Journal of Pharmaceutical and Biomedical Analysis 49 (2009) 32–41

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

journal homepage: www.elsevier.com/locate/jpba

Quantifying low levels of polymorphic impurity in clopidogrel bisulphate by vibrational spectroscopy and chemometrics

Zoltán Német[a], Ádám Demeter[a,*], György Pokol[b]

[a] Drug Polymorphism Research Division, Gedeon Richter Plc, H-1475 Budapest, P.O. Box 27, Hungary
[b] Department of Inorganic and Analytical Chemistry, Budapest University of Technology and Economics, H-1111 Budapest Szt. Gellért tér 4., Hungary

ARTICLE INFO

Article history:
Received 7 July 2008
Received in revised form 8 September 2008
Accepted 25 September 2008
Available online 5 October 2008

Keywords:
IR
Raman
Spectroscopy
Quantitative analysis
Chemometrics
Clopidogrel bisulphate

ABSTRACT

Vibrational spectroscopic methods were developed for quantitative analysis of Form II of clopidogrel bisulphate in Form I and Form II polymorphic mixtures. Results show that both IR and Raman spectroscopy combined with chemometrics are suitable to quantify low levels of Form II in Form I, down to 2 and 3%, respectively, with less than 1% limit of detection. Different preprocessing and multivariate methods were applied for spectral processing and were compared to find the best chemometric model. Common problems of quantitative vibrational spectroscopy in the solid phase are discussed; and procedures appropriate to eliminate them are proposed.

1. Introduction

It has long been known that polymorphic forms of active pharmaceutical ingredients (APIs) may have different bioavailability [1]; the issue has, however, attracted particular attention in recent decades [2–4]. At the present time regulatory authorities require pharmaceutical companies to investigate and control polymorphism of drug substances to ensure product quality, safety and performance [5–7]. Manufacturers have to declare that their API and product does not suffer solid phase transformation within the shelf life, which could affect bioavailability. Stability relationships between different solid forms of the substance and storage conditions avoiding phase transitions have to be established. Gaining such information needs suitable solid state analytical methods to be able to differentiate polymorphic forms and solvates of the substance and often, methods of quantifying these solid forms. Finding a new patentable solid form can prolong the duration of originator's patent protection or provide an opportunity for generics to enter the market without infringement [8]. Since both the originator and generic companies try to protect pure polymorphic forms and define limits to protect polymorphic mixtures, each contestant needs to have reliable solid state analytical methods to prove the polymorphic purity or composition of their product.

Techniques for differentiating polymorphic forms and solvates of drug substances have been reviewed by several authors [3,4,9–11]. These techniques can also be used for quantification of the different forms in their mixtures. The potential and limitations of quantitative analysis are usually discussed in textbooks and there is also a comprehensive review about pharmaceutical applications [12].

It is well-known that vibrational spectroscopy such as infrared (IR) and Raman spectroscopy is able to differentiate solid forms of drug substances [10,13–15]. Textbook discussions are usually confined to liquid and gas phase analysis rather than solid state quantitation. Although principles of the measurement are the same and the instrumental arrangements are sometimes identical, sample preparation is rather different and has a more fundamental effect on quantitation. A comprehensive review was published recently on the quantitative aspects of Raman spectroscopy [16]. The classical univariate calibration, i.e. the use of areas or heights of spectral bands unique for the components according to the Lambert–Beer equation, has long been used for quantification. There are numerous IR [17–22] and Raman spectroscopic [23–25] studies of pharmaceutical applications, where satisfactorily accurate and precise quantification were achieved with univariate calibration. However, if the molecular conformation or different hydrogen bonding pattern is similar in the different polymorphic

* Corresponding author. Tel.: +36 1 431 5922.
E-mail address: a.demeter@richter.hu (Á. Demeter).

0731-7085/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2008.09.042

forms, superficially similar vibrational spectra may be observed. In these cases the chemometric approach is inevitable for quantitative analysis.

Chemometrics is commonly used with near-infrared (NIR) spectroscopy, because overtones and combinations of different vibrational modes of the molecules appear as broad overlapping bands. Several authors have published encouraging applications of NIR spectroscopy in quantitative polymorph analysis [26]. However, spectral variation is relatively small, thus generally large training and test sets are required. As stretching and deformation frequencies of polymorphs are usually distinctive in wavelength or relative intensity, mid-IR spectroscopy is almost always able to differentiate them. Diamond ATR and diffuse reflectance infrared spectroscopy (DRIFTS) have been recommended over the KBr disc technique, because of the reduced possibility of polymorphic transformation during grinding and pressing [2,27,28]. However, ATR is a surface technique which can be affected by sample inhomogeneity in measuring polymorphic mixtures. More parameters influence the intensities in DRIFTS [10], and need to be controlled; otherwise significant analytical error will result [19]. For those cases for which transmission IR is applicable there is no advantage in using other methods which involve controlling additional parameters [29].

The advantage of Raman over IR spectroscopy has been emphasized by several authors [10,13,14]. The technique requires no sample preparation; samples can be measured in sealed glass vials, aqueous slurries, tablets in blisters, and the vibrational bands are usually narrower, which improves the possibility of distinguishing the solid forms. Quantification is based on the direct proportionality of the intensity of the specific band of the component and its concentration in the mixture. Chemometrics is equally applicable for the analysis of either Raman or IR spectral data when univariate analysis is inappropriate [30–36].

In our current study the combination of chemometrics with transmission FT-IR spectroscopy and FT-Raman spectroscopy is compared for the quantification of polymorphic forms in binary mixtures. Clopidogrel bisulphate was selected as a model compound, which is known to be effective in reduction of atherosclerotic myocardial infarction, stroke and death. Form II as well as several solvate forms and amorphous form of the substance are patented [37], but Form I is now open for generic development. As clopidogrel bisulphate polymorphs comprise an enantiotropic system, and Form II is the thermodynamically more stable form at room temperature [38,22], there is a potential for the occurrence of Form II in Form I both in the production steps and during the storage period. This fact requires a suitable analytical technique for the detection and quantification of the stable form in the metastable one. A recent publication has reported on the quantitative analysis of clopidogrel bisulphate polymorphs by X-ray powder diffraction [39]. The authors used whole powder pattern decomposition as well as classical direct methods for quantitation in the range of 10–80% Form I in Form II. The limit of detection using both methods is in the range of 1–2% of phase content in the mixture. A previous study utilized transmission FT-IR spectroscopy for the quantitative measurement of clopidogrel bisulphate polymorphs [22], by using unique peaks of the forms in the analytical range of 10–90% Form I in Form II. Low levels of Form II in Form I cannot be detected because characteristic bands of Form II are not visible in the IR spectra of mixture below 30%. The aim of our present study was to develop methods for the quantification of low levels of Form II in Form I by transmission FT-IR and FT-Raman spectroscopy. The use of different multivariate methods and spectral processing procedures were compared in the case of both techniques. Common problems of quantitative vibrational spectroscopy in solid phase are discussed; and procedures appropriate to eliminate them are proposed.

