International Journal of Pharmaceutics 385 (2010) 29–36

Contents lists available at ScienceDirect

International Journal of Pharmaceutics

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

# Simultaneous quantitative analysis of ternary mixtures of D-mannitol polymorphs by FT-Raman spectroscopy and multivariate calibration models

Doris E. Braun[a], Stephan G. Maas[b], Neslihan Zencirci[a], Christoph Langes[a], Nora A. Urbanetz[c], Ulrich J. Griesser[a,*]

[a] Institute of Pharmacy, University of Innsbruck, Innrain 52, 6020 Innsbruck, Austria
[b] Institute of Pharmaceutics and Biopharmaceutics, Heinrich-Heine-University Düsseldorf, Universitaetsstrasse 1, 40225 Duesseldorf, Germany
[c] Research Center Pharmaceutical Engineering GmbH, Inffeldgasse 21a/II, 8010 Graz, Austria

ARTICLE INFO

*Article history:*
Received 6 April 2009
Received in revised form
11 September 2009
Accepted 7 October 2009
Available online 14 October 2009

*Keywords:*
FT-Raman spectroscopy
D-Mannitol
Crystal polymorphism
Multivariate calibration
PLS regression
Artificial neural networks
Chemometry
Quantitative analysis

## ABSTRACT

D-Mannitol is known to exist in five solid-state forms, a hemihydrate, an amorphous form and three polymorphic forms (I°, II and III), which tend to crystallize concomitantly. Therefore, a fast and simple method for the simultaneous quantification of these polymorphs in powder mixtures was developed on the basis of FT-Raman spectroscopic data, partial least-squares (PLS) regression and artificial neural networks (ANNs). A combination of the first derivative and orthogonal signal correction (OSC) was found to be the optimal data pretreatment that significantly increased the predictive performance of the models. The RMSEPs (root-mean-squared errors of prediction) obtained by PLS for the modifications (mods.) I°, II and III were 0.44%, 0.34% and 0.36% respectively. The estimated limits of detection are ~0.5% (mod. I°) and <1% (mods. II and III). The ANNs model yielded slightly higher RMSEP values of 0.51%, 0.39% and 0.41%. In contrast to related previous studies, calibration was performed with carefully prepared ternary mixtures of all polymorphs, which is one of the reasons for the high precision and accuracy of the presented multivariate models.

## 1. Introduction

In many scientific disciplines multivariate statistical methods have gained importance for analyzing, evaluating and understanding huge data sets, consisting of numerous variables (Forina et al., 2007). Among the many subtopics and applications of multivariate extraction of information from chemical data (chemometrics) (Lavine and Workman, 2006) the development of quantitative assays based on multivariate calibration models and spectroscopic techniques is one of the most common applications in pharmaceutical and chemical disciplines. The majority of analytical tasks clearly concern problems related to diverse chemical entities (e.g. impurities). However, there is an increasing number of applications that focus on solid-state problems (Aaltonen et al., 2008), such as the analysis of solid-state forms (polymorphs, hydrates, solvates, amorphous form) of the pure active pharmaceutical ingredient in powder mixtures or formulations (Aboul-Enein et al., 2002; Dandeu et al., 2006; Heinz et al., 2007; Kachrimanis et al., 2007; Kipouros et al., 2005, 2006; Mazurek and Szostak, 2008; McGoverin et al., 2006; Pratiwi et al., 2002; Poellaenen et al., 2005; Strachan et al., 2004; Suda et al., 2008).

Powder X-ray diffraction (PXRD) and spectroscopy (IR and Raman) are the most commonly applied techniques for the identification and characterization of solid-state forms. PXRD is the gold standard for this task, since true polymorphs must exhibit structural differences (exactly the same chemical composition but differences in crystal packing). However, the most critical limitation for quantitative assays is the possible variability of the peak intensities due to preferred orientation of the crystals. Therefore, special sample preparation is required (Jenkins and Snyder, 1996). Spectroscopic methods (IR, Raman, solid-state NMR) are highly sensitive to changes in the short-range interactions of the molecular assemblies in different polymorphs and this makes them an ideal complement to PXRD (Auer et al., 2003). Infrared spectroscopy, particularly near infrared (NIR), is the dominating analytical technique for such applications. However, one can recognize an increasing preference for Raman spectroscopy, which is obvious from the rising number of publications reporting the use of this method. Apart from the vast technical advancement of Raman spectrometers this trend can be attributed to the fact

---

* Corresponding author at: Institute of Pharmacy/Pharmaceutical Technology, University of Innsbruck, Innrain 52, 6020 Innsbruck, Austria. Tel.: +43 512 507 5309; fax: +43 512 507 2939.
*E-mail address:* ulrich.griesser@uibk.ac.at (U.J. Griesser).

0378-5173/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijpharm.2009.10.019

that Raman spectroscopy offers a variety of advantages compared to other spectroscopic techniques. One benefit is the convenience in sample preparation, which enables recording Raman spectra during ongoing chemical processes, in or through closed reaction vessels. Furthermore, unlike with NIR spectroscopy, the method is not sensitive to particle size variations, and Raman spectra provide more spectral information (higher number of bands and much sharper bands) than NIR spectra. Thus, the information content one can exploit with univariate or multivariate methods can be significantly higher compared to NIR spectroscopy.

Though there are virtually no crystal forms, which cannot be distinguished by Raman spectroscopy (Auer et al., 2003) the differences of Raman spectra can be subtle. Therefore, the development of quantitative assays, based on conventional univariate (peak height/area) calibration methods, may fail. Consequently, the application of more complex multivariate methods, which allow an evaluation of the smallest reproducibly measurable differences, is required. Univariate methods very often fail when a higher number of polymorphs (three or more) are quantified simultaneously. Because of the fact that multivariate algorithms such as PLS (partial least-squares) regression, ANNs (artificial neural networks), SVMs (support vector machines) and others are capable to deal with many thousand variables, they can be employed for these tasks. Moreover, multivariate methods can improve the accuracy of a quantification method since the variation (wavenumber shifts, intensity) over the entire spectral range is considered and not only the intensity of selected bands (Anderton, 2003 and refs. therein). Thus, the high reproducibility of Raman spectra combined with a sophisticated multivariate algorithm is a promising basis for the development of a successful quantitative model, even for rather complex mixtures.

