Original Article

Mahidol University Journal of Pharmaceutical Sciences 2014; 41 (4), 18-28

Development of Chemometric Tool for Determination of Peptide Samples Comparing with High-Performance Liquid Chromatography

C. M. Phechkrajang, S. Jarusintanakorn and O. Vajragupta*
Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Mahidol University, 447 Sri-Ayuthaya Rd., Payathai, Rachatevi, Bangkok, 10400. Thailand

Abstract

In this study, a chemometric tool, partial least square regression (PLS-1), was developed for quantitative determination of synthetic peptide samples, L-alanyl-L-tryptophan (AT) and glycyl-L-phenylalanine in the presence of dansyl-L-phenylalanine. Two sets of mixture solutions, calibration and test sets, were prepared and measured for their UV absorption. The absorbance data of calibration set and the Unscrambler® program were employed for PLS-1 models set up. The obtained models were validated by determination the test set solutions which were not contributed in model construction. On the other hand, a high- performance liquid chromatography (HPLC) method was developed and validated. In comparison, the developed HPLC method was used to determine the same test set as PLS-1 models. By using *t-test* statistical evaluation, the results obtained from two methods, PLS-1 and HPLC, were found comparable and statistical comparison showed no significant difference at 95% confident interval, indicated replacement ability of PLS-1 to a separation method, HPLC for the determination of peptide samples.

Keyword: Chemometrics, PLS-1, Peptides

INTRODUCTION

Nowadays, peptides and peptidomimics played an important role in new drug discovery and drug development due to their several advantages, comparing to other small molecules, including higher affinity, specific to target side, low toxicity and better tissue penetration.[1-7] Quantitative analysis is a key step in drug development. For peptide drugs, their structures sometimes lack of UV chromophore, therefore, quantitative determinations of peptides are usually performed by LC/MS method.[8-13] Although, LC/MS is a highly effective instrument, the operation cost is still a limitation. In the past few years, quantitative determinations of a mixture without chromatographic separation were mainly performed by chemometric tools such as multivariate calibrations.

Chemometrics was introduced in 1972 by Wold. By definition, chemometrics is the chemical discipline that uses mathematical and statistical methods, to design or select optimal measurement procedures and experiments. Additionally, they are used to provide maximum chemical information by analyzing chemical data. Nowadays, the important of chemometric method to chemistry field is increased due to the accessibility of modern analytical instruments, which can generate multivariate responses and overflow data for every analyzed sample. Moreover, traditional univariate statistics is not adequate to resolve multivariate data. Consequently, multivariate data analysis methods based on chemometrics are concentrated by analytical chemists to extract meaningful information from noise containing flood data. [14-16]

***Corresponding author:** Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Mahidol University, 447 Sri-Ayuthaya Rd., Payathai, Rachatevi, Bangkok, 10400. Thailand
Tel:+66-2-644-8695; Fax: +66-2-644-8695; E-mail: chutima.mat@mahidol.ac.th

UV-VIS spectrophotometry is widely used analytical technique with high selectivity, good accuracy, robustness, quick response, simplicity to operate and cost effectiveness, not only for purchasing but also to maintaining the instrument. Therefore, data are easily accessible and convenient from spectrophotometric measurements. However, mixtures of active ingredients and excipients in UV region produce strongly overlapping spectral bands which are not solved by conventional spectrophotometric measurements. These drawbacks prevent the usefulness of such method. Fortunately, the problem can be overcome with the aid of chemometrics.[14-16]

In chemometrics-assisted spectrophotometric measurement, calibration set must be used for calibration, while, determination or prediction can be performed via test set. Calibration set is the objects used to create models by means of multivariate calibration method and test set is objects of unknown samples subjected to predict of its contents. There are at least three types of multivariate calibration methods, multiple linear regression or MLR, principle component regression or PCR and partial least square regression (PLS or PLSR). There are also two versions of PLSR, PLS-1 and PLS-2. However, PLS-1 is more satisfied than PLS-2 since the multivariate calibration model of each interest substance is performed separately from each others for PLS-1 algorithm. This leads to more accurately prediction model compares with PLS-2. [14-16]

In this study, chemometric-assisted spectrophotometric method, PLS-1, was developed for determination of peptide samples. The study was performed to prove whether assay method based on chemometrics can be applied to peptidomimic drug development process. Commercial available small peptides containing structural UV chromophors were selected and used as modeled sample in the present study. To serve the aim of the study, a mixture of L-alanyl-L-tryptophan (AT) and glycyl-L-phenylalanine in the matrix of dansyl-L-phenylalanine was prepared. Dansyl-L-phenylalanine was also added in order to study effect of precursor impurity on the assay method. Chemical structures of studied peptides were illustrated in Figure 1. An HPLC method was also developed and validated in comparison with chemometric methods.

