Carbohydrate Polymers 115 (2015) 582–588

Contents lists available at ScienceDirect

Carbohydrate Polymers

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

Development, validation and influence factor analysis of a near-infrared method for the molecular weight determination of xanthan gum

Zhigang Song[a,b], Peixue Ling[a,b,*], Hengchang Zang[a], Lian Li[a], Jinfeng Wang[a], Yan Jin[b], Huarong Shao[a,b], Xiqiang Zhu[b], Fei Liu[b], Fengshan Wang[a,**]

[a] School of Pharmaceutical Sciences and National Glycoengineering Research Center, Shandong University, Jinan 250012, China
[b] Post-doctoral Scientific Research Workstation, Shandong Academy of Pharmaceutical Sciences, Jinan 250101, China

ARTICLE INFO

Article history:
Received 29 April 2014
Received in revised form 21 July 2014
Accepted 9 August 2014
Available online 3 September 2014

Keywords:
Xanthan gum
Near-infrared spectroscopy
Molecular weight
Partial least squares
Method validation

ABSTRACT

A practical molecular weight determination model of xanthan gum (XG), based on near-infrared (NIR) spectroscopy, was built in this study. Two sample measurement modules, integrating sphere module and fiber-optic probe module, were compared, and the best partial least square (PLS) regression model was based on fiber-optic probe module. The values of coefficient of determination in calibration (R^2_c), coefficient of determination in prediction (R^2_p), residual predictive deviation (RPD) and root mean square error of prediction (RMSEP) were 0.967, 0.975, 6.028 and 0.250×10^6 Da, respectively. The molecular weight range, linearity, accuracy and precision of the established method were also validated. Furthermore, influence factors on this method were discussed in order to establish an appropriate measurement protocol. Results showed that the proposed NIR method may be suitable for practical applications in manufacturing plants and probably be accepted as a good alternative approach for fast determination of molecular weight of XG in production process.

1. Introduction

Xanthan gum (XG) is a high molecular weight microbial heteropolysaccharide with a primary structure consisting of repeated pentasaccharide units formed by two D-glucose, two D-mannose, and one D-glucuronic acid. The molecular weight distribution of XG ranges from 2×10^6 to 20×10^6 Da (García-Ochoa, Santos, Casas, & Gómez, 2000). Its main chain contains a cellulosic backbone of linear linked (β-1,4) D-glucose. The trisaccharide side chains on every alternate glucose at C-3 contain a glucuronic acid residue linked (β-1,4) to a terminal mannose unit and (β-1,2) to a second mannose that connects to the backbone (Palaniraj & Jayaraman, 2011). XG is widely used in many industries mainly in the food, cosmetic, textile, oil, agricultural and pharmaceutical fields (García-Ochoa et al., 2000; Llamas-Moreno, Baiza-Durán, Saucedo-Rodríguez, & Alaníz-De la, 2013; Palaniraj & Jayaraman, 2011; Shalviri, Liu, Abdekhodaie, & Wu, 2010). Recently, studies have been carried out on the XG intra-articular injection in treating experimental osteoarthritis, including its pharmacodynamics and action mechanism. Intra-articular injection of XG had protective effect on the articular cartilage in a rabbit osteoarthritis model (Han et al., 2012b), significantly reduced osteoarthritis pain and alleviated the joint cartilage degradation in a rat osteoarthritis model (Shao et al., 2013). XG also exhibited protective effect on rabbit chondrocytes in the presence of interleukin-1β (Han et al., 2012a). The results suggested that intra-articular injection of XG was probably a new therapeutic method for osteoarthritis. Recent advances in medical and pharmaceutical fields have expanded its applications, therefore, high demand for purified XG is anticipated.

Molecular weight, one of the most fundamental parameters in XG characterization, is a key industrial output control variable for many applications of end products. XG of different molecular weights displays different physical and chemical properties, which may determine its final applications (Born, Langendorff, & Boulenguer, 2005). The current methods used for molecular weight determination of XG include intrinsic viscosity, size exclusion chromatography combined with multiangle laser light scattering (SEC-MALLS), asymmetrical flow field fractionation with multiangle laser light scattering (AFFF-MALLS) and electron microscopy

* Corresponding author at: Shandong Academy of Pharmaceutical Sciences, Jinan 250101, China. Tel.: +86 531 81213002; fax: +86 531 88524738.
** Corresponding author. Tel.: +86 531 88382589; fax: +86 531 88382548.
E-mail addresses: lpx@sdfmg.com, lpxsdf@163.com (P. Ling), fswang@sdu.edu.cn (F. Wang).

http://dx.doi.org/10.1016/j.carbpol.2014.08.079
0144-8617/

(Born et al., 2005). However, these methods may be laborious and time consuming, which limit them as tools for routine analysis online in manufacturing plants. Hence, a rapid, straightforward and still accurate method is very desirable for the molecular weight determination of XG.

