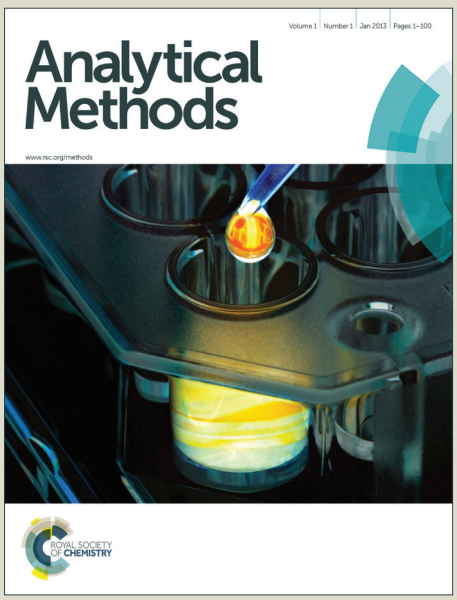
Analytical Methods

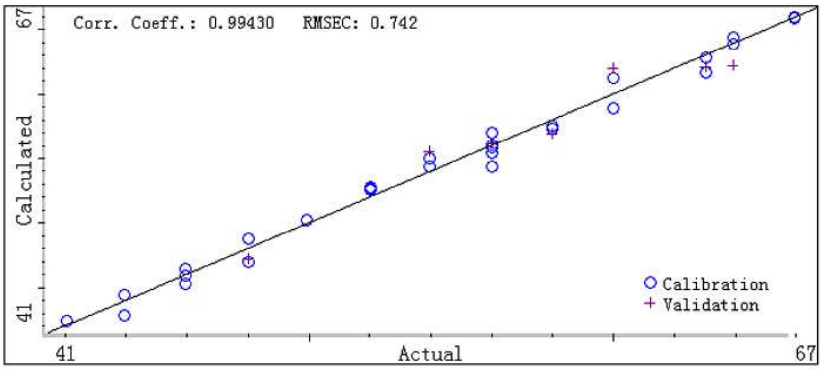
Accepted Manuscript



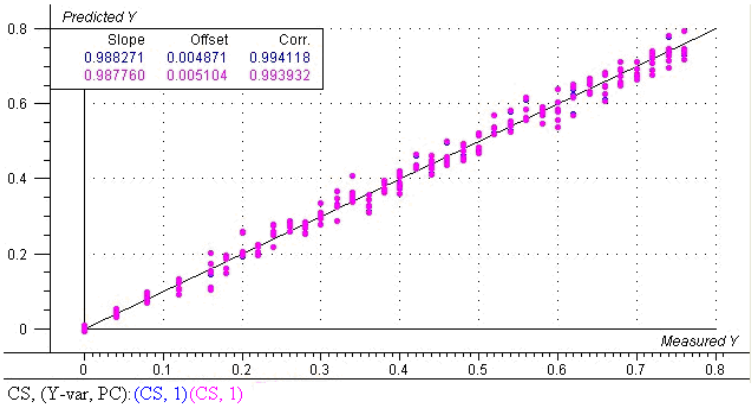
ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/methods

NIR can obtain high accuracy within a wider concentration range. Raman can obtain relatively high accuracy only within a narrower concentration range.



Raman



NIR

Determination of chondroitin sulfate in tablet by Raman spectroscopy and near infrared spectroscopy combined with chemometrics methods

Xiumei Liu[a*], Lian Li[a], Ting Zhao[a], Haiping Dong[b]

[a] *School of Pharmaceutical Sciences, Shandong University, Jinan250012, Shandong , P. R. China*

[b] *Jinhongli Apparatus Company of Jinan, Jinan250012, Shandong , P. R. China*

[*] Corresponding author. Tel.: +86-531-88382013; fax: +86-531-88382548.
E-mail address: liuxium@sdu.edu.cn (X.M. Liu).

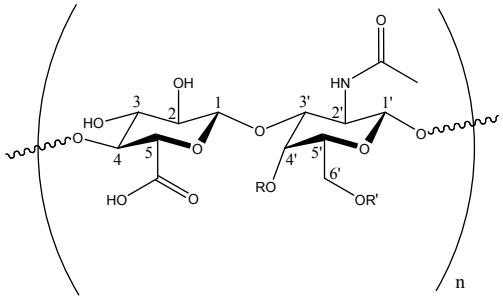
Abstract

Chondroitin sulfate (CS) is one type of acidic mucopolysaccharides, which consist of repeating disaccharide units of glucuronic acid and galactosamine. Since CS has no UV chromophore, it is usually detected by the terminal absorption due to the N-acetyl function at a wavelength of 200 nm, resulting in lower sensitivity. Raman spectroscopy (Raman) and near infrared spectroscopy (NIR) coupled with partial least squares (PLS) can provide rapid, simple, reproducible and non-destructive quantitative analysis of CS, and no sample pre-treatment and pre-separation are required. In this study, we predicted the CS content in tablet using Raman and NIR combined with PLS approaches. Our results showed that the predicted values obtained by NIR were in good agreement with the real values, and the correlation coefficient (Corr. Coeff.) was 0.994. In Raman spectroscopy studies, when the CS content in tablet was in the range of 7%-39%, the Corr. Coeff. and root mean square error of calibration (RMSEC) were 0.998 and 0.578, respectively. When the CS content in tablet was in the range of 41%-67%, the Corr. Coeff. and RMSEC were 0.994 and 0.742, respectively. Therefore, high accuracy could be achieved within a wider concentration range when using NIR, whereas a relatively high accuracy could be obtained only within the certain concentration range when using Raman.

Keywords Chondroitin sulfate; Raman spectrum; Near infrared spectrum; Partial least squares

1. Introduction

Chondroitin sulfate (CS) is one type of acidic glycosaminoglycans (GAGs) consisting of variable number of repeating disaccharide units of alduronic acid and aminohexose, extracted from cartilage of animals. It has been used as drugs for treating neuralgia, arthritis, tinnitus, canker, hyperlipemia and so on.[1,2] The outcome of treatments is directly related to the quality of CS preparations. The structural characteristic of CS include heterogeneity of molecular mass and charge density due to different sites and degree of sulfation, much types depending on the extraction source, such as chondroitin sulfate A (CS-A, 4-sulfation in N-acetyl-D-galactosamine), chondroitin sulfate B (CS-B, C-5 epimerization to iduronic acid), chondroitin sulfate C (CS-C, 6-sulfation in N-acetyl-D-galactosamine) etc and no UV or fluorescence chromophores and so on. The structure of one disaccharide unit of CS is shown in Fig.1. So it is difficult to quantify the intact CS using common spectrum and chromatographic methods. Therefore, to develop reliable and accurate methods of CS content determination is exclusively important for the quality control of CS preparations.