The usefulness and power of the PLS method (Martens and Naes, 1991) in the spectroscopic quantification of polymorphic powder mixtures was demonstrated in numerous applications. For example, Poellaenen et al. (2005) quantitatively characterized the polymorphic composition of sulfathiazole, McGoverin et al. (2006) quantified binary polymorphic mixtures of ranitidine hydrochloride using NIR spectroscopy, and Kachrimanis et al. (2007) recently quantified paracetamol polymorphs in powder mixtures by FT-Raman spectroscopy and PLS regression. ANNs have been successfully applied for cases where no linearity between the input variable and the determined properties exist (Despagne and Massart, 1998). The prediction power of different multivariate calibration methods (including PLS and ANNs) in the analysis of polymorph mixtures was demonstrated by Kipouros et al. (2006). This study aimed at the quantification of three different crystal forms of carbamazepine in powder mixtures based on data that were obtained with diffuse reflectance FTIR spectroscopy (DRIFTS).

In order to enhance the predictive power of a multivariate calibration model it is important to establish the best model and the best methodology for the raw data pretreatment. The purpose of a preprocessing is to remove any variation in the input data, not related to the concentration of a particular given component (Wold et al., 1998). Commonly used spectral preprocessing methods are multiplicative scatter correction (MSC) (Geladi et al., 1985), standard normal variate (SNV) transformation (Barnes et al., 1989), first and second order derivative and orthogonal signal correction (OSC) (Wold et al., 1998; Eriksson et al., 2001). Furthermore, all these preprocessing methods can be applied alone or in combination.

Our aim was to develop a suitable quantitative model for ternary mixtures of D-mannitol polymorphs. This task emerged in connection with a project that aimed at the production and evaluation of spray-dried D-mannitol as carrier (alternative to lactose) for pulmonary administered drugs. Depending on the spray drying parameters, especially on the drying temperature, different morphologies of the spray-dried products were obtained, allowing the tailoring of interparticle interactions between the carrier and the drug in ordered mixtures and the performance of the dry powder inhalate. Hot-stage microscopy revealed that a low drying temperature (60 °C) favors the crystallization of D-mannitol from a supersaturated solution, whereas at high temperature (120 °C) a largely water free melt is formed that crystallizes spontaneously, when the temperature drops to about 90 °C. As different polymorphs may be obtained, depending on the rate of solvent evaporation and the cooling rate of D-mannitol solutions or the melt, we were interested in analyzing the phase composition of the spray-dried products.

Due to its safety and its compatibility with other drugs (Kibbe, 2000) D-mannitol is one of the most important excipients in pharmaceutical and food industries, where it is used as sugar substitute, as additive in tablets, capsule formulations and granules, for parenteral formulations, etc. The characterization of the solid-state behavior of D-mannitol has been the topic of many studies. Burger et al. (2000) summarized past literature and performed a comprehensive study on the solid-state properties of the polymorphs including their powder compaction properties and thermodynamic stability. Here the nomenclature of Burger et al. (2000) will be used, following the order of decreasing melting points. Modification (mod.) $I^\circ$ is additionally marked with a superscript zero to indicate that this is the thermodynamically stable polymorph at room temperature. The mods. $I^\circ$ (beta form) and II (alpha form) are energetically very similar and enantiotropically related to mod. III (delta form), which is the least stable form at and above room temperature. Mod. III ($T_{fus} = 155\,^\circ C$) melts about 10 K below the melting point of the two other forms ($T_{fus} = 166\,^\circ C$). A specific feature of this polymorphic system is the high kinetic stability of the metastable forms (mods. III and II). In fact, we observed that samples of these forms that have been kept in our storage for far more than 10 years showed virtually no change. Another feature of the three D-mannitol polymorphs is their propensity to crystallize concomitantly. This makes the production of the phase pure polymorphs rather challenging. Commercial products mostly consist of mod. $I^\circ$, but they may also be composed of mainly mod. III or mixtures of these two forms. Lian Yu and coworkers (Yu, 2007 and refs. therein) gave an insight into the complex crystallization behavior of the D-mannitol polymorphs. The same authors (Yu et al., 1999) detected also the hemihydrate of D-mannitol, which occurs in freeze-drying processes. Due to its low stability the hydrate transforms on warming to a mixture of the mods. $I^\circ$ and III. The crystal structure of the hemihydrate and details on the solid-state behavior were reported by Nunes et al. (2004). Fronczek et al. (2003) discussed and compared the structural features of the three D-mannitol polymorphs.

The fact that D-mannitol occurs mostly as a mixture of stable polymorphs naturally encouraged many groups to develop quantification methods. Table 1 summarizes quantitative or semi-quantitative studies of D-mannitol forms published in the literature. Campbell Roberts et al. applied univariate methods to analyze binary mixtures of the mods. $I^\circ$ and III with either FT-Raman spectroscopy (Campbell Roberts et al., 2002a) or X-ray powder diffractometry (Campbell Roberts et al., 2002b). A multivariate calibration method using PLS and Raman spectroscopy was applied by Auer et al. (2003) to analyze mixtures of these two forms ($I^\circ$ and III). Raman spectroscopy was also used by Xie et al. (2008), who developed a multivariate curve resolution (MCR) method for the quantitative analysis of the three D-mannitol polymorphs, the hemihydrate and the amorphous form in lyophilized protein formulations. In this study the second derivative of the spectra was used, and the performance of the calibration model was evaluated with binary mixtures of the mods. $I^\circ$ and III and the prediction residue. Not based on standard reference mixtures are the studies reported by Vehring (2005) and Beattie et al. (2007). Vehring (2005) identified and measured the mass fractions of D-mannitol in spray-dried mixtures