2. Materials and methods

2.1. Materials

Clopidogrel base was obtained from the camphorsulphonate salt of the compound. To produce the pure polymorphic Form I of clopidogrel bisulphate, clopidogrel base was dissolved in 9 parts of methyl *tert*-butyl ether and added to a mixture of equimolecular amount of conc. H_2SO_4 and two parts of n-decanol at 25 °C. Pure Form II was precipitated by the addition of equimolecular amount of conc. H_2SO_4 to the solution of clopidogrel base in 4.5 parts of acetone at 25 °C.

Identification of the forms was carried out by IR and Raman spectroscopy, as well as X-ray powder diffraction. According to our own experiments, which agree with the published results [39], X-ray powder diffraction is able to detect 1% of each form as an impurity in the other one. The two starting samples proved to be pure polymorphs and therefore were appropriate for the preparation of calibration mixtures.

2.2. Preparation of polymorphic mixtures

Polymorphic mixtures containing 1, 2, 5, 8, 10 and 15% of Form II, were prepared by geometric mixing of the pure forms in agate mortar with a pestle. Although it has been reported [22] that grinding does not induce polymorph transition of clopidogrel bisulphate polymorphs, it was also verified by the above mentioned techniques. Accurately weighted amount of previously ground pure forms were mixed by stepwise addition of Form I to the minor component.

Validation mixtures containing 3, 7 and 11% of Form II were prepared in the same way. These mixtures were not incorporated into calibration models, but were used for assessing the validity of them.

2.3. FT-IR spectroscopy

Infrared spectra were measured by Thermo Nicolet 6700 FT-IR spectrometer accumulating 100 scans at 4 cm^{-1} resolution. About 2 mg of sample was slightly ground with about 200 mg of KBr and pressed to a pellet of 13 mm in diameter in hydraulic press at about 700 MPa for 20 s. Each sample was measured in triplicate, i.e. repeating the sample preparation three times.

2.4. FT-Raman spectroscopy

Raman spectra were collected by Thermo Nicolet NXR-9650 FT-Raman spectrometer equipped with Nd-YAG laser source at 1064 nm wavelength and liquid nitrogen cooled Ge detector; 64 scans were co added at 750 mW exciting power and 4 cm^{-1} spectral resolution. A special sample holder accessory (see below) was used in the MicroStage™ with laser spot focused on 50 μm.

2.5. X-ray powder diffraction

Diffractograms of pure clopidogrel bisulphate forms were measured on a PANalytical X'Pert PRO diffractometer using Cu Kα radiation with 40 kV accelerating voltage and 40 mA anode current at a scanning rate of 0.031° 2θ min^{-1} over the range of 2–40° 2θ with 0.013° step size in reflection mode. Quantification of Form II in samples of unknown composition was performed by the auto scale routine of PANalytical X'Pert HighScore Plus 2.2b software using the diffractograms of pure forms.

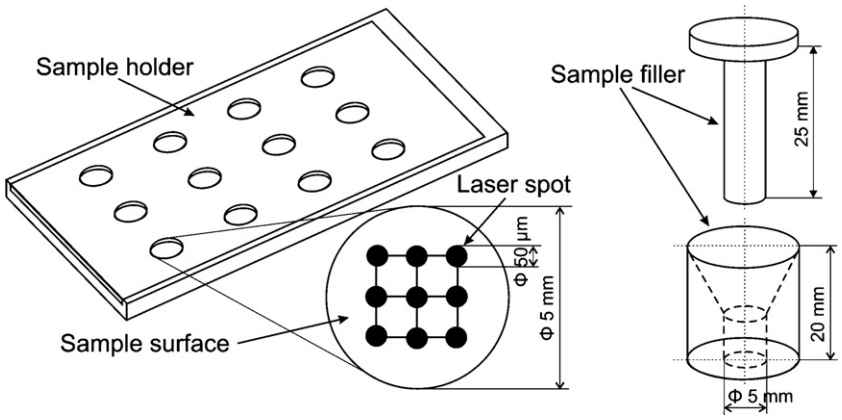

Fig. 1. Sampling accessory used for collecting Raman spectra (on the left), and the tool for sample filling (on the right).

2.6. Multivariate analysis

Multivariate data analysis was carried out by Thermo Nicolet TQ Analyst 7.2 software. Classical least squares (CLS), principal component regression (PCR) and partial least squares (PLS) methods were tested with different pathlength types and spectral preprocessing steps over different spectral ranges. The results were compared using linear correlation coefficients, root-mean-squared errors of calibration (RMSEC), cross-validation (RMSECV) and prediction (RMSEP) values, as well as the relative difference (Rel.Diff.) between the nominal concentration and the predicted one for the indication of model accuracy, and relative standard deviation (RSD) of multiple measurements assessing model precision. The definitions are the followings:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - Y_i)^2}{n}}$$

where y_i, Y_i and n are the calculated value, the nominal value and the number of measurements, respectively, for calibration (RMSEC), cross-validation (RMSECV) and prediction (RMSEP)

$$\mathrm{Rel.Diff.} = 100\left|\frac{\bar{y}_i - \bar{Y}_i}{\bar{Y}_i}\right|$$

where $\bar{y}_i$ and $\bar{Y}_i$ are the mean calculated and mean nominal value for each concentration, respectively

$$\mathrm{RSD} = \frac{100}{\bar{y}_i}\sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}{n-1}}$$

3. Results

3.1. Preliminary examinations

It is pointed out by several authors [10,19,21,23] that the major source of error in solid state quantitative measurements is usually sample inhomogeneity. To prepare well homogenized mixtures for calibration purposes is always a great challenge. Solid forms with considerably different habit and particle size, as in the case of clopidogrel bisulphate polymorphs, can also be prone to segregation during mixing. In order to decrease the particle size difference between the samples, the pure forms were ground in a porcelain mortar for 3 min. Then accurately weighted amounts of the two forms were co-ground in agate mortar to give 200 mg of polymorph mixtures with the desired composition. Form I was added to Form II in three steps, so that in each step roughly equal amount of powders were mixed; and each addition was followed by about 1 min of gentle grinding.

Since each mixture was measured in triplicate, each sample was sub-sampled three times. In IR measurements the observed absorbance varied not only with the exact amount of polymorph mixture in the matrix material (KBr), but also with the time and intensity of grinding the sample with the matrix; moreover it was also dependent on the actual pressure and dwell-time of the compression. This variation greatly exceeded the absolute intensity change originating from instrument instability, and that caused by variation in composition. Therefore, weighing the sample and matrix material seemed to be unnecessary. KBr discs were prepared in such a way as to approximate the absorbance of the strongest band to unity (the actual values varied from 0.63 to 1.56).

