Chinese Chemical Letters 30 (2019) 111–114

Contents lists available at ScienceDirect

Chinese Chemical Letters

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

Communication

Determination of triglycerides in human serum by near-infrared diffuse reflectance spectroscopy using silver mirror as a substrate

Shuyu Wang[a], Jin Zhang[a], Cuicui Wang[a], Xiaoming Yu[b], Wensheng Cai[a], Xueguang Shao[a,c,d,e,*]

[a] Research Center for Analytical Sciences, College of Chemistry, Nankai University, Tianjin 300071, China
[b] Laboratory of Clinic, People's Hospital of Gaomi City, Gaomi 261500, China
[c] Tianjin Key Laboratory of Biosensing and Molecular Recognition, Tianjin 300071, China
[d] State Key Laboratory of Medicinal Chemical Biology, Tianjin 300071, China
[e] Collaborative Innovation Center of Chemical Science and Engineering, Tianjin 300071, China

ARTICLE INFO

Article history:
Received 9 November 2017
Received in revised form 7 January 2018
Accepted 9 January 2018
Available online 10 January 2018

Keywords:
Near-infrared diffuse reflectance spectroscopy
Silver mirror
Variable selection
Quantitative model
Serum triglycerides

ABSTRACT

Near infrared diffuse reflectance spectroscopy (NIRDRS) has gained wide attention due to its convenience for rapid quantitative analysis of complex samples. A method for rapid analysis of triglycerides in human serum using NIRDRS with silver mirror as the substrate is developed. Due to the even and high reflectance of the silver mirror, the spectral response is enhanced and the background interference is reduced. Furthermore, both linear and nonlinear modeling strategies were investigated adopting the partial least squares (PLS) and least squares support vector regression (LS-SVR), continuous wavelet transform (CWT) was used for spectral preprocessing, and variable selection was tried using Monte Carlo uninformative variable elimination (MC-UVE), randomization test (RT) and competitive adaptive reweighted sampling (CARS) for optimization the models. The results show that the determination coefficient (R) between the predicted and reference concentration is 0.9624 and the root mean squared error of prediction (RMSEP) is 0.21. The maximum deviation of the prediction results is as low as 0.473 mmol/L. The proposed method may provide an alternative method for routine analysis of serum triglycerides in clinical applications.

The measurement of biochemical serum parameters is very important in clinical analyses. Because the concentration of the components in serum varies with the conditions of human health, diagnoses can be achieved based on the serum parameters. The triglyceride level in serum is an important indicator for clinical evaluation of coronary heart disease and vascular diseases. The excess triglycerides in blood may result in hypertriglyceridemia and it is related to the occurrence of coronary artery disease [1,2]. Therefore, accurate and precise measurement of serum triglycerides is essential in clinical medicine. The commonly used methods for determination of serum triglycerides include enzymatic reaction [3,4] and colorimetric methods [5]. These methods exhibit the high sensitivity and accuracy, but there are deficiencies that the triglycerides need to be hydrolyzed before the determination. On the other hand, sample pretreatment is needed and the procedure is time-consuming. Therefore, methods of less cost and easy operation are still needed for routine analysis of serum triglycerides in practical uses.

As a reagent-less, nondestructive and noninvasive measurement technique, near infrared (NIR) spectroscopy has been proved to be useful for determination of the biochemical parameters of blood, *e.g.*, glucose, cholesterol, urea and triglyceride [6–9]. Arnold *et al.* [10] proposed a method for the quantification of six components (glucose, urea, lactate, ascorbate, triacetin and alanine) in aqueous solutions using multivariate calibration models. By the same researchers, the measurement of triglycerides in undiluted human serum was also achieved [7]. Furthermore, works for determination of triglycerides in human plasma or serum were reported [2,11,12]. In these studies, NIR transmission spectra were used for the analysis. Due to the strong absorption of water and the complexity of serum matrix, however, it is still difficult to establish an accurate model for determination of triglycerides in serum directly using the transmission spectra.

