ANEXO III

NIR libraries in the pharmaceutical industry: a solution for identity confirmation

M. Blanco y M. A. Romero

Analyst 126, 2001, 2212-2217.

FULL PAPER
THE ANALYST
www.rsc.org/analyst

Near-infrared libraries in the pharmaceutical industry: a solution for identity confirmation

M. Blanco* and M. A. Romero

Departament de Química, Unitat de Química Analítica, Facultat de Ciències, Universitat Autónoma de Barcelona E-08193. Bellaterra, Spain

Received 6th June 2001, Accepted 7th September 2001
First published as an Advance Article on the web 14th November 2001

The construction of near-infrared spectral libraries as an alternative to qualitative analysis methods for identifying pharmaceutical raw materials is proposed. Various conceptual and practical aspects of library construction are assessed and discussed. The procedure is demonstrated by constructing a library including NIR spectra for 125 different raw materials using the correlation coefficient as the discriminating criterion. Compounds with very similar spectra can be identified by constructing sub-cascading libraries branching off the main one that are developed by using chemometric procedures with higher discriminating ability. The construction of sub-libraries and their performance and discriminating power in three different situations are illustrated. The proposed methodology affords the expeditious unequivocal identification of all the compounds included in a library.

Introduction

The identification of active pharmaceutical ingredients (API), raw materials used to manufacture products and the end products themselves are among the routine tests used to control pharmaceutical manufacturing products. In this context, identification encompasses all those processes conducted with a view to verifying the identity of a given substance. The need for this identification tests is acknowledged in various ICH guidelines for pharmaceuticals.[1–3] Although compliance with their contents is not mandatory, many pharmaceutical manufacturers adhere to them in response to either the recommendations of the local authorities or market demands.

A vast number of methods involving a wide variety of techniques are available for identifying pharmaceuticals. Pharmacopoeias[4–6] have compiled a wide range of analytical methods for the identification of pharmaceutical active principles. Usually, several tests per product are recommended that involve either classical (precipitation, colorimetric or group-specific reactions) or instrumental analytical methods (IR spectroscopy, optical rotation, melting-point, HPLC).

Occasionally, providing an objective identification result with these methods is extremely difficult as it relies heavily on the analyst's knowledge and expertise. This usually leads to more than one test being performed in order to confirm the identity of the substance concerned, which lengthens the identification process. Subjectiveness in analytical results tends to be avoided by using objective, conclusive identification methods such as HPLC-MS, GC-MS or enzymatic type, for example. However, the equipment required is expensive to purchase and maintain and the ensuing methods are labour intensive and time consuming and use reagents (solvents).

Near-infrared spectroscopy (NIRS) is one of the instrumental techniques with the brightest prospects in pharmaceutical analysis as it surpasses other instrumental identification alternatives. In its diffuse reflectance variant (NIRRS), it even dispenses with the need for prior dissolution or dilution of the analyte as it affords direct recording of spectra from solid products. This results in substantial resource and time savings and hence in increased productivity. In addition, the chemical information needed can be obtained without losing any physical information (*e.g.*, particle size, density, hardness), which may be very useful and interesting. The construction of NIR libraries containing the spectra for all the compounds potentially handled by a pharmaceutical manufacturer avoids the need to develop a specific method for each and allows the compounds to be identified by using a single, straightforward, expeditious, inexpensive method. In addition, the whole process can be automated and computer controlled, so identifications call for no skilled personnel. The use of the NIR technique for qualitative analysis is illustrated in the European Pharmacopoeia,[4] where it is deemed 'a technique particularly useful for identifying organic substances'; this pharmacopoeia even provides some suggestions on how to construct NIR libraries for pharmaceuticals.

Although NIRS has been used to identify and classify a variety of substances, in most instances it was employed to discriminate among a small number of them (*e.g.*, similar compounds or members of the same family).[7–9] There are, in fact, few references to general identification libraries[10] containing many compounds or describing aspects such as the characteristics of the libraries, the methods used to construct them, the requirements to be met or their validation. This has been a recurrent subject of debate by Pharmeuropa,[11–13] which has addressed general aspects of the identification process but issued no practical recommendations.

This paper provides some general recommendations with a view to identifying pharmaceuticals using NIR libraries and discusses their development and identification potential, as well as the quality of their results. The proposed methodology is illustrated by developing a library and examining various situations and procedures.

