مرکز مطالعات و همکاری های علمی بین المللی وزارت علوم، تحقیقات و فناوری

5th Iranian Biennial Chemometrics Seminar

پنجمین سمینار دوسالانه کمومتریکس ایران

۵-۴ آذر ماه ۱۳۹۴

دانشکده شیمی پردیس ، علوم دانشگاه تهران

آدرس دبیرخانه سمینار: تهران، خیابان انقلاب، دانشگاه تهران، پردیس علوم

تلفن: ۶۶۴۰۳۶۷۲-۶۱۱۱۲۴۶۱ فکس: ۶۶۴۰۵۱۴۱

Email: chemometrics.ut@gmail.com

به نام خداوند جان و خرد

کزین برتر اندیشه برنگذرد

مجموعه مقالات پنجمین سمینار کمومتریکس ایران

دانشکده شیمی، دانشگاه تهران، آذرماه ۹۴

The Proceeding of 5th Iranian Biennial Chemometrics Seminar

Chemistry Faculty, University of Tehran, Tehran, Iran, 25-26 Nov 2015

فهرست مطالب

اعضای کمیته علمی

- دکتر جهانبخش قاسمی
- دکتر علی مقاری
- دکتر سجاد قرقانی
- دکتر تقی خیامیان
- دکتر حمید عبدالهی
- دکتر محمدحسین فاطمی
- دکتر محسن کمپانی زارع
- دکتر بهرام همتی نژاد
- دکتر مرتضی بهرام
- دکتر عبدالحسین ناصری
- دکتر علی نیازی
- دکتر حسن سرشتی
- دکتر هادی پرستار
- دکتر محمدحسین کریمی
- دکتر مهدی موسوی

اعضای کمیته اجرایی

- دکتر جهانبخش قاسمی
- دکتر علی مقاری
- دکتر سجاد قرقانی
- دکتر هادی پرستار
- دکتر حسن سرشتی
- دکتر محمدرضا گنجعلی
- دکتر پرویز نوروزی
- دکتر فرزانه شمیرانی
- دکتر طاهر علیزاده
- دکتر فرنوش فریدبد
- دکتر محمدحسین کریمی

اعضای کمیته دانشجوئی

- علی فراهانی
- راحله رستمیان
- مهرنوش بیات
- سمیه پیرهادی
- زهرا حیدری
- محمدرضا مقدم
- اسماعیل کریمی

حمایت کنندگان

Sponsors

- دانشگاه تهران
- مرکز مطالعات و همکاری های علمی بین‌المللی وزارت علوم، تحقیقات و فناوری
- انجمن شیمی ایران
- بنیاد وان هومبولت

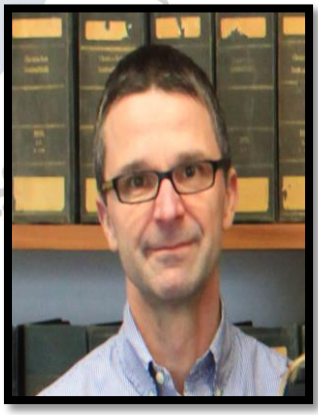
- University of Tehran
- Center for International Scientific Studies & Collaboration
- Iranian Chemical Society
- Alexander van Humboldt Fondation, Germany

Invited Speakers

- ***Federico Marini***

- ***Marcel Maeder***

- ***Kunt Boumann***



Seminar Programs

Wednesday

- *7:30- 8:30, On desk Registration and Reception*
- *8:30- 9:30, Opening Ceremony*
- *9:30-10:00 Coffee Break*
- *10:00- 12:10, Lecture session 1*

Chair: **Prof. Abdollahi and Prof. Ghasemi**

Time	Lecturer	Title
10:00 – 10:35	Taghi Khayamian	Challenges in drug discovery and development
10:35 – 11:10	Marcel Maeder	Chemometrics and Chemistry
11:10– 11:30	Mahsa Akbari	An Investigation on Meaning and Reliability of Local Rank Constraint in Self-Modeling Curve Resolution of Chemical Data
11:30-11:50	Mohammad Ahmadvand	A systematic study on the effect of noise and shift on multivariate figures of merit of second-order calibration algorithms
11:50-12:10	Kobra Samghani	3D-QSAR studies of the inhibition efficiency of phenolic herbicides on photosynthesis by using CoMFA and CoMSIA

- *12:10-14:00, Praying, Launch and Resting*
- *14:30- 15:50, Lecture session 2*

Chair: Prof. Maeder and Prof. Fatemi

Time	Lecturer	Topic
14:00 – 14:35	Hamid Abdollahi	Calculation of Feasible Bands for Visualizing the Constraint Effects in Soft Modeling Methods
14:35 - 15:10	Hadi Parastar	How to Expand the Use of Independent Component Analysis in Analytical Chemistry?
15:10 – 15:30	Saeed Yousefinejad	Quantitative sequence activity relationship of bitter tasting threshold peptides: a comparison study between some two-way and three-way modeling methods
15:30-15:50	Elham Aghaei	Influence of bonded-phase density in chiral chromatography via molecular simulation

- ***15:50 -17:00, Poster session 1 Numbers:1003-1065***
- ***17:00 -18:50, Lecture session 3***

Chair: Prof. Hemmateenejad and Dr. Vosough

Time	Lecturer	Topic
17:00 – 17:35	Mohammad Hossein Fatemi	Present and future of QSAR methodology
17:35 – 18:10	Mehdi Mousavi	Novel feature selection methods and CAIS resolution software
18:10 – 18:30	Masoumeh Rashvand	Investigation of a sewage treatment plant for removal of selective pharmaceutical and personal care products using HPLC-DAD and second-order calibration
18:30-18:50	Zeinab Saberi Dehkordi	Design and construct a small fluoremeter with a smartphone as the detector associated with image processing

Thursday

8:00 - 9:50, Lecture session 4

Chair: Prof. Marini and Dr. Parastar

Time	Lecturer	Topic
8:00 – 8:35	Knut Baumann	Defining the Applicability Domain for Classification Models by Hedging Predictions
8:35 – 9:10	Ali Niazi	Geochemometrics
9:10 – 9:30	Parisa Izadiyan	Metabolite profiling and chemometric classification of two varieties of Ocimum basilicum medicinal plant
9:30– 9:50	Samira Beyramysoltan	Analytical view on uniqueness

9:50-11:00, Poster session 2 Numbers:1068-1122

Thursday, 11:00 - 12:30, Lecture session 5

Chair: Prof. Bauman and Prof. Khayamian

Time	Lecturer	Topic
11:00 – 11:35	Federico Marini	Applications of Particle swarm optimization in chemometrics
11:35 – 12:10	Maryam Vosough	Multi-way assisted chromatographic methodologies for analysing highly complex samples
12:10 – 12:30	Saheleh Sheykhizadeh	Introducing of invasive weed optimization as a new variable selection method

12:30-14:00, Praying, Launch and Resting

14:00 - 15:50, Lecture session 6

Chair: Dr. Naseri and Dr. Mousavi

Time	Lecturer	Topic
14:00 – 14:35	Saeed Masoum	Identification of potential antimicrobial and antioxidant constituents in some medicinal herbs using hephenated chromatograrhic methods and multivariate calibration techniques
14:35 – 15:10	Maryam Salahinezhad	Quantitative Nanostructure Activity Relationship Modeling (Nano-QSAR)
15:10 – 15:30	Akram Rostami	Using global and local sensitivity analysis for variable sorting and selection
15:30-15:50	Afsane Heidari	A theoretical approach to model and predict the adsorption coefficients of some small aromatic molecules on carbon nanotube

- ***15:50-17:00, Poster session 3 Numbers:1123-1171***
- ***17:00 - 18:50, Lecture session 7***

Chair: Dr. Bahram and Dr. Masoum

Time	Lecturer	Topic
17:00 – 17:35	Sajjad Gharaghani	Computational Chemogenomics: an emerging strategy for rapid target and drug discovery
17:35 – 18:10	Ahmad Mani	What is in Chemometrics for Chemoinformatics: Validating Classical Relativity in Chemical Space
18:10 – 18:30	Nematollah Omidikia	Direct Triadic Decomposition

- ***19:00-20:00, Free Discussion Panel***

سخنرانی ها

Oral Presentations

صفحه ۱۵ تا ۶۴

Challenges in drug discovery and development

T. Khayamian
Department of Chemistry, Isfahan University of Technology, Isfahan 84156, Iran

taghi@cc.iut.ac.ir

Drug discovery is a multidimensional method and this work gives a view of it. Drug discovery starts by a selection of a biological target, which is generally a protein. The procedure to a select a validated target is complex, but it is necessary to point out that it also integrates with the selection of drugs. After that, drug delivery begins with screening of thousands or millions of compounds from databases. This defines virtual screening (VS), which is performed in silico and can be performed either ligand based or structure based VS. Nowadays, a combination of ligand and structure based VS is also applied. In structure based BS, the structure of target (protein) is required and selection of the most feasible conformation of protein is still facing demands for doing a lot of computational research (some of our works will be presented). In structure based BS, binding affinity (K_b) between compounds and target is selected as the criterion for screening of the compounds. In the next step, generally an assay (in vitro) is conducted for screening the selected compounds from the previous step VS. In this step, few (possibly one) compound are selected. It would be possible some structural modifications are performed on the selected compounds to enhance bioactivity of the compounds. Then, a lot experiments should be performed to evaluate the potential of the selected compounds. One of these experiments is investigation of the pharmacokinetic (ADME) and toxicology parameters of the selected compounds (some of our works will be presented). For example the residence time of a drug –receptor complex ($K_b = k_{off}/ k_{on}$ and residence time defines as $t_R = 1/k_{off}$), volume of distribution (V_d), half-life of the compounds in the plasma ($t_{1/2}$), clearance (CL), bioavailability (F), and etc. (some of our works will be presented). In addition, interactions of the selected compounds with hERG potassium channel, cytochrome p450, DNA should be investigated. Furthermore, a lot of experiments related to toxicology and physicochemical properties of the compounds should be examined. For example if the selected compounds should enter into the cell nucleus to perform their biological activities (e.g. apoptosis of cancer cell), many experiments should be performed to show this effect. These experiments can be conducted using focal laser spectroscopy and flow cytometry (some works will be presented).

If the compounds passed these assays, go to the next step (in vivo) for investigating and performing furthermore assays. These experiments are several phases and conduct on animals and human being. Finally, if a compound passes all these assays it will introduce as a safe, selective and efficacious drug.

Nowadays, some modifications are conducted to improve the efficacy of drugs. One of these is drug delivery system, which used nanoparticles as the carrier for drugs. This type of delivery system has a lot of advantages relative to using free drugs, because it improves pharmaceutical and pharmacological properties of drugs (some works from our group will be presented). Firstly, it enhances the half-life of the compounds in the plasma ($t_{1/2}$). Secondly, the nanoparticles associated with drug enter into the cell via endocytosis and bypass multidrug resistance MDR) mechanisms, which pumps drugs out of cells. The MDR involves cell surface proteins such as p-glycoproteins (a work from our group is presented). The drug delivery system more improves if targeted drugs are used. In this system, in addition to nanoparticles, a ligand, which interacts with a specific target, is added to the drug delivery system. The ligand can be antibodies or aptamers and leads drug to interact with a specific target and enhances the selectivity of drug.

Multiple target drugs, which means a drug interacts with multiple targets (a work is presented) and protein drugs are being progressed.

3D-QSAR studies of the inhibition efficiency of phenolic herbicides on photosynthesis by using CoMFA and CoMSIA

Kobra Samghani*, Mohammad Hossein Fatemi and Seyedeh Mozhgan Behgozin

Chemometrics Laboratory, Faculty of Chemistry, University of Mazandaran, Babolsar, Iran

E-mail address: *k.samghani@gmail.com*

Abstract

Three-dimensional quantitative structure activity relationship (3D-QSAR) studies were performed to modeling of inhibition efficiency of some phenolic herbicides by using comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA). The models generated by CoMFA and CoMSIA give noncross-validated R^2_{ncv} of 0.97 and 0.97, respectively. The value of leave-one-out (LOO) cross-validated correlation coefficients of CoMFA and CoMSIA model equal to 0.67 and 0.69, respectively, revealed that the models are useful tools for the prediction of test sets. The external validation indicated that CoMFA and CoMSIA possess good predictive power with R^2_{test} values of 0.95 and 0.97 respectively. Relative contributions fields of the CoMFA and CoMSIA models have shown that steric effects govern the bioactivity of the compounds and the bulky substituents will increase the inhibitory activity of the substituted phenols on photosynthesis process.

Key words: QSAR; CoMFA; CoMSIA; PLS; Contour map; Herbicide.

Introduction

Herbicides are including abroad classes of chemicals which used to eliminate the weeds. They destroy the plants by various mechanism[1]. Some of these mechanisms of herbicides are inhibition the amino acid biosynthesis, disorganization the lipid synthesis processes, inhibition or disruption cell division, blockage the carotenoid synthesis and inhibition the photosynthesis process. A class of herbicide that inhibit the photosynthesis process is substituted phenols that block photosynthetic electron transport in photosystem II [1,2].

Quantitative structure activity relationships (QSAR) are the most important applications of chemometrics giving useful relationship between biological activity and molecular properties [3]. In this work, three-dimensional quantitative structure–activity relationships (3D-QSAR) was applied for modeling the inhibition activity of 51 substituted phenols on photosystem II in photosynthetic reaction center by using comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) [4].

Results and discussion

The Structures of phenol derivatives were constructed by SYBYL-X 1.1 program package. Partial atomic charges were calculated by the Gasteiger–Huckel method, and energy minimization was performed using the tripos force field with 0.001 kcal/mol energy gradient convergence criterion and maximum iteration of 5000. Compound 15 which has the most potent inhibitory activity was chosen as a template molecule then other compounds were aligned on the basis of the common structure as rigid body by using Distill in SYBYL-X 1.1.

The steric and electrostatic fields as CoMFA descriptors and the steric, electrostatic, hydrophobic, hydrogen-bond donor and hydrogen-bond acceptor fields as CoMSIA descriptors were calculated at each point of 3D cubic lattice with grid spacing of 2 A° by using a sp^3 hybridized carbon atom with +1 charge as probe atom.

For constructing the 3D-QSAR models, the method of partial least squares (PLS) regression was used to correlate the variations in biological activities of chemicals with variations in their CoMFA and CoMSIA descriptors. The statistical parameters of these models for 41 and 10 chemicals in the training and test set, respectively, are presented in Table 1. The values of the leave one out cross-validation correlation coefficient (q^2) indicate the robustness of CoMFA and CoMSIA models (q^2>0.5). The statistical parameters of non-cross-validated correlation coefficient (R^2_{ncv}) and standard error of stimate (SEE) indicate high predictive ability of the models.

Acording to the contribution of each feild in CoMFA and CoMSIA models, the steric field is the most effective field on the bioactivity of the compounds. As shown in Figure 1 the bulky substituent in all positions will increase the inhibitory activity of the substituted phenols on photosynthesis process.

Table1. Statistical quality parameters of built models.

	CoMFA	CoMSIA
q^2	0.67	0.69
R^2_{ncv}	0.97	0.97
SEE	0.199	0.196
F	119	122
R^2_{test}	0.95	0.97
SEE_{test}	0.33	0.34
$R^2_{bootstraping}$	0.99	0.950
#Component	9	9

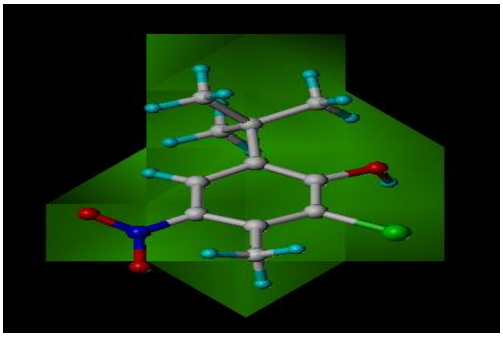


Figure 1. The StDev*coeff. contour map of the CoMFA steric field based on the most acvtine compound.

