Journal of Food and Drug Analysis, Vol. 13, No. 4, 2005, Pages 301-311 藥物食品分析 第十三卷 第四期

Determination of Benazepril HCl and Hydrochlorothiazide in Pharmaceutical Preparations Using UV-Visible Spectrophotometry and Genetic Multivariate Calibration Methods

DURMUS ÖZDEMIR[1*] AND ERDAL DINC[2]

[1.] *Department of Chemistry, Faculty of Science, İzmir Institute of Technology, 35430 Urla-İzmir, Turkey*
[2.] *Department of Analytical Chemistry, Faculty of Pharmacy, University of Ankara, 06100 Tandoğan-Ankara, Turkey*

(Received: February 16, 2005; Accepted: June 14, 2005)

ABSTRACT

Simultaneous determination of binary mixtures of benazepril and hydrochlorothiazide in pharmaceutical tablets using UV-visible spectrophotometry, classical least squares (CLS) and three genetic algorithms (GA) based multivariate calibration methods was demonstrated. The three genetic multivariate calibration methods are Genetic Classical Least Squares (GCLS), Genetic Inverse Least Squares (GILS) and Genetic Regression (GR). The sample data set contains the UV- spectra of 28 synthetic mixtures of benazepril (12~36 μg/mL) and hydrochlorothiazide (10~22 μg/mL) and 16 tablets containing both compounds. The spectra cover the range from 210 to 360 nm in 0.1 nm intervals. Several calibration models were built with the four methods. The root mean square error of calibration (RMSEC) and validation (RMSEV) for the synthetic data were in the range of 0.19 and 0.34 μg/mL for all the genetic algorithm based methods. The root mean square error of Prediction (RMSEP) values for the tablets were in the range of 0.04~0.20 mg/tablets. A comparison of genetic algorithm selected wavelengths for each component was also included.

Key words: UV-visible spectrophotometry, multivariate calibration, genetic algorithms, genetic regression

INTRODUCTION

Benazepril hydrochloride has been known as an angiotensin converting enzyme inhibitor that is used in the treatment of essential hypertension. Hydrochlorothiazide has been a widely used thiazide diuretic. The binary mixture of the two drugs is used in the treatment of hypertension. The resolution of the mixture systems containing two or more compounds without any separation procedure in the presence of excipients in samples is one of the main issues of the simultaneous quantitative determination. The simultaneous quantitative determination of both drugs in pharmaceutical tablets using various methods including spectrophotometry[1-5], HPLC[6,7], potentiometric[8], and capillary electrophoresis[9] have been described for several mixtures.

Modern spectroscopic instruments are so fast that they can produce hundreds of spectra in a few minutes for a given sample that contains multiple components. Unfortunately, univariate calibration methods are not suitable for this type of data, as they require an interference-free system. Multivariate calibration deals with data containing instrument responses measured on multiple wavelengths for a sample that usually contains more than one component. In recent years, advances in chemometrics and computers have lead to the development of several multivariate calibration methods[10-13] for the analysis of complex chemical mixtures.

Genetic regression (GR) is a calibration technique that optimizes linear regression models using a genetic algorithm (GA) and has been applied to a number of multi-instrument calibration and wavelength selection problems[14-17]. GAs are non-local search and optimization methods that are based upon the principles of natural selection[18,19]. For a given full spectrum data, GR selects an optimum linear combination of wavelengths and simple mathematical operators to build a linear calibration model using simple least squares method.

Classical Least Squares (CLS) extends the classical Beer's Law model in which the absorbance at each wavelength is directly proportional to the component concentrations. Inverse Least Squares (ILS) is based on the inverse Beer's Law where concentrations of an analyte are modelled as a function of absorbance measurements. Genetic Classical Least Squares (GCLS) and Genetic Inverse Least Squares (GILS) are modified versions of original CLS and ILS methods in which a small set of wavelengths are selected from a full spectral data matrix and evolved to an optimum solution using a genetic algorithm.

In this work, CLS and three different genetic algorithms based calibration methods GCLS, GILS and GR were tested with the aim of establishing calibration models that have a high predictive capacity for the simultaneous determination of benazepril and hydrochlorothiazide in their

* Author for correspondence. Tel: +90-232-498-7534;
Fax: +90-232-498-7509; E-mail: durmusozdemir@iyte.edu.tr

binary mixtures and in pharmaceutical tablet preparations using the UV-visible spectrophotometry.

I. *Genetic Regression*

Genetic Algorithms (GA) are global search and optimization methods based upon the principles of natural evolution and selection as developed by Darwin[20]. Computationally, the implementation of a typical GA is quite simple and consists of five basic steps including initialization of a gene population, evaluation of the population, selection of the parent genes for breeding and mating, crossover and mutation, and replacing parents with their offspring. These steps have taken their names from the biological foundation of the algorithm.

Genetic Regression (GR) is an implementation of a GA for selecting wavelengths and mathematical operators to build linear calibration models. GR is a hybrid calibration between univariate and multivariate calibration techniques in which it optimizes simple linear regression models through an evolving selection of wavelengths and simple mathematical operators (+, –, ×, /). GR follows the same basic initialize/breed/mutate/evaluate algorithm as other GAs but differs in the way it encodes genes. A gene is a potential solution to a given problem and the exact form may vary from application to application. Here, the term gene is used to describe the collection of instrument response pairs combined with the above mentioned operators. These pairs, called "base pairs", are then combined with an addition operator to produce a score, which relates the instrument response to component concentration. The term "population" is used to describe the collection of individual genes in the current generation.

In the initialization step, first generation of genes is created randomly with a fixed population size. Although random initialization helps to minimize bias and maximize the number of possible recombinations, GR is designed to select initial genes in a somewhat biased random fashion in order to start with genes better suited to the problem than those that would be randomly selected. Biasing is done with a correlation coefficient by plotting the scores of initial genes against the component concentrations. The size of the gene pool is a user defined even number in order to allow breeding of each gene in the population. It is important to note that the larger the population size, the longer the computation time. The number of base pairs in a gene is determined randomly between a fixed low limit and high limit. The lower limit was set to 2 to allow single point crossover whereas the higher limit was set to eliminate overfitting problems and reduce the computation time. Once the initial gene population is created, the next step is to evaluate and rank the genes using a fitness function, which is the inverse of the root mean square error of calibration (RMSEC).

