Thermochimica Acta 507–508 (2010) 146–149

Contents lists available at ScienceDirect

Thermochimica Acta

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

Short communication

Recognition of tablet content by chemometric processing of differential scanning calorimetry curves—An acetaminophen example

Wojciech Miltyk[b], Emilia Antonowicz[b], Łukasz Komsta[a,*]

[a] Department of Medicinal Chemistry, Medical University of Lublin, Jaczewskiego 4, 20-090 Lublin, Poland
[b] Department of Pharmaceutical Analysis, Medical University of Białystok, Kilińskiego 1, 15-089 Białystok, Poland

ARTICLE INFO

Article history:
Received 24 February 2010
Received in revised form 10 May 2010
Accepted 24 May 2010
Available online 1 June 2010

Keywords:
Pharmaceuticals
Differential scanning calorimetry
Chemometrics
Partial least squares

ABSTRACT

Chemometric processing of differential scanning calorimetry (DSC) curves is rarely approached in the literature. This paper discusses and presents an example of building discriminant models against the presence of a compound in pharmaceutical formulations. The dataset containing 11 curves of pure substances and 21 signals of various OTC tablets was processed by partial least squares discriminant analysis (PLS-DA) against detection of these substances in the formulations. Good model was found in the case of acetaminophen (5 PLS factors included), due to good proportion between positive and negative training samples. The model was validated on external prediction set and was found to have following parameters: RMS equal to 0.1068 (98.8% of explained variance), RMSECV equal to 0.148 (97.7% of explained variance) and external predictive error (RMSEP) equal to 0.3918 (86.5% of explained variance).

1. Introduction

The differential scanning calorimetry (DSC) has been used in pharmaceutical analysis since 1960s when its application in quantitative estimation of purity, the relative determination of purity and the detection of polymorphism was first described [1]. The subject has stated appearing regularly since 1980s, such as quantitative determination of cholic acids in pharmaceuticals [2], investigation of drug interactions in solid phase [3], thermal characteristics study of several drug-like substances [4], comprehensive study on purity control [5], water content estimation in pharmaceutical formulations [6] or interaction studies between drugs and excipients [7–10] or determination of the mobility of the drugs [11] and their eutectics [12]. The experimental design aspects of DSC in pharmaceutical analysis were also studied [13] and the method has recently been found to perform well in purity testing of pharmaceutical reference standards [14]. Some modifications of DSC, such as pressurized variant [15,16], modulated temperature [17] or high-speed DSC [18] are also promising alternatives in pharmaceutical analysis.

Although the DSC is a very well established technique, the chemometric processing of digitized DSC curves is an approach rarely mentioned. The first paper related to pure chemometric processing of DSC curves was published by Bergner and Albano [19]. They performed DSC analysis of peat and used the collected data to predict 15 chemical and physical constituents of the samples by partial least squares (PLS) regression. Pomerantsev and Rodionova [20,21] used hard modelling approach to predict antioxidant activity from DSC curve. Bertram et al. [22] performed simultaneous PLS modelling of DSC and NMR data in their study of processes ocurring during meat heating.

To our best knowledge, there are no papers regarding chemometric processing of DSC data in analysis of pharmaceutical formulations. Such approach seems to be attractive alternative to the other analytical methods due to following reasons:

(1) Almost no previous sample treatment (besides grinding) is needed. Therefore, routine analysis does not require much manual work.
(2) The cost of DSC analysis is related mainly to energy used for equipment running, as no other materials or reagents (such as solvents, columns, plates) are needed.
(3) The analysis can be performed almost automatically with easy data collection.

Therefore, our aim was to investigate a possibility to construct discriminant models indicating a presence (or absence) of particular compound in tablets from DSC curve.

2. Experimental

The pure substances: acetaminophen (ACE), aspirin (ASP), caffeine (CAF), codeine phosphate (COD), dipyrone (DIP), ethoxybenzamide (ETO), ibuprofen (IBU), phenylephrine hydrochloride (PHE), propyphenazone (PRO), pseudoephedrine hydrochloride

* Corresponding author. Tel.: +48 81 7423692; fax: +48 81 7423691.
E-mail address: lukasz.komsta@umlub.pl (Ł. Komsta).

0040-6031/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.tca.2010.05.017

(PSE), thiocol (THI) and vitamin C (VIT) were of analytical grade (Sigma–Aldrich, USA).