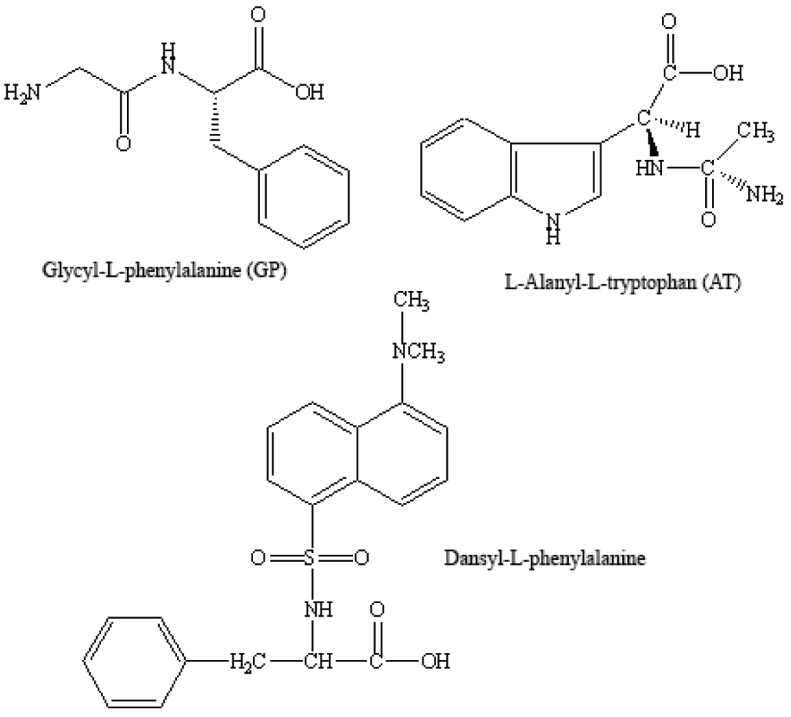

Figure 1. Chemical structures of studied peptides.

MATERIALS AND METHODS

Materials

Standard peptides, L-alanyl-L-tryptophan (AT), glycyl-L-phenylalanine (GP) and dansyl-L-phenylalanine were purchased from Tokyo Chemical Industry Co.Ltd., Tokyo, Japan. Acetonitrile (HPLC grade) was purchased from Lab-Scan, Bangkok, Thailand. Trifluoroacetic acid (Analytical grade) was obtained from Sigma-Aldrich, Missouri, USA. The peptides sample used in this study was the synthetic mixture of L-alanyl-L-tryptophan (AT) and glycyl-L-phenylalanine (GP) in the matrix of dansyl-L-phenylalanine.

Apparatus and software

The absorbance spectra were recorded by a Shimadzu UV-Vis spectrophotometer (UV-160A, Shimadzu Corporation, Kyoto, Japan) using a 1 cm quartz cell. Chromatography was performed on a high-performance liquid chromatography system (Shimadzu Corporation, Kyoto, Japan) consisting of a degasser DGU-12A, liquid chromatograph LC-10 AD, communications bus module CBM-10A, a UV-Visible detector SPD-10A and a data processing system (class LC-10). The analytical column was a Symmetry C18, 150 × 3.9 mm i.d., 5µm (Waters, Ireland). Manual injection was made by using a Rheodyne model 7725 injector with a 20-µL loop. Data analysis, PLS-1 calibrations were performed by the Unscrambler® program, purchased from CAMO Software (Oslo, Norway).

HPLC experiments

Chromatographic system

A mixture of 0.1% trifluoroacetic acid and acetonitrile (14:86, v/v) was used as a mobile phase for elution of AT and GP. The presences of two peptides were detected by UV detector at 210 nm. Flow rate of the mobile phase was at 1.2 mL/min.

Development and validation of HPLC method

Standard mixture solution containing 10.0 µg/mL of AT, GP and 1.5 µg/mL dansyl-L-phenylalanine was employed for HPLC method development. Type and concentration of organic solvents in mobile phase and pH of mobile phase were evaluated to obtain a suitable separation condition. Performance characteristics selected for method validation were linearity, accuracy, repeatability and specificity.

Linearity

Linearity was evaluated in the concentration range of 2.5-20.0 µg/mL for AT and GP. Standard mixtures of AT (2.5-20.0 µg/mL) and GP (2.5-20.0 µg/mL) were prepared and dansyl-L-phenylalanine was added to the final concentration of 1.5 µg/mL. Each standard mixture solution was injected into chromatographic system described above. Calibration curve of each peptide was separately plotted between concentrations (x-axis) versus corresponding peak areas (y-axis). The data were analyzed by least-squares linear regression method.