The third step is where the basic principle of natural evolution is put to work for GR. This step involves the selection of the parent genes from the current population for breeding using a roulette wheel selection method according to their fitness values. The goal is to give a higher chance to those genes with high fitness so that only the best performing members of the population will survive in the long run and will be able to pass their information to the next generations. Because of the random nature of the roulette wheel selection method, however, genes with low fitness values will also have some chance to be selected. Also, there will be genes that are selected multiple times and some genes will not be selected at all and will be thrown out of the gene pool. After the selection procedure is completed, the selected genes are allowed to mate top-down without ranking whereby the first gene mates with the second gene and the third one with the fourth one and so on as illustrated in the following example:

parents

$$S_1 = (A_{347} \times A_{251})^{\#} + (A_{379} + A_{218}) \quad (1)$$

$$S_2 = (A_{225} \times A_{478})^{\#} + (A_{343} / A_{250}) + (A_{451} - A_{358}) + (A_{231} - A_{458}) \quad (2)$$

The points where the genes are cut for mating are indicated by #.

offspring

$$S_3 = (A_{347} \times A_{251}) + (A_{343} / A_{250}) + (A_{451} - A_{358}) + (A_{231} - A_{458}) \quad (3)$$

$$S_4 = (A_{225} \times A_{478}) + (A_{379} + A_{218}) \quad (4)$$

Here A_{347} corresponds to the raw absorbance at 347 nm wavelength. The first part of the S_1 is combined with the second part of the S_2 to give the S_3, likewise the second part of the S_1 combined with the first part of the S_2 to give S_4. This process is called the single point crossover and is the one used in GR. The single point crossover will not provide different offspring if both parent genes are identical, which may happen in the roulette wheel selection, and broken at the same point. Also note that mating can increase or decrease the number of base pairs in the offspring genes. After crossover, the parent genes are replaced by their offspring and the offspring are evaluated. The ranking process is based on their fitness values following the evaluation step. Then the selection for breeding/mating starts all over again. This is repeated until a predefined number of iterations are reached.

Mutation which introduces random deviations into the population was also introduced into the GR during the mating step at a rate of 1% as is typical in GAs. This is usually done by replacing one of the base pairs in an existing gene with a randomly generated new base pair. Mutation allows the GR to explore the search space and incorporate new material into the genetic population. It helps to keep the search moving and can eject GR from a local minimum on the response surface. However, it is important not to set the mutation rate too high since it may keep the GA from being able to exploit the existing population.

Because the GR method is ended with a number of iteration, it is likely that a highly over fitted model may result. To avoid this problem cross validation approach is used during the initial gene selection and iteration cycles. Cross validation is done in way that each sample in the

calibration set is left outside once and the model is built with *m-1* number of samples and this model is used to predict the left over sample. The predicted error of sum of squares (PRESS) is calculated. Then, the RMSEC value is determined after calculating all PRESS values for the samples in the calibration set. In the end, the gene with the lowest RMSEC (highest fitness) is selected for the model building which is done by simple least squares. This model is used to predict the concentrations of component being analyzed in the validation sets. The success of the model in the prediction of the validation sets are evaluated using root mean square error of validation (RMSEV). Because the random processes are heavily involved in the GR as in all the GAs, the program has been set to run several times for each component in a given multi-component mixture during the course of this study. The best run, (i.e. the one generating the lowest RMSEC for the calibration set and at the same time producing RMSEV's for validation sets that are in the same range with the RMSEC) was subsequently selected for evaluation and further analysis. The termination of the algorithm can be done in many ways. The easiest way is to set a predefined iteration number for the number of breeding/mating cycles.

GR has some major advantages over classical univariate and multivariate calibration methods. It is a hybrid calibration method that uses the full spectral information and reduces it to a single score upon which simple calibration models are built. First of all, it is as simple as univariate calibration in terms of the mathematics involved in the model building and prediction steps, but at the same time it has the advantages of the multivariate calibration methods since it uses the full spectrum to extract genetic scores. It automatically corrects baseline fluctuations using simple mathematical operators while forming the base pairs. Also note that no data pretreatment is necessary before calibration, which saves the extra time in the data processing.

II. *Genetic Classical Least Squares*

The classical least squares (CLS) method extends the classical Beer's Law model in which the absorbance at each wavelength is directly proportional to the component concentrations. Model errors are assumed to be in the measurement of the instrument responses as it was in the classical univariate method. In matrix notation, the CLS model for *m* calibration samples containing *l* chemical components whose spectra contain *n* wavelengths is described as:

$$\mathbf{A} = \mathbf{CK} - \mathbf{E_A} \tag{5}$$

where **A** is the $m \times n$ matrix of the calibration spectra, **C** is the $m \times l$ matrix of the component concentrations, **K** is the $l \times n$ matrix of absorptivity-pathlength constants and $\mathbf{E_A}$ is the $m \times n$ matrix of the spectral errors or residuals not fit by the model. Here the **K** matrix represents the first order estimates of the pure component spectra at unit concentration and unit pathlength. The method of least-squares can be used to estimate the **K** matrix. The least-squares estimate of the **K** is defined as:

$$\hat{\mathbf{K}} = (\mathbf{C}'\mathbf{C})^{-1}\,\mathbf{C}'\mathbf{A} \tag{6}$$

Once the estimated $\hat{\mathbf{K}}$ matrix obtained, the concentrations of an unknown sample can be predicted from its spectrum by:

$$\hat{\mathbf{c}} = (\hat{\mathbf{K}}\,\hat{\mathbf{K}}')^{-1}\hat{\mathbf{K}}\,\mathbf{a} \tag{7}$$

where **a** is the spectrum of the unknown sample and $\hat{\mathbf{c}}$ is the vector of the predicted component concentrations. Genetic Classical Least Squares (GCLS) is a modified version of the original CLS method in which a small set of wavelengths are selected from a full spectral data using a genetic algorithm. The algorithm used to select the optimum number of wavelengths in GCLS is quite similar to the GR algorithm, but differs in the way it encodes the gene. In GCLS, the term "gene" describes a vector whose elements are randomly selected wavelengths. The size of the vector is also determined in a random fashion with an upper limit to reduce computation time.

In the initialization step, an even number of genes are formed from full a spectral data matrix and each gene is used to form a CLS model. These models are then evaluated and ranked using the fitness function described in GR. The roulette wheel method is then used to select the gene population for breeding. After the selection procedure is completed, the selected genes are allowed to mate top-down without ranking whereby the first gene mates with the second gene and the third one with the fourth one and so on as described above with one difference. Since the genes used in GCLS are only vector of wavelengths and contain no base pairs as described in GR, for each gene a random number is generated between 1 and the length of the gene and the single point crossover process is performed using this number. After crossover, the parent genes are replaced by their offspring and the offspring are evaluated. The ranking process is based on their fitness values and follows the evaluation step. Then the selection for breeding/mating starts all over again. This is repeated until a predefined number of iterations are reached. In each iteration, the best gene with the lowest RMSEC is stored in order to compare it with the best gene of the next generation. If the next generation produces a better gene then it is replaces the older one; otherwise the older one is kept for further iterations. At the end, the gene with the lowest RMSEC is selected for model building. This model is used to predict the concentrations of component being analyzed in the validation sets as described in GR.