21 OTC tablets with different ingredients and their amounts were bought in local drugstore:

(1) Antidol 15 – tablets (500 mg acetaminophen, 15 mg codeine phosphate), produced by Lek (Ljubliana, Slovenia).
(2) Apap C plus – effervescent tablets (500 mg acetaminophen, 300 mg ascorbic acid), produced by US Pharmacia (Warsaw, Poland).
(3) Aspirin C – effervescent tablets (400 mg aspirin, 240 mg ascorbic acid), produced by Bayer (Oberneukirchen, Germany).
(4) Cefalgin – tablets (250 mg acetaminophen, 150 mg prophyphenazone, 50 mg caffeine), produced by Polfa (Pabianice, Poland).
(5) Coffepirine – tablets (450 mg aspirin, 50 mg caffeine), produced by Marcmed (Lublin, Poland).
(6) Coldrex HotRem – sachets (750 mg acetaminophen, 10 mg phenylephrine hydrochloride, 60 mg ascorbic acid), produced by GlaxoSmithKline (München, Germany).
(7) Coldrex MaxGrip – sachets (1000 mg acetaminophen, 10 mg phenylephrine hydrochloride, 40 mg ascorbic acid), produced by GlaxoSmithKline (München, Germany).
(8) Dafalgan Codeine – tablets (500 mg acetaminophen, 30 mg codeine phosphate), produced by UPSA (Agen, France).
(9) Efferalgan – effervescent tablets (330 mg acetaminophen, 200 mg ascorbic acid), produced by UPSA (Agen, France).
(10) Etopiryna – tablets (300 mg aspirin, 100 mg ethoxybenzamide, 50 mg caffeine), produced by Polpharma (Starogard, Poland).
(11) Gardan P – tablets (200 mg prophyphenazone, 300 mg dipyrone), produced by Polfa (Pabianice, Poland).
(12) Ibuprom – tablets (200 mg ibuprofen), produced by USP (Warsaw, Poland).
(13) Modafen – tablets (200 mg ibuprofen, 30 mg pseudoephedrine hydrochloride), produced by Zentiva (Warsaw, Poland).
(14) Nurofen Plus – tablets (200 mg ibuprofen, 12.8 mg codeine phosphate), produced by Boots Healthcare (Warsaw, Poland).
(15) Panadol Extra – tablets (500 mg acetaminophen, 65 mg caffeine), produced by GlaxoSmithKline (München, Germany).
(16) Saridon – tablets (250 mg acetaminophen, 150 mg prophyphenazone, 50 mg caffeine), produced by Roche (Basel, Switzerland).
(17) Solpadeine – tablets (500 mg acetaminophen, 30 mg caffeine, 8 mg codeine phosphate), produced by SmithKline Beecham.
(18) Solpadeine – effervescent tablets (500 mg acetaminophen, 30 mg caffeine, 8 mg codeine phosphate), produced by GlaxoSmithKline (München, Germany).
(19) Solpadeine – capsules (500 mg acetaminophen, 30 mg caffeine, 8 mg codeine phosphate), produced by SmithKline Beecham (Dresden, Germany).
(20) Talvosilen forte – capsules (500 mg acetaminophen, 30 mg codeine phosphate), produced by Bene-Arzneimittel GmbH (München, Germany).
(21) Thiocodin – tablets (15 mg codeine phosphate, 300 mg thiocol), produced by Unia (Warsaw, Poland).

A "Jade" (Perkin-Elmer) DSC calorimeter was used in all measurements. The weight of each sample was about 3 mg. The curves were registered in range of 30–300 °C with temperature gradient 5 °C × min^{-1} in nitrogen atmosphere. The equipment was calibrated against zinc and indium. Each curve was averaged from 3 measurements and exported as vector of length equal to 3298.

The curves of pure substances were collected together with curves of tablets into one dataset. Partial least squares discriminant analysis (PLS-DA) of resulted matrix (dimensions: 32 × 3298) was performed with Matlab R2009b, using TOMCAT toolbox for multivariate calibration [23].

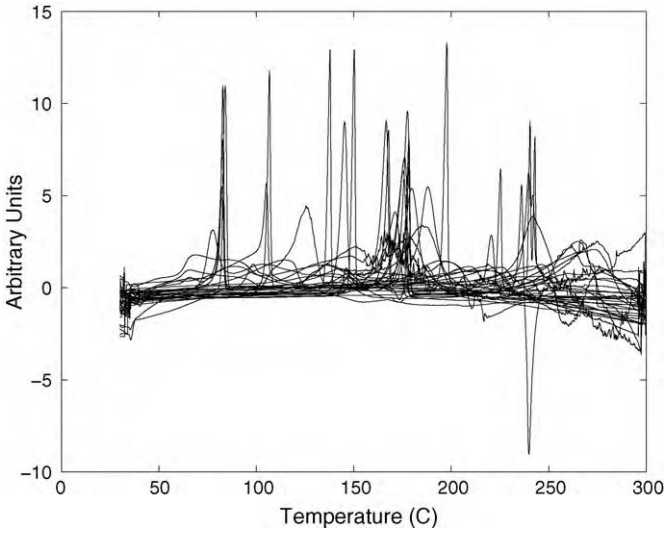

Fig. 1. The DSC dataset after standard normal variate (SNV) preprocessing.

