

AO–MW–PLS method applied to rapid quantification of teicoplanin with near-infrared spectroscopy

Jiemei Chen*, Tian Ai*, Tao Pan^{†,§}, Lijun Yao[†] and Fenggeng Xia[‡]

**Department of Biological Engineering, Jinan University
Huangpu Road West 601 Tianhe District, Guangzhou 510632, P. R. China*

*†Department of Optoelectronic Engineering, Jinan University
Huangpu Road West 601, Tianhe District, Guangzhou 510632, P. R. China*

*‡Guangzhou Institutes of Microbiology, Jian Tashan Road 1
Science City, Luogang District, Guangzhou 510530, P. R. China*

§tpan@jnu.edu.cn

Received 12 November 2015

Accepted 18 January 2016

Published 14 March 2016

Teicoplanin (TCP) is an important lipoglycopeptide antibiotic produced by fermenting *Actinoplanes teichomyceticus*. The change in TCP concentration is important to measure in the fermentation process. In this study, a reagent-free and rapid quantification method for TCP in the TCP–Tris–HCl mixture samples was developed using near-infrared (NIR) spectroscopy by focusing our attention on the fermentation process for TCP. The absorbance optimization (AO) partial least squares (PLS) was proposed and integrated with the moving window (MW) PLS, which is called AO–MW–PLS method, to select appropriate wavebands. A model set that includes various wavebands that were equivalent to the optimal AO–MW–PLS waveband was proposed based on statistical considerations. The public region of all equivalent wavebands was just one of the equivalent wavebands. The obtained public regions were 1540–1868 nm for TCP and 1114–1310 nm for Tris. The root-mean-square error and correlation coefficient for leave-one-out cross validation were 0.046 mg mL⁻¹ and 0.9998 mg mL⁻¹ for TCP, and 0.235 mg mL⁻¹ and 0.9986 mg mL⁻¹ for Tris, respectively. All the models achieved highly accurate prediction effects, and the selected wavebands provided valuable references for designing specialized spectrometers. This study provided a valuable reference for further application of the proposed methods to TCP fermentation broth and to other spectroscopic analysis fields.

Keywords: Teicoplanin; near-infrared spectroscopic analysis; absorbance optimization partial least squares; moving window partial least squares; equivalent model set.

[§]Corresponding author.

This is an Open Access article published by World Scientific Publishing Company. It is distributed under the terms of the Creative Commons Attribution 4.0 (CC-BY) License. Further distribution of this work is permitted, provided the original work is properly cited.

1. Introduction

Antibiotics have played a significant role in relieving bacterial infections in human beings in recent decades. However, overuse of antibiotics causes a strong selective pressure. Therefore, antibiotic resistance spreads rapidly in common pathogens. Teicoplanin (TCP) is a novel lipoglycopeptide antibiotic produced by *Actinoplanes teichomyceticus*.¹ TCP is used to treat a variety of aerobic and anaerobic gram-positive infections, especially those caused by methicillin-resistant *Staphylococcus aureus*.² Vancomycin is an antibiotic treatment for multiple drug-resistant bacteria infection, which was recognized by the international community. Compared with vancomycin, TCP displays a similar antibacterial activity, a similar mechanism of action, and similar or better clinical efficacy, but it has a lower adverse drug reaction rate, lower toxicity, especially lower renal toxicity, and a more convenient and efficient route of administration by intravenous or intramuscular injection.³ Thus, TCP has important economic values and application prospects.

In the fermentation process of TCP, real-time measurement of various substances (such as biomass, nutrients, metabolites, etc.) in fermentation broth can contribute to understanding fermentation dynamic and controlling production quality. The change in TCP concentration is important to measure. The relationship between TCP concentration and medication time in experimental animals (e.g., mice) based on different injection ways is also crucial in the pharmacokinetics of TCP and can provide a basis for optimal clinical drug dosing.⁴ For critically ill patients in intensive care units, therapeutic drug monitoring of TCP concentration is necessary to determine whether the serum trough levels (C_{\min}) achieve the effective therapeutic levels and to avoid the accumulation of drug toxicity in the bodies of patients.⁵

Hence, measuring the concentration of TCP in complex systems (e.g., fermentation broth or blood) rapidly is necessary. The existing available methods for assaying TCP are high-performance liquid chromatography (HPLC), microbiological assay and liquid chromatography-mass spectrometry.⁶⁻⁸ Among them, HPLC is the main measurement method for TCP because of its high separation efficiency and good selectivity. However, these methods involve a tedious variety of reagents or

lengthy run times, which are unsuitable for rapid measurement of TCP in the production process or in blood. Therefore, a reagent-free and rapid method for TCP measurement has an important significance and applied values.

Near-infrared (NIR) spectroscopy primarily reflects the absorption of overtones and the combination of the vibrations of the X-H functional groups (such as C-H, O-H and N-H). Given that NIR absorption strength is weak, most sample types can be measured directly without preprocessing and reagents. The NIR spectroscopic analysis is a rapid and reagentless technique that has been proven to be a powerful analytical tool for use in agriculture,⁹⁻¹³ food,^{14,15} environment,^{16,17} medicine,^{18,19} pharmaceuticals,²⁰ and so on. However, to the best of our knowledge, a quantification method for TCP using NIR spectroscopy has not been developed yet.

TCP is a complex compound (with a molecular weight of approximately 1900), short for TA₂, and is made of five components (TA₂₋₁, TA₂₋₂, TA₂₋₃, TA₂₋₄ and TA₂₋₅) with similar chemical structures. TA₂₋₂ is its main component, which accounts for 40% of all samples. The sixth active component, named TA₃₋₁, is the production of deacylation glucosamine. Two more lipophilic and designated related substances (RS-1 and RS-2) and two less lipophilic related substances (RS-3 and RS-4) exist in TCP crude products and purified products.² TCP molecular contains the C-H, O-H and N-H functional groups (Fig. 1) that have significant NIR absorption. In this study, we aimed to confirm the feasibility of the reagent-free quantitative analysis of TCP with NIR spectroscopy.

The Tris-HCl buffer, with a stable nature, is often used in biochemistry and molecular biology experiments. This buffer is suitable for use in simulating the physiological environment of a living body, such as the enzyme reaction in cell sap.^{21,22} In this study, the Tris-HCl buffer was used to simulate the background liquid of TCP. A quantitative analysis of a mixture of TCP with the Tris-HCl buffer (denoted as the TCP-Tris-HCl mixture) was performed to explore the feasibility of the quantitative analysis of TCP in complex systems (such as fermentation broth or blood). The mixture samples were prepared with different concentrations of Tris to make them suitable for different backgrounds. The simultaneous quantification of TCP and Tris was achieved.