III. *Genetic Inverse Least Squares*

The major drawback of the CLS is that all of the

interfering species must be known and their concentrations are included in the model. This need can be eliminated by using the inverse least squares (ILS) method which uses the inverse of Beer's Law. In the ILS method, concentrations of an analyte are modelled as a function of absorbance measurements. Because modern spectroscopic instruments are very stable and provide excellent signal-to-noise (S/N) ratios, it is believed that the majority of errors lie in the reference values of the calibration samples, not in the measurement of their spectra. The ILS model for *m* calibration samples with *n* wavelengths for each spectrum is described by:

$$\mathbf{C} = \mathbf{AP} - \mathbf{E_C} \tag{8}$$

where **C** and **A** are the same as in CLS, **P** is the $n \times l$ matrix of the unknown calibration coefficients relating *l* component concentrations to the spectral intensities and $\mathbf{E_C}$ is the $m \times l$ matrix of errors in the concentrations not fit by the model. In the calibration step, ILS minimizes the squared sum of the residuals in the concentrations. The biggest advantage of ILS is that equation 8 can be reduced for the analysis of single component at a time since analysis is based on an ILS model is invariant with respect to the number of chemical components included in the analysis. The reduced model is given as:

$$\mathbf{c} = \mathbf{Ap} - \mathbf{e_c} \tag{9}$$

where **c** is the $m \times 1$ vector of concentrations for the analyte that is being analyzed, **p** is $n \times 1$ vector of calibration coefficients and $\mathbf{e_c}$ is the $m \times 1$ vector of concentration residuals not fit by the model. During the calibration step, the least-squares estimate of **p** is:

$$\hat{\mathbf{p}} = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}' \cdot \mathbf{c} \tag{10}$$

where $\hat{\mathbf{p}}$ is the estimated calibration coefficients. Once $\hat{\mathbf{p}}$ is calculated, the concentration of the analyte of interest can be predicted with the equation below.

$$\hat{c} = \mathbf{a}' \cdot \hat{\mathbf{p}} \tag{11}$$

where $\hat{c}$ is the scalar estimated concentration and **a** is the spectrum of the unknown sample. The ability to predict one component at a time without knowing the concentrations of interfering species has made ILS one of the most frequently used calibration methods. However, the identity of interfering species still needs to be known to prepare a good calibration sample set.

The major disadvantage of ILS can be seen in equation (10) where the matrix, which must be inverted, has dimensions equal to the number of wavelengths in the spectrum and this number should not exceed the number of calibration samples. This is a big restriction since the number of wavelengths in a spectrum will generally be more than the number of calibration samples and the selection of wavelengths that provide the best fit for the model is not a trivial process. Several wavelength selection strategies, such as stepwise wavelength selection and all possible combination searches, are available to build an ILS model that fits the data best. Here we used the same genetic algorithm described in GCLS to build genetic inverse least squares (GILS) models with one difference. This difference is in the way the mating and single point crossover operations are carried out. Because the number of wavelengths is restricted in response matrix A in the ILS, the size of the largest gene is restricted to one less than the number of calibration samples in the concentration vector. However, if the single point crossover is set to take place in any point of a gene, then the mating step could produce new genes that have a larger number of wavelengths than the number of calibration samples even though all the genes in the initial gene pool were set to have smaller number of wavelengths than the size of the concentration vector. In order to avoid this problem, the crossover operation is only performed somewhere around the middle of each gene in GILS so that the new generations will not have larger sizes than the number of calibration samples. The rest of the algorithm is the same as the one used in GCLS.

MATERIALS AND METHODS

I. *Materials*

In this work, two commercial pharmaceutical formulations, Cibadrex and Divitab – 5/6.25 (I) and 10/12.5 (II) tablets (produced by Novartis Pharm., Turkey, Batch No. 13 and 18, respectively) containing 5 and 10 mg of benazepril hydrochloride (BE) and 6.25 and 12.5 mg of hydrochlorothiazide (HCT) were investigated. Stock solutions of 100 mg/100 mL HCT and BE were prepared in 0.1 M NaOH. The standard solutions in 25-mL volumetric flasks containing 0~22 mg/mL HCT and 0~36 mg/mL BE were obtained from their stock solutions by appropriate dilution. The concentration profiles of calibration and validation samples were designed in a way that minimizes colinearity problem since a binary system has been studied. For the commercial vitamin, 16 tablets were accurately weighed and powdered in a mortar. An amount equivalent to one tablet was dissolved in 0.1 M NaOH in a 100-mL calibrated flask by sonication. The solution was filtered into a 100-mL calibrated flask through Whatman No. 42 filter paper and diluted to appropriate volume with the same solvent.

II. *Methods*

Sample spectra were measured in a Shimadzu UV-1600 double beam UV-visible spectrophotometer from 210 to 360 nm with 0.1 nm intervals. Quartz cells with 1 cm pathlengths were used. The CLS and the three new genetic algorithms based multivariate calibration

Journal of Food and Drug Analysis, Vol. 13, No. 4, 2005

methods (GCLS, GILS and GR) were written in MATLAB programming language using Matlab 5.3 (MathWorks Inc, Natick, MA, USA). The text files for calibration, validation and prediction sets were generated with the use of Microsoft Excel (MS office 97, Microsoft Corporation, CA, USA).

RESULTS AND DISCUSSION

To generate the calibration models, a total of 20 samples were selected to be included in the calibration set and 8 samples were used to construct the validation set as shown in Table 1. In addition, two different commercial tablets (each consists of 8 samples) were used to build prediction set. The first contains 5 mg benazepril per tablet and 6.25 mg hydrochlorothiazide per tablet. The second contains 10 mg benazepril per tablet and 12.5 mg hydrochlorothiazide per tablet. After dissolving, the tablet samples were diluted to have 16 μg/mL of benazepril and 20 μg/mL of hydrochlorothiazide.