Accuracy

Accuracy of the developed HPLC procedure was studied by standard addition method. The sample containing 5.0-15.0 µg/mL of AT and GP was added to matrix of dansyl-L-phenylalanine (1.5 µg/mL). Three replicates were employed for each concentration. The accuracy of the method was expressed in terms of percent recovery between amount of standard found added and amount of standard added.

Repeatability

A mixture at the concentration of 10.0 µg/mL of each peptide in the matrix of dansyl-L-phenylalanine (1.5 µg/mL) was utilized for repeatability study and six replicates were employed. The repeatability of the method was expressed as the percentage of relative standard deviation (%RSD).

Specificity

To demonstration the specificity of the developed HPLC method, a mixture solu-

tion of peptides at concentration of 10.0 µg/mL in the presence of dansyl-L-phenylalanine (1.5 µg/mL) was determined and compared the result with a mixture solution of peptides at the concentration of 10 µg/mL without matrix and a matrix solution at the concentration of 1.5 µg/mL.

Chemometric experiments

Linear dynamic concentration ranges

To find the linear dynamic concentration range of each compound, one component calibration was performed. Linear dynamic ranges were studied in the concentration range of 3.0-15.0 µg/mL for AT and GP. Absorbance values were recorded at λ_{max} of each compound (220 nm for AT and 210 nm for GP) in 1-cm. quartz cell and used DI water as blank. Linear dynamic range for each compound was determined by least-square linear regression of concentration and the corresponding absorbance values.

Preparing of calibration set and test set

Two sets of standard solutions, calibration set and test set were prepared. According to Table 1, 16 and 7 mixtures solutions were used in calibration set and test set, respectively. The concentrations of calibration set were selected by mean of central composite design (CCD) [17] and those of test set were randomly selected. In addition, matrix dansyl-L-phenylalanine was added to calibration and test solutions to concentration of 1.5 µg/mL. All concentrations in calibration set and test set were within the linear dynamic concentration ranges, which were previously investigated.

Table 1. Compositions of calibration set and test set

Calibration set			Test set		
Mixture no.	AT (µg/mL)	GP (µg/mL)	Mixture no.	AT (µg/mL)	GP (µg/mL)
1	3.00	9.00	1	5.00	10.00
2	15.00	9.00	2	6.00	8.00
3	9.00	3.00	3	10.00	5.00
4	9.00	15.00	4	7.00	7.00
5	4.76	4.76	5	10.00	12.00
6	13.24	4.76	6	12.00	7.00
7	13.24	13.24	7	6.00	9.00
8	4.76	13.24			
9	9.00	9.00			
10	9.00	9.00			
11	9.00	9.00			
12	9.00	9.00			
13	9.00	0.00			
14	9.00	0.00			
15	0.00	9.00			
16	0.00	9.00			

PLS-1 models building and models interpretation

The solution in calibration set and test set were measured for absorbance data in the wavelength interval of 200-400 nm. The absorbance data of calibration set were then subjected to the Unscrambler® program for PLS-1 models building. The optimum models were selected upon parameters, including root mean square error of calibration (RMSEC), root mean square error of prediction (RMSEP), slope and correlation coefficient (r^2) of measured and predicted concentrations in calibration set and test set. RMSEC and RMSEP were calculated by the following equations.

$$RMSEC = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_{i,cal} - y_{i,cal})^2}{n}}$$

Where; RMSEC = Root mean square error of calibration

$y_{i,cal}$ = Measured concentration (i) in calibration set

$\hat{y}_{i,cal}$ = Predicted concentration (i) in calibration set

n = Number of samples in calibration set

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_{i,val} - y_{i,val})^2}{n}}$$

Where; RMSEP = Root mean square error of prediction

$y_{i,val}$ = Measured concentration (i) in validation set (test set)

$\hat{y}_{i,val}$ = Predicted concentration (i) in validation set (test set)

n = Number of samples in validation set (test set)

For validation of the resulted models, the models were used to assay test set samples which were not contributed in models building.

Comparison of PLS-1 and HPLC method

The developed HPLC method and PLS-1 models were used to determine test set samples. The results obtained from the two methods were compared using *t-test* statistic.

RESULTS AND DISCUSSION

Development of HPLC method

To compare the results obtained from PLS-1 models, a HPLC method for simultaneous determination of AT and GP was developed. The optimum condition for separation of AT and GP on C18 column was achieved from the mobile phase mixture of acetonitrile and 0.1% trichloroacetic acid (14:86, v/v). The flow rate of the mobile phase was 1.2 mL/min and the presences of two peptides were monitored by a UV detector at 210 nm. As illustrated in Figure 2, the retention times of GP and AT from the optimum HPLC condition were 3.3 and 5.7 min, respectively.