Background

NIR bands are broad and strongly overlapped, which makes substance identifications by direct comparison of their spectra with a standard spectrum, the procedure usually employed in mid-IR spectroscopy, virtually impossible. NIR identifications rely on the use of pattern recognition methods (PRMs), a wide variety of which exist that have found application in specific fields. Below are briefly described the most usual PRMs and one possible classification (Fig. 1).

 DOI: 10.1039/b105012p

This journal is © The Royal Society of Chemistry 2001

Pattern recognition methods

The wide variety of PRMs available is frequently expanded with new choices. Most PRMs rely on similarity measurements. Similarity here is taken to be the extent to which an object (spectrum) is identical with another. Most often, similarity is expressed in terms of correlation[14] or distance.[15]

PRMs can be of two types depending on whether the objects are known to belong to specific classes that are called supervised and unsupervised methods.

Unsupervised methods search for clustering in an *N*-dimensional space without knowing the class to which the sample belongs. Cluster analysis,[16] the minimal spanning tree (MST)[17] and unsupervised (or Kohonen) neural networks[18] are among the most common unsupervised PRMs.

Supervised methods rely on the prior training of the system, using a set of objects belonging to specific, previously known classes. These methods can be of the discriminant or the modelling type.[19] Discriminant methods split the space pattern into as many regions as classes are included in the training set, thereby creating bounds that are shared by the spaces. The most commonly used among them are discriminant analysis (DA),[20] the *k*-nearest neighbour (KNN) function[21,22] and potential function methods (PFMs).[23,24] Modelling methods create volumes in the pattern space that possess different bounds for each class. Such bounds can be established in the form of correlation coefficients, distances (whether Euclidean, as in the PRIMA method,[25] or of the Mahalanobis type, as in the UNEQ method[26]), the residual variance[27,28] or supervised artificial neural networks such as the multi-layer perceptron (MLP).[29]

Not all pattern recognition methods are suitable for constructing product identification libraries. A purpose such as that addressed in this work requires the use of supervised modelling methods.

Library construction procedure

This section describes the steps involved in developing a library for the identification of pharmaceuticals using NIRS, and the characteristics and identification potential of the library.

Instrument qualification

Assessing whether the instrument operates as scheduled is the first step in developing any instrumental methodology. The European Pharmacopoeia[4] recommends following the manufacturer's instructions for this purpose, checking for wavelength scale, wavelength repeatability, response repeatability and photometric noise.

Approaches to library construction

The identification library should contain all the raw materials used by the pharmaceutical manufacturer concerned in order to be able to identify all possible substances and avoid or reduce errors. In addition, it should be able to distinguish between very similar compounds used in different applications (*e.g.*, products in different particle sizes, product polymorphs, different product grades or suppliers, *etc.*).

The correlation coefficient is especially suitable for constructing the general library as it has the advantage that it is independent of its size, uses only a few spectra to define each product and is scarcely sensitive to slight instrumental oscillations. This parameter allows the library to be developed and validated more expeditiously than others and also to be expanded with new products or additional spectra for an existing product in order to incorporate new variability sources, also in a rapid manner.

One of the crucial factors with a view to ensuring adequate selectivity in constructing a spectral library is the choice of an appropriate threshold, which is the lowest value (for correlation) required to assign unequivocally a given spectrum to a specific class. Too low a threshold can lead to confusion between substances with similar spectra. By contrast, too high a threshold can result in spectra belonging to the same class being incorrectly classified. Choosing an appropriate threshold entails examining the spectra included in the library in an iterative manner: the threshold is successively changed until that resulting in the smallest number of identification errors and confusion are achieved. In some cases, the threshold thus selected may not allow one to distinguish some compounds if

