MULTIVARIATE REGRESSION MODELS BASED ON UV FULL SPECTRA FOR THE SIMULTANEOUS DETERMINATION OF TETRACYCLINE, PENICILLIN G AND CEPHALEXIN IN DIFFERENT DOSAGE FORMS

Đến tòa soạn 21-05-2024

Nguyen Duc Phong[1], Nguyen Manh Son[1], Nguyen Dieu Linh[1], Nguyen Thi Minh Loi[2], Bui Xuan Thanh[1], Pham Gia Bach[1], Le Si Hung[3], Ta Thi Thao[1*]

1. Faculty of Chemistry, VNU University of Science, Hanoi

2. Quang Binh Univerity

3. Thanglong instruments

*Email: tathithao@hus.edu.vn

TÓM TẮT

XÁC ĐỊNH ĐỒNG THỜI TETRACYCLINE, PENICILLIN G VÀ CEPHALEXIN TRONG CÁC DẠNG BÀO CHẾ KHÁC NHAU BẰNG MÔ HÌNH HỒI QUY ĐA BIẾN DỰA TRÊN PHỔ UV TOÀN PHẦN

Nghiên cứu này trình bày phương pháp phân tích nhanh, đơn giản và hiệu quả dựa trên dữ liệu quang phổ UV toàn phần kết hợp với mô hình hồi qui tuyến tính đa biến đã được phát triển và xác nhận giá trị sử dụng để xác định tetracycline (TET) penicillin G (PGP) và cephalexin (CEX) trong các dạng thuốc kháng sinh viên nén, sử dụng nền mẫu giả dược (placebo) chứa các chất phân tích và nền mẫu thực có chứa một chất phân tích và thêm 2 chất còn lại. Khoảng nồng độ tuyến tính của TET, PGP và CEX lần lượt là 12-28 µg/mL, 7-20 µg/mL và 5-18 µg/mL, với giá trị bước sóng cực đại lần lượt là 276 nm, 290 nm và 262 nm. Để xác định đồng thời TET, PGP và CEX, mẫu thuốc được nghiền nhỏ, hòa tan trong nước cất 2 lần, rung siêu âm và đo phổ UV của các chất phân tích trong khoảng từ 230 đến 350 nm với các khoảng $\Delta\lambda$ = 2 nm ở 61 bước sóng. Bộ mẫu dùng để luyện mô hình gồm 31 mẫu chứa ba thành phần (cả mẫu thương mại chứa 1 chất và mẫu thêm chuẩn) và bộ mẫu kiểm tra gồm 9 mẫu đã được sử dụng. Dữ liệu độ hấp thụ quang kết hợp với các thuật toán học máy (bao gồm hồi quy thành phần chính (principal component regression), bình phương tối thiểu từng phần (partial least squares), cây quyết định (decision tree), rừng ngẫu nhiên (random forest) và sự kết hợp của hai trong số các thuật toán này), đã được phát triển và tối ưu hóa để xác định mô hình phù hợp nhất. Các kết quả tốt nhất đạt được bằng cách sử dụng thuật toán PLS, với căn bậc hai của sai số trung bình bình phương (RMSE) từ 0,682 đến 1,132 và hệ số xác định giữa kết quả hàm lượng xác định theo mô hình và kết quả đúng từ giá trị chứng nhận trên bao bì của hãng cũng như kết quả phân tích bằng phương pháp HPLC đạt từ 0,75 đến 0,88. Mô hình PLS tối ưu đã được áp dụng thành công để phân tích đồng thời TET, PGP và CEX trong mẫu dược phẩm cho kết quả phù hợp (độ thu hồi từ 81,0% đến 110,9% và độ chụm khi phân tích lặp lại đạt RSD < 2%) với kết quả xác định theo phương pháp HPLC qui định bởi Dược điển Việt Nam.