Raman measurements were carried out on the MicroStage™ of Thermo Nicolet instrument in a special stainless-steel sample holder designed for this purpose. A high power laser beam focused in very small volume can damage the sample, and can also lead to significant sub-sampling in quantitative analysis. To overcome sub-sampling problems of inhomogeneous samples a rotating sample holder, originally developed to avoid sample heating problems [40], was improved to increase the sampled volume for quantitative analysis [25]. In our experimental setting relatively high excitation power was used on 50 μm spot size to acquire the highest quality data possible. This power, applied for much more time than used for the actual measurements, led to no sample degradation. Defocusing the laser beam greatly decreased the measured intensity and hence the signal-to-noise. To investigate sample inhomogeneity and possible sub-sampling each mixture was measured in three independent samples and nine measurements were performed on each sample. A similar procedure is used in dispersive Raman microscopy to get representative distribution of different components in tablets [41]. Holes, 5 mm in diameter, were filled with about 20 mg of sample on a home-made sample holder plate, and this was placed in the stage of the spectrometer. Nine spectra (64 co-added scans for each) were measured from different points of the holes in a predefined 1.5 mm × 1.5 mm grid. Fig. 1 shows the sampling accessory. The plate with the holes fits into a rimmed outer plate which fastens it into the stage of the spectrometer. Using a simple filling accessory, the holes in the plate can easily be filled from the back: while the plate is held on an even surface, sample powder is fed into the opening and then compacted into the hole by hand using the rod, so obtaining a smooth surface. It was not possible to obtain such a smooth and compact surface in commercially available well plates. The accessory also facilitates data accumulation from predefined positions of multiple samples by moving the MicroStage controlled through the Omnic software.

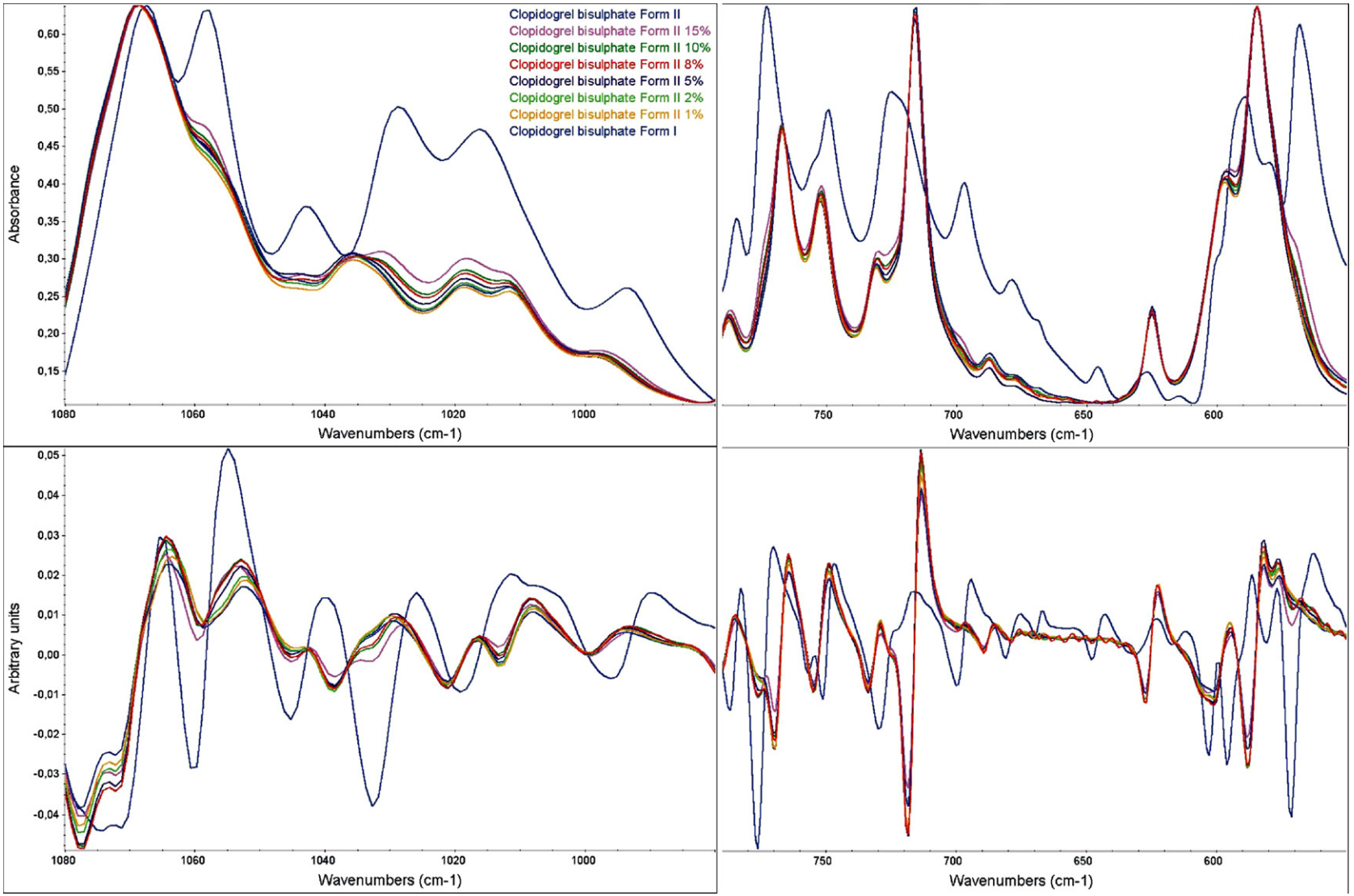

Fig. 2. IR spectra (top) and their 1st derivative (bottom) of polymorphic mixtures. From bottom to top, in both figures: pure Form I, 1%, 2%, 5%, 8%, 10%, 15% of Form II and pure Form II.

Table 1
Characteristics of various chemometric models built on IR data.

1st derivative IR spectra	Full range[a]			Meaningful range[b]			Selected ranges[c]		
	CLS	PCR	PLS	CLS	PCR	PLS	CLS	PCR	**PLS**
r	0.9952	0.9970	0.9997	0.9967	0.9989	0.9995	0.9979	0.9974	**0.9976**
RMSEC	0.497	0.394	0.126	0.411	0.241	0.154	0.328	0.363	**0.354**
RMSECV	0.679	0.795	0.562	0.562	0.661	0.541	0.537	0.553	**0.553**
RMSEP	0.416	0.335	0.247	0.278	0.172	0.174	0.205	0.134	**0.124**
CV residuals[d]	–	– –	–	–	–	–	+	+	+
No. of factors		4	4		3	3		1	**1**
% PC 1[e]	60.6	60.6	60.6	69.62	69.62	69.62	95.88	95.88	**95.88**

[a] Spectral range of 4000–400 cm^{-1}.
[b] Spectral ranges of 3150–2700 and 1550–405 cm^{-1}.
[c] Spectral ranges of 1440–1430, 1063–1029, 904–870, 780–767, 575–565 cm^{-1}.
[d] Symbols (–) and (– –) marks mean that the distribution of the residuals is not, and expressly not random around zero, respectively (curvature and/or drift can be observed); (+) means random distribution.
[e] The percentage contribution of fist principal component to the spectral variation in the selected spectral range.