Validation of the developed HPLC method

The developed HPLC method was validated for its linearity, accuracy, repeatability and specificity. Linearity of the method, investigated by least square regression method, was achieved in the concentration range of 2.5 -20.0 µg/mL for both peptides with the square of correlation coefficients (r^2) greater than 0.999. Accuracy was assessed from standard addition method at three concentration levels, 5.0-15.0 µg/mL and three replicates for each concentration. Accuracy of the method, displayed as recovery percents of amount standard added and amount found, were between 100.1-100.8 %. Repeatability of the method was performed at 100% concentration level and six replicates were employed. Repeatability,

expressed as %RSD of less than 2.0% indicated precision of the developed HPLC method. Summary of validation results for the developed HPLC method were illustrated in Table 2. Specificity of the method was also evaluated. By comparison the chromatograms of peptides in DI water, peptides sample in matrix of dansyl-L-phenylalanine and dansyl-L-phenylalanine solution, it was obvious that the two peptides were well separated from each other and was not interfered by the matrix (Figure 3).

As shown in Figure 4, UV spectrum of AT and GP completely overlapped in the UV region, therefore univariate calibration could not be applied to assay this mixture. Multivariate calibration, PLS-1, was challenged for this task.

Table 2. Validation results of HPLC method

Parameters (Unit)	L-Alanyl-L-tryptophan	Glycyl-L-phenylalanine
Linearity range (μg/mL)	2.5-20.0	2.5-20.0
r^2	0.9999	0.9995
Slope	99254	46244
Intercept	3521	2606
Repeatability (%RSD, n = 6)	0.29	1.59
Recovery (%)	100.1-100.8	99.7-99.8

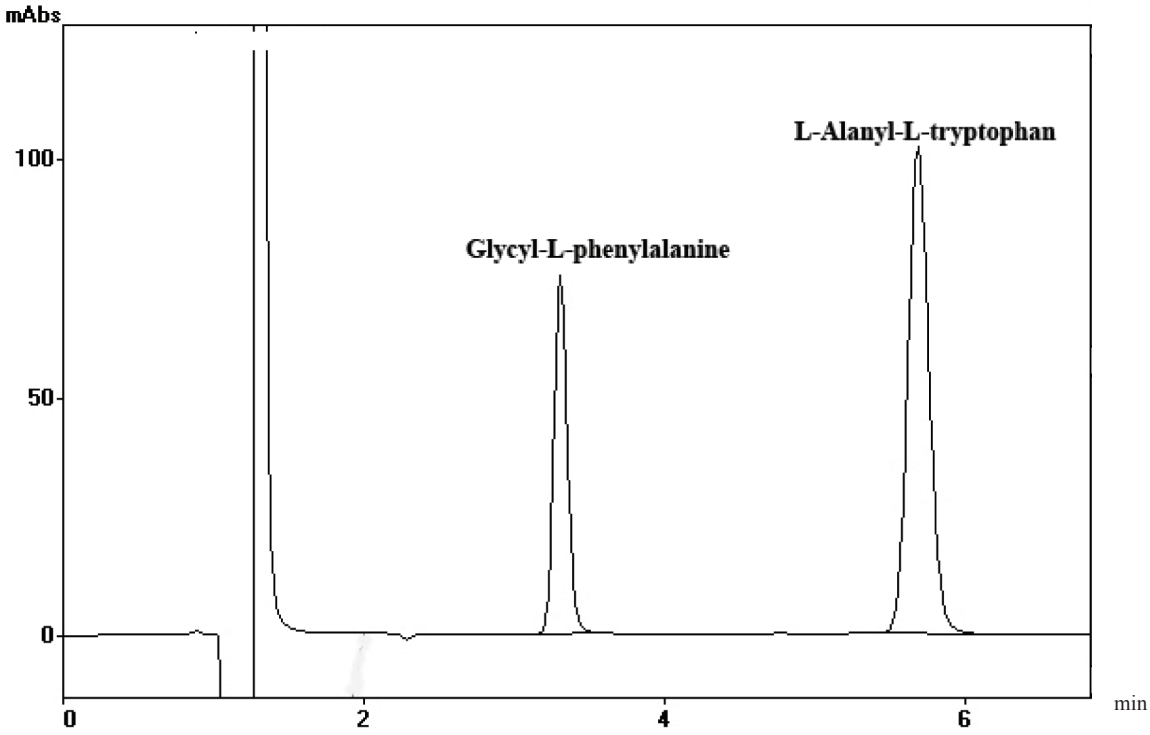

Figure 2. Chromatogram of peptide samples under the optimum HPLC condition. Chromatographic conditions: the column was a Symmetry C18 column (150 × 3.9 mm, i.d., 5 μm); the flow rate was 1.2 mL/min; the mobile phase was trifluoroacetic acid and acetonitrile (14:86, v/v); the injection volume was 20 μL; UV detection was at 210 nm.