CS

Fig. 1. The structure of one disaccharide unit of CS. CS-A: R=SO_3H, R'=H; CS-C: R=H, R'=SO_3H; CS-B: C-5 epimerization to iduronic acid.

Currently, spectrophotometry,[3,4] chromatography[5-11] and electrochemistry method[12] are the most commonly used methods for crude CS quantification. Spectrophotometry has a complex operation process and is often affected by many factors, leading to a

poor reproducibility. Chromatography, such as capillary electrophoresis (CE)[9], has a strong baseline noise due to the end absorption at 200 nm with a tailing peak or wide peak, resulting in an imprecise quantification. While strong-anion exchange (SAX)-HPLC[7] is used for analysis unsaturated disaccharides produced by the action of chondroitin ABC lyase, the analytical process is complicated and reagent is expensive. Although with a high sensitivity, electrochemistry method has a lot of strict determination conditions and difficulties in operations, and its specificity is not good. Therefore, it is necessary to develop an appropriate and precise method for the quantitative analysis of CS.

Both Raman spectroscopy (Raman) and near infrared spectroscopy (NIR) are branches of vibrational spectroscopy and have been applied in many areas of analytical chemistry today. They can provide rapid, simple, reproducible and non-destructive qualitative and quantitative analysis, and no sample pre-treatment and pre-separation are required.

Raman is based on the scattering of light from near infrared or visible radiation due to the vibrational energy of the molecules in the sample. It has several general advantages, such as non-interference of water in the sample, ease of sampling and measurement, and minimal fluorescence interference.[13,14] Because the Raman signal is scattering spectrum signal, the Raman spectrum can be affected by the molecular structure and stability of light irradiation of analyte. Meanwhile, the optical path of irradiation can also affect the determination. Sometimes, the peaks of several analytes can overlap, leading to an inaccurate quantification. Therefore, it is necessary to establish the proper calibration and validation procedures with data acquisition protocols for Raman methods. Chemometrics is often used in Raman spectrum for quantitative analysis in order to extract the information from the complex spectra containing overlapping absorption peaks, interference effects and instrumental artifacts. Feld et al.[15] quantitatively analyzed the histochemical composition of human artery using NIR Raman spectroscopy. They found that the Raman signal behaves linearly with the component concentration, even in a highly scattering medium such as in powder. The calculated fit coefficients from the spectra are in agreement with the

measured values within experimental uncertainties.

NIR technique can quantitatively analyze one or several components in a sample using the optical property of analyte in near infrared spectrum. It has several advantages, such as speediness, simplicity, non-destructiveness and pollution-free. Moreover, NIR is a simultaneous multi-component analytical method, and it can also determine a single chemical compound among a great number of other substances, especially by means of NIR spectrophotometer combined with chemometrics methods. In addition, the original NIR spectrum is difficult to analyze because the NIR spectrum band is wide and often overlaps with other bands. Therefore, it is necessary to obtain available information by chemometrics.

The construction of calibration model is the basis of quantitative analysis by spectrum. The stand or fall of calibration model directly affects the quantitative analysis. Generally, quantitative analysis includes several steps,[16] and calibration set must be also accorded several requirements. For a simple sample, samples in the validation set can be directly prepared. The concentration range of samples in the validation set should cover 95% of those in the calibration set.[17]

Pre-treatment of crude spectral data is important because the NIR or Raman spectra are often affected by the instrumental variation and measurement conditions, resulting in background noise and baseline drift. There are different spectral pre-treatment methods, among which Savitzky-Golay (S-G) filter is effective for smoothing high-frequency noise and elevating signal-to-noise ratio. The first derivative can eliminate the translation and baseline drift, wipe out the interference of other background, discriminate the overlapping peaks and improve the resolution and sensitivity of spectra. Second derivative makes it easier to see the peak feature in the raw spectrum. Both standard normal variate (SNV) and multiple scatter correction (MSC) can remove the slope variation and correct the light scatter due to different particle sizes.[18,19]

The most commonly used multivariate statistical methods in spectrum analysis include principal component analysis (PCA), principal component regression (PCR), partial least squares (PLS) and so on. Data compression, calibration and validation are

the basis of these methods.[20,21] For a successful application of these methods, certain factors should be taken into consideration, such as the proper selection of spectral range, stability of the spectra and the number of variables employed in the calibration model.

At present, NIR is a relative mature technique and is applied in qualitative analysis, quantitative analysis and on-line quality control for polysaccharide.[22-26] While Raman spectrum is deemed to only do semiquantitative analysis for a long time. Recently, Raman spectrum was also used for quantitative analysis and less applications have been reported.[27-29] Mrozek et al.[27] analyzed oligosaccharides using Raman spectrum and PLS, and they obtained good results with an average error of less than 2.7%. However, no one has reported the quantitative analysis of CS using Raman approach. The aim of the study is to determine the CS content in CS tablets using Raman and NIR approaches and to validate the accuracy and suitability of Raman approach in quantitative analysis by means of comparison of the results obtained by Raman and NIR approaches. The main contents in this study include: investigate the potentials of Raman spectroscopy; characterize the preparation with different excipients; quantify the CS content in tablets; develop the calibration and validation models for predicting the CS content in unknown samples; and compare the results obtained by Raman and NIR.

2. Materials and methods

2.1. Chemicals

CS from shark cartilage (purity > 92.0%) was purchased from Chongqing Imperial Bio-Chem. Co., Ltd. (Chongqing, China). Soluble starch (SS), magnesium stearate (MS), talcum (T), dextrin (D) and crystallite cellulose (CC) were provided by Tianjin Bodi Chemical Ltd. (Tianjin, China).

2.2. Sample preparation and partition

A total of 45 groups of samples with different compositions were prepared according to the general tablet composition. Furthermore, 35 groups of them were used for the calibration set (Table 1), and the remaining was used for the validation set (Table 2). In order to investigate the effect of different backgrounds on spectrum analysis, the excipients in samples 5* and 6* in validation set are different from excipients in other samples. The constituents of samples were precisely weighed. The mixtures were ground into a fine powder in a mortar, filtered to a fineness of 80 mesh and then tableted. The CS content in the calibration set and validation set was in ranges of 0%-76% and 20%-65% (increase in orderly), respectively.

