Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Pharmaceutical and Biomedical Analysis 38 (2005) 397–407
JOURNAL OF PHARMACEUTICAL AND BIOMEDICAL ANALYSIS
www.elsevier.com/locate/jpba

Classification of chondroitin sulfate A, chondroitin sulfate C, glucosamine hydrochloride and glucosamine 6 sulfate using chemometric techniques

M. Foot [a,*], M. Mulholland [b]

[a] *Department of Chemistry Materials and Forensic Science, University of Technology, P.O. Box 123, Broadway 2007, Sydney, NSW, Australia*
[b] *School of Chemical Sciences, University of New South Wales, NSW, Australia*

Accepted 20 January 2005
Available online 8 March 2005

Abstract

Chondroitin sulfate A, chondroitin sulfate C, glucosamine hydrochloride and glucosamine sulfate are natural products that are becoming increasingly popular in the treatment of arthritis. They belong to a class of compounds known as glycosaminoglycans (GAGs). They are available over the counter as nutritional supplements. However, increasing use has led to increasing scrutiny of the quality of products on the market. There is also interest in the pharmacological properties of these compounds. To facilitate this, there is a need for better qualitative and quantitative methods of analysis. This paper describes methods for achieving the qualitative identification of chondroitin sulfate A, chondroitin sulfate C, glucosamine hydrochloride or glucosamine sulfate. Fourier transform infrared spectroscopy coupled with a variety of chemometric methods successfully classified these compounds. Using soft independent modeling of class analogies (SIMCA), hierarchical cluster analysis (HCA) and principal components analysis (PCA) samples were classified as either chondroitin sulfate A, chondroitin sulfate C, glucosamine hydrochloride or glucosamine sulfate. This work also examined the discriminating ability of different sections of the spectrum. It was found that for the classification of these compounds that using the finger print region of the spectrum (below 2000 cm^{-1}) gave the best discrimination.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Glycosaminoglyans; Glucosamine; Chondroitin sulphate; FTIR; PCA; SIMCA

1. Introduction

Osteoarthritis (OA) is a degenerative disease of the cartilage in the joints of the human body [1,2]. OA is generally characterized by pain and or swelling in the affected joint and has an increased occurrence in women and those who are overweight [3]. It has been suggested by Mankin et al. that OA is not one disease but a number of diseases through which similar symptoms arise [1]. Chemically, arthritis is characterized by a change in the composition of the extra cellular matrix of the joint [3]. Ishiguro notes there is a distinct change in the ratio of chondroitin sulfate 6 to chondroitin sulfate 4.

There are a number of treatment options available to sufferers of arthritis, ranging from simple lifestyle changes to the use of pharmaceuticals to treat pain and inflammation for example non steroidal anti-inflammatory drugs (NSAIDs) and natural products (Nutraceuticals) [1,4,5]. Extreme cases require surgery including total joint replacement [4]. Increasingly, there has been a move towards the use of so-called natural products for the treatment of OA. The reasons for the increased uptake of these products include cost, availability and a perception of greater safety with the use of natural products [6].

Chondroitin sulfates are one type of "Nutraceuticals" that are being used in the treatment of the symptoms of arthritis [4,7]. Chondroitin sulfates are naturally occurring glycosaminoglycans (GAGs). GAGs share some similar structural properties with polysaccharides. Chondroitin sulfates

* Corresponding author. Tel.: +61 2 9514 1983; fax: +61 2 9514 1460.
E-mail address: matthew.j.foot@uts.edu.au (M. Foot).

0731-7085/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2005.01.026

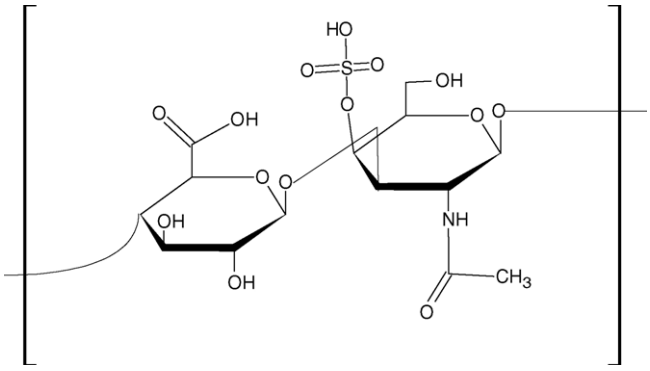

Fig. 1. The disaccharide unit of chondroitin sulfate A.

occur as a part of the proteoglycan Aggrecan, which is considered as an important part of the composition of joint cartilage. Hardingham et al. suggests that the role of aggrecan is to draw water into the collagen matrix of cartilage to give it some of its structural properties [8]. Chondroitin sulfate is present in a number of forms the most common forms are chondroitin 4 sulfate (CS4) also known as chondroitin sulfate A (CSA) seen in Fig. 1, chondroitin 6 sulfate (CS6) also known as chondroitin sulfate C (CSC) seen in Fig. 2 and dermatan sulfate also known as chondroitin sulfate B (CSB). Chondroitins are glycosaminoglycans made up of alternating uronic acid and *N*-acetyl-D galactosamine residues. The disaccharide units are joined to one another by a β1-4 linkage the residues are joined by a β1-3 linkage. Dermatan sulfate is similar, but contains iduronic acid as opposed to uronic acid. There are a number of different sources of chondroitin sulfate. In the nutraceutical market the three main sources are bovine cartilage derived, shark cartilage derived and porcine derived.

It has been reported that treatment with preparations that contain chondroitin sulfates, are more effective than placebo preparations and have a longer lasting, yet slower action than some NSAIDs [5,9]. One of the claimed advantages of chondroitin sulfate therapy over conventional NSAID therapy is that there is a reduced risk of gastrointestinal ulcers, which can be a side effect of NSAID treatments [9–11]. With the

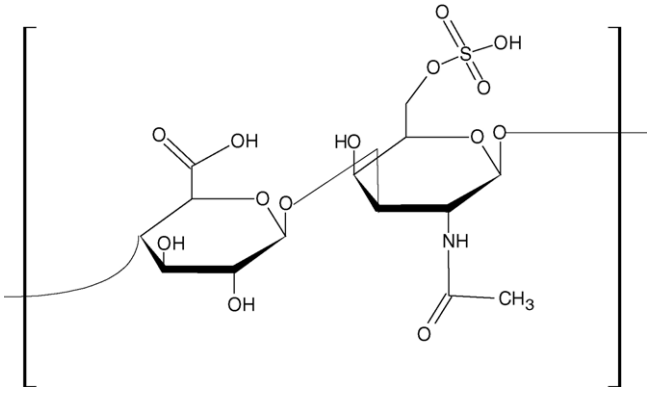

Fig. 2. The disaccharide unit of chondroitin sulfate C.

