)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 290 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 60 spectra of 1 reference samples from the substance/substance group *Beta carotene 20% (R0008)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Beta carotene 20% (R0008)	14001399	85218	60

- 10 029 spectra from a total of 236 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Beta carotene 20% (R0008)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Beta carotene 20% (R0008)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Beta carotene 20% (R0008)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	120	0	13 290
Type B	0	60	0	10 029
Type C	0	0	0	3

The substance/substance group *Beta carotene 20% (R0008)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9307 %)	100.0000 % (> 95.0000 %)
Type B	100.0000 % (> 99.9011 %)	100.0000 % (> 90.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Beta carotene 20% (R0008)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Pea protein (R0092)	25.35	—
Sunflower oil powder 25 (R0319)	36.55	—
L-glutamine (R0031)	45.33	—
Calcium D panthothenate (R0012)	47.17	—
Isoflavones 10% (R0042)	48.09	—
Cyanocobalamin (R0024)	62.76	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Beta carotene 20% (R0008)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85216	85216	0.00	25.35
85312	85312	0.00	34.88

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **beta-glucan 28% (oat fibre) (R0334)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 10002317-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

beta-glucan 28% (oat fibre) (R0334)

Special notes

When selecting the *beta-glucan 28% (oat fibre) (R0334)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
beta-glucan 28% (oat fibre) (R0334)	1	1	0

Second-stage model

For differentiation of the substance/substance group *beta-glucan 28% (oat fibre) (R0334)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *beta-glucan 28% (oat fibre) (R0334)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	beta-glucan 28% ...	20001740	85859	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 40 spectra of 1 reference samples from the substance/substance group *beta-glucan 28% (oat fibre) (R0334)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 370 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *beta-glucan 28% (oat fibre) (R0334)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	beta-glucan 28% (oat fibre) ...	19001853	85622	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *beta-glucan 28% (oat fibre) (R0334)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *beta-glucan 28% (oat fibre) (R0334)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *beta-glucan 28% (oat fibre) (R0334)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	40	0	13 370
Type B	0	40	0	10 049
Type C	0	0	0	3

The substance/substance group *beta-glucan 28% (oat fibre) (R0334)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9317 %)	100.0000 % (> 85.0000 %)
Type B	100.0000 % (> 99.9016 %)	100.0000 % (> 85.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *beta-glucan 28% (oat fibre) (R0334)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Acacia gum (encapsia) / acacia gum (R0...	34.58	—
Magnesium glycerophosphate (R0114)	37.12	—
Sodium chondroitin sulfate (R0061)	42.72	—
Lutein 20% (R0052)	45.03	—
Calcium lactate pentahydrate (R0013)	51.69	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *beta-glucan 28% (oat fibre) (R0334)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85859	85859	0.00	42.72

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Betaine HCl (R0211)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85283-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Betaine HCl (R0211)

Special notes

When selecting the *Betaine HCl (R0211)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Betaine HCl (R0211)	2	2	0

Second-stage model

For differentiation of the substance/substance group *Betaine HCl (R0211)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Betaine HCl (R0211)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Betaine HCl (R02...	15000746	85283	60	not required
Hepart	Betaine HCl (R02...	17000805	85492	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 100 spectra of 2 reference samples from the substance/substance group *Betaine HCl (R0211)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 310 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 2 reference samples from the substance/substance group *Betaine HCl (R0211)*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Betaine HCl (R0211)	16000706	85370	40
Hepart	Betaine HCl (R0211)	18001363	85571	40

- 10 009 spectra from a total of 235 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Betaine HCl (R0211)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Betaine HCl (R0211)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Betaine HCl (R0211)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	100	0	13 310
Type B	0	80	0	10 009
Type C	0	0	0	3

The substance/substance group *Betaine HCl (R0211)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9308 %)	100.0000 % (> 94.0000 %)
Type B	100.0000 % (> 99.9009 %)	100.0000 % (> 92.5000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Betaine HCl (R0211)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-proline (R0070)	165.00	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Betaine HCl (R0211)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85283	85283	0.00	165.00
85492	85492	0.00	175.29

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Biotin (R0010)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85048-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Biotin (R0010)

Special notes

When selecting the *Biotin (R0010)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Biotin (R0010)	1	1	0

Second-stage model

For differentiation of the substance/substance group *Biotin (R0010)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Biotin (R0010)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Biotin (R0010)	15000241	85255	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 60 spectra of 1 reference samples from the substance/substance group *Biotin (R0010)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 350 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *Biotin (R0010)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Biotin (R0010)	14000292	85179	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Biotin (R0010)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Biotin (R0010)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Biotin (R0010)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	60	0	13 350
Type B	0	39	1	10 049
Type C	0	0	0	3

The substance/substance group *Biotin (R0010)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9312 %)	100.0000 % (> 90.0000 %)
Type B	100.0000 % (> 99.9016 %)	97.5000 % (> 90.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Biotin (R0010)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Zinc citrate trihydrate (R0089)	115.91	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Biotin (R0010)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85255	85255	0.00	124.12

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Black cumin extract (R0193)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85278-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Black cumin extract (R0193)

Special notes

When selecting the *Black cumin extract (R0193)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Black cumin extract (R0193)	2	0	0

Second-stage model

For differentiation of the substance/substance group *Black cumin extract (R0193)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Black cumin extract (R0193)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Black cumin extr...	15000616	85278	60	not required
Hepart	Black cumin extr...	16000595	85385	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 100 spectra of 2 reference samples from the substance/substance group *Black cumin extract (R0193)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 310 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Black cumin extract (R0193)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Black cumin extract (R0193)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Black cumin extract (R0193)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Black cumin extract (R0193)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	100	0	13 310
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Black cumin extract (R0193)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9308 %)	100.0000 % (> 94.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Black cumin extract (R0193)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Hyssop extract (R0205)	11.23	—
Oregano extract (R0206)	14.69	—
Coriolus extract (R0207)	20.61	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Garcinia Cambogia (R0028)	27.76	—
Chili extract (R0198)	30.11	—
Citrus bioflavonoids 60% (R0021)	37.69	—
Magnesium glycerophosphate (R0114)	37.99	—
Artichoke extract (R0208)	40.76	—
Hydrangea extract (R0203)	41.93	—
Calmegh extract (R0201)	42.06	—
Osha root extract (R0220)	42.26	—
Walnut leaf extract (R0197)	43.81	—
Calcium lactate pentahydrate (R0013)	49.15	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Black cumin extract (R0193)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85385	85385	0.00	11.23
85278	85278	0.00	20.61

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group	Calcium caseinate (Lactonate RCE) (R0239)
Substance class	Hepart raw materials
Report date	22/03/2021
Report number	10004600-2021-03-22
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Calcium caseinate (Lactonate RCE) (R0239)

Special notes

When selecting the *Calcium caseinate (Lactonate RCE) (R0239)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Calcium caseinate (Lactonate RCE) (R0239)	1	0	0

Second-stage model

For differentiation of the substance/substance group *Calcium caseinate (Lactonate RCE) (R0239)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Calcium caseinate (Lactonate RCE) (R0239)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Calcium caseinat...	20002505	85870	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 40 spectra of 1 reference samples from the substance/substance group *Calcium caseinate (Lactonate RCE) (R0239)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 370 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Calcium caseinate (Lactonate RCE) (R0239)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Calcium caseinate (Lactonate RCE) (R0239)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Calcium caseinate (Lactonate RCE) (R0239)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Calcium caseinate (Lactonate RCE) (R0239)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	40	0	13 370
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Calcium caseinate (Lactonate RCE) (R0239)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9317 %)	100.0000 % (> 85.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Calcium caseinate (Lactonate RCE) (R0239)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-glutamine (R0031)	29.08	—
Cyanocobalamin (R0024)	34.86	—
Methylcobalamin (R0190)	38.09	—
Beta carotene 20% (R0008)	43.98	—
Pea protein (R0092)	53.65	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Calcium caseinate (Lactonate RCE) (R0239)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85870	85870	0.00	34.86

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group	Calcium D panthothenate (R0012)
Substance class	Hepart raw materials
Report date	22/03/2021
Report number	85000-2021-03-22
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Calcium D panthothenate (R0012)

Special notes

When selecting the *Calcium D panthothenate (R0012)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	<i>European Pharmacopoeia 10th Edition, Basic Version 2020</i> [3]
Komm2.2.40	<i>Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident</i> [4]
AA004	<i>Erstellung und Validierung eines IdentModul-Updates</i>

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Calcium D panthothenate (R0012)	2	2	0

Second-stage model

For differentiation of the substance/substance group *Calcium D panthothenate (R0012)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Calcium D panthothenate (R0012)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Calcium D pantho...	16000746	85375	40	not required
Hepart	Calcium D pantho...	18001795	85579	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Calcium D panthothenate (R0012)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 330 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 2 reference samples from the substance/substance group *Calcium D panthothenate (R0012)*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Calcium D panthothenate (R00...	17001023	85494	40
Hepart	Calcium D panthothenate (R00...	19001514	85601	40

- 10 009 spectra from a total of 235 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Calcium D panthothenate (R0012)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Calcium D panthothenate (R0012)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Calcium D panthothenate (R0012)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	13 330
Type B	0	80	0	10 009
Type C	0	0	0	3

The substance/substance group *Calcium D panthothenate (R0012)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9309 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9009 %)	100.0000 % (> 92.5000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Calcium D panthothenate (R0012)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Beta carotene 20% (R0008)	33.52	—
Methylcobalamin (R0190)	59.58	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Calcium D panthothenate (R0012)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85375	85375	0.00	42.74
85579	85579	0.00	40.27

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Calcium lactate pentahydrate (R0013)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85234-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Calcium lactate pentahydrate (R0013)

Special notes

When selecting the *Calcium lactate pentahydrate (R0013)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Calcium lactate pentahydrate (R0013)	3	4	0

Second-stage model

For differentiation of the substance/substance group *Calcium lactate pentahydrate (R0013)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Calcium lactate pentahydrate (R0013)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Calcium lactate ...	14001909	85234	60	not required
Hepart	Calcium lactate ...	15000242/0	85265	60	not required
Hepart	Calcium lactate ...	15000242/1	85266	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 180 spectra of 3 reference samples from the substance/substance group *Calcium lactate pentahydrate (R0013)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 13 230 spectra from a total of 270 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 180 spectra of 4 reference samples from the substance/substance group *Calcium lactate pentahydrate (R0013)*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Calcium lactate pentahydrate...	14001908	85235	60

continued on the next page

continued from previous page

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Calcium lactate pentahydrate...	18000930	85534	40
Hepart	Calcium lactate pentahydrate...	18000931	85535	40
Hepart	Calcium lactate pentahydrate...	19001229	85602	40

- 9909 spectra from a total of 233 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Calcium lactate pentahydrate (R0013)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Calcium lactate pentahydrate (R0013)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Calcium lactate pentahydrate (R0013)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	180	0	13 230
Type B	0	180	0	9909
Type C	0	0	0	3

The substance/substance group *Calcium lactate pentahydrate (R0013)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9305 %)	100.0000 % (> 96.6667 %)
Type B	100.0000 % (> 99.9005 %)	100.0000 % (> 96.6667 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Calcium lactate pentahydrate (R0013)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Garcinia Cambogia (R0028)	37.58	—
Magnesium glycerophosphate (R0114)	69.39	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Calcium lactate pentahydrate (R0013)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85234	85234	0.00	37.58
85265	85265	0.00	54.01
85266	85266	0.00	52.82

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group	Calmegh extract (R0201)
Substance class	Hepart raw materials
Report date	22/03/2021
Report number	85300-2021-03-22
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Calmegh extract (R0201)

Special notes

When selecting the *Calmegh extract (R0201)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Calmegh extract (R0201)	2	0	0

Second-stage model

For differentiation of the substance/substance group *Calmegh extract (R0201)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Calmegh extract (R0201)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Calmegh extract ...	15000785	85300	60	not required
Hepart	Calmegh extract ...	16000716	85398	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 100 spectra of 2 reference samples from the substance/substance group *Calmegh extract (R0201)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 310 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Calmegh extract (R0201)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Calmegh extract (R0201)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Calmegh extract (R0201)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Calmegh extract (R0201)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	100	0	13 310
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Calmegh extract (R0201)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9308 %)	100.0000 % (> 94.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Calmegh extract (R0201)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Calcium lactate pentahydrate (R0013)	20.25	—
Hydrangea extract (R0203)	22.03	—
Chili extract (R0198)	22.49	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Magnesium glycerophosphate (R0114)	30.32	—
Garcinia Cambogia (R0028)	34.13	—
Citrus bioflavonoids 60% (R0021)	36.01	—
Coriolus extract (R0207)	37.71	—
Lycopene 5% extract (R0053)	41.04	—
Black cumin extract (R0193)	41.90	—
Ginseng extract (R0029)	42.27	—
Osha root extract (R0220)	44.00	—
Juniper extract (R0192)	44.30	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Calmegh extract (R0201)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85300	85300	0.00	22.03
85398	85398	0.00	27.49

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50% greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Chaga extract (R0016)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85219-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Chaga extract (R0016)

Special notes

When selecting the *Chaga extract (R0016)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Chaga extract (R0016)	1	1	0

Second-stage model

For differentiation of the substance/substance group *Chaga extract (R0016)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Chaga extract (R0016)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Chaga extract (R...	18000831	85532	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 40 spectra of 1 reference samples from the substance/substance group *Chaga extract (R0016)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 370 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *Chaga extract (R0016)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Chaga extract (R0016)	16000964	85384	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Chaga extract (R0016)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Chaga extract (R0016)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Chaga extract (R0016)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	40	0	13 370
Type B	0	0	40	10 049
Type C	0	0	0	3

The substance/substance group *Chaga extract (R0016)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9317 %)	100.0000 % (> 85.0000 %)
Type B	100.0000 % (> 99.9016 %)	0.0000 % (≥ 0.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Chaga extract (R0016)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Garcinia Cambogia (R0028)	22.26	—
Calcium lactate pentahydrate (R0013)	37.58	—
Walnut leaf extract (R0197)	43.40	—
Artichoke extract (R0208)	58.25	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Chaga extract (R0016)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85532	85532	0.00	22.26

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Chili extract (R0198)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85291-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Chili extract (R0198)

Special notes

When selecting the *Chili extract (R0198)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Chili extract (R0198)	1	1	0

Second-stage model

For differentiation of the substance/substance group *Chili extract (R0198)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Chili extract (R0198)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Chili extract (R...	15000788	85291	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 60 spectra of 1 reference samples from the substance/substance group *Chili extract (R0198)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 350 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *Chili extract (R0198)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Chili extract (R0198)	16000713	85367	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Chili extract (R0198)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Chili extract (R0198)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Chili extract (R0198)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	60	0	13 350
Type B	0	0	40	10 049
Type C	0	0	0	3

The substance/substance group *Chili extract (R0198)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9312 %)	100.0000 % (> 90.0000 %)
Type B	100.0000 % (> 99.9016 %)	0.0000 % (≥ 0.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Chili extract (R0198)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Calmegh extract (R0201)	21.34	—
Hydrangea extract (R0203)	25.48	—
Coriolus extract (R0207)	28.51	—
Juniper extract (R0192)	28.80	—
Magnesium glycerophosphate (R0114)	31.64	—
Garcinia Cambogia (R0028)	41.60	—
Osha root extract (R0220)	41.65	—
Ginseng extract (R0029)	43.32	—
Calcium lactate pentahydrate (R0013)	47.08	—
Oregano extract (R0206)	53.38	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Chili extract (R0198)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85291	85291	0.00	21.34

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Cholecalciferol (vitamin D3) (R0018) / Menaquinone-7 0.2% (R0134)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85256-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Cholecalciferol (vitamin D3) (R0018) / Menaquinone-7 0.2% (R0134); Cholecalciferol (vitamin D3) (R0018); Menaquinone-7 0.2% (R0134)

Special notes

When selecting the *Cholecalciferol (vitamin D3) (R0018) / Menaquinone-7 0.2% (R0134)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Cholecalciferol (vitamin D3) (R0018)	3	3	0
Menaquinone-7 0.2% (R0134)	4	1	0

Second-stage model

For differentiation of the substance/substance group *Cholecalciferol (vitamin D3) (R0018) / Menaquinone-7 0.2% (R0134)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Cholecalciferol (vitamin D3) (R0018) / Menaquinone-7 0.2% (R0134)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Cholecalciferol ...	15000387	85256	60	not required
Hepart	Cholecalciferol ...	15001075	85310	60	not required
Hepart	Cholecalciferol ...	17001099	85497	40	not required
Hepart	Menaquinone-7 0...	15000367	85280	60	not required
Hepart	Menaquinone-7 0...	16001501	85454	40	not required
Hepart	Menaquinone-7 0...	17000959	85487	40	not required
Hepart	Menaquinone-7 0...	2000179	85865	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 340 spectra of 7 reference samples from the substance/substance group *Cholecalciferol (vitamin D3) (R0018) / Menaquinone-7 0.2% (R0134)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 7 different batches.
- 13 070 spectra from a total of 266 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 160 spectra of 4 reference samples from the substance/substance group *Cholecalciferol (vitamin D3) (R0018) / Menaquinone-7 0.2% (R0134)*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Cholecalciferol (vitamin D3)...	16000539	85389	40
Hepart	Cholecalciferol (vitamin D3)...	18001183	85542	40
Hepart	Cholecalciferol (vitamin D3)...	19001483	85600	40
Hepart	Menaquinone-7 0.2% (R0134)	16000527	85386	40

- 9929 spectra from a total of 233 batches from further 104 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Cholecalciferol (vitamin D3) (R0018) / Menaquinone-7 0.2% (R0134)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Cholecalciferol (vitamin D3) (R0018) / Menaquinone-7 0.2% (R0134)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Cholecalciferol (vitamin D3) (R0018) / Menaquinone-7 0.2% (R0134)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	340	0	13 070
Type B	0	160	0	9929
Type C	0	0	0	3

The substance/substance group *Cholecalciferol (vitamin D3) (R0018) / Menaquinone-7 0.2% (R0134)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9303 %)	100.0000 % (> 98.2353 %)
Type B	100.0000 % (> 99.9005 %)	100.0000 % (> 96.2500 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Cholecalciferol (vitamin D3) (R0018) / Menaquinone-7 0.2% (R0134)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Retinyl palmitate (vitamin A palmitate...	13.04	—
Acacia gum (encapsia) / acacia gum (R0...	20.13	—
Ginger extract (R0199)	22.19	—
Sunflower oil powder 25 (R0319)	24.36	—
Galactomannans (R0035/R0094)	30.07	—
Potato starch, pregelatinized (R0047)	36.83	—
Inulin (R0041)	41.67	—
beta-glucan 28% (oat fibre) (R0334)	42.08	—
Gentian extract (R0200)	43.77	—
Lutein 20% (R0052)	44.17	—
Juniper extract (R0192)	44.25	—
Calcium lactate pentahydrate (R0013)	49.88	—
Ginseng extract (R0029)	59.50	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Cholecalciferol (vitamin D3) (R0018) / Menaquinone-7 0.2% (R0134)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85497	85497	0.00	27.68
85280	85280	0.00	32.53
85256	85256	0.00	24.72
85487	85487	0.00	13.04
85454	85454	0.00	145.44
85865	85865	0.00	62.30
85310	85310	0.00	25.73

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical

variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group	Choline chloride (R0019)
Substance class	Hepart raw materials
Report date	22/03/2021
Report number	85014-2021-03-22
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Choline chloride (R0019)

Special notes

When selecting the *Choline chloride (R0019)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Choline chloride (R0019)	3	4	0

Second-stage model

For differentiation of the substance/substance group *Choline chloride (R0019)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Choline chloride (R0019)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Choline chloride...	1209009	85110	40	not required
Hepart	Choline chloride...	15000266	85263	60	not required
Hepart	Choline chloride...	16000866	85413	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 140 spectra of 3 reference samples from the substance/substance group *Choline chloride (R0019)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 13 270 spectra from a total of 270 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 180 spectra of 4 reference samples from the substance/substance group *Choline chloride (R0019)*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Choline chloride (R0019)	16000187	85356	60
Hepart	Choline chloride (R0019)	16001694	85464	40
Hepart	Choline chloride (R0019)	18001224	85543	40
Hepart	Choline chloride (R0019)	19002133	85822	40

- 9909 spectra from a total of 233 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Choline chloride (R0019)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Choline chloride (R0019)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Choline chloride (R0019)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	140	0	13 270
Type B	0	180	0	9909
Type C	0	0	0	3

The substance/substance group *Choline chloride (R0019)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9306 %)	100.0000 % (> 95.7143 %)
Type B	100.0000 % (> 99.9005 %)	100.0000 % (> 96.6667 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Choline chloride (R0019)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Taurine (R0078)	55.81	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Choline chloride (R0019)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85263	85263	0.00	47.21
85110	85110	0.00	93.10
85413	85413	0.00	97.48

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Citric acid monohydrate (R0143)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85304-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Citric acid monohydrate (R0143)

Special notes

When selecting the *Citric acid monohydrate (R0143)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Citric acid monohydrate (R0143)	1	0	0