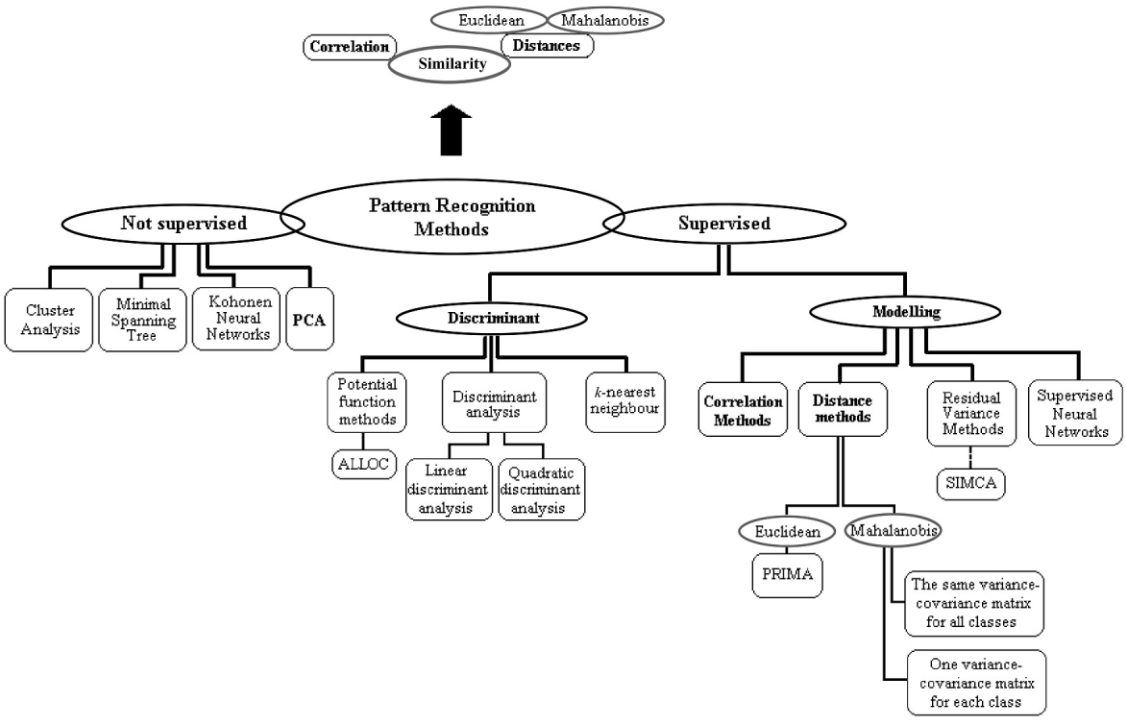

Fig. 1 Pattern recognition methods classification.

their spectra are too similar. This problem can be overcome by identifying the compounds concerned in two steps, by using a general library and a second, smaller library (a subset of the general one constructed with a higher discriminating PRM such as residual variance or Mahalanobis distance). This methodology can be labelled "cascading identification" as it involves identifying the unknown sample against the general library and, if the result is inconclusive, using a sub-library for qualification.

Library construction

The procedure to be followed in constructing the library involves five steps, namely: 1. Recording the NIR spectra by using a set of samples of known identity. 2. Choosing spectra. For each substance, the spectra used to construct the calibration set should belong to various batches so that physico-chemical variability can be effectively incorporated. 3. Constructing the library. First, one must choose the pattern recognition method to be used (*viz*., correlation or wavelength distance). Then, one must choose construction parameters such as the spectral pre-treatment (SNV, derivatives), wavelength range and threshold to be used. The next step involves internal validation of the library in order to check for incorrectly labelled spectra, spectra yielding identification errors or unidentified (or ambiguously identified) substances. Based on the validation results, whether some change in the threshold, spectral range, pre-treatment, *etc*., must be introduced is decided upon iteratively until obtaining the desired specificity level. 4. Constructing sub-cascading libraries. Each sub-library should include all those mutually related substances that result in ambiguous identification with the general library. The number of spectra used should exceed that of spectra required to define an individual class in the general library. The construction procedure is similar to that followed to construct the general library: select the PRM (Mahalanobis distance or residual variance), the characteristics of the library (which now include the number of latent variables or the explained variance, since these methods involves previous PCA) and perform the calculations, varying the parameters until all compounds can be unequivocally distinguished. 5. External validation. The general library and its sub-libraries must be validated by checking that external spectra (validation set) are correctly, unambiguously identified. Likewise, samples not present in the library should not be identified with any of the compounds included in it.

Experimental

Samples

A total of 125 solid substances, all used as raw materials for manufacturing pharmaceuticals, were studied. At least three batches per substance were used to record triplicate spectra; in this way, each class was defined on the basis of at least nine spectra.