A reagent-free and rapid measurement method has obvious advantages for drug monitoring, but it is also a difficult methodology. For a complex analyte with multiple components (e.g., fermentation broth), spectroscopic analysis of some components must mitigate the noise disturbance of other components. Therefore, appropriate chemometric methods are necessary to extract information and to eliminate noise.

Partial least squares (PLS) is widely used for comprehensively screening spectroscopic data, extracting information variables and overcoming spectral co-linearity. However, waveband selection is necessary because the prediction effect of PLS is difficult to improve when the signal-to-noise ratio of a waveband is not adequately high.^{10,12,17,18} PLS analysis combined with waveband selection is important for spectral analysis of complex systems. Moving window PLS (MW-PLS), a well-performed and PLS-based method, is proven as an effective method for waveband selection in the analysis of many objects.^{10-12,17,18,23}

Absorbance is another important factor to consider because high absorbance leads to strong spectral absorption and noise interference, whereas low absorbance results in missing certain information. Each wavelength corresponds to an absorbance value in a spectrum. Therefore, performing wavelength selection based on absorbance is crucial. In this study, a wavelength selection method called absorbance optimization PLS (AO-PLS) was proposed on the basis of selected absorbance ranges. Given that different wavelengths may correspond to almost the same absorbance value, a range of absorbance values may correspond to a combination of multiple wavebands. The use of AO-PLS helped us easily select an appropriate combination of several wavebands.

In this study, AO-PLS was integrated with MW-PLS to select appropriate wavebands to analyze TCP and Tris. AO-PLS was first employed to eliminate noise wavebands. Within the waveband combinations determined by AO-PLS, MW-PLS was performed to select information wavebands.

2. Materials and Methods

2.1. *Experimental materials, instruments and measurement methods*

TCP standard products was purchased from National Institutes for Food and Drug Control

(Beijing, China). Tris was analytical reagent. The structures of TCP and Tris were shown in Figs. 1 and 2, respectively.

In this study, TCP-Tris-HCl mixtures were prepared to simulate TCP fermentation broth. Samples were formulated considering that the concentration of TCP can reach 3.2 mg mL⁻¹ and that the pH values were approximately 7.2 in fermentation broth.^{24,25} The specific steps were as follows:

- (1) Ten Tris-HCl buffer solutions (100 mL⁻¹) were prepared using Tris, double-distilled water and 1 N hydrochloric acid. Their pH values were regulated to 7.2, and the concentration of Tris was 13–40 mg mL⁻¹. These solutions were numbered 1–10 according to ascending order of Tris concentration.
- (2) Seventy two TCP aqueous solutions were prepared using TCP standard products and double-distilled water, in which the concentration of TCP was 0.681–19.626 mg mL⁻¹.
- (3) These TCP aqueous solutions were numbered 1–72 according to descending order of TCP concentration. From the first solution to the 70th solution, one solution was extracted every ninth. The extracted seven solutions composed the first group and the second group until the ninth one; they were prepared in the same way. The remaining nine solutions composed the 10th group, and the TCP concentration of each group distributed from high to low.
- (4) Seventy-two TCP-Tris-HCl mixtures were prepared by TCP aqueous solutions (0.4 mL) from one group and Tris-HCl buffer solutions (0.4 mL) with the corresponded serial number. The statistical analysis and two-dimensional (2D) distribution of the concentrations of TCP and Tris in the 72 mixture samples are shown in Table 1 and Fig. 3, respectively. Figure 3 shows that the concentrations of TCP and Tris were uniformly distributed in the 72 mixture samples.

The samples were then used for spectrometry measurement. The spectroscopy instrument that was used was an XDS Rapid ContentTM Liquid Grating Spectrometer (FOSS, Denmark) equipped with transmission accessory and a 2 mm cuvette. The scanning scope of the spectrum was 400–2498 nm with a 2 nm wavelength interval. The scanning

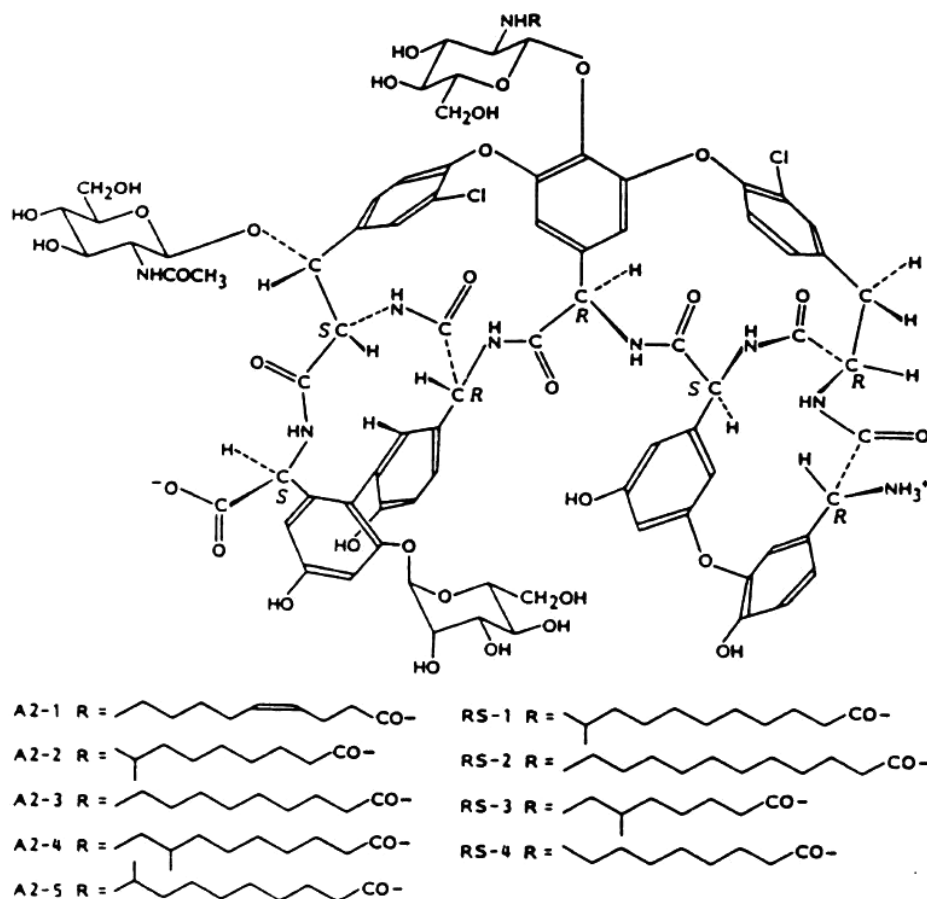


Fig. 1. Structure of TCP.

scope included the entire NIR region and a large part of the visible region. Wavebands of 400–1100 nm and 1100–2498 nm were used for Si and PbS detection, respectively. Each sample was

scanned thrice, and the mean value of the three measurements was used for modeling. Spectra were obtained at $25 \pm 1^\circ\text{C}$ and a relative humidity of $45 \pm 1\%$.