Refrences

[1]Duke SO. Overview of herbicide mechanisms of action. *Environ. Health. Persp.*1990; **87**: 263-271.

[2]Trebst A. Hilp U. Draber W. Response in the inhibitor efficiency of substituted phenols on PS II activity in six mutants of the D1 protein subunit in Chlamydomonas reinhardtii. *Phytochemistry*. 1993**; 33**: 969-977.

[3]Cunha EFF. Sippl W. Castro Ramalho T. Antunes, OAC. Alencastro, RB. Albuquerque MG. 3D-QSAR CoMFA/CoMSIA models based on theoretical active conformers of HOE/BAY-793 analogs derived from HIV-1 protease inhibitor complexes. *Eur. J. Med. Chem.* 2009; **44**: 4344-4352.

[4] Cramer RD., Patterson DE. Bunce JD. Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins; *J. Am. Chem. SOC.* 1988; **110**: 5959-5967.

Identification of potential antimicrobial and antioxidant constituents in some medicinal herbs using hephenated chromatograrhic methods and multivariate calibration techniques

Saeed Masoum

Department of Analytical Chemistry, Faculty of Chemistry, University of Kashan, I. R. Iran

E-mail address: masoum@kashanu.ac.ir

In the last decades, much emphasis has been placed on herbal and traditional medicine and natural products have become a major area of scientific research [1]. Nowadays traditional and herbal medicine for production of new drugs is under investigation by many scientists. The use of aromatic plants in phytotherapy is mostly due to their various biological activities such as antioxidant, antimicrobial, hepatoprotective, carminative, anticarcinogenic and antiviral properties [2]. As these herbal medicines usually are prescribed traditionally and based on empirical information, therefore, it is necessary to set up standard methods for assessment of their quality.

Much research has been documented about antimicrobial and antioxidant properties of different species of medicinal plants but there is no plenary study about antimicrobial and antioxidant active compounds and the importance of them.

Antimicrobial activity of each sample was evaluated by determining the diameters of inhibition zones and minimum inhibitory concentration (MIC) against *Candida albicans*, *Shigelladysanteriae* and *Klebsiella pneumonia* for each microorganism. Disk diffusion assay was performed by agar disk diffusion method (National Committee for Clinical Laboratory Standard). The disks (6 mm in diameter) were impregnated with 10 µl of the sample and the inoculated plates were incubated for 24 h at 37 ^{o}C. The diameters of inhibition zones were used as a measure of antimicrobial activity.

The antioxidant activity was determined with a 1,1-diphenyl-2-picrylhydrazyl (DPPH) radical scavenging activity test. The radical has an absorption band at 515 nm which disappears upon reduction by the antioxidant compounds present in the samples, resulting in an inverse correlation ofthe remaining absorbance at 515 nm to the antioxidant activity of the sample.

Liquid chromatography coupled with mass spectrometry (LC-MS) and gas chromatography hyphenated with mass spectrometry (GC-MS) are frequently used for metabolomic profiling. Because of the complexities of LC-MS and GC-MS, chemometric methods are developed for scrutiny and interpretation of huge chromatographic data.

Combination of chromatographic fingerprints with antimicrobial and antioxidant activities assisted by multivariate calibration techniques such as principal component regression (PCR), partial least square (PLS), orthogonal projections to latent structures (OPLS) and independent component regression (ICR) can identify the peaks potentially responsible for antimicrobial and antioxidant activity [3, 4]. The results of these techniques were displayed as regression plots. The main goal is finding regression coefficients with the greatest potential of interpretability. The lowest value of MIC shows the most power against microorganisms. Therefore, the most negative peak shows the most potential

antimicrobial activity. Because the DPPH scavenging test value decreases by increasing antioxidant activity of a sample, the negative peaks are corresponding to components with antioxidant activities.

Finally, LC-MS and GC–MS were used to indicate the structure of chemical components, which are responsible for antimicrobial and antioxidant activities.

For example, evaluation of the antioxidant activity and identification of the main components that are responsible for antioxidant activity of the thyme samples was investigated [4]. The regression coefficients that are obtained by PCR, PLS, OSC-PLS, and OPLS have been plotted in Figure 1 along with the GC chromatograms of antioxidant samples. The best model is considered as the most interpretable model. According to these plots, compounds that appear at 25.78, 27.12, 39.93, and 40.53 min are responsible for the antioxidant activity behavior. Based on the GC–MS results, four components such as *p*-cymen, γ-terpinen, thymol, and carvacrol are located in the areas that correspond to negative peaks in the regression plot.

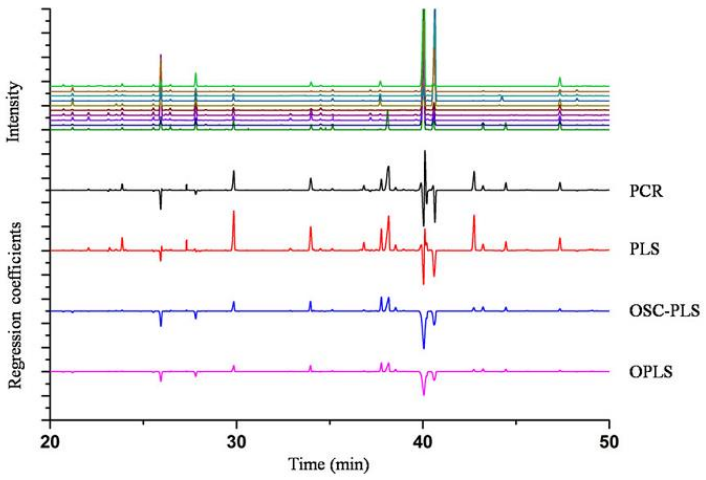


Figure 1. Fingerprints plot of the antioxidant samples. The bottom plots show the regression coefficients obtained from PCR, PLS, OSCPLS, and O-PLS, preprocessed with normalization and mean centering.

References

[1] Mimica-Dukić N, Bugarin D, Grbović S, Mitić-Ćulafić D, Vuković-Gačić B, Orčić D, Jovin E, Couladis M, Essential oil of *Myrtus communis* L. as a potential antioxidant and antimutagenic agents, *Molecules* 2010; **15**: 2759-2770.

[2] F. Bakkali; S. Averbeck; D. Averbeck; M. Idaomar, Biological effects of essential oils – A review, Food Chem. Toxicol. *2008; **46***: 446-475.

[3] Tistaert C, Dejaegher B, Nguyen Hoai N, Chataigné G, Rivière C., Nguyen Thi Hon V, Chau Van M, Quetin-Leclercq J, Vander Heyden Y, Potential antioxidant compounds in Mallotusspecies fingerprints. Part I: indication, using linear multivariate calibration techniques, *Anal. Chim. Acta* 2009; **649**: 24-32.

[4] Masoum S, Mehran M, Ghaheri S, Identification of potential antioxidant compounds in the essential oil of thyme by gas chromatography with mass spectrometry and multivariate calibration techniques, *J. Sep. Sci.* 2015; **38**: 410-417.

Calculation of Feasible Bands for Visualizing the Constraint Effects in Soft Modeling Methods

Hamid Abdollahi,
Institute for Advanced Studies in Basic Sciences, Department of Chemistry, PO Box 4, Zanjan, 5195-1159 Iran

The curve resolution (CR) methods magnetize substantial research efforts aimed to discover knowledge of multicomponent systems. The soft methods establish a specified part in CR methods, as they allow going into insight the data without prior knowledge of the studied system. Using soft methods, the outcome of the analysis of one system is not usually the true one due to rotational ambiguity. The rotational ambiguity leads to a range of possible solutions that fulfill constraints and represent the measured data correctly. Consequently, taking into consideration of all feasible solutions can provide useful information about the system and the process under study when there is no unique solution for the system. Additionally, the calculation of feasible bands is dramatically advantageous not only for quantitative and qualitative application of the resolved profiles but also for visualizing and deep understanding the effects of constraints on results of using soft-modeling methods. Different methods have been proposed in the literature to calculate the feasible solutions. Analytical methods for calculating the feasible solutions in two and three-component systems developed as Lwton-Sylvester and Borgen-Rajko plots. Then the exact and explicit calculations of range of feasible solutions were approximated by numerical methods as grid search procedures up to four component systems. The micro structure of chemical data can be visualized by showing the feasible solution regions in abstract spaces under considered constraints. The effects of non-negativity, equality, selectivity-zero regions and hard modeling constraints on micro structure of chemical data have been investigated basically and some new conclusions have been obtained from interpretation of observations.

Computational Chemogenomics: an emerging strategy for rapid target and drug discovery

Sajjad Gharaghani

Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran

E-mail address: s.gharaghani@ut.ac.ir

Chemogenomics as a scientific discipline at the interface between chemistry and biology searches for all molecules that are capable of interacting with any biological target. Traditional approaches for the identification of bioactive compounds use a chemical library, a single target protein, and an assay, which allows us to measure the activity of these compounds against the selected target. In contrast chemogenomics aims at the identification of the bioactivity of all these compounds against multiple targets and even beyond: in a very general sense the ultimate goal of chemogenomics is characterization of all possible compound-target interactions. As such, the chemogenomics concept is related to chemical genetics and other aspects of chemical biology. Although chemogenomics is a priori an experimentally grounded concept, the systematic assessment of ligand-target interactions provides an attractive playground for computational analysis and predictive modeling. In fact, chemogenomics represents a conceptual framework highly suitable for complementary experimental and computational studies, which has been widely recognized in the field. Without doubt, chemogenomics is one of the current and future growth areas for computational analysis and design in pharmaceutical research and the life sciences. The key principle of Chemogenomics is that similar compounds bind to similar targets. The major opportunity for Chemogenomics is in the close cooperation with other highly active and novel scientific disciplines. Especially, the appearance of Next Generation Sequencing technologies, extensive expression profiling, epigenetics and phosphoproteomics, among others, allow for better understanding of biological systems.

How to Expand the Use of Independent Component Analysis in Analytical Chemistry?

Hadi Parastar*

Chemometrics Research Lab., Department of Chemistry, Sharif University of Technology, Tehran, Iran
Email address: h.parastar@sharif.edu

Abstract

In the present contribution, the potential of newly developed saprsity constraint for independent component analysis (ICA) by Winding and Keenan as a new concept to expand the use of ICA in analytical chemistry is evaluated in terms of independence and the extent of rotational ambiguity in the solutions compared to other ICA algorithms. The prerequisite of statistical independence severely limits the application of ICA. However, increasing the sparsity of a data set increases the independence of components for some classes of data, which enables the successful application of ICA. The sparsity can be increased by simply adding zeros to the data set or by applying a Haar wavelet transform. In this regard, the solutions obtained by ICA with sparsity constraint are compared with different ICA algorithms such as MFICA, SNICA and MILCA. Mutual information (MI) is used as the quantitative measure of independence between ICA solutions. Also, the extent of rotational ambiguity in the solutions is calculated using polygon inflation algorithm (FAC-PACK). It is concluded that ICA with sparsity is important in noisy data. When zero values are present in the solutions, one has to expect positive and negative values, since noise cannot be modeled exactly.

Key words: Independent component analysis; Multivariate curve resolution; Independence; Sparsity.

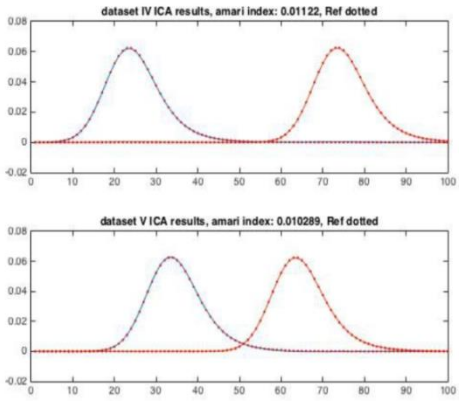
Introduction

Independent Component Analysis (ICA) has been proposed as a powerful signal processing (i.e., blind source separation, BSS) technique during the last years. One of the most important filed of ICA application is for multivariate resolution purposes [1]. From this point of view, ICA has the same goals as multivariate curve resolution (MCR) methods for the analysis of complex multi-component mixtures. However, ICA is based on the statistical independence of the component profiles whereas MCR tries to maximize the explained variance in the data with component profiles fulfilling a set of more natural and physically meaningful constraints. Unfortunately, the pre-requisite of statistical independence severely limits the application of ICA in analytical chemistry [2]. For this purpose, increasing the sparsity of a data set increases the independence of components, which enables the successful application of ICA has been proposed by Windig and Keenan [3]. In this study, the potential of this new development in ICA algorithms is evaluated compared to other ICA algorithms using simulated chromatographic data sets.

Results and Discussion

Different simulated chromatographic data with different number of components (i.e., two and three components), noise and degree of overlap were simulated and used to evaluate the ICA algorithms. The simulated data (individual data matrices and augmented ones) were analyzed by fastICA algorithm using sparsity. The obtained results were compared

with well-known ICA algorithms, such as mean-field ICA (MFICA), stochastic non-negative ICA (SNICA) and mutual-information based on least dependent component analysis (MILCA) [2]. The ultimate objective of all ICA algorithms is statistical independence (or least dependence) of derived components. Figure 1 shows an example of the use of sparsity combined to fastICA for resolution of a two-component chromatographic system with different degree of overlap without application of non-negativity constraint. As it can be seen the ICA solutions are matched with the true profiles (dotted profiles) with acceptable amari indices.



The mutual information (MI) [4] was used to measure the indepence between resolved sources and Amari index was used to calculate the similarity between resolved elution profiles [2]. In addition, to be sure that the ICA solutions are in range of feasible solutions, polygon inflation algorithm using FAC-PACK toolbox [5] was used to calculate the extent of rotational ambiguity in data. It is concluded that increasing the independence of the chemical components by sparsity clearly improves the resolution in terms of chemical meaning using ICA which is really important for noisy data compared to other ICA algorithms. It is important to note that meaningful solutions can be obtained without applying non-negativity constraint which destroys the independence prerequisite (for example in MFICA). Intuitively, it makes sense that complete knowledge about a very sparse component provides relatively little information about the overall state of the system, thus, conforming more closely to the assumptions underlying ICA. Adding zeros increases the sparsity of the data set. When the concentration matrix consists of truly independent components, ICA works on the original data set. In this case, adding zeros will decrease the independence of the chemical components. This also means that for certain applications ICA can be made possible by deleting sparsity.

References

[1] Hyvarinen A., Oja E., Independent component analysis: algorithms and applications, *Neural Networks, 2000:* ***13****: 411-430.*

[2] Parastar H., Jalali-Heravi M., Tauler, R., Is independent component analysis is appropriate for multivariate resolution in analytical chemistry?. *Trends Anal. Chem., 2012:* ***31****: 134-143*

[3] Windig w., Keenan M.R., Homeopathic ICA: A simple approach to expand the use of independent component analysis (ICA), *Chemom. Intell. Lab. Syst., 2015:* ***142****: 54-63.*

[4] Moon Y.I., Rajagopalan B., Lall U., Estimation of mutual information using kernel density estimators, *Phys. Rev. E, 1995:* ***52****: 2318-2321.*

[5] Sawall M., Fischer C., Heller D., and Neymeyr K.. Reduction of the rotational ambiguity of curve resolution technqiues under partial knowledge of the factors. Complementarity and coupling theorems. *J. Chemometrics, 2012:* ***26****: 526–537.*

Quantitative sequence activity relationship of bitter tasting threshold peptides: a comparison study between some two-way and three-way modeling methods

Bahram Hemmateenejad *, Saeed Yousefinejad

Laboratory for Chemometrics and Chemoinformatics, Department of Chemistry, Shiraz University, Shiraz, Iran.

* E-mail address (corresponding author): hemmatb@shirazu.ac.ir

Abstract

According to the nature of peptides as a polymerized structure composed by amino acids (AAs), the local indices of amino acids were arranged as a cube. N-way PLS (N-PLS) and unfolded PLS (U-PLS) as Multi-way calibration methods were used to construct the quantitative sequence acitivity relationship model. The cube of peptide's indices were unfolded and were analyzed by common PLS and MLR method as well. The best model were obtained with two latent variables using Unfolded PLS (R^2cal= 0.895(±0.016 and Q^2cv= 0.837(±0.022)). The external test set was predicted better with U-PLS (R^2p= 0.886(±0.017)) than with bilinear PLS (R2p= 0.809(±0.012)) or MLR (R^2_p= 0.813(±0.014)). The results showed the stability and effectiveness of the used multi-way methods in the QSAM of BTT peptides.