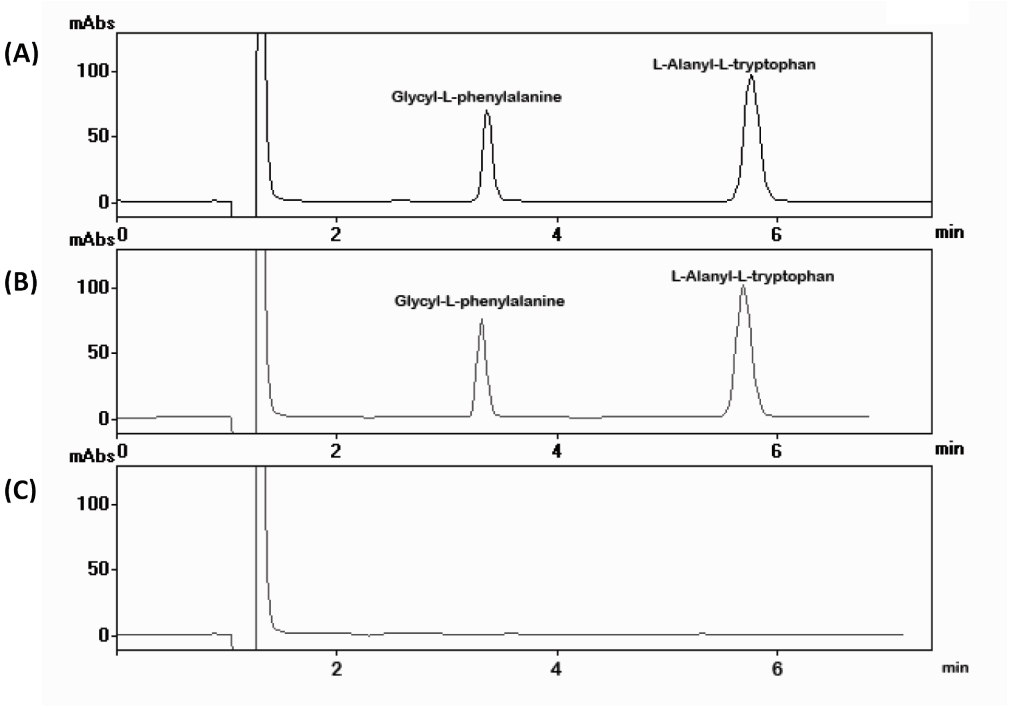

Figure 3. Specificity results of the developed HPLC method. (A) Chromatogram of peptides in water, (B) chromatogram of peptides in dansyl-L-phenylalanine and (C) chromatogram of dansyl-L-phenylalanine. Chromatographic conditions: the column was a Symmetry C18 column (150 × 3.9 mm, i.d., 5 μm); the flow rate was 1.2 mL/min; the mobile phase was trifluoroacetic acid and acetonitrile (14:86, v/v); the injection volume was 20 μL; UV detection was at 210 nm.

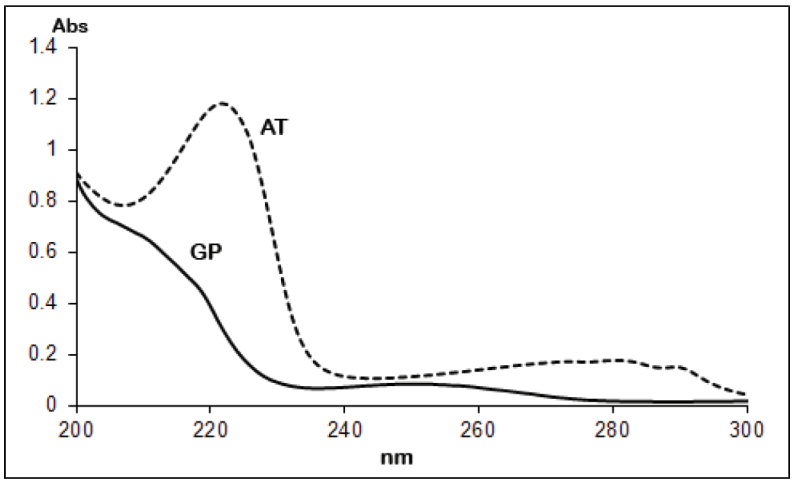

Figure 4. UV spectra of AT and GP.

Linear dynamic concentration ranges for chemometric

To set up PLS-1 models for determination of AT and GP, linear dynamic concentration range of each peptide was firstly evaluated. Linear dynamic concentration range was the concentration range that gave the linear relationship between the concentrations and their corresponding absorbance values. Linear dynamic concen-

trations of the two peptides were investigated in the concentration ranges of 3.0-15.0 µg/mL. Linear response curves were plotted between the concentrations and their corresponding absorbance at the maximum wavelength (λ_{max}). As shown in Figure 5, linear relationships were achieved in the investigated concentration ranges of AT and GP. Therefore, these concentrations interval were further used in calibration step.