Near-infrared (NIR) spectroscopy is a rapid, cost-effective and non-destructive technique that requires little or no sample preparation. NIR spectroscopy offers many possibilities for a broad range of industrial analysis applications (Huang & Qu, 2011; Luypaert, Massart, & Vander Heyden, 2007; Xiao et al., 2012; Zamora-Rojas, Garrido-Varo, De Pedro-Sanz, Guerrero-Ginel, & Pérez-Marín, 2013). It has been adopted as a practical quality control method both in laboratories and in the production plants for online process control (Luypaert et al., 2007). With the development of NIR spectrometer, innovative handheld devices with small size, light weight, low cost and ease of use are now available in the market. Thereby, instead of taking the sample from the process and presenting it to the instruments, the NIR devices are placed on the process lines, making non-destructive in situ NIR spectroscopy analysis a reality (Zamora-Rojas et al., 2013). The NIR spectrum, with the wavelength ranging from 780 to 2500 nm, is mainly composed of overtones and the combinations of fundamental molecular vibrations caused by hydrogen bonds of X—H (e.g., N—H, C—H, O—H) (Luypaert et al., 2007). Many of these hydrogen bonds are present in XG molecules, making NIR spectroscopy suitable for investigating the physical or chemical properties of XG. The main difficulty of NIR technique in analysis is the complexity of the spectra due to their nature (overtones and combination bands of vibrational energy levels) (Luypaert et al., 2007). Overtones can be thought of as harmonics. And the fundamental will produce a series of absorptions at different wavelength. Combinations are more complicated. The fundamental/overtone absorptions are combined to make the NIR spectra to look rather uninteresting and to consist of a few broad absorption peaks. All these make it difficult to acquire relevant information about the feature of interest using measurements at only one wavelength from the raw spectra as is the case, for instance, like UV spectrophotometric or colorimetric analytical methods (Luypaert et al., 2007). So indirect calibration methods are needed to solve this problem and one of the most commonly used chemometric regression methods is partial least squares (PLS) regression. The PLS regression method, defined as a predictive two-block regression method based on estimated latent variables, is a method for relating two data matrices, X (matrix containing NIR spectra) and Y (matrix of response variables), by a linear multivariate mode (Wold, Sjöström, & Eriksson, 2001).

In recent years, NIR spectroscopy has shown its successful potential for monitoring or controlling polymer molecular weights (Cherfi, Fevotte, & Novat, 2002; Othman, Févotte, Peycelon, Egraz, & Suau, 2004), and for non-destructive determination of cellulose molecular weight in pulp hand sheets and historic papers (Henniges, Schwanninger, & Potthast, 2009). Furthermore, NIR spectroscopy has been proved in the molecular weight determination of hyaluronic acid using sample cup as a sampling accessory (Dong et al., 2010), which was based on the presupposition that hyaluronic acid of different molecular weights might possess different amount of hydroxyl end-groups per unit, which would be reflected in the NIR spectra (Dong et al., 2010). However, there is little discussion associated with the validation characteristics and influence factors of this method. Moreover, diffuse reflection (load sample into a sample cup) analysis may have limitations for practical applications in manufacturing plants. It may not be suitable for the in situ analysis of the sample. Industries are interested in simplification of the sampling method for in situ measurement on the process line. NIR spectroscopy has been studied for these purposes by using fiber-optic instrumentations (Claps & Virojanapa, 2013; Rodionova, Balyklova, Titova, & Pomerantsev, 2013; Xiang et al., 2009). To the best of our knowledge, there was no report on using NIR spectroscopy to determine the molecular weight of XG.

Table 1
Weight-average molecular weight (M_w) of the 54 samples.

Sample number	M_w ($\times 10^6$ Da) (relative error, %)	Sample number	M_w ($\times 10^6$ Da) (relative error, %)
1	5.962 (5)	28	2.908 (5)
2	5.656 (5)	29	2.805 (3)
3	5.571 (5)	30	2.730 (5)
4	4.913 (5)	31	2.694 (4)
5	4.874 (5)	32	2.617 (4)
6	4.755 (4)	33	2.592 (3)
7	4.754 (4)	34	2.384 (4)
8	4.536 (4)	35	2.209 (4)
9	4.384 (4)	36	2.099 (3)
10	4.205 (5)	37	2.087 (4)
11	3.564 (4)	38	1.951 (3)
12	3.367 (5)	39	1.906 (5)
13	5.074 (5)	40	1.761 (4)
14	5.059 (5)	41	1.736 (4)
15	4.968 (5)	42	1.631 (3)
16	4.586 (4)	43	1.522 (3)
17	4.227 (4)	44	1.515 (3)
18	4.145 (4)	45	1.512 (3)
19	3.848 (5)	46	1.452 (3)
20	3.802 (4)	47	1.267 (3)
21	3.698 (4)	48	1.265 (3)
22	3.530 (4)	49	1.146 (3)
23	3.514 (4)	50	1.097 (3)
24	3.462 (4)	51	1.083 (2)
25	3.275 (4)	52	1.057 (2)
26	3.164 (4)	53	0.968 (2)
27	3.062 (3)	54	0.844 (2)

In this paper, the feasibility to determine the molecular weight of XG by NIR spectroscopy coupled with PLS algorithm was attempted. Through PLS method the useful information of the NIR spectra was extracted and a new orthogonal matrix was set to represent the raw spectral data. The new data set was related to the molecular weight, which was used to create the PLS model. Two sampling modules, integrating sphere and fiber-optic probe, were used and compared according to the performance of the best PLS regression models. Several validation characteristics (range, linearity, accuracy and precision) of the method were evaluated. The influence factors of the method were also discussed.

2. Materials and methods

2.1. Sample preparation

Twelve batches of purified XG powder samples (99.6% purity) were prepared according to the previously described method (Han et al., 2012b). These samples were collected over a period of one year, so that the batches included could be thought to be sufficiently representative to cover the normal variation of these samples. To acquire more samples with different molecular weights, we prepared another 42 samples by ultrasonic degradation with a modified method of Milas, Rinaudo, and Tinland (1985). Four batches of samples were randomly selected and dissolved to 0.5% (w/v) in 0.4% (w/v) aqueous NaCl solutions. The native samples were then degraded by ultrasonicating for different times from 10 min to 4 h using a JY92-2D Sonifier (Xinzhi Scientific Corp., China) with ultrasonic power of 100 W. Each of the ultrasonicated solutions was poured to a large quantity of isopropanol to re-precipitate the XG. In this way, 42 samples of different molecular weights, designated as sample 13–54, were prepared (Table 1). All samples were ground and passed through a 40 mesh screen, then dried in vacuum oven at 45 °C for 24 h before analysis.