Preconcentration has been proved to be an effective technique for enhancing the sensitivity of near infrared diffuse reflectance

* Corresponding author at: Research Center for Analytical Sciences, College of Chemistry, Nankai University, Tianjin 300071, China.
E-mail address: xshao@nankai.edu.cn (X. Shao).

https://doi.org/10.1016/j.cclet.2018.01.016

spectroscopy (NIRDRS) during the quantitative determination of micro-components. Aromatic hydrocarbons at concentrations of 100 ppm in water were detected by using silicone sensing phase [13,14]. Employing an enrichment device of a silica-based monolithic material, Du *et al.* [15] improved the detection limits of NIRDRS in the determination of trace lead in water. With the help of the enrichment device, lead at a concentration of 0.2 mg/L could be determined. In our previous works, resin, thiol-functionalized magnesium phyllosilicate clay, thiol-functionalized silica porous particles and nano-hydroxyapatite were used as adsorbents to preconcentrate metal ions (including copper, cobalt, nickel, mercury, lead, cadmium, zinc and chromium) [16–19], phenol and *p*-nitrophenol [20] in wastewater. To investigate the biological systems, amino-modified silica particle, β-cyclodextrins immobilized and thiourea-functionalized silica nanoparticles were adopted for concentrating fish sperm deoxyribonucleic acid and bilirubin in water solutions, and albumin in urine [21–23]. A method for fast determination of bovine serum albumin using a filter paper as the substrate was also developed [24]. It was found in these studies that the absorbents have strong absorption in NIR region and the spectra overlap with the responses of the target analyte. Therefore, silver mirror was used as the substrate in our recent studies [25–27]. The substrate has low background absorbance and can enhance the spectral feature of the analyte due to the high reflectance of silver to the NIR light. Dropping a certain volume of liquid onto the silver mirror and then evaporating the water, the spectrum of NIRDRS can be measured without the interference of water. Using such an approach, the maximum deviation of the prediction in absolute quantity can be as low as 12.8 μg for lysozyme samples [25]. For further optimization of the method, experimental and chemometric effort were made to obtain a satisfactory quantitative model. The maximum deviation for the prediction samples is only 2.98 μg [26]. When the method applied to routine analysis in clinical application, an acceptable result was obtained by the optimized model for directly analyzing the content of urea in human serum without any sample pretreatment [27].

In this work, quantitative determination of triglycerides in human serum by NIRDRS technique with silver mirror as the substrate was studied. 325 serum samples supplied by People's Hospital of Gaomi City (Shandong, China) were used in this study. The samples were collected from patients in clinical laboratory with the approval of the Ethics Committee of the hospital. The serum samples (400 μL) were dropped onto the prepared silver mirror grooves. After evaporation of the water, liquid film with 30 mm diameter was formed. The silver mirror with the dried serum sample was measured by NIRDRS. The details of the experiment are described in the Supporting information.

Fig. 1 shows the spectra of the 325 samples in wavenumber range of 9000–4000 cm^{-1}. Each of the spectra is the average of three parallel measurements. Because the silver mirror substrate was used as the background, all the peaks in the spectra come from the samples and the spectral features can be clearly seen in the spectra. Therefore, compared with the substrates with strong NIR absorption, silver mirror may be a good substrate to enhance the diffuse reflection and to reduce the background, as discussed in our previous work [25]. However, the spectra are composed of broad and overlapping peaks and variant background still exists in the spectra, which may be caused by the complexity of the serum samples. Therefore, it is difficult to assign the peaks to the structures of the components. Chemometric modeling is needed for analyzing the spectra.

Before the calculation of the modeling, the spectra were divided into a calibration and a validation set. The former was used for building the model and the latter was used for validation. According to the sequence sorted by triglyceride concentration, three of every four samples were selected as calibration set, and the others were used as validation set. The partial least squares (PLS) and least squares support vector regression (LS-SVR) were used for building the linear and non- linear models in this study, and Monte Carlo cross validation (MCCV) with adjusted Wold's R criterion [28] was used for determination of the number of latent variable (LV). Cross-validation was used for evaluation of the models and the performance was evaluated by the correlation coefficient between the predicted and reference values (*R*cv), root mean squared error of cross validation (RMSECV) and residual predictive deviation (RPD). In addition, the parameters of determination coefficient (*R*), root mean squared error of prediction (RMSEP) and maximal deviation of the validation set were adopted to evaluate the prediction results of the models. The model with higher *R*cv or *R* and lower RMSECV or RMSEP value is considered as a better one, and the higher the RPD value is, the better the model is. A RPD value greater than 3.0 or even 5.0 is taken as a criterion for an acceptable model in quantitative analysis [29,30].