Recording of NIR spectra

A total of 3000 NIR spectra were recorded, using an NIRSystems 5000 spectrophotometer from Foss NIRSystems (Silver Spring, MD, USA) equipped with a reflectance detector and an AP6641ANO4P fibre-optic probe. The instrument was governed via a PC, using the software Vision 2.22, also from NIRSystems, both to acquire data and to process and validate the identification library. Each spectrum was the average of 32 scans performed at 2 nm intervals over the wavelength range 1100–2500 nm.

Results and discussion

We constructed a general library for the identification of pharmaceutical raw materials using correlation as the discriminating criterion and three sub-libraries for that of compounds that could not be unambiguously identified with the general library.

General library

Preliminary tests revealed the optimum conditions for constructing the library to be the use of second-derivative spectra and a threshold of 0.97, which avoided confusion between compounds with similar spectra without altering the results for other substances. These conditions allowed the correct identification of virtually every compound included in the library and the rejection of those not pertaining to it. Some compounds, among which we chose three to illustrate the case, were ambiguously identified as they exhibited correlation coefficients above the threshold for more than one class. Accurate identification of these compounds required the development of three sub-cascading libraries.

Sub-cascading libraries

The cases discussed here illustrate typical problems encountered by the pharmaceutical industry in using a general library and also the effectiveness of the proposed methodology, based on the use of sub-cascading libraries, for this purpose.

We used two different discriminating criteria (*viz*., the Mahalanobis distance and the residual variance), which proved similarly effective. Fig. 2 summarizes the features of each sub-library. Unless the general library, the thresholds for the sub-libraries were upper bounds (*i.e*., identification was positive if the result was below the threshold).

Particle size sub-library. We addressed the distinction of five types of sucrose with as many different particle sizes, namely crystal sucrose, powdered sucrose A, semolina sucrose,

RAW MATERIALS LIBRARY		
AIM: correct identification of all the raw materials of a pharmaceutical manufacturer	LIBRARY PARAMETERS: Wavelength range: 1100-2200 nm Pretreatment: 2nd derivative spectra Method: Correlation coefficients Threshold: 0.97	
SUB-CASCADING-LIBRARIES		
PARTICLE SIZE	POLYMORPHISM	
	Diacetylmidecamycin	Ketoprofen
AIM: classification of five different particle size of sucrose	AIM: classification of amorphous samples of diacetylmidecamycin impurified with the crystalline form	AIM: classification of dextro and levoketoprofen
SUBLIBRARY PARAMETERS: 1100-2200 nm 1st derivative spectra Mahalanobis Distance in PC Space 4 PCs Threshold: 0.89	SUBLIBRARY PARAMETERS: 1100-2200 nm 2nd derivative spectra Residual Variance 95 % Explained Variance Threshold: 0.88	SUBLIBRARY PARAMETERS: 1100-2200 nm 2nd derivative spectra Residual Variance 95 % Explained Variance Threshold: 0.87

Fig. 2 Aim and parameters of the general library and sub-cascading-libraries.

powdered sucrose B and granulated sucrose. Their spectra are shown in Fig. 3. Table 1 gives the correlation coefficient for each sucrose type when identified in its own and in the other four types. In many cases, the coefficient exceeded 0.97 (the threshold for the general library), so it would have led to ambiguous identification with the general library. This shortcoming was circumvented by developing the sucrose classification sub-library, the main characteristics of which are shown in Fig. 2. The sub-library was constructed using the Mahalanobis distance; the residual variance, however, provided similar results.

Table 2 shows the results obtained by using this sub-library to identify the different types of sucrose. The values for the identification of each type of sucrose with the other classes were higher than the threshold, so the identification was negative and all five types of sucrose, which differed in particle size, were correctly identified in all the samples studied.

Polymorphism sub-libraries. Two different sub-libraries were constructed, the diacetylmidecamycin sub-library and ketoprofen sub-library, which allowed us to solve the ambiguities found in the general library.

Diacetylmidecamycin. Diacetylmidecamycin was studied in two different forms, amorphous and crystalline. Distinguishing them *via* the correlation coefficient was fairly easy as their spectra are markedly different (Fig. 4). However, samples of the amorphous form "impurified" with the crystalline form were assigned to the amorphous class, even in the presence of as much as 10% of the crystalline form, with a correlation coefficient higher than 0.97. A maximum "impurity" content of 2% was established to distinguish amorphous diacetylmidecamycin.