2.3. Apparatus and parameters

Manual and single punch tablet press with 80-mesh fineness was provided by Shandong Medical Appliance Factory (Shandong, China).

Raman spectra were recorded using a LabRAM HR UV-800 NIR Confocal Laser MicroRaman spectrometer **(**HORIBA Jobin Yvon, France). He-Ne laser operating at

Table 1 The content percentage of every constituent in CS tablets in the calibration set

Sample number	CS (%)	CC (%)	MS (%)
1	0.000	88.34	11.66
2	4.029	85.25	10.72
3	8.009	81.62	10.37
4	12.00	78.32	9.680
5	16.00	74.92	9.080
6	18.00	73.33	8.670
7	20.02	70.94	9.040
8	22.00	69.29	8.710
9	24.00	67.46	8.540

10	26.00	65.01	8.990
11	28.00	63.91	8.090
12	29.99	62.14	7.870
13	32.00	60.50	7.500
14	34.02	58.61	7.370
15	36.00	56.78	7.220
16	37.90	54.98	7.120
17	39.97	53.29	6.740
18	41.95	50.02	8.030
19	43.94	44.99	11.07
20	45.94	48.35	5.710
21	47.99	46.68	5.330
22	49.95	43.36	6.690
23	52.02	42.65	5.330
24	53.97	40.82	5.210
25	56.00	39.15	4.850
26	57.96	37.34	4.700
27	60.00	35.49	4.510
28	63.02	34.42	2.560
29	63.94	32.02	4.040
30	65.96	30.19	3.850
31	67.97	28.67	3.360
32	69.93	26.86	3.210
33	71.96	25.52	2.520
34	73.93	23.04	3.030
35	76.00	20.98	3.020

167

Table 2 The content percentage of every constituent in CS tablets in the validation set

Sample number	CS (%)	CC (%)	MS (%)
1	20.02	71.64	8.340
2	25.00	67.82	7.180
3	30.01	64.31	5.680
4	34.98	57.18	7.840
5*[a]	40.00	54.48	5.520
6*[b]	44.99	46.99	8.020
7	49.98	45.66	4.360
8	54.99	40.17	4.840
9	60.00	35.50	4.500
10	64.99	31.16	3.850

[a] the excipients in 5*sample are CC, MS and T, (mass ratio, CC:MS:T =54.48:4.14:1.38)

[b] the excipients in 6* samples are CC, MS and T, (mass ratio, CC:MS:T =46.99:6.015:2.005)

632.81 nm with a power of 17 mW was used as the excitation source, and the laser beam was focused to a spot of approximately 100 μm diameter. The exposure time was 20 s, and the aperture was 400 μm equipped with long-focus lens of 50 times. Raman spectra were obtained in the range of 500 ~ 4,000 cm^{-1}. The system was operated using the TQ Analyst software (Thermo Nicolet), and the experiments were performed in triplicate.

NIR measurement was carried out using Brimrose Luminar 5030 AOTF-NIR spectrometer (Brimrose Co., USA) with an InGaAs detector. NIR spectra were collected at about 25℃, and the humidity was well controlled in the laboratory. Each spectrum was the average of 300 scans with a wavelength increment of 2 nm over the wavelength range of 1,100-2,300 nm. The spectrum data were analyzed by the Unscrambler analytical software.

2.4. Methods

2.4.1. Raman spectrum analysis

A total of 35 groups tablets were scanned. Tablets with the same component and content were scanned in triplicate. In order to investigate the precision, the tablets with 56% CS content were scanned in sextuplicate. So 108 (34×3+1×6) tablets were scanned, corresponding to 108 samples. The calibration model was constructed in following steps: (1) sample selection, which was conducted according to leverage and studentized residual; (2) spectrum pre-treatment, the methods used for spectrum pre-treatment included derivative, spectrum smoothness and MSC. Derivative can eliminate the baseline drift and strengthen the spectrum band character. Spectrum smoothness can improve the signal-to-noise of analysis signal. MSC is often used to diffuse the reflection spectrum, and it can reduce the difference of spectrum and reserve the spectrum information related to the chemical ingredient; (3) selection of spectral range, which can reduce the reference of noise from the irrelevant interzone spectrum; (4) selection of main factor number, which can avoid the fit-not-enough or overfit; (5) construction of mathematical model; and (6) model estimation and optimization. The calibration set and validation set were stochastically selected by the TQ Analyst software.

2.4.2. NIR spectrum analysis

NIR spectra were obtained from 45 groups CS tablets, among which 35 groups were used as the calibration set and 10 groups were used as the validation set. Every surface (side) of one tablet was scanned once, and three CS tablets with the same component and content were scanned. Therefore, the same CS content produced six spectra and 270 (45×6=270) spectra were obtained, corresponding to 270 samples. Moreover, the averaged spectrum was calculated. The diffuse reflectance spectra were collected in a ratio mode, which can reduce the effect caused by the change of background. Leave-one-out cross-validation strategy was used to identify the optimum factors during the model development.

In Raman and NIR quantitative analysis, the PLS was used to construct models. In addition, the first derivative was used for spectrum pre-treatment, and the optimum

number of calibration factors was selected based on the root mean square error of cross validation (RMSECV). Cross validation was used to estimate the performance of the developed models.

3. Results and discussion

3.1. Quantitative analysis of CS tablet by Raman spectrum

5 samples were ignored, 84 samples were used as calibration set and 29 samples were used as validation set. The correlative parameters of calibration models are showed in Table 3. The results suggested that the calibration model was bad when the CS concentration of samples in the calibration set ranged between 0.0 and 76.0%. Therefore, 35 groups tablets were divided into two groups, with CS concentration ranges of 0%-40% and 41%-76%, respectively. Accordingly, two models, model A with a CS content of 41%-67% and model B with a CS content of 7%-39%, were constructed in order to obtain good predicted results. The correlative parameters are listed in Table 3.