2.2. Reference method for the determination of XG molecular weight

The weight-average molecular weight (M_w) was analyzed by SEC-MALLS system consisting of a HP 1100 HPLC (Agilent Technologies Corp., USA), a DAWN EOS 18 angle laser light scattering detector and a OPTILAB refractive index detector (both from Wyatt Technologies Corp., USA). The samples were analyzed by a TSK GMPWXL size exclusion chromatography column (300 mm × 7.8 mm, TosoHaas, Montgomeryville, PA). XG was dissolved in the mobile phase of 0.1 M NaCl to a concentration of 0.1 mg/ml and complete dissolution was achieved by gentle stirring for 24 h. A sample volume of 200 μl was injected at a flow rate of 0.6 ml/min and the analysis time was 30 min. The refractive index increment (dn/dc) used for XG was 0.163. Measurement for each sample was repeated twice and the average data was calculated for accurate results. For data collection from the MALLS instrument, the Wyatt ASTRA software version 5.3 was used.

2.3. NIR measurement

The measurements were performed in the laboratory with Antaris II Fourier-transform NIR spectrometer (Thermo Fisher Scientific Corp., USA), covering a wavenumber range from 10,000 to 4000 cm^{-1}. The system software RESULT 3.3 (Thermo Fisher) was used for spectra acquisition.

All samples were analyzed by two sampling modules:

(1) Integrating sphere module. Samples were loaded into a sample cup (Antaris micro cup kit, 10 mm, Thermo Fisher) and compressed by a cap, then scanned with an integration sphere module. All samples were measured with the same sample cup.
(2) Fiber-optic probe module. Samples were respectively loaded into 54 glass vials (6 mm × 50 mm, Thermo Fisher) and covered with tightly fitting lids, then scanned by placing a fiber-optic probe into the glass vial.

Each spectrum was the average of 32 scans with a resolution of 4 cm^{-1}. The background spectrum (32 scans) was taken before the measurement of every sample in order to eliminate the effect caused by background. Each sample was reloaded for three times to eliminate the effect caused by sample loading. The mean of the three spectra which were collected from the same sample was used for data analysis in the following step. Therefore, a total of 54 spectra were obtained. The temperature was kept at room temperature and the relative humidity was kept at a steady level in the laboratory. The measuring time of each sample was less than 1 min.

2.4. Spectral pre-processing

Pre-processing methods, such as multiplicative signal correction (MSC) and standard normal variate (SNV), were applied to remove unwanted systematic variation to get the signal related to molecular weight (Huang & Qu, 2011). MSC was developed to reduce the effect of scattered light on diffuse reflection and transmission NIR spectra (Burns & Ciurczak, 2007). SNV is used to correct the effects for multiplicative interference of scatter and particle size (Zang et al., 2012). Both MSC and SNV can be used to remove physical spectral information (due to particle size), so that PLS model was performed based on mainly chemical spectral information (Chen, Zhao, Zhang, & Wang, 2006). In this study, both MSC and SNV were used to eliminate most variations respectively, followed by mean centering.

2.5. Model building

All the data were processed by Chemometric PLS Toolbox software (Version 7.5.2, Eingenvector Research, Inc., USA) based on MATLAB 2010a (Version 7.10.0, Mathworks Inc., USA).

All samples were divided into two sets by principal component analysis (PCA) method. The two sets included calibration set and validation set, which were used for model construction and validation, respectively. The leave-one-out cross-validation (LOOCV) was utilized to determine the optimal number of latent variables. The performance of calibration models was evaluated by the coefficient of determination (R^2), the root mean square error of calibration (RMSEC) and the root mean square error of prediction (RMSEP). In addition, the residual predictive deviation (RPD), defined as the ratio of standard deviation (SD) in validation set to RMSEP was also used to indicate the robustness and predictive accuracy of the models (Zang et al., 2013).

3. Results and discussion

3.1. SEC-MALLS results

With the development of MALLS technique, SEC-MALLS is widely accepted as a preferred method for determining polymer absolute molecular weight today (Chen, Ilasi, & Sekulic, 2011). Nonetheless, accurate determination of the M_w of XG is difficult because of its very high molecular weight, the stiffness of the molecules and the presence of aggregates between chains (Born et al., 2005). The molecular weights of the 54 samples are shown in Table 1. A wide range of molecular weights was covered, from low molecular weight of 0.844×10^6 Da to high molecular weight of 5.962×10^6 Da. Table 1 also showed the calculated value of the relative errors, denoting the accuracy of molecular weight determination, which is acceptable in laboratories (≤5%).

3.2. Spectral investigation

The raw NIR spectra of all the 54 samples are shown in Fig. 1(A and B). It is difficult to select information related to the molecular weight on the basis of specific bands. As seen from Fig. 1, there were strong water absorption bands around 5400 cm^{-1}, which were caused by the combination of the O—H stretching and the O—H bending band (Luypaert et al., 2007). These bands were excluded for the following analysis due to their significant influence on the PLS model created. Spectral region between 10,000 cm^{-1} and 9000 cm^{-1} exhibiting a high noise level and little chemical information was also excluded (Chen et al., 2006). Several combinations of pre-processing methods and the selected spectral regions were studied in order to construct the optimal calibration models. Spectral regions including 4900–4000 cm^{-1}, 7100–5465 cm^{-1} and 8240–7345 cm^{-1} were finally selected. The most useful feature

Table 2
Absorption bands of XG raw NIR spectra.