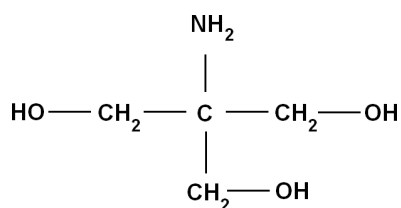


Fig. 2. Structure of Tris.

Table 1. Statistical analysis of actual concentrations of TCP and Tris in 72 TCP–Tris–HCl mixture samples.

Indicator	Min (mg mL ⁻¹)	Max (mg mL ⁻¹)	Mean (mg mL ⁻¹)	SD
TCP	0.338	9.805	4.114	2.225
Tris	6.272	20.561	13.438	4.505

Note: SD is the abbreviation of standard deviation.

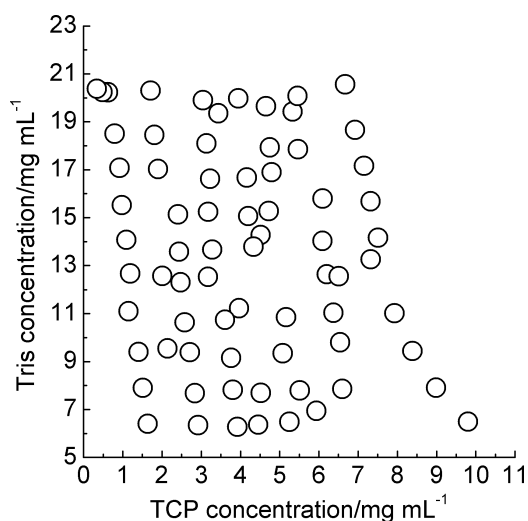


Fig. 3. 2D distribution of TCP and Tris concentrations in 72 TCP–Tris–HCl mixture samples.

2.2. Leave-one-out cross validation based on the PLS model

All n samples was subjected to leave-one-out cross validation (LOOCV) based on the PLS model. The model parameters, such as the waveband combination and the number of PLS factor (F), were optimized according to the prediction effect. The specific procedure was as follows:

First, each sample was left out from all n samples to predict. The remaining $n-1$ samples were used to construct a calibration model. The prediction value of the leave sample was calculated by the calibration model. The prediction values of all n samples were calculated based on the same process. The actual and predicted values for i th sample were denoted as C_i, \tilde{C}_i , respectively, where $i = 1, 2, \dots, n$. The mean actual and predicted values of all samples were denoted as C_{Ave}, \tilde{C}_{Ave} , respectively. The prediction accuracy was evaluated by the root-mean-square error (SECV), the correlation coefficient (R_{CV}) and relative root-mean-square error (RSECV), respectively. The calculation formulas were as follows:

$$SECV = \sqrt{\frac{\sum_{i=1}^n (\tilde{C}_i - C_i)^2}{n}}, \quad (1)$$

$$R_{CV} = \frac{\sum_{i=1}^n (C_i - C_{Ave})(\tilde{C}_i - \tilde{C}_{Ave})}{\sqrt{\sum_{i=1}^n (C_i - C_{Ave})^2(\tilde{C}_i - \tilde{C}_{Ave})^2}}, \quad (2)$$

$$RSECV = \frac{SECV}{C_{Ave}} \times 100\%. \quad (3)$$

In this paper, in order to avoid overfitting, the number of PLS factor (F) was set to $F \in \{1, 2, \dots, 10\}$.

2.3. Proposed AO-PLS method

Lambert Beer's law is described by the following equation:

$$A(\lambda) = -\lg\left(\frac{I_1(\lambda)}{I_0(\lambda)}\right) = \lg\left(\frac{1}{T(\lambda)}\right), \quad (4)$$

where λ is the wavelength; $A(\lambda)$ is the absorbance; $I_0(\lambda)$ and $I_1(\lambda)$ are the intensity of incident light and the intensity of transmitted light through the

sample, respectively; and $T(\lambda)$ is the transmittance, i.e., the ratio of transmitted light intensity and incident light intensity. Equation (4) can then be expressed as follows:

$$T(\lambda) = \frac{I_1(\lambda)}{I_0(\lambda)} = 10^{-A(\lambda)} \times 100\%. \quad (5)$$

According to the above equation, e.g., when $A(\lambda) = 4$, the transmitted light intensity was merely one-ten 1000th of the incident light intensity, i.e., the 99.99% of the incident light was absorbed by the sample. In this case, the transmitted light was weak and difficult to detect; it would thus be likely to cause noise in the spectrum. When $A(\lambda) \approx 0$, the transmitted light intensity was approximately equal to the incident light intensity, i.e., almost no incident light was absorbed; consequently, the sample information cannot be detected. Therefore, wavelength selection with appropriate absorbance values, which correspond to a high quality of sample information and low levels of noise, is necessary. In this study, a PLS-based method (AO-PLS) for wavelength selection was proposed on the basis of the selection of the upper bound of absorbance, which can appropriately remove noise bands. The specific steps were as follows:

Step 1. A region of wavelength screening (Δ) was set in advance for the entire scanning region according to the physical and chemical characteristics of the measured objects and the instrument properties. Meanwhile, in the average spectrum for all samples within the region Δ , the minimum and maximum values of absorbance were denoted as A_{\min} and A_{\max} , respectively. An appropriate step of absorbance (ε) was set.

Step 2. Some value of A^* was set, i.e., $A^* \geq A_{\min}$, and the upper bound of absorbance A_{upper} was changed from A^* to A_{\max} with the step ε . According to the relationship between the wavelength and the absorbance within the region Δ , for each A_{upper} , the absorbance interval ($A_{\min}, A_{\text{upper}}$) corresponded to a waveband combination.

Step 3. Every obtained waveband combination was employed to establish LOOCV based on the PLS model. The corresponding SECV and R_{CV} values were then calculated.

Step 4. According to the minimum SECV, the optimal A_{upper} was determined, and the corresponding waveband combination ($A_{\min}, A_{\text{upper}}$) was selected.