Key words: peptide, QSAM, multilinear PLS; unfolded PLS, multiway calibration

Introduction

Nowadays, quantitative structure–activity relationship (QSAR) modeling is a reliable and practical method in chemometrics for studying the relationship between molecular structures of therapeutic compounds and biological activities. Quantitative sequence-activity modeling (QSAM) is an attractive field which has proposed by Jonsson et al. [1] and employs QSAR strategies to quantify biosequence-activity/function relationship for the peptides and also for proteins and nucleic acids.

Different multivariate calibration methods could be applied to model bioactivity of bio-compounds as a function of their molecular structures [1]. Two-way calibration methods as well as multi-way methods [1,2] like n-way partial least squares (NPLS) was used in QSAR modeling. In QSAR of peptides statistical two-way regression analysis such as like multiple linear regressions (MLR) and partial least squares (PLS) have been widely used to develop models [**Error! Bookmark not defined.**], two other new references]. But, to the best of our knowledge the reports on using muli-way methods in QSAR of peptides are very rare [1]. In this work we used multiway partial least square (N-PLS), unfolded PLS (U-PLS) to model the activity of 48 bitter tasting threshold di-peptides.

Results and discussion

According to the QSAM approach of peptides, AAs could be considered as the base bricks of peptides. So in a series of N-peptide data set, in addition to unfolding the descriptors of the sequences of N-peptide molecules beside each other and obtaining a row vector of descriptors, it is also possible to putting the common descriptors of the sequences of a peptide below each other to result a matrix of descriptors for that peptide. In the first case, the row vector of peptides of a data set could obtain a matrix of data and could be analyzed by first order calibration method. In the second case, the matrix of peptides of a data set could put behind each other to form a cube of data and could be analyzed by second order calibration methods (See Fig. 1).

In this work the possibility of using some calibration methods for three way data (e.g. N-PLS and U-PLS) were tested in QSAM of peptides. The results showed that U-PLS with R^2_{cal} equal to 0.895(±0.016) and Q^2_{cv} equal to 0.837(± 0.022) obtained good and robust models in the case of BBT di-peptides data set. It seems that in the case of more complex peptides with more sequences, multi-way methods could also be applicable. Multiway methods have some advantages like more stability, lower sensitivity to the noise and also could improve the interpretation of the results.

By a brief comparison of calibration on two-way data e.g. MLR and PLS methods and calibration methods on three way data array like N-PLS and U-PLS in modeling of BTT activities, it was observed that N-PLS and U-PLS models were more robust and less sensitive to the molecules in the training and test sets.

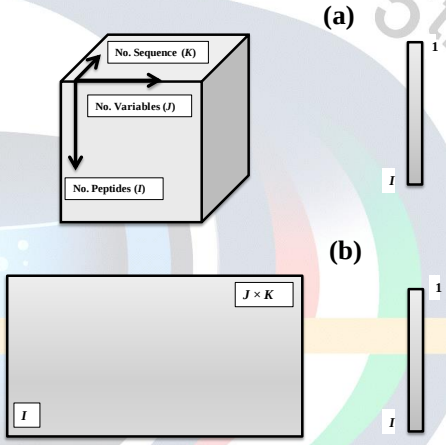


Figure 1. Graphical representation of the structure of independent and dependent arrays used in (a) the three-way and (b) unfolded models.

References

[1] Jonsson J, Norberg T, Carlsson L, Gustafsson C, Wold S, Quantitative sequence-activity models (QSAM) tools for sequence design. Nucleic Acids Res 1993, **21**:733–739.

[2] Ferreira, M. M. CMultivariate QSAR, J. Braz. Chem. Soc., 2002, **13**: 742-753.

[3] Nilsson J., De Jong, S., Smilde A. K, Multiway Calibration in 3D QSAR, J.Chemom., 1997, **11**:511–524.

[4] Eriksson L., Andersson P.L., Johansson E., Tysklind M., Megavariate analysis of environmental QSAR data. Part II – Investigating very complex problem formulations using hierarchical, non-linear and batch-wise extensions of PCA and PLS, Molecular Diversity 2006, **10**: 187–205

[5] Bhonsle J.B., Zi-xuan W., Hirokazu T., Fujii N., Peiper S.C., Trent J.O., A simple, Automated Quasi-4D-QSAR, Quasi-multi Way PLS Approach to Develop Highly Predictive QSAR Models for Highly Flexible CXCR4 Inhibitor Cyclic Pentapeptide Ligands Using Scripted Common Molecular Modeling Tools, QSAR Comb. Sci. 2005, **24**: 620-630.

Influence of bonded-phase density in chiral chromatography via molecular simulation

Elham Aghaee [1]*, Jahan B. Ghasemi [2]

[1]*Chemistry Department, Faculty of Sciences, K. N. Toosi University of Technology, Tehran, Iran*

[2] *Department of Chemistry, Faculty of Science , University of Tehran, Tehran, Iran*

E-mail: elham_aghaee@yahoo.com

Abstract

Molecular dynamics simulations were performed with the Gromacs software package using GROMOS96 force-field to examine the effects of bonding density on molecular structure in chiral chromatography. 1, 2-diphenyl-1,2-ethylenediamine (DPEDA) stationary phase with four different bonding densities (0.26, 0.83, 1.06, and 1.78 µmol/m^2) in contact with an n-hexane /dichloromethane /methanol 75:20:5 (v/v/v) mobile phase were simulated. The simulations indicate that the selector chains become more aligned and interact with themselves as coverage is increased. At lower densities, significant amounts of the silica surface are exposed leading to an enhanced wetting of the stationary phase and increase of interaction between enantiomers and selectors and finally better separation. This approach has successfully predicted reasonable separation factors on this chiral stationary phase.

Key words: Molecular dynamics simulations, Brush-type chiral stationary phases (CSPs), bonding density

Introduction

Chiral separations are now an essential part of the drug development process. Over the years, high performance liquid chromatography using chiral stationary phases (CSPs) has become an important tool for the separation of racemates. Brush-type chiral stationary phases (CSPs) have received considerable interest as a possible basis for new selectors in chiral chromatography [1,2]. 1, 2-diphenyl-1,2-ethylenediamine (DPEDA) was synthesized and introduced by Kotoni et al. in 2013 as a new chiral selector which was attached on the silica gel [3]. The interest in understanding fundamental mechanisms underlying chromatography seriously increased over the past decades [4, 5]. In this study, we employ molecular dynamics simulations to examine the interfacial structure of this chiral selector for separation of enantiomers.

Results and discussion

In order to visualize the interfacial structure of separation process, geometry optimizations of the ligand, enantiomers and solvent molecules were performed for DFT B3LYP/6-31g** level using Gaussian 09 package. The ligands where then distributed evenly with 0.05-1.2 nm spacing to create a surface plane with four different bonding densities. Trimethylsilane groups with 0.08nm spacing were added between the ligands and a layer of silanol groups with 0.01nm spacing was added under ligand layer. The structure of simulation box is shown in Fig 1. To keep the ligands in a layer and to avoid a breakdown of the surface during the simulation, the Cartesian coordinates of the first three atoms of each chain of ligand were restrained by a harmonic potential. Then the box was filled with enantiomers and solvent molecules. The simulation box includes 20 analytes: 10 of each enantiomer which are randomly distributed

over the surface. Molecular dynamics simulations were performed with the Gromacs software package using GROMOS96 force-field. Then 30 ns MD run (NVT type) was performed.

Table 1. Interaction summary for the crab-like CSP extracted from MD snapshot analysis. A. calculated and experimental separation factors on DPEDA surface with bonding densities of 1.06 μmol/m^2. B. the achieved results for lorazepam compound.

A

Compound	α^{calc}	α^{exp} [3]
1,1-Binaphthol	1.72	1.57
Med-13	1.01	1.13
Oxazepam	1.08	1.00
Temazepam	1.37	1.21
Lorazepam	1.20	1.13

B

Selector density μmol/m^2	chirality	Docked percent (2 or more than 2 hydrogen bonds)	Undocked percent (Less than 2 hydrogen bonds)	α
0.26	**R**	48.78	51.22	1.65
	S	36.85	63.15	
0.83	**R**	35.62	64.38	1.49
	S	45.36	54.64	
1.06	**R**	21.59	78.41	1.20
	S	18.64	81.36	
1.78	**R**	6.59	93.41	1.03
	S	6.36	93.64	

The separation factor in chiral chromatography is given by:

$$\alpha = \frac{K_2}{K_1} = \frac{[CSP.A^2][A^1]}{[A^2][CSP.A^1]}$$

where K_2 and K_1 are the equilibrium constants for the more, and the least, retained enantiomers, A^2 and A^1 are the more, and the least retained enantiomers, $CSP.A^1$ and $CSP.A^2$ are for docking enantiomers on the CSP, respectively. This approach has successfully predicted reasonable separation factors on this new crab-like chiral stationary phase.

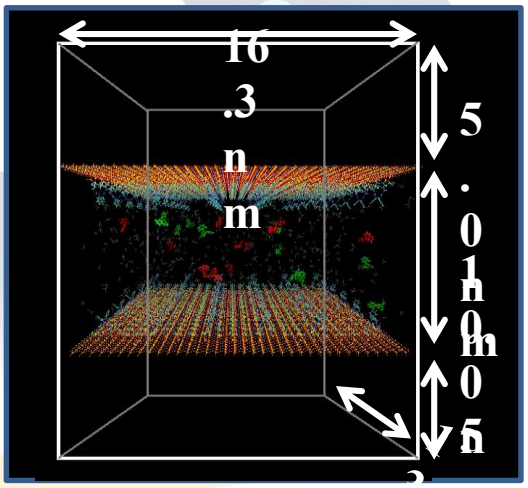


Fig 1. The structure of simulation box

The simulations indicate that the selector chains become more aligned and interact with themselves as coverage is increased. At lower densities, significant amounts of the silica surface are exposed leading to an enhanced wetting of the stationary phase and increase of interaction between enantiomers and selectors and finally increase of separation factor.

References

[1] C.F. Zhao, S. Diemert, N.M. Cann, Journal of Chromatography A, 1216 (2009) 5968–5978.
[2] M. Ashtari, N.M. Cann, Journal of Chromatography A, 1218 (2011) 6331– 6347.
[3] Dorina Kotoni, Claudio Villani , David S. Bell , Donatella Capitani, Pietro Campiglia, Francesco Gasparrini, Journal of Chromatography A, 1297 (2013) 157– 167
[4] Florian Dismer, Juergen Hubbuch, Journal of Chromatography A, 1217 (2010) 1343–1353.
[5] Lin Zhang, Yan Sun, Biochemical Engineering Journal 48 (2010) 408–415.

Present and future of QSAR methodology

M. H. Fatemi

Chemometric Laboratory, Faculty of Chemistry, University of Mazandaran, Babolsar, Iran

Quantitative structure activity relationship (QSAR), and/or quantitative structure–property relationships (QSPRs), are mathematical modeling approaches that attempt to relate the structure-derived features of a compound to its biological or physicochemical activity. The basic strategy of QSAR is to find the optimum quantitative relationship, which can then be used for the prediction of the properties of molecular structures including those unmeasured or even unknown. QSPR became more attractive for chemists with development of new software tools, which allowed them to discover and to understand how molecular structure influences properties, and very importantly, to predict and prepare the optimum structure. However, throughout its entire history it has drawn both praise and criticism concerning its reliability, limitations, successes, and failures. In this discussion, we try to explain (i) the development and evolution of QSAR; (ii) the current trends, unsolved problems, and pressing challenges; and (iii) several novel and emerging applications of QSAR modeling. Throughout this discussion, we provide guidelines for QSAR development, validation, and application, which are summarized in best practices for building rigorously validated and externally predictive QSAR models. Finally a bibliometric analysis based on the Science Citation Index Expanded was conducted to provide insights into the publication performance and research trend of quantitative structure-activity relationship and quantitative structure-property relationship.

Novel feature selection methods and CAIS resolution software

M. Mousavi[1,*], M.H. Nezamabadi-pour[2], B. Mohseni[1], H. Asadollahzadeh[3], E. Pakneeyat[1], T. Baghgoli[1], H. Nikpour[3], N. Peeroznejad[1], A. Nasiri[1], S. Nazari[1], M. Rahmani[1]

[1]*Department of Chemistry, Faculty of Sciences, Shahid Bahonar University of Kerman, Kerman, Iran*
[2]*Department of Electrical Engineering, Shahid Bahonar University of Kerman, Kerman, Iran*
[3]*Department of Chemistry, faculty of Sciences, Azad University, Kerman, Iran.*

*E-mail address:mmousavi@uk.ac.ir

Finding reliable descriptors among a large number of structural parameters that can be linked to the biological or chemical activity in a QSAR model has been a serious challenge and many feature selection algorithms such as genetic algorithm (GA), ant colony optimization (ACO) and particle swarm optimization (PSO) have been used in QSAR studies. The idea behind feature selection is that, only a limited number of structural parameters affected our desired activity. In these algorithms, different criteria are implemented for selecting appropriate sets of descriptors. Exploration and exploitation are the most important factors that control the efficiency and accuracy of an optimization algorithm.

Gravitational search algorithm (GSA) have been established based on the metaphor of gravity and motion laws [1-2]. In this algorithm, the searcher agents are a collection of masses that can determine the position and status of the other masses via gravitational force. Each agent is a solution of the problem and according to its fitness a mass attributes to it. The force acting on an agent from other masses will change the position of the agent. Agents with higher performance have greater masses, so that an agent with a heavier mass has a larger effective attraction radius and hence greater intensity of attraction. Consequently, agents tend to move toward the best agent (biggest mass) resembling an optimum.

The quantum-behaved particle swarms optimization (QPSO) algorithm [3] considers the search space as a system with quantum particles by inspiration of Heisenberg's uncertainty principle and then scans it. In fact, this algorithm is probability-based version of PSO algorithm in which particles moves in quantum manner instead of Newtonian mode. A potential well in each dimension

attracts the particles based on their fitness. Particles are transfered to their new positions after exchanging their information with each other.

We have recently developed a new software, called CAIS (Chemometrics Analyzer Interface System), to carry out MCR-ALS and PARAFAC calculations on two-way chromatographic data. The CAIS software is a graphical interface to use MATLAB functions that wrapped in DLL files by MATLAB Builder. The calculation steps and data transformation are inspired from MCRC software [4]. The CAIS software is able to open ASCII format chromatographic data and resolve them, easily. The CAIS output is a TXT file for each identified component, which can be used directly by NIST MS library.

References

[1] E. Rashedi, H. Nezamabadi-pour, S. Saryazdi, BGSA: binary gravitational search algorithm, Nat. Comput. 2010; 9: 727–745.

[2] B. Mohseni-bababdani, M. Mousavi, Gravitational search algorithm: A new feature selection method for QSAR study of anticancer potency of imidazo [4,5-b] Pyridine derivatives. Chemometrics and Intelligent Laboratory Systems.2013; 122: 1-11.

[3] J. Sun, B. Feng, W.B. Xu, Particle swarm optimization with particles having quantum behavior, in: IEEE Congress on Evolutionary Computation. 2004; pp. 325–331.

[4] M. Jalali-Heravi, H. Parastar, M. Kamalzadeh, R. Tauler, J. Jaumot: MCRC software: A tool for chemometric analysis of two-way chromatogtaphic data. Chemometrics and Intelligent Laboratory Systems.2010; 104: 155-171.