Wavenumber (cm^{-1})	Remark
8226	C—H stretch second overtone of CH_2
7355	C—H stretch first overtone/asym C—H bend combination of CH_3
6916	O—H stretch first overtone of H_2O and C—H stretch combination
6330	O—H stretch first overtone of hydrogen bond
5935	C—H stretch second overtone
5650	C—H stretch first overtone of CH_2 and CH_3
5181	O—H stretch/HOH twist combination
4770	O—H scissor third overtone of H_2O
4303	O—H stretch/CH_2 twist combination

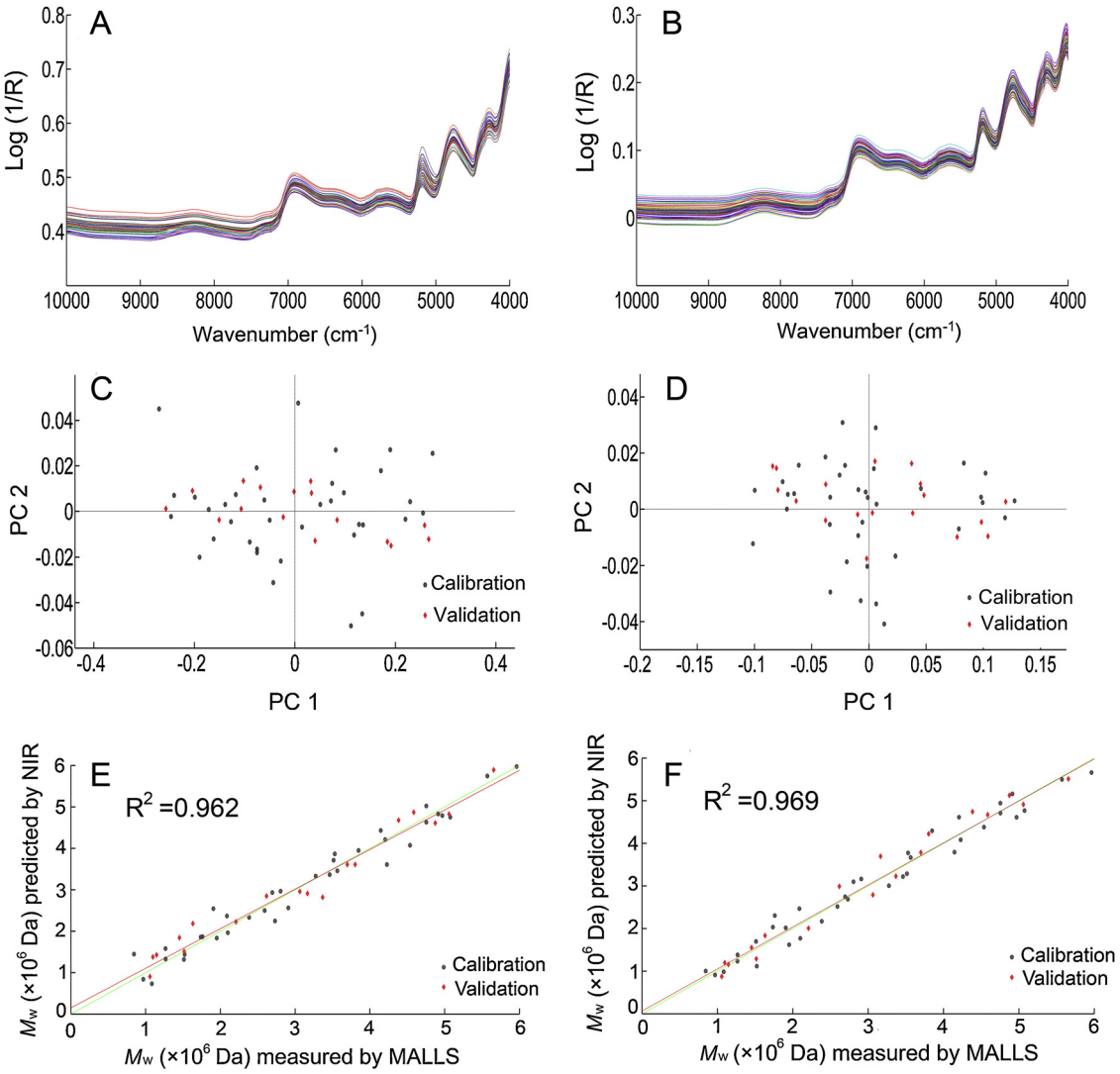

Fig. 1. Comparison between integrating sphere module and fiber-optic probe module. Original spectra of 54 samples measured by integrating sphere (A) and fiber-optic probe module (B). PCA results of integrating sphere (C) and fiber-optic probe module (D). Relationship of the predicted M_w values with the reference values from the model established with integrating sphere (E) and fiber-optic probe module (F). The red line indicates the linear trend obtained using both calibration and validation sets. The green line indicates the best linear relationship (1:1). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

information (Workman Jr & Weyer, 2007) of the absorption bands of XG in the selected spectral regions is shown in Table 2.

3.3. PLS model results

A 2/1 division of calibration/validation samples was obtained, so that there were 36 samples in calibration set and the remained 18 samples in validation set. All samples in validation were selected based on PCA method (Fig. 1(C and D)) and the M_w range of validation set (1.057×10^6–5.656×10^6 Da) was covered in that of calibration set (0.844×10^6–5.962×10^6 Da), therefore the distribution of the samples was appropriate in calibration set and validation set.