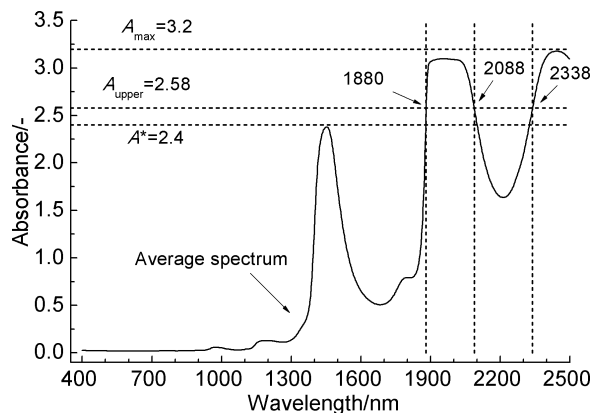


Fig. 4. Sketch map for the relationship between wavelength and absorbance.

In this study, the region Δ was set to be the entire scanning region (400–2498 nm) with 1050 wavelengths (also shown in Fig. 4). A_{\min} was greater than or close to 0, and A_{\max} value was less than or close to 3.2. Therefore, A_{\min} and A_{\max} were set to 0 and 3.2, respectively. Another obvious absorption peak with an absorbance value of 2.4 existed around 1458 nm, which corresponded to the absorption of water molecules and other components in samples. A^* value was set as 2.4 (namely, $A_{\text{upper}} > 2.4$) to retain the relevant information of the region because the main purpose was to remove the noise bands with saturate absorption. The absorbance step ε was set to 0.01, and the number of PLS factors (F) was set to $F \in \{1, 2, \dots, 10\}$. Figure 4 shows a sketch map of the relationship between wavelength and absorbance for the case in which the absorbance value $A_{\text{upper}} = 2.58$ and the corresponding waveband combinations were 400–1880 nm and 2088–2338 nm.

2.4. Integrated AO–MW–PLS method

For MW-PLS, consecutive spectral data on N adjacent wavenumbers were designated as a window. For all the windows in a predetermined search region of the spectrum, PLS models were established, and the optimal analytical wavebands were selected by moving and varying the window size (see also Fig. 5). By considering the position and the length of the wavebands, as well as the PLS factor, we set the search parameters of MW-PLS as follows: (1) the initial wavenumber (I), (2) the number of wavenumbers (N) and (3) the number of PLS factors (F). The search range of the parameters I , N and F

can be selected according to the actual chemical, physical and statistical significance. The LOOCV based on the PLS model was established for any combinations (I , N , F). The corresponding SECV and R_{CV} values were then calculated. The optimal waveband with minimum SECV was selected.

In this study, the waveband combination determined with the AO-PLS method was used as the screening region of the MW-PLS method, which was called the AO-MW-PLS method. The parameters I , N and F were set to $I \in \{400, 402, \dots, 1880\} \cup \{2088, 2090, \dots, 2338\} \triangleq I$, $N \in \{1, 2, \dots, 867\} \triangleq N$ and $F \in \{1, 2, \dots, 10\} \triangleq F$. Figure 5 shows a sketch map of the waveband and the screening mode for MWs for the case $I = 1494$ nm and $N = 188$. The computer algorithms were designed using MATLAB 7.6.

Selection of the number of PLS factors: The number of PLS factors (F) is an important parameter that corresponds to the number of integrated spectral variables that correspond to sample information. The selection of a reasonable F is necessary but difficult.^{10,12,17,18,23} Each waveband corresponded to a unique combination of parameters (I , N) = (I_0 , N_0); the optimal F was determined according to the following expression:

$$\text{SECV}(I_0, N_0) = \min_F \text{SECV}(I_0, N_0, F). \quad (6)$$

Global optimal model: The global optimal model was selected according to the following equation:

$$\text{SECV}_* = \min_{I,N,F} \text{SECV}(I, N, F). \quad (7)$$

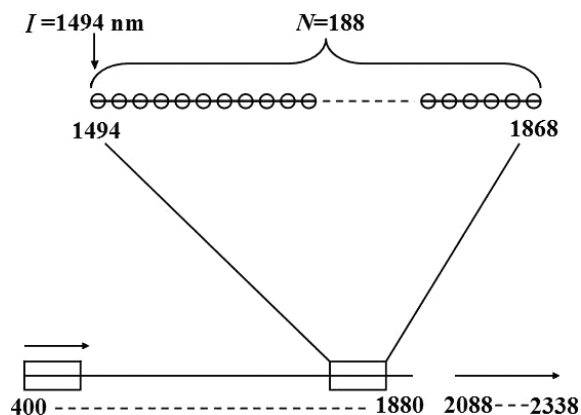


Fig. 5. Schematic for the waveband and screening mode for MWs.

Local optimal model corresponding to a single parameter: The instrument design typically involves some limitations of the position and the number of wavenumbers (such as costs and material properties). At some instances, the demand of actual conditions is not met by the global optimal waveband. Therefore, local optimal wavebands that correspond to different positions and numbers of wavelengths are significant. For any fixed $I = I_0$, the local optimal model was selected according to the following equation:

$$\text{SECV}(I_0) = \min_{N,F} \text{SECV}(I_0, N, F). \quad (8)$$

Meanwhile, for any fixed $N = N_0$, the local optimal model was selected according to the following expression:

$$\text{SECV}(N_0) = \min_{I,F} \text{SECV}(I, N_0, F). \quad (9)$$

Proposed equivalence model set: As previously mentioned, the global optimal waveband can be selected according to the minimum SECV; however, the models with insignificantly fluctuating prediction accuracy are statistically considered equivalent because the samples are random and limited. Therefore, the optimal SECV value can float upward a certain percentage. A model set that includes various wavebands that were equivalent to the optimal AO–MW–PLS waveband was also proposed. The specific steps were as follows:

All parameter combinations (I, N, F) , called the entire model set, can be expressed as follows:

$$\Omega = \{(I, N, F) | I \in I, N \in N, F \in F\}. \quad (10)$$

The equivalence model set that corresponds to a certain percentage (α) can be expressed as follows:

$$\begin{aligned} \Omega_\alpha &= \{(I, N, F) | \text{SECV}(I, N, F) \\ &\leq \text{SECV}_*(1 + \alpha)\}. \end{aligned} \quad (11)$$

α value was set according to the actual situation, in which a low α corresponded to a good prediction effect for the equivalent models.