The measurements therefore may be automated, so utilizing the equipment maximally, and enabling data to be collected for the whole calibration series in one night.

Each hole was filled with particular care; thereby samples were equally compacted and they have smooth and even surface. In addition, measurements at each point were preceded by auto focusing, i.e. finding the interferogram maximum. Nevertheless, probably because of refraction and polarization effects at the microscopically unevenly oriented crystal surfaces, the absolute spectral intensities varied significantly point by point (in certain cases by 30%). This intensity variation, however, does not inevitably affect the quantitative analysis, because normalization generates spectra practically identical with those of samples of the same composition.

In a quantitative Raman spectroscopic assay another important aspect can be the variation of particle size in mixtures of different compositions. There is no full agreement in the literature about the effect of particle size on the measured Raman intensity, because some experimental observations [42] seem to contradict theory [43]. Recent investigations support the recommendation that care must be taken to achieve particle size uniformity [44]. To assess this effect, clopidogrel bisulphate Form II was measured both as received and after the above described grinding procedure. The initial substance has large crystals with irregular shape; most of the particles were greater than 100 μm. Grinding decreased the particle size significantly; most of the particles become smaller than 10 μm and the relative amount of fines (particles smaller than 1 μm) also increased. Actually, it was difficult to form a smooth surface in the sample holder from the crude initial material as opposed to the ground sample that was a very fine powder. The measured spectra of both samples were identical; the absolute intensity of the most intense band at 1029 cm^{-1} was 97.4 (8.6) and 95.6 (10.0) for the initial and ground sample, respectively. The values in parenthesis are the standard deviations from 9 consecutive measurements. This is in accord with our previous findings with other pharmaceutical materials. It may be assumed that our compaction procedure minimises the effect of particle size on Raman signal intensity. This is in agreement with the observations discussed in a recent thorough study [45], where the Raman intensity was also independent of particle size if a large spot size Raman probe was used.

3.2. FT-IR examinations

It has been noted previously that the presence of less than 30% of Form II in Form I of clopidogrel bisulphate has no visible spectral signature [22]. In our opinion this statement is rather pessimistic. The visual detection limit of Form II based on the peak shift of the 1035 cm^{-1} band characteristic of Form I toward 1029 cm^{-1} characteristic of form II is around 15%. A lower percentage of Form II in Form I remains invisible. Characteristic vibrational bands at 1439, 1058, 1029, 883, 867, 773 and 568 cm^{-1} appear as shoulders on increasing the amount of Form II in the mixture. This relative change is unequivocally observable in certain spectral ranges, especially on the derivative spectra (Fig. 2).

Multivariate calibrations were applied to raw data as well as 1st or 2nd derivatives, either using the whole spectrum or selecting specific regions, among which 1st derivative spectra gave significantly better models. Table 1 summarizes the results of chemometric models built from 1st derivative spectral data. Mean centering and variance scaling proved to be useful preprocessing operations, as well as multiplicative scatter correction (MSC) to account for pathlength differences. However, normalizing spectra for the net intensity of 1753 cm^{-1} band, common for both polymorphic forms, and standard normal variate (SNV) pathlength correction did not give significantly different results.

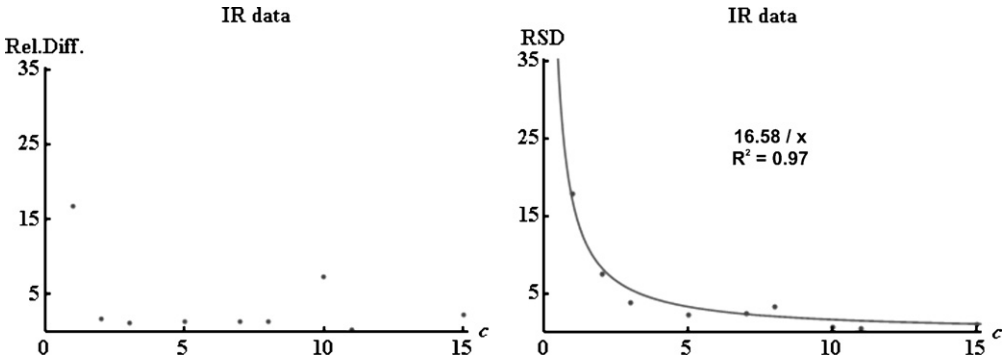

Fig. 3. Relative estimated error (left) and relative standard deviation (right) as the function of concentration in IR chemometric method.