Table 3 The correlative parameters of models of Raman and NIR analysis

	C_{CS}[a]	N_1[b]	N_2[c]	Method of treatment	PLS factors	R_c[d]	RMSEC[f]	RMSECV	R_p[e]	RMSEP[g]
Raman	0-76%	84	29	MSC, 1st, S-G7[h]	3	0.780	13.4	14.9	-	-
	7-39%	23	6	MSC, 1st, S-G7	9	0.998	0.578	3.79	0.760	2.56
	41-67%	29	7	MSC, 1st, S-G7	6	0.994	0.742	3.95	0.966	1.42
NIR	0-76%	210	60	1st, S-G9	1	0.994	2.28	0.994	0.999	8.45×10^{-3}

[a] C_{CS}, the percentage concentration of CS

[b] N_1, the number of samples in calibration set

[c] N_2, the number of samples in validation set

[d] R_c, correlation coefficient of calibration model

[e] R_p, correlation coefficient of validation model

[f] RMSEC, root mean square error of samples in calibration set

[g] RMSEP, root mean square error of samples in validation set

[h] S-G7, Savitzky-Golay smoothing with 7 points

3.1.1. Model A

3.1.1.1. Outlier detection

An important step in building a PLS model is the identification of outliers because PLS calibration method is strongly influenced by the presence of outliers.[30] Studentized residual and leverage methods are usually used to detect and remove outliers.[31-34] Those samples which have higher studentized residual or leverage or which have obvious difference from others are regarded as outliers. Fig.2 shows the selected 29 samples whose concentration percentage range of CS were from 42% to 66% in calibration set after removing the outliers.

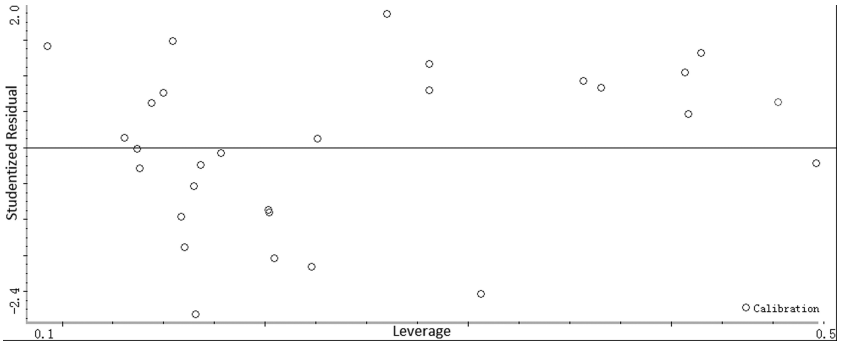


Fig. 2. Outlier detection. 29 samples were selected and those dots that have higher studentized residual or leverage had been deleted.

3.1.1.2. Spectrum treatment

Fig. 3 shows the original spectra of the selected 29 samples. We found that good results could be obtained after the spectra were dealt with MSC, first-order

derivative and S-G smoothing with 7 points (S-G7, the polynomial order: 3). Fig. 4 shows the spectra after dealt with MSC, first-order derivative and S-G7.

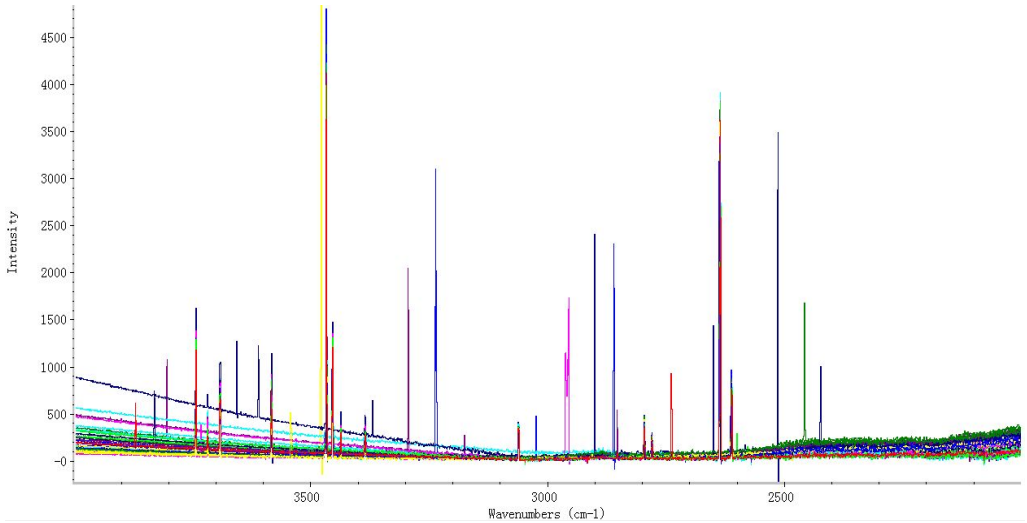


Fig.3. The Raman spectra of 29 samples with CS percentage concentration ranges between 41% and 67% in calibration set.

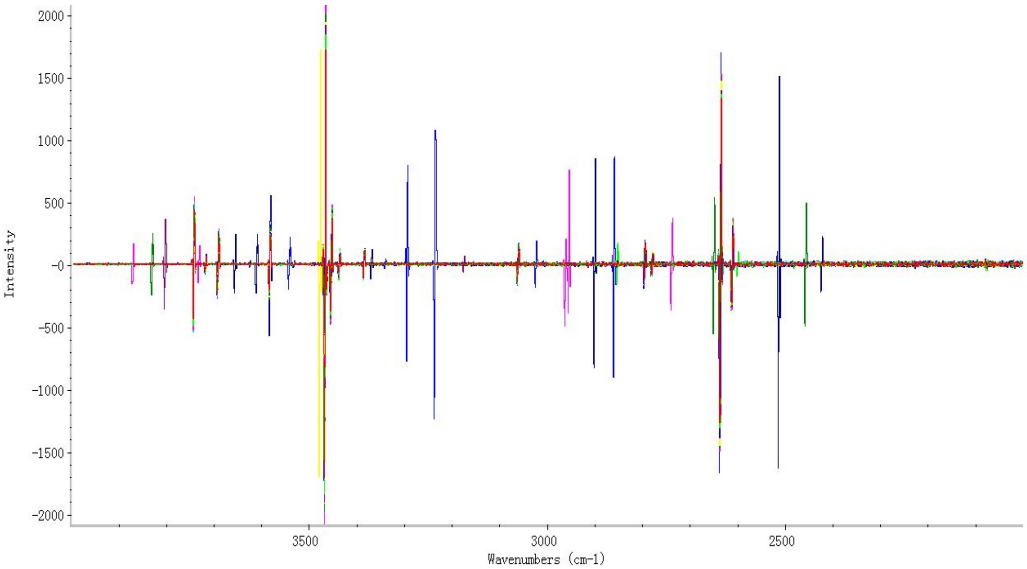


Fig.4. The pre-treatment spectrum via MSC treatment, first-order derivative and S-G7 smoothing.