3. Results and Discussion

The NIR spectra of 72 TCP–Tris–HCl mixture samples in the entire scanning region (400–2498 nm) are shown in Fig. 6(a). For comparison, the NIR spectrum of double-distilled water is shown in Fig. 6(b), and it shows that the absorption of water

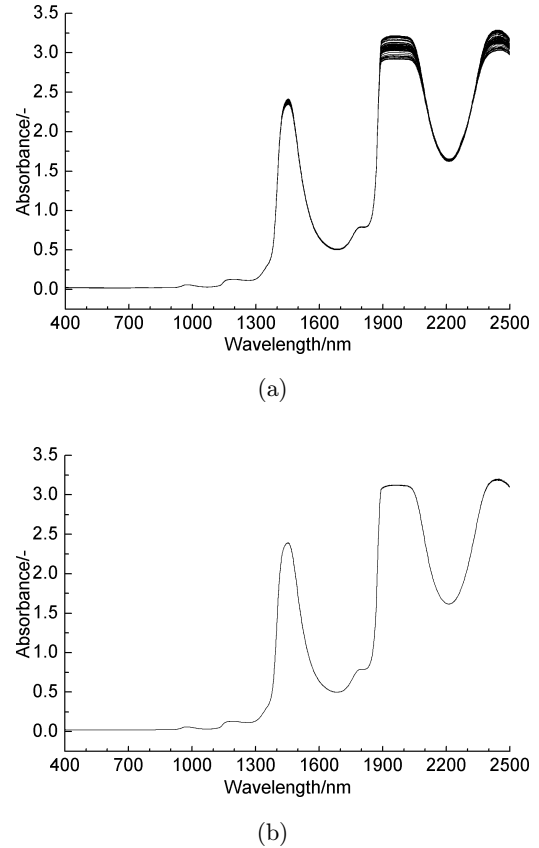


Fig. 6. Spectra in the entire scanning region (400–2498 nm) for (a) 72 TCP–Tris–HCl mixture samples and (b) double-distilled water.

molecules is located around 1400 nm and 1900 nm. The same absorption peaks can be found in Fig. 6(a). As indicated in Figs. 6(a) and 6(b), two saturate absorption regions appeared at around 2000 nm and 2430 nm, which were caused by strong absorption and could lead to strong spectral noise interference.

3.1. Full PLS models

Quantitative analyses for the TCP and Tris in a TCP–Tris–HCl mixture were conducted. For comparison, PLS models based on the entire scanning region, called full PLS model, were first established. The corresponding model parameters and the prediction effects (SECV, R_{CV} and RSECV) are summarized in Table 2 for the two indicators. Results showed that the R_{CV} values all achieved above 0.9895 for the two indicators. The RSECV values of TCP was greater than 7.9%, which indicated that the prediction error was significant. The full PLS models employed 1050 wavelengths that contained noise wavebands with high absorption, which led to

Table 2. Parameter selection and prediction effects of the full PLS models for TCP and Tris.

Indicator	Waveband (nm)	N	F	SECV (mg mL ⁻¹)	R_{CV}	RSECV (%)
TCP	400–2498	1050	5	0.325	0.9895	7.9
Tris			7	0.355	0.9969	2.6

high model complexity and low prediction effect. Therefore, AO-PLS was performed to eliminate the noise wavebands with high absorption, and appropriate waveband combinations were selected to improve prediction effects.

3.2. Elimination of noise waveband with AO-PLS

The AO-PLS mentioned in Sec. 2.3 was performed to eliminate noise wavebands. For TCP and Tris, the SECV values for each upper bound of absorbance (A_{upper}) are shown in Fig. 7. Results showed that the minimum SECV was achieved when $A_{upper} = 2.58$ for Tris. For TCP, SECV values were all small and close to others when $A_{upper} < 3.0$, and they gradually became larger when $A_{upper} > 3.0$. Consequently, the minimum SECV was not clearly observed for TCP. The minimum SECV was achieved when $A_{upper} = 2.58$ approximately for both TCP and Tris, and the corresponding waveband combinations were 400–1880 nm and

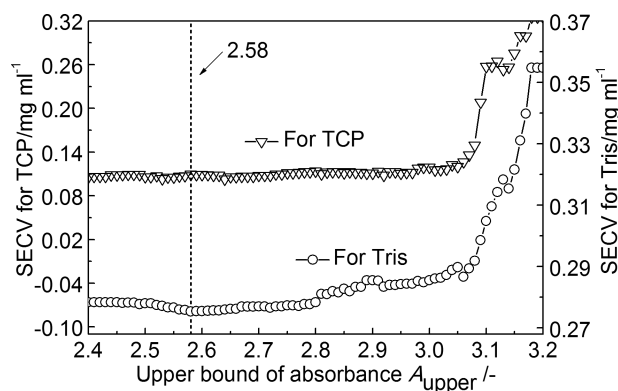


Fig. 7. SECV values for each upper bound of absorbance A_{upper} with the AO-PLS method for TCP and Tris.

2088–2338 nm. The corresponding model parameters and prediction effects for TCP and Tris are summarized in Table 3.

Compared with Table 2, Table 3 indicates that the R_{CV} values all increased to above 0.9981 for the two indicators, and the RSECV values were all less than 2.7%, which indicated that the prediction errors were significantly reduced. The number of the adopted wavelengths was reduced to 867. Results showed that the prediction effects were all significantly improved for the two indicators.

3.3. Information waveband selection with AO-MW-PLS

In Sec. 3.2, AO-PLS was performed to eliminate the noise wavebands in the vertical direction. Selecting the information wavebands for TCP and Tris in the horizontal direction was then necessary. Within the waveband combinations selected by AO-PLS, the MW-PLS mentioned in Sec. 2.4 was performed to select the information waveband. The obtained optimal parameters I and N were 1494 nm and 188 nm for TCP and 1102 nm and 106 nm for Tris, respectively. The corresponding wavebands were 1494–1868 nm for TCP and 1102–1312 nm for Tris. The optimal model parameters I , N and F and the prediction effects for the two indicators are summarized in Table 4.

Compared with Table 3, Table 4 indicates that the R_{CV} values further increased to above 0.9987 for the two indicators, and the RSECV values were all less than 1.7%, which indicated that the prediction errors were significantly reduced. The optimal AO-MW-PLS models employed 188 or less wavelengths. Results showed that the prediction effects were all further improved for the two indicators,

Table 3. Parameter selection and prediction effects of the AO-PLS models for TCP and Tris.

Indicator	Waveband (nm)	A_{upper}	N	F	SECV (mg mL ⁻¹)	R_{CV}	RSECV (%)
TCP	(400–1880)&(2088–2338)	2.58	867	8	0.110	0.9988	2.7
Tris				7	0.275	0.9981	2.0

Table 4. Parameter selection and prediction effects of the optimal AO-MW-PLS models for TCP and Tris.