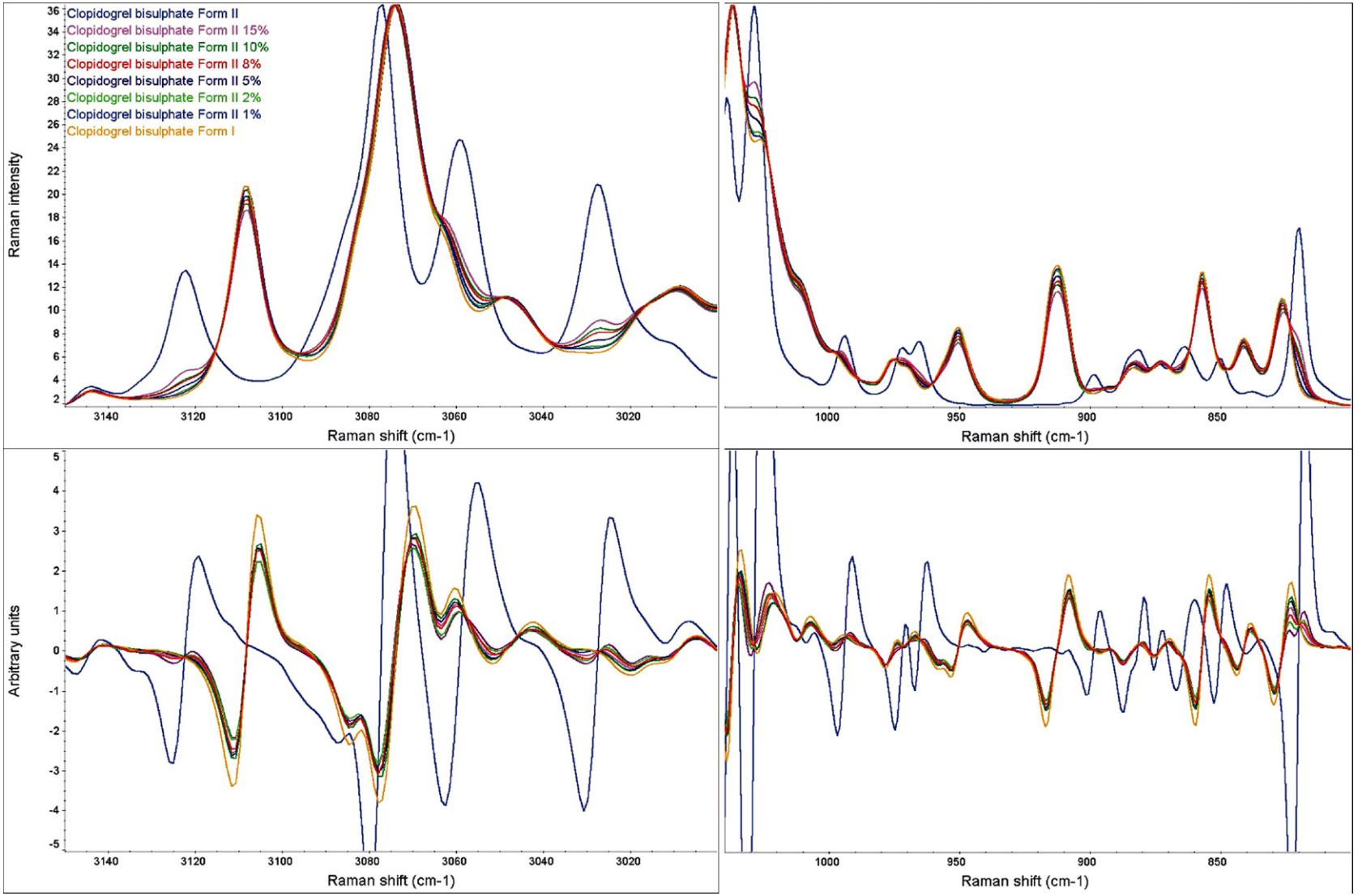

Fig. 4. Raman spectra (top) and their 1st derivative (bottom) of polymorphic mixtures. From bottom to top, in both figures: pure Form I, 1%, 2%, 5%, 8%, 10%, 15% of Form II and pure Form II.

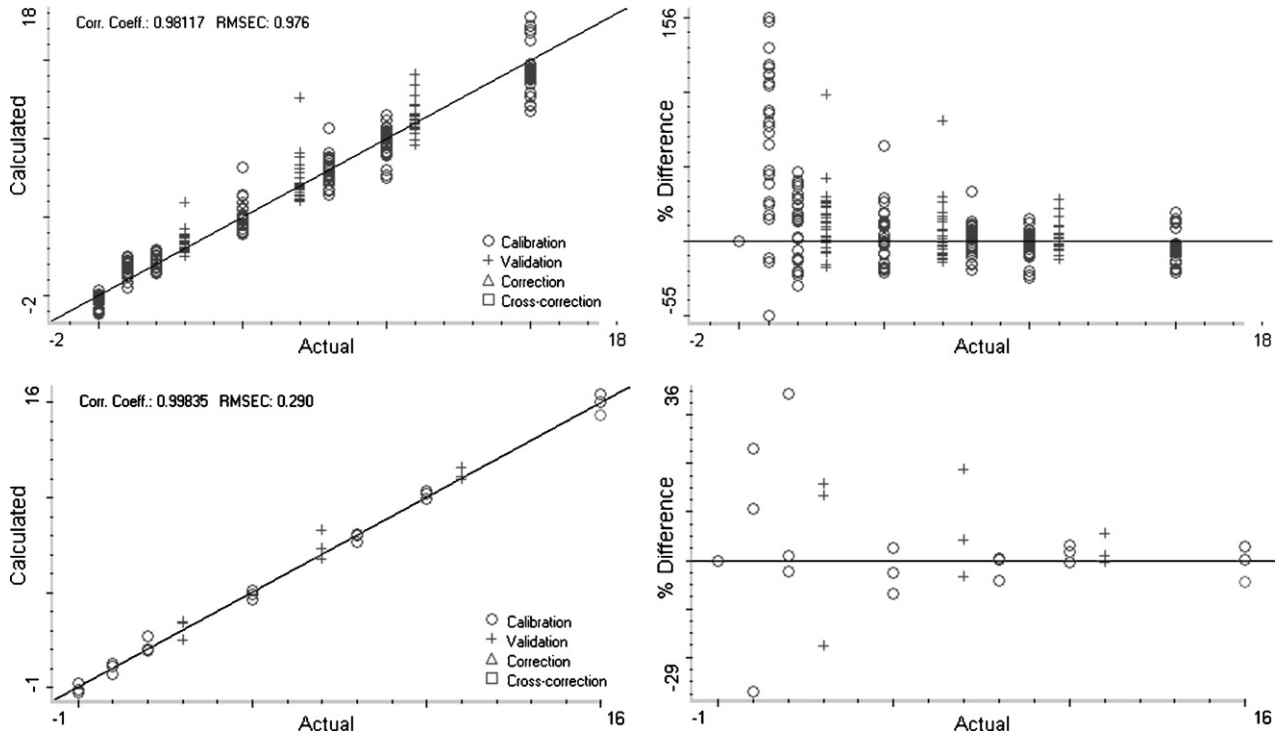

Fig. 5. Calculated vs. actual concentration (left) and the distribution of residuals (right) from the model using all measured Raman spectra (top) and composition spectra (bottom).

Results show that similarly high correlation coefficient and low RMSEC values can be obtained either on the whole spectral data or on those selected ranges more indicative of concentration variation. Using the full spectral range or meaningful spectral ranges (where vibrational bands of the substance are located, in the case of clopidogrel bisulphate 3150–2700 and 1550–405 cm^{-1}) the PLS model gives better results than CLS or PCR. In these cases, however, low RMSECV and RMSEP values are obtained only by using 3 or 4 factors. In addition, residuals, i.e. the difference between theoretical composition and calculated one based on cross-validation equation, are not randomly distributed around zero. This can be a sign of over-fitting. On the contrary, using only spectral data from carefully selected ranges, CLS, PCR and PLS models give equally high correlation coefficient and low RMSEC, randomly distributed cross-validation residuals, and RMSEP around 0.1–0.2. In addition, the optimal number of factors is 1 as determined by finding the minimum of RMSECV. This indicates that spectral variation used for constructing the chemometric model originates only from concentration changes. The first step in selection of the most appropriate spectral ranges was to identify the regions showing the most substantial differences between Form I and Form II. Then the derivative spectra of calibration mixtures were inspected to find the limits for the regions where the variation of band shape with concentration was most pronounced. Finally the performance characteristics of models built on different combination of ranges were compared to find the best. Including more or less regions than this resulted in weaker model characteristics; smoothing the data proved to be little disadvantageous.