A sub-library was developed to determine whether a sample of amorphous diacetylmidecamycin was acceptable (see Fig. 2). An 'impurified' sample was identified as amorphous diacetylmidecamycin by the general library and as adequate in quality by the sub-library. Table 3 shows the results of the identification of amorphous diacetylmidecamycin samples containing various proportions of the crystalline form. As can be seen, all samples containing more than 2% of such a form exceeded the established threshold, so they were classified as unacceptable.

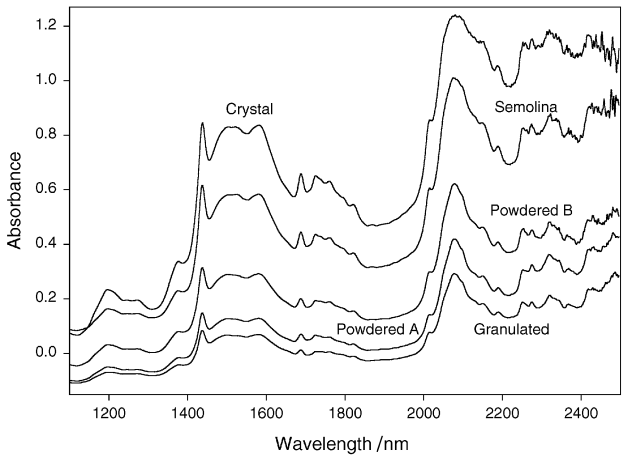

Fig. 3 Spectra of five different types of sucrose. Each spectrum is the average of five spectra of different batches from each type of sucrose.

Ketoprofen. It is known that chiral substances can occur in different crystal structures (or polymorphs) whether they are racemic or pure enantiomers.[30] Pure enantiomers crystallize in a non-centrosymmetric space group while 90% of racemates do so in a centrosymmetric space group, so differing crystal structures can be obtained for the same chemical compound. Different structures yield differences in NIR spectra and these differences can be used for qualitative or quantitative analysis. Buchanan *et al*., for example, used the differences in NIR spectra to determine the enantiomeric purity of valine.[31]

Ketoprofen is an anti-inflammatory agent, which has a chiral centre. The most often used in the pharmaceutical industry are the dextrorotatory and the racemic forms. Fig. 5 shows the NIR spectra for both of those and the levorotatory form. As can be seen, the spectra for the two enantiomers differed slightly between them but markedly from that for the racemate, probably owing to differences in crystal structure. The differences were large enough to allow the general library to distinguish between the pure enantiomers and their racemate, since they had a correlation coefficient of 0.92.

Although NIR reflectance spectroscopy cannot differentiate enantiomers, the dextro and levo forms of ketoprofen show slightly different NIR spectra, which cannot be attributed to impurities or instrumental noise, probably owing to slightly different packing in the crystal structure. A sub-library constructed using the residual variance as discriminating criterion (Fig. 2) afforded the discrimination between the two forms. Thus, a ketoprofen sample was identified as either racemate or pure enantiomer by the general library and the latter form as dextro- or levorotatory by the sub-library. Table 4 shows the results obtained in the identification of samples of both enantiomers. All samples were correctly identified.

Once this sub-library had been checked to distinguish both pure enantiomers effectively, it was refined to allow the detection of small amounts of one enantiomer in the other, similarly as with diacetylmidecamycin. The eutomer, dextroketoprofen, was used to prepare samples containing 2 or 5% of the levorotatory form (the acceptable contamination limit was 2%). The process by which a dextroketoprofen sample impurified with levoketoprofen is obtained gives rise to an equivalent amount of racemate during crystallization, the

Table 1 Typical correlation values obtained in crossed identification of different types of sucrose. The threshold for acceptance is 0.97

	Identified as					
Sample	Crystal	Powdered A	Semolina	Powdered B	Granulated	Comments
Crystal	**0.9987**	0.9287	0.9767	0.9178	0.9504	Ambiguous with semolina
Powdered A	0.9212	**0.9999**	0.9840	0.9995	0.9972	Ambiguous with semolina, powdered B and granulated
Semolina	0.9727	0.9842	**0.9995**	0.9790	0.9935	Ambiguous with crystal, powdered A, powdered B and granulated
Powdered B	0.9078	0.9986	0.9780	**0.9996**	0.9944	Ambiguous with powdered A, semolina and granulated
Granulated	0.9459	0.9968	0.9937	0.9948	**0.9997**	Ambiguous with powdered A, semolina and powdered B