Indicator	Waveband (nm)	N	F	SECV (mg mL^{-1})	R_{CV}	RSECV (%)
TCP	1494–1868	188	6	0.046	0.9998	1.1
Tris	1102–1312	106	7	0.227	0.9987	1.7

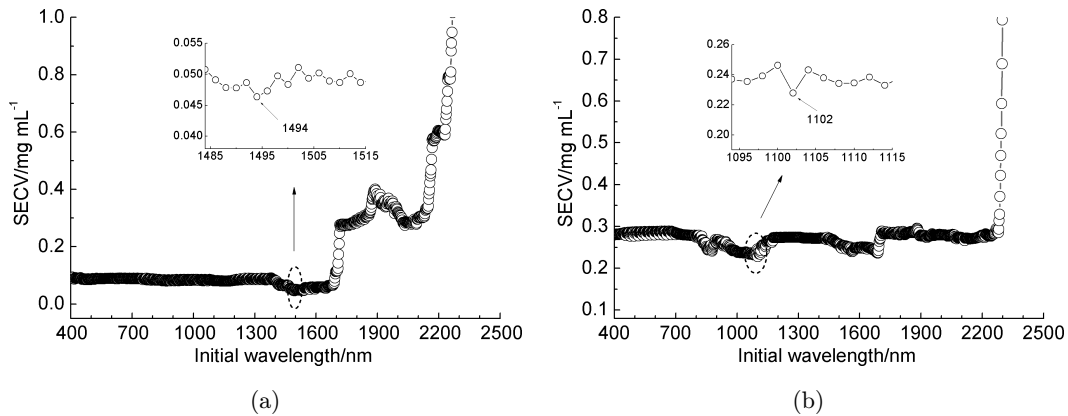


Fig. 8. Optimal SECV corresponding to the initial wavelength for (a) TCP and (b) Tris.

and the model complexity was significantly reduced.

For the two indicators, the SECV values of a local optimal model corresponding to each initial wavelength I are shown in Figs. 8(a) and 8(b). In Fig. 8, the SECV values of some I were large [such as $I > 1750$ nm in Fig. 8(a)], and the prediction effects were poor. By contrast, the SECV values of some I were small and close to the global optimal value [such as $1485 < I < 1515$ nm in Fig. 8(a)].

For the two indicators, the SECV values of the local optimal model corresponding to each number of wavelengths N are shown in Figs. 9(a) and 9(b). In Fig. 9, the global optimal SECV values were

achieved when N were 188 and 106. The SECV values were large for the large or small N values. For N values between 50 and 200, the SECV values were all small and close to the global optimal value.

These local optimal models that correspond to a single parameter may serve as a valuable reference for designing splitting systems for spectroscopic instruments. The local optimal models with prediction effects close to one of the global optimal models remain good choices, and they address restrictions, such as cost and material properties, as well as the position and the number of wavelengths in instrument design.

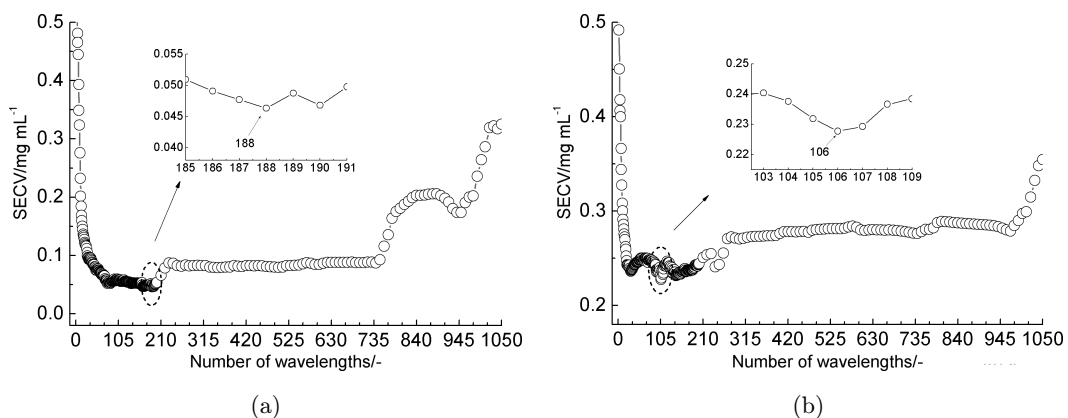


Fig. 9. Optimal SECV corresponding to the number of wavelengths for (a) TCP and (b) Tris.

3.4. Equivalence model set

The method mentioned in Sec. 2.4 was performed to select an equivalence model set for the two indicators and to evaluate the model equivalence clearly.

For TCP, the global optimal SECV value (denoted as $SECV_*$) was 0.046, α value was set as 6%, and the corresponding equivalence model set was as follows:

$$\Omega_\alpha = \{(I, N, F) | SECV(I, N, F) \leq 0.049\}. \quad (12)$$

The model set included 12 wavebands that were equivalent to the optimal AO–MW–PLS waveband. Their positions are shown in Fig. 10(a) starting from left to right in the order of initial wavelengths. The corresponding minimum and maximum values of N were 165 and 190, respectively. The public range of the 12 equivalent wavebands was 1540–1868 nm, and it was just one of the equivalent wavebands. Thus, the waveband 1540–1868 nm can be employed instead of other equivalent wavebands, which contained enough information. And TCP molecular contains the ArCH and CH₂ functional groups (seen in Fig. 1). The NIR absorption band of these groups located in 1500–1900 nm region²⁶ which are almost consistent with the selected waveband 1540–1868 nm for TCP. The corresponding model parameters (I, N, F) and the prediction effects are summarized in Table 5.

For Tris, the global optimal SECV value ($SECV_*$) was 0.227, α value was set as 3.5%, and the corresponding equivalence model set was as follows:

$$\Omega_\alpha = \{(I, N, F) | SECV(I, N, F) \leq 0.235\}. \quad (13)$$

The model set included 25 wavebands. Their positions are shown in Fig. 10(b) starting from left to right in the order of initial wavelengths. The corresponding minimum and maximum values of N were 99 and 158, respectively. The public range of the 25 equivalent wavebands was 1114–1310 nm, and it was just one of the equivalent wavebands. Thus, the waveband 1114–1310 nm can be employed instead of other equivalent wavebands, which contained enough information. And Tris molecular contains the CH₂ functional groups (seen in Fig. 2). The NIR absorption band of this group is also located in 1100–1300 nm region²⁶ which are almost consistent with the selected waveband 1114–1310 nm for Tris. The corresponding model parameters and the prediction effects are also summarized in Table 5.