As the indicators of model accuracy and precision, Fig. 3 shows the relative error of estimation (Rel.Diff.) and relative standard deviation (RSD) from the best model found (typed bold in Table 1) as a function of actual concentration. Because absolute error and standard deviation are relatively constant in the whole concentration range, the RSD decreases more or less regularly as a reciprocal function. Both Rel.Diff. and RSD are smaller than 5% above a 3 wt% concentration of Form II, except for the 10% calibration mixture. The estimation of limit of detection (LOD) and limit of quantitation (LOQ) was carried out by multiplying the standard deviation of the blank value by 3.3 and 10 and dividing it by the slope of calibration curve [46]. With the slope not significantly different than 1, and 0.2% standard deviation determined for the pure Form I this resulted LOD = 0.7% and LOQ = 2.0%. Regression was applied by Mathematica 7.0 for least-squares fit of one parameter reciprocal function to the RSD vs. concentration data set. From the high correlation coefficient it may be assumed that the equation obtained accurately describes the concentration dependence of relative standard deviation. Using the fitted RSD curve the measured RSD value for every mixture can be used for the estimation of LOD and LOQ. Solving the equation, shown in Fig. 3, for RSD = 33% and 10% resulted in LOD = 0.5% and LOQ = 1.7%, respectively. Considering that the measured RSD for the validation mixture containing 3 wt% of Form II was 3.8%, well below the 10% which is usually considered as satisfactory in similar assays, the above estimation for LOQ seems to be correct, or even precautious.

3.3. FT-Raman examinations

Increasing concentrations of Form II resulted in changes of band shape in the Raman spectra. Characteristic vibrations appeared as shoulders, for example, at 3122, 3027, 1029, 972, 965, 820, 647 and 306 cm^{-1}. However, unique bands, the intensity of which could be appropriate for quantification, are not present. As in the case of IR spectroscopy, the variation is distinctly visible on the 1st derivative spectra (Fig. 4).

The use of 1st derivative spectral data as input in chemometric models provided better results than spectra or 2nd derivatives, but the difference was not so outstanding than in the case of IR spectroscopy. Mean centering was necessary to obtain low RMSEC, RMSECV and RMSEP; variance scaling, however, decreased the merit of fit. MSC proved to be the best pathlength correction

Table 2
Characteristics of various chemometric models built on averaged Raman data.

1st derivative Raman spectra	Full range[a]			Meaningful range[b]			Selected ranges[c]		
	CLS	PCR	PLS	CLS	PCR	PLS	CLS	PCR	PLS
r	0.9962	0.9887	0.9934	0.9962	0.9894	0.9937	0.9979	0.9983	**0.9984**
RMSEC	0.440	0.758	0.581	0.444	0.733	0.565	0.331	0.291	**0.290**
RMSECV	0.585	1.130	0.814	0.556	1.070	0.764	0.347	0.315	**0.314**
RMSEP	0.502	0.753	0.683	0.512	0.750	0.684	0.682	0.569	**0.570**
CV residuals[d]	+	– –	+	+	–	+	+	+	**+**
No. of factors		1	1		1	1		1	**1**
% PC 1[e]	61.12	61.12	61.12	64.02	64.02	64.02	96.37	96.37	**96.37**

[a] Spectral range of 3500–200 cm^{-1}.
[b] Spectral ranges of 3200–2900 and 1800–225 cm^{-1}.
[c] Spectral ranges of 3135–3117, 3050–3021, 1033–1024, 823–810, 705–692, 651–641, 319–299 cm^{-1}.
[d] Symbols (–) and (– –) mean that the distribution of the residuals is not, and expressly not random around zero, respectively (curvature and/or drift can be observed); (+) means random distribution.
[e] The percentage contribution of fist principal component to the spectral variation in the selected spectral range.

procedure; applying SNV or normalization gave a bit worse results. Using all the collected spectra (9 per samples, each re-packed 3 times), i.e. 189 calibration and 81 validation spectra, relatively large variation was found within the spectra of a given composition. This could be caused by the inhomogeneous distribution of two forms in the mixtures on the 50 μm scale of laser spot size. The problem was solved by averaging the nine spectra collected from different sample positions of one sample. These averaged 9 spectra are referred to as the composition spectrum for each mixture. This spatial averaging significantly improved the model characteristics as shown in Fig. 5. The linear fit is very good using only one factor and the residuals are evenly distributed around zero in both cases. After averaging (creating composition spectra), however, significantly higher correlation coefficient as well as lower RMSEC, RMSECV and RMSEP values were obtained.

Table 2 shows the comparison of chemometric models built on different ranges of averaged Raman spectra. Unlike IR data, using full spectral range or with spectral data from a meaningful range (in this case 3200–2900 and 1800–225 cm^{-1}) every method gives satisfactory results with only one factor. However, the most appropriate spectral ranges, which were selected in the same way as in the case of IR spectral data, assure the lowest error values and highest contribution of the 1st factor (or principal component) to the spectral variation. CLS, PCR and PLS models give similar results, suggesting that the merit of calibration is independent of the model used.

Fig. 6 shows the relative error of estimation (Rel.Diff.) and relative standard deviation (RSD) as the function of actual concentration from the best model (PLS on 1st derivative of selected spectral ranges, typed bold in Table 2). Rel.Diff. is smaller than 15% in the whole concentration range, and drops to about 5% above 3 wt% of Form II. RSD is changing approximately by reciprocal function and become smaller than 10% above 5 wt% concentration of Form II. With the same estimation as in the case of IR data, using 0.3% standard deviation measured for pure Form I, LOD = 1.0%, and LOQ = 3.0% were obtained. The fitted equation of Fig. 6 resulted in LOD = 0.9% and LOQ = 3.1%.

4. Discussion

The above results clearly show that quantification of low levels of Form II of clopidogrel bisulphate in polymorphic mixtures is possible by both IR and Raman spectroscopy through chemometric data analysis. Correlation coefficients and error values in the tables indicate that both perform equally well. This is also confirmed by similar relative estimated error and relative standard deviation seen in Figs. 3 and 6.

In order to analyse the data for chemometric model building one useful approach can be to use different processing methods on the same data set [16]. In our case CLS, PCR and PLS algorithms give similarly good models both utilizing the full measured Raman spectral data and only selected spectral ranges. This is more or less also true for the models built from IR spectral data, but there is borderline significant effect of changing the structure of the input: by using more spectral information unrelated to concentration variation, additional factors enter into the model, and the statistics of residuals become poorer. The relative ratio of the 1st principal component (% PC 1) in explaining the spectral variation in selected ranges significantly increases by applying more adequate data processing (careful range selection): from about 61 to 96% in both IR and Raman methods (cf. Table 1 and 2). 83.5% of Raman spectral variation is explained by PC 1 on the full spectrum and its loading

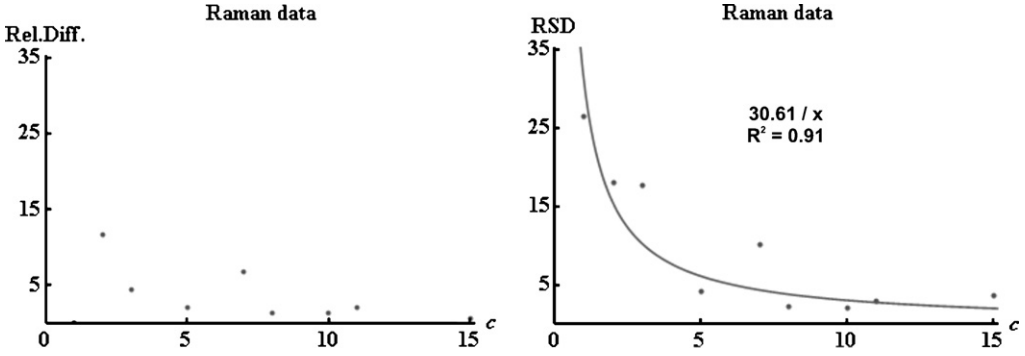

Fig. 6. Relative estimated error (left) and relative standard deviation (right) as the function of concentration in Raman chemometric method.