The PLS model with the best predictive capacity was selected according to conventional classical parameters such as the R^2, the number of latent variables, RMSEC and RMSEP (Tomuta, Iovanov, Vonica, & Leucuta, 2010). Fig. 1(E and F) shows the predictive M_w values of XG determined by NIR method versus the reference values. The parameters of PLS models applying to two sampling modules (integrating sphere and fiber-optic probe) are listed in Table 3. As shown in Table 3, MSC is as good as SNV for fiber-optic probe module, but is better than SNV for integrating sphere module. Therefore, MSC was used as the best pre-processing method.

Table 3
Statistical parameters of the PLS models produced.

Sampling modules	Pre-processing method	nLV[a]	R^2_c	R^2_p	RMSEC ($\times 10^6$ Da)	RMSEP ($\times 10^6$ Da)	RPD
Integrating sphere (sample cup)	MSC	8	0.962	0.963	0.277	0.285	5.288
	SNV	8	0.946	0.941	0.329	0.377	3.997
Fiber-optic probe	MSC	8	0.967	0.975	0.259	0.250	6.028
	SNV	8	0.967	0.975	0.258	0.249	6.052

[a] nLV, number of latent variables.

Table 4
Demonstration of accuracy.

Sample set	Method	n[a]	Mean ($\times 10^6$ Da)	SD ($\times 10^6$ Da)	F-test (0.05)		t-test (0.05)	
					F	P	t	P
Calibration	NIR	36	3.067	1.414	0.000	1.000	0.000	1.000
	SEC-MALLS	36	3.067	1.438				
Validation	NIR	18	3.088	1.560	0.017	0.896	−1.151	0.266
	SEC-MALLS	18	3.021	1.507				

[a] n, number of samples.

As seen from Table 3, both sampling modules are acceptable for quantitative determination, because their R^2 values are above 0.95 and RPD values are above 5. Generally, a model with RPD over 5.0 is considered suitable for accurate quantitative analysis (Li, Zhang, Cai & Shao, 2011). Apparently, the model established with fiber-optic probe module is better than that of the integrating sphere module, with higher values of R^2 and RPD. For this model, the coefficient of determination in prediction (R^2_p), RPD and RMSEC are 0.975, 6.028 and 0.259×10^6 Da, respectively. The optimal number of latent variables was found to be 8, and the performance of the model based on fiber-optic probe module may be more suitable for this application. Finally, the RMSEP value of fiber-optic probe module is 0.250×10^6 Da. In consideration of the very high molecular weight and the difficulties of accurate molecular weight determination of XG, they were acceptable from a manufacturing perspective.

3.4. Method validation

The method validation was based on the fiber-optic module. According to current guideline (ICH Harmonised Tripartite Guideline, 2005), the following validation characteristics were evaluated.

3.4.1. Range

Ensuring an appropriate calibration range is one of the main limitations of NIR spectroscopy applications in molecular weight determination of XG. A wider calibration range than that expected for production samples is necessary. In this study, the range of the model was determined by the minimum and maximum molecular weight values of samples in the calibration set, which was from 0.844×10^6 to 5.962×10^6 Da, respectively. The NIR model was only valid within this calibration range.

3.4.2. Linearity

The linearity of the method was evaluated by establishing the correlation between the molecular weight values predicted by the NIR model and those determined by SEC-MALLS. As seen from Fig. 1(F), the predicted M_w values were highly correlated with the standard reference values. The line of best fit (green line) for both the calibration and validation sets almost overlaps with the correlation line (red line). The R^2_c and R^2_p of the model were 0.967 and 0.975 (Table 3), respectively, demonstrating a good linearity of the model.

3.4.3. Accuracy

Accuracy of the NIR method was evaluated by performing a paired t-test for independent samples by variables between the molecular weight values predicted by the model and those determined by SEC-MALLS (Peinado, Hammond, & Scott, 2011). The analysis was carried out both for samples in the calibration and in the validation set. Before the paired t-test, the variance of both methods was compared using an F-test to assess whether they differed significantly (Peinado et al., 2011). The results obtained were listed in Table 4. For a significance level of 0.05, the F-experimental was lower than F-critical for both the calibration and validation sets, which indicated that there were no significant differences in variance between the standard deviation of these two methods. The t-experimental was also lower than t-critical for a significant level of 0.05 for both sets. As a result, the differences between the molecular weight values determined by the NIR method and SEC-MALLS were not significant. The accuracy was thus validated.

Table 5
Repeatability data of NIR spectroscopy for molecular weight determination.

Assay	NIR predicted ($\times 10^6$ Da)				
	Sample 5	Sample 12	Sample 14	Sample 35	Sample 50
1	4.931	3.315	4.986	2.299	1.061
2	4.889	3.270	5.135	2.067	1.102
3	5.034	3.109	4.811	2.288	1.015
4	4.867	3.438	4.911	2.263	1.021
5	4.739	3.219	4.959	2.116	1.132
6	4.906	3.311	5.211	2.152	1.052
7	4.717	3.259	5.161	2.243	0.986
8	4.909	3.288	5.189	2.059	1.046
9	4.744	3.409	5.043	2.327	1.095
10	5.053	3.249	4.992	2.118	1.136
Mean value	4.879	3.287	5.040	2.193	1.065
SD	0.1169	0.0930	0.1316	0.1015	0.0505
RSD (%)	2.395	2.829	2.611	4.629	4.741
χ^2	8.295				

3.4.4. Precision

The precision of the method was evaluated at two levels: repeatability and intermediate precision.