UV spectra of pure benazepril (36 μg/mL) and hydrochlorothiazide (22 μg/mL) along with the binary mixture of the two components between 210 and 360 nm wavelength range are shown in Figure 1. As seen from the figure, benazepril gives a broad peak with the maximum around 242 nm and hydrochlorothiazide has a maximum absorbance around 271 nm. Their mixture spectrum, however, indicates some overlap over the entire region which indicates that the use of multivariate methods would be needed to resolve these components. Throughout the genetic multivariate calibration process, it is expected that these overlaps will be resolved and reveal the information necessary to build successful calibration models otherwise almost impossible with univariate calibration methods.

Several calibration models were generated with the four methods and Table 2 shows the results of binary mixtures for calibration and validation sets. Here, the CLS method was applied to the whole spectrum data set and in

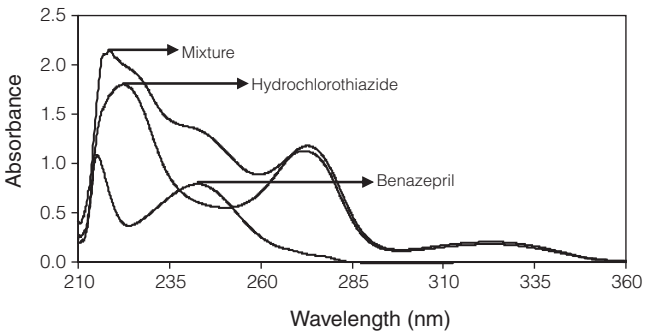

Figure 1. UV spectra of benazepril (36 μg/mL) and hydrochlorothiazide (22 μg/mL) along with a binary mixture of the two components between 220 and 360 nm wavelength range.

Table 1. Concentration profiles of benazepril (BE) and hydrochlorothiazide (HCT) binary mixtures in the calibration, the validation and the prediction (actual tablets) sets

Sample number	Calibration (μg/mL)		Validation (μg/mL)		Prediction (μg/mL)	
	BE	HCT	BE	HCT	BE	HCT
1	12	0	16	0	16	20
2	20	0	32	0	16	20
3	24	0	20	20	16	20
4	28	0	24	20	16	20
5	36	0	0	12	16	20
6	12	20	0	18	16	20
7	16	20	16	14	16	20
8	28	20	16	16	16	20
9	32	20			16	20
10	36	20			16	20
11	0	10			16	20
12	0	14			16	20
13	0	16			16	20
14	0	20			16	20
15	0	22			16	20
16	16	10			16	20
17	16	12				
18	16	18				
19	16	20				
20	16	22				

Table 2. Results of the mixtures in the calibration and the validation sets containing benazepril and hydrochlorothiazide obtained from the four multivariate calibration methods (CLS, GCLS, GILS and GR)

Parameters	Methods			
	CLS	GCLS	GILS	GR
Benazepril				
Calibration				
RMSEC[a] (μg/mL)	0.48	0.33	0.19	0.34
Average recovery (%)	99.91	100.22	99.89	99.98
RSD[b]	3.49	2.63	1.33	2.48
Validation				
RMSEV[c] (μg/mL)	0.38	0.34	0.35	0.43
Average recovery (%)	99.60	99.43	100.44	99.19
RSD	2.60	1.96	1.91	3.03
Hyrochlorothiazide				
Calibration				
RMSEC (μg/mL)	0.63	0.29	0.24	0.31
Average recovery (%)	99.04	99.62	99.89	99.74
RSD	5.38	2.11	1.36	1.91
Validation				
RMSEV (μg/mL)	0.65	0.26	0.18	0.23
Average recovery (%)	98.49	98.70	99.18	98.84
RSD	5.30	1.21	0.99	1.15

[a]Root mean square error of calibration.
[b]Relative standard deviation.
[c]Root mean square error of validation.

Table 3. Results of commercial tablets (I) containing benazepril HCl (5 mg/tablet) and hydrochlorothiazide (6.25 mg/tablet) obtained from the four multivariate calibration methods (CLS, GCLS, GILS and GR)

Component	Predicted (mg/tablet)							
	Benazepril HCl				Hydrochlorothiazide			
Method	CLS	GCLS	GILS	GR	CLS	GCLS	GILS	GR
Mean	5.47	4.90	5.00	4.89	6.87	6.26	6.21	6.31
SD[a]	0.04	0.07	0.09	0.07	0.05	0.06	0.05	0.06
RSD[b]	0.81	1.37	1.82	1.50	0.76	0.89	0.74	0.92
RMSEP[c]	0.48	0.12	0.09	0.13	0.62	0.05	0.06	0.08

[a]Standard deviation.
[b]Relative standard deviation.
[c]Root mean square error of prediction.

Table 4. Results of commercial tablets (II) containing benazepril HCl (10 mg/tablet) and hydrochlorothiazide (12.5 mg/tablet) obtained from the four multivariate calibration methods (CLS, GCLS, GILS and GR)

Component	Predicted (mg/tablet)							
	Benazepril HCl				Hydrochlorothiazide			
Method	CLS	GCLS	GILS	GR	CLS	GCLS	GILS	GR
Mean	10.50	10.05	9.80	9.93	14.02	12.54	12.52	12.64
SD[a]	0.01	0.02	0.04	0.04	0.03	0.02	0.03	0.02
RSD[b]	0.08	0.24	0.43	0.38	0.20	0.18	0.27	0.15
RMSEP[c]	0.50	0.05	0.20	0.08	1.52	0.05	0.04	0.14

[a]Standard deviation.
[b]Relative standard deviation.
[c]Root mean square error of prediction.

Journal of Food and Drug Analysis, Vol. 13, No. 4, 2005

the case of genetic algorithm based methods (GCLS, GILS and GR) the algorithms were set to run 30 times with 20 genes and 100 iterations in each run. The results given in table for GCLS, GILS and GR are from the runs that generate the lowest RMSEC and RMSEV combination. Then these models were used later to predict the actual tablet samples and compared with each other based on their success of predicting actual samples as shown in Tables 3 and 4.