Keywords: *tetracycline, penicillin g, cephalexin, hồi qui đa biến, UV-VIS.*

1. INTRODUCTION

Tetracyclines hydrochloride, penicillin G procaine, and cephalexin monohydrate (Figure 1) are among the three major antibiotics groups used for veterinary purposes, human therapy, and agricultural purposes [1]. Tetracyclines hydrochloride (TET), chemically

(4S,4aS,5aS,6S,12aR)-4-(dimethylamino)-1,6,10,11,12a-pentahydroxy-6-methyl-3,12-dioxo-4,4a,5,5a-tetrahydrotetracene-2-carboxamide hydrochloride, is an antibiotic which belongs to the group of tetracyclines. It is a broad-spectrum antibiotic produced semisynthetically from chlortetracycline, an antibiotic isolated from the bacterium *Streptomyces aureofaciens*. It is used to treat urinary tract infections, acne, gonorrhea, and other conditions [2]. Penicillin G Procaine (PGP**)**, chemically (2S,5R,6R)-3,3-dimethyl-7-oxo-6-(2-phenylacetyl)amino-4-thia-1-azabicyclo(3.2.0)heptane-2-carboxylic acid, belongs to the group of β-lactam antibiotics [3]. Penicillins constitute about 50% of the antimicrobial agents currently in use and they are the first choice drugs in the treatment of nosocomial infections[1,3]. Cephalexin monohydrate (CEX), chemically (7R)-7-(D-a-Amino-a-phenylacetamido)-3-methyl-3-cephem-4-carboxylic acid hydrate or (6R,7R)-7-{((2R)-2-amino-2-phenylacetyl)amino}-3-methyl-8-oxo-5-thia-1-azabicycio(4.2.0)oct-2-ene-2-carboxylic acid hydrate, is a first-generation cephalosporin antibiotic[4]. It is used in the treatment of susceptible infections of the respiratory tract, urinary tract, and skin.

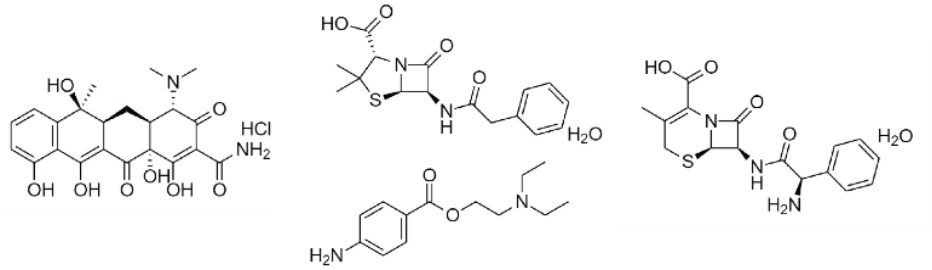

Figure 1. Chemical structures of three antibiotics, including tetracycline hydrochloride, penicillin G procaine, and cephalexin

TET, PGP, and CEX and most other active pharmaceutical compounds dosage form are officially listed in the Vietnamese Pharmacopoeia as high-performance liquid chromatography (HPLC)[5]. Howerver, in order to meet the need for rapid analysis in production processing with a single measurement, the full- UV spectrum method proves to be quite effective. To overcome the difficulties related to matrix effects, chemometrics-assisted methods were also developed for the simultaneously spectrophotometric analysis of several active compounds in pharmaceutical dosage forms without any pretreatment such as determination of ciprofloxacin and doxycycline hyclate [6], simultaneous determination of two antibiotics in tablets [7] or certain β-lactam antibiotics combinations [8].

The application of chemometrics, particularly multivariate calibration methods, is recently an important role in the multicomponent analysis of mixture [9]. Multivariate calibration methods, such as principal component regression (PCR) and partial least squares (PLS), have been used to analyze spectra data without the use of separation technique [10,11], which provide a cheap, fast and simple method for detecting active compound in pharmaceutical mixtures. In this study, one analytical procedure was established for simultaneous determination based on UV-Vis spectrometry coupled with several multivariate linear regress solved by Python – an open language programme to quantify three antibiotics, including TET, PGP, and CEX, in laboratory-prepared synthetic and pharmaceutical mixtures. This study also provided a universally analytical procedure for determining the three antibiotics in their pharmaceutical forms without the need for optimization for each analyte.

2. EXPERIMENTALS

2.1. Apparatus and Software

Spectrophotometric measurements were performed on a UV-1601PC (Shimadzu) connected to a computer loaded with UV-Win PC software. All absorption spectra were saved and subsequently exported UV-Win software to Microsoft Excel program for statistical manipulation. Python 3.9.7 was utilized on the Windows 11 system equipped with a 2.4 GHz Intel Core i5–1135G7 processor was employed to establish a multivariate linear model for concentration determination of each antibiotic.