3.4.4.1. Repeatability tests. Repeatability tests were performed by repeated measurement of spectra over a short interval of time (Moffat, Trafford, Jee, & Graham, 2000). Five samples (sample 5, 12, 14, 35 and 50) were chosen from the validation set to represent low, medium and rather high molecular weight. Ten spectra of each sample were acquired using fiber-optic probe as described above. The molecular weight of each sample was predicted using the established model, and then the mean of each sample's predicted values and repeatability were calculated. A Chi-square (χ^2) test was used to investigate whether the repeatability described by standard deviations of 5 samples belongs to the same population (Dong et al., 2010). The predicted values and statistical results are shown in Table 5. For a given 95% confidence level, the critical value of χ^2 (0.05, 4) is 9.49. The χ^2 of NIR model of molecular weight is 8.295, lower than the critical value, indicating that the standard deviations of repeatability belong to the same population. As the relative standard deviation (RSD) from NIR analysis was below 5%, the repeatability was validated.

3.4.4.2. Intermediate precision. Intermediate precision expresses within-laboratories variations, such as different days, different analysts and different equipment (Moffat et al., 2000). In our study, the variable parameters were days and operators. In this work, the intra-day precision was measured by determining a single sample (sample 12) in six different times on the same day. The inter-day

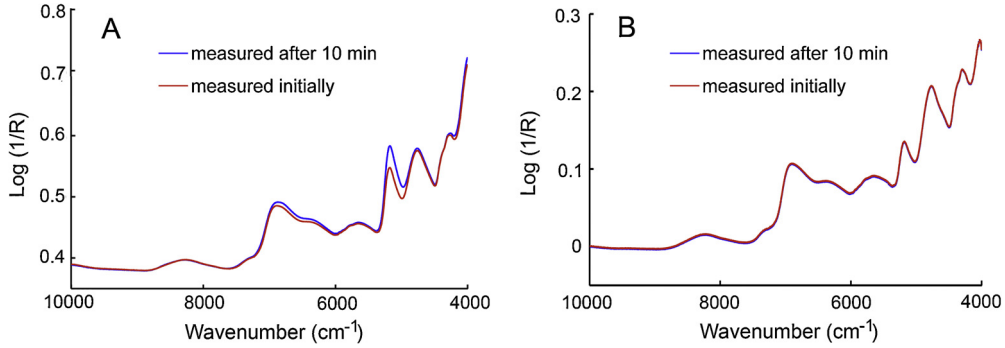

Fig. 2. Spectra of the same sample measured before and after 10 min by integrating sphere (A) and fiber-optic probe module (B). Each spectrum was the average of three measurements.

Table 6
Intermediate precision data of NIR spectroscopy for molecular weight determination.

Assay	NIR predicted ($\times 10^6$ Da)		
	Intra-day	Inter-day	Analyst
1	3.221	3.427	3.219
2	3.213	3.226	3.206
3	3.375	3.189	3.227
4	3.394	3.278	3.287
5	3.188	3.422	3.335
6	3.349	3.292	3.406
Mean value	3.290	3.306	3.280
SD	0.0922	0.0990	0.0786
RSD (%)	2.803	2.995	2.397

precision was measured by determining the same sample on six consecutive days. The analyst precision was measured by six different analysts testing the same sample six times in the same day. Results are shown in Table 6. As seen from Table 6, the RSD is below the normally accepted level of 5%, therefore the intermediate precision of the NIR method is validated. Moreover, as to sample 12, comparison of the repeatability (see Table 5) and intermediate precision indicated that the variations introduced by different operators or days did not increase the RSD values.

3.5. Influence factors on molecular weight determination of XG by NIR spectroscopy

3.5.1. Samples used in NIR analysis

In order to construct a model with high accuracy, sufficient samples with a wide range of molecular weight are needed. Frequently, under commercial production, the variability in the process parameter is too limited to provide samples with a wide range of molecular weight. To overcome this problem, the set of samples was enlarged with self-prepared laboratory samples. Ultrasonic degradation is a well established procedure, which has been applied to XG by several authors (Kim, Choi, Kim, & Jhon, 1998; Milas et al., 1985; Sohn, Kim, Choi, & Jhon, 2001). Ultrasonic degradation is generally accepted to prepare XG with a controlled molecular weight without chemical structure changes (Milas, Rinaudo, & Tinland, 1986; Sohn et al., 2001). Therefore, ultrasonic degradation was selected to provide XG samples of different molecular weights. Nevertheless, in order to establish a more practical model in production, as many production samples as possible should be involved in the future work.

3.5.2. The sampling modules

Integrating sphere and fiber-optic modules were performed and compared under the same spectral conditions. As expected, sampling modules may produce the predictive ability of the models. Although both modules could satisfy the needs of molecular weight determination, the best model was obtained when fiber-optic probe was applied directly to the samples. Fiber-optic sampling devices are widely used in NIR analysis for easy sampling of raw materials or for in situ measurement of process streams. The fiber-optic probe is more convenient than integrating sphere module. Thus, the NIR method for molecular weight determination of XG based on fiber-optic instrumentations is more practical in manufacturing plants.