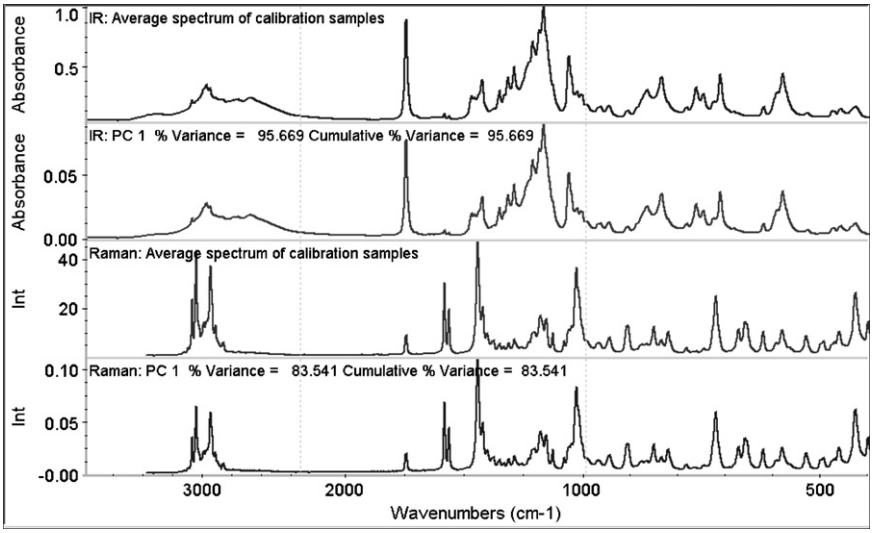

Fig. 7. Comparison of average spectra and 1st principal component spectra calculated for IR (top) and Raman method (bottom).

spectrum is identical with the average spectrum of the calibration mixtures (Fig. 7). Furthermore the loading spectrum of the 2nd factor (PC 2), describing 10.0% of variation, is very similar to the difference spectrum of Form I and Form II. These characteristics also point to the fact that the most significant factor in the model is concentration dependence. PC 1 in IR methods is even more significant; it describes 95.7% of full spectral variation and is also a perfect match for the average spectrum (see Fig. 7). In spite of the fact that in our study the highest percentage of Form II is only 15% in the mixtures, the pure IR and Raman spectra calculated by TQ Analyst for Form II are almost identical with the corresponding Form II spectra of clopidogrel bisulphate.

It is well-known in industrial practice and usually stressed by experts [10,12,18] that models without proved real-life applicability have minute value. Using validation samples can assure that the models obtained will be free from over-fitting. The question remains open as to whether these models will accurately quantify the polymorph composition of samples of different physical properties (particle size, habit, etc.), originating from different manufacturing batches and/or synthetic routes. In order to test the applicability of the developed quantitative models on general samples, clopidogrel bisulphate polymorphic mixtures of unknown composition, crystallized under different conditions, were measured. For comparison, the polymorph composition was also assessed by X-ray powder diffraction. Least-squares full pattern fitting was performed by X'Pert HighScore Plus software, auto scaling the diffractograms of pure forms to the pattern of the unknown samples after baseline adjustment. As can be seen in Table 3, there is a good agreement between IR and Raman results; and the determined Form II percentages are close to the values estimated by XRPD. Since from other considerations it is very likely that sample 1 contains about 1% of Form II, it seems that the detection limit of IR and Raman methods are rather higher than 1 wt%, which was calculated from the calibration and validation results. The accuracy above 5%, however, can be considered satisfactory.

Table 3
Composition of unknown samples measured by different analytical methods.

% Form II	XRPD	IR	Raman
Sample 1	1.0	−1.0	0.3
Sample 2	11.0	12.9	11.4
Sample 3	7.0	6.5	5.6

5. Conclusions

IR and Raman spectroscopy alone are unsuitable for the detection of low levels of clopidogrel bisulphate Form II in polymorphic mixtures with Form I. However, with proper chemometric data processing and model building quantification is possible down to 3 wt% by both IR and Raman methods. The accuracy is better than 1 wt% and RSD is smaller than 10% in the analysis range, i.e. 3–15 wt%. The best chemometric model utilizes PLS analysis of selected ranges of the first derivative spectrum after preprocessing operations (mean centering and multiplicative scatter correction). The first principal component covers 96% of the spectral variation of selected ranges, which indicates that the variation is predominantly due to the change in analyte concentration. The models obtained lead to correct measurements of the polymorph composition of API samples completely independent of those used for model construction.

Clopidogrel bisulphate model system allowed quantitative methods to be developed by using transmission infrared spectroscopy in potassium bromide matrices, which provides a simple method for the measurement of calibration and unknown spectra with easy sample preparation. Neither accurately measuring the weight of samples and matrix material nor any special mixing and pellet preparation protocol was necessary. The only prerequisite for constructing adequate chemometric model was to collect good quality spectra.

Raman spectroscopy requires minimal sample preparation. Despite sub-sampling due to the small laser spot size even in carefully homogenized calibration samples, an adequate description of the bulk sample analyte concentration can be assured by collecting and averaging spectra from different positions of the sample in an appropriate sample holder. It was proved that a relatively good calibration model can be constructed by chemometric evaluation either using the whole measured spectra or their derivatives, either by normalizing for intensity variation or by correcting the pathlength with MSC and SNV, providing the quality of the data set is high. The latter requirement implies the collection of high quality spectra from calibration samples with accurate and statistically representative composition within the analysis scale. Furthermore, by careful selection of the analytically meaningful spectral ranges which vary most significantly with concentration, better accuracy and precision as well as robustness can be obtained.

Acknowledgements

Technical assistance of Zsuzsanna Hoffmann in measuring IR and Raman spectra and of István Torma in constructing the Raman sampling accessory is greatly acknowledged. Authors thank Dr. Terry Threlfall for his valuable suggestions and correction of the manuscript.