3.1.1.3. The selection of main factor number

The main factor is also named the main component, it is very important to select the main factor number (n) for the accuracy of calibration model. If n is too little, there will lose more useful information of original spectrum, resulting in underfit. If n is too much, there will include more noise of original spectrum, resulting in overfit.

Appropriate n not only can make the most of spectrum information but also eliminate noise.[16] The proper number of main factors was selected by the leave-one-out cross-validation method, and RMSECV was used to optimize the parameters. When a least value of RMSECV is obtained, the corresponding number of main factors is optimum. In this study the optimum number of main factors was 6.

3.1.1.4. The selection of spectral region

The spectral regions of 2,616-2,643 cm^{-1} and 2,800-2,780 cm^{-1} were automatically selected by the software. The calibration model was constructed based on the selected spectrum range.

3.1.1.5. Model construction and evaluation

In order to characterize the prediction ability of a created PLS model, the correlation coefficient of calibration model (R_c), the correlation coefficient of validation model (R_p), root mean square error of calibration (RMSEC) and root mean square error of prediction (RMSEP) were used for estimation.[30] Fig. 5 shows the pertinence between predicted results obtained by calibration model and real values of percentage concentration of CS between 41% and 67%. Samples in the calibration set were distributed in two sides of tropic, R_c and RMSEC were 0.994 and 0.742, respectively. Samples in the validation set were also distributed in two sides of tropic, R_p and RMSEP were 0.966 and 1.42, respectively. The results suggested that the predicted values of samples in the validation set were consistent with real values. Table 4 lists the predicted values and real values of samples with a CS content range of 41%-67%.

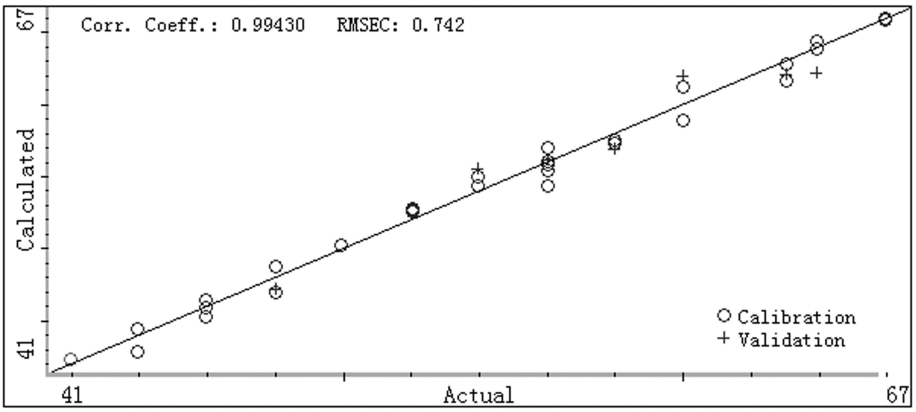


Fig. 5. The pertinence between predicted results obtained by calibration model and real values of CS contents with percentage concentration range of 41%-67%.

Table 4 The predicted values obtained by Raman and the real values of samples within the CS content range of 41%-67%

Real value (%)	Predicted value (%)	Absolute error	Relative error
41.95	42.31	0.36	0.86%
43.94	44.48	0.54	1.08%
43.94	42.88	-1.06	-2.41%
45.94	45.89	-0.05	-0.11%
45.94	46.47	0.53	1.15%
45.94	45.26	-0.68	-1.48%
47.99	48.75	0.76	1.58%
47.99	47.02	-0.97	-2.02%
47.99*	47.18	-0.81	-1.69%
49.95	50.23	0.28	0.56%
52.02	52.60	0.58	1.11%
52.02	52.64	0.62	1.19%
52.02	52.80	0.78	1.50%
53.97*	55.58	1.61	2.98%
53.97	55.00	1.03	1.91%
53.97	54.41	0.44	0.82%
56.00	55.38	-0.62	-1.11%
56.00	54.37	-1.63	-3.00%
56.00	57.05	1.05	1.88%
56.00	56.10	0.10	0.18%
56.00	55.83	-0.17	-0.30%
56.00*	56.19	0.19	0.34%
57.96	57.59	-0.37	-0.64%

57.96	56.93	-1.03	-1.78%
57.96	57.37	-0.59	-1.02%
60.00	58.88	-1.12	-1.87%
60.00	61.23	1.23	2.05%
60.00*	61.95	1.95	3.25%
63.02	62.81	-0.21	-0.33%
63.02*	62.09	-0.93	-1.48%
63.02	61.71	-1.31	-2.08%
63.94	64.45	0.51	0.80%
63.94*	62.18	-1.76	-2.75%
63.94	63.93	-0.01	-0.02%
65.96	65.84	-0.12	-0.18%
65.96	66.05	0.09	0.00%

* the samples in validation set

3.1.2. Model B

The same operation process and treatment methods used in model A were applied to model B. The optimum number of main factors was 9 for model B. The wave bands of spectra automatically selected by the software were 2,779-2,460 cm^{-1}, 3,476-3,463 cm^{-1} and 3,777-3,742 cm^{-1}. Fig. 6 shows the pertinence between predicted values obtained by calibration model and real values of percentage concentration of CS between 7% and 39%. The R_c was 0.998 and the RMSEC was 0.578, R_p and RMSEP were 0.760 and 2.56, respectively. The results suggested that the predicted values of samples were consistent with real values. Table 5 lists the predicted values and real values of samples with a CS content range of 7%-39%.

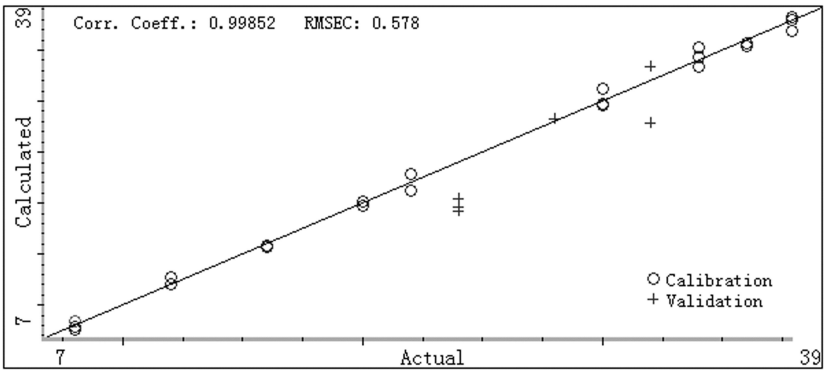


Fig. 6. The pertinence between predicted values obtained by calibration model and real values of CS contents with percentage concentration range of 7%-39%.