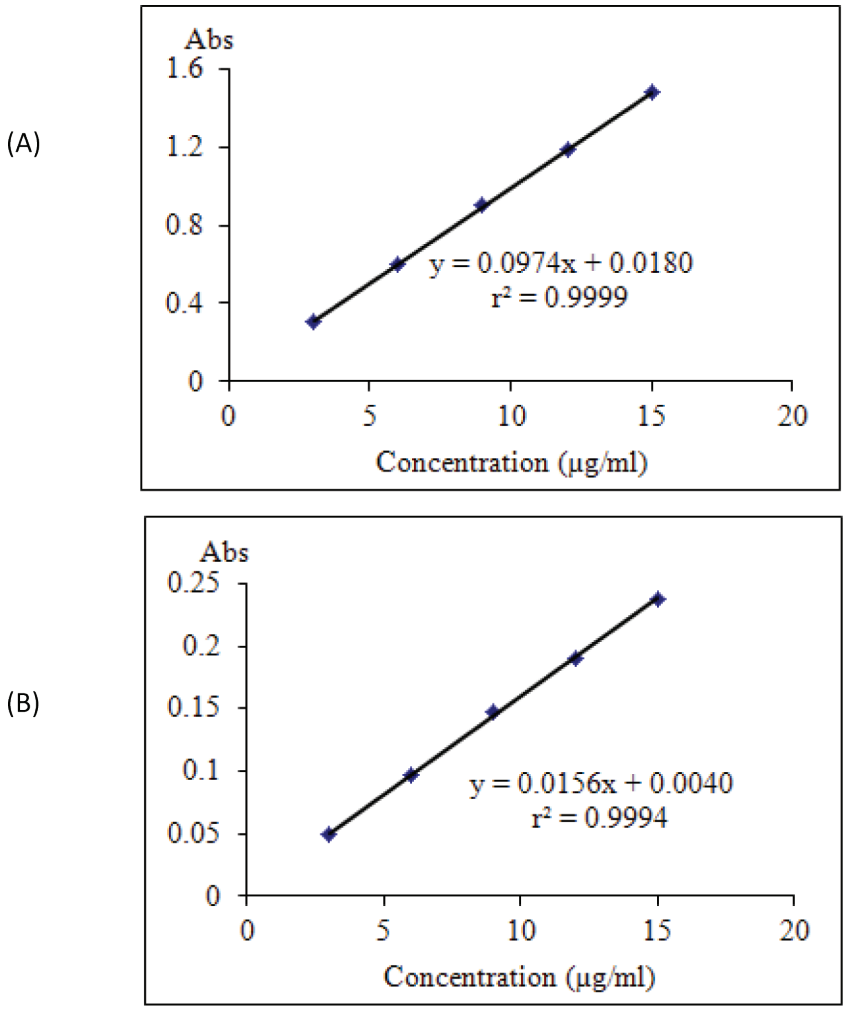

Figure 5. Linear dynamic concentration ranges of AT (A) and GP (B).

PLS-1 models construction

The absorbance data of calibration set were subjected to the Unscrambler® program for PLS-1 models construction. Summary of PLS-1 models parameters were displayed in Table 3. Absorbance data in the wavelength regions of 260-300 nm and 200-270 nm were utilized for PLS-1 models of AT and GP, respectively. The small values of RMSEC, the corresponding measure for the model fit, were obtained for the developed PLS-1 models. Slopes of the plots between true and predicted concentrations in calibration set were closed to 1 indicating that predicted concentrations were not deviated from real concentrations. In addition, there was high correlation between true and predicted concentrations with the square of correlation coefficient (r^2) greater than 0.999 for AT and GP. Prediction ability of the resulted models in the test set, which was not contributed in models construction,

showed the small values of root mean square error of prediction (RMSEP). The plots of true versus predicted concentrations in the test set were displayed in Figure 6. From Figure 6, slopes of two plots were closed to 1 and the square of correlation coefficients (r^2) were higher than 0.999. These results clearly illustrated that new samples were well predicted by the resulted models.