**Table 1**
List of reports on (semi-)quantitative studies of D-mannitol solid-state forms.

| Author | Year | Technique | Quantification method | Quantified forms | Aims/results |
|---|---|---|---|---|---|
| Campbell Roberts et al. | 2001a | Raman | UV (peak intensity ratios) | I° and III | • Quantitation of mod. I° in binary mannitol mixtures<br>• Levels down to ∼2% can be quantified |
| Campbell Roberts et al. | 2001b | PXRD | UV (peak area) | I° and III | • Quantification of mod. III in binary mannitol mixtures<br>• LOD: 1%, LOQ: 3% |
| Auer et al. | 2003 | Raman | MV (PLS) | I° and III | • Quantitative analysis of mannitol mods. I° and III<br>• S.E. of CV: 1.039% |
| Vehring | 2005 | Raman | UV (deconvolution) | Calictonin/mannitol (I°, II, III, amorph.) | • Identification/mass fraction measurement of mannitol (I°, II, III, amorphous) in spray-dried salmon calcitonin mixtures<br>• Absolute error of measured mass fraction calcitonin/mannitol: <3% (0–100% range)<br>• Relative error of measured mass fraction calcitonin/mannitol: 10% (5–10% range). |
| Beattie et al. | 2007 | Raman | MV (LCE) | (I°, II, III, hydrate, amorph.) | • Semi-quantitative determination of all mannitol forms in frozen system<br>• Determination of the predominant form |
| Xie et al. | 2008 | Raman | MV (MCR) | I°, II, III, hydrate, amorph. | • Characterize/quantify the five known mannitol forms in lyophilized protein formulations<br>• Quantitation based on binary mixtures of mods. I° and III combined with evaluation of the prediction residue<br>• LOQ: 5% |
| This study | | Raman | MV (PLS and ANNs) | I°, II and III | • Simultaneous quantification of ternary polymorphic powder mixtures<br>• LOD: ∼1.5% (I°) and <1% (II and III)<br>• RMSECV (PLS): <0.5% for all three forms<br>• RMSEP (PLS): <0.5% for all three forms |

UV: univariate; MV: multivariate; PLS: partial least-squares regression; LCE: linear combination of element analysis; MCR: multivariate curve resolution; ANNs: artificial neural networks; LOD: limit of detection; LOQ: limit of quantitation; S.E.: standard error; CV: cross-validation; RMSECV/P: root-mean-squared error of cross-validation/prediction.

with salmon calcitonin. The D-mannitol fractions consisted of mixtures of the three polymorphs and the amorphous phase. Beattie et al. (2007) published a semi-quantitative calibration method, based on a linear combination of element (LCE) analysis and involving all five solid-state forms, along with the crystallization behavior of frozen D-mannitol samples. However, there has been no study involving all three polymorphs in a calibration set.

Therefore, the aim of the present study was to develop a FT-Raman method for the simultaneous quantification of ternary powder mixtures of D-mannitol polymorphs on the basis of standard reference mixtures. Furthermore, the predictability power of different multivariate calibration algorithms, namely PLS and ANNs should be compared. Since the hemihydrate (Yu et al., 1999) is not stable under common storage conditions, only the three D-mannitol polymorphs were considered in this study. It was anticipated that the preparation of mixtures is simple, since the modifications do not undergo phase changes upon grinding and mixing.

## 2. Materials and methods

### 2.1. Preparation of pure D-mannitol polymorphs

The commercial D-mannitol product (Apoka ACM, Austria), consisting of pure mod. I°, was used for the quantification. Mod. II was prepared by dissolving 50 g D-mannitol in 450 g 70% EtOH under heating. The hot, but not saturated solution was slowly cooled to room temperature. After 8 h, needles of pure mod. II precipitated, which were filtered off and dried at 40 °C. Mod. III was produced according to the instructions of Burger et al. (2000) by rapidly cooling a hot saturated solution in water to 0 °C in an ice bath.

### 2.2. Preparation of polymorphic mixtures

Three groups, each containing eleven ternary mixtures, were prepared by geometrically mixing a starting mixture (36:32:32) with either a pure form or a second mixture, resulting in following mass ratios—36:32:32, 52:24:24, 68:16:16, 76:12:12, 84:8:8, 88:6:6, 92:4:4, 96:2:2, 97:1.5:1.5, 98:1:1 and 99:0.5:0.5 for each polymorph. Including the three pure polymorphs (mods. I°, II and III), a set of 36 different samples was obtained.

Before mixing the calibration and validation sets the pure polymorphs were milled in a cross-beater mill (Primax, Troisdorf, Germany) and particles >300 μm were removed by sieving the powder. The homogenization procedure showed that especially mod. II is very electrostatic, which complicated the preparation of the mixtures. Equal amounts (w/w) of the polymorphs or polymorph mixtures were weighed and mixed in a mortar using a pestle or a spatula, until a homogenous blend was obtained. The homogeneity of the mixtures was controlled with FT-Raman spectroscopy, in 5 min intervals until identical Raman spectra were obtained (minimum mixing time: 15–20 min). For every mixture an amount of at least 1000 mg was prepared.

### 2.3. X-ray powder diffraction (PXRD)

The diffractograms were obtained with an X'Pert PRO diffractometer (PANalytical, Almelo, The Netherlands) with the following setup: theta/theta coupled goniometer, transmission geometry, programmable XYZ stage with well plate holder, Cu-K$\alpha_{1,2}$ radiation (wavelength 0.15419 nm), focussing mirror, 0.5° divergence slit and 0.02° Soller slit collimator on the incident beam side, 2 mm antiscattering slit and 0.02° Soller slit collimator on the diffracted beam side, solid-state PIXcel detector. The patterns were recorded at a tube voltage of 40 kV, tube current of 40 mA, applying a step size of $2\theta = 0.013°$ with 40 s per step in the $2\theta$ range between 2° and 40°.

### 2.4. FT-Raman spectroscopy

FT-Raman spectra were recorded with a Bruker RFS 100 FT-Raman spectrometer, equipped with a diode pumped Nd:YAG laser (1064 nm) as the excitation source, and a liquid nitrogen cooled, high sensitivity Ge detector (Bruker Optik GmbH, Ettlingen, Germany). A few milligrams of the samples were packed into aluminum sample cups. For each spectrum 128 scans were performed at a resolution of 4 cm$^{-1}$ over the range 30–4000 cm$^{-1}$. A Blackman-Harris B4 term was used as apodization function. Raman measurements of the mixtures were repeated five times to exclude the effect of imperfect mixing and guarantee a representative sampling of the mixtures.

### 2.5. Data preprocessing

In a first step the whole spectral range (4000–30 cm$^{-1}$) was used for the multivariate calibration. The raw spectra were normalized to the strongest peak (886–874 cm$^{-1}$) using the Bruker Opus software (version 5.5), in order to remove inhomogeneities due to sample preparation. Different preprocessing algorithms, SNV (Barnes et al., 1989), MSC (Geladi et al., 1985), first and second derivative, and OSC (Wold et al., 1998) were either applied alone or in combination. For all preprocessing steps, except the calculation of the derivatives (calculated with the Bruker Opus software), the Simca-P11.0 software (Umetrics AB) was used. To evaluate the best models, cross-validation by the leave-one-out method was applied.