2.2. Reagents and Samples

The primary standard of tetracycline hydrochloride (TET), penicillin G (PGP), and

cephalexin m(CEX) (purity > 99.9%) were obtained from the National Institute of Drug Quality Control (Hanoi, Vietnam). Double distilled water was used throughout the study. The commercial samples containing of one antibiotic with and without adding two other components were used for calibration, validation set. The samples were purchased from pharmacies in Hanoi, Vietnam.

2.3. Preparation of stock and working Standard Solutions

Standard stock solutions of tetracycline monohydrate, penicillin G procaine, and cephalexin monohydrate were prepared by dissolving the appropriate amounts of each analytical reagent in pure water to get a concentration of 200 µg/mL. The solutions were stored and protected from light at 4°C. Working standard solutions were prepared daily by appropriate dilution in HCl medium.

Suitable aliquots of the stock standard solutions of TET, PGP, and CEX were diluted with distilled water to obtain concentration. The mixture of the three components was also prepared in a concentration of 25 µg/mL. These solutions were then scanned in the range of 230 nm – 350 nm.

2.4. Construction of Training Set and Test Set

The linear concentration ranges of TET, PGP and CEX in UV spectrophotometry were 12-28 µg/mL, 5-18 µg/mL and 7-20 µg/mL, respectively. Absorbance maximum values were recorded at λ max of each drug (276nm for TET, 290nm for PGP, and 262nm for CEX) against distilled water as a calibration blank.

The training and validation mixtures were prepared by combining sets of working standard solutions and commercial pharmaceutical samples certified by manufactures in the linear range of concentrations. Each solution contained three of TET, PGP, and CEX (available and spiked) in different ratios in their concentration linearity ranges. Five concentration levels of each analyte were chosen to construct both training and validation sets. A total set of 31 mixtures and 12 mixtures were independently prepared for training and validation sets, respectively.

The absorption spectra of all mixtures were recorded over the range 230-350nm with a 2nm interval.

2.5. Analysis of the pharmaceutical formulations

Five tablets were accurately weighed and finely powdered. Tablet powder equivalent to TET (250 mg), PGP (150 mg), and CEX (120 mg) was accurately weighed and transferred into a 100 mL volumetric flask, and 50 mL of distilled water was added. The solution was well shaken and ultrasonicated for 15 min. Then, the solution was filtrated in a 100 mL volumetric flask through a filter paper. The residue was washed three times with 10 mL water, and the volume was adjusted volume with water to 100 mL. The stock solutions then were diluted with the solvent to obtain the appropriate working sample solution for UV measurements at the specified range.

2.6. Accuracy Study

The accuracy of the method was evaluated as the percent recovery by the standard addition method at three levels: 80, 100, and 120% of the known concentration of the analyte in the sample. Known amounts of the standard solutions of TET, PGP, and CEX were spiked into the sample solution, and the resulting solutions were scanned in the range of 230 – 350 nm. The accuracy of the method was assessed based on the percent recovery of the added amounts of the standard to the previously analyzed samples. The developed method was validated according to the International Conference on Harmonization (ICH) guidelines.

3. RESULTS AND DISCUSSION

3.1. Multivariate Calibration Analysis

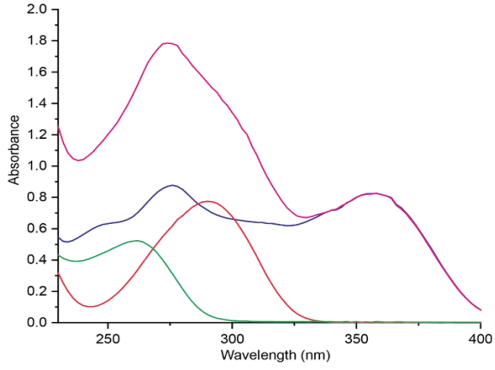

Figure 2. The representative UV absorption spectra of TET (blue line), PGP (red line), CEX (green line), and their mixture (magenta). The spectra were recorded from 230 to 400 nm.

The UV absorption spectra of TET, PGP, CEX and their mixture in standard solution are given in Figure 2. The absorbance spectra which showed a significant overlap were recorded between 230 and 350 nm. TET showed high absorbance from 350 nm to 380 nm while both PGP and CEX did not

absorb light at these wavelengths. Thus, to avoid the bias of constructed multivariate model, only UV absorbance data between 230 and 350 nm was selected to build the multivariate linear regression.