A close examination of the results given in Table 2 indicates that all four methods generate approximately the same results for benazepril whereas genetic algorithm based methods produces somewhat better results than CLS for hydrochlorothiazide in the synthetic mixtures. However, this could be very misleading conclusion if one considers the results given in Tables 3 and 4 where the results of actual tablets are shown for the first and second type tablets, respectively. The RMSEP values generated with conventional CLS methods are much larger than the ones generated by genetic algorithm based methods. On

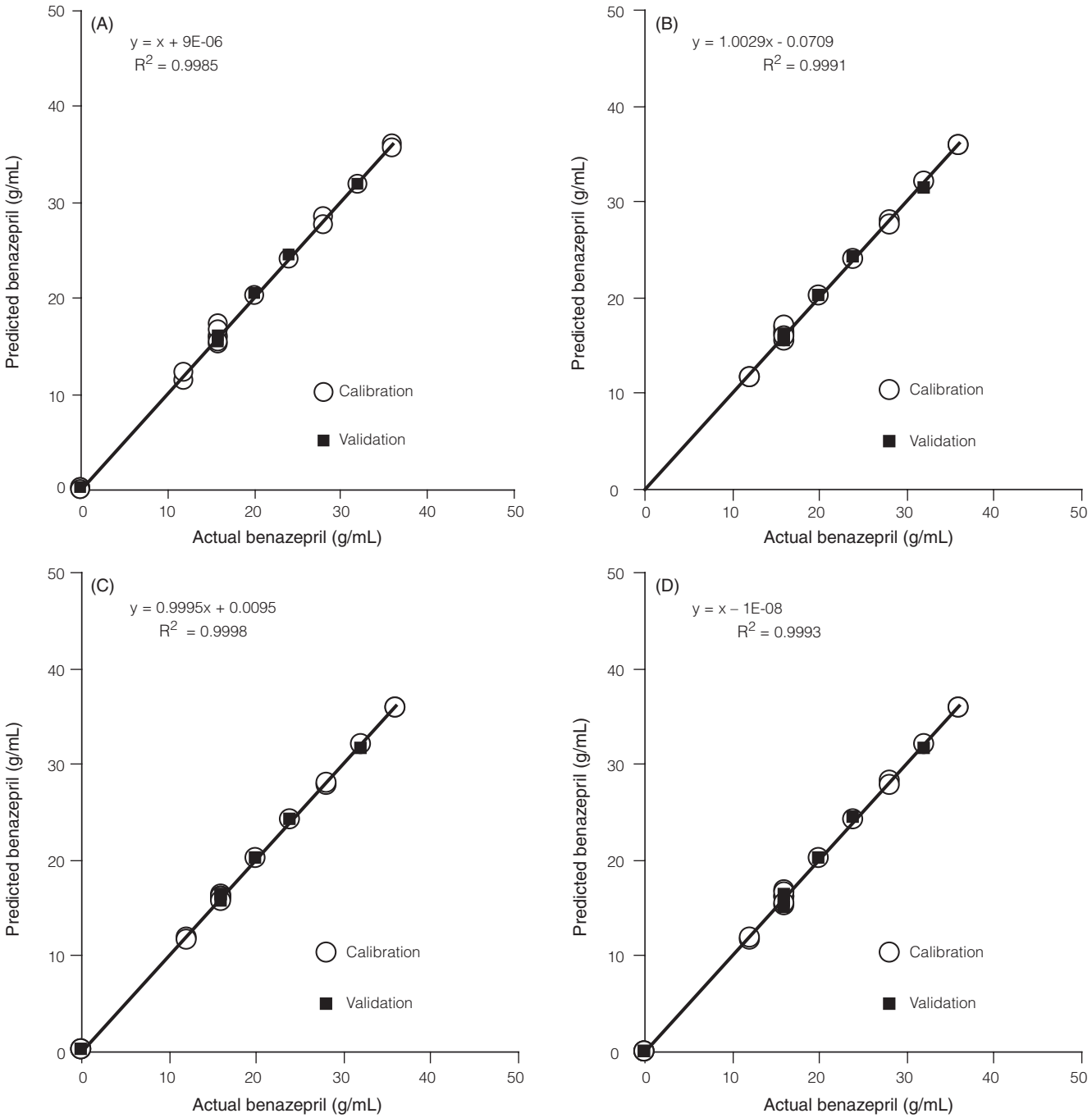

Figure 2. Plots of the actual vs. the predicted benazepril concentrations for the calibration and the validation sets obtained from the four multivariate calibration methods: (A) CLS, (B) GCLS, (C) GILS and (D) GR method.

the other hand, the three genetic multivariate methods were generated very similar results where GILS seems to have slightly better than the other two but the differences does not indicate a significant difference. It is evident that the hard modelling method CLS is unable to predict the composition of actual tablets as good as the genetically modified multivariate methods. In addition, the mean tablet results obtained with CLS were significantly different from actual values which might be the indication of overfit for the model. In terms of the overall performance of the four methods it can be said that the genetically modified methods improves the prediction ability of models for actual tablet samples.

Figures 2 and 3 show the plot of actual vs. predicted benazepril and hydrochlorothiazide concentrations, respectively for the calibration and validation sets obtained with the four methods in the second experiment. The R square (R^2) values of regression were ranged between 0.9950 and 09998 indicating very good fit between actual and predicted concentrations for the synthetic samples.

In order to determine whether the genetic algorithm selected wavelengths correspond to the particular