OSC is a PLS based method and removes from the X-data variation which is not related to the concentration (in general output data) (Wold et al., 1998). For each fitted model two OSC components were calculated. To minimize the observation risk of over-fitting two spectral regions were excluded from the dataset. The first of these concerns a part of the region of the symmetric and antisymmetric stretch vibrations of the R–CH$_2$–R′ and CH(R)$_3$ groups (2947–2939 cm$^{-1}$), the second the broad range from 951 to 777 cm$^{-1}$, which includes, among others, the in-phase C–C–O stretch vibrations of the OH groups. Furthermore, the region starting at 89 cm$^{-1}$ (primary beam of the laser) to 30 cm$^{-1}$ was excluded. The final dataset comprised 1932 X variables (wavenumbers), and three Y variables (the concentration of each polymorph). This dataset was then split into a training/calibration and test/validation subset, containing 24 and 12 samples respectively, applying the Kennard–Stone design (Kennard and Stone, 1969). The validation subset contained following mixtures: I36, I76, I99, I100, II36, II84, II96, II98, III52, III88, III97, III99 (concentration of the dominating polymorph given), also illustrated in the triangular diagram in Fig. 1.

### 2.6. Model fitting and validation

PLS regression was performed with the Simca-P11.0 software. The optimal number of PLS components was automatically determined by the leave-one-out cross-validation (CV) procedure (Wold, 1978). In each calibration model a single model was developed for the three modifications.

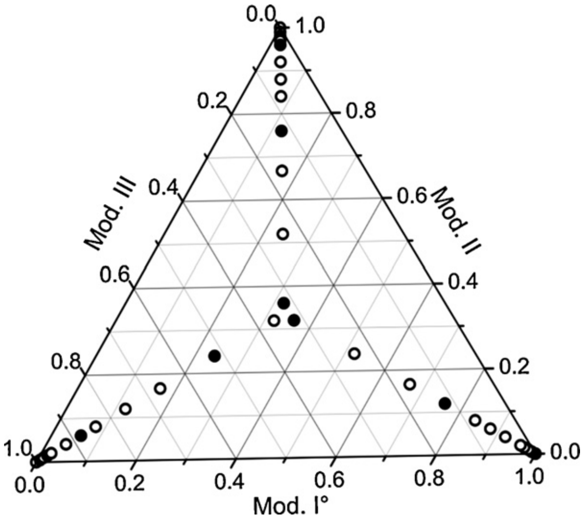

**Fig. 1.** Triangular diagram presenting the partitioning of input data into training/calibration (empty symbols) and test/prediction subset (filled symbols).

Furthermore, artificial neural networks (ANNs) of the feed-forward architecture were fitted to the training subset (1932 input units, 1000 hidden units and three output units). Each input unit represented the preprocessed Raman unit value (wavenumber) and the output units the concentration. In all cases the logistic sigmoid transfer function was used for the input to hidden layer connections, and the linear function was used for the hidden layer to output connections. The networks were trained using the scaled conjugate gradient method. Training was repeated five times and the averaged concentrations were used. The netlab toolbox (Nabney, 2001) for matlab was employed for the calculation of the ANNs. All calculations were performed on a PC running MS Windows XP, equipped with an AMD Turion64 2 GHz processor and 2 GB RAM.

The predictive performance was assessed on the basis of the root-mean-squared error of cross-validation (RMSECV), calibration (RMSEC) and prediction (RMSEP), calculated by Eq. (1).

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

The limit of detection (LOD) of the polymorphs in the ternary mixtures was calculated from the standard deviation (SD) of samples containing the analyte in the range of the LOD (0–2%) and the slope of the linear regression of the actual versus predicted concentration plot, using LOD = 3.3 SD/slope (ICH, 1996).

## 3. Results and discussion

### 3.1. Phase purity of the polymorphs

The phase purity of the starting polymorphs was confirmed with various analytical techniques including powder X-ray diffraction, infrared and Raman spectroscopy as well as differential scanning calorimetry (DSC) and hot-stage microscopy. The overlay of the experimental and calculated powder X-ray diffractograms of the three D-mannitol modifications in Fig. 2 demonstrates the phase identity and purity, as each modification shows characteristic and unique peak positions. The theoretical PXRD patterns, calculated from the single crystal structure data (Berman et al., 1968; Fronczek et al., 2003) using the PowderCell v. 2.4 software are in good agree-

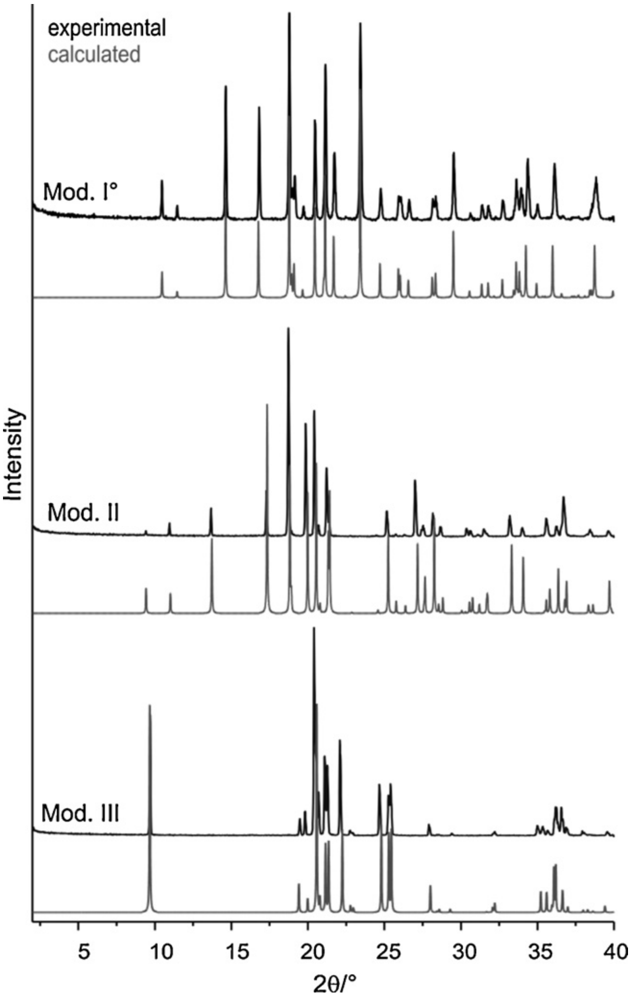

**Fig. 2.** Overlay of experimental versus calculated PXRD patterns (PowderCell v. 2.4) of the three D-mannitol modifications. Experimental patterns were recorded at room temperature.