Table 5 The predicted values obtained by Raman and the real values of samples within the CS content range of 7%-39%

Real value (%)	Predicted value (%)	Absolute error	Relative error
8.01	8.48	0.47	5.87%
8.01	7.54	-0.47	-5.87%
8.01	7.85	-0.15	-1.87%
12.00	12.09	0.09	0.75%
12.00	12.07	0.07	0.58%
16.00	15.79	-0.21	-1.31%
16.00	15.73	-0.27	-1.69%
16.00	15.84	-0.16	-1.00%
20.02	20.13	0.11	0.55%
20.02	19.77	-0.25	-1.25%
22.00	22.86	0.86	3.91%
22.00	21.18	-0.82	-3.73%
24.00*	21.55	-2.45	10.2%
24.00*	21.24	-2.76	11.5%
24.00*	21.12	-2.88	12%
28.00*	28.32	0.32	1.14%
29.99	31.19	1.20	4.00%

29.99	29.57	-0.42	-1.40%
29.99	29.71	-0.28	-0.93%
32.00*	29.24	-2.96	-9.25%
32.00*	33.38	1.38	4.31%
34.02	33.38	-0.64	-1.88%
34.02	34.30	0.28	0.82%
34.02	35.26	1.24	3.64%
36.00	37.47	-0.53	-1.47%
36.00	35.68	-0.32	-0.89%
37.70	38.31	0.41	1.09%
37.70	38.03	0.13	0.34%
37.70	36.84	-1.06	-2.81%

* the samples in validation set

3.1.3. Discussion

The interference experiment showed that the excipients in CS tablet exerted a serious interference to Raman spectrum. Therefore, it was impossible to quantitatively analyze the CS content by routine methods without the pre-treatment. However, the quantitative analysis for CS tablet could be directly performed without the complicated pre-treatment process by Raman spectrum combined with PLS modeling method.

The relative errors of predicted values obtained by models A and B were less than 4% and 12%, respectively. It suggested that the quantitative analysis by Raman spectrum combined with PLS was quite accurate in the certain concentration range. The relative standard deviation (RSD) of six precited values for 56% CS content was 1.64%.

However, Raman spectrum also had some shortcomings when analyzing solid mixture. For example, the uniformity of solid samples is highly required by Raman spectrum. Bad sample uniformity results in poor reproducibility and much exceptional samples.

3.2. Quantitative analysis of CS tablet by NIR spectrum

3.2.1. Spectrum treatment

Fig. 7 shows the original spectra of all 270 (45×6) samples between 1,100-2,300 nm. The band at 1,900-1,950 nm was caused by the strong absorption of water.[35] Light scattering caused by different particle sizes and densities affected the raw spectra, resulting in the baseline drift. The spectra were more tightly arranged, suggesting that the comparability among spectra was good.

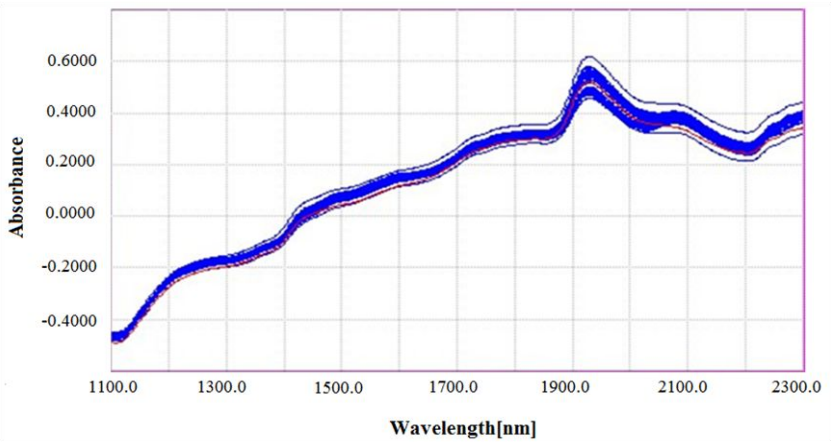


Fig. 7. The original NIR spectra of 45 groups CS tablets, total 270 samples. Red line: a selected spectrum, has not specific meaning.

In order to eliminate the effect of noise and baseline, the original spectra need to be pretreated prior to the model construction. The adopted pre-treatment methods included first-order derivative and S-G9 smoothing treatments. First-order derivative treatment can eliminate the baseline excursion and drift caused by the color difference of samples. Fig. 8 shows the spectra after the pre-treatment.

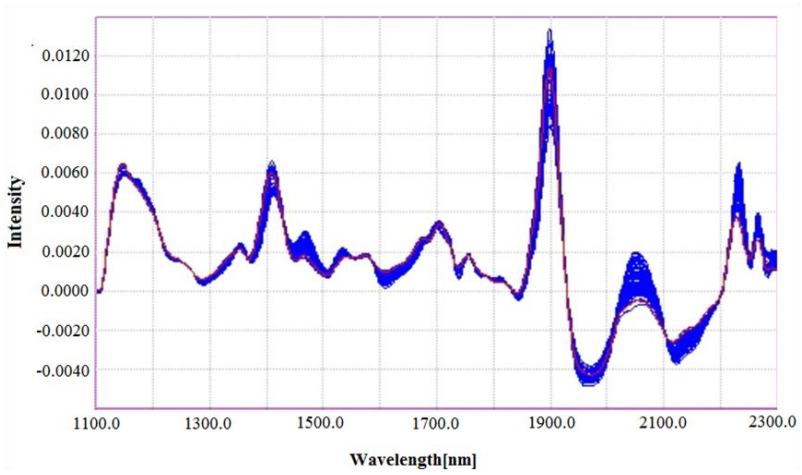


Fig. 8. The spectra obtained after treatment by first-order derivative and S-G9 smoothing. Red line: a selected spectrum, has not specific meaning.

3.2.2. Construction of PLS model

PLS method was used to construct the pertinence between the spectrum data after the pre-treatment and sample content. Fig. 9 shows the perfect calibration model obtained by the quantitative analysis software of Unscrambler. The blue dots were from 210 samples (35×6) and were used as calibration set, R_c=0.9941. The red rots were also from the 210 samples and were used as leave-one-out cross-validation set, RMSECV=0.9939. So these blue and red rots were almost overlapped completely. The predicted values of all samples in calibration model had good pertinence with the real CS content, and correlative parameters of model are showed in Table 3.

