Pergamon
0731-7085(95)01304-0

Journal of Pharmaceutical & Biomedical Analysis, Vol. 13, No. 4/5, pp. 415–417, 1995

0731-7085/95 $9.50 + 0.00

Quantitative determination of Sulfasalazine by near-infrared spectroscopy and multivariate analysis in reflectance mode with a fibre-optic probe*

SVEN P. JACOBSSON,† MARGARETA CARLSSON, ULF JÖNSSON and GUN NILSSON

Pharmacia AB, S-751 82 Uppsala, Sweden

Keywords: *Sulfasalazine; near-infrared spectroscopy; multivariate analysis; partial least squares.*

Introduction

Near-infrared spectroscopy (NIR) as a rapid and general technique for quantitative analysis has received considerable attention for some time [1, 2]. In particular the use of NIR in combination with multivariate analysis methods such as, for example, principal component analysis and partial least squares, has been shown to be a feasible approach for classification and calibration. The present study examines the use of NIR, in reflectance mode, and partial least squares (PLS) regression for the quantitative determination of sulfasalazine, a drug used for treatment of intestinal inflammations. The determinations have been performed on batches of technical quality. Various strategies for data preprocessing as derivatives, multiplicative scatter correction and spectral transforms have been examined with the objective to establish a method based on NIR/PLS with good prediction capabilities, which could complement or replace the existing method based on liquid chromatography (LC).

Experimental

A collection of 44 batches of Sulfasalazine of technical quality has been included in this study. The concentration range of Sulfasalazine was 81.9–90.4%, with a mean of 86.7%, thus containing batches of standard technical quality as well as outliers (i.e. concentrations too low). The concentrations of Sulfasalazine were assayed by reversed-phase liquid chromatography.

NIR spectra in reflectance mode, in triplicate, were collected in the wavelength range 4545–10 000 cm^{-1} by use of the InfraProver instrument (Bran and Luebbe GmbH, Germany). The spectra were collected remotely by use of fibre optics (2 m) at a resolution of ± 6 cm^{-1}, with the probe-head slightly immersed in the drug powder.

Multivariate analysis was performed by use of PLS regression utilizing the whole spectral data range (459 variables/spectrum). To assess the prediction capabilities of the NIR/PLS method training and testing were performed on different data sets. The original data set was divided into a training set, consisting of 33 batches chosen at random, while the remaining 11 batches constituted the test (prediction) set. The optimal model, for the training set, after spectral transforms and preprocessing, was established by the use of cross-validation using the 'leave-one-out at a time' procedure.

All data anlysis was performed by using Matlab 4.1 (The MathWorks, Inc., Natick, MA, USA) using the extended functionality of the Chemometric Toolbox 2.0.

Results and Discussion

It is generally considered that for NIR reflectance data, especially in combination with multivariate analysis methods, the larger

*Presented at the Fifth International Symposium on Pharmaceutical and Biomedical Analysis, Stockholm, Sweden, September 1994.

†Author to whom correspondence should be addressed.

part of the spectral variance is related to optical scatter [2]. This scattering arises predominantly as an effect of differences in sample morphology, but also specular reflectance plays a role. To overcome this scattering effect various preprocessing strategies have been devised. A general approach for baseline correction in multivariate analysis is to use the second derivative of the spectra [3]. Another approach is to use the multiplicative scatter correction (MSC) method [2]. In MSC the correction parameters (b) for each spectrum (y) are calculated by the least squares algorithm, cf equation (1).

$$[1\ y]\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = [y_{\text{Referens}}] + [y_{\text{Residual}}], \quad (1)$$

where y_{Referens} is the average spectral reading. The deviation from the linear model, y_{Residual}, is then considered to carry the chemical information.

By analogy with MSC, piece-wise MSC (PMSC) has been introduced [4]. In contrast to MSC which generates the correction parameters from the complete spectrum, PMSC generates the correction parameters from a proportion of the spectrum by moving a window of fixed size over the entire spectrum. This method has been shown to yield lower prediction errors by comparison with MSC [4].

The prediction errors of the test (prediction) set were calculated as the root mean square error of prediction (RMSEP) defined as

$$\text{RMSEP} = \left[I_p^{-1} \sum_{i=1}^{I_p} (y_i - \hat{y}_i)^2\right]^{1/2}, \quad (2)$$

where $I_p[1,2,3,\ldots,I_p]$ denotes the total number of spectra in the test set used, y_i the result obtained by the LC analysis reference method and $\hat{y}_i$ denotes the NIR predicted value of the spectrum for sample i. It should be noted that RMSEP is calculated using the total number of triplicate spectra included in the test set.

The results for the various data pretreatments and spectral transforms used in this study are summarized in Table 1.