References

[1] J.K. Haleblian, W.C. McCrone, J. Pharm. Sci 58 (1969) 911–929.
[2] T.L. Threlfall, Analyst 120 (1995) 2435–2460.
[3] H.G. Brittain (Ed.), Polymorphism in Pharmaceutical Solids, 95, Marcel Dekker, New York, 1999.
[4] J. Bernstein, Polymorphism of Molecular Crystals, Clarendon, Oxford, 2002.
[5] S. Byrn, R. Pfeiffer, M. Ganey, C. Hoiberg, G. Poochikian, Pharm. Res. 12 (1995) 945–954.
[6] A.S. Raw, M.S. Furness, D.S. Gill, R.C. Adams, F.O. Holcombe Jr., L.X. Yu, Adv. Drug. Deliv. Rev. 56 (2004) 397–414.
[7] Food and Drug Administration, Center for Drug Evaluation and Research, Guidance for Industry, ANDAs: Pharmaceutical Solid Polymorphism: Chemistry, Manufacturing, and Controls Information, http://www.fda.gov/cder/guidance/index.htm, 2007, June.
[8] J. Lucas, P. Burgess, PharmaVoice (2004) 54–57.
[9] A. Zakrzewski, M. Zakrzewski (Eds.), Solid State Characterization of Pharmaceuticals, Assa International Inc, Danbury, Connecticut, USA, 2006.
[10] D.E. Bugay, Adv. Drug. Deliv. Rev. 48 (2001) 43–65.
[11] D.A. Giron, Thermochim. Acta 248 (1995) 1–59.
[12] G.A. Stephenson, R.A. Forbes, S.-M. Reutzel-Edens, Adv. Drug. Deliv. Rev. 48 (2001) 67–90.
[13] W.P. Findlay, D.E. Bugay, J. Pharm. Biomed. Anal. 16 (1998) 921–930.
[14] M.E. Auer, U.J. Griesser, J. Sawatzki, J. Mol. Struct. 661–662 (2003) 307–317.
[15] T. Vankiersblick, A. Vercauteren, W. Baeyens, G. van der Weken, F. Verpoort, G. Vergote, J.P. Remon, Trends Anal. Chem. 21 (2002) 869–877.
[16] M.J. Pelletier, Appl. Spectrosc. 57 (2003) 20A–42A.
[17] D.B. Black, E.G. Lovering, J. Pharm. Pharm. 29 (1977) 684–687.
[18] D.A. Roston, M.C. Walters, R.R. Rhinebarger, L.J. Ferro, J. Pharm. Biomed. Anal. 11 (1993) 293–300.
[19] D.E. Bugay, A.W. Newman, W.P. Findlay, J. Pharm. Biomed. Anal. 15 (1996) 49–61.
[20] G. Sertsou, S. Agatonovic-Kustrin, T. Rades, Internet J. Vib. Spec. [www.ijvs.com] 3, 5, 5 (1999).
[21] N. Al-Zoubi, J.E. Koundourellis, S. Malamataris, J. Pharm. Biomed. Anal. 29 (2002) 459–467.
[22] V. Koradia, G. Chawla, A.K. Bansal, Acta Pharm. 54 (2004) 193–204.
[23] S.N. Campbell Roberts, A.C. Williams, I.M. Grimsey, S.W. Booth, J. Pharm. Biomed. Anal. 28 (2002) 1135–1147.
[24] L.S. Taylor, G. Zografi, Pharm. Res. 15 (1998) 755–761.
[25] F.W. Langkilde, J. Sjöblom, L. Tekensberg-Hjelte, J. Mrak, J. Pharm. Biomed. Anal. 15 (1997) 687–696.
[26] Y. Roggo, P. Chalus, L. Maurer, C. Lema-Martinez, A. Edmond, N. Jent, J. Pharm. Biomed. Anal. 44 (2007) 683–700.
[27] K. Pöllänen, A. Häkkinen, M. Huhtanen, S. Reinikainen, M. Karjalainen, J. Rantanen, M. Louchi-Kultanen, L. Nyström, Anal. Chim. Acta 544 (2005) 108–117.
[28] K. Kipouros, K. Kachrimanis, I. Nikolakakis, S. Malamataris, Anal. Chim. Acta 550 (2005) 191–198.
[29] Z. Krivácsy, J. Hlavay, J. Mol. Struct. 349 (1995) 289–292.
[30] C.M. Deeley, R.A. Spragg, T.L. Threlfall, Spectrochim. Acta 47A (1991) 1217–1223.
[31] A.M. Tudor, S.J. Church, P.J. Hendra, M.C. Davies, C.D. Melia, Pharm. Res. 10 (1993) 1772–1776.
[32] G. Jalsovszky, O. Egyed, S. Holly, B. Hegedűs, Appl. Spectrosc. 49 (1995) 1142–1145.
[33] D. Pratiwi, J.P. Fawcett, K.C. Gordon, T. Rades, Eur. J. Pharm. Biopharm. 54 (2002) 337–341.
[34] C.J. Strachan, D. Pratiwi, K.C. Gordon, T. Rades, J. Raman Spectrosc. 35 (2004) 347–352.
[35] B. De Spiegeleer, B. Baert, N. Diericx, D. Seghers, F. Verpoort, L. Van Vooren, C. Burvenich, G. Slegers, J. Pharm. Biomed. Anal. 44 (2007) 254–257.
[36] F. Tian, F. Zhang, N. Sandler, K.C. Gordon, C.M. McGoverin, C.J. Strachan, D.J. Saville, T. Rades, Eur. J. Pharm. Biopharm. 66 (2007) 466–474.
[37] R. Lifshitz-Liron, E. Kovalevski-Ishai, S. Wizel, S. Avhar-Maydan, R. Lidor-Hadas, Polymorphs of clopidogrel hydrogensulfate, WO 03/051362 (18.12.2002).
[38] A. Bousquet, B. Castro, J. Saint-Germain, Polymorphic form of clopidogrel hydrogen sulphate US 6504030 (21.06.2002).
[39] V. Uvarov, I. Popov, J. Pharm. Biomed. Anal. 46 (2008) 676–682.
[40] S. Kaminaka, T. Kitagawa, Appl. Spectrosc. 49 (1995) 685–687.
[41] S.E.J. Bell, J.R. Beattie, J.J. McGravey, K.L. Peters, N.M. Sirimuthu, J. Speer, J. Raman Spectrosc. 35 (2004) 409–417.
[42] M.V. Pellow-Jarman, P.J. Hendra, R.J. Lehnert, Vibr. Spectrosc. 12 (1996) 257–261.
[43] M. D'Orazio, B. Schrader, J. Raman Spectrosc. 4 (1976) 253–268.
[44] H. Wang, C.K. Mann, T.J. Vickers, Appl. Spectrosc. 56 (2002) 1538–1544.
[45] Y. Hu, H. Wikström, S.R. Byrn, L.S. Taylor, Appl. Spectrosc. 60 (2006) 977–984.
[46] Food and Drug Administration, Center for Drug Evaluation and Research, Guidance for Industry, ICH Topic Q2B. Validation of analytical procedures: methodology, http://www.fda.gov/cder/Guidance/1320fnl.pdf, November, 1996.