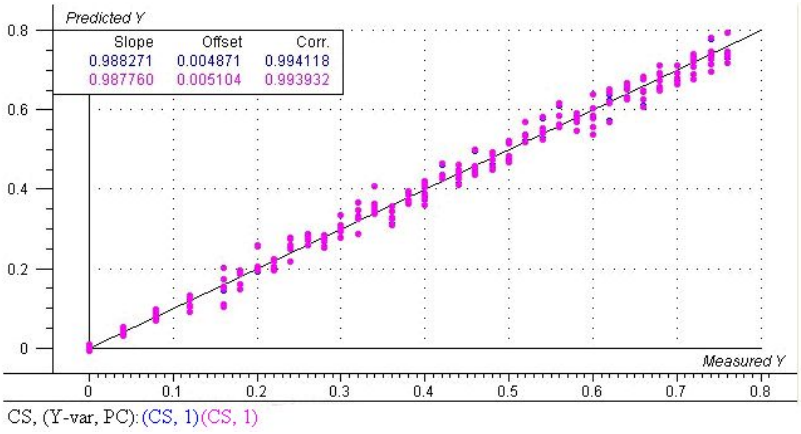


Fig. 9. The PLS regression model of CS tablet. Blue dots: calibration set, R_c=0.9941, red dots: leave-one-out cross-validation set, RMSECV=0.9939. (CS,1): the main factor number of PLS model is 1.

3.2.3. Exterior validation

A total of 10 groups samples in the exterior validation set were predicted using the constructed calibration model (Table 6). Because each CS content has six spectra, the predicted value of each CS content is the average value of six predicted values. The results suggested that predicted values are very near to real values except for two samples of 5* and 6*, the R_p is 8.45×10^{-3}.

Table 6 The predicted average values obtained by NIR and real values of 10 groups samples in the validation set

Sample number	Predicted value	Real value	Relative error (%)	Average error (%)
1	0.20	0.20	0.00	
2	0.26	0.25	4.00	
3	0.30	0.30	0.00	
4	0.34	0.35	2.86	
5*[a]	*0.42*	*0.40*	*5.00*	1.54
6*[a]	*0.49*	*0.45*	*8.89*	
7	0.51	0.50	2.00	
8	0.54	0.55	1.82	
9	0.59	0.60	1.67	
10	0.65	0.65	0.00	

[a] the excipients in 5* and 6* are different from those in other samples in validation set.

3.2.4. Discussion

According to the modeling and predicted results, the NIR method is better than the Raman method because the NIR spectrum can eliminate the effect of tablet uniformity to a certain extent. Table 3 lists the correlative parameters of Raman and NIR analysis.

The average relative error of the 8 predicted values (except for two samples of 5* and 6*) was 1.54%. Large errors were observed from the samples of 5* and 6* (Table

6). The main reason could be that the accessory materials of these two samples were different from those of other samples in the calibration set. Therefore different background have large effect on the precited ability of PLS model.

4. Conclusion

Raman spectrum combined with PLS is a new analytical technique for quantitative analysis of GAG. This method requires no the sample pre-treatment and pre-separation, and the obtained results have good accuracy within a relatively narrow concentration range. Our data showed that the uniformity of tablet and the stability of experimental conditions have great effect on predicted results. The most promising finding in this study is that Raman spectroscopy can also be used to detect the CS content in tablet. Prediction can be improved if standard calibration and validation models are developed for samples within a narrow concentration range.

NIR is a relatively mature method for the quantitative analysis, and its determination is quite accurate in a very wide concentration range of given compounds. However, in order to obtain the high accuracy, the premise is that the chemical components of samples in the calibration set must be consistent with those of samples in the validation set. Therefore, NIR is very suitable for the sample with known chemical components, which is very important for on-line quality control of drugs in enterprises.

NIR and Raman spectroscopy coupled with multivariate calibration are promising techniques for the quantitative analysis of GAG. Furthermore, both methods are well suited for rapid screening procedures, by which a large number of samples can be quickly evaluated. NIR is a rapid, non-destructive and fluorescence-insensitive technique. However, water in the sample can affect the NIR spectrum. Raman spectroscopy is also a rapid and non-destructive method, but it is sensitive to the fluorescence in sample or sample container. Therefore, NIR and Raman spectroscopies are complementary methods for the quantitative analysis.

Acknowledgement

This work was financially supported by " National Natural Science Foundation of China (No. 21205069)", "the Independent Innovation Fund of Shandong University (No. 2012TS101)" and "the Doctoral Program of Higher Education of Special Research Foundation (The Class of New Teacher) (No. 20110131120039)".

References

[1] H.X. Yang, Y. Liu, C.X. Liu, S.M. Yu, Research progress in the detection methods of chondroitin sulfate, Chin. J. Biochem. Pharm., 2009, **30**, 58-60.

[2] C.J. Malavaki, A.P. Asimakopoulou, F.N. Lamari, A.D. Theocharis, G.N. Tzanakakis, N.K. Karamanos, Capillary electrophoresis for the quality control of chondroitin sulfates in raw materials and formulations, Anal. Biochem., 2008, **374**, 213-220.

[3] Z. Lu, Y.N. Yang, Content determination of chondroitin sulfate in pig by carbazole spectrophotometry, J. Anhui Agfi. Sci., 2008, **36**, 14368 and 14371.

[4] H. Gao, K. Liu, Z.D. Yu, F.S. Geng, Spectrophotometrlc determinatiun of chondroitin sulfate with phlorglucinol, Chin. J. Biochem. Pharm., 2000, **21**, 247-248.

[5] D.W. Choi, M.J. Kim, H.S. Kim, S.H. Chang, G.S. Jung, K.Y. Shin, S.Y. Chang, A size-exclusion HPLC method for the determination of sodium chondroitin sulfate in pharmaceutical preparations, J. Pharm. Biomed. Anal., 2003, **31**, 1229-1236.

[6] T. Tyler, B. Khandelwal, D. Norden, F.R. Rolle, Determination of chondroitin sulfate in raw materials by liquid chromatography, J. AOAC Intl., 2002, **85**, 567-571.

[7] N. Volpi, Analytical aspects of pharmaceutical grade chondroitin sulfates, J. Pharm. Sci., 2007, **96**, 3168-3180.

[8] F.N. Lamari, R. Kuhn, N.K. Karamanos, Derivatization of carbohydrates for chromatographic, electrophoretic, and mass spectrometric structure analysis, J. Chromatogr. B, 2003, **793**, 15-36.