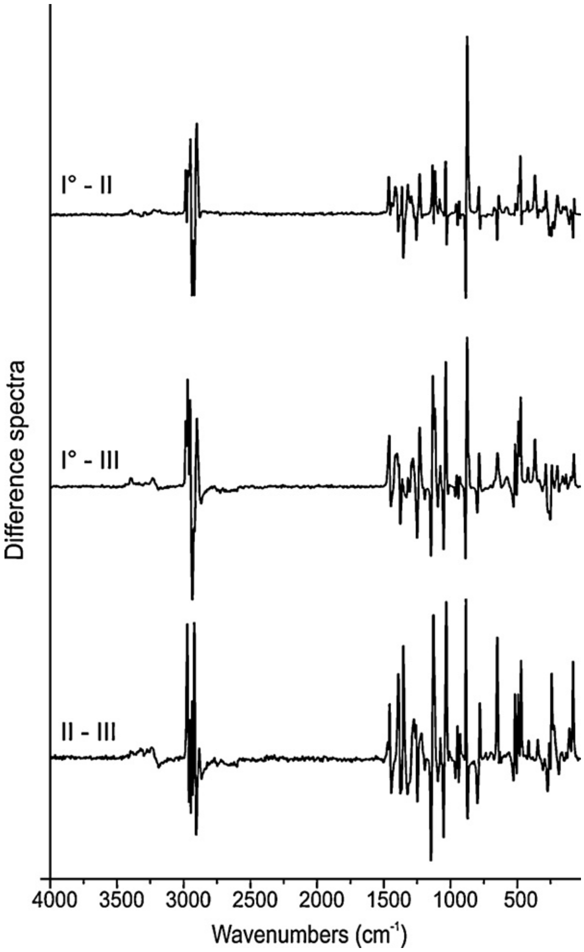

**Fig. 4.** Difference spectra between the mods. I°–II, I°–III and II–III.

ment with the respective experimental patterns. Furthermore, a Rietveld refinement was performed, confirming a high phase purity of the starting materials.

Even though the D-mannitol samples were ground with mortar and pestle prior to the measurements, one can identify intensity differences between the experimental and calculated powder X-ray diffraction patterns that arise from preferred orientation (Jenkins and Snyder, 1996), which is hard to avoid for the acicular mannitol crystals. This problem has been previously described by other authors for D-mannitol (Campbell Roberts et al., 2002b) and makes spectroscopic methods such as Raman spectroscopy more advantageous compared to PXRD.

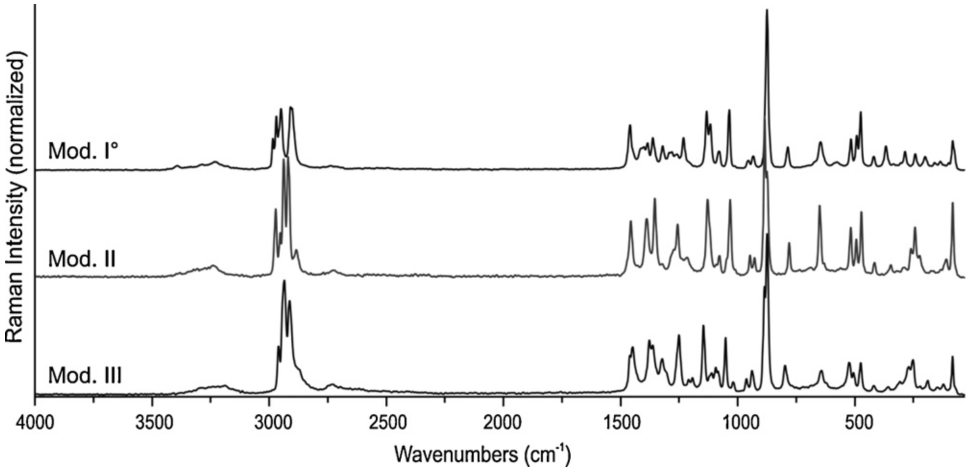

**Fig. 3.** FT-Raman spectra of D-mannitol mods. I°, II and III used in this study.

**Table 2**
Actual contents (%) of the three D-mannitol polymorphs and values predicted by the multivariate methods. The spectra were preprocessed by calculating the first derivative and Orthogonal Signal Correction (OSC).

| Mod. I° | | | Mod. II | | | Mod. III | | |
|---|---|---|---|---|---|---|---|---|
| Actual content (%) | Predicted content (%) by | | Actual content (%) | Predicted content (%) by | | Actual content (%) | Predicted content (%) by | |
| | PLS | ANN | | PLS | ANN | | PLS | ANN |
| 0.5 | 0.7 | 0.9 | 0 | 0.0 | 0.6 | 0 | 0.3 | 0.6 |
| 1 | 1.1 | 1.4 | 0.5 | 0.9 | 1.0 | 0.5 | 0.7 | 1.0 |
| 1.5 | 1.7 | 1.5 | 0.5 | 0.5 | 0.9 | 1 | 1.8 | 1.5 |
| 2 | 2.6 | 2.6 | 1.5 | 1.3 | 1.4 | 2 | 2.0 | 1.9 |
| 6 | 5.4 | 5.6 | 6 | 6.2 | 5.7 | 8 | 8.1 | 8.1 |
| 8 | 8.0 | 8.1 | 12 | 12.0 | 12.1 | 12 | 11.4 | 11.9 |
| 24 | 23.4 | 24.2 | 24 | 24.2 | 23.9 | 32 | 32.1 | 32.1 |
| 32 | 31.5 | 32.0 | 32 | 31.9 | 31.9 | 32 | 31.7 | 32.2 |
| 36 | 36.4 | 36.2 | 36 | 36.3 | 35.9 | 52 | 52.4 | 51.9 |
| 76 | 76.6 | 75.9 | 84 | 84.0 | 83.8 | 88 | 88.4 | 88.7 |
| 99 | 98.5 | 98.0 | 96 | 95.3 | 96.1 | 97 | 97.0 | 97.1 |
| 100 | 99.7 | 98.7 | 98 | 97.1 | 97.1 | 99 | 98.9 | 98.3 |

ANN: artificial neural networks; PLS: partial least-square regression.