Laboratory synthetic mixtures in training and validation sets (Table 1) (including principle component regression (PCR), partial least squares (PLS), decision tree (DT), random forest (RF) and combined two of them) were used in the analysis to prove the suitability of the calibration model for the determination of TET, PGP, and CEX in pharmaceutical samples. To select the optimum number of principal components for PCR and PLS models, a cross-validation method was used for the training set. The predictive abilities of the models were evaluated by the root-mean-square error of cross-validation (RMSECV), root-mean-square error of prediction (RMSEP), and coefficient of determination (R^2) (Table 1). In the cross-validation method, the same set of mixtures used for both model training and testing. The model was then validated by prediction of concentration of analytes in a sample set which was not used for the model development. In contrast, RMSEP calculated from the validation set was the estimated prediction calibration error that accurately reflected all sources of variability in the calibration method. In general, the values of RMSECV and RMSEP─had to be as low as possible while R^2 value should be as close as 1 for an accurate model. Coefficient of determination (R^2) values of validation set, obtained for each antibiotic in mixtures by PLS models, were from 0.756 to 0.887, which shows good predictive abilities of the models whereas other models exhibited overfitting in the results (R^2=1 and RMSE=0) due to the limited number of samples.

Table 1. Results of the training set and validation set

Model	Datatype	TET		PGP		CEX	
		R^2	RMSE	R^2	RMSE	R^2	RMSE
PCR	**Validation**	*0.843*	*1.332*	*0.822*	*0.795*	*0.843*	*0.776*
PLS	**Validation**	***0.887***	***1.132***	***0.869***	***0.682***	***0.756***	***0.966***
Random-Forest (RF)	**Validation**	*0.921*	*0.946*	*0.25*	*1.634*	*0.473*	*1.422*
Decision Tree (DT)	**Validation**	*0.921*	*0.946*	*0.25*	*1.633*	*0.473*	*1.422*
PCA-RF	**Validation**	*0.659*	*1.968*	*0.105*	*1.784*	*0.791*	*0.894*
PCA-DT	**Validation**	*0.331*	*2.756*	*-1.359*	*2.898*	*0.271*	*1.673*
PLS-RF	**Validation**	*0.518*	*2.339*	*0.329*	*1.545*	*0.725*	*1.027*
PLS-DT	**Validation**	*-0.267*	*3.794*	*-0.797*	*2.529*	*0.27*	*1.673*

3.2. Accuracy

The validation of the optimized model was examined by the standard addition technique at 80%, 100%, and 120% of the test concentration. The percent recoveries range from 91.6% to 132.7% (Table 2) with n = 3. Both TET and PGP do not had a good recovery, which might be caused by the interference of excipients in pharmaceutical products. These excipients could absorb the wavelength from 230 nm to 350 nm; thus, it was necessary to include a sample treatment step to remove the interferences for accurate TET and PGP detection. Otherwise, the percent recovery of CEX, from 91.6% to 110.6%, indicated that there was minimal interference of excipients included in pharmaceutical products to the CEX detection.

Table 2. Accuracy data of TET, PGP, and CEX by PLS model

Antibiotic	TET			PGP			CEX		
Level (%)	80	100	120	80	100	120	80	100	120
Amount (available and taken) (µg/mL)	13.6	17	20.4	7.2	9	10.8	4.8	6	7.2
Predicted amount (µg/mL)	15.8	19.9	26.7	8.8	10.8	14.3	4.8	5.5	8
% Recovery	116	117	131	122	121	133	99.6	91.6	112
%RSD	2.7	3.2	3.8	1.5	0.9	2.2	2.1	4.7	7.4

The intraday precision of the method was examined by repeating the assay of four replicate dilutions of the same concentration. The results that the percent relative standard deviations were all below 2% and the recovery ranged from 97-101% and showed good accuracy of the method.

3.3 Analysis of Tablet Formulation

The developed method was applied to determine TET, PGP, and CEX in laboratory-prepared mixtures of their pharmaceutical doses and comercial products. The assay result of CEX detection showed a good agreement with the concentration taken for the formulation. On the other hand, the results of TET and PGP detection showed around a 10% difference with the label concentration. This revealed that the sample matrices and/or excipients significantly interfered with the quantification of TET and PGP. As the discussion above, it was necessary to include an extra sample treatment step for accurate detection.