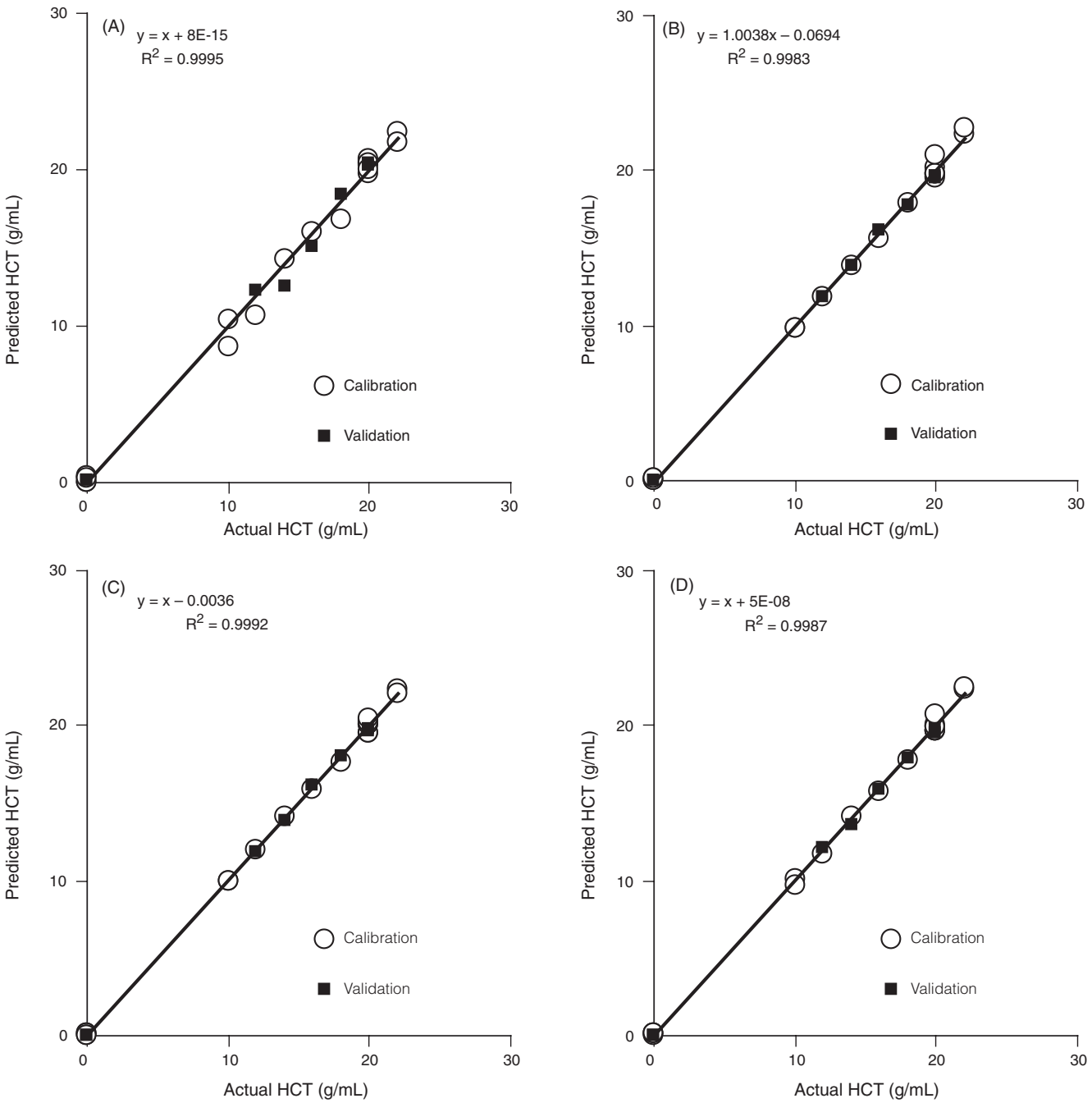

Figure 3. Plots of the actual vs. the predicted hydrochlorothiazide (HCT) concentrations for the calibration and the validation sets obtained from the four multivariate calibration methods: (A) CLS, (B) GCLS, (C) GILS and (D) GR method.

Journal of Food and Drug Analysis, Vol. 13, No. 4, 2005

component absorbance region, frequency of the selected wavelengths in the 30 runs for each genetic algorithm based method were plotted against wavelength range along with a mixture spectrum in Figures 4~6 for each component. As seen from the figures, the frequency of the selected wavelengths is significantly higher around the peak maximum of each component. This shows that the genetic multivariate calibration methods select the wavelengths that correspond to the each component absorption range even though the algorithm starts with the whole spectrum information at the beginning of each run and each wavelength has equal chance of being selected. The explanation is in the evolutionary nature of genetic algorithm where the wavelengths suited for the particular component survives in the long run of iterations and other do not. This gives an advantage to the genetic algorithm based methods where only the information related to the particular component are used to construct the model thereby reducing the noise in the overall information.

For example, there are three regions of high selection frequency for benazepril in Figure 4A. One of the regions corresponds to the main absorbance peak of benazepril as shown in Figure 1 and the other two correspond to the baseline area. This trend is also observed in the corresponding Figures 5A and 6A. On the other hand, frequency distribution of hydrochlorothiazide seems to spread more over the entire wavelength region. The possible explanation could be the strong dominance of hydrochlorothiazide spectrum over benazepril.

CONCLUSIONS

This study illustrates the application of the hard modelling technique CLS and three genetic algorithm based multivariate calibration methods to simultaneous determination of pharmaceuticals in synthetic and actual tablet formulations. It can be said that all genetic algorithm based methods generate acceptable results in the given concentration range of the components. These methods coupled with spectrophotometry could be an alternative to other methods such as chromatography, which is more expensive and time-consuming.