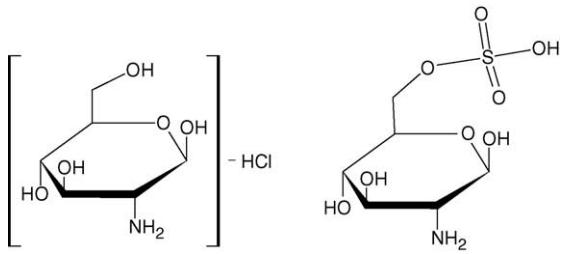

Fig. 3. The structure of glucosamine hydrochloride (left) and glucosamine 6 sulfate (right).

increase in use of these drugs there will be an increased need for new and better analysis techniques.

Glucosamine is a naturally occurring compound present in many of the body's tissues being used in the treatment of arthritis [7]. The chemical structure of glucosamine can be seen in Fig. 3. Glucosamine for arthritis products is usually formulated as the hydrochloride salt or glucosamine sulfate and often combined with chondroitin sulphate. It is notable that while both the hydrochloride salt and glucosamine sulfate are used in pharmaceutical preparations, glucosamine sulfate is thought to have a higher biological activity due to the presence of the sulfate [12]. It should also be noted that there is a large cost difference between the two salts, with the hydrochloride salt being significantly less expensive.

The chemical structure of GAG compounds make them suitable for study with infrared spectroscopy (IR). It has been shown that IR spectroscopy can be used to gather information about GAGs in aqueous and deuterium chloride solutions by Cabassi et al and Casu et al., respectively [13,14]. Other work with IR spectroscopy has included the analysis of the breakdown products of the polymeric components of cartilage, when reacted with enzymatically produced hypochlorous acid in order to better understand the pathology of arthritis [15]. Servaty et al. have studied the hydration profiles of GAG compounds to better understand how they absorb water, an important chemical property that relates to the physical properties of cartilage [16].

Chemometric pattern recognition techniques have been successfully applied with infrared spectroscopy to the analysis of polysaccharides in wine using principal components analysis (PCA) and canonical correlation analysis (CCA) [17], the analysis of sugar adulterants in honey using linear discriminate analysis (LDA) [18], esterification in pectin samples using PCA and principal variables analysis (PV) [19] and the classification of the GAG NaPPS using PCA and SIMCA [20].

This work looked at using principal components analysis as a data reduction techniques and linking that with hierarchical cluster analysis (HCA) and soft independent modeling of class analogies (SIMCA developed by S. Wold in the 1970s [21]). SIMCA revolves around the use of PCA as a data reduction and modeling technique [22].

Hierarchical cluster analysis is a multivariate analysis technique that is used to sort samples into groups, with the

goal of having each group or cluster containing objects which are similar to each other. Often cluster analysis occurs after a visual inspection of the data using a technique such as PCA [22]. The similarity or dissimilarity between samples (objects) is usually represented in a dendrogram for ease of interpretation.

2. Experimental

2.1. Chemicals and reagents

Chondroitin sulfate A sodium salt (70%), chondroitin sulfate C sodium salt (90%), glucosamine hydrochloride and glucosamine 6 sulfate were purchased from Sigma–Aldrich (Sigma–Aldrich, Australia, Castle hill, Australia). The samples were pressed into KBr disks (Spectroscopic Grade KBr, Merck Darmstadt, Germany).

2.2. Instrumentation

All samples were analysed by transmission fourier transform infrared spectroscopy (FTIR) spectroscopy on a Nicolet Magna IR 760 spectrometer (Nicolet, WI, USA).

2.3. Data analysis

Data analysis was undertaken using Microsoft Excel (Microsoft, Redmond, USA) and SIRIUS (PRS, Bergen, Norway). The data was first converted into an excel file, excel was then used to calculate the first derivative spectra. Both the original data and the first derivative data were transposed into SIRIUS 6.5. SIRIUS was then used to perform principal components analysis, hierarchical cluster analysis and SIMCA on the data.

3. Results and discussion

Ten spectra from each sample were collected. All spectra of one sample were combined into a data matrix containing the mid infrared transmission values for each of the repeats at each of the recorded wave numbers. From this matrix, a matrix of first derivative values was calculated thus the data was presented as two matrices for 40 samples. The first derivative transformation greatly reduced the intra-sample variation. The spectra of chondroitin sulfate A, chondroitin sulfate C, glucosamine hydrochloride and glucosamine sulfate can be seen in Fig. 4.

In the chondroitin sulfate A and chondroitin sulfate C spectra the region above 2000 cm^{-1} was dominated by the OH stretching vibration. The band at 1630 and 1650 cm^{-1} was due to the amide 1 band for chondroitin sulfate A and chondroitin sulfate C, respectively. The band at 1250 cm^{-1} for CSA and at 1240 cm^{-1} for CSC has been assigned to S=O corresponding to the band assignment by Cabassi et al. [13]. The peak at 850 cm^{-1} (CSA) and 825 cm^{-1} (CSC) were due to the C–O–S vibration according to Honda et al. [23]. Peaks at 1150 and 1050 cm^{-1} (CSA) and 1140 and 1070 cm^{-1} (CSC) were probably due to carbon–carbon vibrations, while peaks at 1420 and 1380 cm^{-1} in both compounds were from carbon–hydrogen vibrations.