Investigation of a sewage treatment plant for removal of selective pharmaceutical and personal care products using HPLC-DAD and second-order calibration

M. Rashvand[a], M. Vosough [*, a], K. Kargosha[a]

a Department of Green chemistry, Chemistry and Chemical Engineering Research Center of Iran

E-mail address (corresponding author): Vosough@ccerci.ac.ir

Abstract

This study focus on the quantification of selected pharmaceutical and personal care products (PPCPs), Paracetamol, Methyl paraben, Carbamazepin, Naproxen, Buthyl paraben and Diclofenac in influent and effluent of a sewage treatment plant (STP) by application of multivariate curve resolution - alternating least squares (MCR-ALS). The samples were preconsentrated on C-18 cartriges and were analyzed using high performance liquid chromatography with diode-array detection in optimized situation. The qualitative and quantitative results showed the appropriate operation of MCR-ALS algorithm for solving co-elution of analytes with matrix interferences. The recoveries and figures of merits values of the optimized procedure depicted acceptable results for proposed method.

Key words: Pharmaceutical and personal care product, Wastewater, Second order calibration, solid phase extraction

Introduction

Pharmaceutical and personal care products (PPCPS) are widely disseminated into the environment [1]. The trace amounts of these organic compounds, as a class of emerging contaminants (ECs), are released continuously into aquatic environment through different routes such as sewages and domestic wastes, disposing in landfills, storm water overflow from residual sources and recharge of ground water by tertiary treated wastewater [2]. In spite of substantial removal during conventional treatment processes, they can be detected in the receiving water samples at low concentrations [3]. Among pharmaceuticals, NSAIDS, because of their long-term therapy and severe side effects, and antiepileptics, due to their extensive usage in neuropathic pains, catch more concern for investigation. In additions, parabens, as ingredients in personal care products, are employed extensively as the antimicrobial preservatives in cosmetic, pharmaceuticals and food products. It has been proved that parabens show estrogenic activity and easily react with free chlorine after mixing with chlorinated tap water, so mono- and di-chlorinated derivatives can be produced [4]. The main goal of this study, is quantifying the selected PPCPs and investigating the turn over of STP cleaning through a simple pretreatment and analysis procedure benefiting second-order calibration algorithm.

Results and discussion

The pharmaceutical standards were kindly provided by Tehran Daroo and the parabens standards were purchased from Aldrich. For pretreatment, after preparation the sorbent, 400 mL of samples pass through the cartridge by 5 mL/min and 4 mL acetonitril used for cleaning the analytes by 0.5 mL/min. Then, after evaporating the solvent and adding the injection phase, separations by HPLC-DAD were done. The analytes were separated in less than 5.5 min during a gradiant programming consisting of a binary mixture of solvents: (A) water (containing 0.3% acid acetic) and (B) acetonitrile as amobile phase. The method was validated via univariate methodology and the figures of merit are summarized in table1.

Table 1. Analytical figures of merit for the determination of the six antibiotics by multivariate calibration

Analyte	R^2	RSD %[a]	LODs(μgL^{-1})	LOQs(μgL^{-1})
		Repeatability		
PARAC	0.995	4.4	1.8	6
MET-P	0.983	6.52	0.08	0.27
CBZ	0.991	5.7	0.17	0.60
NAP	0.987	5.9	0.19	0.63
BUT-P	0.985	2.86	0.1	0.35
DIC	0.982	6.3	0.04	0.13

Regarding validation of modeling process, the influent and effluent of selected STP were spiked at three levels of concentration. Six regions were chosen for processing and the exportaed data was modeled by MCR-ALS for chromatographic subsets. The resolved spectra showed complete compatibility with the pure's one. In addition, the predicted chromatograms clearly depicted the co-elution of interferences with each analytes. Furthermore, good agreement between predicted and spiked values in all validation sets, represent acceptable modeling.

The comparison between the concentration of analytes in influent and effluent shows complete cleaning for methyl paraben and carbamazepine. In the case of naproxen and diclofenac, the cleaning process was successful up to 80% and for the others (buthyl paraben and paracetamol), just 50% were purified in this STP.

References

[1] Lapworth DJ, Baran N, Stuart ME,Ward RS, Emerging organic contaminants in groundwater: A review of sources, fate and occurrence, Environ Pollut. 2012; **163**; 287–303.

[2] C.G. Daughton, T.A. Ternes, Environ. Pharmaceuticals and personal care products in the environment: agents of subtle change?, Health Perspect. 1999; **107**; 907-938.

[3] Lee, H.B., Peart, T.E., Svoboda, M.L., Determination of endocrine-disrupting phenols, acidic pharmaceuticals, and personal-care products in sewage by solid-phase extraction and gas chromatography-mass spectrometry, Journal of Chromatography A. 2005; **1094**; 122-129.

[4] Villaverde-de-Sáa, E., González-Mariñ o, I., Quintana, J.B., Rodil, R., Rodríguez, I., Cela, R., In-sample acetylation-non-porous membrane-assisted liquid-liquid extraction for the determination of parabens and triclosan in water samples, Analytical and Bioanalytical Chemistry, 2010; **397**; 2559-2568.

Design and construct a small fluoremeter with a smartphone as the detector associated with image processing

Morteza Sadeghiye[a], Zeinab Saberi Dehkordi[a], Taghi Khayamian[a,*]

[a]*Deparment of chemistry, Isfahan Universiy of Technology, Isfahan, Iran*.

E-mail address: taghi@cc.iut.ac.ir

Abstract

A novel and simple fluoremeter was designed. The home-made system composed of a dark box, a laser, smartphone and image processing which is a written algorithm in Matlab. The data of each concentration are three matrices R, G, B and each of the matrix or combination of them were examined to find out the best response, considered in calibration cure. Linear calibration curve for Rhodamine was found 10^{-10} to 10^{-7} mol l^{-1} with detection limits $1.23*10^{-11}$ mol l^{-1}. These results were compared to Shimatzu fluorescence spectrometer that concentration range and detection limits were found to be 10^{-11} to $2*10^{-6}$ mol l^{-1} with $2.13*10^{-12}$ mol l^{-1}.

Furthermore, this fluoremeter was also applied of analysis simultaneously of two fluorescent compounds in a tablet. The Prineciple Component Regression (PCR) was used for deconvolotion of the overlapping peaks.

Key words: Image processing, Fluoremeter, Rhodamine, PCR

1. Introduction

The new fluoremeter composed of a dark box, a laser and a samartphone. The recorded fluoresce images were analyzed using image processing. The images in Matlab are converted to color values that three matrices R, G, B are output. These matrices or combination of them were examined for finding out the best response [1].

Rhodamine B dye was chosen for testing device that can be detected easly and inexpensively. It is used in biotechnology application. B_2 and B_9 vitamines were chosen as mixtures for analysis simultaneously of two compounds. B vitamins exist in an extended range of food and tablets.

2. Results and discussion

The linear calibration curve of analyzed materials was found by new fluoremeter and results were campared by Shimatzu fluorescence spectrometer. For Rhodamine, the linear calibration curve was achieved 10^{-10} to 10^{-7} mol l^{-1} with the detection limits $1.23*10^{-11}$ for new fluoremeter also linear concentration range and detection limits were found to be 10^{-11} to $2*10^{-6}$ mol l^{-1} with $2.13*10^{-12}$ mol $l^{-1}$by Shimatzu. The linear calibration curve and detection limits of B_2 vitamin were found by new fluoremeter and Shimatzu .02 mg l^{-1} to 0.1 mg l^{-1} and 0.001 mg l^{-1} to 1 mg l-1 with 0.0046 mg l^{-1} and 0.00023 mg l^{-1} respectively. Results of B_9 vitamin were obtained 10 mg l^{-1} to 100 mg l^{-1} and 0.5 mg l^{-1} to 100 mg l^{-1} with 1.89 mg l^{-1} and 0.043 mg l^{-1} for new and Shimatzu fluoremeter respectively.

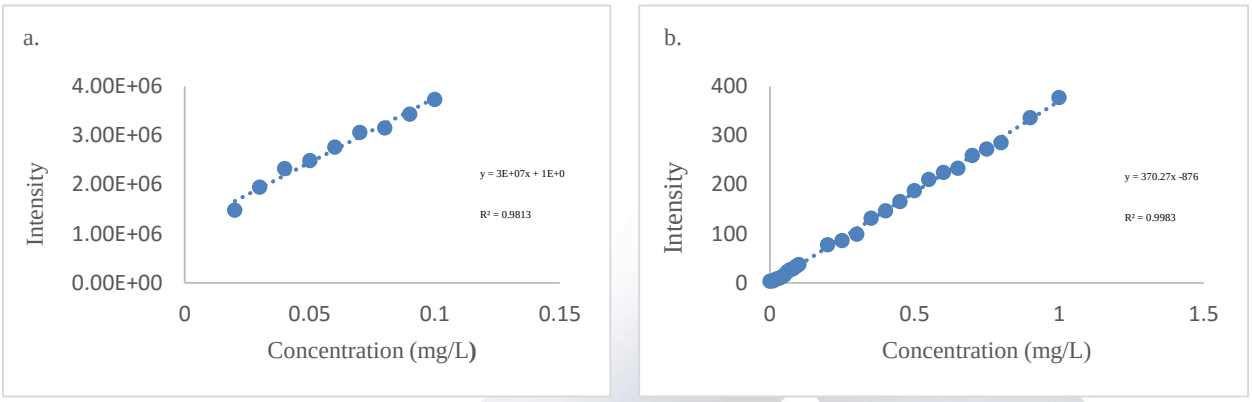


Figure 1. Calibration curve of B_2 vitamin. New fluoremeter (a) and Shimatzu spectrofluorimeter (b)

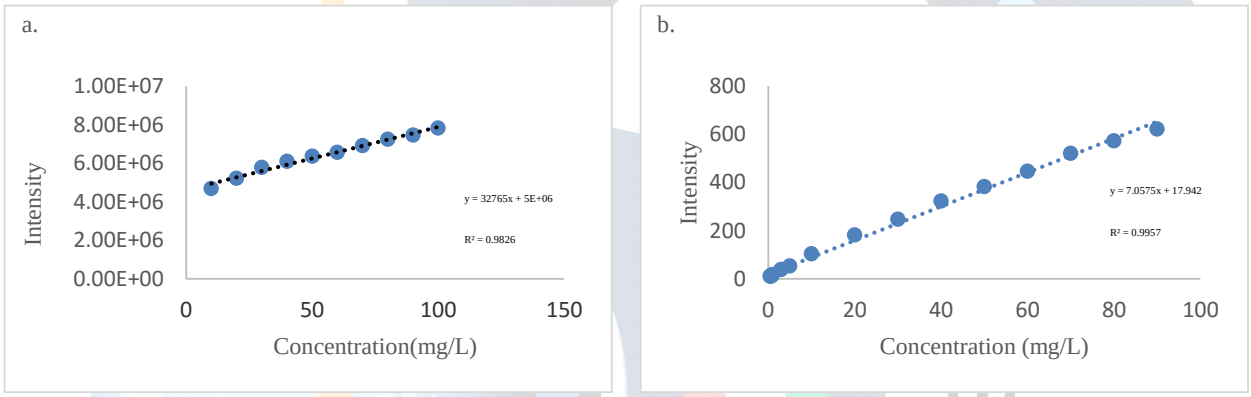


Figure 2. Calibration curve of B_9 vitamin. New fluoremeter (a) and Shimatzu spectrofluorimeter (b)

References

[1] T. Hayakawa, M. Hirai, An Assay of Ganglioside Using FluorescenceImage Analysis on a Thin-Layer Chromatography Plate; Anal. Chem.2003, 75, 6728-6731

Defining the Applicability Domain for Classification Models by Hedging Predictions

K. Baumann, M. Mathea, W. Klingspohn

Institute of Medicinal and Pharmaceutical Chemistry, University of Technology Braunschweig, Beethovenstrasse 55, 38106 Braunschweig, Germany
k.baumann@tu-braunschweig.de

Classification rules are often used in chemoinformatics to predict categorical properties of drug candidates related to bioactivity from explanatory variables, which encode the respective molecular structures (i.e. molecular descriptors). The predictive performance of the developed classifier is assessed by testing the classifier on unseen objects for which the class labels are known. This allows estimating the prediction error, which characterizes the average error for the unseen objects. The classifier is expected to work with a similar error rate for future objects. However, it is never possible to train the classifier on all possible object varieties (e.g. chemotypes) that are fed into the classifier during its productive phase. It makes intuitively sense that the error rate for future objects that are not well embedded in the training data set may be larger than for future objects that are similar to the training data. Hence, to avoid predictions with an unduly large error probability, the domain the classifier is applied to should be restricted to the domain covered by the training set objects. This latter domain is commonly referred to as applicability domain in chemoinformatics. Conceptually, the applicability domain defines the region in space where the “normal” objects are located. Defining the border of the applicability domain may then be viewed as detecting anomalous or novel objects or as detecting outliers.[1] Currently two different types of measures are in use. The first one defines the applicability domain solely in terms of the molecular descriptor space, which is referred to as novelty detection. The second type defines the applicability domain in terms of the expected reliability of the predictions which is referred to as confidence estimation or hedging predictions. Both types are systematically differentiated here and the most popular measures are presented. It will be shown that all common chemoinformatic classifiers have built-in confidence scores to hedge predictions. Surprisingly, these are rarely used for defining the applicability domain. Since confidence estimation uses information of the class labels for computing the confidence scores, it is more efficient in reducing the error rate than novelty detection, which solely uses the information of the explanatory variables. A large benchmark study was carried out to show that confidence estimation consistently outperforms novelty detection in reducing the error rate of future predictions.

References

[1] V. Chandola, A. Banerjee, V. Kumar, *ACM Comput. Surv.*, **41** (2009) 1-58.

Geochemometrics

Somayeh Veyseh [1,2], Ali Niazi *[1,2]

[1]*Department of Chemistry, Faculty of Science, Arak Branch, Islamic Azad University, Arak, Iran*
[2]*Applied Geological Research Center, Geological Survey of Iran*

E-mail address (corresponding author): ali.niazi@gmail.com & a-niazi@iau-arak.ac.ir

Abstract

Geochemometrics is defined as the application of statistical, mathematical, graphic and computer sciences for extraction and better understanding of geological science data. Chemometrics was first introduced in 1972 by Wold and several articles and books have been presented in the last decades. And always considering it as operational chemometrics methods in different data has attracted great attention. Therefore according to the applied studies for utilizing chemometrics methods in geological science data, Geochemometrics term was first introduced as an interbranch scince. Achivements and abilities of geochemometrics can be used a useful tools in elimination geological challenges. As an example, in the field of exploring inorganic compounds, increasing the number of exploratory samples and sampling times have made researchers be confused. Therefore according to the obtained data from different advanced analytical instruments, it is a need to extractuseful information from these data.

So, for this purpose Geochemometrics will be effective for modeling and classification of analytical data, and the most usefuldata will be obtained by spanding the least cost and time. Geochemometrics can have wide application in the fields of quality control, determination the type and the origins of the samples modeling, instrumental calibration in order to analyze the geology, and medical and naval geology. As an example, for determining the jewelry to be origin and distinguishing between them as beingoriginal or fake, evaluating the age of geological samples, modeling and qualitative analyze of underground waters, hydrogeochemical water analyzing, processing satellite photos that are captured from earth surface using principal component analysis (PCA), two and three multivariate calibration and artificial intelegence can be refered.

In the recent studies for classification and pattern recognition for toxic metals, radioisotops, fundamentals and rare earth elements, proper and good data are obtained by instrumental analysis of ICP-MS and PCA in different regions of Iran, according to the performed classification, that one of the applications of this study in the field of medical geology, anomaly investigation of metals and their interaction effects in purposed data.

Key words: Geochemometrics, Chemometrics, Classification, Principal Componet Analysis

References

[1] Verma S.P, *Revsita Mexicana de Ciencias Geologicas*, 2012; **29**: 276-298.

[2] Zarate D, Santoyo E, Guevara M, *Applied Thermal Engineering*, 2015; **75**: 1244-1261.

[3] Veyseh S, Niazi A, *Talanta*, 2016; **147**: 117-123.