Second-stage model

For differentiation of the substance/substance group *Citric acid monohydrate (R0143)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Citric acid monohydrate (R0143)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Citric acid mono...	15000903	85304	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 60 spectra of 1 reference samples from the substance/substance group *Citric acid monohydrate (R0143)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 350 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Citric acid monohydrate (R0143)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Citric acid monohydrate (R0143)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Citric acid monohydrate (R0143)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Citric acid monohydrate (R0143)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	60	0	13 350
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Citric acid monohydrate (R0143)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9312 %)	100.0000 % (> 90.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Citric acid monohydrate (R0143)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Green tea extract (R0034)	36.99	—
OPC (grape seed) extract (R0065)	42.10	—
Rhodiola rosea extract 3% (R0159)	55.00	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Citric acid monohydrate (R0143)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85304	85304	0.00	36.99

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Citric acid powder (R0258)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 10004598-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Citric acid powder (R0258)

Special notes

When selecting the *Citric acid powder (R0258)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Citric acid powder (R0258)	1	0	0

Second-stage model

For differentiation of the substance/substance group *Citric acid powder (R0258)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Citric acid powder (R0258)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Citric acid powd...	20002527	85867	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 40 spectra of 1 reference samples from the substance/substance group *Citric acid powder (R0258)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 370 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Citric acid powder (R0258)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Citric acid powder (R0258)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Citric acid powder (R0258)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Citric acid powder (R0258)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	40	0	13 370
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Citric acid powder (R0258)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9317 %)	100.0000 % (> 85.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Citric acid powder (R0258)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
D-galactose (R0135)	55.59	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Citric acid powder (R0258)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85867	85867	0.00	92.33

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Citrus bioflavonoids 60% (R0021)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85004-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Citrus bioflavonoids 60% (R0021)

Special notes

When selecting the *Citrus bioflavonoids 60% (R0021)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Citrus bioflavonoids 60% (R0021)	7	3	0

Second-stage model

For differentiation of the substance/substance group *Citrus bioflavonoids 60% (R0021)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Citrus bioflavonoids 60% (R0021)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Citrus bioflavon...	1208018	85122	40	not required
Hepart	Citrus bioflavon...	14000663	85176	40	not required
Hepart	Citrus bioflavon...	14001194	85206	40	not required
Hepart	Citrus bioflavon...	16001143	85406	40	not required
Hepart	Citrus bioflavon...	17001296	85499	40	not required
Hepart	Citrus bioflavon...	19002111/0	85615	40	not required
Hepart	Citrus bioflavon...	20001463	85860	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 280 spectra of 7 reference samples from the substance/substance group *Citrus bioflavonoids 60% (R0021)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 7 different batches.
- 13 130 spectra from a total of 266 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 3 reference samples from the substance/substance group *Citrus bioflavonoids 60% (R0021)*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Citrus bioflavonoids 60% (R0...	14001176	85207	40
Hepart	Citrus bioflavonoids 60% (R0...	17000142	85473	40
Hepart	Citrus bioflavonoids 60% (R0...	17000958	85491	40

- 9969 spectra from a total of 234 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Citrus bioflavonoids 60% (R0021)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Citrus bioflavonoids 60% (R0021)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Citrus bioflavonoids 60% (R0021)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	280	0	13 130
Type B	0	105	15	9969
Type C	0	0	0	3

The substance/substance group *Citrus bioflavonoids 60% (R0021)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9304 %)	100.0000 % (> 97.8571 %)
Type B	100.0000 % (> 99.9006 %)	87.5000 % (> 85.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Citrus bioflavonoids 60% (R0021)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Sodium borate (R0191)	22.79	—
Garcinia Cambogia (R0028)	33.05	—
Magnesium glycerophosphate (R0114)	43.51	—
Acacia gum (encapsia) / acacia gum (R0...	52.73	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Citrus bioflavonoids 60% (R0021)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85615	85615	0.00	36.68
85860	85860	0.00	32.09
85499	85499	0.00	33.05
85176	85176	0.00	27.82
85206	85206	0.00	26.55
85122	85122	0.00	22.79
85406	85406	0.00	45.22

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group	Coenzyme Q10 (R0022)
Substance class	Hepart raw materials
Report date	22/03/2021
Report number	85063-2021-03-22
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Coenzyme Q10 (R0022)

Special notes

When selecting the *Coenzyme Q10 (R0022)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	<i>European Pharmacopoeia 10th Edition, Basic Version 2020</i> [3]
Komm2.2.40	<i>Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident</i> [4]
AA004	<i>Erstellung und Validierung eines IdentModul-Updates</i>

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Coenzyme Q10 (R0022)	4	4	0

Second-stage model

For differentiation of the substance/substance group *Coenzyme Q10 (R0022)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Coenzyme Q10 (R0022)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Coenzyme Q10 (R0...	1007-018	85063	60	not required
Hepart	Coenzyme Q10 (R0...	14001313	85229	60	not required
Hepart	Coenzyme Q10 (R0...	15001651	85332	60	not required
Hepart	Coenzyme Q10 (R0...	16000193	85337	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 4 reference samples from the substance/substance group *Coenzyme Q10 (R0022)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 4 different batches.
- 13 170 spectra from a total of 269 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 180 spectra of 4 reference samples from the substance/substance group *Coenzyme Q10 (R0022)*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Coenzyme Q10 (R0022)	16000185	85336	60

continued on the next page

continued from previous page

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Coenzyme Q10 (R0022)	16001481	85444	40
Hepart	Coenzyme Q10 (R0022)	17000783	85489	40
Hepart	Coenzyme Q10 (R0022)	19000669/0	85605	40

- 9909 spectra from a total of 233 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Coenzyme Q10 (R0022)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Coenzyme Q10 (R0022)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Coenzyme Q10 (R0022)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	240	0	13 170
Type B	0	180	0	9909
Type C	0	0	0	3

The substance/substance group *Coenzyme Q10 (R0022)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9304 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9005 %)	100.0000 % (> 96.6667 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Coenzyme Q10 (R0022)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Progesterone, micronized (R0108)	36.36	—
a-lipoic acid (R0003)	98.38	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Coenzyme Q10 (R0022)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85063	85063	0.00	44.42
85332	85332	0.00	42.29
85229	85229	0.00	42.55
85337	85337	0.00	41.00

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Coriolus extract (R0207)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85288-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Coriolus extract (R0207)

Special notes

When selecting the *Coriolus extract (R0207)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Coriolus extract (R0207)	2	0	0

Second-stage model

For differentiation of the substance/substance group *Coriolus extract (R0207)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Coriolus extract (R0207)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Coriolus extract...	15000795	85288	60	not required
Hepart	Coriolus extract...	16000722	85363	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 100 spectra of 2 reference samples from the substance/substance group *Coriolus extract (R0207)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 310 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Coriolus extract (R0207)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Coriolus extract (R0207)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Coriolus extract (R0207)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Coriolus extract (R0207)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	100	0	13 310
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Coriolus extract (R0207)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9308 %)	100.0000 % (> 94.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Coriolus extract (R0207)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Chili extract (R0198)	17.83	—
Hyssop extract (R0205)	19.13	—
Hydrangea extract (R0203)	20.14	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Black cumin extract (R0193)	23.99	—
Oregano extract (R0206)	29.77	—
Artichoke extract (R0208)	31.47	—
Calmegh extract (R0201)	32.49	—
Garcinia Cambogia (R0028)	32.74	—
Magnesium glycerophosphate (R0114)	35.61	—
Juniper extract (R0192)	38.07	—
Osha root extract (R0220)	39.26	—
Katuka herb extract (R0210)	43.73	—
Calcium lactate pentahydrate (R0013)	48.67	—
Citrus bioflavonoids 60% (R0021)	48.83	—
Ginseng extract (R0029)	50.40	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Coriolus extract (R0207)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85363	85363	0.00	19.13
85288	85288	0.00	17.83

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Creatine monohydrate (R0048)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85067-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Creatine monohydrate (R0048)

Special notes

When selecting the *Creatine monohydrate (R0048)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Creatine monohydrate (R0048)	3	2	0

Second-stage model

For differentiation of the substance/substance group *Creatine monohydrate (R0048)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Creatine monohydrate (R0048)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Creatine monohyd...	060701	85067	60	not required
Hepart	Creatine monohyd...	17000597	85504	40	not required
Hepart	Creatine monohyd...	18000867	85525	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 140 spectra of 3 reference samples from the substance/substance group *Creatine monohydrate (R0048)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 13 270 spectra from a total of 270 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 2 reference samples from the substance/substance group *Creatine monohydrate (R0048)*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Creatine monohydrate (R0048)	18000868	85526	40
Hepart	Creatine monohydrate (R0048)	18000869	85527	40

- 10 009 spectra from a total of 235 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Creatine monohydrate (R0048)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Creatine monohydrate (R0048)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Creatine monohydrate (R0048)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	140	0	13 270
Type B	0	80	0	10 009
Type C	0	0	0	3

The substance/substance group *Creatine monohydrate (R0048)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9306 %)	100.0000 % (> 95.7143 %)
Type B	100.0000 % (> 99.9009 %)	100.0000 % (> 92.5000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Creatine monohydrate (R0048)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
D-galactose (R0135)	39.94	–
Tripotassium citrate monohydrate (R0117)	55.07	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Creatine monohydrate (R0048)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85504	85504	0.00	54.83
85525	85525	0.00	47.27
85067	85067	0.00	48.00

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group	Curcumin 95% extract (R0023)
Substance class	Hepart raw materials
Report date	22/03/2021
Report number	85191-2021-03-22
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Curcumin 95% extract (R0023)

Special notes

When selecting the *Curcumin 95% extract (R0023)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	<i>European Pharmacopoeia 10th Edition, Basic Version 2020</i> [3]
Komm2.2.40	<i>Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident</i> [4]
AA004	<i>Erstellung und Validierung eines IdentModul-Updates</i>

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Curcumin 95% extract (R0023)	2	2	0

Second-stage model

For differentiation of the substance/substance group *Curcumin 95% extract (R0023)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Curcumin 95% extract (R0023)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Curcumin 95% ext...	14001177	85214	60	not required
Hepart	Curcumin 95% ext...	16000741	85405	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 100 spectra of 2 reference samples from the substance/substance group *Curcumin 95% extract (R0023)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 310 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 94 spectra of 2 reference samples from the substance/substance group *Curcumin 95% extract (R0023)*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Curcumin 95% extract (R0023)	130816500	85191	34
Hepart	Curcumin 95% extract (R0023)	16000422	85348	60

- 9995 spectra from a total of 235 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Curcumin 95% extract (R0023)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Curcumin 95% extract (R0023)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Curcumin 95% extract (R0023)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	100	0	13 310
Type B	0	94	0	9995
Type C	0	0	0	3

The substance/substance group *Curcumin 95% extract (R0023)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9308 %)	100.0000 % (> 94.0000 %)
Type B	100.0000 % (> 99.9008 %)	100.0000 % (> 93.6170 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Curcumin 95% extract (R0023)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-5-hydroxytryptophan (R0050)	128.48	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Curcumin 95% extract (R0023)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85405	85405	0.00	138.22
85214	85214	0.00	128.48

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Cyanocobalamin (R0024) / Methylcobalamin (R0190)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85274-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Cyanocobalamin (R0024) / Methylcobalamin (R0190); Cyanocobalamin (R0024); Methylcobalamin (R0190)

Special notes

When selecting the *Cyanocobalamin (R0024) / Methylcobalamin (R0190)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Cyanocobalamin (R0024)	2	1	0
Methylcobalamin (R0190)	2	1	0

Second-stage model

For differentiation of the substance/substance group *Cyanocobalamin (R0024)* / *Methylcobalamin (R0190)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Cyanocobalamin (R0024)* / *Methylcobalamin (R0190)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Cyanocobalamin (...)	15000619	85274	60	not required
Hepart	Cyanocobalamin (...)	17000988	85488	40	not required
Hepart	Methylcobalamin ...	15000620	85302	60	not required
Hepart	Methylcobalamin ...	16000704	85362	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 200 spectra of 4 reference samples from the substance/substance group *Cyanocobalamin (R0024)* / *Methylcobalamin (R0190)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 4 different batches.
- 13 210 spectra from a total of 269 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 2 reference samples from the substance/substance group *Cyanocobalamin (R0024)* / *Methylcobalamin (R0190)*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Cyanocobalamin (R0024)	16001695	85458	40
Hepart	Methylcobalamin (R0190)	16001701	85465	40

- 10 009 spectra from a total of 235 batches from further 104 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Cyanocobalamin (R0024) / Methylcobalamin (R0190)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Cyanocobalamin (R0024) / Methylcobalamin (R0190)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Cyanocobalamin (R0024) / Methylcobalamin (R0190)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	200	0	13 210
Type B	0	55	25	10 009
Type C	0	0	0	3

The substance/substance group *Cyanocobalamin (R0024) / Methylcobalamin (R0190)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9305 %)	100.0000 % (> 97.0000 %)
Type B	100.0000 % (> 99.9009 %)	68.7500 % (> 65.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Cyanocobalamin (R0024) / Methylcobalamin (R0190)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-glutamine (R0031)	35.85	—
Calcium caseinate (Lactonate RCE) (R02...	53.76	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Cyanocobalamin (R0024) / Methylcobalamin (R0190)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85362	85362	0.00	96.33
85302	85302	0.00	54.37
85488	85488	0.00	53.76
85274	85274	0.00	54.75

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **D-alpha tocopheryl acetate 50% dry CWS (R0081)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85317-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

D-alpha tocopheryl acetate 50% dry CWS (R0081)

Special notes

When selecting the *D-alpha tocopheryl acetate 50% dry CWS (R0081)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
D-alpha tocopheryl acetate 50% dry CWS (R0081)	2	0	0

Second-stage model

For differentiation of the substance/substance group *D-alpha tocopheryl acetate 50% dry CWS (R0081)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *D-alpha tocopheryl acetate 50% dry CWS (R0081)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	D-alpha tocopher...	15001162	85317	60	not required
Hepart	D-alpha tocopher...	16001030	85414	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 100 spectra of 2 reference samples from the substance/substance group *D-alpha tocopheryl acetate 50% dry CWS (R0081)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 310 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *D-alpha tocopheryl acetate 50% dry CWS (R0081)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *D-alpha tocopheryl acetate 50% dry CWS (R0081)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *D-alpha tocopheryl acetate 50% dry CWS (R0081)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *D-alpha tocopheryl acetate 50% dry CWS (R0081)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	100	0	13 310
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *D-alpha tocopheryl acetate 50% dry CWS (R0081)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9308 %)	100.0000 % (> 94.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *D-alpha tocopheryl acetate 50% dry CWS (R0081)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-glutamine (R0031)	90.63	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *D-alpha tocopheryl acetate 50% dry CWS (R0081)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85317	85317	0.00	157.98
85414	85414	0.00	143.61

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **D-galactose (R0135)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85096-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

D-galactose (R0135)

Special notes

When selecting the *D-galactose (R0135)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
D-galactose (R0135)	2	0	0

Second-stage model

For differentiation of the substance/substance group *D-galactose (R0135)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *D-galactose (R0135)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	D-galactose (R01...	061201	85096	40	not required
Hepart	D-galactose (R01...	18001111	85540	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *D-galactose (R0135)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 330 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *D-galactose (R0135)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group D-galactose (R0135).
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group D-galactose (R0135) can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with D-galactose (R0135) and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	13 330
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group D-galactose (R0135) can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9309 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group D-galactose (R0135) in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Creatine monohydrate (R0048)	64.85	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group D-*galactose (R0135)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85096	85096	0.00	69.20
85540	85540	0.00	68.27

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **DL-methionine (R0209)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85303-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

DL-methionine (R0209)

Special notes

When selecting the *DL-methionine (R0209)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
DL-methionine (R0209)	1	0	0

Second-stage model

For differentiation of the substance/substance group DL-methionine (R0209) the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group DL-methionine (R0209):

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	DL-methionine (R...	15000799	85303	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 60 spectra of 1 reference samples from the substance/substance group DL-methionine (R0209). These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 350 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group DL-methionine (R0209).
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group DL-methionine (R0209).
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group DL-methionine (R0209) can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with DL-methionine (R0209) and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	60	0	13 350
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group DL-methionine (R0209) can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9312 %)	100.0000 % (> 90.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group DL-methionine (R0209) in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-leucine (R0051)	60.33	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *DL-methionine (R0209)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85303	85303	0.00	60.33

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **D-ribose (R0137)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85104-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

D-ribose (R0137)

Special notes

When selecting the *D-ribose (R0137)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
D-ribose (R0137)	2	4	0

Second-stage model

For differentiation of the substance/substance group D-ribose (R0137) the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group D-ribose (R0137):

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	D-ribose (R0137)	061001	85104	40	not required
Hepart	D-ribose (R0137)	17001673	85507	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group D-ribose (R0137). These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 330 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 180 spectra of 4 reference samples from the substance/substance group D-ribose (R0137).
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	D-ribose (R0137)	15001757	85329	60
Hepart	D-ribose (R0137)	17001672	85506	40
Hepart	D-ribose (R0137)	18000287	85517	40
Hepart	D-ribose (R0137)	20000529/0	85839	40

- 9909 spectra from a total of 233 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *D-ribose (R0137)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *D-ribose (R0137)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *D-ribose (R0137)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	13 330
Type B	0	180	0	9909
Type C	0	0	0	3

The substance/substance group *D-ribose (R0137)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9309 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9005 %)	100.0000 % (> 96.6667 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group D-ribose (*R0137*) in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Citrus bioflavonoids 60% (R0021)	51.11	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *D-ribose (R0137)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85104	85104	0.00	51.11
85507	85507	0.00	53.21

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50% greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Galactomannans (R0035/R0094) / Potato starch, pregelatinized**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85105-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Galactomannans (R0035/R0094) / Potato starch, pregelatinized (R0047); Galactomannans (R0035/R0094);
Potato starch, pregelatinized (R0047)

Special notes

When selecting the *Galactomannans (R0035/R0094) / Potato starch, pregelatinized (R0047)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Galactomannans (R0035/R0094)	17	23	0
Potato starch, pregelatinized (R0047)	2	3	0

Second-stage model

For differentiation of the substance/substance group *Galactomannans (R0035/R0094) / Potato starch, pregelatinized (R0047)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Galactomannans (R0035/R0094) / Potato starch, pregelatinized (R0047)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Galactomannans (...)	1208006	85105	40	not required
Hepart	Galactomannans (...)	1211002	85107	40	not required
Hepart	Galactomannans (...)	1208002	85115	40	not required
Hepart	Galactomannans (...)	1208007	85120	40	not required
Hepart	Galactomannans (...)	14000084	85174	40	not required
Hepart	Galactomannans (...)	15000407	85259	60	not required
Hepart	Galactomannans (...)	15000685	85286	60	not required
Hepart	Galactomannans (...)	15000673	85287	60	not required
Hepart	Galactomannans (...)	15001293	85307	60	not required
Hepart	Galactomannans (...)	15001463	85319	60	not required
Hepart	Galactomannans (...)	15001362	85320	60	not required
Hepart	Galactomannans (...)	15001581	85326	60	not required
Hepart	Galactomannans (...)	15001729	85330	60	not required
Hepart	Galactomannans (...)	15001683	85331	60	not required
Hepart	Galactomannans (...)	16000652	85380	40	not required
Hepart	Galactomannans (...)	16000775	85418	40	not required
Hepart	Galactomannans (...)	16001114	85420	40	not required
Hepart	Potato starch, p...	14001694	85231	60	not required
Hepart	Potato starch, p...	17000989	85490	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 960 spectra of 19 reference samples from the substance/substance group *Galactomannans (R0035/R0094) / Potato starch, pregelatinized (R0047)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 19 different batches.
- 12 450 spectra from a total of 254 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples.

Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 1040 spectra of 26 reference samples from the substance/substance group *Galactomannans (R0035/R0094)* / *Potato starch, pregelatinized (R0047)*.
- Among them are spectra of independent samples from 26 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Galactomannans (R0035/R0094)	16000994	85419	40
Hepart	Galactomannans (R0035/R0094)	16001195	85424	40
Hepart	Galactomannans (R0035/R0094)	16001131	85425	40
Hepart	Galactomannans (R0035/R0094)	16001405	85438	40
Hepart	Galactomannans (R0035/R0094)	16001579	85450	40
Hepart	Galactomannans (R0035/R0094)	16001510	85455	40
Hepart	Galactomannans (R0035/R0094)	16001642	85456	40
Hepart	Galactomannans (R0035/R0094)	17001248	85502	40
Hepart	Galactomannans (R0035/R0094)	17001725	85511	40
Hepart	Galactomannans (R0035/R0094)	17001727	85512	40
Hepart	Galactomannans (R0035/R0094)	18000493	85522	40
Hepart	Galactomannans (R0035/R0094)	18000924	85536	40
Hepart	Galactomannans (R0035/R0094)	18001511	85575	40
Hepart	Galactomannans (R0035/R0094)	18002050	85576	40
Hepart	Galactomannans (R0035/R0094)	18001708	85577	40
Hepart	Galactomannans (R0035/R0094)	19000514/0	85597	40
Hepart	Galactomannans (R0035/R0094)	19001138/0	85598	40
Hepart	Galactomannans (R0035/R0094)	19002029/1	85612	40
Hepart	Galactomannans (R0035/R0094)	19002029/0	85614	40
Hepart	Galactomannans (R0035/R0094)	19001996/0	85616	40
Hepart	Galactomannans (R0035/R0094)	19002539	85826	40
Hepart	Galactomannans (R0035/R0094)	20001155	85861	40
Hepart	Galactomannans (R0035/R0094)	20001467	85862	40
Hepart	Potato starch, pregelatinize...	16000526	85392	40
Hepart	Potato starch, pregelatinize...	18000631	85523	40
Hepart	Potato starch, pregelatinize...	19001623	85610	40

- 9049 spectra from a total of 211 batches from further 104 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Galactomannans (R0035/R0094)* / *Potato starch, pregelatinized (R0047)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within

this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Galactomannans (R0035/R0094) / Potato starch, pregelatinized (R0047)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Galactomannans (R0035/R0094) / Potato starch, pregelatinized (R0047)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	960	0	12 450
Type B	0	1040	0	9049
Type C	0	0	0	3

The substance/substance group *Galactomannans (R0035/R0094) / Potato starch, pregelatinized (R0047)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9302 %)	100.0000 % (> 99.3750 %)
Type B	100.0000 % (> 99.9002 %)	100.0000 % (> 99.4231 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Galactomannans (R0035/R0094) / Potato starch, pregelatinized (R0047)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Acacia gum (encapsia) / acacia gum (R0...	16.01	—
Konjac glucomannan 80% (R0293)	20.92	—
Gentian extract (R0200)	22.10	—
Juniper extract (R0192)	30.67	—
Marshmallow root extract (R0204)	32.57	—
Menaquinone-7 0.2% (R0134)	53.66	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Galactomannans (R0035/R0094) / Potato starch, pregelatinized (R0047)*

is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85490	85490	0.00	31.64
85259	85259	0.00	35.40
85107	85107	0.00	30.14
85331	85331	0.00	36.81
85380	85380	0.00	30.25
85420	85420	0.00	27.30
85105	85105	0.00	30.30
85115	85115	0.00	27.54
85120	85120	0.00	32.12
85286	85286	0.00	32.49
85287	85287	0.00	27.87
85307	85307	0.00	21.01
85320	85320	0.00	31.05
85326	85326	0.00	34.47
85174	85174	0.00	26.00
85330	85330	0.00	22.22
85319	85319	0.00	36.67
85231	85231	0.00	19.07
85418	85418	0.00	31.82

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group	Garcinia Cambogia (R0028)
Substance class	Hepart raw materials
Report date	22/03/2021
Report number	85119-2021-03-22
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Garcinia Cambogia (R0028)

Special notes

When selecting the *Garcinia Cambogia (R0028)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	<i>European Pharmacopoeia 10th Edition, Basic Version 2020</i> [3]
Komm2.2.40	<i>Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident</i> [4]
AA004	<i>Erstellung und Validierung eines IdentModul-Updates</i>

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Garcinia Cambogia (R0028)	2	4	0

Second-stage model

For differentiation of the substance/substance group *Garcinia Cambogia (R0028)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Garcinia Cambogia (R0028)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Garcinia Cambogi...	1003-010	85119	40	not required
Hepart	Garcinia Cambogi...	16000421	85350	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 100 spectra of 2 reference samples from the substance/substance group *Garcinia Cambogia (R0028)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 310 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 200 spectra of 4 reference samples from the substance/substance group *Garcinia Cambogia (R0028)*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Garcinia Cambogia (R0028)	130816501	85193	40
Hepart	Garcinia Cambogia (R0028)	14001695	85351	60
Hepart	Garcinia Cambogia (R0028)	15000482	85353	60
Hepart	Garcinia Cambogia (R0028)	19002074	85609	40

- 9889 spectra from a total of 233 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Garcinia Cambogia (R0028)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Garcinia Cambogia (R0028)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Garcinia Cambogia (R0028)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	100	0	13 310
Type B	0	110	90	9889
Type C	0	0	0	3

The substance/substance group *Garcinia Cambogia (R0028)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9308 %)	100.0000 % (> 94.0000 %)
Type B	100.0000 % (> 99.9004 %)	55.0000 % (> 53.5000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Garcinia Cambogia (R0028)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Magnesium glycerophosphate (R0114)	14.51	—
Osha root extract (R0220)	20.81	—
Calcium lactate pentahydrate (R0013)	27.08	—
Citrus bioflavonoids 60% (R0021)	36.31	—
Oregano extract (R0206)	40.60	—
Artichoke extract (R0208)	52.91	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Garcinia Cambogia (R0028)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85350	85350	0.00	39.97
85119	85119	0.00	17.81

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Ginger extract (R0199)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85290-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Ginger extract (R0199)

Special notes

When selecting the *Ginger extract (R0199)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Ginger extract (R0199)	2	1	0

Second-stage model

For differentiation of the substance/substance group *Ginger extract (R0199)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Ginger extract (R0199)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Ginger extract (...)	16000714	85372	40	not required
Hepart	Ginger extract (...)	20001026/1	85845	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Ginger extract (R0199)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 330 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 60 spectra of 1 reference samples from the substance/substance group *Ginger extract (R0199)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Ginger extract (R0199)	15000800	85290	60

- 10 029 spectra from a total of 236 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Ginger extract (R0199)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Ginger extract (R0199)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ginger extract (R0199)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	13 330
Type B	0	0	60	10 029
Type C	0	0	0	3

The substance/substance group *Ginger extract (R0199)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9309 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9011 %)	0.0000 % (≥ 0.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ginger extract (R0199)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Sunflower oil powder 25 (R0319)	21.85	—
Menaquinone-7 0.2% (R0134)	27.89	—
Retinyl palmitate (vitamin A palmitate...	28.25	—
Lutein 20% (R0052)	36.47	—
Ginseng extract (R0029)	39.68	—
Acacia gum (encapsia) / acacia gum (R0...	39.82	—
Cholecalciferol (vitamin D3) (R0018)	40.85	—
Sodium chondroitin sulfate (R0061)	45.14	—
beta-glucan 28% (oat fibre) (R0334)	50.98	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ginger extract (R0199)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85372	85372	0.00	28.25
85845	85845	0.00	21.85

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Ginkgo extract (R0093)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85114-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Ginkgo extract (R0093)

Special notes

When selecting the *Ginkgo extract (R0093)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Ginkgo extract (R0093)	3	1	0

Second-stage model

For differentiation of the substance/substance group *Ginkgo extract (R0093)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Ginkgo extract (R0093)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Ginkgo extract (...)	050101	85150	40	not required
Hepart	Ginkgo extract (...)	18000830	85531	40	not required
Hepart	Ginkgo extract (...)	19002109	85611	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 120 spectra of 3 reference samples from the substance/substance group *Ginkgo extract (R0093)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 13 290 spectra from a total of 270 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *Ginkgo extract (R0093)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Ginkgo extract (R0093)	1010006	85114	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Ginkgo extract (R0093)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Ginkgo extract (R0093)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ginkgo extract (R0093)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	120	0	13 290
Type B	0	0	40	10 049
Type C	0	0	0	3

The substance/substance group *Ginkgo extract (R0093)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9307 %)	100.0000 % (> 95.0000 %)
Type B	100.0000 % (> 99.9016 %)	0.0000 % (≥ 0.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ginkgo extract (R0093)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Olive leaf extract (R0202)	43.61	—
Citrus bioflavonoids 60% (R0021)	54.91	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ginkgo extract (R0093)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85611	85611	0.00	45.66
85531	85531	0.00	43.61
85150	85150	0.00	82.13

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group	Ginseng extract (R0029) / Lutein 20% (R0052) / Lycopene 5%
Substance class	Hepart raw materials
Report date	22/03/2021
Report number	85098-2021-03-22
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Ginseng extract (R0029) / Lutein 20% (R0052) / Lycopene 5% extract (R0053); Ginseng extract (R0029); Lutein 20% (R0052); Lycopene 5% extract (R0053)

Special notes

When selecting the *Ginseng extract (R0029) / Lutein 20% (R0052) / Lycopene 5% extract (R0053)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	<i>European Pharmacopoeia 10th Edition, Basic Version 2020</i> [3]
Komm2.2.40	<i>Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident</i> [4]
AA004	<i>Erstellung und Validierung eines IdentModul-Updates</i>

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Ginseng extract (R0029)	2	0	0
Lutein 20% (R0052)	2	2	0
Lycopene 5% extract (R0053)	2	1	0

Second-stage model

For differentiation of the substance/substance group *Ginseng extract (R0029)* / *Lutein 20% (R0052)* / *Lycopene 5% extract (R0053)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Ginseng extract (R0029)* / *Lutein 20% (R0052)* / *Lycopene 5% extract (R0053)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Ginseng extract ...	068001	85151	30	not required
Hepart	Ginseng extract ...	14001403	85224	60	not required
Hepart	Lutein 20% (R0052)	14001179	85228	60	not required
Hepart	Lutein 20% (R0052)	16001144	85401	40	not required
Hepart	Lycopene 5% extr...	050901	85098	40	not required
Hepart	Lycopene 5% extr...	14000658	85181	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 270 spectra of 6 reference samples from the substance/substance group *Ginseng extract (R0029)* / *Lutein 20% (R0052)* / *Lycopene 5% extract (R0053)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 6 different batches.
- 13 140 spectra from a total of 267 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 3 reference samples from the substance/substance group *Ginseng extract (R0029)* / *Lutein 20% (R0052)* / *Lycopene 5% extract (R0053)*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Lutein 20% (R0052)	130708600	85177	40
Hepart	Lutein 20% (R0052)	14000664	85182	40
Hepart	Lycopene 5% extract (R0053)	1102-016	85180	40

- 9969 spectra from a total of 234 batches from further 104 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Ginseng extract (R0029) / Lutein 20% (R0052) / Lycopene 5% extract (R0053)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Ginseng extract (R0029) / Lutein 20% (R0052) / Lycopene 5% extract (R0053)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ginseng extract (R0029) / Lutein 20% (R0052) / Lycopene 5% extract (R0053)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	270	0	13 140
Type B	0	93	27	9969
Type C	0	0	0	3

The substance/substance group *Ginseng extract (R0029) / Lutein 20% (R0052) / Lycopene 5% extract (R0053)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9304 %)	100.0000 % (> 97.7778 %)
Type B	100.0000 % (> 99.9006 %)	77.5000 % (> 75.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ginseng extract (R0029) / Lutein 20% (R0052) / Lycopene 5% extract (R0053)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Acacia gum (encapsia) / acacia gum (R0...	19.77	—
Magnesium glycerophosphate (R0114)	20.69	—
Juniper extract (R0192)	20.91	—
Hydrangea extract (R0203)	22.29	—
Instant yeast (dry yeast) (R0112)	24.76	—
Marshmallow root extract (R0204)	24.81	—
Gentian extract (R0200)	28.81	—
Katuka herb extract (R0210)	32.15	—
beta-glucan 28% (oat fibre) (R0334)	33.23	—
Chili extract (R0198)	33.64	—
Garcinia Cambogia (R0028)	35.21	—
Artichoke extract (R0208)	35.50	—
Ginger extract (R0199)	35.66	—
Calmegh extract (R0201)	37.86	—
Calcium lactate pentahydrate (R0013)	40.20	—
Beta carotene 10% (R0237)	41.34	—
Pumpkin seed protein (R0306)	42.00	—
Konjac glucomannan 80% (R0293)	43.78	—
Osha root extract (R0220)	43.87	—
Sunflower oil powder 25 (R0319)	47.11	—
Citrus bioflavonoids 60% (R0021)	47.34	—
Coriolus extract (R0207)	48.69	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ginseng extract (R0029) / Lutein 20% (R0052) / Lycopene 5% extract (R0053)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85228	85228	0.00	43.11
85098	85098	0.00	33.23
85151	85151	0.00	35.66
85181	85181	0.00	24.76
85401	85401	0.00	41.34
85224	85224	0.00	20.91

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Glucosamine sulfate 2 KCl (R0030)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85170-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Glucosamine sulfate 2 KCl (R0030); Glucosamine sulphate 2 KCl (R0030)

Special notes

When selecting the *Glucosamine sulfate 2 KCl (R0030)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Glucosamine sulfate 2 KCl (R0030)	3	7	3

Second-stage model

For differentiation of the substance/substance group *Glucosamine sulfate 2 KCl (R0030)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Glucosamine sulfate 2 KCl (R0030)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Glucosamine sulf...	14000030	85170	40	not required
Hepart	Glucosamine sulf...	16001125	85427	40	not required
Hepart	Glucosamine sulf...	17000804	85484	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 120 spectra of 3 reference samples from the substance/substance group *Glucosamine sulfate 2 KCl (R0030)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 13 290 spectra from a total of 270 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 280 spectra of 7 reference samples from the substance/substance group *Glucosamine sulfate 2 KCl (R0030)*.
- Among them are spectra of independent samples from 7 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Glucosamine sulfate 2 KCl (R...	16000867	85428	40

continued on the next page

continued from previous page

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Glucosamine sulfate 2 KCl (R...	16001696	85452	40
Hepart	Glucosamine sulfate 2 KCl (R...	16001667	85453	40
Hepart	Glucosamine sulfate 2 KCl (R...	17001309	85501	40
Hepart	Glucosamine sulfate 2 KCl (R...	19000254	85580	40
Hepart	Glucosamine sulfate 2 KCl (R...	19002345/0	85817	40
Hepart	Glucosamine sulfate 2 KCl (R...	20000447	85838	40

- 9809 spectra from a total of 230 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 3 spectra from 3 *Apo-Ident* customers from 3 batches from the substance/substance group *Glucosamine sulfate 2 KCl (R0030)*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Caelo	Glucosamine sulfate 2 KCl (R...	1122	1
Euro OTC	Glucosamine sulfate 2 KCl (R...	L1307010-01	1
Koisser	Glucosamine sulfate 2 KCl (R...	120719K371	1

- 0 spectra from 0 *Apo-Ident* customers from a total of 0 batches from a further 0 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Glucosamine sulfate 2 KCl (R0030)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Glucosamine sulfate 2 KCl (R0030)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	120	0	13 290
Type B	0	280	0	9809
Type C	0	3	0	0

The substance/substance group *Glucosamine sulfate 2 KCl (R0030)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9307 %)	100.0000 % (> 95.0000 %)
Type B	100.0000 % (> 99.9004 %)	100.0000 % (> 97.8571 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Glucosamine sulfate 2 KCl (R0030)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Inositol (myo-) (R0039)	74.44	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Glucosamine sulfate 2 KCl (R0030)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85427	85427	0.00	74.44
85170	85170	0.00	92.07
85484	85484	0.00	78.99

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Glutathione (reduced form) (R0032)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85261-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Glutathione (reduced form) (R0032)

Special notes

When selecting the *Glutathione (reduced form) (R0032)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Glutathione (reduced form) (R0032)	2	0	0

Second-stage model

For differentiation of the substance/substance group *Glutathione (reduced form)* (R0032) the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Glutathione (reduced form)* (R0032):

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Glutathione (red...	15000418	85261	60	not required
Hepart	Glutathione (red...	16000695	85379	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 100 spectra of 2 reference samples from the substance/substance group *Glutathione (reduced form)* (R0032). These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 310 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Glutathione (reduced form)* (R0032).
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Glutathione (reduced form) (R0032)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Glutathione (reduced form) (R0032)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Glutathione (reduced form) (R0032)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	100	0	13 310
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Glutathione (reduced form) (R0032)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9308 %)	100.0000 % (> 94.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Glutathione (reduced form) (R0032)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Zinc citrate trihydrate (R0089)	74.97	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Glutathione (reduced form) (R0032)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85379	85379	0.00	77.60
85261	85261	0.00	74.97

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Glycine (R0033)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85050-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Glycine (R0033)

Special notes

When selecting the *Glycine (R0033)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Glycine (R0033)	2	2	0

Second-stage model

For differentiation of the substance/substance group *Glycine (R0033)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Glycine (R0033)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Glycine (R0033)	1212001	85131	40	not required
Hepart	Glycine (R0033)	14000031	85173	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Glycine (R0033)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 330 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 100 spectra of 2 reference samples from the substance/substance group *Glycine (R0033)*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Glycine (R0033)	15000641	85273	60
Hepart	Glycine (R0033)	16001698	85472	40

- 9989 spectra from a total of 235 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Glycine (R0033)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Glycine (R0033)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Glycine (R0033)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	13 330
Type B	0	66	34	9989
Type C	0	0	0	3

The substance/substance group *Glycine (R0033)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9309 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9007 %)	66.0000 % (> 63.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Glycine (R0033)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-tyrosine (R0086)	61.35	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Glycine (R0033)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85131	85131	0.00	96.18
85173	85173	0.00	82.90

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Green tea extract (R0034)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85019-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Green tea extract (R0034)

Special notes

When selecting the *Green tea extract (R0034)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Green tea extract (R0034)	3	1	0

Second-stage model

For differentiation of the substance/substance group *Green tea extract (R0034)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Green tea extract (R0034)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Green tea extrac...	058901	85095	40	not required
Hepart	Green tea extrac...	1208017	85117	40	not required
Hepart	Green tea extrac...	130430-GT12	85190	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 120 spectra of 3 reference samples from the substance/substance group *Green tea extract (R0034)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 13 290 spectra from a total of 270 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *Green tea extract (R0034)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Green tea extract (R0034)	1008-004	85188	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Green tea extract (R0034)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Green tea extract (R0034)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Green tea extract (R0034)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	120	0	13 290
Type B	0	21	19	10 049
Type C	0	0	0	3

The substance/substance group *Green tea extract (R0034)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9307 %)	100.0000 % (> 95.0000 %)
Type B	100.0000 % (> 99.9016 %)	52.5000 % (> 45.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Green tea extract (R0034)* in the substance class *Heptaraw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
OPC (grape seed) extract (R0065)	13.68	—
Citric acid monohydrate (R0143)	56.18	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Green tea extract (R0034)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85190	85190	0.00	13.68
85117	85117	0.00	17.65
85095	85095	0.00	32.74

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group	Guarana extract (R0289)
Substance class	Hepart raw materials
Report date	22/03/2021
Report number	85495-2021-03-22
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Guarana extract (R0289)

Special notes

When selecting the *Guarana extract (R0289)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Guarana extract (R0289)	1	1	0

Second-stage model

For differentiation of the substance/substance group *Guarana extract (R0289)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Guarana extract (R0289)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Guarana extract ...	17000482	85495	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 40 spectra of 1 reference samples from the substance/substance group *Guarana extract (R0289)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 370 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *Guarana extract (R0289)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Guarana extract (R0289)	19001513	85603	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Guarana extract (R0289)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Guarana extract (R0289)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Guarana extract (R0289)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	40	0	13 370
Type B	0	39	1	10 049
Type C	0	0	0	3

The substance/substance group *Guarana extract (R0289)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9317 %)	100.0000 % (> 85.0000 %)
Type B	100.0000 % (> 99.9016 %)	97.5000 % (> 90.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Guarana extract (R0289)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Citrus bioflavonoids 60% (R0021)	28.10	—
Acacia gum (encapsia) / acacia gum (R0...	28.41	—
Marshmallow root extract (R0204)	33.39	—
Gentian extract (R0200)	41.66	—
Ginseng extract (R0029)	47.82	—
Konjac glucomannan 80% (R0293)	49.42	—
Vanilla flavoring (natural flavor) (R0...	51.33	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Guarana extract (R0289)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85495	85495	0.00	28.10

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Hydrangea extract (R0203)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85295-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Hydrangea extract (R0203)

Special notes

When selecting the *Hydrangea extract (R0203)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Hydrangea extract (R0203)	2	0	0

Second-stage model

For differentiation of the substance/substance group *Hydrangea extract (R0203)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Hydrangea extract (R0203)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Hydrangea extrac...	15000792	85295	60	not required
Hepart	Hydrangea extrac...	16000718	85368	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 100 spectra of 2 reference samples from the substance/substance group *Hydrangea extract (R0203)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 310 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Hydrangea extract (R0203)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Hydrangea extract (R0203)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Hydrangea extract (R0203)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Hydrangea extract (R0203)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	100	0	13 310
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Hydrangea extract (R0203)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9308 %)	100.0000 % (> 94.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Hydrangea extract (R0203)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Coriolus extract (R0207)	17.70	—
Chili extract (R0198)	19.24	—
Ginseng extract (R0029)	20.20	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Calmegh extract (R0201)	25.49	—
Calcium lactate pentahydrate (R0013)	34.39	—
Juniper extract (R0192)	37.17	—
Citrus bioflavonoids 60% (R0021)	42.32	—
Magnesium glycerophosphate (R0114)	42.97	—
Acacia gum (encapsia) / acacia gum (R0...	44.04	—
Marshmallow root extract (R0204)	45.45	—
Gentian extract (R0200)	47.09	—
Artichoke extract (R0208)	47.96	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Hydrangea extract (R0203)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85368	85368	0.00	20.20
85295	85295	0.00	17.70

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50% greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Hyssop extract (R0205)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85292-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Hyssop extract (R0205)

Special notes

When selecting the *Hyssop extract (R0205)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Hyssop extract (R0205)	2	0	0

Second-stage model

For differentiation of the substance/substance group *Hyssop extract (R0205)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Hyssop extract (R0205)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Hyssop extract (...)	15000787	85292	60	not required
Hepart	Hyssop extract (...)	16000719	85378	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 100 spectra of 2 reference samples from the substance/substance group *Hyssop extract (R0205)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 310 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Hyssop extract (R0205)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Hyssop extract (R0205)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Hyssop extract (R0205)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Hyssop extract (R0205)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	100	0	13 310
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Hyssop extract (R0205)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9308 %)	100.0000 % (> 94.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Hyssop extract (R0205)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Oregano extract (R0206)	15.50	—
Artichoke extract (R0208)	16.59	—
Coriolus extract (R0207)	16.87	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Black cumin extract (R0193)	21.30	—
Walnut leaf extract (R0197)	24.61	—
Katuka herb extract (R0210)	26.44	—
Garcinia Cambogia (R0028)	27.53	—
Magnesium glycerophosphate (R0114)	31.12	—
Osha root extract (R0220)	31.33	—
Chili extract (R0198)	41.76	—
Juniper extract (R0192)	43.81	—
Citrus bioflavonoids 60% (R0021)	44.83	—
Hydrangea extract (R0203)	45.05	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Hyssop extract (R0205)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85292	85292	0.00	16.87
85378	85378	0.00	15.50