3. Results and discussion

The dataset was formed of pure substances and tablets containing different drugs and unknown excipients in different qualitative and quantitative combinations. In tablets, acetaminophen is present in 13 of them, caffeine in 8, codeine phosphate in 8, aspirin in 4, vitamin C in 4, propyphenazone in 3, ibuprofen in 3, ethoxybenzamide in 2, phenylephrine in 2, dipyrone in 1, thiocol in 1. The curves were preprocessed by standard normal variate (SNV) treatment – each signal was scaled and centered to have mean equal to zero and unit variance. This preserves only "shape" of the signal and removes unwanted variability in absolute signal values. The dataset after preprocessing is shown in Fig. 1.

Preliminary dataset exploration by unscaled principal component analysis showed that PC1/PC2 plot represents uniformly distributed cloud. The first PC explains only 22.9% of total variance, the second explains only 13% of additional variance. This confirms the complex nature of DSC signal and quite high complexity of PLS-DA models should be suspected.

Theoretically, a manual interpretation of DSC curves should be simple (each compound should give a main peak at the temperature of its melting). In practice, as in the case of the presented dataset, it is not that easy. Several compounds can melt at very similar temperatures (for example acetaminophen 169 °C, dipyrone 173 °C). Moreover, the melting point of each substance can be shifted due to some physical interactions between ingredients and/or excipients. Moreover, additional peaks (indicating some complex processes in solid phase) can be present.

The PLS-DA is a flexible approach of discriminant analysis using partial least squares (PLS) regression method. Having a X matrix, containing n rows (signals) and p columns (each DSC curve of length p) and a column vector y containing 1 values for "positive" and −1 for "negative" samples, the method tries to find row-vector estimator of p length, called β, satisfying following equation in matrix form:

$$y = X\beta + \varepsilon \quad (1)$$

where ε is a matrix of normally distributed measurement errors. Once the estimator is found, one can multiply DSC curve by it and automatically check for a presence of a compound (when $y > 0$) or its absence (when $y < 0$). A reader interested in theoretical consider-

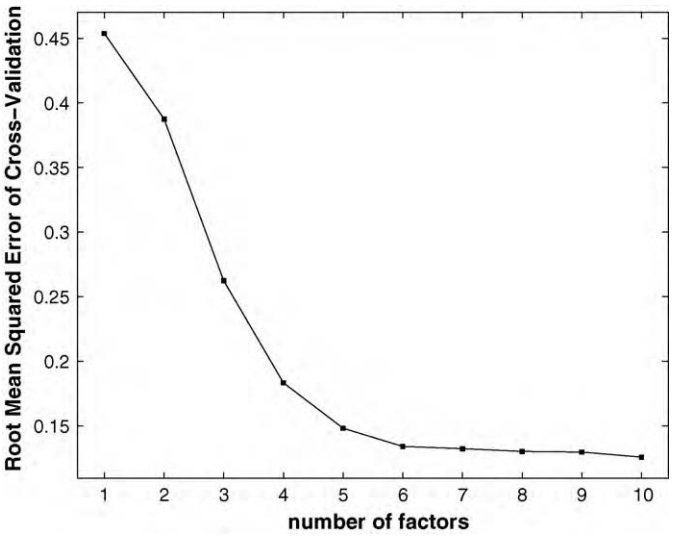

Fig. 2. The leave-one-out crossvalidation results of PLS-DA model building.

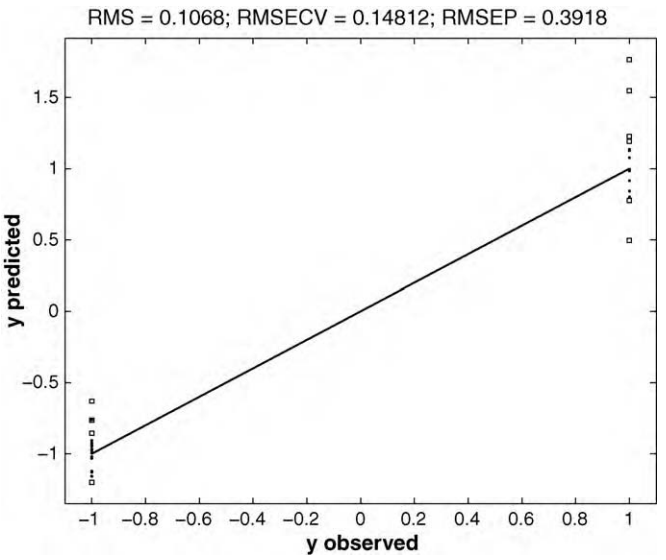

Fig. 4. Model validation results. LOO crossvalidation predictions are marked as dots, external validation values are marked as squares.

ations of PLS itself and application of this method in chemometrics is referred to excelent tutorial by Wold et al. [24].