Metabolite profiling and chemometric classification of two varieties of *Ocimum basilicum* medicinal plant

Parisa Izadiyan[1,3], Bahram Hemmateenejad[1, 2*]

[1] *Medicinal and Natural Products Chemistry Research Center, Shiraz University of Medical Sciences, Shiraz, Iran*
[2] *Chemistry Department, Shiraz University. Shiraz, Iran*
[3]*Student Research Committee, Shiraz University of Medical Sciences, Shiraz, Iran*

E-mail address (corresponding author): hemmatb@sums.ac.ir

Abstract

In this work, metabolite fingerprints of twenty four accessions of Ocimum basilicum L. (basil) belonging to two different varieties (var. purpurascens and var. basilicum L) grown in Iran were assessed by HPLC-DAD analysis and chemometric methods. Qualitative analysis of the major phenolic compounds was also conducted using high-performance liquid chromatography coupled to mass spectrometry (HPLC-MS). Evaluation of the phenolic profile of basil varieties revealed that two phenolic acids of rosmarinic acid and chicoric acid are the major compounds found in Iranian basil. Using their HPLC fingerprints the samples were classified according to their varieties. Different chemometric algorithms were examined to differentiate basil varieties. According to the obtained results, PLS-DA algorithm was the most promising chemometric algorithm for classification of basil accessions to the variety level.

Key words: Metabolite fingerprinting, HPLC-DAD, PLS-DA, *Ocimum basilicum*

Introduction

The medicinal plant of *Ocimum basilicum* from genus *Ocimum*, a member of the *Lamiaceae* family, contains 200 species of herbs and shrubs (1). Chemical constituents possessing antioxidant properties in this plant can play an important role in the prevention of various degenerative diseases (2). Although, basil is one of the most important and effective edible medicinal herbs in Iran, few studies have been carried out regarding identification and chemical classification of this plant (3). In this work, high performance liquid chromatography coupled with diode array detector (HPLC-DAD) fingerprinting analysis was carried out for evaluation of the phenolic compounds of two different basil varieties. The aim of this work was firstly to investigate the chemical composition of two varieties of Iranian basil and secondly to compare different discrimination and classification techniques to distinguish the two basil varieties.

Results and discussion

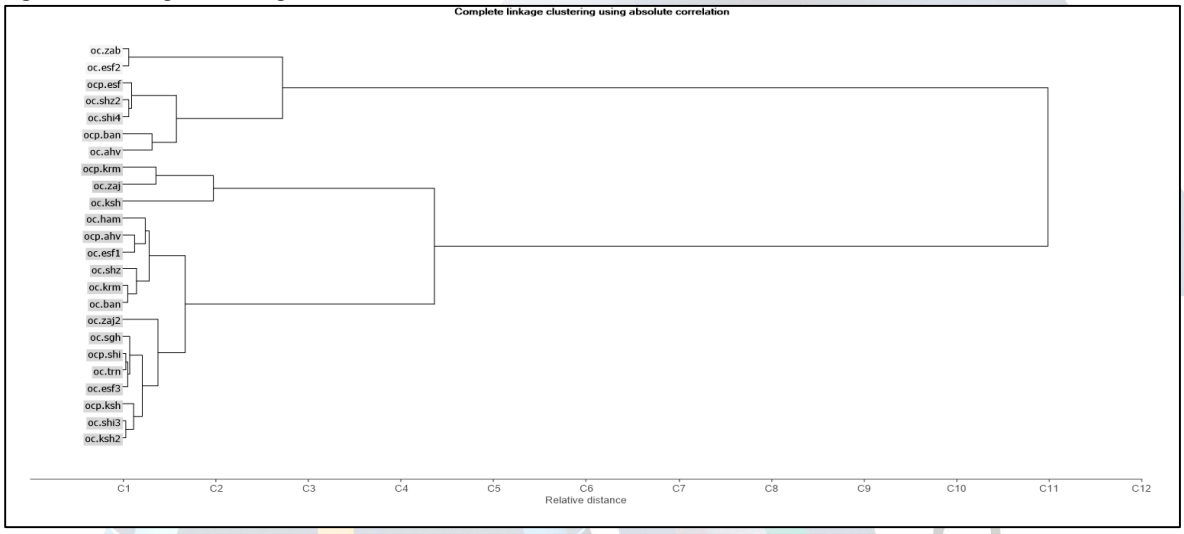
Ocimum basilicum accessions were collected during different seasons from different geographical origins including northern, southern, eastern and western cities of Iran. Fingerprints of 24 samples from the 2 varieties were developed using high-performance liquid chromatography with diode array detection (HPLC-DAD).

At first, fingerprint data pretreatment was carried out. Different preprocessing methods were examined such as: standard normal variate (SNV), multiple scatter correction (MSC) and mean centering. As the next step, preprocessed data were

subjected to two exploratory data analysis using Principal Component Analysis (PCA) and hierarchical cluster analysis (HCA) which revealed a clear but not precise discrimination pattern among samples (Figure 1).

Classification models were built by means of soft independent modeling of class analogy (SIMCA), Support Vector Machines (SVM) and Partial Least Square Discriminant Analysis (PLS-DA). Cross validation was also used for model validation. After performing SNV the best discrimination among two varieties was obtained using PLS-DA classification model with model and cross validated non-error rate of 1 and 0.95, respectively. Three samples were not assigned.

Figure 1. Denderogram resulting from HCA



Although, classification of plants according to their genera and species using their chromatography fingerprints bas been successfully demonasterated in earlier studies (4) but in this study it was shown that chemometrics methods could be considered promising regarding classification of lower plant class orders such as “variety” .

References

1. Song H, Kumar P, Arivazhagan G, Lee S-I, Yoon H-M, Kim I-H, et al. Antioxidant property of leaves and calluses extracts of in-vitro grown 5 different Ocimum species. Journal of Plant Biotechnology. 2012;39(3):146-53.
2. Scalbert A, Johnson IT, Saltmarsh M. Polyphenols: antioxidants and beyond. The American journal of clinical nutrition. 2005;81(1):215S-7S.
3. Javanmardi J, Khalighi A, Kashi A, Bais H, Vivanco J. Chemical characterization of basil (Ocimum basilicum L.) found in local accessions and used in traditional medicines in Iran. J Agric Food Chem. 2002;50(21):5878-83.
4. Viaene J, Goodarzi M, Dejaegher B, Tistaert C, Hoang Le Tuan A, Nguyen Hoai N, et al. Discrimination and classification techniques applied on Mallotus and Phyllanthus high performance liquid chromatography fingerprints. Anal Chim Acta. 2015;877:41-50.

Analytical view on uniqueness

Róbert Rajkó[1,*], Hamid Abdollahi[2] and Samira Beyramysoltan[1,2]

[1] *Institute of Process Engineering, Faculty of Engineering, University of Szeged, H-6725 Szeged, Moszkvai krt. 5-7, Hungary*

[2] *Department of Chemistry, Institute for Advanced Studies in Basic Sciences, P.O. Box 45195-1159, Zanjan, Iran*

E-mail address: rajko@sol.cc.u-szeged.hu

Abstract

Multivariate curve resolution methods, frequently used in analyzing bilinear data sets, result in ambiguous decomposition in general. Implementing the adequate constraints may lead to reduce the so-called rotational ambiguity drastically, and in the most favorable cases to the unique solution. Several studies on exploring the uniqueness of the bilinear non-negatively constrained multivariate curve resolution methods have been made in the literature. More in this study, a general procedure is suggested for detection of uniquely recovered profile(s) on the basis of data set structure in the abstract space and a new term, i.e., data-based uniqueness is defined practically and investigated in details. Close inspection of Borgen plots of these data sets leads to realize the comprehensive information of local rank, and these argumentations furnish a basis for data-based uniqueness theorem. The reported phenomenon and its exploration is a new stage in understanding and describing the bilinear (matrix-type) chemical data in general.

Keywords: bilinear data set; uniqueness; rotational ambiguity; Borgen plot; self-modeling/multivariate curve resolution.

Introduction

Self-modeling and multivariate curve resolution (SMCR and MCR) techniques focus on decomposing the unresolved mixtures to profiles of components which are interpretable in physical and chemical terms. In general, the estimated profiles are not unique for a two-way problem because of the permanent presence of rotational ambiguity [1]. Rotational ambiguity complicates the interpretation of the profiles to discover the specific information of the system. Ahmadi and Abdollahi [2] warned about the accuracy of chemical quantitative analysis that is affected by rotational ambiguity. Many efforts have been taken to reduce the rotational ambiguity applying useful knowledge of the systems as constraints during unraveling factors. Certainly, enough information of the system aids to determine the true underlying causes of data variation. The constraints such as selectivity [3] and local rank [4] may have drastic effects on the bands of feasible solutions and in favorite cases feasible solutions turn into unique.

Results and discussion

Close inspection of Borgen plots of these data sets leads to realize the comprehensive information of local rank, and these reasoning furnish a basis for data-based uniqueness theorem. The reported phenomenon and its exploration is a new stage (it can be said fundament) on understanding and on describing the bilinear (matrix-type) analytical data in

general. Data-based uniqueness condition was defined as the following theorem in two equivalent forms in general, i.e., the theorem is formulated in such way that is independent of the number of the components.

***Data-based uniqueness theorem*:**

Form 1: If the Borgen plot (or its generalization for arbitrary dimensions) contains a point belonging to the coincidence of vertices of the inner and outer polygon (inner and outer polyhedrons), and, as a result, this point will be the matching vertices of the two Borgen triangles (all Borgen simplices), then the point signifies the unique solution for the component.

Form 2: If a component has a selective window in one direction/way/mode and considering the other direction/way/mode, this component has a sub-window in which the contribution of this component is zero while all interfering compounds appear inside this window, then the component can be uniquely decomposed in this second direction/way/mode.

Figure 1. An illustration of abstract space (derived by Borgen method) of data set (shown in Fig. 2) when a unique profile occurred. a) U-space (abstract concentration space). b) V-space (abstract spectral space).

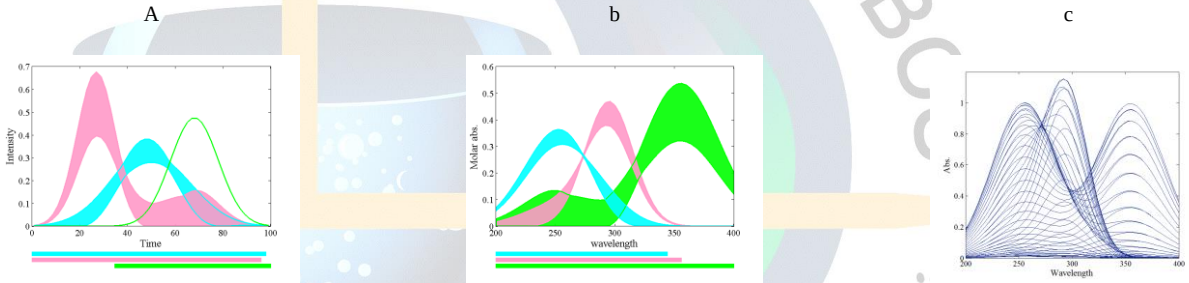


Figure 2. Simulated data set for illustrations of data-based uniqueness. a) True concentration profiles (**C**). b) Spectral bands (**S**). c) The generated data sets (**R**) of multiplication of **C** and **S**. The horizontal lines below the plots having the same color as the bands of the components display the presence window of the components.

Two forms of theorem were illustrated schematicly in abstract spaces (Fig. 1a, 1b) and feasible bands (Fig. 2a, 2b) of a HPLC-DAD data set as an example (Fig. 2c). Conditions in *form 1* and *2* of theorem can be seen in abstract spaces and feasible bands of data set, respectively.

Our proposed detection tool is restricted to three-component systems because of the visual limitations of Borgen plot, but the theorem is general for systems with more than three components.

References:

[1] Tauler R, Smilde A.K, Kowalski B.R. Selectivity, local rank, three-way data analysis and ambiguity in multivariate curve resolution. *J. Chemom.* 1995; **9**: 31-58.
[2] Ahmadi G, Abdollahi H. A systematic study on the accuracy of chemical quantitative analysis using soft modeling methods. *Chemom. Intell. Lab. Syst.* 2013; **120**: 59-70.
[3] Omidikia N, Abdollahi H, Kompany-Zareh M. On uniqueness and selectivity in three-component parallel factor analysis. *Anal. Chim. Acta* 2013; **782**: 12-20.
[4] Maeder M. Evolving factor analysis for the resolution of overlapping chromatographic peaks. *Anal. Chem.* 1987; **59**: 527-530.

Multi-way assisted chromatographic methodologies for analysing highly complex samples

Maryam Vosough

Chemistry and Chemical Engineering Research Center of Iran, P.O. Box 14335-186 Tehran, Iran

E-mail address: vosough@ccerci.ac.ir

Abstract

With the increase in the complexity of analytical objects, sophistication in analytical instrumentation has also been enhanced. Application of different high-order calibration methods for handling multi-dimensional chromatographic data is considered as an efficient strategy to extract qualitative and quantitative information from complex analytical data. In recent years, our group working in Chemistry and Chemical Engineering Research Center of Iran has carried out scientific research for development and validation of different analytical systems for multi-target determination in highly complex samples. Most of the developed analytical strategies in the mentioned works have been conducted to achieve a short run time and simplified chromatographic and extraction conditions regardless of the changes in the sample matrices. In this presentaion, select works in the context of second-order calibrations in various fields, such as environmental samples, biological fluids, food, illicit drugs and cosmetics are highlighted.

Keywords: Multi-way methods, Second-order calibration, Hyphenated chromatography, Complex samples

Introduction

Hyphenated chromatographic systems such as HPLC-DAD, GC-MS and LC-MS are being extensively used for detection, quantification of compounds in samples such as biological, environmental and food samples. However, these samples are typically have a complex nature containing baseline drift, overlapping and unwanted peaks, peak shifts and low signal-to-noise ratios. The need for reducing the analytical time and cost has added more analytical challenges for complex samples, too. Matrix effect as the most important problem for analysis of these samples has been faced when the sample preparation steps are not selective enough for target analytes and also the resolution of the eluting compounds is not perfect, too [1]. However, these chromatographic systems have provided the possibility of employing high-order calibration methods for handling unknown interferences by exploiting second-order advantage.

Multiway calibration has become an important frontier in chemometric research [2-3]. There are different multivariate algorithms that benefit from the second-order advantage and can cope-with the three-way hyphenated chromatographic data, such as direct trilinear decomposition (DTLD), parallel factor analysis (PARAFAC), PARAFAC2, alternating trilinear decomposition (ATLD), self-weighted alternating trilinear decomposition (SWATLD), alternating penalty trilinear decomposition (APTLD), multivariate curve resolution alternating least squares (MCR-ALS), bilinear least squares (BLLS), and also unfolded partial least squares (U-PLS) and multi-way PLS (N-PLS), with residual bilinearization (RBL). With the increasing capabilities of higher-order instruments, such as LC×LC–DAD and LC×LC–MS, it is necessary to develop higher-order data analysis methods. So, several algorithms

for analyzing four-way data have been recently proposed, such as four-way PLS, four-way PARAFAC, U-PLS/residual trilinearization (U-PLS/RTL) and etc. [4].

Results and discussion

Many of the mentioned second-order calibration methods have found real applications in different sample matrices. Our proposed methodologies have been employed to enhance the selectivity and attain predicted concentration of analyte(s) of interest free from interference of potential interfering matrix, fully exploiting the "second-order advantage". (1) ***Food samples.*** The applications in this field cover the analysis of mono and poly unsatuarated fatty acids (PUFAs) using GC-MS and aflatoxins in different pistachio samples. (2) ***Environmental samples.*** In this field, the emerging pollutants such as UV-filters, parabens, antibiotics, illicit drugs and NSAIDs in river water, effeluent and influent waste water sample, hospital wastewater, sediments and sewage sludge have been analysed using different extraction process or direct HPLC injection. These samples are generally complex and the number of analytes to be determined is usually high. (3) ***Biological fluids.*** In the analysis of biological fluids, matrix-free analysis of anti-epileptic drugs, benzodiazepines and immunusuppresants have been performed using HPLC-DAD after a protein precipitation or a simple solvent extraction step. The selected analytes are mainly organic contaminants detrimental to human health. The methods of choice (on raw or pre-processed data) were MCR/ALS, U-PLS/RBL, PARAFAC, PARAFAC2, ATLD and SWATLD. (4) ***Cosmetics.*** In this field, combination of HPLC-DAD with ATLD and MCR/ALS was proposed for analysis of UV-filters and parabens in more than 50 sunscreen samples.