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group	Inositol (myo-) (R0039)
Substance class	Hepart raw materials
Report date	22/03/2021
Report number	85009-2021-03-22
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Inositol (myo-) (R0039)

Special notes

When selecting the *Inositol (myo-) (R0039)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	<i>European Pharmacopoeia 10th Edition, Basic Version 2020</i> [3]
Komm2.2.40	<i>Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident</i> [4]
AA004	<i>Erstellung und Validierung eines IdentModul-Updates</i>

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Inositol (myo-) (R0039)	2	3	0

Second-stage model

For differentiation of the substance/substance group *Inositol (myo-) (R0039)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Inositol (myo-) (R0039)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Inositol (myo-) ...	1209006	85091	40	not required
Hepart	Inositol (myo-) ...	15000233	85258	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 100 spectra of 2 reference samples from the substance/substance group *Inositol (myo-) (R0039)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 310 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 140 spectra of 3 reference samples from the substance/substance group *Inositol (myo-) (R0039)*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Inositol (myo-) (R0039)	16000188	85358	60
Hepart	Inositol (myo-) (R0039)	18001225	85556	40
Hepart	Inositol (myo-) (R0039)	19002139/0	85821	40

- 9949 spectra from a total of 234 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Inositol (myo-) (R0039)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Inositol (myo-) (R0039)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Inositol (myo-) (R0039)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	100	0	13 310
Type B	0	140	0	9949
Type C	0	0	0	3

The substance/substance group *Inositol (myo-) (R0039)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9308 %)	100.0000 % (> 94.0000 %)
Type B	100.0000 % (> 99.9006 %)	100.0000 % (> 95.7143 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Inositol (myo-) (R0039)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
D-galactose (R0135)	52.65	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Inositol (myo-) (R0039)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85258	85258	0.00	72.00
85091	85091	0.00	71.84

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group	Inositol hexanicotinate (R0040)
Substance class	Hepart raw materials
Report date	22/03/2021
Report number	85044-2021-03-22
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Inositol hexanicotinate (R0040)

Special notes

When selecting the *Inositol hexanicotinate (R0040)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	<i>European Pharmacopoeia 10th Edition, Basic Version 2020</i> [3]
Komm2.2.40	<i>Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident</i> [4]
AA004	<i>Erstellung und Validierung eines IdentModul-Updates</i>

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Inositol hexanicotinate (R0040)	2	4	0

Second-stage model

For differentiation of the substance/substance group *Inositol hexanicotinate (R0040)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Inositol hexanicotinate (R0040)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Inositol hexanic...	14001130	85205	40	not required
Hepart	Inositol hexanic...	16000696	85382	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Inositol hexanicotinate (R0040)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 330 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 160 spectra of 4 reference samples from the substance/substance group *Inositol hexanicotinate (R0040)*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Inositol hexanicotinate (R00...	063501	85186	40
Hepart	Inositol hexanicotinate (R00...	18001226	85549	40
Hepart	Inositol hexanicotinate (R00...	19000057	85578	40
Hepart	Inositol hexanicotinate (R00...	20001025/0	85846	40

- 9929 spectra from a total of 233 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Inositol hexanicotinate (R0040)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Inositol hexanicotinate (R0040)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Inositol hexanicotinate (R0040)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	13 330
Type B	0	146	14	9929
Type C	0	0	0	3

The substance/substance group *Inositol hexanicotinate (R0040)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9309 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9005 %)	91.2500 % (> 89.3750 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Inositol hexanicotinate (R0040)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-tryptophan (R0085)	43.32	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Inositol hexanicotinate (R0040)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85205	85205	0.00	34.63
85382	85382	0.00	36.27

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Instant yeast (dry yeast) (R0112)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 10004604-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Instant yeast (dry yeast) (R0112)

Special notes

When selecting the *Instant yeast (dry yeast) (R0112)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Instant yeast (dry yeast) (R0112)	1	0	0

Second-stage model

For differentiation of the substance/substance group *Instant yeast (dry yeast) (R0112)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Instant yeast (dry yeast) (R0112)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Instant yeast (d...	20001441	85875	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 40 spectra of 1 reference samples from the substance/substance group *Instant yeast (dry yeast) (R0112)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 370 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Instant yeast (dry yeast) (R0112)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Instant yeast (dry yeast) (R0112)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Instant yeast (dry yeast) (R0112)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Instant yeast (dry yeast) (R0112)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	40	0	13 370
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Instant yeast (dry yeast) (R0112)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9317 %)	100.0000 % (> 85.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Instant yeast (dry yeast) (R0112)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Lutein 20% (R0052)	17.53	—
Lycopene 5% extract (R0053)	31.06	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ginger extract (R0199)	42.05	—
Isoflavones 10% (R0042)	45.79	—
Pumpkin seed protein (R0306)	45.89	—
Magnesium glycerophosphate (R0114)	48.52	—
Citrus bioflavonoids 60% (R0021)	51.01	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Instant yeast (dry yeast) (R0112)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85875	85875	0.00	17.53

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Inulin (R0041)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85109-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Inulin (R0041)

Special notes

When selecting the *Inulin (R0041)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Inulin (R0041)	3	4	0

Second-stage model

For differentiation of the substance/substance group *Inulin (R0041)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Inulin (R0041)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Inulin (R0041)	1209021	85109	40	not required
Hepart	Inulin (R0041)	14001196	85226	60	not required
Hepart	Inulin (R0041)	15000234	85242	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 3 reference samples from the substance/substance group *Inulin (R0041)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 13 250 spectra from a total of 270 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 180 spectra of 4 reference samples from the substance/substance group *Inulin (R0041)*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Inulin (R0041)	15000747	85281	60
Hepart	Inulin (R0041)	16001666	85441	40
Hepart	Inulin (R0041)	18000514	85516	40
Hepart	Inulin (R0041)	18002044	85583	40

- 9909 spectra from a total of 233 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Inulin (R0041)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Inulin (R0041)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Inulin (R0041)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	13 250
Type B	0	180	0	9909
Type C	0	0	0	3

The substance/substance group *Inulin (R0041)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9305 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9005 %)	100.0000 % (> 96.6667 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Inulin (R0041)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Acacia gum (encapsia) / acacia gum (R0...	19.73	—
Juniper extract (R0192)	21.50	—
Gentian extract (R0200)	23.02	—
Cholecalciferol (vitamin D3) (R0018)	24.74	—
Marshmallow root extract (R0204)	26.81	—
Potato starch, pregelatinized (R0047)	26.89	—
Galactomannans (R0035/R0094)	32.86	—
Konjac glucomannan 80% (R0293)	34.24	—
Menaquinone-7 0.2% (R0134)	35.43	—
Ginger extract (R0199)	42.10	—
beta-glucan 28% (oat fibre) (R0334)	53.14	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Inulin (R0041)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85109	85109	0.00	23.63
85226	85226	0.00	22.29
85242	85242	0.00	21.50

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Isoflavones 10% (R0042)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85092-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Isoflavones 10% (R0042)

Special notes

When selecting the *Isoflavones 10% (R0042)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Isoflavones 10% (R0042)	2	1	0

Second-stage model

For differentiation of the substance/substance group *Isoflavones 10% (R0042)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Isoflavones 10% (R0042)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Isoflavones 10% ...	057001	85092	40	not required
Hepart	Isoflavones 10% ...	16001515	85445	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Isoflavones 10% (R0042)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 330 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 100 spectra of 2 reference samples from the substance/substance group *Isoflavones 10% (R0042)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Isoflavones 10% (R0042)	14001396	85225	60
Hepart	Isoflavones 10% (R0042)	16001515	85457	40

- 9989 spectra from a total of 235 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Isoflavones 10% (R0042)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Isoflavones 10% (R0042)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Isoflavones 10% (R0042)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	13 330
Type B	0	40	60	9989
Type C	0	0	0	3

The substance/substance group *Isoflavones 10% (R0042)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9309 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9007 %)	40.0000 % (> 37.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Isoflavones 10% (R0042)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Citrus bioflavonoids 60% (R0021)	49.86	—
Instant yeast (dry yeast) (R0112)	50.15	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Isoflavones 10% (R0042)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85092	85092	0.00	74.64
85445	85445	0.00	50.15

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Katuka herb extract (R0210)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85297-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Katuka herb extract (R0210)

Special notes

When selecting the *Katuka herb extract (R0210)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Katuka herb extract (R0210)	2	0	0

Second-stage model

For differentiation of the substance/substance group *Katuka herb extract (R0210)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Katuka herb extract (R0210)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Katuka herb extr...	15000793	85297	60	not required
Hepart	Katuka herb extr...	16000721	85366	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 100 spectra of 2 reference samples from the substance/substance group *Katuka herb extract (R0210)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 310 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Katuka herb extract (R0210)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Katuka herb extract (R0210)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Katuka herb extract (R0210)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Katuka herb extract (R0210)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	100	0	13 310
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Katuka herb extract (R0210)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9308 %)	100.0000 % (> 94.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Katuka herb extract (R0210)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Artichoke extract (R0208)	11.34	—
Hyssop extract (R0205)	21.83	—
Marshmallow root extract (R0204)	22.59	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Juniper extract (R0192)	24.87	—
Gentian extract (R0200)	28.09	—
Magnesium glycerophosphate (R0114)	30.11	—
Acacia gum (encapsia) / acacia gum (R0...	31.80	—
Ginseng extract (R0029)	35.03	—
Osha root extract (R0220)	35.90	—
Chili extract (R0198)	45.27	—
Coriolus extract (R0207)	45.88	—
Hydrangea extract (R0203)	47.46	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Katuka herb extract (R0210)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85297	85297	0.00	20.69
85366	85366	0.00	11.34

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50% greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Konjac glucomannan 80% (R0293)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 10003564-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Konjac glucomannan 80% (R0293)

Special notes

When selecting the *Konjac glucomannan 80% (R0293)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Konjac glucomannan 80% (R0293)	2	1	0

Second-stage model

For differentiation of the substance/substance group *Konjac glucomannan 80% (R0293)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Konjac glucomannan 80% (R0293)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Konjac glucomann...	18001364	85574	40	not required
Hepart	Konjac glucomann...	19002347	85816	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Konjac glucomannan 80% (R0293)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 330 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *Konjac glucomannan 80% (R0293)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Konjac glucomannan 80% (R0293)	18002046	85589	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Konjac glucomannan 80% (R0293)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Konjac glucomannan 80% (R0293)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Konjac glucomannan 80% (R0293)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	13 330
Type B	0	40	0	10 049
Type C	0	0	0	3

The substance/substance group *Konjac glucomannan 80% (R0293)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9309 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9016 %)	100.0000 % (> 85.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Konjac glucomannan 80% (R0293)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Galactomannans (R0035/R0094)	16.13	—
Acacia gum (encapsia) / acacia gum (R0...	20.94	—
Gentian extract (R0200)	24.66	—
Juniper extract (R0192)	33.69	—
Marshmallow root extract (R0204)	36.48	—
Potato starch, pregelatinized (R0047)	37.88	—
Citrus bioflavonoids 60% (R0021)	50.70	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Konjac glucomannan 80% (R0293)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85574	85574	0.00	25.93
85816	85816	0.00	26.51

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **L-5-hydroxytryptophan (R0050)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85208-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

L-5-hydroxytryptophan (R0050)

Special notes

When selecting the *L-5-hydroxytryptophan (R0050)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
L-5-hydroxytryptophan (R0050)	3	1	0

Second-stage model

For differentiation of the substance/substance group *L-5-hydroxytryptophan (R0050)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *L-5-hydroxytryptophan (R0050)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	L-5-hydroxytrypt...	1103004	85208	40	not required
Hepart	L-5-hydroxytrypt...	15000794	85309	60	not required
Hepart	L-5-hydroxytrypt...	15000794	85349	60	not required
Hepart	L-5-hydroxytrypt...	16000693	85390	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 200 spectra of 4 reference samples from the substance/substance group *L-5-hydroxytryptophan (R0050)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 13 210 spectra from a total of 270 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *L-5-hydroxytryptophan (R0050)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	L-5-hydroxytryptophan (R0050)	17001298	85503	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *L-5-hydroxytryptophan (R0050)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *L-5-hydroxytryptophan (R0050)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *L-5-hydroxytryptophan (R0050)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	200	0	13 210
Type B	0	40	0	10 049
Type C	0	0	0	3

The substance/substance group *L-5-hydroxytryptophan (R0050)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9305 %)	100.0000 % (> 97.0000 %)
Type B	100.0000 % (> 99.9016 %)	100.0000 % (> 85.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *L-5-hydroxytryptophan (R0050)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Ascorbic acid (R0007)	187.34	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *L-5-hydroxytryptophan (R0050)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85390	85390	0.00	201.27
85208	85208	0.00	206.27
85309	85309	0.00	187.34
85349	85349	0.00	190.34

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **L-arginine base (R0006)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85012-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

L-arginine base (R0006)

Special notes

When selecting the *L-arginine base (R0006)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
L-arginine base (R0006)	2	2	0

Second-stage model

For differentiation of the substance/substance group *L-arginine base (R0006)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *L-arginine base (R0006)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	L-arginine base ...	1209011	85094	40	not required
Hepart	L-arginine base ...	18001222	85544	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *L-arginine base (R0006)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 330 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 2 reference samples from the substance/substance group *L-arginine base (R0006)*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	L-arginine base (R0006)	16001692	85461	40
Hepart	L-arginine base (R0006)	19000250	85590	40

- 10 009 spectra from a total of 235 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *L-arginine base (R0006)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *L-arginine base (R0006)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *L-arginine base (R0006)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	13 330
Type B	0	80	0	10 009
Type C	0	0	0	3

The substance/substance group *L-arginine base (R0006)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9309 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9009 %)	100.0000 % (> 92.5000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *L-arginine base (R0006)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Citrus bioflavonoids 60% (R0021)	83.15	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *L-arginine base (R0006)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85094	85094	0.00	89.79
85544	85544	0.00	88.99

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **L-carnitine-L-tartrate (R0014)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85086-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

L-carnitine-L-tartrate (R0014)

Special notes

When selecting the *L-carnitine-L-tartrate (R0014)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
L-carnitine-L-tartrate (R0014)	5	0	0

Second-stage model

For differentiation of the substance/substance group *L-carnitine-L-tartrate* (*R0014*) the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *L-carnitine-L-tartrate* (*R0014*):

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	L-carnitine-L-ta...	058101	85086	50	not required
Hepart	L-carnitine-L-ta...	14001129	85198	40	not required
Hepart	L-carnitine-L-ta...	16000165	85345	60	not required
Hepart	L-carnitine-L-ta...	18000445	85518	40	not required
Hepart	L-carnitine-L-ta...	19001446	85604	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 230 spectra of 5 reference samples from the substance/substance group *L-carnitine-L-tartrate* (*R0014*). These samples are listed above in the [calibration samples](#) section. The reference samples come from 5 different batches.
- 13 180 spectra from a total of 268 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *L-carnitine-L-tartrate* (*R0014*).
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the

substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group L-carnitine-L-tartrate (*R0014*).
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group L-carnitine-L-tartrate (*R0014*) can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with L-carnitine-L-tartrate (*R0014*) and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	230	0	13 180
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group L-carnitine-L-tartrate (*R0014*) can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9304 %)	100.0000 % (> 97.3913 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group L-carnitine-L-tartrate (*R0014*) in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Glutathione (reduced form) (R0032)	82.07	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *L-carnitine-L-tartrate (R0014)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85198	85198	0.00	82.07
85604	85604	0.00	85.41
85518	85518	0.00	99.48
85345	85345	0.00	94.79
85086	85086	0.00	98.12

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **L-carnosine (R0015)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85097-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

L-carnosine (R0015)

Special notes

When selecting the *L-carnosine (R0015)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
L-carnosine (R0015)	1	1	0

Second-stage model

For differentiation of the substance/substance group *L-carnosine (R0015)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *L-carnosine (R0015)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	L-carnosine (R00...	050801	85097	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 40 spectra of 1 reference samples from the substance/substance group *L-carnosine (R0015)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 370 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *L-carnosine (R0015)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	L-carnosine (R0015)	18000730	85530	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group L-carnosine (R0015).
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group L-carnosine (R0015) can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with L-carnosine (R0015) and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	40	0	13 370
Type B	0	23	17	10 049
Type C	0	0	0	3

The substance/substance group L-carnosine (R0015) can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9317 %)	100.0000 % (> 85.0000 %)
Type B	100.0000 % (> 99.9016 %)	57.5000 % (> 50.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group L-carnosine (R0015) in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-lysine HCl (R0054)	112.70	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *L-carnosine (R0015)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85097	85097	0.00	112.70

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **L-glutamine (R0031)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85016-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

L-glutamine (R0031)

Special notes

When selecting the *L-glutamine (R0031)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
L-glutamine (R0031)	4	6	0

Second-stage model

For differentiation of the substance/substance group *L-glutamine (R0031)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *L-glutamine (R0031)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	L-glutamine (R00...	1209007	85118	40	not required
Hepart	L-glutamine (R00...	14001120	85200	40	not required
Hepart	L-glutamine (R00...	15001426	85318	60	not required
Hepart	L-glutamine (R00...	15001604	85327	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 200 spectra of 4 reference samples from the substance/substance group *L-glutamine (R0031)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 4 different batches.
- 13 210 spectra from a total of 269 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 255 spectra of 6 reference samples from the substance/substance group *L-glutamine (R0031)*.
- Among them are spectra of independent samples from 6 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	L-glutamine (R0031)	16000178	85343	55

continued on the next page

continued from previous page

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	L-glutamine (R0031)	16000868	85422	40
Hepart	L-glutamine (R0031)	16001697	85469	40
Hepart	L-glutamine (R0031)	19000253	85588	40
Hepart	L-glutamine (R0031)	19002346	85830	40
Hepart	L-glutamine (R0031)	20000846/0	85843	40

- 9834 spectra from a total of 231 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *L-glutamine (R0031)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *L-glutamine (R0031)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *L-glutamine (R0031)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	190	10	13 210
Type B	0	255	0	9834
Type C	0	0	0	3

The substance/substance group *L-glutamine (R0031)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9305 %)	95.0000 % (> 93.5000 %)
Type B	100.0000 % (> 99.9004 %)	100.0000 % (> 97.6471 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *L-glutamine (R0031)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Magnesium glycerophosphate (R0114)	60.45	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *L-glutamine (R0031)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85200	85200	0.00	66.80
85318	85318	0.00	79.98
85327	85327	0.00	82.90
85118	85118	0.00	74.44

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group	L-histidine base (R0037)
Substance class	Hepart raw materials
Report date	22/03/2021
Report number	85099-2021-03-22
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

L-histidine base (R0037)

Special notes

When selecting the *L-histidine base (R0037)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
L-histidine base (R0037)	1	1	0

Second-stage model

For differentiation of the substance/substance group *L-histidine base (R0037)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *L-histidine base (R0037)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	L-histidine base...	051801	85099	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 40 spectra of 1 reference samples from the substance/substance group *L-histidine base (R0037)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 370 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *L-histidine base (R0037)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	L-histidine base (R0037)	14001055	85192	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *L-histidine base (R0037)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *L-histidine base (R0037)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *L-histidine base (R0037)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	40	0	13 370
Type B	0	0	40	10 049
Type C	0	0	0	3

The substance/substance group *L-histidine base (R0037)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9317 %)	100.0000 % (> 85.0000 %)
Type B	100.0000 % (> 99.9016 %)	0.0000 % (≥ 0.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *L-histidine base (R0037)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-carnosine (R0015)	91.19	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *L-histidine base (R0037)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85099	85099	0.00	91.19

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group	L-isoleucine (R0043)
Substance class	Hepart raw materials
Report date	22/03/2021
Report number	85076-2021-03-22
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

L-isoleucine (R0043)

Special notes

When selecting the *L-isoleucine (R0043)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	<i>European Pharmacopoeia 10th Edition, Basic Version 2020</i> [3]
Komm2.2.40	<i>Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident</i> [4]
AA004	<i>Erstellung und Validierung eines IdentModul-Updates</i>

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
L-isoleucine (R0043)	3	2	0

Second-stage model

For differentiation of the substance/substance group *L-isoleucine (R0043)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *L-isoleucine (R0043)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	L-isoleucine (R0...	1109011	85076	50	not required
Hepart	L-isoleucine (R0...	15000638	85272	60	not required
Hepart	L-isoleucine (R0...	18000832	85529	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 150 spectra of 3 reference samples from the substance/substance group *L-isoleucine (R0043)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 13 260 spectra from a total of 270 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 140 spectra of 3 reference samples from the substance/substance group *L-isoleucine (R0043)*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	L-isoleucine (R0043)	16000179	85347	60
Hepart	L-isoleucine (R0043)	18001228	85548	40
Hepart	L-isoleucine (R0043)	18001228	85818	40