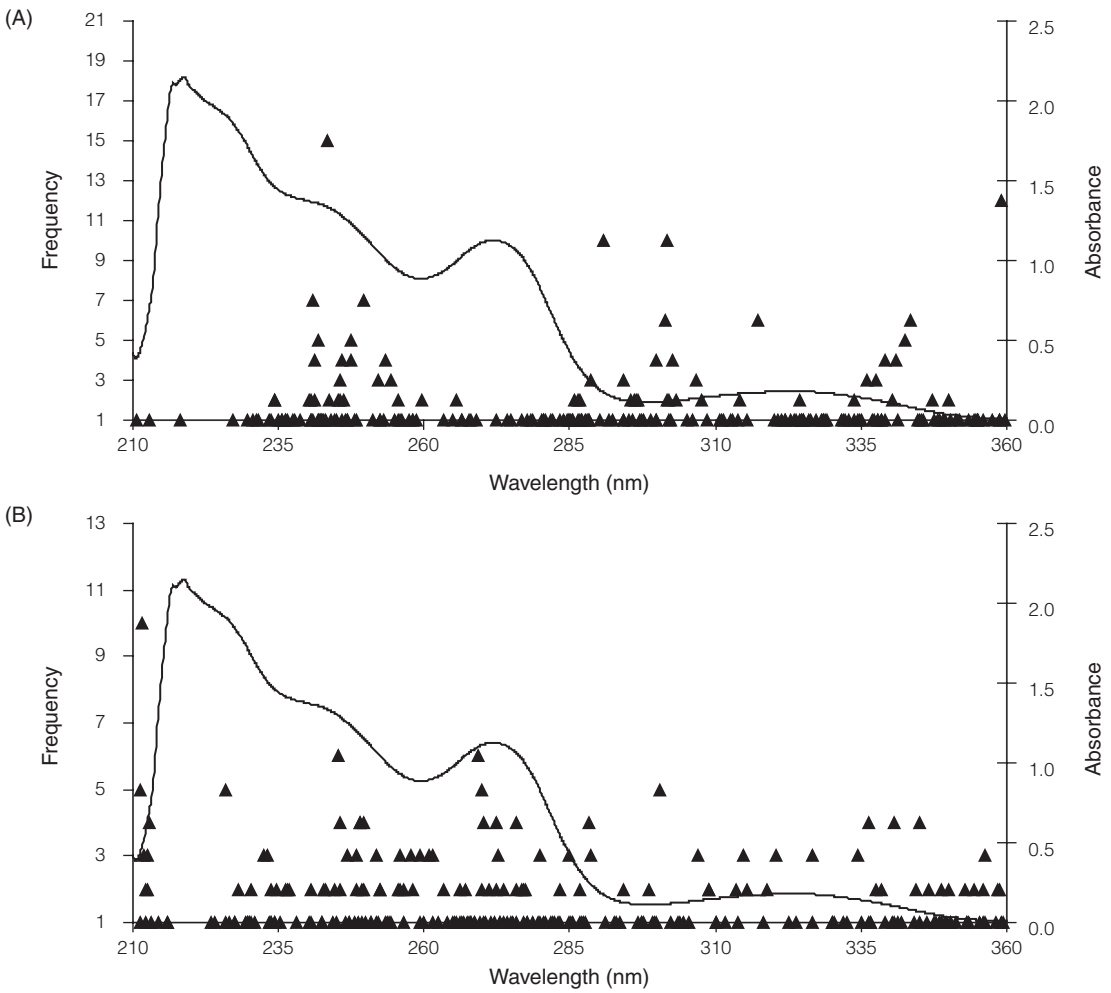

Figure 4. Distribution of the selected wavelengths by genetic algorithm in GCLS method for a total of 50 runs with 20 genes and 100 iterations along with a spectrum of binary mixture: (A) benazepril and (B) hydrochlorothiazide.

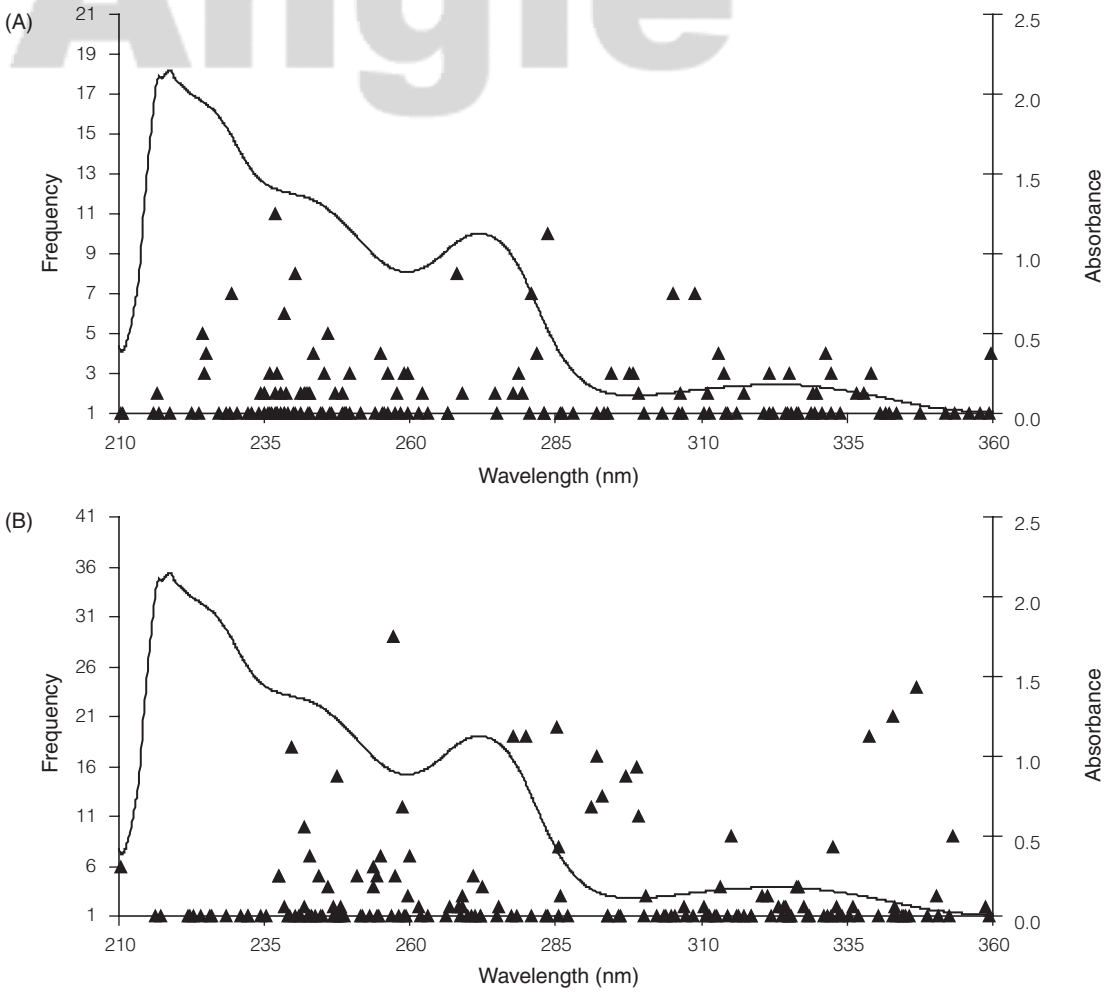

Figure 5. Distribution of the selected wavelengths by genetic algorithm in GILS method for a total of 50 runs with 20 genes and 100 iterations along with a spectrum of binary mixture: (A) benazepril and (B) hydrochlorothiazide.

REFERENCES

1. Dinç, E. and Üstündağ, Ö. 2003. Spectrophotometric quantitative resolution of hydrochlorothiazide and spironolactone in tablets by chemometric analysis methods. Il Fafmaco 58: 1151-1161.
2. Dinç, E. 2002. Spectral analysis of benazepril hydrochloride and hydrochlorothiazide in pharmaceutical formulations by three chemometric techniques. Anal. Letters 35: 1032-1039.
3. El-Gindy, A., Ashour, A., Abdel-Fattah, L. and Shabana, M. M. 2001. Spectrophotometric determination of benazepril hydrochloride and hydrochlorothiazide in binary mixture using second derivative, second derivative of the ratio spectra and chemometric methods. J. Pharm. Biomed. Anal. 25: 299-307.
4. Erk, N. 1999. Determination of active ingredients in the pharmaceutical formulations containing hydrochlorothiazide and its binary mixtures with benazepril hydrochloride, triamterene and cilazapril by ratio spectra derivative spectrophotometry and vierordt's method. J. Pharm. Biomed. Anal. 20: 155-167.
5. Panderi, I. E. 1999. Simultaneous determination of benazepril hydrochloride and hydrochlorothiazide in tablets by second order derivative spectrophotometry. J. Pharm. Biomed. Anal. 21: 257-265.
6. Banoğlu, E., Özkan, Y. and Atay, O. 2000. Dissolution tests of benazepril–HCl and hydrochlorothiazide in commercial tablets: Comparison of spectroscopic and high performance liquid chromatography methods. Il Farmaco 55: 477-483.
7. Panderi, I. E. and Parissi-Poulou, M. 1999. Simultaneous determination of benazepril hydrochloride and hydrochlorothiazide by micro-bore liquid chromatography. J. Pharm. Biomed. Anal. 21: 1017-1024.
8. Khalil, S. and El-Aliem, S. A. 2002. Potentiometric and thermal studies of a coated-wire benazepril-selective electrode. J. Pharm. Biomed. Anal. 27: 25-29.
9. Hillaert, S. and Van den Bossche, W. 2001. The quantitative determination of several inhibitors of the angiotensin-converting enzyme by CE. J. Pharm. Biomed. Anal. 25: 775-783.
10. Haaland, D. M. and Thomas, E. V. 1988. Partial least-