For building an optimized model, chemometric techniques for signal preprocessing and variable selection were used [31,32]. To remove the variant background and enhance the resolution of the spectra, continuous wavelet transform (CWT) with the wavelet 'sym2′ and scale parameter 20 was employed [33–36]. The preprocessed spectra of the samples were shown in Fig. 2. The dash line is obtained from the spectrum of a blank sample (the silver substrate). Clearly, the resolution of the spectra is greatly improved and the variant background is removed. However, it is still impossible to associate the peaks with the structures of the components. Variable selection has been proved powerful to select the informative variables for building the optimized model. The variables selected by three methods, *i.e.*, Monte Carlo uninformative variable elimination (MC-UVE) [37], randomization test (RT) [38] and competitive adaptive reweighted sampling (CARS) [39], are shown in Fig. 2 by the short vertical bars in different colors. The

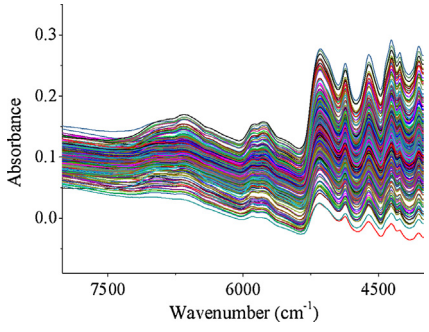

Fig. 1. NIRDRS spectra of the serum samples.

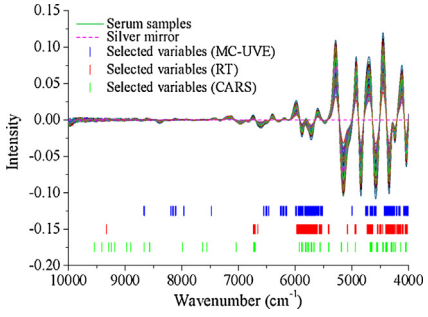

Fig. 2. CWT-preprocessed spectra of silver mirror (dot line) and serum samples (straight line), and selected variables by MC-UVE (blue bars), RT (red bars) and CARS (green bars).

CWT-preprocessed spectra of the serum samples and the silver mirror were plotted in the figure for reference. Only 78 variables are selected by CARS, but 229 variables are selected by MC-UVE method and 252 variables are selected by RT method. The number of the selected variables is determined by cross–validation of the calibration set with the criterion of RMSECV [37–39]. Although the number is different, most of the selected variables are identical, as shown in the figure. Furthermore, it can be observed that most of the variables related to the spectral feature of C—H groups were selected except for the variables around 5000 cm^{-1}, which may come from the absorption of the remained water in the samples. It should be reasonable for analyzing the target of triglycerides. However, more detail analysis of these variables is too difficult because the spectra are composed of the responses from the complicated composition of the serum.

To compare the models obtained by the selected variables, the statistics for the results obtained by the models were listed in Table 1. Both the results of the cross-validation (left part) and the prediction (right part) were summarized in the table. The first line shows the results of the model built with raw spectra from 4000 cm^{-1} to 9000 cm^{-1} (1555 variables) without any preprocessing and the results of the model with only CWT preprocessing was listed in the second line.

For PLS models, it is clear that the raw spectral model is not qualified for quantitative prediction. Comparing the raw spectra model and CWT model, CWT preprocessing can improve model obviously. However, the model can still not be acceptable for quantitative prediction according to the value of RPD. When the variable selection methods were used for building the model, much better results were obtained, although the models are still not good enough for quantitative analysis.

As for LS-SVR models, the raw spectra model is better than the PLS model. The CWT preprocessing significantly improved the model, obtaining a RPD value as high as 5.0. The result may suggest that CWT can remove the noise and the variant background in the spectra, but it cannot change the linearity of the spectral data. Non-linearity may be the reason for the poorness of the PLS models. As for the models with the selected variables, further improvements can be found, indicating that variable selection is necessary for the data to build a better model. Comparing the three methods of variable selection, the MC-UVE model is slightly better than the others, but the difference between the models is not significant although different spectral variables are used. The result is common in the comparison of the variable selection methods, because each method can only select a subset of the informative variables [40].

The right part of Table 1 shows the results for the validation set, which can be used to investigate the practicability of the models. The last parameter shows the largest prediction error among the samples in the validation set. If the largest error is acceptable, all the predicted concentrations can be trusted in the application. Comparing the results for the calibration and validation set, the parameters *R* and RMSEP are in the same level with *R*cv and RMSECV, and the relative values between the models are also similar. This indicates that the models are reasonably established and the validation set are similar with the calibration set. For the last parameter, max deviation, the same conclusion can be obtained. Clearly, the best model is obtained by the spectra with CWT preprocessing and the selected variables with MC-UVE. The maximum deviation of the optimal model is as low as 0.473 mmol/L. It is difficult to compare the results of the method with that of the conventional methods, because detection limit and the precision cannot be obtained for multivariate calibration such as PLS and SVR. The maximum deviation, which is the largest error among all the studied samples, may be a parameter to show the performance of the method. The detection limits reported in literatures are generally between 0.1 mmol/L and 4.2 mmol/L and the recoveries can be as low as 81% [3]. The deviation of the result obtained by asymmetric flow field-flow fractionation (AF4) method from the standard enzymatic test can be as large as 0.6 mmol/L [4]. Compared these data with the maximum deviation in this study, the accuracy of the method should be acceptable. Furthermore, the practicability of the method can be shown by the maximum deviation, which is almost the lowest concentration of the triglyceride content in the serum samples.