### 3.2. Raman spectroscopy

The FT-Raman spectra of the three pure D-mannitol polymorphs are displayed in Fig. 3, and the difference spectra (mods. I°–II, I°–III and II–III) are shown in Fig. 4. The spectra are in good agreement with those published by Burger et al. (2000). The three modifications show clear differences over almost the entire spectral range. The most significant differences occur within the symmetric and antisymmetric stretch vibrations of the R–$CH_2$–R′ and $CH(R)_3$ groups (3000–2850 $cm^{-1}$), the out-of-phase C–C–O stretch (1150–1000 $cm^{-1}$) and in-phase C–C–O stretch (900–800 $cm^{-1}$) vibrations of the primary and secondary alcohol groups. However, for a simultaneous quantification of all three modifications there are no suitable, non-overlapping (single) peaks with a high intensity, which would be the prerequisite for a simple univariate quantification method based on peak intensities or peak area. The univariate Raman quantification method developed by Campbell Roberts et al. (2002a) for binary mixtures of mods. I° and III would not be applicable for the ternary mixture, because the bands used for the analysis of mod. I° (1037 $cm^{-1}$) overlap with bands of mod. II (1032 $cm^{-1}$). Therefore multivariate models represent the most promising alternative for a successful quantitative analysis of the ternary polymorph mixtures with Raman spectroscopy.

### 3.3. Multivariate quantification: partial least-squares (PLS) regression and artificial neural networks (ANNs)

Two different algorithms, PLS and ANNs, were used in the present study. Table 2 lists the actual contents (%) of the mannitol modifications in the prediction subset and the values that have been predicted with the two multivariate models (PLS and ANNs) using preprocessed Raman data (first derivative and orthogonal signal correction), and Fig. 5 illustrates the results. The corresponding RMSEP (root-mean-squared error of prediction, %) values together with the linear regression parameters of observed versus predicted concentrations (coefficients of determination, slopes and intercepts) are listed in Table 3.

#### 3.3.1. Partial least-squares regression

The most important step in the development of multivariate calibration models is the choice of a sufficient number of well selected observations and predictor(s), the choice of the correct regression technique (or alternative technique), a suitable pretreatment of the data, and the careful validation of the model (Forina et al., 2007). Another important point that is particularly relevant for PLS models, is the number of latent variables. The leave-one-out cross-validation procedure was used for calculating the number of principal components (PCs). Table 4 also lists the number of PLS components and the prediction errors. Except for the model where no preprocessing was applied, only 2 PCs were found to be significant.

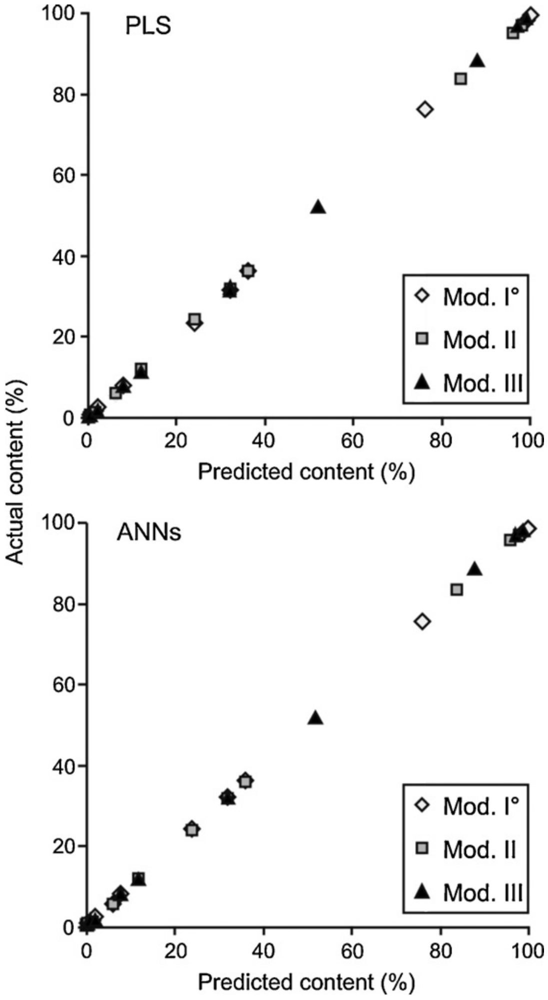

**Fig. 5.** Predicted versus actual content of D-mannitol mods. I°, II and III in the test data set, calculated for partial least-squares (PLS) regression and artificial neural networks (ANNs).

**Table 3**
Root-Mean-Squared Error of Prediction (RMSEP %) with the Linear Regression Parameters of Observed versus Predicted Concentrations (coefficient of Determination, $R^2$, Slope and Intercept). Spectra were preprocessed by calculating the first Derivative and Orthogonal Signal Correction (OSC).

| Algorithm | RMSEP (%) | $R^2$ | Slope | Intercept |
|---|---|---|---|---|
| Mod. I° | | | | |
| PLS | 0.440 | 0.9999 | 1.002 | 0.041 |
| ANN | 0.510 (0.027) | 0.9999 (<0.0001) | 1.011 (0.020) | −0.202 (0.020) |
| Mod. II | | | | |
| PLS | 0.376 | 0.9999 | 1.007 | −0.158 |
| ANN | 0.386 (0.008) | 0.9999 (<0.0001) | 1.006 (<0.001) | −0.105 (0.019) |
| Mod. III | | | | |
| PLS | 0.360 | 0.9999 | 1.001 | −0.127 |
| ANN | 0.411 (0.008) | 0.9999 (<0.0001) | 1.003 (<0.001) | −0.269 (0.010) |

ANN: artificial neural networks; PLS: partial least-square regression.

**Table 4**
Comparison of the Root-Mean-Squared Error of Cross-Validation (RMSECV, %), Calibration (RMSEC, %) and Prediction (RMSEP, %) of the differently preprocessed PLS and ANN models. For the PLS models the number of principal components (PCs) is stated.

| Method | PLS | | | | ANN |
|---|---|---|---|---|---|
| Preprocessing | PCs | RMSECV (%) | RMSEC (%) | RMSEP (%) | RMSEP (%) |
| Mod. I° | | | | | |
| Without | 4 | 2.817 | 3.068 | 4.187 | 1.558 (0.055) |
| First der. | 2 | 3.687 | 4.063 | 6.830 | 1.148 (0.023) |
| OSC | 2 | 1.270 | 1.524 | 0.875 | 0.950 (0.039) |
| First der. + OSC | 2 | 0.545 | 0.426 | 0.440 | 0.510 (0.027) |
| Mod. II | | | | | |
| Without | 4 | 3.685 | 4.253 | 3.367 | 0.892 (0.015) |
| First der. | 2 | 3.806 | 4.282 | 8.466 | 1.222 (0.024) |
| OSC | 2 | 0.855 | 1.051 | 0.737 | 0.804 (0.018) |
| First der. + OSC | 2 | 0.269 | 0.249 | 0.376 | 0.386 (0.008) |
| Mod. III | | | | | |
| Without | 4 | 1.902 | 2.014 | 3.520 | 1.246 (0.023) |
| First der. | 2 | 2.614 | 3.020 | 5.009 | 1.087 (0.016) |
| OSC | 2 | 2.076 | 2.538 | 1.507 | 0.996 (0.030) |
| First der. + OSC | 2 | 0.426 | 0.452 | 0.360 | 0.411 (0.008) |

ANN: artificial neural networks; PLS: partial least-square regression; first der.: first derivative; OSC: orthogonal signal correction.