3.5.3. Moisture content of the samples

As known, water absorbs strongly in the NIR region. The presence or absence of water in a sample will influence the extent of hydrogen bonding within the sample (Burns & Ciurczak, 2007). For powder samples, moisture content is one of the most important influence factors, and ideally all samples should be at a reasonably uniform moisture content (Burns & Ciurczak, 2007). XG has a high affinity for water due to its polyanionic nature. Fig. 2 illustrates the difference in the water band at 5400 cm^{-1} for the same XG sample measured before and after 10 min. During the measurement, the relative humidity was kept at a steady level of 40%. As seen from Fig. 2(A), an evident increase in the water band at 5400 cm^{-1} was observed in the repeated measurement after 10 min. The moisture exchange with the surrounding atmosphere was inevitable if the moisture content of the sample was not in equilibrium with the relative humidity of the atmosphere surrounding the sample. Because of its high attraction for water, a fast and stable sample measurement module is required for XG determination. Fiber-optic module would reduce the interference caused by moisture absorption during analysis (Fig. 2(B)). It may be why fiber-optic module is better than integrating sphere module for molecular weight determination of XG. Although water absorption bands were excluded prior to construction of the calibration model, the presence of water in a sample could still affect the whole spectral range, which might induce significant errors to the results. Therefore, it is very important to standardize the sampling in order to get an accurate and representative measurement.

In this study, all samples were dried in vacuum oven at 45 °C for 24 h to reach equilibrium, then stored in glass vials and covered with tightly fitting lids to prevent moisture absorption. The final moisture content of the sample should not exceed 10%. The measurement time of each sample was limited to less than 1 min. These procedures should be considered as part of the measurement protocol for XG molecular weight determination by NIR spectroscopy.

4. Conclusions

A simple, rapid, non-destructive and practical method based on NIR spectroscopy and PLS multivariate calibration for molecular

weight determination of XG was developed. Fiber-optic probe was selected as the optimum sampling module, which yielded a simpler operation, faster analysis time and increased flexibility. The NIR method developed in this work was proved to be accurate and precise. The RMSEP value was 0.250×10^6 Da, which was within the permissible error limits in manufacturing plants. The method was applicable over a wide range of molecular weight, from 0.844×10^6 to 5.962×10^6 Da. The proposed method may be a good alternative for fast determination of XG molecular weight in production process, and probably be accepted as a potential quality control method for online process control in the production plants. The moisture content of the samples was considered to have significant influence on the method, therefore, an appropriate measurement protocol should be established to minimize the adverse effect. Further work is in progress to enlarge the data sets and build a more practical model in production.

Acknowledgments

We thank Prof. Lei Nie for his help of technical assistance and language revision. This work was supported by Promotive Research Fund for Excellent Young and Middle-Aged Scientists of Shandong Province, China (No. BS2011YY050) and Jinan University and Institute Independent Innovation Project (No. 201202038).

References

Born, K., Langendorff, V., & Boulenguer, P. (2005). Xanthan. *Biopolymers Online*, 259–291. http://onlinelibrary.wiley.com/doi/10.1002/3527600035.bpol5011/figures

Burns, D. A., & Ciurczak, E. W. (2007). *Handbook of near-infrared analysis* (3rd ed.). Boca Raton: CRC press (Chapter 13).

Chen, Q., Zhao, J., Zhang, H., & Wang, X. (2006). Feasibility study on qualitative and quantitative analysis in tea by near infrared spectroscopy with multivariate calibration. *Analytica Chimica Acta*, *572*(1), 77–84.

Chen, R., Ilasi, N., & Sekulic, S. S. (2011). Absolute molecular weight determination of hypromellose acetate succinate by size exclusion chromatography: Use of a multi angle laser light scattering detector and a mixed solvent. *Journal of Pharmaceutical and Biomedical Analysis*, *56*(4), 743–748.

Cherfi, A., Fevotte, G., & Novat, C. (2002). Robust on-line measurement of conversion and molecular weight using NIR spectroscopy during solution polymerization. *Journal of Applied Polymer Science*, *85*(12), 2510–2520.

Claps, R., & Virojanapa, J. (2013). Fast, broadband, rugged spectrometer for near-infrared measurements of organic solutions and multicomponent mixtures. *Journal of Biomedical Optics*, *18*(11), 117001.

Dong, Q., Zang, H., Liu, A., Yang, G., Sun, C., Sui, L., et al. (2010). Determination of molecular weight of hyaluronic acid by near-infrared spectroscopy. *Journal of Pharmaceutical and Biomedical Analysis*, *53*(3), 274–278.

García-Ochoa, F., Santos, V., Casas, J., & Gómez, E. (2000). Xanthan gum: production, recovery, and properties. *Biotechnology Advances*, *18*(7), 549–579.

Han, G., Shao, H., Zhu, X., Wang, G., Liu, F., Wang, F., et al. (2012). The protective effect of xanthan gum on interleukin-1β induced rabbit chondrocytes. *Carbohydrate Polymers*, *89*(3), 870–875.

Han, G., Wang, G., Zhu, X., Shao, H., Liu, F., Yang, P., et al. (2012). Preparation of xanthan gum injection and its protective effect on articular cartilage in the development of osteoarthritis. *Carbohydrate Polymers*, *87*(2), 1837–1842.

Henniges, U., Schwanninger, M., & Potthast, A. (2009). Non-destructive determination of cellulose functional groups and molecular weight in pulp hand sheets and historic papers by NIR-PLS-R. *Carbohydrate Polymers*, *76*(3), 374–380.

Huang, H., & Qu, H. (2011). In-line monitoring of alcohol precipitation by near-infrared spectroscopy in conjunction with multivariate batch modeling. *Analytica Chimica Acta*, *707*(1–2), 47–56.