Journal of Food and Drug Analysis, Vol. 13, No. 4, 2005

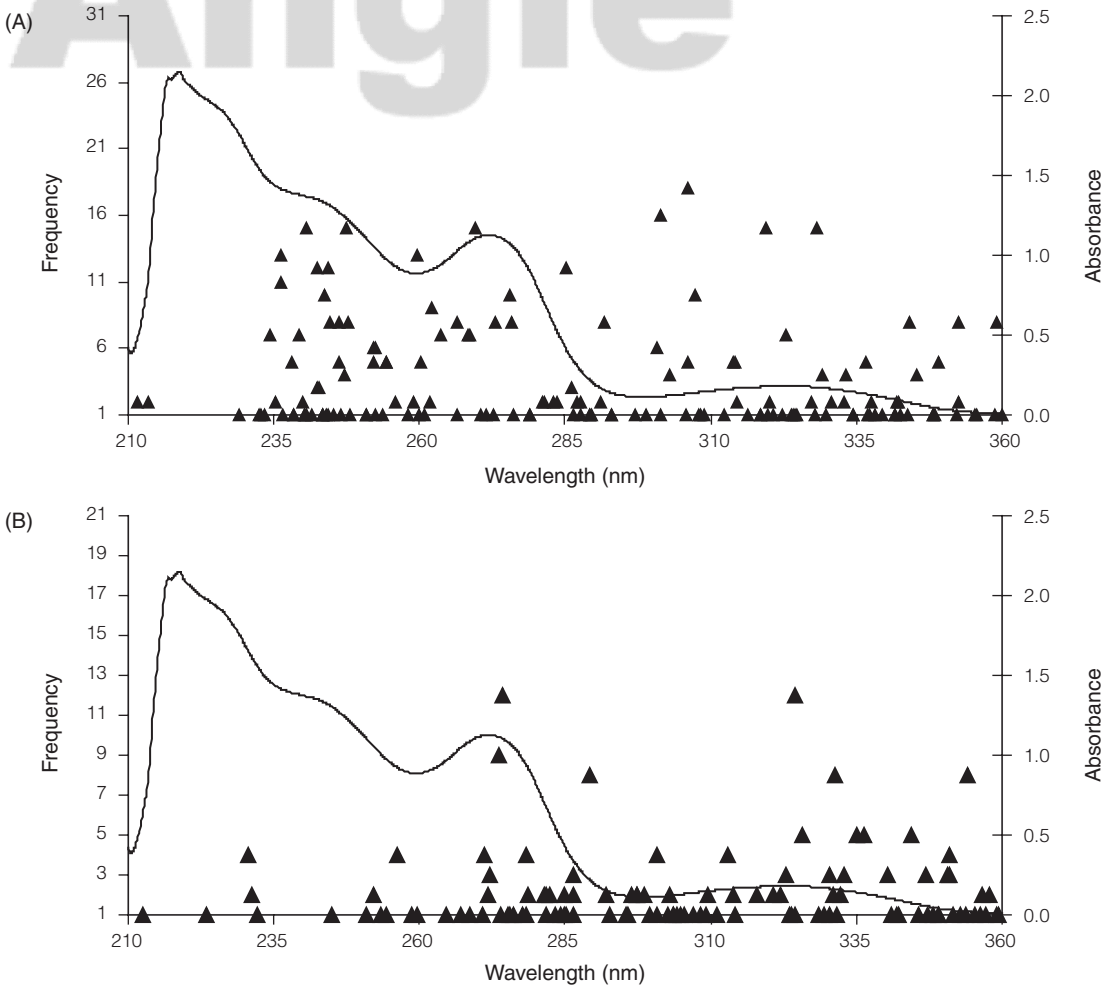

Figure 6. Distribution of the selected wavelengths by genetic algorithm in GR method for a total of 50 runs with 20 genes and 100 iterations along with a spectrum of binary mixture: (A) benazepril and (B) hydrochlorothiazide.

squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. Anal. Chem. 60: 1193-1202.

11. Geladi, P. and Kowalski, B. R. 1986. Partial least-squares regression: A tutorial. Anal. Chim. Acta 186: 1-17.
12. Wentzell, P. D., Andrews, D. T. and Kowalski, B. R.1997. Maximum likelihood multivariate calibration. Anal. Chem. 69: 2299-2311.
13. Esbensen, K., Geladi, P. and Wold, S. 1987. Principal component analysis. Chem. Intell. Lab. Syst. 2: 37-52.
14. Özdemir, D. and Öztürk, B. 2004. Genetic multivariate calibration methods for near infrared (NIR) spectroscopic determination of complex mixtures. Turk. J. Chem. 28: 497-514.
15. Özdemir, D. and Dinç, E. 2004. Determination of thiamine HCl and pyridoxine HCl in pharmaceutical preparations using UV–visible spectrophotometry and genetic algorithm based multivariate calibration methods. Chem. Pharm. Bull. 52: 810-817.
16. Özdemir, D. and Williams, R. R. 1999. Multi-instrument calibration with genetic regression in UV-visible spectroscopy. Appl. Spectrosc. 53: 210-217.
17. Özdemir, D., Mosley, R. M. and Williams, R. R. 1998. Hybrid calibration models: An alternative to calibration transfer. Appl. Spectrosc. 52: 599-603.
18. Hörchner, U. and Kalivas, J. H. 1995. Further investigation on a comparative study of simulated annealing and genetic algorithm for wavelength selection. Anal. Chim. Acta 311: 1-13.
19. Lucasius, C. B. and Kateman, G. 1993. Understanding and using genetic algorithms. Part 1. Concepts, properties and context. Chem. Intell. Lab. Syst. 19: 1-33.
20. Murray, J. 1859. The preservation of favored races in the struggle for life. In “The Origin of Species by Means of Natural Selection”. 1st ed. Darwin, C. ed. Pinguin Books. London, U. K.