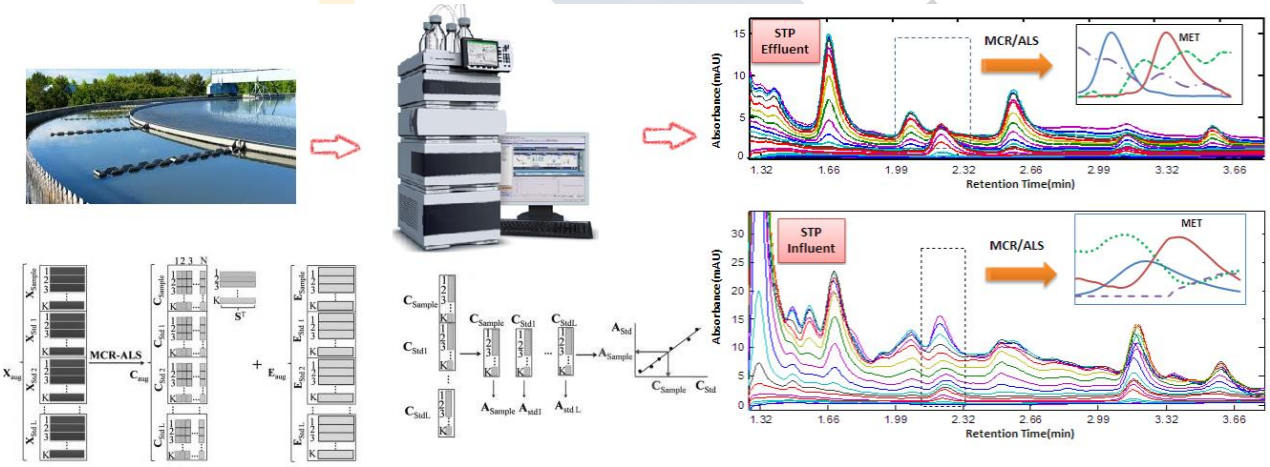


Fig. 1. Representation of the second-order calibration methodology for LC–DAD data obtained from wastewater samples.

References

[1] M.M. Galera, M.D. García, H. C. Goicoechea, The application to wastewaters of chemometric approaches to handling problems of highly complex matrices TrAC, Trends Anal. Chem. 26 (2007) 1032-1041.

[2] S.D. Brown, R. Tauler, B. Walzak, Comprehensive Chemometrics: Chemical and Biochemical Data Analysis. Elsevier Science Ltd.: Amsterdam, 2009.

[3] G.M. Escandar, H.C. Goicoechea, A.M. de la Peña, A.C. Olivieri, Second- and higher-order data generation and calibration: A tutorial, Anal. Chim. Acta 806 (2014) 8-26.

Introducing of invasive weed optimization as a new variable selection method

Saheleh Sheykhizadeh, Abdolhossein Naseri*

Department of Analytical Chemistry, Faculty of Chemistry, University of Tabriz, Tabriz, Iran

E-mail address: a_naseri@tabrizu.ac.ir

Abstract

Variable selection plays an important role in classification or multivariate calibration. The variable selection methods aim is to choose variables, from a large number of available, those relevant to the estimation of analyte concentrations or to achieve to a better classification results. Up to now, many variable selection techniques have been developed. Among all different types of variable selection techniques, those based on swarm intelligence optimization methodologies are more interesting since they are usually simulated based on animal and insect life behavior. In this work a new variable selection algorithm is described, based on invasive weed optimization (IWO). IWO is a bio-spired metaheuristic that mimics the ecological behavior of weeds in colonizing and finding suitable place for growth and reproduction which is shown to be very robust and adaptive to changes in the environment.Thus, capturing their properties would lead to a powerful optimization algorithm. In this paper we wish to report the first application of IWO to variable selection in different experimental data sets including FTIR and NIR data.

Key words: Invasive Weed Optimization; Feature selection; Metaheuristics; Swarm intelligence

Introduction

One of the greatest problems in classification and multivariate calibration is to select the subset of variables that produces the best result, beacause it is clear that not all the features that are stored in the resulting data set are necessary or sufficient to learn the concept of interest There can be many reasons for selecting only a subset of the variables instead of the whole set of candidate variables, obtain (a) improvement of the model predictions, (b) a better interpretation or (c) lower measurement costs [1]. So many different procedures for feature selection are proposed. One type of feature selection methods is based of heuristic techniques, A random search generates random subsets within the search space, which can find good solutions (although they cannot guarantee the optimum) in a reasonable amount of time [2]. Different metaheuristic strategies such as genetic algorithms [3] and swarm algorithms have been used to the variable selection problem. Swarm intelligence (SI) is briefly defined as the collective behaviour of decentralized and self-organized swarms include Ant Colony Optimization (ACO) [4], Particle Swarm Optimization (PSO) [5], Artificial Bee Colony [6] and etc.

Invasive weed optimization (IWO) as a new general purpose swarm intelligence meta-heuristic algorithm was proposed in 2006 by mehrabian and lucas [7]. A weed is any plant growing where it is not wanted. Weeds have shown very robust and adaptive nature which turns them to undesirable plants in agriculture. A common belief in agronomy is that “The Weeds Always Win”. The harder people try, the better they get [7]. It is tried to mimic robustness, adaptation and

randomness of colonizing weeds in a simple but effective optimizing algorithm. To simulate the colonizing behavior of weeds some basic properties of the process is considered below [8]:1) A finite number of seeds are being spread out over the search area. 2) Every seed grows to a flowering plant and produces seeds depending on its fitness. 3) The produced seeds are being randomly dispersed over the search area and grow to new plants. 4) This process continues until maximum number of plants is reached; now only the plants with lower fitness can survive and produce seeds, others are being eliminated. The process continues until maximum number of iterations is reached and hopefully the plant with the best fitness is closest to the optimal solution.

To the best of our knowledge, the application of IWO to the variable selection problem in analytical chemistry is an unexplored field of research. Thus, in this paper, we wish to report the first application of IWO to variable selection in different experimental data sets including FTIR and NIR data.

References

[1] Andersen, CM, Bro R. Variable selection in regression—a tutorial. *Journal of Chemometrics*, 2010; **24**: 728-737.

[2] Yusta, SC, Bonrostro JP, Letamendia LN. A new variable selection method for classification. 2007.

[3] Yang, J, Honavar V. Feature subset selection using a genetic algorithm, *Feature extraction, construction and selection*. 1998; Springer: 117-136.

[4] Al-Ani, A. Ant Colony Optimization for Feature Subset Selection. in WEC (2). 2005; Citeseer: 35-38.

[5] Xue, B, Zhang M, Browne WN. Particle swarm optimization for feature selection in classification: A multi-objective approach. *Cybernetics, IEEE Transactions on*, 2013; **43**: 1656-1671.

[6] Schiezaro, M, Pedrini H. Data feature selection based on Artificial Bee Colony algorithm. *EURASIP Journal on Image and Video Processing*, 2013; **2013**: 1-8.

[7] Mehrabian, AR, Lucas C. A novel numerical optimization algorithm inspired from weed colonization. *Ecological informatics*, 2006; **1**: 355-366.

[8] Mallahzadeh, ARR, Oraizi H, Davoodi-Rad Z. Application of the invasive weed optimization technique for antenna configurations. *Progress In Electromagnetics Research*, 2008; **79**: 137-150.

Quantitative Nanostructure Activity Relationship Modeling (Nano-QSAR)

M. Salahinejad

Environmental Laboratory, NSTRI, Tehran, Iran

E-mail: salahinejad@gmail.com

With the rapid growth of the nanotechnology industries and the increasing use of nanomaterials in our life, the investigation of favorable and unfavorable effects of biological and physico-chemical properties of nanomaterials is very crucial for the future of nanotechnology. Experimental studies, particularly in the area of nanotoxicology are time-consuming and expensive. So, the use of efficient computational tools that can predict the desired properties are particularly important. Modeling and simulation of nanoparticles, in order to predicting their biological effects and estimate their physicochemical properties is facing important challenges.

The high structural complexity and diversity of nanoparticles is the first challenge to develop numerical parameters capable of characterizing the structural. The second problem relates to the lack of structural studies, biological and physico-chemical nanoparticles that can make it difficult to provide statistical models.

Nano-QSAR or Quantitative Nanostructure-Activity Relationships (QNAR) which was first introduced in 2009 by Puzyn and his colleagues, refers to the application of QSAR methods to build and develop of predictive models to estimate activity or property of nanomaterials.

Although the process of nano-QSAR is similar to QSAR approach, but in each step of modeling the limitation and difficulties related to the field of nanomaterials should be considered. Along with all problems and challenges in nano-QSAR modeling, there are a growing interest in using these methods because if the quick and accurate estimation of nanomaterials properties and reduce the cost and time of laboratory efforts.

Here, Nano-QSAR method was introduced and the difficulties of each step of nanoparticles properties modeling will discuss in more details. The achievements and applications of nano-QSAR models will be demonstrated.

Using global and local sensitivity analysis for variable sorting and selection

Akram Rostami[a] , Hamid Abdollahi[a]*, Marcel Maeder[b]

[a] *Department of Chemistry, Institute for Advanced Studies in Basic Sciences, P.O. Box 45195-1159, Zanjan, Iran,* [b] *Department of Chemistry, University of Newcastle, Callaghan, New South Wales 2308, Australia*

Email: abd@iasbs.ac.ir

Abstract

In this study we use the concept of sensitivity analysis (SA) for variable sorting based on their participation in the data variance. It is for the first time that the SA concept is used for variable sorting. We illustrate that Local SA just can sorted the extreme points of response vectors. But Global SA can do this for the entire response vectors. We show the results by an environmental simulated data.

Key words: Variable Selection, Variable Sorting, Lockal Sensitivity Analysis, Global Sensitivity Analysis

Introduction

Chemists routinely create models of reaction systems to understand reaction mechanisms, kinetic properties, process yields under various operating conditions. These models are built in the presence of uncertainties of various levels, in the pathway, in the order of the kinetics associated to the pathway, in the numerical value of the kinetic and thermodynamic constants for that pathway, and so on [1,2]. Sensitivity analysis (SA) contributes to the assessment of how variations in the output of a model can be apportioned, qualitatively or quantitatively, to different sources of variations [3]. Based on the factor space of interest, these techniques can be divided into two categories: local sensitivity analysis and global sensitivity analysis. Local SA detects the net effects of single parameters. Local SA methods do not identify interactions among factors [4]. A class of global methods of interest is that of the variancebased measures. Variance based sensitivity analysis want to rank factors according to how much the unconditional variance V(Y) of Y is reduced by fixing the various factors to their true value. The factors could then be ranked according to $V(Y \mid X_i = X^*_i)$, the variance being taken over all factors but X_i. In conclusion, $S_i = V_i/V(Y)$ is a proper measure of sensitivity to use to rank the input factors in order of importance. The total un-conditional variance for a model with k factors can be decomposed as:

$$V(Y) = \sum_{i} V_i + \sum_{i}\sum_{j>i} V_{ij} + \ldots V_{12\ldots k} \quad \text{Where} \quad V_i = V(E(Y \mid X_i)) \quad \text{and} \quad V_{ij} = V(E(Y \mid X_i, X_j)) - V_i - V_j$$

The V_{ij} terms capture that part of the effect of X_i and X_j that is not described by the first order terms. In this study we used this concept for variable sorting based on their participation in the data variance.

Results and discussion

In this contribution an environmental data is simulated. This data has 30 objects and 50 variables. The mentioned data is illustrated in figure 1. As shown in Figure 2a some of the columns of data placed in the corner of inner polygon (numbers show the variables number) among them variables number 12, 26 and 27 are the purest columns of data matrix (SIMPLISMA results). Pure variable methods just find the purest profiles equal the number of components that are in the corners of inner polygon (Enclosed space using extreme points of response vectors). Additionally others variables can be sorted based on sensitivity of data structure to them or based on participation in the data variance. Variance based local sensitivity analysis just could sort the variables based main effect, so just the corners of inner polygon is sorted. But variance based global SA could sort all of the variables based on their contribution on the data variance (figure 2b and 2c).

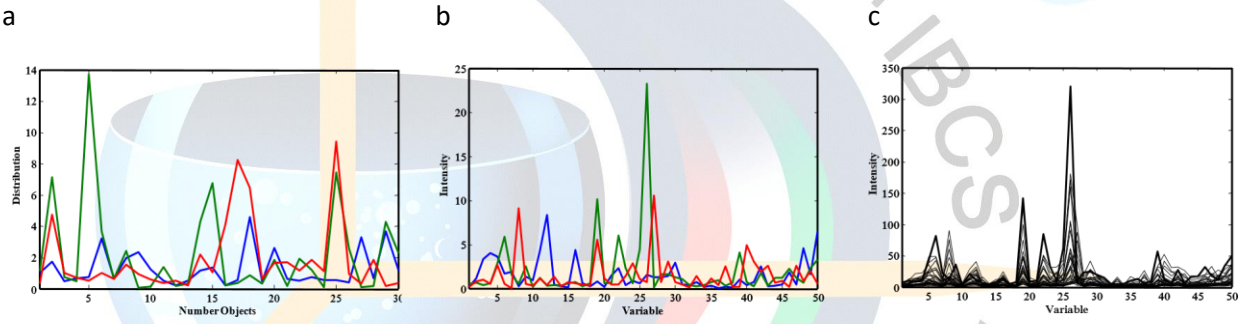


Figure 1: Demonstration of a) Concentration profiles, b) Spectral profiles, c) Data matrix.

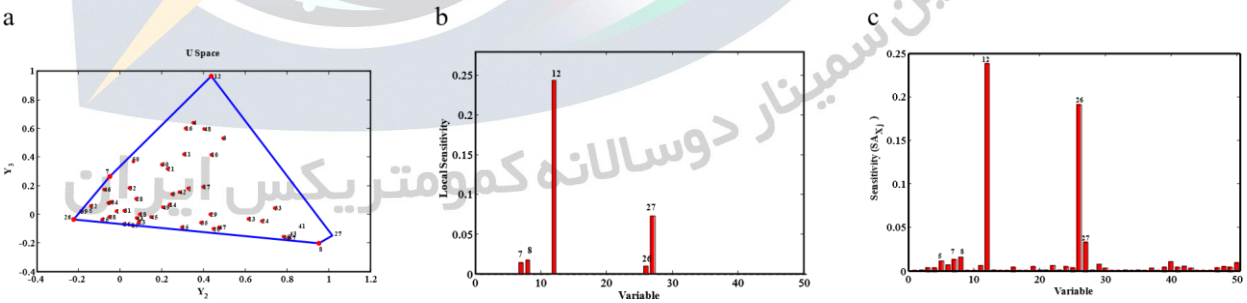


Figure 2: a) Illustration of sample space. Each point in this space is one column of data matrix. The blue closed area is the inner polygone. b) Local and c) Global sensitivity Analysis of the mentioned data.

References

[1]Rosen, R. A Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life; Columbia University Press: New York, 1991.

[2]Saltelli, A.; Chan, K; Scott, M. Sensitivity Analysis; John Wiley & Sons: Chichester, 2000.

[3]Saltelli, A.; Tarantola, S.; Campolongo, F. Stat. Sci. 2000, 15, 377.

[4]Saltelli, A.; Annoni, P. Environ. Modell. Soft. 2010, 25, 1508.