- 9949 spectra from a total of 235 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *L-isoleucine (R0043)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *L-isoleucine (R0043)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *L-isoleucine (R0043)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	150	0	13 260
Type B	0	140	0	9949
Type C	0	0	0	3

The substance/substance group *L-isoleucine (R0043)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9306 %)	100.0000 % (> 96.0000 %)
Type B	100.0000 % (> 99.9006 %)	100.0000 % (> 95.7143 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *L-isoleucine (R0043)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-leucine (R0051)	35.44	—
Valine (R0087)	41.80	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *L-isoleucine (R0043)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85529	85529	0.00	38.51
85272	85272	0.00	35.44
85076	85076	0.00	36.94

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **L-leucine (R0051)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85126-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

L-leucine (R0051)

Special notes

When selecting the *L-leucine (R0051)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
L-leucine (R0051)	4	2	0

Second-stage model

For differentiation of the substance/substance group *L-leucine (R0051)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *L-leucine (R0051)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	L-leucine (R0051)	110055M04C	85126	40	not required
Hepart	L-leucine (R0051)	14001122	85203	20	not required
Hepart	L-leucine (R0051)	15000636	85284	60	not required
Hepart	L-leucine (R0051)	16000180	85346	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 180 spectra of 4 reference samples from the substance/substance group *L-leucine (R0051)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 4 different batches.
- 13 230 spectra from a total of 269 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 2 reference samples from the substance/substance group *L-leucine (R0051)*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	L-leucine (R0051)	18000833	85528	40
Hepart	L-leucine (R0051)	19000248/0	85586	40

- 10 009 spectra from a total of 235 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *L-leucine (R0051)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *L-leucine (R0051)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *L-leucine (R0051)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	180	0	13 230
Type B	0	80	0	10 009
Type C	0	0	0	3

The substance/substance group *L-leucine (R0051)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9305 %)	100.0000 % (> 96.6667 %)
Type B	100.0000 % (> 99.9009 %)	100.0000 % (> 92.5000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *L-leucine (R0051)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-isoleucine (R0043)	34.18	—
Valine (R0087)	52.22	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *L-leucine (R0051)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85126	85126	0.00	34.18
85346	85346	0.00	38.85
85203	85203	0.00	37.94
85284	85284	0.00	44.12

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **L-lysine HCl (R0054)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85033-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

L-lysine HCl (R0054)

Special notes

When selecting the *L-lysine HCl (R0054)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
L-lysine HCl (R0054)	3	6	0

Second-stage model

For differentiation of the substance/substance group *L-lysine HCl (R0054)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *L-lysine HCl (R0054)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	L-lysine HCl (R0...	1007010	85106	40	not required
Hepart	L-lysine HCl (R0...	15000637	85294	60	not required
Hepart	L-lysine HCl (R0...	16000181	85344	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 3 reference samples from the substance/substance group *L-lysine HCl (R0054)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 13 250 spectra from a total of 270 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 260 spectra of 6 reference samples from the substance/substance group *L-lysine HCl (R0054)*.
- Among them are spectra of independent samples from 6 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	L-lysine HCl (R0054)	14001829	85248	60
Hepart	L-lysine HCl (R0054)	17000480	85479	40

continued on the next page

continued from previous page

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	L-lysine HCl (R0054)	18001229	85573	40
Hepart	L-lysine HCl (R0054)	19001515	85593	40
Hepart	L-lysine HCl (R0054)	20000446	85837	40
Hepart	L-lysine HCl (R0054)	20000847	85844	40

- 9829 spectra from a total of 231 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *L-lysine HCl (R0054)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *L-lysine HCl (R0054)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *L-lysine HCl (R0054)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	13 250
Type B	0	260	0	9829
Type C	0	0	0	3

The substance/substance group *L-lysine HCl (R0054)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9305 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9004 %)	100.0000 % (> 97.6923 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *L-lysine HCl (R0054)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Valine (R0087)	76.65	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *L-lysine HCl (R0054)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85106	85106	0.00	76.65
85344	85344	0.00	79.20
85294	85294	0.00	84.71

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50% greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group	L-ornithine HCl (R0066)
Substance class	Hepart raw materials
Report date	22/03/2021
Report number	85008-2021-03-22
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

L-ornithine HCl (R0066)

Special notes

When selecting the *L-ornithine HCl (R0066)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
L-ornithine HCl (R0066)	2	2	0

Second-stage model

For differentiation of the substance/substance group *L-ornithine HCl (R0066)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *L-ornithine HCl (R0066)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	L-ornithine HCl ...	15000639	85276	60	not required
Hepart	L-ornithine HCl ...	15000998	85308	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 120 spectra of 2 reference samples from the substance/substance group *L-ornithine HCl (R0066)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 290 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 2 reference samples from the substance/substance group *L-ornithine HCl (R0066)*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	L-ornithine HCl (R0066)	16000871	85416	40
Hepart	L-ornithine HCl (R0066)	17001821	85510	40

- 10 009 spectra from a total of 235 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *L-ornithine HCl (R0066)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *L-ornithine HCl (R0066)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *L-ornithine HCl (R0066)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	120	0	13 290
Type B	0	80	0	10 009
Type C	0	0	0	3

The substance/substance group *L-ornithine HCl (R0066)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9307 %)	100.0000 % (> 95.0000 %)
Type B	100.0000 % (> 99.9009 %)	100.0000 % (> 92.5000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *L-ornithine HCl (R0066)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-tyrosine (R0086)	77.55	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *L-ornithine HCl (R0066)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85308	85308	0.00	82.94
85276	85276	0.00	101.35

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **L-phenylalanine (R0067)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85045-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

L-phenylalanine (R0067)

Special notes

When selecting the *L-phenylalanine (R0067)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
L-phenylalanine (R0067)	2	1	0

Second-stage model

For differentiation of the substance/substance group *L-phenylalanine (R0067)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *L-phenylalanine (R0067)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	L-phenylalanine ...	14000033	85169	40	not required
Hepart	L-phenylalanine ...	19000249	85587	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *L-phenylalanine (R0067)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 330 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *L-phenylalanine (R0067)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	L-phenylalanine (R0067)	16001699	85470	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *L-phenylalanine (R0067)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *L-phenylalanine (R0067)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *L-phenylalanine (R0067)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	13 330
Type B	0	40	0	10 049
Type C	0	0	0	3

The substance/substance group *L-phenylalanine (R0067)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9309 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9016 %)	100.0000 % (> 85.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *L-phenylalanine (R0067)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-tryptophan (R0085)	42.36	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *L-phenylalanine (R0067)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85169	85169	0.00	35.25
85587	85587	0.00	33.56

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **L-proline (R0070)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85217-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

L-proline (R0070)

Special notes

When selecting the *L-proline (R0070)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
L-proline (R0070)	1	0	0

Second-stage model

For differentiation of the substance/substance group *L-proline (R0070)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *L-proline (R0070)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	L-proline (R0070)	14001322	85217	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 60 spectra of 1 reference samples from the substance/substance group *L-proline (R0070)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 350 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 60 spectra of 1 reference samples from the substance/substance group *L-proline (R0070)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	L-proline (R0070)	14001322	85220	60

- 10 029 spectra from a total of 236 batches from further 105 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *L-proline (R0070)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *L-proline (R0070)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *L-proline (R0070)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	60	0	13 350
Type B	0	47	13	10 029
Type C	0	0	0	3

The substance/substance group *L-proline (R0070)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9312 %)	100.0000 % (> 90.0000 %)
Type B	100.0000 % (> 99.9011 %)	78.3333 % (> 73.3333 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *L-proline (R0070)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Betaine HCl (R0211)	47.75	—
L-phenylalanine (R0067)	133.66	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *L-proline (R0070)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85217	85217	0.00	47.75

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **L-threonine (R0080)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85006-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

L-threonine (R0080)

Special notes

When selecting the *L-threonine (R0080)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
L-threonine (R0080)	1	0	0

Second-stage model

For differentiation of the substance/substance group *L-threonine (R0080)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *L-threonine (R0080)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	L-threonine (R00...	1209013	85093	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 40 spectra of 1 reference samples from the substance/substance group *L-threonine (R0080)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 370 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *L-threonine (R0080)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *L-threonine (R0080)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *L-threonine (R0080)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *L-threonine (R0080)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	40	0	13 370
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *L-threonine (R0080)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9317 %)	100.0000 % (> 85.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *L-threonine (R0080)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-tyrosine (R0086)	52.63	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *L-threonine (R0080)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85093	85093	0.00	87.61

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **L-tryptophan (R0085)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85068-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

L-tryptophan (R0085)

Special notes

When selecting the *L-tryptophan (R0085)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
L-tryptophan (R0085)	3	3	0

Second-stage model

For differentiation of the substance/substance group *L-tryptophan (R0085)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *L-tryptophan (R0085)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	L-tryptophan (R0...	060801	85068	60	not required
Hepart	L-tryptophan (R0...	15000417	85260	60	not required
Hepart	L-tryptophan (R0...	16000183	85342	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 180 spectra of 3 reference samples from the substance/substance group *L-tryptophan (R0085)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 13 230 spectra from a total of 270 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 3 reference samples from the substance/substance group *L-tryptophan (R0085)*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	L-tryptophan (R0085)	16000876	85407	40
Hepart	L-tryptophan (R0085)	18001233	85546	40
Hepart	L-tryptophan (R0085)	19000251/0	85585	40

- 9969 spectra from a total of 234 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *L-tryptophan (R0085)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *L-tryptophan (R0085)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *L-tryptophan (R0085)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	180	0	13 230
Type B	0	120	0	9969
Type C	0	0	0	3

The substance/substance group *L-tryptophan (R0085)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9305 %)	100.0000 % (> 96.6667 %)
Type B	100.0000 % (> 99.9006 %)	100.0000 % (> 95.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *L-tryptophan (R0085)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-phenylalanine (R0067)	74.78	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group L-tryptophan (R0085) is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85342	85342	0.00	74.19
85260	85260	0.00	73.57
85068	85068	0.00	86.13

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **L-tyrosine (R0086)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85215-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

L-tyrosine (R0086)

Special notes

When selecting the *L-tyrosine (R0086)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
L-tyrosine (R0086)	2	2	0

Second-stage model

For differentiation of the substance/substance group *L-tyrosine (R0086)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *L-tyrosine (R0086)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	L-tyrosine (R0086)	070801	85215	60	not required
Hepart	L-tyrosine (R0086)	14001321	85223	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 120 spectra of 2 reference samples from the substance/substance group *L-tyrosine (R0086)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 290 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 2 reference samples from the substance/substance group *L-tyrosine (R0086)*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	L-tyrosine (R0086)	16000877	85411	40
Hepart	L-tyrosine (R0086)	19002343	85824	40

- 10 009 spectra from a total of 235 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *L-tyrosine (R0086)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *L-tyrosine (R0086)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *L-tyrosine (R0086)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	120	0	13 290
Type B	0	80	0	10 009
Type C	0	0	0	3

The substance/substance group *L-tyrosine (R0086)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9307 %)	100.0000 % (> 95.0000 %)
Type B	100.0000 % (> 99.9009 %)	100.0000 % (> 92.5000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *L-tyrosine (R0086)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-threonine (R0080)	97.07	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *L-tyrosine (R0086)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85215	85215	0.00	97.07
85223	85223	0.00	100.62

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Magnesium citrate nonahydrate (R0056)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85194-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Magnesium citrate nonahydrate (R0056)

Special notes

When selecting the *Magnesium citrate nonahydrate (R0056)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Magnesium citrate nonahydrate (R0056)	2	6	0

Second-stage model

For differentiation of the substance/substance group *Magnesium citrate nonahydrate (R0056)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Magnesium citrate nonahydrate (R0056)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Magnesium citrat...	14000396	85340	60	not required
Hepart	Magnesium citrat...	14001234	85341	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 120 spectra of 2 reference samples from the substance/substance group *Magnesium citrate nonahydrate (R0056)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 290 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 280 spectra of 7 reference samples from the substance/substance group *Magnesium citrate nonahydrate (R0056)*.
- Among them are spectra of independent samples from 6 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Magnesium citrate nonahydrat...	16001519	85447	40
Hepart	Magnesium citrate nonahydrat...	18000588/0	85514	40

continued on the next page

continued from previous page

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Magnesium citrate nonahydrat...	18000588/1	85515	40
Hepart	Magnesium citrate nonahydrat...	18000929	85539	40
Hepart	Magnesium citrate nonahydrat...	19001820	85608	40
Hepart	Magnesium citrate nonahydrat...	20001457	85858	40
Hepart	Magnesium citrate nonahydrat...	16001519	85468	40

- 9809 spectra from a total of 231 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Magnesium citrate nonahydrate (R0056)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Magnesium citrate nonahydrate (R0056)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Magnesium citrate nonahydrate (R0056)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	120	0	13 290
Type B	0	280	0	9809
Type C	0	0	0	3

The substance/substance group *Magnesium citrate nonahydrate (R0056)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9307 %)	100.0000 % (> 95.0000 %)
Type B	100.0000 % (> 99.9004 %)	100.0000 % (> 97.8571 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Magnesium citrate nonahydrate (R0056)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Creatine monohydrate (R0048)	51.25	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Magnesium citrate nonahydrate (R0056)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85340	85340	0.00	59.11
85341	85341	0.00	58.89

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Magnesium glycerophosphate (R0114)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85052-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Magnesium glycerophosphate (R0114)

Special notes

When selecting the *Magnesium glycerophosphate (R0114)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Magnesium glycerophosphate (R0114)	3	1	0

Second-stage model

For differentiation of the substance/substance group *Magnesium glycerophosphate (R0114)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Magnesium glycerophosphate (R0114)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Magnesium glycer...	16000888	85417	40	not required
Hepart	Magnesium glycer...	18000362	85521	40	not required
Hepart	Magnesium glycer...	19000359	85591	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 120 spectra of 3 reference samples from the substance/substance group *Magnesium glycerophosphate (R0114)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 13 290 spectra from a total of 270 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *Magnesium glycerophosphate (R0114)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Magnesium glycerophosphate (...)	18001369	85570	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Magnesium glycerophosphate (R0114)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Magnesium glycerophosphate (R0114)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Magnesium glycerophosphate (R0114)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	120	0	13 290
Type B	0	40	0	10 049
Type C	0	0	0	3

The substance/substance group *Magnesium glycerophosphate (R0114)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9307 %)	100.0000 % (> 95.0000 %)
Type B	100.0000 % (> 99.9016 %)	100.0000 % (> 85.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Magnesium glycerophosphate (R0114)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Garcinia Cambogia (R0028)	22.36	—
Osha root extract (R0220)	26.58	—
Citrus bioflavonoids 60% (R0021)	48.75	—
L-glutamine (R0031)	52.08	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Magnesium glycerophosphate (R0114)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85417	85417	0.00	24.09
85521	85521	0.00	22.36
85591	85591	0.00	22.45

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Magnesium stearate (R0140)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85149-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Magnesium stearate (R0140)

Special notes

When selecting the *Magnesium stearate (R0140)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Magnesium stearate (R0140)	5	3	0

Second-stage model

For differentiation of the substance/substance group *Magnesium stearate (R0140)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Magnesium stearate (R0140)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Magnesium steara...	130503-MS1	85149	30	not required
Hepart	Magnesium steara...	131120001	85156	40	not required
Hepart	Magnesium steara...	131129505	85160	40	not required
Hepart	Magnesium steara...	15001477	85324	60	not required
Hepart	Magnesium steara...	17000218	85460	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 210 spectra of 5 reference samples from the substance/substance group *Magnesium stearate (R0140)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 5 different batches.
- 13 200 spectra from a total of 268 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 110 spectra of 3 reference samples from the substance/substance group *Magnesium stearate (R0140)*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Magnesium stearate (R0140)	16001231	85439	40
Hepart	Magnesium stearate (R0140)	17001092	85483	30
Hepart	Magnesium stearate (R0140)	18000919	85537	40

- 9979 spectra from a total of 234 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Magnesium stearate (R0140)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Magnesium stearate (R0140)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Magnesium stearate (R0140)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	210	0	13 200
Type B	0	101	9	9979
Type C	0	0	0	3

The substance/substance group *Magnesium stearate (R0140)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9304 %)	100.0000 % (> 97.1429 %)
Type B	100.0000 % (> 99.9007 %)	91.8182 % (> 89.0909 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Magnesium stearate (R0140)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Zinc citrate trihydrate (R0089)	125.20	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Magnesium stearate (R0140)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85460	85460	0.00	180.01
85149	85149	0.00	177.10
85324	85324	0.00	176.71
85160	85160	0.00	178.28
85156	85156	0.00	177.78

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Marshmallow root extract (R0204) / Gentian extract (R0200) /**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85275-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Marshmallow root extract (R0204) / Gentian extract (R0200) / Juniper extract (R0192); Gentian extract (R0200); Juniper extract (R0192); Marshmallow root extract (R0204)

Special notes

When selecting the *Marshmallow root extract (R0204) / Gentian extract (R0200) / Juniper extract (R0192)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Gentian extract (R0200)	2	1	0
Juniper extract (R0192)	2	0	0
Marshmallow root extract (R0204)	2	1	0

Second-stage model

For differentiation of the substance/substance group *Marshmallow root extract (R0204)* / *Gentian extract (R0200)* / *Juniper extract (R0192)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Marshmallow root extract (R0204)* / *Gentian extract (R0200)* / *Juniper extract (R0192)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Gentian extract ...	15000798	85299	60	not required
Hepart	Gentian extract ...	16000715	85373	40	not required
Hepart	Juniper extract ...	15000615	85275	60	not required
Hepart	Juniper extract ...	16000596	85383	40	not required
Hepart	Marshmallow root...	15000796	85301	60	not required
Hepart	Marshmallow root...	16000724	85377	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 300 spectra of 6 reference samples from the substance/substance group *Marshmallow root extract (R0204)* / *Gentian extract (R0200)* / *Juniper extract (R0192)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 6 different batches.
- 13 110 spectra from a total of 267 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 2 reference samples from the substance/substance group *Marshmallow root extract (R0204)* / *Gentian extract (R0200)* / *Juniper extract (R0192)*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Gentian extract (R0200)	18001108	85541	40
Hepart	Marshmallow root extract (R0...	18001107	85547	40

- 10 009 spectra from a total of 235 batches from further 104 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Marshmallow root extract (R0204) / Gentian extract (R0200) / Juniper extract (R0192)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Marshmallow root extract (R0204) / Gentian extract (R0200) / Juniper extract (R0192)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Marshmallow root extract (R0204) / Gentian extract (R0200) / Juniper extract (R0192)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	300	0	13 110
Type B	0	80	0	10 009
Type C	0	0	0	3

The substance/substance group *Marshmallow root extract (R0204) / Gentian extract (R0200) / Juniper extract (R0192)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9304 %)	100.0000 % (> 98.0000 %)
Type B	100.0000 % (> 99.9009 %)	100.0000 % (> 92.5000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Marshmallow root extract (R0204) / Gentian extract (R0200) / Juniper extract (R0192)* in the substance class *Herb raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Konjac glucomannan 80% (R0293)	10.35	—
Acacia gum (encapsia) / acacia gum (R0...	11.46	—
Ginseng extract (R0029)	21.10	—
Chili extract (R0198)	25.07	—
Galactomannans (R0035/R0094)	27.89	—
Artichoke extract (R0208)	28.15	—
Hydrangea extract (R0203)	30.25	—
Potato starch, pregelatinized (R0047)	31.04	—
Katuka herb extract (R0210)	31.54	—
Magnesium glycerophosphate (R0114)	33.46	—
Coriolus extract (R0207)	43.23	—
Osha root extract (R0220)	43.89	—
Guarana extract (R0289)	44.53	—
Calcium lactate pentahydrate (R0013)	46.86	—
Calmegh extract (R0201)	54.42	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Marshmallow root extract (R0204) / Gentian extract (R0200) / Juniper extract (R0192)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85299	85299	0.00	25.95
85373	85373	0.00	26.18
85275	85275	0.00	17.79
85301	85301	0.00	28.04
85377	85377	0.00	27.87
85383	85383	0.00	21.10

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Methionine (R0059)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85112-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Methionine (R0059)

Special notes

When selecting the *Methionine (R0059)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Methionine (R0059)	4	2	0

Second-stage model

For differentiation of the substance/substance group *Methionine (R0059)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Methionine (R0059)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Methionine (R0059)	1209010	85112	40	not required
Hepart	Methionine (R0059)	14001123	85212	40	not required
Hepart	Methionine (R0059)	15000268	85245	60	not required
Hepart	Methionine (R0059)	14001830	85246	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 200 spectra of 4 reference samples from the substance/substance group *Methionine (R0059)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 4 different batches.
- 13 210 spectra from a total of 269 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 2 reference samples from the substance/substance group *Methionine (R0059)*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Methionine (R0059)	16000869	85421	40
Hepart	Methionine (R0059)	19002132/0	85827	40

- 10 009 spectra from a total of 235 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Methionine (R0059)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Methionine (R0059)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Methionine (R0059)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	200	0	13 210
Type B	0	80	0	10 009
Type C	0	0	0	3

The substance/substance group *Methionine (R0059)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9305 %)	100.0000 % (> 97.0000 %)
Type B	100.0000 % (> 99.9009 %)	100.0000 % (> 92.5000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Methionine (R0059)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-leucine (R0051)	45.16	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Methionine (R0059)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85246	85246	0.00	55.34
85245	85245	0.00	51.35
85212	85212	0.00	45.16
85112	85112	0.00	52.70