The spectrums of glucosamine 6 sulfate and glucosamine hydrochloride are similar to that of glucose as they share very similar structures. Both samples showed exceptions, due to the presence of the sulfate and amine functional groups. The S=O band in glucosamine 6 sulfate was seen at around 1230 cm^{-1}. Glucosamine hydrochloride showed two N–H peaks above 3000 cm^{-1}. Both samples showed a number of peaks between 1500 and 1700 cm^{-1} that were due to the amine group on each ring structure. Notable here were the peaks at 1585 and 1540 cm^{-1} in the glucosamine hydrochloride that are due to the asymmetric and symmetric NH_3^+ bending, respectively.

The effect of selecting various wavenumber ranges was investigated. First, the full spectrum for both the raw spectrums and the first derivative data, second the area below 2000 cm^{-1} for both the raw data and the first derivative and finally the area above 2000 cm^{-1} for both the raw and first derivative data.

3.1. Principal components analysis

The score plot of the first two principal components (Fig. 5.), for the full spectrum showed the four groups clearly separated. One of the chondroitin sulfate C samples was excluded from the group as it was determined to be an outlier. As expected, the chondroitin sulfate groups were closer to each other than the glucosamine groups. The variable loading plot (Fig. 6.), of the first principal component indicates that a great deal of the information was derived from the area above 2300 cm^{-1}. The loading plot of the second principal component (Fig. 7.), is dominated by peaks below 2300 cm^{-1}. The plot of the first two principal components also indicated a clear division between the chondroitin and the glucosamine samples. The plot of the first two principal components of the raw data below 2000 cm^{-1} looked highly similar to that of the first two principal components of the full spectrum, however, it appeared to have been shifted through approximately a quarter turn anticlockwise this is seen in Fig. 8. Again one chondroitin sulfate C sample was removed because it was an outlier. This suggests that the area above 2000 cm^{-1}, while greatly affecting the principal components analysis had little effect on the differentiation of the samples and that most of the differences between the samples were occurring in the fingerprint region of the spectrum. Notably, the highest loading came from the area of the spectrum around 1230 cm^{-1}. PCA performed on the area above 2000 cm^{-1} showed similar trends to the full spectrum, however there was less discrimination between the both groups of compounds.

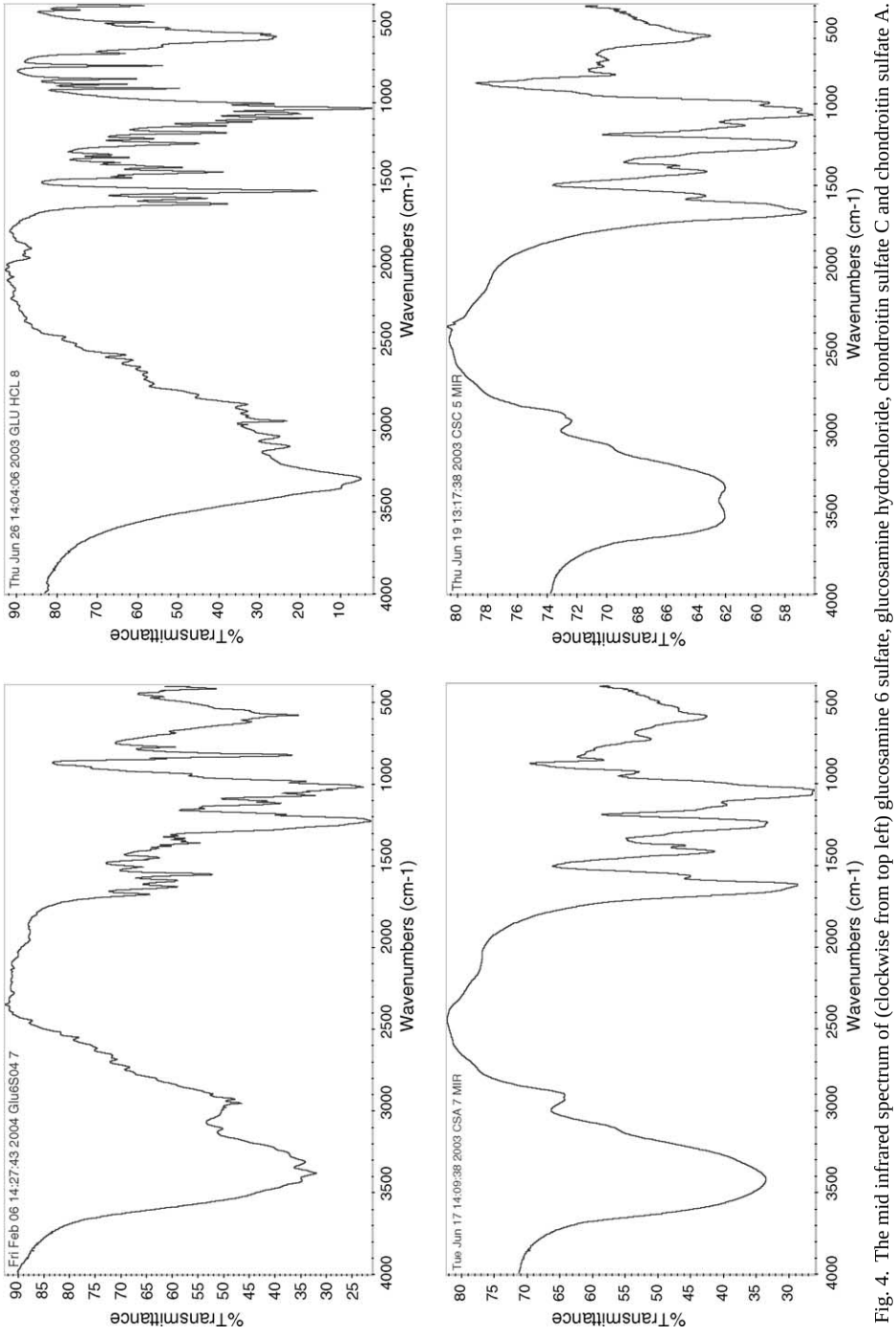

Fig. 4. The mid infrared spectrum of (clockwise from top left) glucosamine 6 sulfate, glucosamine hydrochloride, chondroitin sulfate C and chondroitin sulfate A.