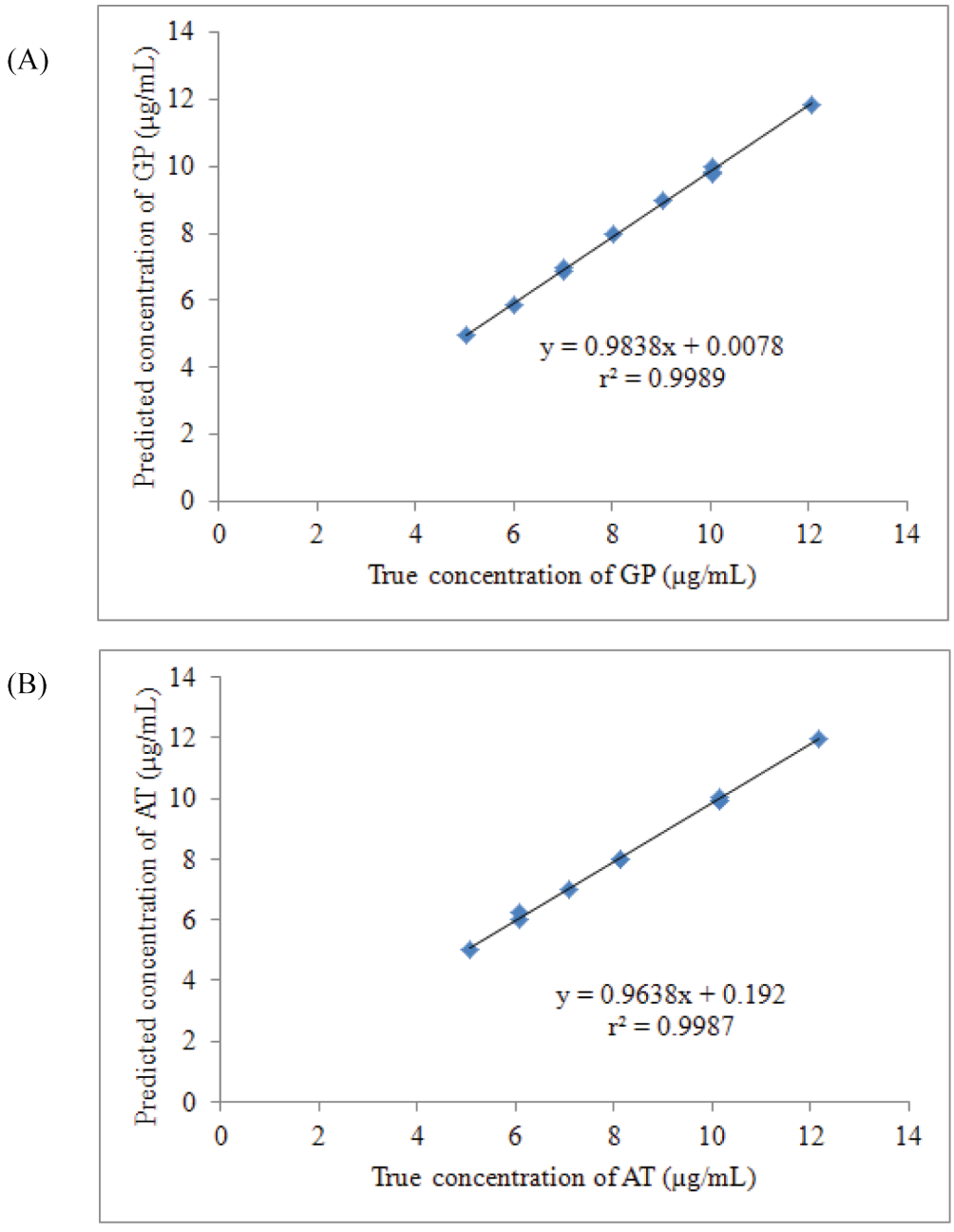

Figure 6. Plots of predicted and true concentrations of test set: (A) glycyl-L-phenylalanine (GP), (B) L-alanyl-L-tryptophan (AT).

Table 3. Summary of statistical parameters of the developed PLS-1 models

Parameters*	L-Alanyl-L-tryptophan	Glycyl-L-phenylalanine
Spectral range (nm)	260-300	200-270
Latent factors	1	2
Slope	0.9990	0.9994
r^2	0.9995	0.9997
RMSEC	0.1205	0.0883
RMSEP	0.1508	0.1052

* r2 is the square of correlation coefficient
RMSEP is root mean square error of prediction.
RMSEC is root mean square error of calibration.

Determination of test set by PLS-1 models and comparison with HPLC method

The resulted PLS-1 models were used to determine the test set samples for model validation. The determination results, expressed in term of average recovery percent between amount added and amount found, were closed to one hundred percent indicating the accuracy of the resulted models (Table 4). The developed HPLC method was used to determine the same test set as PLS-1 models. The results obtained from the two methods were compared by *t-test* statistic. As shown in Table 4, the determination results obtained from HPLC method and PLS-1 models were not significantly different at 95% confidence limit. This implied that the standard method as HPLC may be replaced by simpler and cheaper method such as PLS-1 as shown in this study.