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Methylsulfonylmethane (R0060)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85240-2021-03-22
Executing company HiperScan GmbH
 Weißeritzstraße 3
 01067 Dresden
 Germany

Relevant substance names

Methylsulfonylmethane (R0060)

Special notes

When selecting the *Methylsulfonylmethane (R0060)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Methylsulfonylmethane (R0060)	5	6	0

Second-stage model

For differentiation of the substance/substance group *Methylsulfonylmethane (R0060)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Methylsulfonylmethane (R0060)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Methylsulfonylme...	14001831	85240	60	not required
Hepart	Methylsulfonylme...	15000635	85282	60	not required
Hepart	Methylsulfonylme...	15001758	85334	60	not required
Hepart	Methylsulfonylme...	15001811	85335	60	not required
Hepart	Methylsulfonylme...	16000697	85388	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 280 spectra of 5 reference samples from the substance/substance group *Methylsulfonylmethane (R0060)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 5 different batches.
- 13 130 spectra from a total of 268 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 260 spectra of 6 reference samples from the substance/substance group *Methylsulfonylmethane (R0060)*.
- Among them are spectra of independent samples from 6 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Methylsulfonylmethane (R0060)	16000184	85355	60
Hepart	Methylsulfonylmethane (R0060)	16001669	85451	40
Hepart	Methylsulfonylmethane (R0060)	17000957	85485	40
Hepart	Methylsulfonylmethane (R0060)	17001803	85513	40
Hepart	Methylsulfonylmethane (R0060)	18001230	85553	40
Hepart	Methylsulfonylmethane (R0060)	19000670/0	85594	40

- 9829 spectra from a total of 231 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Methylsulfonylmethane (R0060)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Methylsulfonylmethane (R0060)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Methylsulfonylmethane (R0060)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	280	0	13 130
Type B	0	260	0	9829
Type C	0	0	0	3

The substance/substance group *Methylsulfonylmethane (R0060)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9304 %)	100.0000 % (> 97.8571 %)
Type B	100.0000 % (> 99.9004 %)	100.0000 % (> 97.6923 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Methylsulfonylmethane (R0060)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-proline (R0070)	168.90	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Methylsulfonylmethane (R0060)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85388	85388	0.00	176.22
85335	85335	0.00	174.58
85282	85282	0.00	176.38
85240	85240	0.00	183.57
85334	85334	0.00	168.90

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50% greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **N-acetyl-L-cysteine (R0002)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85057-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

N-acetyl-L-cysteine (R0002)

Special notes

When selecting the *N-acetyl-L-cysteine (R0002)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
N-acetyl-L-cysteine (R0002)	4	4	0

Second-stage model

For differentiation of the substance/substance group *N-acetyl-L-cysteine (R0002)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *N-acetyl-L-cysteine (R0002)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	N-acetyl-L-cyste...	1007-013	85057	60	not required
Hepart	N-acetyl-L-cyste...	14001825	85250	60	not required
Hepart	N-acetyl-L-cyste...	15000267	85251	60	not required
Hepart	N-acetyl-L-cyste...	15000640	85279	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 4 reference samples from the substance/substance group *N-acetyl-L-cysteine (R0002)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 4 different batches.
- 13 170 spectra from a total of 269 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 160 spectra of 4 reference samples from the substance/substance group *N-acetyl-L-cysteine (R0002)*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	N-acetyl-L-cysteine (R0002)	14001118	85199	40
Hepart	N-acetyl-L-cysteine (R0002)	16000863	85410	40
Hepart	N-acetyl-L-cysteine (R0002)	19002344/0	85820	40
Hepart	N-acetyl-L-cysteine (R0002)	19002137/0	85836	40

- 9929 spectra from a total of 233 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *N-acetyl-L-cysteine (R0002)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *N-acetyl-L-cysteine (R0002)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *N-acetyl-L-cysteine (R0002)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	240	0	13 170
Type B	0	160	0	9929
Type C	0	0	0	3

The substance/substance group *N-acetyl-L-cysteine (R0002)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9304 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9005 %)	100.0000 % (> 96.2500 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *N-acetyl-L-cysteine (R0002)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Zinc citrate trihydrate (R0089)	91.64	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *N-acetyl-L-cysteine (R0002)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85250	85250	0.00	91.64
85057	85057	0.00	98.22
85251	85251	0.00	98.98
85279	85279	0.00	98.07

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Nicotinamide (R0064)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85020-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Nicotinamide (R0064)

Special notes

When selecting the *Nicotinamide (R0064)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Nicotinamide (R0064)	2	4	0

Second-stage model

For differentiation of the substance/substance group *Nicotinamide (R0064)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Nicotinamide (R0064)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Nicotinamide (R0...	14001124	85204	40	not required
Hepart	Nicotinamide (R0...	15000235	85254	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 100 spectra of 2 reference samples from the substance/substance group *Nicotinamide (R0064)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 310 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 160 spectra of 4 reference samples from the substance/substance group *Nicotinamide (R0064)*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Nicotinamide (R0064)	16000870	85415	40
Hepart	Nicotinamide (R0064)	16001700	85442	40
Hepart	Nicotinamide (R0064)	18001231	85551	40
Hepart	Nicotinamide (R0064)	19002140/0	85829	40

- 9929 spectra from a total of 233 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Nicotinamide (R0064)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Nicotinamide (R0064)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Nicotinamide (R0064)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	100	0	13 310
Type B	0	160	0	9929
Type C	0	0	0	3

The substance/substance group *Nicotinamide (R0064)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9308 %)	100.0000 % (> 94.0000 %)
Type B	100.0000 % (> 99.9005 %)	100.0000 % (> 96.2500 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Nicotinamide (R0064)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-5-hydroxytryptophan (R0050)	65.35	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Nicotinamide (R0064)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85204	85204	0.00	65.35
85254	85254	0.00	69.85

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Olive leaf extract (R0202)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85298-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Olive leaf extract (R0202)

Special notes

When selecting the *Olive leaf extract (R0202)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Olive leaf extract (R0202)	2	0	0

Second-stage model

For differentiation of the substance/substance group *Olive leaf extract (R0202)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Olive leaf extract (R0202)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Olive leaf extra. ...	15000789	85298	60	not required
Hepart	Olive leaf extra. ...	16000717	85364	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 100 spectra of 2 reference samples from the substance/substance group *Olive leaf extract (R0202)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 310 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Olive leaf extract (R0202)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Olive leaf extract (R0202)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Olive leaf extract (R0202)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Olive leaf extract (R0202)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	100	0	13 310
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Olive leaf extract (R0202)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9308 %)	100.0000 % (> 94.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Olive leaf extract (R0202)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Citrus bioflavonoids 60% (R0021)	20.84	—
Rhodiola rosea extract 3% (R0159)	21.72	—
Calcium lactate pentahydrate (R0013)	51.48	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Olive leaf extract (R0202)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85364	85364	0.00	21.72
85298	85298	0.00	27.58

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50% greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **OPC (grape seed) extract (R0065)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85034-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

OPC (grape seed) extract (R0065)

Special notes

When selecting the *OPC (grape seed) extract (R0065)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
OPC (grape seed) extract (R0065)	2	0	0

Second-stage model

For differentiation of the substance/substance group *OPC (grape seed) extract (R0065)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *OPC (grape seed) extract (R0065)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	OPC (grape seed)...	059601	85121	40	not required
Hepart	OPC (grape seed)...	130430-TE13	85230	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 100 spectra of 2 reference samples from the substance/substance group *OPC (grape seed) extract (R0065)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 310 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *OPC (grape seed) extract (R0065)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *OPC (grape seed) extract (R0065)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *OPC (grape seed) extract (R0065)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *OPC (grape seed) extract (R0065)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	100	0	13 310
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *OPC (grape seed) extract (R0065)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9308 %)	100.0000 % (> 94.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *OPC (grape seed) extract (R0065)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Green tea extract (R0034)	20.06	—
Citric acid monohydrate (R0143)	49.14	—
Citrus bioflavonoids 60% (R0021)	80.83	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *OPC (grape seed) extract (R0065)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85230	85230	0.00	48.31
85121	85121	0.00	20.06

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50% greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group	Oregano extract (R0206)
Substance class	Hepart raw materials
Report date	22/03/2021
Report number	85293-2021-03-22
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Oregano extract (R0206)

Special notes

When selecting the *Oregano extract (R0206)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Oregano extract (R0206)	2	0	0

Second-stage model

For differentiation of the substance/substance group *Oregano extract (R0206)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Oregano extract (R0206)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Oregano extract ...	15000790	85293	60	not required
Hepart	Oregano extract ...	16000720	85365	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 100 spectra of 2 reference samples from the substance/substance group *Oregano extract (R0206)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 310 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Oregano extract (R0206)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Oregano extract (R0206)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Oregano extract (R0206)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Oregano extract (R0206)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	100	0	13 310
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Oregano extract (R0206)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9308 %)	100.0000 % (> 94.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Oregano extract (R0206)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Hyssop extract (R0205)	13.32	—
Black cumin extract (R0193)	19.95	—
Walnut leaf extract (R0197)	28.41	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Garcinia Cambogia (R0028)	31.19	—
Osha root extract (R0220)	32.32	—
Magnesium glycerophosphate (R0114)	33.50	—
Coriolus extract (R0207)	34.12	—
Citrus bioflavonoids 60% (R0021)	35.57	—
Artichoke extract (R0208)	46.19	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Oregano extract (R0206)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85365	85365	0.00	13.32
85293	85293	0.00	16.62

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group	Osha root extract (R0220)
Substance class	Hepart raw materials
Report date	22/03/2021
Report number	85314-2021-03-22
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Osha root extract (R0220)

Special notes

When selecting the *Osha root extract (R0220)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Osha root extract (R0220)	2	0	0

Second-stage model

For differentiation of the substance/substance group *Osha root extract (R0220)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Osha root extract (R0220)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Osha root extrac...	15000939	85314	60	not required
Hepart	Osha root extrac...	16000546	85393	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 100 spectra of 2 reference samples from the substance/substance group *Osha root extract (R0220)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 310 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Osha root extract (R0220)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Osha root extract (R0220)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Osha root extract (R0220)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Osha root extract (R0220)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	100	0	13 310
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Osha root extract (R0220)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9308 %)	100.0000 % (> 94.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Osha root extract (R0220)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Magnesium glycerophosphate (R0114)	11.37	—
Garcinia Cambogia (R0028)	13.48	—
Chili extract (R0198)	31.30	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Juniper extract (R0192)	44.08	—
Calmegh extract (R0201)	45.09	—
Citrus bioflavonoids 60% (R0021)	48.44	—
Coriolus extract (R0207)	49.51	—
Oregano extract (R0206)	50.37	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Osha root extract (R0220)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85314	85314	0.00	19.23
85393	85393	0.00	13.48

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **p-aminobenzoic acid (R0004)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85065-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

p-aminobenzoic acid (R0004)

Special notes

When selecting the *p-aminobenzoic acid (R0004)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
p-aminobenzoic acid (R0004)	4	3	0

Second-stage model

For differentiation of the substance/substance group *p-aminobenzoic acid (R0004)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *p-aminobenzoic acid (R0004)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	p-aminobenzoic a...	0912-11	85065	60	not required
Hepart	p-aminobenzoic a...	14001832	85247	60	not required
Hepart	p-aminobenzoic a...	15000235	85257	60	not required
Hepart	p-aminobenzoic a...	15001755	85333	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 4 reference samples from the substance/substance group *p-aminobenzoic acid (R0004)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 4 different batches.
- 13 170 spectra from a total of 270 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 3 reference samples from the substance/substance group *p-aminobenzoic acid (R0004)*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	p-aminobenzoic acid (R0004)	16000864	85412	40
Hepart	p-aminobenzoic acid (R0004)	16001691	85467	40
Hepart	p-aminobenzoic acid (R0004)	19002136/0	85825	40

- 9969 spectra from a total of 234 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *p-aminobenzoic acid (R0004)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *p-aminobenzoic acid (R0004)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *p-aminobenzoic acid (R0004)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	240	0	13 170
Type B	0	120	0	9969
Type C	0	0	0	3

The substance/substance group *p-aminobenzoic acid (R0004)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9304 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9006 %)	100.0000 % (> 95.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *p-aminobenzoic acid (R0004)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-carnitine-L-tartrate (R0014)	133.23	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *p-aminobenzoic acid (R0004)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85065	85065	0.00	165.76
85247	85247	0.00	161.24
85257	85257	0.00	167.41
85333	85333	0.00	163.52

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Pea protein (R0092)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 10003715-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Pea protein (R0092)

Special notes

When selecting the *Pea protein (R0092)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Pea protein (R0092)	1	1	0

Second-stage model

For differentiation of the substance/substance group *Pea protein (R0092)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Pea protein (R0092)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Pea protein (R00...	19001584	85620	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 40 spectra of 1 reference samples from the substance/substance group *Pea protein (R0092)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 370 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *Pea protein (R0092)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Pea protein (R0092)	20001060	85842	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Pea protein (R0092)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Pea protein (R0092)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Pea protein (R0092)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	40	0	13 370
Type B	0	40	0	10 049
Type C	0	0	0	3

The substance/substance group *Pea protein (R0092)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9317 %)	100.0000 % (> 85.0000 %)
Type B	100.0000 % (> 99.9016 %)	100.0000 % (> 85.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Pea protein (R0092)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Beta carotene 20% (R0008)	26.93	—
L-glutamine (R0031)	33.08	—
Calcium caseinate (Lactonate RCE) (R02...	49.54	—
Cyanocobalamin (R0024)	56.41	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Pea protein (R0092)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85620	85620	0.00	29.75

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Piperine 95% extract (R0069)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85189-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Piperine 95% extract (R0069)

Special notes

When selecting the *Piperine 95% extract (R0069)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Piperine 95% extract (R0069)	1	1	0

Second-stage model

For differentiation of the substance/substance group *Piperine 95% extract (R0069)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Piperine 95% extract (R0069)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Piperine 95% ext...	130816502	85189	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 40 spectra of 1 reference samples from the substance/substance group *Piperine 95% extract (R0069)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 370 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *Piperine 95% extract (R0069)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Piperine 95% extract (R0069)	16000742	85403	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Piperine 95% extract (R0069)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Piperine 95% extract (R0069)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Piperine 95% extract (R0069)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	40	0	13 370
Type B	0	39	1	10 049
Type C	0	0	0	3

The substance/substance group *Piperine 95% extract (R0069)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9317 %)	100.0000 % (> 85.0000 %)
Type B	100.0000 % (> 99.9016 %)	97.5000 % (> 90.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Piperine 95% extract (R0069)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-proline (R0070)	78.59	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Piperine 95% extract (R0069)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85189	85189	0.00	78.59

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Progesterone, micronized (R0108)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85021-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Progesterone, micronized (R0108)

Special notes

When selecting the *Progesterone, micronized (R0108)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Progesterone, micronized (R0108)	2	0	0

Second-stage model

For differentiation of the substance/substance group *Progesterone, micronized (R0108)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Progesterone, micronized (R0108)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Progesterone, mi...	11177211	85079	50	not required
Hepart	Progesterone, mi...	19000561	85599	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 90 spectra of 2 reference samples from the substance/substance group *Progesterone, micronized (R0108)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 320 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Progesterone, micronized (R0108)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Progesterone, micronized (R0108)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Progesterone, micronized (R0108)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Progesterone, micronized (R0108)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	90	0	13 320
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Progesterone, micronized (R0108)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9308 %)	100.0000 % (> 93.3333 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Progesterone, micronized (R0108)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Coenzyme Q10 (R0022)	35.10	—
a-lipoic acid (R0003)	98.19	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Progesterone, micronized (R0108)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85079	85079	0.00	39.23
85599	85599	0.00	41.15

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Pumpkin seed protein (R0306)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 10002314-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Pumpkin seed protein (R0306)

Special notes

When selecting the *Pumpkin seed protein (R0306)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Pumpkin seed protein (R0306)	1	1	0

Second-stage model

For differentiation of the substance/substance group *Pumpkin seed protein (R0306)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Pumpkin seed protein (R0306)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Pumpkin seed pro...	20002519	85871	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 40 spectra of 1 reference samples from the substance/substance group *Pumpkin seed protein (R0306)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 370 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *Pumpkin seed protein (R0306)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Pumpkin seed protein (R0306)	19002077/0	85618	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Pumpkin seed protein (R0306)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Pumpkin seed protein (R0306)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Pumpkin seed protein (R0306)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	40	0	13 370
Type B	0	0	40	10 049
Type C	0	0	0	3

The substance/substance group *Pumpkin seed protein (R0306)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9317 %)	100.0000 % (> 85.0000 %)
Type B	100.0000 % (> 99.9016 %)	0.0000 % (≥ 0.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Pumpkin seed protein (R0306)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Beta carotene 20% (R0008)	32.80	—
L-glutamine (R0031)	33.26	—
Lutein 20% (R0052)	44.21	—
Isoflavones 10% (R0042)	48.84	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Pumpkin seed protein (R0306)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85871	85871	0.00	39.80

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Pyridoxal-5-phosphate (R0195)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85269-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Pyridoxal-5-phosphate (R0195)

Special notes

When selecting the *Pyridoxal-5-phosphate (R0195)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Pyridoxal-5-phosphate (R0195)	2	0	0

Second-stage model

For differentiation of the substance/substance group *Pyridoxal-5-phosphate (R0195)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Pyridoxal-5-phosphate (R0195)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Pyridoxal-5-phos...	15000618	85269	60	not required
Hepart	Pyridoxal-5-phos...	16000705	85400	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 100 spectra of 2 reference samples from the substance/substance group *Pyridoxal-5-phosphate (R0195)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 310 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Pyridoxal-5-phosphate (R0195)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Pyridoxal-5-phosphate (R0195)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Pyridoxal-5-phosphate (R0195)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Pyridoxal-5-phosphate (R0195)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	100	0	13 310
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Pyridoxal-5-phosphate (R0195)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9308 %)	100.0000 % (> 94.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Pyridoxal-5-phosphate (R0195)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Ascorbic acid (R0007)	28.08	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Pyridoxal-5-phosphate (R0195)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85400	85400	0.00	28.08
85269	85269	0.00	32.11

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Pyridoxine hydrochloride (vitamin B6) (R0071)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85022-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Pyridoxine hydrochloride (vitamin B6) (R0071)

Special notes

When selecting the *Pyridoxine hydrochloride (vitamin B6) (R0071)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Pyridoxine hydrochloride (vitamin B6) (R0071)	3	5	0

Second-stage model

For differentiation of the substance/substance group *Pyridoxine hydrochloride (vitamin B6) (R0071)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Pyridoxine hydrochloride (vitamin B6) (R0071)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Pyridoxine hydro...	1209016	85111	40	not required
Hepart	Pyridoxine hydro...	15000236	85238	60	not required
Hepart	Pyridoxine hydro...	14001827	85249	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 3 reference samples from the substance/substance group *Pyridoxine hydrochloride (vitamin B6) (R0071)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 13 250 spectra from a total of 270 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 200 spectra of 5 reference samples from the substance/substance group *Pyridoxine hydrochloride (vitamin B6) (R0071)*.
- Among them are spectra of independent samples from 5 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Pyridoxine hydrochloride (vi...	16000872	85409	40

continued on the next page

continued from previous page

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Pyrodoxine hydrochloride (vi...	16001137	85429	40
Hepart	Pyrodoxine hydrochloride (vi...	16001689	85471	40
Hepart	Pyrodoxine hydrochloride (vi...	18001232	85555	40
Hepart	Pyrodoxine hydrochloride (vi...	19002130/0	85831	40

- 9889 spectra from a total of 232 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Pyrodoxine hydrochloride (vitamin B6) (R0071)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Pyrodoxine hydrochloride (vitamin B6) (R0071)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Pyrodoxine hydrochloride (vitamin B6) (R0071)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	13 250
Type B	0	161	39	9889
Type C	0	0	0	3

The substance/substance group *Pyrodoxine hydrochloride (vitamin B6) (R0071)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9305 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9004 %)	80.5000 % (> 79.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Pyridoxine hydrochloride (vitamin B6) (R0071)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-5-hydroxytryptophan (R0050)	174.23	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Pyridoxine hydrochloride (vitamin B6) (R0071)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85238	85238	0.00	175.90
85249	85249	0.00	175.52
85111	85111	0.00	174.23

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Resveratrol (R0339)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 10004601-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Resveratrol (R0339)

Special notes

When selecting the *Resveratrol (R0339)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Resveratrol (R0339)	1	0	0

Second-stage model

For differentiation of the substance/substance group *Resveratrol (R0339)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Resveratrol (R0339)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Resveratrol (R03...	20001466	85872	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 40 spectra of 1 reference samples from the substance/substance group *Resveratrol (R0339)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 370 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Resveratrol (R0339)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Resveratrol (R0339)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Resveratrol (R0339)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Resveratrol (R0339)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	40	0	13 370
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Resveratrol (R0339)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9317 %)	100.0000 % (> 85.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Resveratrol (R0339)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-5-hydroxytryptophan (R0050)	49.29	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Resveratrol (R0339)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85872	85872	0.00	49.29

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group	Retinyl palmitate (vitamin A palmitate) (R0073)
Substance class	Hepart raw materials
Report date	22/03/2021
Report number	85237-2021-03-22
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Retinyl palmitate (vitamin A palmitate) (R0073)