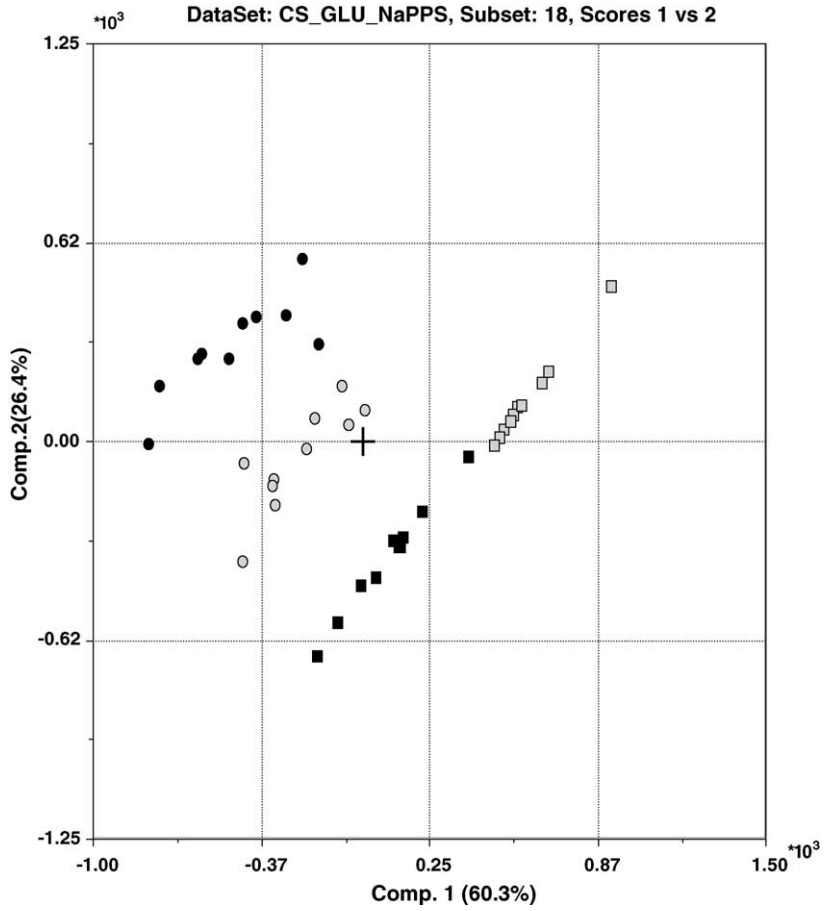

Fig. 5. Score plot of the first two principal components for the full spectrum. The black squares represent chondroitin sulfate A, the grey squares represent chondroitin sulfate C, the black circles represent glucosamine hydrochloride and the grey circles represent glucosamine 6 sulfate.

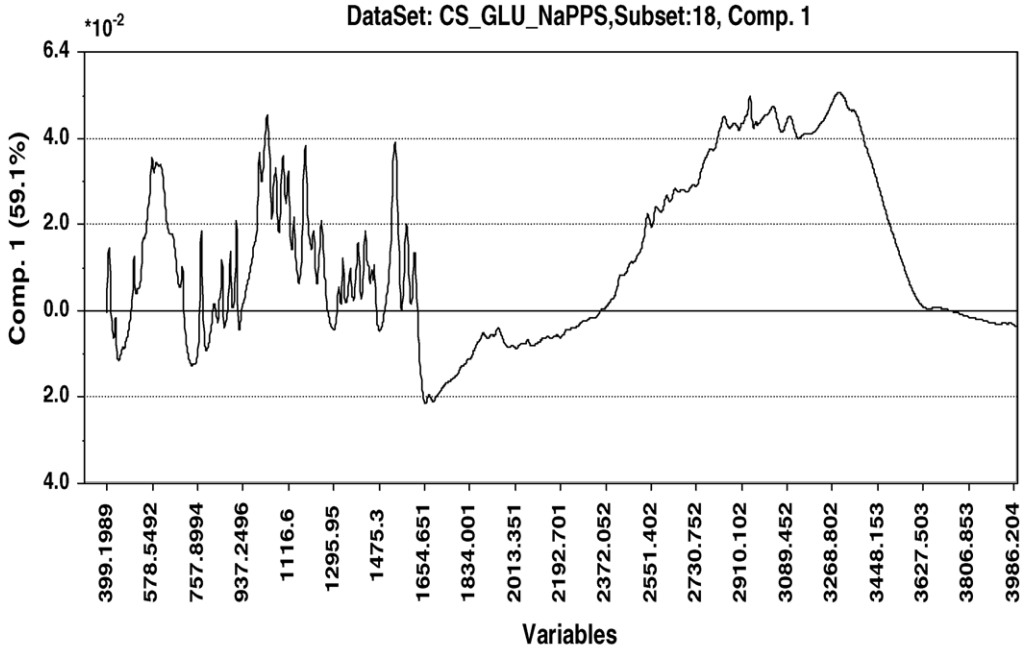

Fig. 6. The loadings plot of the first principal component for the full spectrum using the raw data.