Table 4. Percent recoveries of test set samples obtained from PLS-1 models and HPLC method

L-Alanyl-L-tryptophan			Glycyl-L-phenylalanine		
Added Conc. (μg/mL)	PLS-1	HPLC	Added Conc. (μg/mL)	PLS-1	HPLC
5.00	99.2	100.5	10.00	99.5	98.8
6.00	99.1	100.1	8.00	99.2	99.7
10.00	97.4	99.9	5.00	98.5	98.6
7.00	98.3	99.7	7.00	98.0	99.4
10.00	99.0	100.0	12.00	98.4	100.0
12.00	98.4	99.7	7.00	99.2	100.0
6.00	102.5	99.6	9.00	99.3	99.7
Average	99.1	99.9	Average	98.9	99.5
%RSD	1.63	0.31	%RSD	0.58	0.56

CONCLUSION

Two quantitative determination methods, HPLC and PLS-1, were successfully developed for determination of peptide samples, L-alanyl-L-tryptophan (AT) and glycyl-L-phenylalanine (GP) in the matrix of dansyl-L-phenylalanine. The results obtained from PLS-1 method is comparable with those obtained from HPLC method, indicating to the replacement ability of chemometrics-assisted spectrophotometric method, PLS-1, to the standard method, HPLC.

ACKNOWLEDGEMENT

The authors would like to thank The Innovation in Drug Discovery and Novel Drug Delivery project (NRU-PY 540102) under theNational Research Universities Initiative of the Office of the Higher Education on Commission and Mahidol University for financial support and the Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Mahidol University for research facilities.

REFERENCES

1. Temsamani J, Vidal P. The use of cell-penetrating peptides for drug delivery. *Drug Discov Today* 2004; 9: 1012-1019.
2. Hooft R. Drug discovery and development for metabolic diseases. *Drug Discov Today* 2003**;** 8: 1064-1066.
3. Cheriyan M, Perler FB. Protein splicing: A versatile tool for drug discovery. *Adv Drug Delivery Rev* 2009**;** 61: 899 – 907.
4. Lochmanna D, Jaukb E, Zimmer A. Drug delivery of oligonucleotides by peptides. *Eur J Pharm Biopharm* 2004; 58: 237–251.
5. Jordan AM, Roughley SD. Drug discovery chemistry : a primer for the non-specialist. *Drug Discov Today* 2009; 14: 731-744.
6. Petsalaki E, Russell RB. Peptide-mediated interactions in biological systems: new discoveries and applications. *Curr Opin Biotechnol* 2008; 19: 344–350.
7. Sato AK, Viswanathan M, Kent RB, Wood CR. Therapeutic peptides: technological advances driving peptides into development. *Curr Opin Biotechnol* 2006; 17: 638–642.
8. Koh HL, Yau WP, Ong PS, Hegde A. Current trends in modern pharmaceutical analysis for drug discovery. *Drug Discov Today* 2003; 8: 889-897.
9. Lövgren U, Johansson S, Jensen LS, Ekström C, Carlshaf A. Quantitative determination of peptide drug in human plasma samples at low pg/ml levels using coupled column liquid chromatography–tandem mass spectrometry. *J Pharm Biomed Anal* 2010; 53: 537-545.
10. Bakhtiar R, Majumdar TK. Tracking problems and possible solutions in the quantitative determination of small molecule drugs and metabolites in biological fluids using liquid chromatography – mass spectrometry. *J Pharmacol Toxicol Methods* 2007; 55: 227 – 243.
11. Pereiraa AS, DiLeoneb L, Souzab FH, Lillaa S, Richterb M, Schwartsmannb G, De Nucci G. Quantification of the bombesin/gastrin releasing peptide antagonist RC-3095 by liquid chromatography–tandem mass spectrometry. *J Chromatogr B* 2005; 816: 321-326.
12. Li H, Rose MJ, Tran L, Zhang J, Miranda LP, James CA, Sasu BJ. Development of a method for the sensitive and quantitative determination of hepcidin in human serum using LC-MS/MS. *J Pharmacol Toxicol Methods* 2009; 59: 171– 180.
13. Yang JZ, Bastian KC , Moore RD , Stobaugh JF, Borchardt RT. Quantitative analysis of a model opioid peptide and its cyclic prodrugs in rat plasma using high-performance liquid chromatography with fluorescence and tandem mass spectrometric detection. *J Chromatogr B* 2002; 780: 269-281.
14. Thomas EV. A primer on multivariate calibration. *Anal Chem* 1994; 66: 795A-804A.
15. Geladi P. Chemometrics in spectroscopy. Part 1. Classical chemometrics. *Spectrochim. Acta. A* 2003; 58: 767-782.
16. Lavine B, Workman J. Chemometrics. *Anal Chem* 2006; 78: 4137-4145.
17. Morgan E. (1991) Chemometrics: Experimental design. England: John Wiley & Sons 1991: 238-45.