The dataset used for building PLS-DA should meet some requirements, the main of which is a proper proportion between the number of "positive" and "negative" samples [25]. In the case of the presented tablets, only acetaminophen divides the samples into classes of enough equal length (14 positive, 18 negative). Although building models against presence of other compounds seemed to be good idea and the crossvalidation results seemed to be acceptable, their predictive ability (against external testing set) was very poor.

The acetaminophen model was built in the following way. First, 10 of 32 curves were selected by Kennard–Stone algorithm as a representative external validation dataset and removed from further processing. The remaining 22 curves were used as a training PLS-DA dataset.

The main coefficients characterizing model quality are expressed as a root mean square error:

$$\mathrm{RMS} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i' - y_i)^2} \qquad (2)$$

where y_i is i-th value (-1 or 1 in our case) and y_i' is a value obtained from the model. When the last one is a fitted one, the simple RMS of the model is obtained. However, the most interesting ones, measuring real model performance are RMSECV (root mean square error of crossvalidation, when y_i are cross-validated y values) and RMSEP (root mean square error of prediction, where y_i come from independent external dataset).

The crossvalidation results are presented in Fig. 2. The number of PLS factors equal to 5 was chosen as optimal and its estimator is presented in Fig. 3. The model has RMS equal to 0.1068 (98.8% of explained variance). The RMSECV value is equal to 0.148 (97.7% of explained variance) and external predictive error (RMSEP) is equal to 0.3918 (86.5% of explained variance). The predicted values, shown in Fig. 4, conclude that there is no misclassified sample, both in the crossvalidation results and the external prediction test.

As the RMSECV curve has no clear and visible minimum and RMSECV value is only slightly larger than RMS value, additional check against the optimal number of PLS factors was done by leave-five-out crossvalidation and half/half crossvalidation (1000 Monte-Carlo iterations). In both cases, the complexity equal to 5

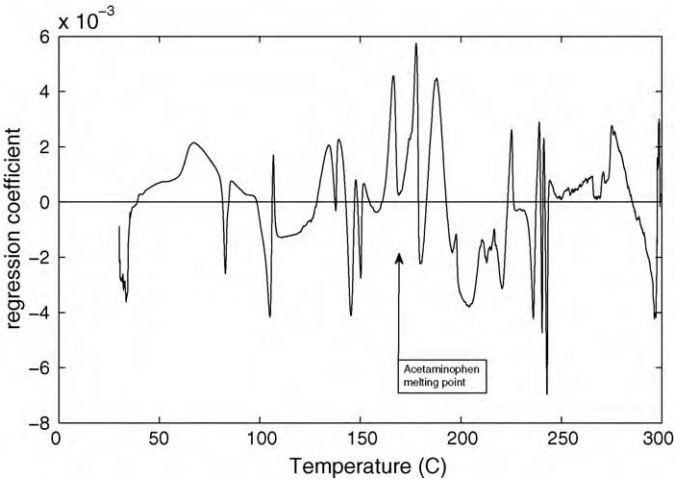

Fig. 3. The discriminant estimator recognizing presence of acetaminophen in investigated dataset.

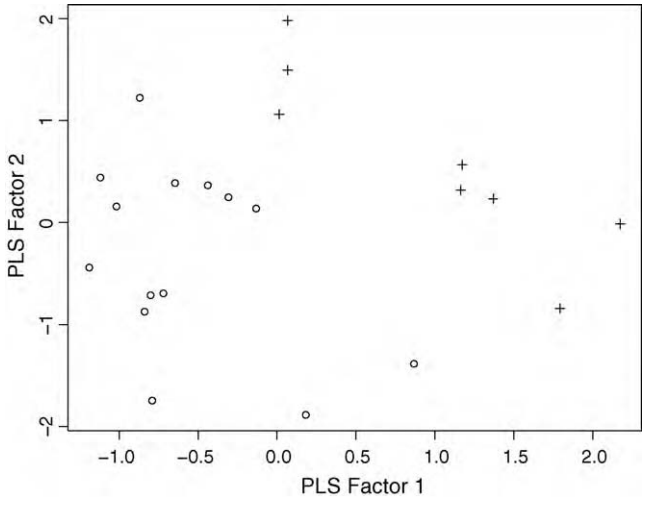

Fig. 5. Two first PLS factors plot. Samples containing acetaminophen are marked as plus signs.

was also chosen as optimal and again there was no misclassified sample during crossvalidation. Additional tests done on external dataset with less number of PLS factors did not show any visible advantage. Although a discrimitative ability is visible among first two factors (Fig. 5), less number of them performs visibly worse.