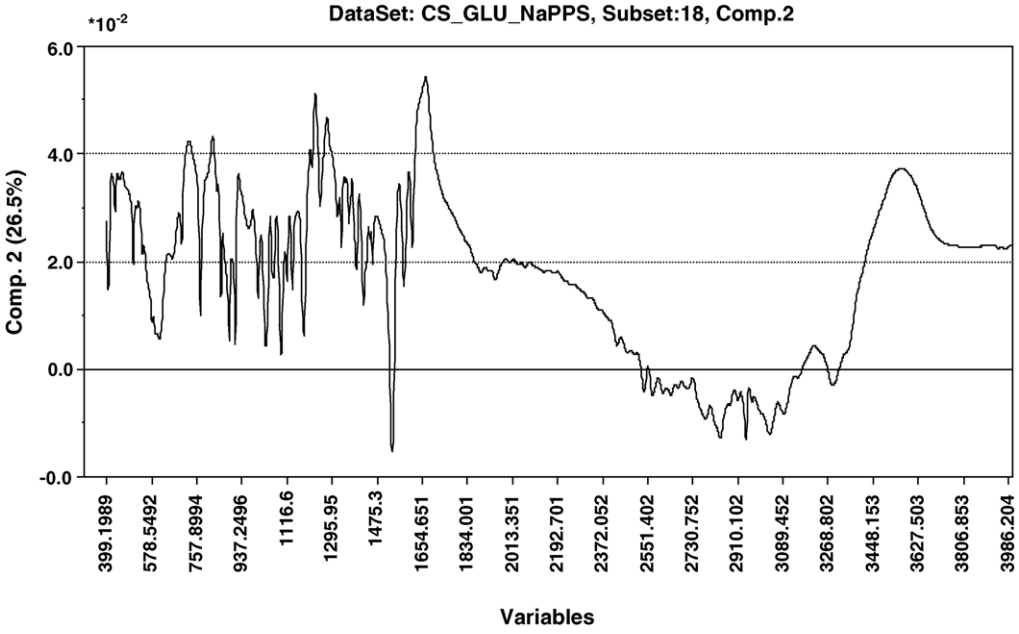

Fig. 7. The loadings plot of the second principal component for the full spectrum using the raw data.

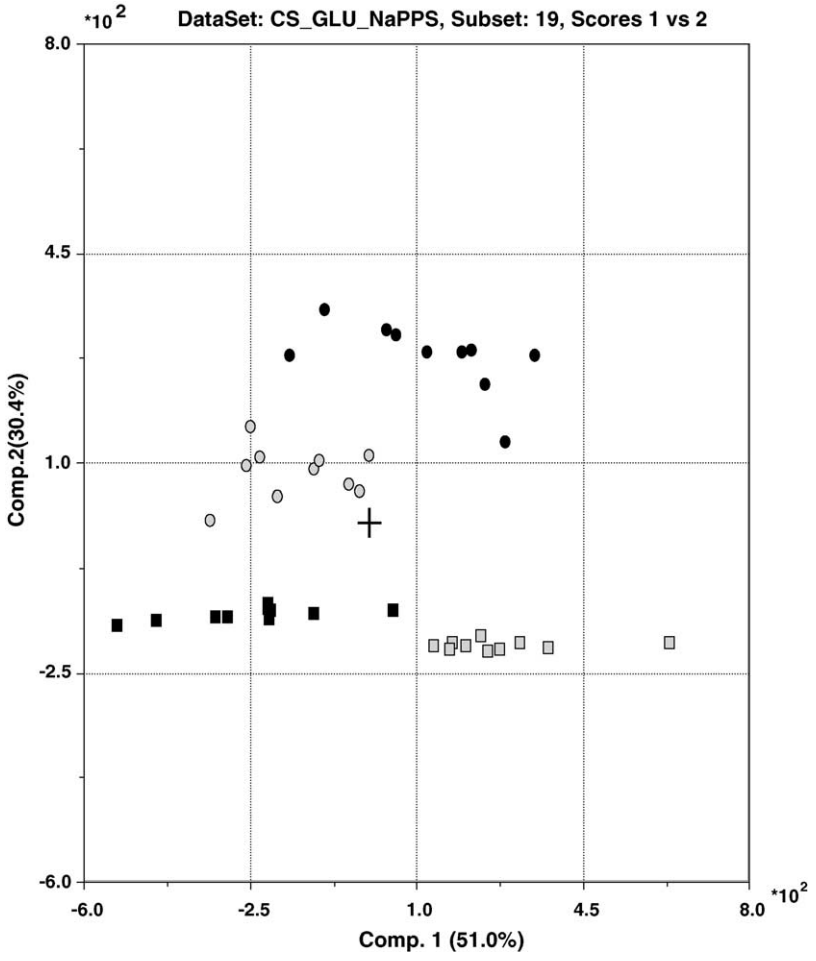

Fig. 8. The score plot of the first two principal components of the raw data for the region of the spectrum below 2000 cm^{-1}. The black squares represent chondroitin sulfate A, the grey squares represent chondroitin sulfate C, the black circles represent glucosamine hydrochloride and the grey circles represent glucosamine 6 sulfate.

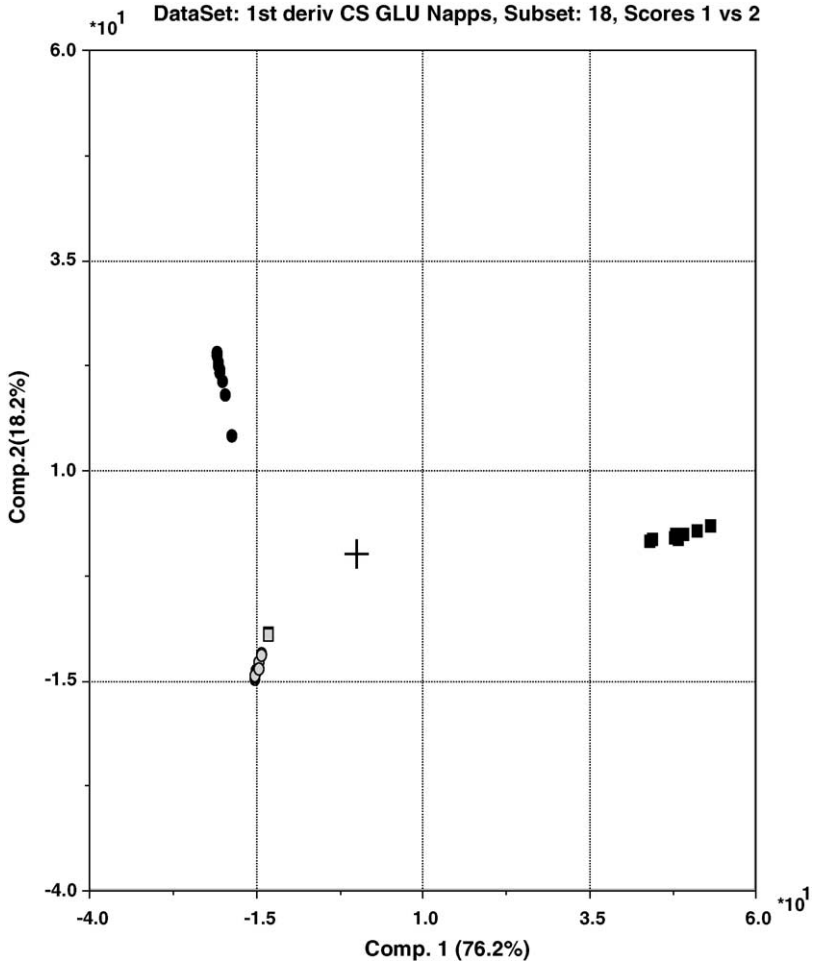

Fig. 9. The score plot of the first two principal components for the full spectrum of the first derivative data. The black squares represent chondroitin sulfate A, the grey squares represent chondroitin sulfate C, the black circles represent glucosamine hydrochloride and the grey circles represent glucosamine 6 sulfate.