ICH Harmonised Tripartite Guideline. (2005). *Validation of analytical procedures: Text and methodology*. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q2_R1/Step4/Q2_R1__Guideline.pdf

Kim, C., Choi, H., Kim, C., & Jhon, M. (1998). Drag reduction characteristics of polysaccharide xanthan gum. *Macromolecular Rapid Communications*, *19*(8), 419–422.

Li, J., Zhang, Y., Cai, W., & Shao, X. (2011). Simultaneous determination of mercury, lead and cadmium ions in water using near-infrared spectroscopy with preconcentration by thiol-functionalized magnesium phyllosilicate clay. *Talanta*, *84*(3), 679–683.

Llamas-Moreno, J. F., Baiza-Durán, L. M., Saucedo-Rodríguez, L. R., & Alaníz-De la, O. J. (2013). Efficacy and safety of chondroitin sulfate/xanthan gum versus polyethylene glycol/propylene glycol/hydroxypropyl guar in patients with dry eye. *Clinical Ophthalmology*, *7*, 995–999.

Luypaert, J., Massart, D. L., & Vander Heyden, Y. (2007). Near-infrared spectroscopy applications in pharmaceutical analysis. *Talanta*, *72*(3), 865–883.

Milas, M., Rinaudo, M., & Tinland, B. (1985). The viscosity dependence on concentration, molecular weight and shear rate of xanthan solutions. *Polymer Bulletin*, *14*(2), 157–164.

Milas, M., Rinaudo, M., & Tinland, B. (1986). Comparative depolymerization of xanthan gum by ultrasonic and enzymic treatments. Rheological and structural properties. *Carbohydrate Polymers*, *6*(2), 95–107.

Moffat, A. C., Trafford, A. D., Jee, R. D., & Graham, P. (2000). Meeting the international conference on harmonisation's guidelines on validation of analytical procedures: Quantification as exemplified by a near-infrared reflectance assay of paracetamol in intact tablets. *Analyst*, *125*(7), 1341–1351.

Othman, N. S., Févotte, G., Peycelon, D., Egraz, J. B., & Suau, J. M. (2004). Control of polymer molecular weight using near infrared spectroscopy. *AIChE Journal*, *50*(3), 654–664.

Palaniraj, A., & Jayaraman, V. (2011). Production, recovery and applications of xanthan gum by *Xanthomonas campestris*. *Journal of Food Engineering*, *106*(1), 1–12.

Peinado, A., Hammond, J., & Scott, A. (2011). Development, validation and transfer of a near infrared method to determine in-line the end point of a fluidised drying process for commercial production batches of an approved oral solid dose pharmaceutical product. *Journal of Pharmaceutical and Biomedical Analysis*, *54*(1), 13–20.

Rodionova, O. Y., Balyklova, K. S., Titova, A. V., & Pomerantsev, A. L. (2013). The influence of fiber-probe accessories application on the results of near-infrared (NIR) measurements. *Applied Spectroscopy*, *67*(12), 1401–1407.

Shalviri, A., Liu, Q., Abdekhodaie, M. J., & Wu, X. Y. (2010). Novel modified starch–xanthan gum hydrogels for controlled drug delivery: Synthesis and characterization. *Carbohydrate Polymers*, *79*(4), 898–907.

Shao, H., Han, G., Ling, P., Zhu, X., Wang, F., Zhao, L., et al. (2013). Intra-articular injection of xanthan gum reduces pain and cartilage damage in a rat osteoarthritis model. *Carbohydrate Polymers*, *92*(2), 1850–1857.

Sohn, J.-I., Kim, C., Choi, H., & Jhon, M. (2001). Drag-reduction effectiveness of xanthan gum in a rotating disk apparatus. *Carbohydrate Polymers*, *45*(1), 61–68.

Tomuta, I., Iovanov, R., Vonica, A. L., & Leucuta, S. E. (2010). High-throughput NIR-chemometric method for meloxicam assay from powder blends for tableting. *Scientia Pharmaceutica*, *79*(4), 885–898.

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, *58*(2), 109–130.

Workman, J., Jr., & Weyer, L. (2007). *Practical guide to interpretive near-infrared spectroscopy*. Boca Raton: CRC press (Appendix 4a).

Xiang, D., LoBrutto, R., Cheney, J., Wabuyele, B. W., Berry, J., Lyon, R., et al. (2009). Evaluation of transmission and reflection modalities for measuring content uniformity of pharmaceutical tablets with near-infrared spectroscopy. *Applied Spectroscopy*, *63*(1), 33–47.

Xiao, X., Hou, Y., Du, J., Liu, Y., Liu, Y., Dong, L., et al. (2012). Determination of main categories of components in corn steep liquor by near-infrared spectroscopy and partial least-squares regression. *Journal of Agricultural and Food Chemistry*, *60*(32), 7830–7835.

Zamora-Rojas, E., Garrido-Varo, A., De Pedro-Sanz, E., Guerrero-Ginel, J. E., & Pérez-Marín, D. (2013). Prediction of fatty acids content in pig adipose tissue by near infrared spectroscopy: At-line versus in-situ analysis. *Meat Science*, *95*(3), 503–511.

Zang, H., Li, L., Wang, F., Yi, Q., Dong, Q., Sun, C., et al. (2012). A method for identifying the origin of chondroitin sulfate with near infrared spectroscopy. *Journal of Pharmaceutical and Biomedical Analysis*, *61*, 224–229.

Zang, H., Wang, J., Li, L., Zhang, H., Jiang, W., & Wang, F. (2013). Application of near-infrared spectroscopy combined with multivariate analysis in monitoring of crude heparin purification process. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *109*, 8–13.