Special notes

When selecting the *Retinyl palmitate (vitamin A palmitate) (R0073)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Retinyl palmitate (vitamin A palmitate) (R0073)	3	2	0

Second-stage model

For differentiation of the substance/substance group *Retinyl palmitate (vitamin A palmitate) (R0073)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Retinyl palmitate (vitamin A palmitate) (R0073)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Retinyl palmitat...	14001530	85237	60	not required
Hepart	Retinyl palmitat...	16000747	85369	40	not required
Hepart	Retinyl palmitat...	17000435	85481	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 140 spectra of 3 reference samples from the substance/substance group *Retinyl palmitate (vitamin A palmitate) (R0073)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 13 270 spectra from a total of 270 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 2 reference samples from the substance/substance group *Retinyl palmitate (vitamin A palmitate) (R0073)*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Retinyl palmitate (vitamin A...	17001022	85480	40
Hepart	Retinyl palmitate (vitamin A...	17001255	85500	40

- 10 009 spectra from a total of 235 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Retinyl palmitate (vitamin A palmitate) (R0073)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Retinyl palmitate (vitamin A palmitate) (R0073)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Retinyl palmitate (vitamin A palmitate) (R0073)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	140	0	13 270
Type B	0	80	0	10 009
Type C	0	0	0	3

The substance/substance group *Retinyl palmitate (vitamin A palmitate) (R0073)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9306 %)	100.0000 % (> 95.7143 %)
Type B	100.0000 % (> 99.9009 %)	100.0000 % (> 92.5000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Retinyl palmitate (vitamin A palmitate) (R0073)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Menaquinone-7 0.2% (R0134)	21.01	—
Sunflower oil powder 25 (R0319)	24.72	—
Ginger extract (R0199)	25.55	—
Lutein 20% (R0052)	42.82	—
Cholecalciferol (vitamin D3) (R0018)	46.18	—
Ginseng extract (R0029)	47.70	—
Beta carotene 10% (R0237)	51.18	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Retinyl palmitate (vitamin A palmitate) (R0073)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85481	85481	0.00	26.25
85369	85369	0.00	23.32
85237	85237	0.00	21.01

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Rhodiola rosea extract 3% (R0159)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85172-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Rhodiola rosea extract 3% (R0159)

Special notes

When selecting the *Rhodiola rosea extract 3% (R0159)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Rhodiola rosea extract 3% (R0159)	5	0	0

Second-stage model

For differentiation of the substance/substance group *Rhodiola rosea extract 3% (R0159)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Rhodiola rosea extract 3% (R0159)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Rhodiola rosea e...	14000090	85172	40	not required
Hepart	Rhodiola rosea e...	16001514	85459	40	not required
Hepart	Rhodiola rosea e...	18001234	85550	40	not required
Hepart	Rhodiola rosea e...	19002135/0	85832	40	not required
Hepart	Rhodiola rosea e...	20001464	85856	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 200 spectra of 5 reference samples from the substance/substance group *Rhodiola rosea extract 3% (R0159)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 5 different batches.
- 13 210 spectra from a total of 268 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *Rhodiola rosea extract 3% (R0159)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Rhodiola rosea extract 3% (R...	16001514	85448	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Rhodiola rosea extract 3% (R0159)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Rhodiola rosea extract 3% (R0159)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Rhodiola rosea extract 3% (R0159)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	200	0	13 210
Type B	0	40	0	10 049
Type C	0	0	0	3

The substance/substance group *Rhodiola rosea extract 3% (R0159)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9305 %)	100.0000 % (> 97.0000 %)
Type B	100.0000 % (> 99.9016 %)	100.0000 % (> 85.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Rhodiola rosea extract 3% (R0159)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Citrus bioflavonoids 60% (R0021)	28.06	—
Olive leaf extract (R0202)	31.27	—
Calcium lactate pentahydrate (R0013)	31.77	—
Hydrangea extract (R0203)	41.81	—
Citric acid monohydrate (R0143)	51.06	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Rhodiola rosea extract 3% (R0159)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85550	85550	0.00	31.27
85856	85856	0.00	43.20
85832	85832	0.00	35.12
85459	85459	0.00	36.09
85172	85172	0.00	34.14

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Riboflavin (vit. B2) (R0074)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85013-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Riboflavin (vit. B2) (R0074)

Special notes

When selecting the *Riboflavin (vit. B2) (R0074)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Riboflavin (vit. B2) (R0074)	6	1	0

Second-stage model

For differentiation of the substance/substance group *Riboflavin (vit. B2) (R0074)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Riboflavin (vit. B2) (R0074)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Riboflavin (vit...	1209014	85116	40	not required
Hepart	Riboflavin (vit...	14001828	85236	60	not required
Hepart	Riboflavin (vit...	15000237	85243	60	not required
Hepart	Riboflavin (vit...	15001035	85311	60	not required
Hepart	Riboflavin (vit...	16000523	85391	40	not required
Hepart	Riboflavin (vit...	19002138	85819	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 300 spectra of 6 reference samples from the substance/substance group *Riboflavin (vit. B2) (R0074)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 6 different batches.
- 13 110 spectra from a total of 267 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *Riboflavin (vit. B2) (R0074)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Riboflavin (vit. B2) (R0074)	16000873	85404	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Riboflavin (vit. B2) (R0074)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Riboflavin (vit. B2) (R0074)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Riboflavin (vit. B2) (R0074)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	300	0	13 110
Type B	0	40	0	10 049
Type C	0	0	0	3

The substance/substance group *Riboflavin (vit. B2) (R0074)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9304 %)	100.0000 % (> 98.0000 %)
Type B	100.0000 % (> 99.9016 %)	100.0000 % (> 85.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Riboflavin (vit. B2) (R0074)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Ginkgo extract (R0093)	65.15	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Riboflavin (vit. B2) (R0074)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85391	85391	0.00	65.15
85236	85236	0.00	90.42
85243	85243	0.00	93.57
85311	85311	0.00	90.52
85116	85116	0.00	77.82
85819	85819	0.00	76.05

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50% greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Sodium borate (R0191)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85271-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Sodium borate (R0191)

Special notes

When selecting the *Sodium borate (R0191)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Sodium borate (R0191)	2	2	0

Second-stage model

For differentiation of the substance/substance group *Sodium borate (R0191)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Sodium borate (R0191)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Sodium borate (R...	15000621	85271	60	not required
Hepart	Sodium borate (R...	15000999	85306	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 120 spectra of 2 reference samples from the substance/substance group *Sodium borate (R0191)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 290 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 3 reference samples from the substance/substance group *Sodium borate (R0191)*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Sodium borate (R0191)	16000597	85381	40
Hepart	Sodium borate (R0191)	17000954	85493	40
Hepart	Sodium borate (R0191)	16000597	85426	40

- 9969 spectra from a total of 235 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Sodium borate (R0191)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Sodium borate (R0191)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Sodium borate (R0191)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	120	0	13 290
Type B	0	120	0	9969
Type C	0	0	0	3

The substance/substance group *Sodium borate (R0191)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9307 %)	100.0000 % (> 95.0000 %)
Type B	100.0000 % (> 99.9006 %)	100.0000 % (> 95.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Sodium borate (R0191)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-arginine base (R0006)	49.55	—
Citrus bioflavonoids 60% (R0021)	62.83	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Sodium borate (R0191)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85271	85271	0.00	54.06
85306	85306	0.00	49.56

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Sodium chondroitin sulfate (R0061)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85002-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Sodium chondroitin sulfate (R0061); Sodium chondroitin sulphate (R0061)

Special notes

When selecting the *Sodium chondroitin sulfate (R0061)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Sodium chondroitin sulfate (R0061)	1	1	0

Second-stage model

For differentiation of the substance/substance group *Sodium chondroitin sulfate (R0061)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Sodium chondroitin sulfate (R0061)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Sodium chondroit...	16000186	85338	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 60 spectra of 1 reference samples from the substance/substance group *Sodium chondroitin - sulfate (R0061)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 350 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *Sodium chondroitin sulfate (R0061)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Sodium chondroitin sulfate (...)	17001310	85505	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Sodium chondroitin sulfate (R0061)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Sodium chondroitin sulfate (R0061)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Sodium chondroitin sulfate (R0061)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	60	0	13 350
Type B	0	0	40	10 049
Type C	0	0	0	3

The substance/substance group *Sodium chondroitin sulfate (R0061)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9312 %)	100.0000 % (> 90.0000 %)
Type B	100.0000 % (> 99.9016 %)	0.0000 % (≥ 0.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Sodium chondroitin sulfate (R0061)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Sodium borate (R0191)	52.92	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Sodium chondroitin sulfate (R0061)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85338	85338	0.00	52.92

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group	Sunflower oil powder 25 (R0319)
Substance class	Hepart raw materials
Report date	22/03/2021
Report number	10004602-2021-03-22
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Sunflower oil powder 25 (R0319)

Special notes

When selecting the *Sunflower oil powder 25 (R0319)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Sunflower oil powder 25 (R0319)	1	0	0

Second-stage model

For differentiation of the substance/substance group *Sunflower oil powder 25 (R0319)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Sunflower oil powder 25 (R0319)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Sunflower oil po...	20001971	85873	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 40 spectra of 1 reference samples from the substance/substance group *Sunflower oil powder 25 (R0319)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 370 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Sunflower oil powder 25 (R0319)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Sunflower oil powder 25 (R0319)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Sunflower oil powder 25 (R0319)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Sunflower oil powder 25 (R0319)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	40	0	13 370
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Sunflower oil powder 25 (R0319)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9317 %)	100.0000 % (> 85.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Sunflower oil powder 25 (R0319)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Menaquinone-7 0.2% (R0134)	21.82	—
Retinyl palmitate (vitamin A palmitate...	31.15	—
Ginger extract (R0199)	44.29	—
Lutein 20% (R0052)	61.84	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Sunflower oil powder 25 (R0319)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85873	85873	0.00	21.82

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Taurine (R0078)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85018-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Taurine (R0078)

Special notes

When selecting the *Taurine (R0078)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Taurine (R0078)	3	5	0

Second-stage model

For differentiation of the substance/substance group *Taurine (R0078)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Taurine (R0078)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Taurine (R0078)	1209008	85102	40	not required
Hepart	Taurine (R0078)	15000276	85252	60	not required
Hepart	Taurine (R0078)	19002131/0	85823	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 140 spectra of 3 reference samples from the substance/substance group *Taurine (R0078)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 13 270 spectra from a total of 270 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 220 spectra of 5 reference samples from the substance/substance group *Taurine (R0078)*.
- Among them are spectra of independent samples from 5 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Taurine (R0078)	16000182	85357	60
Hepart	Taurine (R0078)	16000874	85402	40

continued on the next page

continued from previous page

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Taurine (R0078)	16001668	85440	40
Hepart	Taurine (R0078)	18000515	85520	40
Hepart	Taurine (R0078)	18002045	85582	40

- 9869 spectra from a total of 232 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Taurine (R0078)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Taurine (R0078)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Taurine (R0078)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	140	0	13 270
Type B	0	220	0	9869
Type C	0	0	0	3

The substance/substance group *Taurine (R0078)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9306 %)	100.0000 % (> 95.7143 %)
Type B	100.0000 % (> 99.9004 %)	100.0000 % (> 97.2727 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Taurine (R0078)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Choline chloride (R0019)	22.60	—
L-5-hydroxytryptophan (R0050)	215.16	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Taurine (R0078)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85252	85252	0.00	24.30
85823	85823	0.00	24.96
85102	85102	0.00	22.60

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Thiamine HCl (Vitamin B1) (R0079)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85015-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Thiamine HCl (Vitamin B1) (R0079)

Special notes

When selecting the *Thiamine HCl (Vitamin B1) (R0079)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Thiamine HCl (Vitamin B1) (R0079)	6	3	0

Second-stage model

For differentiation of the substance/substance group *Thiamine HCl (Vitamin B1) (R0079)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Thiamine HCl (Vitamin B1) (R0079)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Thiamine HCl (Vi...	1209015	85113	40	not required
Hepart	Thiamine HCl (Vi...	131203001	85157	40	not required
Hepart	Thiamine HCl (Vi...	14001197	85222	60	not required
Hepart	Thiamine HCl (Vi...	14001826	85239	60	not required
Hepart	Thiamine HCl (Vi...	15000419	85262	60	not required
Hepart	Thiamine HCl (Vi...	15001037	85313	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 320 spectra of 6 reference samples from the substance/substance group *Thiamine HCl (Vitamin B1) (R0079)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 6 different batches.
- 13 090 spectra from a total of 267 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 3 reference samples from the substance/substance group *Thiamine HCl (Vitamin B1) (R0079)*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Thiamine HCl (Vitamin B1) (R...	16000875	85423	40
Hepart	Thiamine HCl (Vitamin B1) (R...	16001534	85446	40
Hepart	Thiamine HCl (Vitamin B1) (R...	19002134/0	85624	40

- 9969 spectra from a total of 234 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Thiamine HCl (Vitamin B1) (R0079)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Thiamine HCl (Vitamin B1) (R0079)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Thiamine HCl (Vitamin B1) (R0079)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	320	0	13 090
Type B	0	120	0	9969
Type C	0	0	0	3

The substance/substance group *Thiamine HCl (Vitamin B1) (R0079)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9303 %)	100.0000 % (> 98.1250 %)
Type B	100.0000 % (> 99.9006 %)	100.0000 % (> 95.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Thiamine HCl (Vitamin B1) (R0079)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Glutathione (reduced form) (R0032)	69.61	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Thiamine HCl (Vitamin B1) (R0079)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85157	85157	0.00	78.23
85113	85113	0.00	90.53
85239	85239	0.00	79.74
85222	85222	0.00	124.82
85313	85313	0.00	130.21
85262	85262	0.00	69.61

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Thiamine mononitrate (R0212)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85305-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Thiamine mononitrate (R0212)

Special notes

When selecting the *Thiamine mononitrate (R0212)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Thiamine mononitrate (R0212)	1	0	0

Second-stage model

For differentiation of the substance/substance group *Thiamine mononitrate (R0212)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Thiamine mononitrate (R0212)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Thiamine mononit...	15000748	85305	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 60 spectra of 1 reference samples from the substance/substance group *Thiamine mononitrate (R0212)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 350 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Thiamine mononitrate (R0212)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Thiamine mononitrate (R0212)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Thiamine mononitrate (R0212)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Thiamine mononitrate (R0212)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	60	0	13 350
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Thiamine mononitrate (R0212)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9312 %)	100.0000 % (> 90.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Thiamine mononitrate (R0212)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Riboflavin (vit. B2) (R0074)	102.73	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Thiamine mononitrate (R0212)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85305	85305	0.00	104.60

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Tripotassium citrate monohydrate (R0117)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85268-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Tripotassium citrate monohydrate (R0117)

Special notes

When selecting the *Tripotassium citrate monohydrate (R0117)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Tripotassium citrate monohydrate (R0117)	3	4	0

Second-stage model

For differentiation of the substance/substance group *Tripotassium citrate monohydrate (R0117)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Tripotassium citrate monohydrate (R0117)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Tripotassium cit...	15000230	85268	60	not required
Hepart	Tripotassium cit...	16000736	85374	40	not required
Hepart	Tripotassium cit...	17000934	85482	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 140 spectra of 3 reference samples from the substance/substance group *Tripotassium citrate monohydrate (R0117)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 13 270 spectra from a total of 270 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 160 spectra of 4 reference samples from the substance/substance group *Tripotassium citrate monohydrate (R0117)*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Tripotassium citrate monohyd...	17000186	85462	40

continued on the next page

continued from previous page

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Tripotassium citrate monohyd...	18000459	85519	40
Hepart	Tripotassium citrate monohyd...	18000377	85538	40
Hepart	Tripotassium citrate monohyd...	18002048	85581	40

- 9929 spectra from a total of 233 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Tripotassium citrate monohydrate (R0117)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Tripotassium citrate monohydrate (R0117)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Tripotassium citrate monohydrate (R0117)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	140	0	13 270
Type B	0	160	0	9929
Type C	0	0	0	3

The substance/substance group *Tripotassium citrate monohydrate (R0117)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9306 %)	100.0000 % (> 95.7143 %)
Type B	100.0000 % (> 99.9005 %)	100.0000 % (> 96.2500 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Tripotassium citrate monohydrate (R0117)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Creatine monohydrate (R0048)	17.06	–
D-galactose (R0135)	76.21	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Tripotassium citrate monohydrate (R0117)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85482	85482	0.00	36.36
85374	85374	0.00	39.71
85268	85268	0.00	39.61

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Valine (R0087)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85069-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Valine (R0087)

Special notes

When selecting the *Valine (R0087)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Valine (R0087)	3	1	0

Second-stage model

For differentiation of the substance/substance group *Valine (R0087)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Valine (R0087)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Valine (R0087)	057601	85069	60	not required
Hepart	Valine (R0087)	057701	85075	50	not required
Hepart	Valine (R0087)	14000230	85221	60	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 170 spectra of 3 reference samples from the substance/substance group *Valine (R0087)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 13 240 spectra from a total of 270 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 1 reference samples from the substance/substance group *Valine (R0087)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Valine (R0087)	16001533	85443	40

- 10 049 spectra from a total of 236 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Valine (R0087)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Valine (R0087)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Valine (R0087)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	170	0	13 240
Type B	0	40	0	10 049
Type C	0	0	0	3

The substance/substance group *Valine (R0087)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9305 %)	100.0000 % (> 96.4706 %)
Type B	100.0000 % (> 99.9016 %)	100.0000 % (> 85.0000 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Valine (R0087)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
L-isoleucine (R0043)	40.34	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Valine (R0087)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85069	85069	0.00	44.61
85221	85221	0.00	40.34
85075	85075	0.00	64.00

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Vanilla flavoring (natural flavor) (R0240)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 10004603-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Vanilla flavoring (natural flavor) (R0240)

Special notes

When selecting the *Vanilla flavoring (natural flavor) (R0240)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Vanilla flavoring (natural flavor) (R0240)	1	0	0

Second-stage model

For differentiation of the substance/substance group *Vanilla flavoring (natural flavor) (R0240)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Vanilla flavoring (natural flavor) (R0240)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Vanilla flavorin...	20001760	85874	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 40 spectra of 1 reference samples from the substance/substance group *Vanilla flavoring (natural flavor) (R0240)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 13 370 spectra from a total of 272 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Vanilla flavoring (natural flavor) (R0240)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Vanilla flavoring (natural flavor) (R0240)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Vanilla flavoring (natural flavor) (R0240)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Vanilla flavoring (natural flavor) (R0240)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	40	0	13 370
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Vanilla flavoring (natural flavor) (R0240)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9317 %)	100.0000 % (> 85.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Vanilla flavoring (natural flavor) (R0240)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Lutein 20% (R0052)	36.50	—
Citrus bioflavonoids 60% (R0021)	41.79	—
Sodium chondroitin sulfate (R0061)	45.42	—
Acacia gum (encapsia) / acacia gum (R0...	47.13	—
beta-glucan 28% (oat fibre) (R0334)	60.20	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Vanilla flavoring (natural flavor) (R0240)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85874	85874	0.00	36.50

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Walnut leaf extract (R0197)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85296-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Walnut leaf extract (R0197)

Special notes

When selecting the *Walnut leaf extract (R0197)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Walnut leaf extract (R0197)	2	0	0

Second-stage model

For differentiation of the substance/substance group *Walnut leaf extract (R0197)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Walnut leaf extract (R0197)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Walnut leaf extr...	15000791	85296	60	not required
Hepart	Walnut leaf extr...	16000712	85376	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 100 spectra of 2 reference samples from the substance/substance group *Walnut leaf extract (R0197)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 310 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Walnut leaf extract (R0197)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Walnut leaf extract (R0197)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Walnut leaf extract (R0197)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Walnut leaf extract (R0197)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	100	0	13 310
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Walnut leaf extract (R0197)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9308 %)	100.0000 % (> 94.0000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Walnut leaf extract (R0197)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Magnesium glycerophosphate (R0114)	30.23	—
Oregano extract (R0206)	31.92	—
Hyssop extract (R0205)	35.95	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Artichoke extract (R0208)	36.45	—
Osha root extract (R0220)	36.64	—
Garcinia Cambogia (R0028)	37.10	—
Coriolus extract (R0207)	44.18	—
Katuka herb extract (R0210)	50.01	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Walnut leaf extract (R0197)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85376	85376	0.00	37.10
85296	85296	0.00	31.92

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Yeast extract A (R0292)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 10004599-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Yeast extract A (R0292)

Special notes

When selecting the *Yeast extract A (R0292)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Yeast extract A (R0292)	2	0	0

Second-stage model

For differentiation of the substance/substance group *Yeast extract A (R0292)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Yeast extract A (R0292)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Yeast extract A ...	20002318	85868	40	not required
Hepart	Yeast extract A ...	20001621	85869	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Yeast extract A (R0292)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 13 330 spectra from a total of 271 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 0 spectra of 0 reference samples from the substance/substance group *Yeast extract A (R0292)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.
- 10 089 spectra from a total of 237 batches from further 106 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Yeast extract A (R0292)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Yeast extract A (R0292)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Yeast extract A (R0292)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	13 330
Type B	0	0	0	10 089
Type C	0	0	0	3

The substance/substance group *Yeast extract A (R0292)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9309 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9021 %)	n/a (n/a)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra ([Rule of Three](#) [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Yeast extract A (R0292)* in the substance class *Hepart raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Calcium lactate pentahydrate (R0013)	80.57	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Yeast extract A (R0292)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85868	85868	0.00	82.44
85869	85869	0.00	83.22

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.3-2021-01

Validated substance/substance group **Zinc citrate trihydrate (R0089)**
Substance class Hepart raw materials
Report date 22/03/2021
Report number 85178-2021-03-22
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Zinc citrate trihydrate (R0089)

Special notes

When selecting the *Zinc citrate trihydrate (R0089)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9 *European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40 *Erfüllung von Ph. Eur. 2.2.40 durch Apo-Ident* [4]
AA004 *Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also “database”) in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. In the *chemometric model* generated, the distances between all separable substances are checked for compliance with the specified safety distances.
3. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance	Type A	Type B	Type C
Zinc citrate trihydrate (R0089)	3	3	0

Second-stage model

For differentiation of the substance/substance group *Zinc citrate trihydrate (R0089)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the [chemometric models](#). The following samples have been obtained from the substance/substance group *Zinc citrate trihydrate (R0089)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
Hepart	Zinc citrate tri...	14000662	85178	40	not required
Hepart	Zinc citrate tri...	15000315	85267	60	not required
Hepart	Zinc citrate tri...	17000567	85486	40	not required

Validation samples

A total of 23 502 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 140 spectra of 3 reference samples from the substance/substance group *Zinc citrate trihydrate (R0089)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 13 270 spectra from a total of 270 batches from further 107 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix A](#). The samples in [Appendix A](#) were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a [†]. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 160 spectra of 3 reference samples from the substance/substance group *Zinc citrate trihydrate (R0089)*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
Hepart	Zinc citrate trihydrate (R00...	14001807	85233	60
Hepart	Zinc citrate trihydrate (R00...	15001427	85321	60
Hepart	Zinc citrate trihydrate (R00...	17001244	85498	40

- 9929 spectra from a total of 234 batches from further 105 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix B](#).