The estimator (Fig. 3) is rather complex, but around the melting point of acetaminophen (indicated by arrow) it forms narrow positive and negative peaks sensitive to acetaminophen and insensitive to dipyrone. Its shape also reflects other peaks, caused by interactions between acetaminophen and other compounds/excipients.

4. Conclusion

The discriminant models can be easily constructed from DSC datasets by PLS-DA technique, even if the dataset does not meet full factorial experimental design (in our case there would be 2048 combinations). The only requirement is to collect similar amount (and enough amount) of positive and negative samples. This suggest a new area of interest for future research – chemometric processing of DSC signals.

References

[1] R. Reubke, J.A. Mollica Jr., J. Pharm. Sci. 56 (1967) 822–825.
[2] U.B. Ceipidor, R. Curini, G. D'Ascenzo, M. Tomassetti, Thermochim. Acta 46 (1981) 279–287.
[3] J.L. Ford, M.H. Rubinstein, Drug Dev. Ind. Pharm. 7 (1981) 675–682.
[4] F.I. Khattab, Thermochim. Acta 61 (1983) 253–268.
[5] V. Grdinic, A. Bezjak, A. Janekovic, B. Cetina-Cizmek, D. Briski, Farmaceutski Glasnik 47 (1991) 195–204.
[6] R.K. Khankari, D. Law, D.J.W. Grant, Int. J. Pharm. 82 (1992) 117–127.
[7] K. Danjo, F. Nishio, B.D. Zhow, A. Otsuka, Chem. Pharm. Bull. 43 (1995) 1958–1960.
[8] J.M. Ginas, M.J. Arias, A.M. Rabasco, C. Novak, A. Ruiz-Conde, P.J. Sanchez-Soto, J. Therm. Anal. 46 (1996) 291–304.
[9] P. Mura, G.P. Bettinetti, M.T. Faucci, A. Manderioli, P.L. Parrini, Thermochim. Acta 321 (1998) 59–65.
[10] P. Mura, M.T. Faucci, A. Manderioli, G. Bramanti, L. Ceccarelli, J. Pharm. Biomed. Anal. 18 (1998) 151–163.
[11] B.C. Hancock, S.L. Shamblin, Thermochim. Acta 380 (2001) 95–107.
[12] A.T. Riga, L.M. Oberoi, K.S. Alexander, Am. Pharm. Rev. 7 (2004) 18–23.
[13] S. Roy, A.T. Riga, K.S. Alexander, Thermochim. Acta 392–393 (2002) 399–404.
[14] S. Mathkar, S. Kumar, A. Bystol, K. Olawoore, D. Min, R. Markovich, A. Rustum, J. Pharm. Biomed. Anal. 49 (2009) 627–631.
[15] J. Han, R. Suryanarayanan, Int. J. Pharm. 157 (1997) 209–218.
[16] J. Han, S. Gupte, R. Suryanarayanan, Int. J. Pharm. 170 (1998) 63–72.
[17] N.J. Coleman, D.Q.M. Craig, Int. J. Pharm. 135 (1996) 13–29.
[18] C. McGregor, E. Bines, Int. J. Pharm. 350 (2008) 48–52.
[19] K. Bergner, C. Albano, Anal. Chem. 65 (1993) 204–208.
[20] A.L. Pomerantsev, O.Ye. Rodionova, Chemom. Intell. Lab. Syst. 79 (2005) 73–83.
[21] A.L. Pomerantsev, O.Ye. Rodionova, Kinet. Catal. 47 (2006) 537–548.
[22] H.C. Bertram, Z. Wu, F. van den Berg, H.J. Andersen, Meat Sci. 74 (2006) 684–689.
[23] M. Daszykowski, S. Serneels, K. Kaczmarek, P. Van Espen, C. Croux, B. Walczak, Chemom. Intell. Lab. Syst. 85 (2007) 269–277.
[24] S. Wold, M. Sjostrom, L. Eriksson, Chemom. Intell. Lab. Syst. 58 (2001).
[25] R.G. Brereton, Trends Anal. Chem. 25 (2006) 1103–1111.