Table 2 Crossed identification values obtained for five sucrose granulations in the sucrose sub-library. The threshold is fixed at 0.87

Sucrose sample	Selected as	Identification result	Identification value
Crystal	**Crystal**	**Positive**	**0.693**
	Powdered A	Negative	1.000
	Powdered B	Negative	1.000
	Granulated	Negative	1.000
	Semolina	Negative	1.000
Powdered A	**Powdered A**	**Positive**	**0.100**
	Powdered B	Negative	0.999
	Crystal	Negative	1.000
	Granulated	Negative	1.000
	Semolina	Negative	1.000
Granulated	**Granulated**	**Positive**	**0.550**
	Powdered A	Negative	1.000
	Powdered B	Negative	1.000
	Crystal	Negative	1.000
	Semolina	Negative	1.000
Powdered B	**Powdered B**	**Positive**	**0.530**
	Powdered A	Negative	1.000
	Crystal	Negative	1.000
	Semolina	Negative	1.000
	Granulated	Negative	1.000
Semolina	**Semolina**	**Positive**	**0.127**
	Powdered A	Negative	1.000
	Powdered B	Negative	1.000
	Granulated	Negative	1.000
	Crystal	Negative	0.999

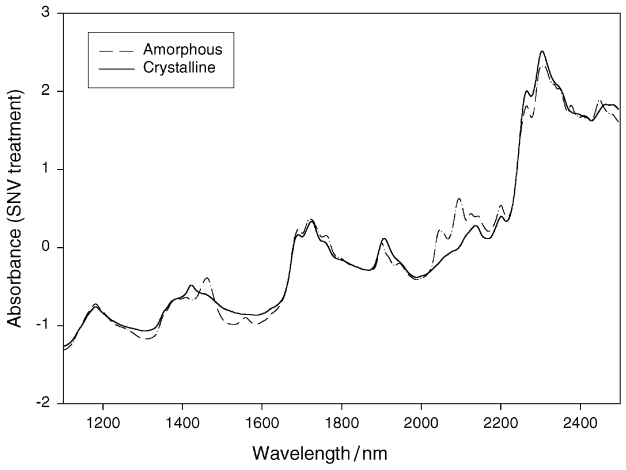

Fig. 4 Spectra of amorphous and crystalline diacetylmidecamycin. Each spectrum is the average of five spectra from different batches of each type of diacetylmidecamycin.

Table 3 Samples consisting of mixtures of amorphous and crystalline diacetylmidecamycin. The maximum value of crystalline diacetylmidecamycin accepted is 2%. The threshold is fixed at 0.88. Positive identifications (ID Value <0.88) are in bold face

Sample	Selected as	Identification result	ID value
Amorphous + 0.1% crystalline	Amorphous	**Positive**	**0.608**
Amorphous + 0.3% crystalline	Amorphous	**Positive**	**0.724**
Amorphous + 0.6% crystalline	Amorphous	**Positive**	**0.686**
Amorphous + 1.5% crystalline	Amorphous	**Positive**	**0.789**
Amorphous + 2.0% crystalline	Amorphous	**Positive**	**0.796**
Amorphous + 3.0% crystalline	Amorphous	Negative	0.932
Amorphous + 4.0% crystalline	Amorphous	Negative	0.975
Amorphous + 6.0% crystalline	Amorphous	Negative	1.000
Amorphous + 8.0% crystalline	Amorphous	Negative	1.000
Amorphous + 10.0% crystalline	Amorphous	Negative	1.000

spectrum for the latter being clearly different from those for the pure enantiomers. In order to confirm this assumption, samples containing 2 or 5% levoketoprofen were recrystallized in methanol and their spectra recorded. A sample of pure dextroketoprofen was also recrystallized using the same procedure. The spectrum for the recrystallized pure dextroketoprofen was found to be identical with that for the original dextroketoprofen; this was not the case for the impurified samples, the NIR spectra for which differ from those of the enantiomers in zones where the racemate spectrum is also different.