Compared with Table 4, Table 5 indicates that, for the two indicators, the prediction effects of the public wavebands were close to those of the optimal AO–MW–PLS wavebands. Thus, the formers contained enough information and avoided redundant wavelengths.

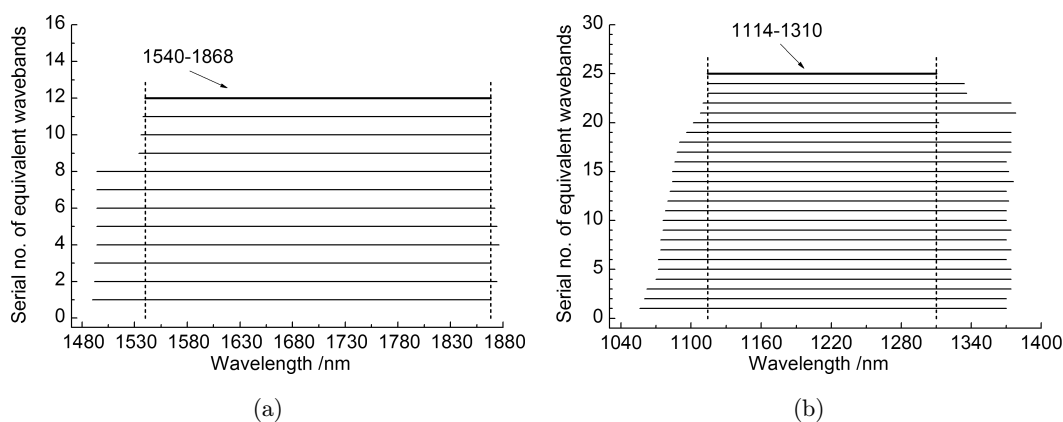


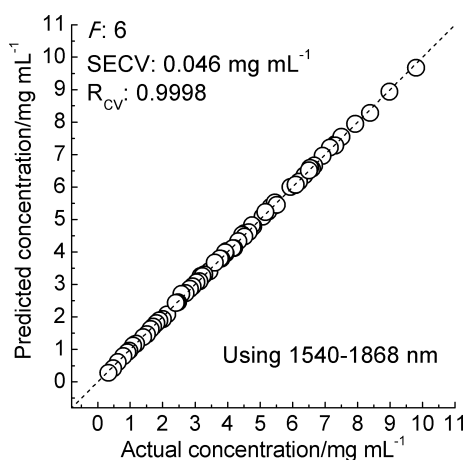
Fig. 10. Positions of equivalent wavebands for (a) TCP and (b) Tris.

Table 5. Parameter selection and prediction effects of the equivalent public wavebands for TCP and Tris.

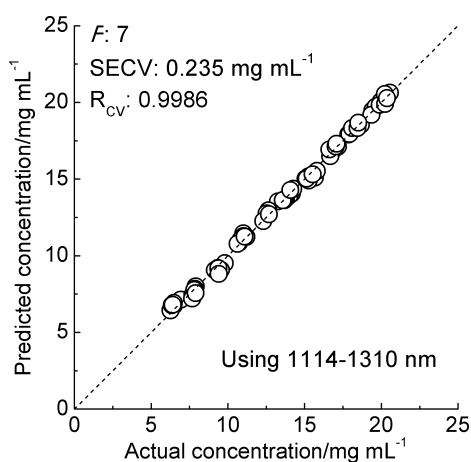
Indicator	Waveband (nm)	N	F	SECV (mg mL^{-1})	R_{CV}	RSECV (%)
TCP	1540–1868	165	6	0.046	0.9998	1.1
Tris	1114–1310	99	7	0.235	0.9986	1.8

3.5. Relationship between predicted and actual concentrations

The PLS models based on two public equivalent wavebands (1540–1868 nm and 1114–1310 nm) were used as examples. For the two indicators, the relationship between the predicted and actual concentrations are shown in Figs. 11(a) and 11(b). Figure 11 shows that the predicted and actual concentrations had a high correlation, and the predicted error was insignificant. Therefore, satisfactory prediction effects were achieved for the two indicators. The result indicated that the prediction was feasible in reagent-free quantitative analysis of TCP with NIR spectroscopy.



(a)



(b)

Fig. 11. Relationship between the predicted and actual concentrations of (a) TCP and (b) Tris.

4. Conclusion

A reagent-free and rapid quantification method for the TCP in the TCP–Tris–HCl mixture was developed based on NIR spectroscopy. The key technique was to apply chemometric tools to improve the NIR model. AO–PLS was proposed and integrated with MW–PLS to select appropriate wavebands. The obtained optimal AO–MW–PLS wavebands were 1494–1868 nm for TCP and 1102–1312 nm for Tris.

On the basis of statistical considerations, we also proposed a model set that includes various wavebands that were equivalent to the optimal AO–MW–PLS waveband. The public region of all equivalent wavebands was just one of the equivalent wavebands. The public waveband contained enough information and avoided redundant wavelengths, and the corresponding prediction effect was close to that of the optimal AO–MW–PLS model. The obtained public equivalent wavebands were 1540–1868 nm for TCP and 1114–1310 nm for Tris. These models all achieved high prediction effects.

The result showed that NIR spectroscopy combined with the proposed equivalence AO–MW–PLS method was successfully applied for reagent-free and accurate quantitative analysis of TCP and Tris. The integrated AO–MW–PLS method can select multi-band combination, it has broader applicability. The selected wavebands provided valuable references for designing specialized spectrometers. TCP–Tris–HCl mixture samples were used to simulate TCP fermentation broth. This study provided a valuable reference for the further application of the proposed methods to TCP fermentation broth and to other spectroscopic analysis fields.

Acknowledgments

This work was supported by the Science and Technology Project of Guangdong Province of China (Nos. 2014A020213016 and 2014A020212445), and the Science and Technology Project of Guangzhou of China (No. 2011Y5-00002).

References

1. F. Parenti, G. Beretta, M. Berti, V. Arioli, "Teichomycins, new antibiotics from *Actinoplanes teichomyceticus*, Nov. Sp. I. Description of the producer strain, fermentation studies and biological properties," *J. Antibiot.* **31**, 276–283 (1978).