To find out which kind of preprocessing gives the best results, different spectral pretreatments were applied and compared. The RMSECV, RMSEC and RMSEP for the PLS analysis and RMSEP for the ANNs models are given in Table 4. The numbers show that the best predictive performance was achieved by calculating the first derivative of the spectra OSC (Wold et al., 1998). Without spectral pretreatment the prediction errors (RMSEP) for the PLS models ranged from 3.4 to 4.2%. Using only the first derivative resulted in even higher prediction errors than with the unprocessed data, whereas the OSC preprocessing lowered the RMSEPs of the original data by 0.7–1.5%. However, the data show that the most effective preprocessing is to apply OSC to the first derivative. This indicates that OSC is a very critical step to enhance the predictability power of the model.

To avoid over-fitting the performance of the models was validated using an external test subset. A comparison of the RMSECV, RMSEC and RMSEP values (see Table 4) showed that over-fitting is not a problem for the best model, since the RMSEs (cross-validation, calibration and prediction) are of the same magnitude. The LODs for the best model were found to be ~1.5% for mod. I° and <1% for mods. II and III respectively.

### 3.3.2. *Artificial neural networks*

In order to compare the predictive power of two algorithms (PLS and ANNs) the data set was also used as an input for ANNs. The ANN models using non-preprocessed data or data preprocessed with one procedure (first derivative or OSC) showed considerable prediction errors of 0.8–1.6% (see Table 4). Unlike the PLS model without pretreatment the first derivative of the spectra alone did not result in higher prediction errors. Similar to the PLS model, the lowest prediction errors were obtained by a combination of the two preprocessing techniques before calculating the ANNs. However, the impact of OSC on the performance of the ANN models is lower than for the PLS models.

The PLS and ANN models allow a satisfactory prediction of low as well as high ratios of each polymorph in the ternary mixtures (see Table 2). The RMSEP values (Tables 3 and 4) obtained for the best model (first derivative plus OSC) indicate that the two multivariate methods yield comparable results. Since the less complex PLS model gives even slightly lower error values it is obvious that this method is more favorable.

## 4. Conclusions

The study demonstrated that PLS and ANNs of Raman spectroscopic data provide sensitive methods for the simultaneous quantification of ternary polymorphic D-mannitol powder mixtures, with RMSEPs ~0.5% and below for each polymorph and quantification method.

Spectral preprocessing by calculating the first derivative followed by OSC improved the predictive performance of the models significantly. With the PLS model a better enhancement was obtained by applying OSC, which is most likely due to the fact that OSC is a PLS based preprocessing technique. Furthermore, the comparison between PLS and ANNs showed that both algorithms were capable of correctly predicting low as well as high concentrations of each D-mannitol polymorph in the presence of the other two modifications. Apart from its slightly better performance compared to the ANNs PLS is also the simpler and faster (computation time) method. Moreover, a big advantage is that, unlike ANNs, there is no dependence of the convergence on the initial weights. Thus, several training and testing cycles are necessary if ANNs are used.

This study also demonstrated that the predictability of a data set depends not only on the selection of the right algorithms, but also on the appropriate preprocessing, which needs to be evaluated for the individual experimental method and also the algorithm itself. ANNs are the quantification methods of choice if no preprocessing is carried out. Our work shows that PLS becomes the best model after the observations (variables) have been subjected to suitable pretreatment. Since it is wise to choose not only the best but also the simplest method, we regard PLS as the first choice for quantifying polymorphic D-mannitol powder mixtures. However, if a system becomes more complex, e.g. when a higher number of crystal forms needs to be quantified, this preference may change.

## Acknowledgment

D.E.B. acknowledges financial support from a "Forschungsstipendium an österreichische Graduierte 2008" of the University of Innsbruck.

## References

Aaltonen, J., Gordon, K.C., Strachnan, C.J., Rades, T., 2008. Perspectives in the use of spectroscopy to characterise pharmaceutical solids. Int. J. Pharm. 364, 159–169.

Aboul-Enein, H.Y., Bunaciu, A.A., Fleschin, S., 2002. Analysis of mebendazole polymorphs by Fourier transform IR spectrometry using chemometric methods. Biopolymers 67, 56–60.

Anderton, C.L., 2003. Vibrational spectroscopy in pharmaceutical analysis. In: Lee, D.C., Webb, M. (Eds.), Pharmaceutical Analysis. Blackwell Publishing, Oxfordshire, pp. 203–239.

Auer, M.E., Griesser, U.J., Sawatzki, J., 2003. Qualitative and quantitative study of polymorphic forms in drug formulations by near infrared FT-Raman spectroscopy. J. Mol. Struct. 661–662, 307–317.

Barnes, R.J., Dhanoa, M.S., Lister, S.J., 1989. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. Appl. Spectrosc. 43, 772–777.

Beattie, J.R., Barrett, L.J., Malone, J.F., McGarvey, J.J., Nieuwenhuyzen, M., Kett, V.L., 2007. Investigation into the subambient behavior of aqueous mannitol solutions using temperature-controlled Raman microscopy. Eur. J. Pharm. Biopharm. 67, 569–578.

Berman, H.M., Jeffrey, G.A., Rosenstein, R.D., 1968. Crystal structures of the α′- and β- forms of D-mannitol. Acta Crystallogr. A24, 442–449.

Burger, A., Henck, J.-O., Hetz, S., Rollinger, J.M., Weissnicht, A.A., Stottner, H., 2000. Energy/temperature diagram and compression behavior of the polymorphs of D-mannitol. J. Pharm. Sci. 89, 457–468.

Campbell Roberts, S.N., Williams, A.C., Grimsey, I.M., Booth, S.W., 2002a. Quantitative analysis of mannitol polymorphs. FT-Raman spectroscopy. J. Pharm. Biomed. Anal. 28, 1135–1147.