For further investigate the predicted results of the optimal model, the scatter plot of the reference and predicted contents for

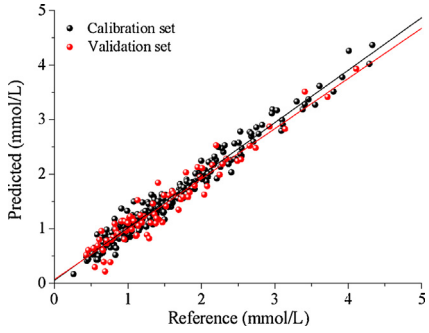

Fig. 3. Relationship between the predicted concentrations and the reference values using the optimal model.

Table 1
Statistics for the calibration and validation performance of PLS and LS-SVR models.

Method	nVar[a]	nLV[b]	Calibration set			Validation set		
			*R*cv	RMSECV	RPD	*R*	RMSEP	Max deviation (mmol/L)
PLS								
None	1555	5	0.3542	0.75	1.1	0.5721	0.65	2.086
CWT	1555	5	0.8346	0.44	2.0	0.8083	0.46	1.420
CWT-MC-UVE	229	5	0.9003	0.35	2.5	0.8986	0.34	1.069
CWT-RT	252	5	0.8892	0.37	2.4	0.8797	0.37	1.051
CWT-CARS	78	5	0.8957	0.36	2.3	0.8903	0.35	1.127
LS-SVR								
None	1555	–	0.7768	0.51	2.8	0.7846	0.48	1.468
CWT	1555	–	0.9144	0.32	5.0	0.9482	0.25	0.739
CWT-MC-UVE	229	–	0.9576	0.23	5.7	0.9624	0.21	0.473
CWT-RT	252	–	0.9575	0.23	5.6	0.9608	0.22	0.514
CWT-CARS	78	–	0.9491	0.25	5.6	0.9584	0.22	0.527

[a] Number of the variables used in the modelling.
[b] Number of latent variables.

the validation set was shown in Fig. 3. The black points represent the results of calibration set obtained by cross-validation, and the red points are the results of the validation set predicted by the CWT-MC-UVE model of LS-SVR. The solid-lines are obtained by least squares fitting of the points of the two sets, respectively. Apparently, all the points are reasonably distributed along the lines, and the two lines are close to each other. The results demonstrate that the model is applicable for the determination of the triglyceride content in serum samples.

In conclusion, a method for determination of triglyceride in human serum was developed by using NIRDRS. In the proposed method, silver mirror was used as the substrate for the spectral measurement. Due to the high reflectance of silver, the background in the spectra was reduced and the spectral feature of samples was enhanced. On the other hand, both linear and non-linear modeling strategies were investigated for building the model, CWT was adopted for further improving the quality of the spectra and variable selection techniques were employed for building the optimized model. CWT was proved to be effective but it cannot change the linearity of the spectral responses. Variable selection can improve the performance of the model, and non-linear model is more suitable to model the spectra. Furthermore, an optimal LS-SVR model with the CWT preprocessing and the selected variable was obtained. The model can quantitative predict the content of triglyceride in serum samples with the maximal deviation 0.473 mmol/L. Therefore, the proposed method may provide an efficient tool for predicting the serum parameters in clinical diagnosis.

Acknowledgment

This work is supported by the National Natural Science Foundation of China (Nos. 21475068, 21775076).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.cclet.2018.01.016.