A theoretical approach to model and predict the adsorption coefficients of some small aromatic molecules on carbon nanotube

Afsane Heidari*, Mohammad H. Fatemi

Chemometrics Laboratory, Faculty of Chemistry, University of Mazandaran, Babolsar 47416-95447, Iran
E-mail address: afsane_heidari@yahoo.com

Abstract

In this study, a hybrid molecular docking-quantitative structure property relationship (QSPR) approach was used as predictive technique for modeling of adsorption coefficients of some small aromatic molecules on carbon nanotube. The structure of single wall carbon nanotube (SWCNT) was optimized using the density functional theory at the B3LYP/6-31G level of theory. Then molecular docking was used to simulate the interactions between 40 small aromatic compounds as ligands and optimized SWCNT as receptor. Then the most stable docked conformation of each ligand was used to extract the molecular descriptors. Some QSPR models were developed to correlate the relevant molecular descriptors of ligands to their adsorption coefficient onto nanotube. Three MLR models with squared correlation coefficient (R^2) values of 0.93, 0.94 and 0.95 were selected. The prediction power of models was evaluated on test set that was not used during the modeling and led to (R^2) values of 0.88, 0.85 and 0.93. The obtained models reveal role of topology and some quantum-electronic features of ligands in quantity of adsorption.*

Key words: adsorption coefficients- density functional theory- molecular docking- SWCNT.

Introduction

Adsorption of chemicals on nanomaterials is an important and effective phenomenon in nanotechnology. Adsorption forces in nanoparticles are dramatically influenced by physicochemical properties of nanoparticles (NPs), target compounds and solvents. The ability to predict the adsorption behavior of various compounds onto NPs leads designing efficient nanomaterials, achieving deeper insight into influencing structural properties and saving time and cost. So recently, computational studies on the adsorption of NPs have drawn more attentions. A theoretical method is quantitative structure-property relationship (QSPR) that is a mathematical model to correlate the molecular structural features of a set of compounds to their properties [1, 2]. These structural features are called descriptors and are calculated from optimized conformations of compounds. In real systems, the conformation of compounds depends on interactions of them with surrounding molecules. Since the main aim of molecular docking is reproduction of the experimental conformation of candidate ligand in interaction of target receptor, extraction of optimized conformers from this method can lead to more informative descriptors and more predictive QSPR models.

Results and discussion

Data set consists of the adsorption coefficients of forty small organic compounds in logarithmic scale on MWCNT [3]. Primary optimization of ligand and receptor structures was performed using Gaussian 09 [4] at the B3LYP/6-31G (d) level of theory. Docking calculations were conducted using the AutoDock Vina program and the

Flexidock module in SYBYL-X 1.1. The correlation coefficient (R) among the experimental log K values with calculated docking scores were 0.68 and 0.85 for AutoDock Vina [5] and the Flexidock, respectively. Therefore, conformations obtained from flexidock were chosen for further investigations and modeling.

Table 1. Statistical parameters of developed models.

	training set			test set					
	N	R^2	RMSE	N	R^2	RMSE	R^2_{LOO}	SPRESS	$R^2_{(Y-}$
Model	28	0.93	0.18	11	0.88	0.63	0.83	0.11	0.18
Model	28	0.94	0.17	11	0.85	0.55	0.78	0.13	0.15
Model	28	0.95	0.15	10	0.93	0.32	0.91	0.05	0.19

BINding ANAlyzer (BINANA) [6], DRAGON 3 [7] and CODESSA 2.7.2 [8] were used for calculation of descriptors. The stepwise MLR procedure was used to develop linear QSPR models. Three 5-parameter QSPR models were developed that their statistical parameters are indicated in Table 1. As can be seen in this table, the models have favorable statistical quality and good prediction power.

References

[1] Shao Y, Liu J, Wang M, Shi L, Yao X, Gramatica P: Integrated QSPR models to predict the soil sorption coefficient for a large diverse set of compounds by using different modeling methods. Atmos. Environ. 2014, 88:212-218.

[2] Cherkasov A, Muratov EN, Fourches D, Varnek A, Baskin II, Cronin M, Dearden J, Gramatica P, Martin YC, Todeschini R: QSAR modeling: where have you been? Where are you going to? J. Med. Chem. 2014, 57:4977-5010.

[3] Xia X-R, Monteiro-Riviere NA, Riviere JE: An index for characterization of nanomaterials in biological systems. Nat. Nanotechnol. 2010, 5:671-675.

[4] Frisch M, Trucks G, Schlegel H, Scuseria G, Robb M, Cheeseman J, Scalmani R, Barone G, Mennucci B, Petersson G: Gaussian 09, version A02, Gaussian. Inc, Wallingford, CT 2009.

[4] Trott O, Olson AJ: AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. 2010, 31:455-461.

[5] Durrant JD, McCammon JA: BINANA: A novel algorithm for ligand-binding characterization. J. Mol. Graphics Modell. 2011, 29:888-893.

[6] Todeschini, R.,Consonni, V.,Mauri, A.,Pavan, M., 2004. DRAGON-Software for the calculation of Molecular Descriptors. Webversion 3.

[7] Katritzky, A., Lobanov,V., Karelson, M., 1994. CODESSA Comprehensive Descriptors for Structural and Statistical Analysis. Reference Manual.

What is in Chemometrics for Chemoinformatics: Validating "classical relativity" in chemical space

Ahmad Mani-Varnosfaderani*[a] Marzieh Mokhtari[b], Alieh Ahdi[b]

[a] Department of Chemistry, Tarbiat Modares University, Tehran, Iran

[b] Department of advanced medical technologies, Isfahan University of Medical Sciences, Isfahan, Iran.

Abstract

The concept of "classical relativity in chemical space" (CRCS), has been previously proposed by our research group [1]. This is a new way for exploring in chemical space and finding molecules with reasonable activity against a particular therapeutic target. The present work paves the way for validating the CRCS methodology. One million random molecules of PubChem and ZINC databases were used as decoys for virtual screening. 98,000 active molecules belong to 110 important biological targets were collected from Binding-DB [2] and used for development of the classifiers. Three strategies have been applied for validating the concept: 1) Development of one-against-all (OAO) classifiers using the genetic algorithm and quadratic discriminant analysis techniques, 2) Assessment of orthogonality in target space for model development and 3) High throughput docking using autodock vina [3]. The results revealed that as the number of comparisons in chemical space increases, better AUC values (area under curve values of receiver operating characteristic curves) obtain. Moreover, the selected therapeutic targets have to be orthogonal from the view of their amino acid sequences. Finally, the results revealed that the enriched subspaces using "CRCS" methodology yield better binding energies rather that the random selected subsets of ZINC and PubChem databases.

Keywords: virtual screening, chemical space, classification, high throughput docking, CS-MINER

Introduction

With proceeding of technology in recent years, the newly developed techniques help the researchers to discover potent drugs and finally cure fatal diseases. The technologies of high throughput screening and database development are among the most basic strategies. Although high throughput screening is being developed, its efficiency remains rather low. As the size of the known chemical space is rather infinite, it is near impossible to completely explore in. In order to overcome this difficulty, the in-silico techniques help researchers to simulate the real experiments. With no doubt, in-silico evaluation of newly suggested molecules saves huge amount of time and money in modern drug discovery projects. Researchers in the field of computational drug design and Chemoinformatics try to find new molecules in chemical space with potential activity toward some biological targets. They try to develop some algorithms for effective screening of the chemical space. The biggest database of molecules made in computers is GDB-17 developed by Reymond et al in Switzerland [4]. This database contains more than 166 billion molecules. This is quite obvious that a thorough screening of such large databases requires advanced computers and technologies [5]. As previously mentioned, our research group proposed the concept of "classical relativity" in chemical space as a new strategy for efficient exploring of large compound databases. This theory implies that using only active sets molecules and multi-class classifiers it is possible to find particular subspaces in chemical space [1]. Figure 1 illustrates a large chemical space occupied by some subspaces. Each subspace is densely populated with specific class of biosimilar molecules. We previously illustrated that for thorough addressing a particular subspace in chemical space, one can compare a subspace with other subspaces. As the dimensionality of the chemical space is usually high, therefore it is important to find discriminative molecular descriptors for addressing medicinal subspaces. We have used a combination of genetic algorithm and quadratic discriminant analysis for dimension reduction and classification purposes, simultaneously. In

our previous paper, "*CS-MINER: A tool for association mining in chemical space*" [1] we illustrated that as the number of comparisons between subspaces increase, better AUC values obtain for the screening procedure. The present work tries to answer this following question: "Witch subspaces should be compared to yield better AUC values?" Suppose a medicinal chemist want to synthesize some Aromatase inhibitors. Therefore he/she has to find a subspace in chemical space, mainly occupied by Aromatase inhibitors. The question is how to find this subspace. Our response is "relativity". If this medicinal chemist starts to compare active Aromatase inhibitors with the active inhibitors of other classes such as Kinases, Histidine and Dopamine, he/she will find some specific characteristics of Aromatase inhibitors. This is simply the main way of learning i.e. *"comparison"*.

Results and Discussion

In order to answer the question "which subspaces should be compared?" a new strategy is proposed in this work. The answer is comparing those subspaces which their biological targets are more orthogonal from the view of their amino acid sequences. Actually, we propose "*orthogonality in target space"* as a new way for virtual screening in *ligand space*. In fact we are trying to build a bridge in between the target space and ligand space. In order to compare the sequence similarity of 110 biological targets the Needleman-Wunsch and Smith-Waterman algorithms have been used. These methods use the PAM and BLOSUM matrices for comparing the sequences of proteins. We sort different sets of proteins based on the degree of their similarities. Then we used the inhibitors of most non-similar biological targets for development of the classifiers. We realized that as the degree of the similarity of proteins decrease, better AUC values obtain for ligand based virtual screening. It implies that it is better to compare the inhibitors of the proteins that their amino acid sequences are more different. This methodology has been tested using 1 million random molecules of GDB-17 as decoys and 98000 active inhibitors taken from Binding-DB. The average AUC values for the inhibitors of proteins from different families were distinctively larger than similar biological targets. We explore the accuracy of this theory using virtual screening of active decoys from DUD database. The average values of the molecules selected using the proposed strategy were more than the randomly selected compounds.

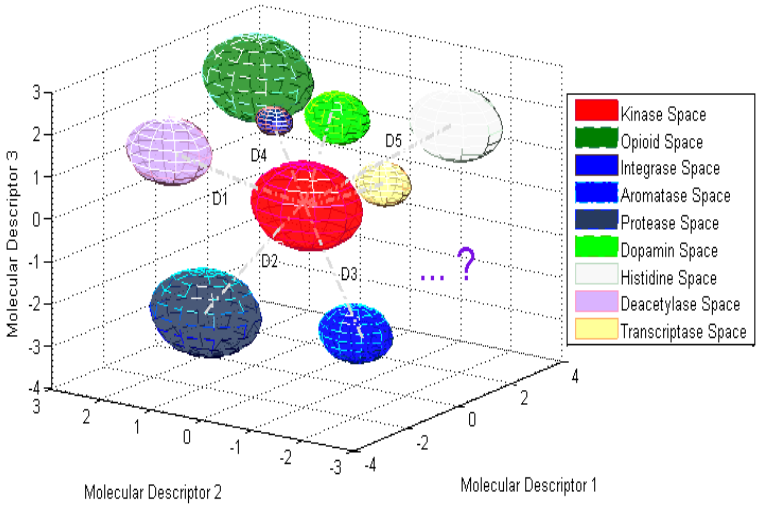


Figure 1.

A schematic view of different medicinal subspaces in a large chemical space. It is important how to address a particular subspace. Our response is "classical

References

[1] A. Mani-Varnosfaderani, A. Valadkhani, M. Jalali-Heravi, *Mol. Inform*. **34** (2015) 185-196.

[2] X. Chen, Y. Lin, M. K. Gilson, *Nucleic Acid Sci*. **61** (2002) 127-141.

[3] O. Trott, A. J. Olson, *J. Comput. Chem*. **31** (2010) 455-461.

[4] L. Ruddigkeit, L. C. Blum, J.-L. Reymond, *J. Chem. Inf. Model*. **53** (2013) 56-65.

[5] J.-L. Reymond, *Acc. Chem. Res*., **48** (2015) 722-730.

Direct Triadic Decomposition

Nematollah Omidikia*[1,2], Róbert Rajkó*[2], Hamid Abdollahi[1], Mohsen Kompany-Zareh[1]

[1] *Department of Chemistry, Institute for Advanced Studies in Basic Sciences, P.O. Box 45195-1159, Zanjan, Iran*

[2] *Institute of Process Engineering, Faculty of Engineering, University of Szeged, Moszkvai krt. 5-7, H-6725 Szeged, Hungary*

E-mail address: rajko@sol.cc.u-szeged.hu

Abstract

Multi-way data are being generated by sophisticated analytical instruments, requiring new and effective data-mining techniques. Although much effort has been directed to the development of numerical methods for decomposition of two and three-way data arrays, direct methods for curve resolution of three-way data arrays has lagged behind. Parallel factor analysis (PARAFAC) is believed to be one of the most widely used trilinear curve resolution techniques. The most common way for calculating PARAFAC decomposition is alternating least squares method (ALS). In this contribusion Direct Triadic Decomposition (DTD) is proposed for direct triadic resolution of two and three-component three-way arrays. The principle behind the algorithm is described and tested for simulated noise free and noisy data sets. Superiority of DTD to PARAFAC decomposition is confirmed by promising decomposition of big data sets.

Key words: Triadic decomposition, Trilinearity, PARAFAC, Uniqueness, Curve resolution, SVD

Introduction

Curve resolution is a class of mathematical techniques concerned with the estimation of pure profiles [1]. For two-way data sets, Lawton and Sylvestre (LS) [2] and Borgen method [3,4] calculated all of possible solutions analytically under non-negativity constraint for two and three-component system respectivily. Despite SMCR methods which determine all of possible solution analytically in MCR using approximation methods, a single solution among feasible profiles is provided. Although much effort have been directed to the development of approximation methods for decomposition of data sets, analytical determination remains unexplored especially for three-way data sets.

In this contribution, direct triadic decomposition, DTD, is introduced based on non-convex linear combination of data rows for two and three-component three-way arrays. Finally, it is demonstrated that the proposed algorithm combines fast computation and successful handling of noisy data sets, like given in the case of massive fallible data sets.

Results and discussion

$\underline{\boldsymbol{R}}_{I,J,K}$ as a two-component EEM three-way array is simulated. Profiles for simulation are depicted in fig. 1.
Simplified steps of DTD implemented as follows:

1) Data array is unfolded in appropriate direction $\boldsymbol{R}_{I,J\times K}$, row-wise here.
2) Two rows (m^{th} and n^{th}) among I rows of $\boldsymbol{R}_{I,J\times K}$ are selected. Profiles related to considered rows with the size of ($1\times JK$), are reshaped to matrix with the size of ($J\times K$). For example $\vec{\boldsymbol{r}}_{m,J\times K}$ reshaped to a matrix, $\boldsymbol{M}_{J\times K}$.
3) Columns of $\boldsymbol{M}_{J\times K}$, *and* $\boldsymbol{N}_{J\times K}$ are normalized using 1-norm ($\|\boldsymbol{x}\|_1 = \boldsymbol{1}^{\boldsymbol{T}}|\boldsymbol{x}|$) for the columns.
4) The task is calculation of α and β using non-convex linear combination of $\boldsymbol{M}_{J\times K}$, *and* $\boldsymbol{N}_{J\times K}$:

$$\alpha\, \boldsymbol{M}_{J\times K} + \beta\, \boldsymbol{N}_{J\times K} = \boldsymbol{TR}_{J\times K} \qquad (1)$$

$\boldsymbol{TR}_{J\times K}$ is matricized trilinear solution.