2. A. Borghi, P. Antonini, M. Zanol, P. Ferrari, L. F. Zerilli, G. C. Lancini, "Isolation and structure determination of two new analogs of teicoplanin, a glycopeptide antibiotic," *J. Antibiot.* **42**, 361–366 (1989).
3. M. J. Wood, "The comparative efficacy and safety of teicoplanin and vancomycin," *J. Antimicrob. Chemother.* **37**, 209–222 (1996).
4. J. D. Knudsen, K. Fuursted, F. Espersen, N. Fridmodt-Moller, "Activities of vancomycin and teicoplanin against penicillin-resistant pneumococci in vitro and in vivo and correlation to pharmacokinetic parameters in the mouse peritonitis model," *Antimicrob. Agents Ch.* **41**, 1910–1915 (1997).
5. H. Li, X. Wang, H. Y. Dong, Y. L. Dong, G. Tian, "Clinical analysis of therapeutic drug monitoring and drug dosage for ICU patients treated with teicoplanin," *Chin. J. Clin. Pharmacol. Therapeut.* **16**, 538–544 (2011).
6. N. Mochizuki, K. Ohno, T. Shimamura, H. Furukawa, S. Todo, S. Kishino, "Quantitative determination of individual teicoplanin components in human plasma and cerebrospinal fluid by high-performance liquid chromatography with electrochemical detection," *J. Chromatogr. B* **847**, 78–81 (2007).
7. H. J. Chang, P. C. Hsu, C. C. Yang, L. K. Siu, A. J. Kuo, J. H. Chia, T. L. Wu, C. T. Huang, M. H. Lee, "Influence of teicoplanin MICs on treatment outcomes among patients with teicoplanin-treated methicillin-resistant *Staphylococcus aureus* bacteraemia: A hospital-based retrospective study," *J. Antimicrob. Chemother.* **67**, 736–741 (2012).
8. F. H. Y. Fung, J. C. Y. Tang, J. P. P. Hopkins, J. J. Dutton, L. M. Bailey, A. S. Davison, "Measurement of teicoplanin by liquid chromatography-tandem mass spectrometry: Development of a novel method," *Ann. Clin. Biochem.* **49**, 475–481 (2012).
9. R. Welle, W. Greten, B. Rietmann, S. Alley, G. Sinnavee, P. Dardenne, "Near-Infrared Spectroscopy on chopper to measure maize forage quality parameters online," *Crop. Sci.* **43**, 1407–1413 (2003).
10. H. Z. Chen, T. Pan, J. M. Chen, Q. P. Lu, "Waveband selection for NIR spectroscopy analysis of soil organic matter based on SG smoothing and MWPLS methods," *Chemometr. Intell. Lab.* **107**, 139–146 (2011).
11. J. M. Chen, T. Pan, G. S. Liu, Y. Hun, D. X. Chen, "Selection of stable equivalent wavebands for near-infrared spectroscopic analysis of total nitrogen in soil," *J. Innov. Opt. Health Sci.* **7**, 1350071–1–1350071–9 (2013).
12. T. Pan, M. M. Li, J. M. Chen, "Selection method of quasi-continuous wavelength combination with applications to the near-infrared spectroscopic analysis of soil organic matter," *Appl. Spectrosc.* **68**, 263–271 (2014).
13. H. S. Guo, J. M. Chen, T. Pan, J. H. Wang, G. Cao, "Vis-NIR wavelength selection for non-destructive discriminant analysis of breed screening of transgenic sugarcane," *Anal. Methods* **6**, 8810–8816 (2014).
14. J. Y. Chen, C. Iyo, F. Terada, S. Kawano, "Effect of multiplicative scatter correction on wavelength selection for near-infrared calibration to determine fat content in raw milk," *J. Near Infrared Spectrosc.* **10**, 301–307 (2002).
15. Z. Y. Liu, B. Liu, T. Pan, J. D. Yang, "Determination of amino acid nitrogen in tuber mustard using near-infrared spectroscopy with waveband selection stability," *Spectrochim. Acta A* **102**, 269–274 (2013).
16. A. C. Sousa, M. M. L. M. Lucio, O. F. Bezerra Neto, G. P. S. Marcone, A. F. C. Pereira, E. O. Dantas, W. D. Fragoso, M. C. U. Araujo, R. K. H. Galvao, "A method for determination of COD in a domestic wastewater treatment plant by using near-infrared reflectance spectrometry of seston," *Anal. Chim. Acta* **588**, 231–236 (2007).
17. T. Pan, Z. H. Chen, J. M. Chen, Z. Y. Liu, "Near-infrared spectroscopy with waveband selection stability for the determination of COD in sugar refinery wastewater," *Anal. Methods* **4**, 1046–1052 (2012).
18. T. Pan, J. M. Liu, J. M. Chen, G. P. Zhang, Y. Zhao, "Rapid determination of preliminary thalassaemia screening indicators based on near-infrared spectroscopy with wavelength selection stability," *Anal. Methods* **5**, 4355–4362 (2013).
19. T. Pan, M. M. Li, J. M. Chen, H. Y. Xue, "Quantification of glycated hemoglobin indicator HbA1c through near-infrared spectroscopy," *J. Innov. Opt. Health Sci.* **7**, 1350060 (2014).
20. J. G. Rosas, H. D. Waard, T. D. Beer, C. Vervaet, J. P. Remon, W. L. J. Hinrichs, H. W. Frijlink, M. Blanco, "NIR spectroscopy for the in-line monitoring of a multicomponent formulation during the entire freeze-drying process," *J. Pharm. Biomed. Anal.* **97**, 39–46 (2014).
21. T. Pan, A. Hashimoto, M. Kanou, K. Nakanishi, T. Kameoka, "Development of a quantification system of ionic dissociative metabolites using an FT-IR/ATR method," *Bioprocess Biosyst. Eng.* **26**, 133–139 (2003).
22. T. Pan, A. Hashimoto, M. Kanou, K. Nakanishi, T. Kameoka, "Mid-infrared spectroscopic quantification of ionic dissociative metabolites based on three spectral extraction methods," *Jpn. J. Food Eng.* **5**, 22–31 (2004).
23. X. L. Long, G. S. Liu, T. Pan, J. M. Chen, "Waveband selection of reagent-free determination

- for thalassemia screening indicators using Fourier transform infrared spectroscopy with attenuated total reflection,” *J. Biomed. Opt.* **19**, 087004 (2014).
24. K. Zou, Q. Zhang, Y. Xia, “Culture medium optimization for teicoplanin production by response surface analysis,” *Chin. J. Antibiot.* **36**, 671–675 (2011).
25. S. R. Cheong, S. Y. Kim, Y. W. Lee, J. K. Lee, H. K. Jung, B. S. Koo, Microorganism and the process for the production of teicoplanin: US, 7432080B2[P]. 2008–10–07.
26. W. Z. Lu, *Modern Near-Infrared Spectroscopy Analytical Technology* (China petrochemical press, Beijing, 2007).