[9] X.M. Liu, C.X. Sun, H.C. Zang, W.H. Wang, R.C. Guo, F.S. Wang, Capillary electrophoresis for simultaneous analysis of heparin, chondroitin sulfate and hyaluronic acid and its application in preparations and synovial fluid, J. Chromatogr. Sci., 2012, **50**, 373-379.

[10] N. Volpi, F. Maccari, Microdetermination of chondroitin sulfate in normal

human plasma by fluorophore-assisted carbohydrate electrophoresis (FACE), Clin. Chim. Acta, 2005, **356**, 125-133.

[11] H. Okamoto, T. Nakajima, Y. Ito, K. Shimada, S. Yamato, Development of a novel analytical method for determination of chondroitin sulfate using an in-capillary enzyme reaction, J. Chromatogr. A, 2004, **1035**, 137-144.

[12] USP29-NF24, United States Pharmacopeial Convention Inc., Rockville, MD, 2006, pp. 2306-2308.

[13] Y. Ozaki, Raman spectroscopy, in: M.M. Mossoba (Eds.), Spectral methods in food analysis instrumentation and applications, New York, Marcel Dekker, 1999, pp. 427-462.

[14] H.G.M. Edwards, A.C. Williams, B.W. Barry, Potential applications of FT-Raman spectroscopy for dermatological diagnostics, J. Mol. Struct., 1995, **347**, 379-388.

[15] R. Manoharan, J.J. Baraga, M.S. Feld, R.P. Rava, Quantitative histochemical analysis of human artery using Raman spectroscopy, J. Photochem. Photobiol. B: Biology, 1992, **16**, 211-233.

[16] W.Z. Lu, Analytical technique of modern near infrared spectrum, China Petrochemical Press, 2006.

[17] Y.B. Wang, X.L. Chu, W.Z. Lu, The introduction of the standard practice for NIR qualitative and quantitative analysis, The contemporary NIR technique--The 1th China NIR Symposium, 2006, 99-105.

[18] K.R. Beebe, R.J. Pell, M.B. Seasholtz, Chemometrics-a practical guide, John Wiley, New York, 1998.

[19] Z. Ni, C. Hu, F. Feng, Progress and effect of spectral data pretreatment in NIR analytical technique, Chin. J. Pharm. Anal., 2008, **28**, 824-829.

[20] K.R. Beebe, B.R. Kowalski, An introduction to multivariate calibration and analysis, Anal. Chem., 1987, **59**, 1007A-1017A.

[21] D. Zhu, B. Ji, C. Meng, B. Shi, Z. Tu, Z. Qing, The application of direct orthogonal signal correction for linear and non-linear multivariate calibration, Chemometr. Intell. Lab. Syst., 2008, **90**, 108-115.

[22] H. Yan, B.X. Han, Q.Y. Wu, M.Z. Jiang, Z.Z. Gui, Rapid detection of Rosa laevigata polysaccharide content by near-infrared spectroscopy, Spectrochim Acta A: Mol. Biomol. Spectrosc., 2011, **79**, 179-184.

[23] D. Lomiwes, M.M. Reis, E. Wiklund, O.A. Young, M. North, Near infrared spectroscopy as an on-line method to quantitatively determine glycogen and predict ultimate pH in pre rigor bovine M. longissimus dorsi, Meat Sci., 2010, **86**, 999-1004.

[24] Q. Dong, H.C. Zang, A. Liu , G. Yang, C.X. Sun, L. Sui, P. Wang, L. Li, Determination of molecular weight of hyaluronic acid by near-infrared spectroscopy, J. Pharm. Biomed. Anal., 2010, **53**, 274-278.

[25] Y. Wang, M.J. Qin, J. Qi, B.Y. Yu, L. Tang, Analysis of polysaccharides contents in Ophiopogon Japonicus by NIR, Spectrosc. Spect. Anal., 2009, **29**, 2677-2680.

[26] J.A. Spencer, J.F. Kauffman, J.C. Reepmeyer, C.M. Gryniewicz, W. Ye, D.Y. Toler, L.F. Buhse, B.J. Westenberger, Screening of heparin API by near infrared reflectance and Raman spectroscopy, J. Pharm. Sci., 2009, **98**, 3540-3547.

[27] M.F. Mrozek, D.M. Zhang, D.B. Amotz, Oligosaccharide identification and mixture quantification using Raman spectroscopy and chemometric analysis, Carbohydr. Res., 2004, **339**, 141-145.

[28] M. F. Mrozek, M.J. Weaver, Detection and identification of aqueous saccharides by using surface-enhanced raman spectroscopy, Anal. Chem., 2002, **74**, 4069-4075.

[29] H. Yang, J. Irudayaraj, Rapid detection of foodbome microorganisms on food surface using Fourier transform Raman spectroscopy, J. Mol. Struct., 2003, **646**, 35-43.

[30] L. Li, B.Y. Ding, Q. Yang, S. Chen, H.Y. Ren, J.F. Wang, H.C. Zang, F.S. Wang, L.X. Zang, The relevance study of effective information between near infrared spectroscopy and chondroitin sulfate in ethanol precipitation process, J. Innov. Opt. Health Sci., 2014, **7**, 1450022 (7 pages).

[31] M.A.M. Silva, M. H. Ferreira, J. W. B. Braga, M.M. Sena, Development and

analytical validation of a multivariate calibration method for determination of amoxicillin in suspension formulations by near infrared spectroscopy, Talanta, 2012, **89**, 342-351.

[32] P. Valderrama, J. W. B. Braga, R. J. Poppi, Variable selection, outlier detection, and figures of merit estimation in a partial least-squares regression multivariate calibration model. A case study for the determination of quality parameters in the alcohol industry by near-infrared spectroscopy, J. Agric. Food. Chem., 2007, **55**, 8331-8338.

[33] B. Walczak, D. L. Massart, Multiple outlier detection revisited, Chemometr. Intell. Lab. Syst., 1998, **41**, 1-15.

[34] M. H. Zhang, J. Luypaert, J. A. Fern andez Pierna, Q. S. Xu, D. L. Massart, Determination of total antioxidant capacity in green tea by near-infrared spectroscopy and multivariate calibration, Talanta, 2004, **62**, 25-35.

[35] Y. Zheng, X. Lai, S.W. Bruun, H. Ipsen, J.N. Larsen, H. Løwenstein, I. Søndergaard, S. Jacobsen, Determination of moisture content of lyophilized allergen vaccines by NIR spectroscopy, J. Pharm. Biomed. Anal., 2008, **46**, 592-596.