The spectra for several samples containing 2% levoketoprofen were included in the calibration set for dextroketoprofen in the sub-library. The parameters for this new library were slightly different (second-derivative spectra, 95% of variance accounted for and a threshold of 0.85). This new sub-library allowed the discrimination of dextroketoprofen and levoketoprofen and even the rejection of samples containing more than 2% of the latter. As can be seen from Table 5, the samples containing 2% of levoketoprofen were all correctly identified, whereas those containing 5% were not (their identification result exceeded the threshold). A similar procedure could be

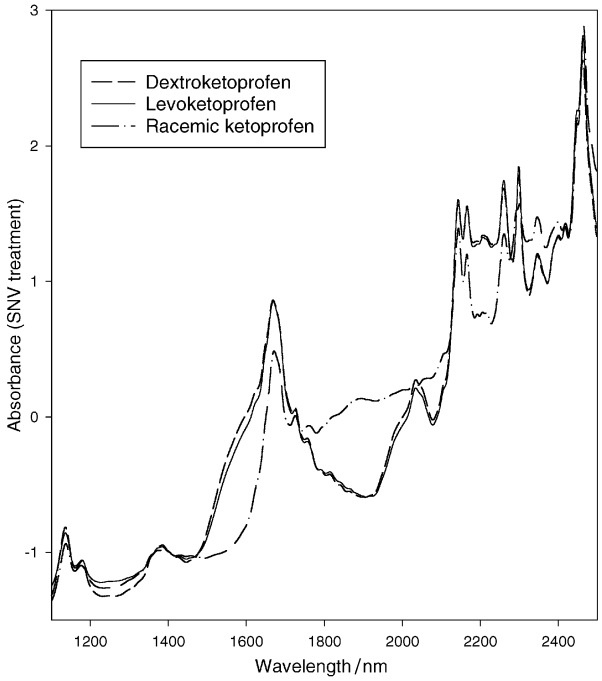

Fig. 5 SNV spectra of different ketoprofen forms. Each spectrum is the average of three spectra from different batches of each type of ketoprofen.

Table 4 Samples of dextroketoprofen and levoketoprofen identified in the ketoprofen sub-library. The threshold is fixed at 0.87. Positive identifications (ID value <0.87) are in bold face

Sample	Identified as dextro form	Identified as levo form
Dextroketoprofen 1	**0.498**	1.000
Dextroketoprofen 2	**0.216**	1.000
Dextroketoprofen 3	**0.446**	0.999
Dextroketoprofen 4	**0.510**	0.999
Dextroketoprofen 5	**0.361**	0.999
Levoketoprofen 1	1.000	**0.611**
Levoketoprofen 2	1.000	**0.474**
Levoketoprofen 3	1.000	**0.326**
Levoketoprofen 4	1.000	**0.333**
Levoketoprofen 5	1.000	**0.425**

Table 5 Samples of dextroketoprofen containing 2 and 5% of levoketoprofen identified in the ketoprofen sub-library. The threshold is fixed at 0.85. Positive identifications (ID value < 0.85) are in bold face

Sample	Selected as	Identification result	ID value
Dex + 2% Lev	Dextroketoprofen	**Positive**	**0.813**
Dex + 2% Lev	Dextroketoprofen	**Positive**	**0.756**
Dex + 5% Lev	Dextroketoprofen	Negative	0.885
Dex + 5% Lev	Dextroketoprofen	Negative	0.874
Dex + 5% Lev	Dextroketoprofen	Negative	0.896

used to identify levoketoprofen impurified with the dextrorotatory form.

Conclusions

This paper describes the most salient conceptual and practical aspects of the construction of NIR libraries for identifying pharmaceutical raw materials. This method of qualitative analysis substantially simplifies the identification of such materials as it integrates all identifications in a single process. NIR libraries provide an expeditious, non-invasive identification method: measurements require no sample treatment; also, they yield a numerical value, which makes identifications more objective and hence also more reliable. The use of a single library containing all the raw materials used by a manufacturer and constructed on the basis of correlation coefficients is highly effective as it provides immediate results and allows the number of compounds in the library to be readily expanded without increasing analysis times. The proposed method using subcascading libraries and more strongly discriminating techniques, based on the reduction of variables for those substances that cannot be distinguished in terms of their correlation coefficients, provides the desired specificity and allows one to discriminate between products that would otherwise require more complex techniques and procedures. NIR libraries are thus highly useful tools for the qualitative analysis of pharmaceutical products.