Campbell Roberts, S.N., Williams, A.C., Grimsey, I.M., Booth, S.W., 2002b. Quantitative analysis of mannitol polymorphs. X-ray powder diffractometry-exploring preferred orientation effects. J. Pharm. Biomed. Anal. 28, 1149–1159.

Dandeu, A., Humbert, B., Carteret, C., Muhr, H., Plasari, E., Bossoutrot, J.M., 2006. Raman spectroscopy—a powerful tool for the quantitative determination of the composition of polymorph mixtures: application to $CaCO_3$ polymorph mixtures. Eng. Technol. 29, 221–225.

Despagne, F., Massart, D.L., 1998. Neural networks in multivariate calibration. Analyst 123, 157–178.

Eriksson, L., Johansson, E., Kettaneh-Wold, N., Wold, S., 2001. Multi- and Megavariate Data Analysis (Principles and Applications). Umetrics Academy, Sweden.

Forina, M., Lanteri, S., Casale, M., 2007. Multivariate calibration. J. Chromatogr. A 1158, 61–93.

Fronczek, F.R., Kamel, H.N., Slattery, M., 2003. Three polymorphs (α, β, and δ) of D-mannitol at 100 K. Acta Crystallogr. C59, o567–o570.

Geladi, P., MacDougall, D., Martens, H., 1985. Linearization and scatter-correction for near-infrared relectance spectra of meat. Appl. Spectrosc. 39, 491–500.

Heinz, A., Savolainen, M., Rades, T., Strachan, C.J., 2007. Quantifying ternary mixtures of different solid-state forms of indomethacin by Raman and near-infrared spectroscopy. Eur. J. Pharm. Sci. 3, 182–192.

ICH Harmonized guideline, 1996. Q2B: Note for Guidance on Validation of Analytical Procedures: Methodology.

Jenkins, R., Snyder, R., 1996. Introduction to X-ray Powder Diffractometry. Wiley, New York.

Kachrimanis, K., Braun, D.E., Griesser, U.J., 2007. Quantitative analysis of paracetamol polymorphs in powder mixtures by FT-Raman spectroscopy and PLS regression. J. Pharm. Biomed. Anal. 43, 407–412.

Kennard, R.W., Stone, L.A., 1969. Computer aided design of experiments. Technometrics 11, 137–148.

Kibbe, A.H., 2000. Handbook of Pharmaceutical Excipients, 3rd ed. Pharmaceutical Press, London, pp. 324–328.

Kipouros, K., Kachrimanis, K., Nikolakakis, I., Malamataris, S., 2005. Quantitative analysis of less soluble form IV in commercial carbamazepine (form III) by diffuse reflectance fourier transform spectroscopy (DRIFTS) and lazy learning algorithm. Anal. Chim. Acta 550, 191–198.

Kipouros, K., Kachrimanis, K., Nikolakakis, I., Tserki, V., Malamataris, S., 2006. Simultaneous quantification of carbamazepine crystal forms in ternary mixtures (I, III, and IV) by diffuse reflectance FTIR spectroscopy (DRIFTS) and multivariate calibration. J. Pharm. Sci. 95, 2419–2431.

Lavine, B., Workman, J., 2006. Chemometrics. Anal. Chem. 78, 4137–4145.

Martens, H., Naes, T., 1991. Multivariate Calibration. Wiley VCH, Chichester.

Mazurek, S., Szostak, R., 2008. Quantitative determination of diclofenac sodium in solid dosage forms by FT-Raman spectroscopy. J. Pharm. Biomed. Anal. 48, 814–821.

McGoverin, C.M., Ho, L.C.H., Zeitler, J.A., Strachan, C.J., Gordon, K.C., Rades, T., 2006. Quantification of binary polymorphic mixtures of ranitidine hydrochloride using NIR spectroscopy. Vib. Spectrosc. 41, 225–231.

Nabney, I.T., 2001. Netlab. Algorithms for Pattern Recognition (Advances in Pattern Recognition). Springer, Berlin.

Nunes, C., Suryanarayanan, R., Botez, C.E., Stephens, P.W., 2004. Characterization and crystal structure of D-mannitol hemihydrate. J. Pharm. Sci. 93, 2800–2809.

Poellaenen, K., Haekkinen, A., Huhtanen, M., Reinikainen, S.-P., Karjalainen, M., Rantanen, J., Louhi-Kultanen, M., Nystroem, L., 2005. DRIFT-IR for quantitative characterization of polymorphic composition of sulfathiazole. Anal. Chim. Acta 544, 108–117.

Pratiwi, D., Fawcett, J.P., Gordon, K.C., Rades, T., 2002. Quantitative analysis of polymorphic mixtures of ranitidine hydrochloride by Raman spectroscopy and principal components analysis. Eur. J. Pharm. Biopharm. 54, 337–341.

Strachan, C.J., Pratiwi, D., Gordon, K.C., Rades, T., 2004. Quantitative analysis of polymorphic mixtures of carbamazepine by Raman spectroscopy and principal components analysis. J. Raman Spectrosc. 35, 347–352.

Suda, M., Takayama, K., Otsuka, M., 2008. An accurate quantitative analysis of polymorphic content by chemometric X-ray powder diffraction. Anal. Sci. 24, 451–457.

Vehring, R., 2005. Red-excitation dispersive Raman spectroscopy is a suitable technique for solid-state analysis of respirable pharmaceutical powders. Appl. Spectrosc. 59, 286–292.

Wold, S., Antti, H., Lindgren, F., Ohman, J., 1998. Orthogonal signal correction of near-infrared spectra. Chemom. Intell. Lab. Syst. 44, 175–185.

Wold, S., 1978. Cross-validatory Estimation of the Number of Components in Factor and Principal Component Models. Technometrics 20, 397–405.

Xie, Y., Cao, W., Krishnan, S., Lin, H., Cauchon, N., 2008. Characterization of mannitol polymorphic forms in lyophilized protein formulations using a multivariate curve resolution (MCR)-based Raman spectroscopic method. Pharm. Res. 25, 2292–2301.

Yu, L., Milton, N., Groleau, E.G., Mishra, D.S., Vansickle, R.E., 1999. Existence of a mannitol hydrate during freeze-drying and practical implications. J. Pharm. Sci. 88, 196–198.

Yu, L., 2007. Survival of the fittest polymorph: how fast nucleater can lose to fast grower. CrystEngComm 9, 847–851.