From the results given in Table 1 it can be concluded that NIR spectroscopy in reflectance mode, in this study, correlates to a satisfactory degree with that of liquid chromatography, provided that multiplicative scatter correction is applied as a data preprocessing method. The two multiplicative scatter correction methods (MSC;PMSC) give closely similar results, provided that the moving window size comprises ±50 spectral readings or more. A minimum in the value of RMSEP for the training set is observed for the PMSC method using a window size of ±120 spectral readings. This result illustrates the empirical nature of PMSC and the necessity to scan a number of different windows sizes in order to establish a satisfactory calibration model with a low RMSEP. The plot of actual vs predicted concentrations of Sulfasalazine for the test set by near-infrared spectroscopy and partial least-squares regression with piece-wise multiplicative scatter correction is given in Fig.1.

In principle, if the liquid chromatographic method could be replaced by that of NIR/PLS a considerable decrease in analysis time could be achieved, as well as a benefit in simplicity of analysis. However, the replacement of LC as the primary quantitative method by the pro-

Table 1
Summary of results from various NIR/PLS models in reflectance mode

Model	Number of PLS components*	RMSEP†	Linear coefficient‡	Correlation coefficient‡
Raw spectra (R)	6	1.831	0.882	0.529
1st derivative of (R)	5	2.100	1.158	0.796
2nd derivative of (R)	3	2.849	1.224	0.689
log ($1/R$)	8	2.979	1.687	0.811
MSC‖ on 1st derivative of (R)	5	0.629	0.809	0.964
MSC(R)	8	0.404	0.946	0.982
MSC($1/R$)	7	0.420	0.855	0.989
PMSC(R)§¶	6	0.386	0.880	0.990

*Optimal numbers of factors determined by cross validation.
†On the test set, RMSEP = root mean square error of prediction (see text).
‡Linear regression analysis on actual vs predicted data for the test set.
§Window size ±120 spectral variables.
‖Multiplicative scatter correction.
¶Piece-wise multiplicative scatter correction

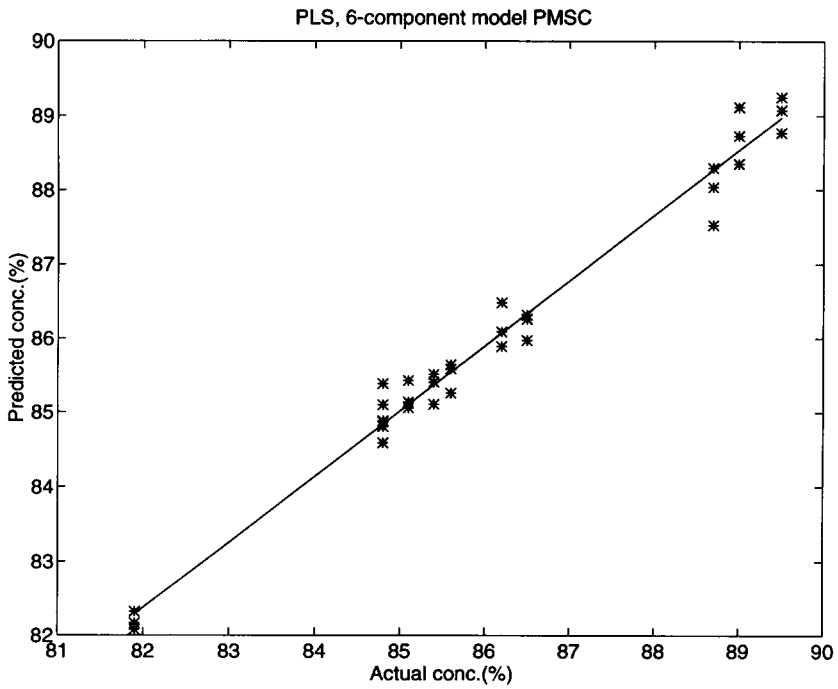

Figure 1
Actual vs predicted Sulfasalazine concentrations (%) in samples of technical quality (for the test set). The plot represents results obtained by near-infrared reflectance spectroscopy and partial least squares regression. Data pre-treatment was by piece-wise multiplicative scatter correction (window size ±120 spectra variables; 6-component model).

posed secondary NIR/PLS method, would require that questions addressing, for example, the robustness of calibration and the need for adequate methods for outlier detection, will have to be answered in depth.

Acknowledgement — Nils Nyberg is acknowledged for implementing the multiplicative scatter correction method and piece-wise multiplicative scatter correction method as Matlab routines.

References

[1] H. Martens and T. Naes, *Multivariate Calibration*. Wiley, Chichester (1989).
[2] P. Geladi, D. MacDougall and H. Martens, *Appl. Spectro.* **39**, 491–500 (1985).
[3] J. Toft, O.M. Kvalheim, T.V. Karstang, A.A. Christy, K. Kleveland and A. Henriksen, *Appl. Spectro.* **46**, 1002–1008 (1992).
[4] T. Isaksson and B. Kowalski, *Appl. Spectro.* **47**, 702–709 (1993).

[*Received for review 23 September 1994; revised manuscript received 29 November 1994*]