Acknowledgement

The authors are grateful to Spain's DGICyT for funding this research within the framework of Project BQU2000-0234. M. A. Romero acknowledges additional funding from Spain's Ministry of Education and Culture in the form of a researcher training grant. Finally, the authors wish to thank Laboratorios Menarini (Badalona, Spain) for supplying the samples.

References

1 *Good Manufacturing Practice Guide for Active Pharmaceutical Ingredients*, Draft Consensus Guideline, International Conference on Harmonisation (ICH), 2000.
2 *ICH Q6 A. Specifications: Test Procedures and Acceptance Criteria for New Drug Substances and New Drug Products: Chemical Substances*, Consensus Guideline, International Conference on Harmonisation (ICH), 1996.
3 *ICH Q2B: Validation of Analytical Procedures: Methodology*, Consensus Guideline, International Conference on Harmonisation (ICH), 1998.
4 *European Pharmacopoeia*, Council of Europe, Strasbourg, 3rd edn., 1997.
5 *British Pharmacopoeia 1993*, H. M. Stationery Office, London, 1993.
6 *United States Pharmacopoeia, USP-24-NF 19*, United States Pharmacopeial Convention, Rockville, MD, 2000.
7 K. B. Bradfield and R. A. Forbes, *J. Near Infrared Spectrosc*., 1997, **5**, 41.
8 C. Van der Vlies, K. J. Kaffka and W. Plugge, *Pharm. Tech. Eur*., 1996, **4**, 43.
9 E. Dreassi, G. Ceramelli, P. Corti, P. L. Perruccio and S. Lonardi, *Analyst*, 1996, **121**, 219.
10 C. I. Geräusser and K. A. Kovar, *Appl. Spectrosc*., 1997, **51**, 1504.
11 *Pharmeuropa*, Council of Europe, Strasbourg, 1998, vol. 10, p. 468.
12 *Pharmeuropa*, Council of Europe, Strasbourg, 1999, vol. 11, p. 275.
13 *Pharmeuropa*, Council of Europe, Strasbourg, 1999, vol. 11, p. 444.
14 M. Blanco, J. Coello, H. Iturriaga, S. Maspoch and C. de la Pezuela, *Anal. Chim. Acta*, 1994, **298**, 183.
15 H. Mark, in *Handbook of Near Infrared Analysis*, ed. D. A. Burnsand E. Ciurczac, Marcel Dekker, NewYork, 1992, ch. 13.
16 N. Bratchell, *Chemom. Intell. Lab. Syst*., 1989, **6**, 105.
17 O. Strouf, *Chemical Pattern Recognition*, Research Studies Press, Letchworth, 1986.
18 P. Cáceres-Alonso and A. García-Tejedor, *J. Near Infrared Spectrosc*., 1995, **3**, 97.
19 M. P. Derde and D. L. Massart, *Anal. Chim. Acta*, 1986, **191**, 1.
20 P. A. Lachenbruch, *Discriminant Analysis*, Hafner Press, New York, 1975.
21 D. Coomans and D. L. Massart, *Anal. Chim. Acta*, 1982, **136**, 15.
22 D. Coomans and D. L. Massart, *Anal. Chim. Acta*, 1982, **138**, 153.
23 D. Coomans and I. Broeckaert, *Potential Pattern Recognition*, Wiley, New York, 1986.
24 D. Coomans, M. P. Derde, D. L. Massart and I. Broeckaert, *Anal. Chim. Acta*, 1981, **133**, 241.
25 I. Jurickskay and G. E. Veress, *Anal. Chim. Acta*, 1981, **171**, 61.
26 M. P. Derde and D. L. Massart, *Anal Chim. Acta*, 1986, **184**, 33.
27 *VISION User Manual*, Foss NIRSystems, Silver Spring, MD, 1998.
28 P. J. Gemperline and L. D. Webber, *Anal. Chem*., 1989, **61**, 138.
29 E. Bertran, M. Blanco, J. Coello, H. Iturriaga, S. Maspoch and I. Montoliu, *J. Near Infrared Spectrosc*., 2000, **8**, 45.
30 H. G. Britain, *Polymorphism in Pharmaceutical Solids*, Marcel Dekker, New York, 1999.
31 B. R. Buchanan, E. W. Ciurczak, A. Q. Grunke and D. E. Honigs, *Spectroscopy*, 1988, **3**, 54.