The developed analytical procedure was much easier than the HPLC method listed in pharmacopoeias[5,6] for determining three antibiotics. This method used a cheap solvent, distilled water, and a simple instrumental, UV-Vis spectrometer. It showed that all three antibiotics could successfully be quantified, especially CEX. Because of the interference of excipients, it is required a sample treatment step for the accurate detection of TET and PGP.

Table 3. Assay results of TET, PGP, and CEX in laboratory-prepared commercial mixtures by developed PLS methods (n = 3).

Sample	Analyte	Amount recorded in drug label (mg/tablet)	Founded (mg /tablet± SD)	Error (%)
Individual	TET	250	206.9 ± 3.1	-17.2
	PGP	150	121.4 ± 1.4	-19.0
	CEX	120	120.0 ± 1.4	0.0
Synthetic mixture	TET	250	238.9 ± 3.4	-4.4
	PGP	150	166.4 ± 1.3	10.9
	CEX	120	117.1 ± 1.3	-2.4

4. CONCLUSION

This environmentally friendly method eliminatesd the use of organic solvents in both sample preparation and analysis. Utilizing water as a solvent and a UV-Vis spectrometer, it offered significant advantages in resource-limited areas. When combined with machine learning models, it enabled the simultaneous determination of tetracycline, procaine benzylpenicillin, and cephalexin with minimal sample pretreatment, ensuring rapid, accurate, and economical analysis. The PLS algorithm with the lowest error value was selected to evaluate the recovery. The recovery results showed that there was no influence of the presence of excipients in the pharmaceutical formulation. This method can thus be used to replace other complicated and costly methods in the case of limited resources.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

References

[1] Daghrir, R.; Drogui, P, (2013). Tetracycline Antibiotics in the Environment: A Review. *Environmental chemistry letters*, **11(3)**, 209–227.

[2] Vardanyan, R. S.; Hruby, V. J, (2006). Antiviral Drugs. In Synthesis of Essential Drugs. *Elsevier*, 549–557.

[3] Dumancas, G. G.; Hikkaduwa Koralege, R. S.; Mojica, E.-R. E.; Murdianti, B. S.; Pham, P. J., (2014). Penicillins. In Encyclopedia of Toxicology. *Elsevier*, 768–772.

[4] Castle, S. S, (2007). In xPharm: The Comprehensive Pharmacology Reference. *Elsevier*, 1–5.

[5] Vietnamese Pharmacopoeia, (1997). Medical Publishing House, Hanoi, Vietnam, 89.

[6] Eticha, T.; Kahsay, G.; Asefa, F.; Hailu, T.; Gebretsadik, H.; Gebretsadikan, T.; Thangabalan, B, (2018). Chemometric-Assisted Spectrophotometric Method for the Simultaneous Determination of Ciprofloxacin and Doxycycline Hyclate in Pharmaceutical Formulations. *Journal of analytical methods in chemistry*, **2018(1)**, 9538435.

[7] Ribone, M.E., Pagani, A.P., Goicoechea, H.C. et al, (2000). Simultaneous Determination of Two Antibiotics in Tablets by Spectrophotometry and Principal Component Regression (PCR) Analysis. An Advanced Undergraduate Experiment Involving Chemometrics. *The Chemical Educator*, **5**, 236–241.

[8] I. Mohamed, Abd El-Maboud; Salem, Hesham; and Maher, Eman, (2007). Chemometrics-assisted spectrophotometric determination of certain β-lactam antibiotics combinations. *The Thai Journal of Pharmaceutical Sciences*, **31(1)**, Article 2.

[9] Z.T. Chowhan, Li-Hua Chi, (1981). Mixing of Pharmaceutical Solids III: Multivariate Statistical Analysis of Multicomponent Mixing. *Journal of Pharmaceutical Sciences*, **70(3),** 247-251.

[10] Ipp, M, (1996). Comparison of PLS, PCR and MLR for the quantitative determination of foreign oils and fats in butter fats of several European countries by their triglyceride composition. *Z Lebensm Unters Forch*, **202**, 193–198.

[11] Hemmateenejad B, Akhond M, Samari F, (2007). A comparative study between PCR and PLS in simultaneous spectrophotometric determination of diphenylamine, aniline, and phenol: Effect of wavelength selection. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, **67(3-4)**, 958-65.