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Zinc citrate trihydrate (R0089)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 3 spectra from 3 *Apo-Ident* customers from a total of 3 batches from a further 1 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this [chemometric model](#), the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in [Appendix C](#).

Validation results

The validation runs checked whether the substance/substance group *Zinc citrate trihydrate (R0089)* can clearly be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Zinc citrate trihydrate (R0089)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive/negative*) and the validation spectrum type (*A/B/C*).

	False positive	True positive	False negative	True negative
Type A	0	140	0	13 270
Type B	0	125	35	9929
Type C	0	0	0	3

The substance/substance group *Zinc citrate trihydrate (R0089)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate (specificity)* and the weighted *true positive rate (recognition rate)* are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9306 %)	100.0000 % (> 95.7143 %)
Type B	100.0000 % (> 99.9005 %)	78.1250 % (> 76.2500 %)
Type C	n/a (n/a)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.

Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Zinc citrate trihydrate (R0089)* in the substance class *Hepat raw materials*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
N-acetyl-L-cysteine (R0002)	79.11	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Zinc citrate trihydrate (R0089)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
85486	85486	0.00	111.79
85178	85178	0.00	83.78
85267	85267	0.00	86.71

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

Appendix A: Additional calibration samples (Type A)

Substance spectra also enter the generation of the [chemometric model](#) which cannot be identified with this model. This serves the purpose that substances which the model is not supposed to recognise are safely rejected. The spectra of these samples are entered into the validation as *Type A*. They were recorded by *HiperScan GmbH*.

The samples originate from 1 batches. From these, 40 spectra were recorded. The calibration spectra of substances which can be identified with the model are listed respectively in the section *calibration samples* for the individual substances and do not appear again elsewhere in this list.

Supplier	Substance	Batch	Spectra	Certificate
Hepart	Potato fiber (SANACEL [®] Pot 90) (R0232)	20000941	40	not required

Appendix B: Additional validation samples (*Type B*)

It is necessary that spectra also enter the validation which cannot be identified with this model. In this manner, it is verified that the model also rejects unknown substances. The spectra for these samples were recorded by *HiperScan GmbH*. They are allocated to the *Type B*. They also include the calibration spectra for other models.

The samples originate from 34 batches. From these 1410 spectra were recorded. The spectra recorded on independent samples of substances which can be identified with the model are listed respectively in the section *Type B* for the individual substances and do not appear again elsewhere in this list.

Supplier	Substance	Batch	Spectra	Certificate
Hepart	Anhydrous sodium selenite (R0187)	15000383	60	not required
Hepart	Apricot fruit powder (R0335)	19002075	40	not required
Hepart	Artemisia annua 30:1 extract (R0290)	19002110/0	40	not required
Hepart	beta-glucan (barley fibre) (R0333)	19001720	40	not required
Hepart	Caffeine (R0286)	17000387	30	not required
Hepart	Chlorophyllin 97% (R0221)	16001000	40	not required
Hepart	Chromium(III) chloride hexahydrate (R0020)	14000606	40	not required
Hepart	D-(+)-mannose (R0136)	061101	40	not required
Hepart	D-(+)-mannose (R0136)	16000524	40	not required
Hepart	Dipotassium phosphate, anhydrous (R0025)	15000146	60	not required
Hepart	Folic acid (R0027)	077301	40	not required
Hepart	Gamma oryzanol (R0194)	16000594	40	not required
Hepart	Grape polyphenols (R0084)	1209022	40	not required
Hepart	Hemp protein powder 50% (R0264)	18001289	40	not required
Hepart	Hemp protein powder 50% (R0264)	19001828	40	not required
Hepart	Hydroxypropylmethylcellulose (R0038)	131028505	40	not required
Hepart	Iron (III) sulfate heptahydrate (R0026)	14000045	40	not required
Hepart	Isomaltulose (R0236)	16000382	40	not required
Hepart	Lupine proteins (R0098)	19002142/0	40	not required
Hepart	Magnesium malate (R0314)	18000782	40	not required
Hepart	Magnesium oxide (R0057)	051701	40	not required
Hepart	Organic yeast powder (R0263)	18001329	40	not required
Hepart	PF1 Extract Carob (ACO-PAN) (R0336)	19002076/0	40	not required
Hepart	Potassium chloride (R0044)	15000147	60	not required
Hepart	Potassium chloride (R0044)	16001518	40	not required
Hepart	Potato flour (R0046)	20001156	40	not required
Hepart	Saccharin sodium (R0260)	16000830	40	not required
Hepart	Saffron flowers (R0315)	18000629	40	not required
Hepart	Silymarine 80% extract (R0077)	131121500	40	not required
Hepart	Silymarine 80% extract (R0077)	19002342/0	40	not required
Hepart	Skimmed milk powder (SMP SI medium-high heat) (R0234)	16000734	40	not required
Hepart	Sodium cyclamate (R0259)	16000829	40	not required

continued on the next page

continued from previous page

Supplier	Substance	Batch	Spectra	Certificate
Hepart	Sunflower protein (55% protein) (R0308)	19001397	40	not required
Hepart	Sweet whey powder S Instant (R0144)	16000686	40	not required

Appendix C: Additional validation samples (Type C)

Not required.

Appendix D: Requirements of validation

In order to ensure adherence to the safe scientific status, the individual methods for manufacturing and testing must be validated under certain circumstances (compare § 34 para. 1 no. 3, § 35 para. 1 no. 4 and para. 4 sentence 1 no. 2 b, para. 6 sentence 3 *ApoBetrO* [Pharmacies Rules and Regulations]). The *ApoBetrO* [Pharmacies Rules and Regulations] incorporates a legal definition in § 1 a para. 16 (quotation translated):

“Validation is the provision of documented proof which with a high degree of safety documents that, via a specific process or standard work process, a medicinal product is manufactured and tested, which is in accordance with previously determined quality features.”

Validation documentation can be used to prove that methods or devices which are not described in the Pharmacopoeia within the meaning of § 6 para. 1 sentence 3 *ApBetrO* [Pharmacies Rules and Regulations] achieve the same results as those in the Pharmacopoeia. On the other hand, with the requirements of the demanded validation it must be observed whether the respective testing method is already incorporated in the Pharmacopoeia.

NIR spectroscopy as a general testing method need not be validated in accordance with the express ruling in the *Ph. Eur. Section 1.1* [3], as it is already described in *Section 2.2.40* of the *Ph. Eur.* as an area of application for the identification of raw materials.

However, a special validation requirement exists for the reference database. This requirement is met with the existing document. Further requirements or rules as to how this proof must be furnished do not exist. It is required that the processes guarantee the same results as the methods and devices in the Pharmacopoeia [17].

Carrying out identity tests with *Apo-Ident* is therefore also possible if the NIR spectroscopy process is not required in the Pharmacopoeia monograph of the substance for identity testing. All NIR analyses with *Apo-Ident* prove several, often all molecule groups and are therefore comparable with a series of individual, targeted chemical proofs [4]. Therefore, the identity proof with *Apo-Ident* replaces the monograph test series (with two or more test combinations).

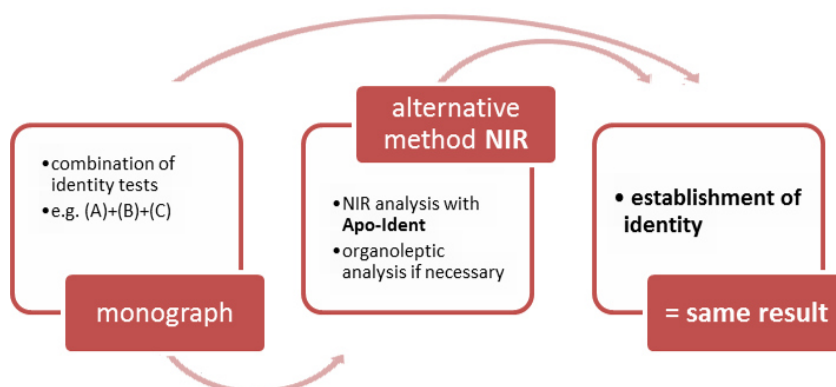


Figure 2: The combination of tests of the monograph is replaced by the alternative method NIR spectroscopy using *Apo-Ident*. This is permissible because both test procedures result in the establishment of the identity of the raw material.

With this validation documentation, proof is furnished that identical results are achieved with *Apo-Ident* and Pharmacopoeia methods, i.e. confirmation of the identity of the raw material [2].

Appendix E: Conformity of Apo-Ident with the European pharmacopeia

According to *Ph. Eur. Section 2.2.40*, NIR spectroscopy is basically suitable for: “Identification of agents, excipients, dosage forms, intermediate manufacturing products, chemical raw materials and packaging materials” ([3], quotation translated).

The fact that *Apo-Ident* meets the further criteria of the European Pharmacopoeia under the headings in *Section 2.2.40*

- Apparatus
- Measurement methods
- Sample preparation and presentation
- Testing the functionality of the instrument
- Identification and characterisation (qualitative analysis)
- Quantitative analysis
- Ongoing model evaluation
- Transfer of databases
- Data storage

can be proven based on the *HiperScan GmbH* documentation of “Meeting 2.2.40 *Ph. Eur.* by *Apo-Ident*” [4].

References

- [1] ABDA – BUNDESVEREINIGUNG DEUTSCHER APOTHEKERVERBÄNDE: Verordnung über den Betrieb von Apotheken (Apothekenbetriebsordnung – ApBetrO), 2012
- [2] REIMANN, B. ; REGIERUNGSPRÄSIDIUM DARMSTADT: Hinweise zur ordnungsgemäßen Prüfung von Arzneimitteln und Ausgangsstoffen (§§ 6 und 11 *ApBetrO*), 2007
- [3] *Europäisches Arzneibuch, Grundwerk 2020*. 10. Ausgabe. Deutscher Apotheker Verlag (978-3-7692-7517-9)
- [4] HIPERSCAN GMBH: Erfüllung von 2.2.40 Ph. Eur. durch Apo-Ident, 2013
- [5] ARBEITSGEMEINSCHAFT DER PHARMAZIERÄTE DEUTSCHLANDS (APD): Resolution in Verbindung mit §§ 6 und 11 *ApBetrO* Verwendung eines Nah-Infrarot-Spektrometers (NIR) zur Identitätsprüfung, 16. 10. 2013, DAZ 48, November 2013
- [6] ARBEITSGEMEINSCHAFT DER PHARMAZIERÄTE DEUTSCHLANDS (APD): Resolution 2014, Arbeitsgemeinschaft der Pharmazieräte Deutschlands (APD), Oktober 2014
- [7] *DAC/NRF*. Govi-Verlag (978-3-7741-0044-2)
- [8] KESSLER, W.: *Multivariate Datenanalyse*. WILEY-VCH Verlag, 2007 (978-3-527-31262-7)
- [9] NÆS, T. ; ISAKSSON, T. ; FEARN, T. ; DAVIES, T.: *Multivariate Calibration and Classification*. NIR Publications, 2002 (978 0 9528666 2 6)
- [10] HANLEY, J. A. ; LIPPMAN-HAND, A.: If nothing goes wrong, is everything all right? In: *Journal of the American Medical Association* 249 (1983), S. 1743–1745
- [11] JOVANOVIĆ, B. D. ; LEVY, P. S.: A Look at the Rule of Three. In: *The American Statistician* 51 (1997), S. 137–139
- [12] BRONSTEIN, I. N. ; SEMENDJAJEW, K. A. ; MUSIOL, G. ; MÜHLIG, H.: *Taschenbuch der Mathematik*. 5. überarbeitete und erweiterte Auflage. Verlag Harri Deutsch, 2000 (3-8171-2015-2)
- [13] MAHALANOBIS, P.: On the generalized distance in statistics. In: *Proc. Nat. Inst. Sci. India (Calcutta)* 2 (1936), S. 49–55
- [14] YAMBOR, B. ; DRAPER, W. ; BEVERIDGE, R.: Analyzing PCA-based face recognition algorithms: Eigenvector selection and distance measures. In: *Second Workshop Empirical Evaluation in Computer Vision* (2000)
- [15] HIPERSCAN GMBH: Identifikationsmethodik Apo-Ident, 2012
- [16] HIPERSCAN GMBH: Datenvorbehandlung des Identifikationssystems Apo-Ident, 2012
- [17] CYRAN ; RÖTTA: Apothekenbetriebsordnung, Kommentar § 6 Rn. 10, 2010

Index

- a-lipoic acid (R0003), [25](#)
Acacia gum (encapsia) / acacia gum (R001/R0141), [21](#)
Apple powder (R0005), [29](#)
Artichoke extract (R0208), [33](#)
Ascorbic acid (R0007), [39](#)

Beta carotene 10% (R0237), [43](#)
Beta carotene 20% (R0008), [47](#)
beta-glucan 28% (oat fibre) (R0334), [51](#)
Betaine HCl (R0211), [55](#)
Biotin (R0010), [59](#)
Black cumin extract (R0193), [63](#)

Calcium caseinate (Lactonate RCE) (R0239), [67](#)
Calcium D panthothenate (R0012), [71](#)
Calcium lactate pentahydrate (R0013), [75](#)
Calmegh extract (R0201), [79](#)
Chaga extract (R0016), [83](#)
Chili extract (R0198), [87](#)
Cholecalciferol (vitamin D3) (R0018), *see* Cholecalciferol (vitamin D3) (R0018) / Menaquinone-7 0.2% (R0134), [91](#)
Cholecalciferol (vitamin D3) (R0018) / Menaquinone-7 0.2% (R0134), [91](#)
Choline chloride (R0019), [97](#)
Citric acid monohydrate (R0143), [103](#)
Citric acid powder (R0258), [107](#)
Citrus bioflavonoids 60% (R0021), [111](#)
Coenzyme Q10 (R0022), [115](#)
Coriolus extract (R0207), [119](#)
Creatine monohydrate (R0048), [123](#)
Curcumin 95% extract (R0023), [127](#)
Cyanocobalamin (R0024), *see* Cyanocobalamin (R0024) / Methylcobalamin (R0190), [131](#)
Cyanocobalamin (R0024) / Methylcobalamin (R0190), [131](#)

D-alpha tocopheryl acetate 50% dry CWS (R0081), [135](#)
D-galactose (R0135), [139](#)
D-ribose (R0137), [147](#)
DL-methionine (R0209), [143](#)

Galactomannans (R0035/R0094), *see* Galactomannans (R0035/R0094) / Potato starch, pregelatinized (R0047), [153](#)
Galactomannans (R0035/R0094) / Potato starch, pregelatinized (R0047), [153](#)
Garcinia Cambogia (R0028), [159](#)
Gentian extract (R0200), *see* Marshmallow root extract (R0204) / Gentian extract (R0200) / Juniper extract (R0192), [319](#)
Ginger extract (R0199), [165](#)
Ginkgo extract (R0093), [169](#)
Ginseng extract (R0029), *see* Ginseng extract (R0029) / Lutein 20% (R0052) / Lycopene 5% extract (R0053), [173](#)
Ginseng extract (R0029) / Lutein 20% (R0052) / Lycopene 5% extract (R0053), [173](#)
Glucosamine sulfate 2 KCl (R0030), [179](#)
Glucosamine sulphate 2 KCl (R0030), *see* Glucosamine sulfate 2 KCl (R0030), [179](#)
Glutathione (reduced form) (R0032), [183](#)
Glycine (R0033), [187](#)
Green tea extract (R0034), [191](#)
Guarana extract (R0289), [195](#)

Hydrangea extract (R0203), [199](#)
Hyssop extract (R0205), [203](#)

Inositol (myo-) (R0039), [207](#)
Inositol hexanicotinate (R0040), [213](#)
Instant yeast (dry yeast) (R0112), [219](#)
Inulin (R0041), [223](#)
Isoflavones 10% (R0042), [229](#)

Juniper extract (R0192), *see* Marshmallow root extract (R0204) / Gentian extract (R0200) / Juniper extract (R0192), [319](#)

Katuka herb extract (R0210), [233](#)
Konjac glucomannan 80% (R0293), [237](#)

L-5-hydroxytryptophan (R0050), [241](#)
L-arginine base (R0006), [245](#)
L-carnitine-L-tartrate (R0014), [249](#)
L-carnosine (R0015), [253](#)
L-glutamine (R0031), [257](#)
L-histidine base (R0037), [261](#)
L-isoleucine (R0043), [265](#)
L-leucine (R0051), [271](#)
L-lysine HCl (R0054), [277](#)
L-ornithine HCl (R0066), [281](#)
L-phenylalanine (R0067), [285](#)
L-proline (R0070), [289](#)
L-threonine (R0080), [293](#)
L-tryptophan (R0085), [297](#)
L-tyrosine (R0086), [303](#)
Lutein 20% (R0052), *see* Ginseng extract (R0029) / Lutein 20% (R0052) / Lycopene 5% extract (R0053), [173](#)

Lycopene 5% extract (R0053), *see* Ginseng extract (R0029) / Lutein 20% (R0052) / Lycopene 5% extract (R0053), [173](#)

Magnesium citrate nonahydrate (R0056), [307](#)
Magnesium glycerophosphate (R0114), [311](#)
Magnesium stearate (R0140), [315](#)
Marshmallow root extract (R0204), *see*
Marshmallow root extract (R0204) / Gentian extract (R0200) / Juniper extract (R0192), [319](#)
Marshmallow root extract (R0204) / Gentian extract (R0200) / Juniper extract (R0192), [319](#)

Menaquinone-7 0.2% (R0134), *see*
Cholecalciferol (vitamin D3) (R0018) / Menaquinone-7 0.2% (R0134), [91](#)

Methionine (R0059), [323](#)
Methylcobalamin (R0190), *see*
Cyanocobalamin (R0024) / Methylcobalamin (R0190), [131](#)
Methylsulfonylmethane (R0060), [329](#)

N-acetyl-L-cysteine (R0002), [333](#)
Nicotinamide (R0064), [337](#)

Olive leaf extract (R0202), [343](#)
OPC (grape seed) extract (R0065), [347](#)
Oregano extract (R0206), [351](#)
Osha root extract (R0220), [355](#)

p-aminobenzoic acid (R0004), [359](#)
Pea protein (R0092), [363](#)
Piperine 95% extract (R0069), [367](#)
Potato starch, pregelatinized (R0047), *see*
Galactomannans (R0035/R0094) / Potato starch, pregelatinized (R0047), [153](#)

Progesterone, micronized (R0108), [371](#)
Pumpkin seed protein (R0306), [375](#)
Pyridoxal-5-phosphate (R0195), [379](#)
Pyridoxine hydrochloride (vitamin B6) (R0071), [383](#)

Resveratrol (R0339), [387](#)
Retinyl palmitate (vitamin A palmitate) (R0073), [391](#)
Rhodiola rosea extract 3% (R0159), [395](#)
Riboflavin (vit. B2) (R0074), [399](#)

Sodium borate (R0191), [403](#)
Sodium chondroitin sulfate (R0061), [409](#)
Sodium chondroitin sulphate (R0061), *see*
Sodium chondroitin sulfate (R0061), [409](#)
Sunflower oil powder 25 (R0319), [413](#)

Taurine (R0078), [417](#)
Thiamine HCl (Vitamin B1) (R0079), [421](#)
Thiamine mononitrate (R0212), [425](#)
Tripotassium citrate monohydrate (R0117), [429](#)

Valine (R0087), [433](#)
Vanilla flavoring (natural flavor) (R0240), [437](#)

Walnut leaf extract (R0197), [441](#)

Yeast extract A (R0292), [445](#)

Zinc citrate trihydrate (R0089), [449](#)