The score plot of the first two principal components of the full first derivative spectrum showed good separation of all the samples, this can be seen in Fig. 9. Indeed the scores of all 10 samples of the chondroitin sulfate C almost completely overlapped each other. It was interesting to note that the chondroitin sulfate C was grouped closer to the glucosamine 6 sulfate than the chondroitin sulfate A. The loading plot of the first principal component showed both positive and negative spikes at 3200 cm^{-1}. Similar to the raw spectrum results, the twisting effect was seen when using only the area below 2000 cm^{-1}(seen in Fig. 10.) suggesting that most of the information contributing to the separation of the samples is coming from the area below 2000 cm^{-1}. The area above 2000 cm^{-1} showed a greatly reduced separation of the samples with overlap of chondroitin sulfate C and glucosamine 6 sulfate.

3.2. Hierarchical cluster analysis

Before performing HCA the data was reduced using principal components analysis. To construct the dendrograms (the output of HCA) the scores of the first three principal components were used this can be seen in Fig. 11. The dendrogram of the full spectrum based on the Euclidian distance showed that the chondroitin sulfate C samples were the most dissimilar compared to the other samples. It was also interesting to note that one of the chondroitin sulfate A samples was incorrectly grouped with the chondroitin sulfate C samples. Another interesting features of this information was that the glucosamine 6 sulfate was closer to the CSA than to the glucosamine hydrochloride. This may have been due to the hydrochloride salt, as opposed to the sulfated structure of the G6S, which was closer to that of the chondroitin sulfate as the sulfate is part of the monosaccharide. Although the glucosamine sulfate which is sulfated in the 6 position was more closely associated with the chondroitin sulfate A, which is sulfated in the 4 position of the *N*-acetyl-galactosamine moiety of the disaccharide unit.

The dendrograms produced from the reduced data set showed similar trends to that of the full wavelength range. One chondroitin sulfate A sample was still grouped with the chondroitin sulfate C samples and the glucosamine 6 sulfate was still shown to be most similar to the chondroitin sulfate A. Examination of the dendrograms produced from

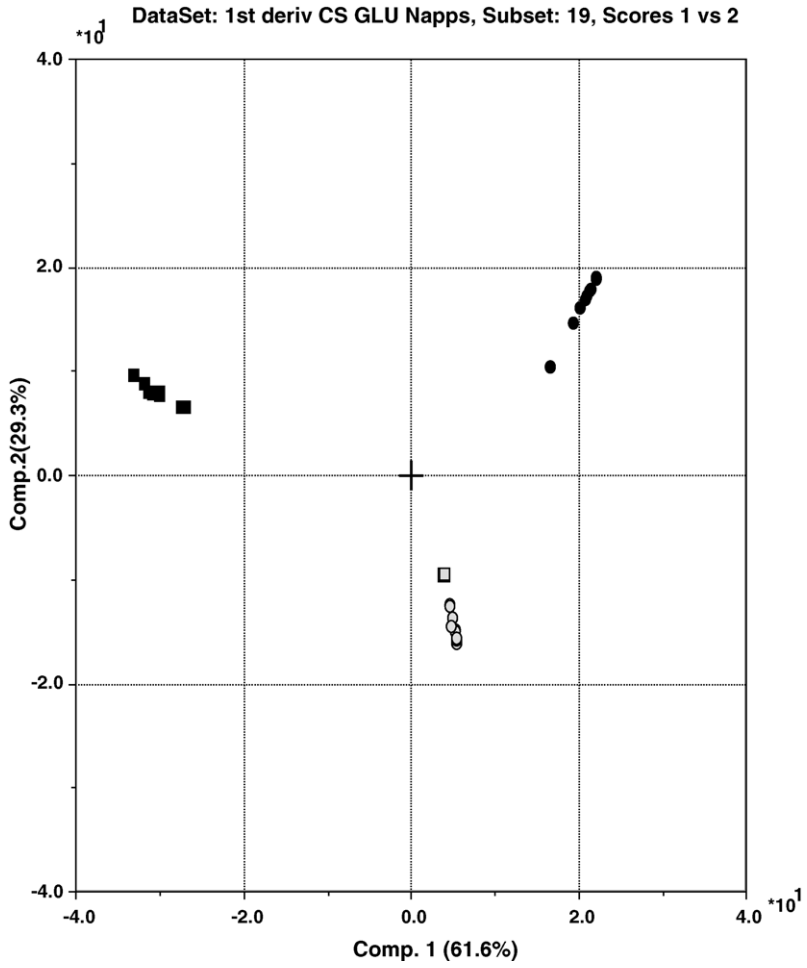

Fig. 10. The score plot of the first two principal components for region below 2000 cm^{-1} for the first derivative data. The black squares represent chondroitin sulfate A, the grey squares represent chondroitin sulfate C, the black circles represent glucosamine hydrochloride and the grey circles represent glucosamine 6 sulfate.

the first derivative data indicated that there was no longer any misclassification of the chondroitin sulfate A (with one sample incorrectly grouped with chondroitin sulfate C using the raw spectra). As with the raw data there was little change in the structure of the dendrogram when using only the area below 2000 cm^{-1}. Both data sets showed the glucosamine 6 sulfate grouped more closely to the chondroitin sulfate C. The chondroitin sulfate A was the most dissimilar group.