References

[1] M. Brandstetter, L. Volgger, A. Genner, et al., Appl. Phys. B 110 (2013) 233–239.
[2] P.A. da Costa Filho, R.J. Poppi, Anal. Chim. Acta 446 (2001) 39–47.
[3] J. Narang, Minakshi, M. Bhambi, C.S. Pundir, Int. J. Biol. Macromol. 47 (2010) 691–695.
[4] R.N. Qureshi, W.Th. Kok, P.J. Schoenmakers, Anal. Chim. Acta 654 (2009) 85–91.
[5] L.B. Foster, R.T. Dunn, Clin. Chem. 19 (1973) 338–340.
[6] J.W. Hall, A. Pollard, Clin. Biochem. 26 (1993) 483–490.
[7] K.H. Hazen, M.A. Arnold, G.W. Small, Anal. Chim. Acta 371 (1998) 255–267.
[8] H.M. Heise, R. Marbach, A. Bittner, J. Near Infrared Spectrosc. 6 (1998) 361–374.
[9] K. Maruo, M. Tsurugi, M. Tamura, et al., Appl. Spectrosc. 57 (2003) 1236–1244.
[10] J. Chen, M.A. Arnold, G.W. Small, Anal. Chem. 76 (2004) 5405–5413.
[11] L.J. Yao, N. Lyu, J.M. Chen, et al., Spectrochim. Acta Part A 159 (2016) 53–59.
[12] T. Kobayashi, Y.H. Kato, M. Tsukamoto, et al., Int. J. Mol. Med. 23 (2009) 75–79.
[13] J.S. Albuquerque, M.F. Pimentel, V.L. Silva, et al., Anal. Chem. 77 (2005) 72–77.
[14] K.M.G. Lima, I.M. Raimundo Jr., M.F. Pimentel, Sens. Actuators B 125 (2007) 229–233.
[15] Z.X. Huang, W. Tao, J.J. Fang, et al., Chemom. Intell. Lab. Syst. 98 (2009) 195–200.
[16] N. Sheng, W.S. Cai, X.G. Shao, Talanta 79 (2009) 339–343.
[17] J.H. Li, Y. Zhang, W.S. Cai, et al., Talanta 84 (2011) 679–683.
[18] Y. Ning, J.H. Li, W.S. Cai, et al., Spectrochim. Acta Part A 96 (2012) 289–294.
[19] F.X. Liu, W.S. Cai, X.G. Shao, Vib. Spectrosc. 68 (2013) 104–108.
[20] Y. Zhang, Y. Hao, W.C. Cai, et al., Anal. Methods 3 (2011) 703–708.
[21] Y.F. Yang, J.R. Tu, W.S. Cai, et al., Talanta 99 (2012) 871–874.
[22] C.C. Wang, W.S. Cai, X.G. Shao, J. Near Infrared Spectrosc. 24 (2016) 345–352.
[23] X. Ma, C.C. Wang, W.S. Cai, et al., Chin. Chem. Lett. 27 (2016) 1597–1601.
[24] C.J. Cui, W.S. Cai, X.G. Shao, Chin. Chem. Lett. 24 (2013) 67–69.
[25] C.C. Wang, S.Y. Wang, W.S. Cai, et al., Talanta 162 (2017) 123–129.
[26] C.C. Wang, W.S. Cai, X.G. Shao, Anal. Lett. 51 (2017) 537–546.
[27] C.C. Wang, X.M. Yu, W.S. Cai, et al., Chem. J. Chin. Univ. 38 (2017) 1947–1952.
[28] S. Wold, Technometrics 20 (1978) 397–405.
[29] D.F. Malley, P.C. Williams, Environ. Sci. Technol. 31 (1997) 3461–3467.
[30] M. Forouzangohar, D. Cozzolino, R.S. Kookana, et al., Environ. Sci. Technol. 43 (2009) 4049–4055.
[31] W. Liu, X.Y. Cui, X.M. Yu, et al., Chin. Chem. Lett. 28 (2017) 1447–1452.
[32] Y.N. Ni, W. Lin, Chin. Chem. Lett. 22 (2011) 1473–1476.
[33] G. Shao, A.K.M. Leung, F.T. Chau, Acc. Chem. Res. 36 (2003) 276–283.
[34] G. Shao, C.X. Ma, Chemom. Intell. Lab. Syst. 69 (2003) 157–165.
[35] C.X. Ma, X.G. Shao, J. Chem. Inf. Comput. Sci. 44 (2004) 907–911.
[36] D. Chen, B. Hu, X.G. Shao, et al., Analyst 129 (2004) 664–669.
[37] W.S. Cai, Y.K. Li, X.G. Shao, Chemom. Intell. Lab. Syst. 90 (2008) 188–194.
[38] H. Xu, Z.C. Liu, W.S. Cai, et al., Chemom. Intell. Lab. Syst. 97 (2009) 189–193.
[39] H.D. Li, Y.Z. Liang, Q.S. Xu, et al., Anal. Chim. Acta 648 (2009) 77–84.
[40] J. Zhang, X.Y. Cui, W.S. Cai, et al., J. Chemom. (2017), doi:http://dx.doi.org/10.1002/cem.2971.