3.3. SIMCA

Soft independent modeling of class analogies has been used to classify the four groups based on a principal components model of each group. This work used the cross validation method to test the goodness of fit as opposed to a test set [24,25]. Using SIRIUS, cross validated principal component models were fitted to one another to calculate a class distance. A class distance of less than one meant that there was very little differences between the groups, a class distance of greater than one but less than thee indicated a partial separation of the groups, a class distance of greater than three indicated complete separation of the two groups.

The raw spectra in the cross validated model for each group was tested for outliers, which were then discarded and the group remodeled. For the raw spectra only one sample across all the groups was discarded, this was from the chondroitin sulfate C group leaving this group with only nine members. Table 1 lists the number samples and the number of principal components calculated in the cross validated model for each group using the different spectral ranges. Table 2 shows the class distances calculated for each group using the three areas of the raw spectrum. It was interesting to note that for each of regions of the spectrums the same samples were not completely resolved, namely the chondroitin sulfate A and both the glucosamine samples. However, it can be seen that the region of the spectrum below 2000 cm^{-1} comes closest to reaching the critical value (3) indicating complete separation of the two groups.

Table 3 shows the number of samples and the number of principal components calculated for the cross validated models using the three regions of the first derivative

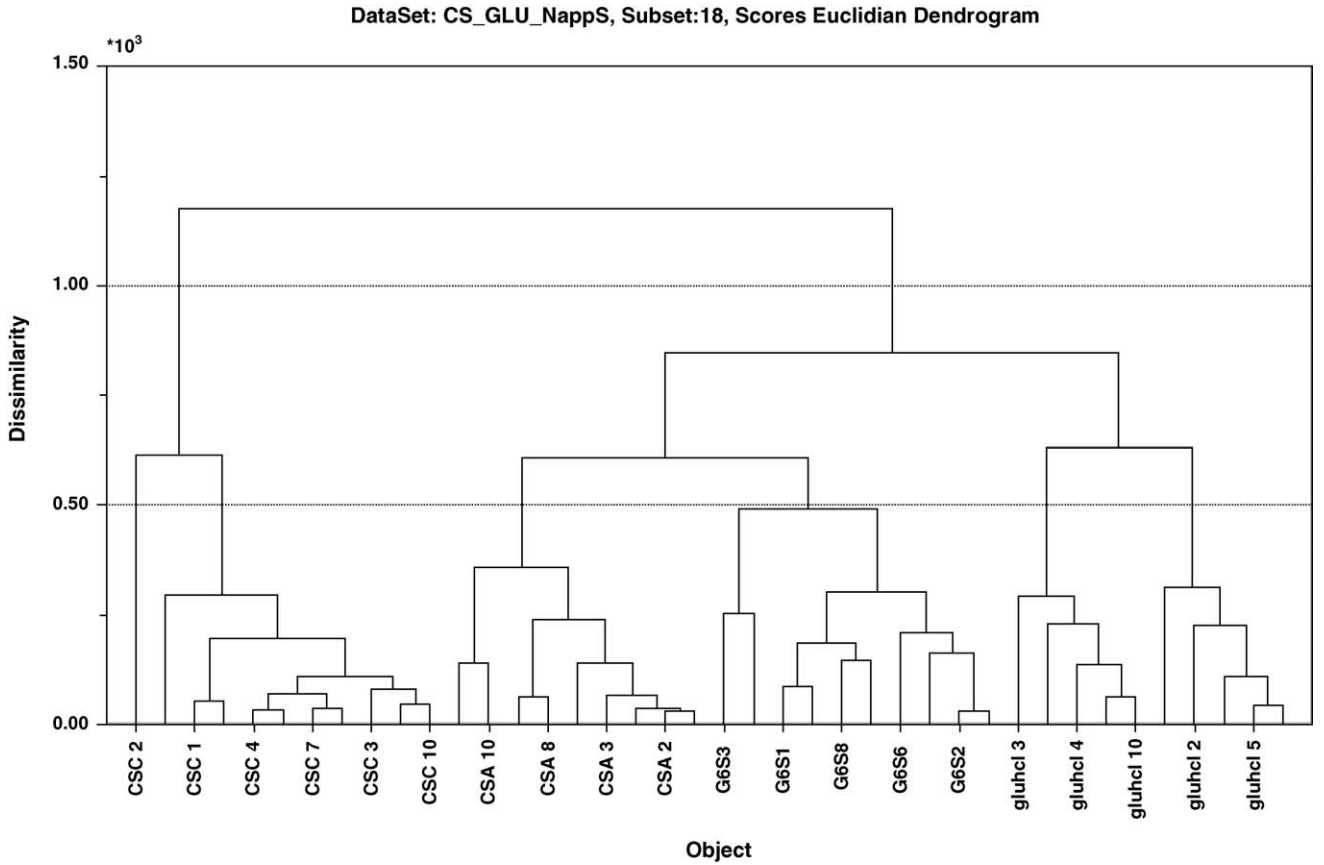

Fig. 11. Dendrogram showing the dissimilarity of the samples based on their Euclidian distance calculated from the first three principal components. The second sample is the misclassified chondroitin sulfate A sample.

Table 1
The number of principal components and the (%) variance explained for each cross validated model using the raw spectra

Sample	Spectral range	Outliers	#PC's	Variation explained (%)
CSA	Full	0	6	99.98
	Below 2000	0	5	99.98
	Above 2000	0	4	99.97
CSC	Full	1	6	99.97
	Below 2000	1	4	99.93
	Above 2000	1	4	99.92
GluHCl	Full	0	6	99.98
	Below 2000	0	5	99.96
	Above 2000	0	5	99.99
$GluSO_4$	Full	0	6	99.94
	Below 2000	0	6	99.94
	Above 2000	0	5	99.97

Table 2
Class distances calculated from the cross validated models using the raw spectrum

	CSA	CSC	GluHCl	$GluSO_4$
Full spectrum				
CSA	***	3.636	3.387	2.253
CSC		4.952	***	3.867
GluHCl			***	2.163
GluSO4				***
Below 2000 cm^{-1}				
CSA	***	3.911	3.318	2.692
CSC		***	3.653	4.279
GluHCl			***	2.902
$GluSO_4$			***	
Above 2000 cm^{-1}				
CSA	***	3.363	3.44	1.98
CSC		***	5.729	3.679
GluHCl			***	1.668
$GluSO_4$				***

Asterisks signify N/A.

spectrum. Table 4 shows the class distances from the first derivative spectrum. Consistent with previous results, greater class distances were achieved with the use of the first derivative spectrum. The results showed that the different regions of the spectrum (full spectrum, below 2000 cm^{-1} and above 2000 cm^{-1}) provided roughly the same class distances. Compared to the raw spectrum the first derivative spectrums provided complete separation for all the models as opposed to partial separation of two pairs of samples seen when using the raw spectra. Across all regions of the spectrum the class

Table 3
The number of principal components and the 9%) variance explained for each cross validated model using the first derivative spectra

Sample	Spectral range	Outliers	#PC's	Variation explained (%)
CSA	Full	0	6	99.34
	Below 2000	0	6	99.31
	Above 2000	0	7	99.69
CSC	Full	0	2	67.22
	Below 2000	0	2	66.08
	Above 2000	0	6	98.65
GluHCl	Full	0	4	98.51
	Below 2000	0	4	98.72
	Above 2000	0	8	99.96
$GluSO_4$	Full	0	3	98.48
	Below 2000	0	6	99.80
	Above 2000	0	7	99.87

Table 4
Class distances calculated from the cross validated models using the first derivative spectra

	CSA	CSC	GluHCl	$GluSO_4$
Full spectrum				
CSA	***	15.397	13.31	10.31
CSC		***	9.523	3.763
GluHCl			***	6.099
$GluSO_4$				***
Below 2000 cm^{-1}				
CSA	***	15.185	12.148	7.816
CSC		***	9.595	3.756
GluHCl			***	6.118
$GluSO_4$				***
Above 2000 cm^{-1}				
CSA	***	15.526	15.08	15.145
CSC		***	7.76	3.937
GluHCl			***	5.557
$GluSO_4$				***

Asterisks signifies N/A.

distances between chondroitin sulfate C and glucosamine 6 sulfate were the lowest.

4. Conclusion

These results showed the ability of FTIR and chemometrics to be used as technique for the classification of chondroitin and glucosamine samples. It was concluded that the samples were best classified using the first derivative spectra. This use of first derivative spectra reduced the variability of the samples in cases, where the spectra had the similar features but varied in concentration this allowed for better classification of the samples. While each of these techniques can be used by itself it is advisable to use them together as in this way they can provide a more complete picture. This work has shown that techniques such as PCA, HCA and SIMCA are able to distinguish between spectra of samples, making this a fast and simple technique for the evaluation of these compounds as raw materials in the pharmaceutical and natural products industry. Further work in this area could include the development of methods to determine the quantity of the active ingredient in finished products.

References

[1] H.J. Mankin, K.D. Brandt, L.E. Shulman, J. Rheumatol. 13 (1983) 1130–1160.
[2] R. Jurmain, Stories from the Skeleton. Behavioral Reconstruction in Human Osteology, Gordon and Breach science publishers, Amsterdam, The Netherlands, 1999.
[3] N. Ishiguro, T. Kojima, A.R. Poole, Nagoya J. Med. Sci. 63 (2002) 73–84.
[4] C. Todd, J. Am. Pharm. Assoc. (Wash) 42 (2002) 74–82.
[5] C.C. da Camara, G.V. Dowless, Ann. Pharmacother. 32 (1998) 580–587.
[6] W. Adbelfattah, T. Hammad, JANA 3 (2001) 16–23.
[7] M. O'Rourke, Nurse Pract. 26 (2001) 44–52.
[8] T.E. Hardingham, A.J. Fosang, J. Dudhia, in: K.E. Kuettner, R. Schleyerback, J.G. Peyron, V.C. Hascall (Eds.), Articular Cartilage and Osteoarthritis, Raven Press, Wiesbaden, Germany, 1991, pp. 5–20.
[9] P. Morreale, R. Manopulo, M. Galati, L. Boccanera, G. Saponati, L. Bocchi, J. Rheumatol. 23 (1996) 1385–1391.
[10] A. Baici, D. Horler, B. Moser, H.O. Hofer, K. Fehr, F.J. Wagenhauser, Rheumatol. Int. 12 (1992) 81–88.
[11] M.S. Cappell, J.R. Schein, Gastroenterol. Clin. North Am. 29 (2000) 97–124.
[12] G.S. Kelly, Altern. Med. Rev. 3 (1998) 27–39.
[13] F. Cabassi, B. Casu, A.S. Perlin, Carbohydr. Res. 63 (1978) 1–11.
[14] B. Casu, G. Scovenna, A.J. Cifonelli, A.S. Perlin, Carbohydr. Res. 63 (1978) 13–27.
[15] R. Servaty, J. Schiller, H. Binder, B. Kohlstrunk, K. Arnold, Bioorg. Chem. 26 (1998) 33–45.
[16] R. Servaty, J. Schiller, H. Binder, K. Arnold, Int. J. Biol. Macromol. 28 (2001) 121–127.
[17] M.A. Coimbra, F. Goncalves, A.S. Barros, I. Delgadillo, J. Agric. Food Chem. 50 (2002) 3405–3411.
[18] S. Sivakesava, J. Irudayaraj, J. Food Sci. Technol. 37 (2002) 351–360.
[19] S.B. Engelsen, L. Norgaard, Carbohydr. Polym. 30 (1996) 9–24.
[20] M. Foot, M. Mulholland, L. Kirkup, Chromatographia 58 (2003) 343–348.
[21] S. Wold, Pattern Recognit. 8 (1976) 127–139.
[22] R.G. Brereton, Multivariate Pattern Recognition in Chemmetrics Illustrated by Case Studies, Elsevier, Amsterdam, 1992.
[23] S. Honda, H. Yuki, K. Takiura, J. Biochem. 76 (1974) 209–211.
[24] A. Kher, M. Mulholland, B. Reedy, P. Maynard, Appl. Spectrosc. 55 (2001) 1192–1198.
[25] H. Panayiotou, S. Kokot, Anal. Chim. Acta 392 (1999) 223–235.