Algebraic description of our tasks can be formulated as follows:

$$\alpha\, \boldsymbol{M}^1_{J\times 1} + \beta\, \boldsymbol{N}^1_{J\times 1} = n_1\, \boldsymbol{TR}^1_{J\times 1} \qquad (2)$$

$$\alpha\, \boldsymbol{M}^2_{J\times 1} + \beta\, \boldsymbol{N}^2_{J\times 1} = n_2\, \boldsymbol{TR}^2_{J\times 1}$$

… … …

$$\alpha\, \boldsymbol{M}^K_{J\times 1} + \beta\, \boldsymbol{N}^K_{J\times 1} = n_K\, \boldsymbol{TR}^K_{J\times 1}$$

According to abovementioned equations there are K series of equations with two unknowns. Although dependent variable is not at hand, there is a strong relation something like this between dependent variables:

$$n_1\, \boldsymbol{TR}^1_{J\times 1} = n_2\, \boldsymbol{TR}^2_{J\times 1} = \ldots = n_K\, \boldsymbol{TR}^K_{J\times 1} \quad (3)$$

Normalized linear combination of corresponding p^{th} and q^{th} columns of $\boldsymbol{M}_{J\times K}$, *and* $\boldsymbol{N}_{J\times K}$ must be equal:

$$\alpha\, \boldsymbol{M}^1_{J\times 1} + \beta\, \boldsymbol{N}^1_{J\times 1} = \alpha\, \boldsymbol{M}^2_{J\times 1} + \beta\, \boldsymbol{N}^2_{J\times 1} = \cdots = \alpha\, \boldsymbol{M}^K_{J\times 1} + \beta\, \boldsymbol{N}^K_{J\times 1} \quad (4)$$

Considering normalization, 1-norm is considered in this part for normalization, and non-convex linear combination, $\beta = 1 - \alpha$, equation (4) can be rewritten as:

$$\frac{[\alpha\, \boldsymbol{M_p} + (1-\alpha)\boldsymbol{N_p}]}{[\alpha \sum \boldsymbol{M_p} + (1-\alpha)\sum \boldsymbol{N_p}]} = \frac{[\alpha\, \boldsymbol{M_q} + (1-\alpha)\boldsymbol{N_q}]}{[\alpha \sum \boldsymbol{M_q} + (1-\alpha)\sum \boldsymbol{N_q}]} \quad (5)$$

Rearrangement of equation 5 resulted in a quadratic equation with respect to α:

$$a\,\alpha^2 + b\,\alpha + c = 0 \quad (6)$$

Using MATLAB function *Roots*, two possible solutions of quadratic equation can be obtained.

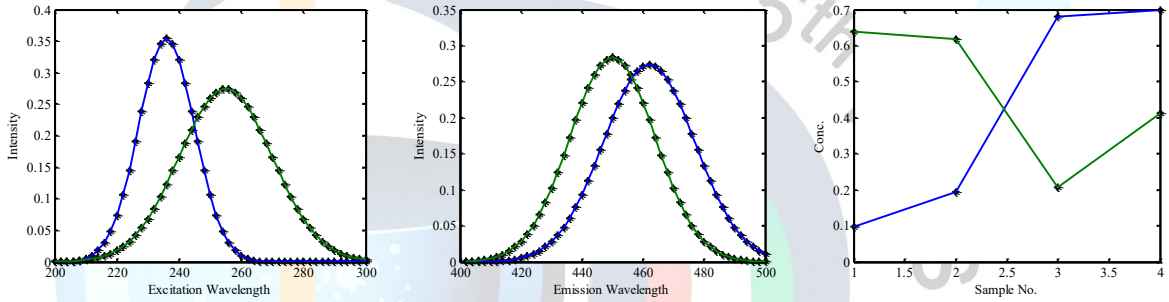


Figure 1. Real and estimated profile for considered data set. Solid lines depict real profiles and black stars represent estimated profiles.

It is worth to highlight that for a simulated noisy two-component data arrays with the size of 401×903×50, proposed strategy takes 4 second while standard PARAFAC-ALS of PLS-toolbox takes 58 seconds. Finally, generalization of proposed method to more than two-component is straightforward.

References

[1] Paulick C, Wright M N, Verleger R, Keller K. Decomposition of 3-way arrays: A comparison of different PARAFAC algorithms. *Chemom. Intell. Lab. Syst*. 2014; **137**; 97–109.

[2] Lawton WH, Sylvestre EA. Self Modeling Curve Resolution. *Technometrics* 1971; **13**; 617–633.

[3] Borgen O S, Kowalski B R. An extension of the multivariate component-resolution method to three components. *Anal. Chim. Acta* 1985; **174**; 1–26.

[4] Rajkó R, István K. Analytical solution for determining feasible regions of self-modeling curve resolution (SMCR) method based on computational geometry *J. Chemom.* 2005; **19**; 1–16.

An Investigation on Meaning and Reliability of Local Rank Constraint in Self-Modeling Curve Resolution of Chemical Data

Hamid Abdollahi* and Mahsa Akbari

Faculty of Chemistry, Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan, Iran

E-mail address: abd@iasbs.ac.ir

Abstract

The local rank information plays a key role in the curve resolution of multicomponent chemical systems. Applying the selectivity and local rank constraints can reduce considerably the extent of rotational ambiguity in the results of MCR methods and in some favorable data sets unique solutions may be obtained. Local rank exploratory methods like EFA provide the presence pattern of components which contains knowledge as essential as selective windows and windows with a number of components smaller than the total rank. The presence windows of components are deduced from the local rank maps on the main assumption that the number of components in each window is equal to its rank.

It is shown in this work that the information related to the local rank can be extracted from the structure of data sets by means of computational and geometry tools (i.e. LS in the case of two- and Borgen plot in the case of three-component systems) in addition to rank exploratory methods. Also, the concordance of local rank windows with the presence windows of components is investigated. Although the proposed detection tools are restricted to two and three component systems up to now, but the obtained results are general for systems with any number of components.

Keywords: Local rank constraint; Borgen plot; Self modeling / multivariate curve resolution; Rotational ambiguity.

1. Introduction

Geladi and Wold used the local rank analysis in chemistry for the first time [1]. In their approach, called local rank mapping, the two-dimensional data set splits into small sections and then the local rank of each section is measured. After that work, a broad range of approaches were developed that determine the local rank information of a system based mostly on PCA method and then relate the determined local rank windows (or rank maps) to the presence windows of components (e.g. elution windows in case of chromatographic data) like Evolving Factor Analysis (EFA) method [2, 3] and other algorithms based on EFA [4-6]. Actually, it is believed that the resolution results of many non-iterative methods rely almost exclusively on the appropriate determination of local rank information and the iterative methods also take the advantage of this information to employ equality and zero region constraints or to build initial estimates.

After 20 years from the famous papers of Borgen *et al.* [7], computational geometry tools were introduced and applied by Rajko and Istvan for the first time to draw the Borgen plot [8]. One year later, Rajko [9] proposed a simpler algorithm based on duality concept for minimal constrained SMCR. Very recently, Rajko *et al.* have proposed the so called "data-based uniqueness theorem" for detection of uniquely recovered profile(s) based on data microstructure, i.e. the structure of data in reduced abstract subspace, again for minimal constrained SMCR [10].

With regard to the points mentioned about the importance of local rank information, two major objectives are pursued in this work: Evaluation and gaining different local rank information from Borgen plot and generally from data microstructure and then investigating the concordance of the local rank information with the presence windows of components in bilinear data matrices.

2. Results and discussion

2.1. Assessment of local rank information by data microstructure

Generally, a row or column window with local rank-r describes an r-dimensional hyperplane in the microstructure of an n-dimensional data and after a proper normalization (Borgen norm), this dimensionality reduces by one (an r-1 dimensional subspace). In example, the position of some rows of a specific data matrix with local rank-one in the V^T subspace should be along a line passing through the origin because all have the same shape. After using mentioned normalization, the points (rows of the data in abstract subspace) will coincide with each other, means they are actually the same. So, one point in the normalized abstract subspace denoting more than one row (or column) indicates the local rank-one window, as it is recognizable in the Borgen plot of the simulated data, fig. (1).

So, it can be concluded that the inner polygon in the case of a tree component system (or polytope in an higher dimensional data) can indicate the local rank windows in either U-space, the space spanned by the columns, or V-space, the space spanned by the rows (depending on considered abstract subspace), though local rank has been used mostly in concentration space.

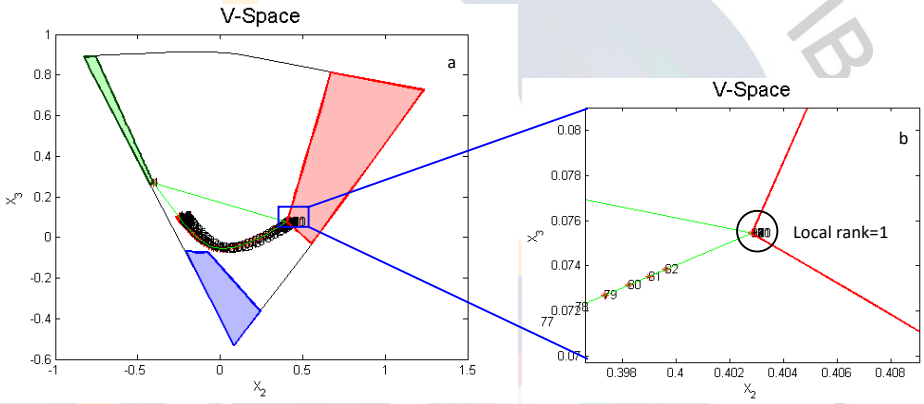


Figure 1. The Borgen plot of the simulated data set. The location of rows with local rank-one in the V-space is indicated by the circle.

2.2. Investigation on the concordance of local rank information and presence windows of components

The agreement of the local rank information with actual presence windows of the components is of great importance because the local rank constraint, which usually has a drastic effect on feasible bands, is applied based on the correctness of this concordance. For instance, when selectivity is applied as a constraint it is considered that just one component should participate to reconstruct the rows or columns with local rank-one. A main question arises here: Is

there any possibility that more than one component participate to build the local rank-one window? Or in general, may rank deficiency be present in local windows of a data matrix?

Data microstructure will give exact guides to this study. Regarding the Borgen plot of the simulated data set, the local rank-one window, which is indicated by a single point, is inside the feasible region of the first component thus the retransformed point (i.e. the spectral profile) can be considered as the pure spectrum for that component. Then there is no wonder to say that the contribution of other components for building these rows is zero (selective window). However, a subtle look will confirm that these rows can be constructed from two or three components, too! This is due to the convex linearity property in the feasible regions. Data microstructure demonstrates several sets of feasible solutions that the linear combination of solutions for the first and second components can build the rows with local rank-one. The local rank-one window, in the same way, can be constructed from the linear combination of three different components. Consequently, the local rank-one window in this case may be considered as a selective window, or not.

It is interesting that different sets of feasible solutions may have different presence patterns but all have the same local rank windows, because all fit the original data set equally well. The sets of feasible solutions that are not in full agreement with local rank information have rank deficiency in some of their local windows due to the equal shape of two or three profiles in that particular section. So, it seems that the presence pattern extracted from local rank exploratory methods is not always in agreement with real chemical presence pattern.

References

[1] Geladi P., Wold S. Local Principal Component Models, Rank Maps and Contextuality for Curve Resolution and Multi-Way Calibration Inference, *Chemom. Intell. Lab. Syst*. 1987; **2**: 273–281.

[2] Gampp H., Meader M., Meyer C. J., Zuberbuhler A. D. MODEL-FREE ANALYSIS OF SPECTROPHOTOMETRIC AND ESR TITRATIONS, *Talanta*, 1985; **32**: 1133-1139.

[3] Meader, M. Evolving Factor Analysis for the Resolution of Overlapping Chromatographic Peaks, *Anal. Chem.* 1987, **59**: 527-530.

[4] Andrew, C., Whitson, A., Maeder M. Exhaustive evolving factor analysis (E-EFA), *J. Chemom*. 2001; **15**: 475–484.

[5] Zeng Z. D., Xu C. J., Liang L. Z., Li B. Y. Sectional moving window factor analysis for diagnosing elution chromatographic patterns, *Chemom. Intell. Lab. Syst*. 2003; **69**: 89-101.

[6] Keller H. R., Massart D. L. Peak purity control in liquid chromatography with photodiode-array detection by a fixed size moving window evolving factor analysis, *Anal. Chim. Acta*, 1991; **246**: 379-390.

[7] Borgen OS, Kowalski BR. An extension of the multivariate component-resolution method to three components, *Anal. Chim. Acta* 1985; **174**: 1–26.

[8] Rajkó R., István K. Analytical solution for determining feasible regions of self-modeling curve resolution (SMCR) method based on computational geometry *J. Chemom.* 2005; **19**: 448-463.

[9] Rajkó R. Natural duality in minimal constrained self modeling curve resolution, *J. Chemom.* 2006; **20**: 164-169.

[10] Rajkó R., Abdollahi H., Beyramysoltan S., Omidikia N. Definition and detection of data-based uniqueness in evaluating bilinear (two-way) chemical measurements, *Anal. Chim. Acta*, 2015; **855**: 21-33.

A systematic study on the effect of noise and shift on multivariate figures of merit of second-order calibration algorithms

Mohammad Ahmadvand[a], HadiParastar[b*], Hassan Sereshti[a]

[a]*Department of Chemistry, Faculty of Science, University of Tehran, Tehran, Iran*
[b]*Department of Chemistry, Sharif University of Technology, Tehran, Iran*
E-mail address: h.parastar@sharif.edu

Abstract

In the present study, multivariate analytical figures of merit (AFOM) for three well-known second-order calibration algorithms (i.e., PARAFAC, PARAFAC2 and MCR-ALS)were investigated in simulated hyphenated chromatographic systems including different artifacts (e.g., noise (0-1%) and shift (0-10%)). In this regard, different two- and three- component systems with calibrated analytes and uncalibrated interferences in test set were simulated. The resolved profiles of target compounds were used to build calibration curves and calculation of multivariate AFOM of sensitivity (SEN), selectivity (SEL) and limit of detection (LOD). Finally, the obtained AFOM for different data sets using different algorithms were used to compare the performance of algorithms and calibration ability. It was concluded that MCR-ALS has better performance for correct resolution and obtaining reliable AFOM in chromatographic systems with various levels of noise, shifts and interferences. However, as the noise level can significantly affect the SEN values, therefore, it is hard to get satisfactory AFOM in real chromatographic systems.

Key words: Second-order calibration; Multivariate analytical figures of merit; MCR-ALS; PARAFAC.

Introduction

During the recent decade, second-order calibration algorithms have been frequently used toget pure qualitative and quantitative information of target analytes in complex mixtures. On the other side, extensive efforts have been made by chemometricians to define appropriate multivariate analytical figures of merit (AFOM) [1-3]. Consequently, sensitivity (SEN) has been defined as the main AFOM to calculate other AFOMs [4]. In fact, the SEN is obtained from proportion of uncertainty in signal of test sample to uncertainty in concentration. Therefore, the ratio of the uncertainty in test sample signal to uncertainty in predicted concentration can be a good measure of SEN. In this new definition, sensitivity is analyte specific, sample specific and algorithm specific that is dependent on the test samples. The other multivariate AFOMs are defined based on SEN, such as analytical sensitivity, selectivity (SEL) and limit of detection (LOD). To the best of our knowledge, there is no report on comparison of different second-order calibration algorithms based on multivariate AFOM. Also, the effects of noise and shift on multivariate AFOM have not yet been studied.

Results and Discussion

Different hyphenated chromatographic data sets with different degree of overlap, shift, noise and various number of components were simulated as follows:(*i*) calibrated two-component system, (*ii*) calibrated three-component system and (*iii*) calibrated two-component system with two uncalibrated components in the test set. The elution profiles were simulated exponentially modified Gaussian (EMG) equation [5] and real spectral profiles were given from NIST library

for some polycyclic aromatic hydrocarbons. The calibration sets were designed in seven concentration levels and three replicates. Data sets for different systems with various levels of shift (0-10%) and noise (0-0.5% homoscedastic and 0-0.5% heteroscedastic) were properly arranged in multi-set and multi-way structure and were analyzed by MCR-ALS, PARAFAC and PARAFAC2. As an istance Table 1 shows the results for target analyte in three different systems (5% shift and 0.1% noise).

Table 1. The obtained results for benzo [a] pyrene with 5% shift and 0.1% noise in three data sets

	LOF[a]	R^{2b}	RE[c]	Analytical SEN	SEL	LOD
System (*i*)						
MCR-ALS	2.76	1.0000	0.0029	38.00	1.00	0.0945
PARAFAC	33.88	0.9982	3.5087	0.09	1.00	39.280
PARAFAC2	N.R[c]	N.R	N.R	N.R	N.R	N.R
System (*ii*)						
MCR-ALS	2.07	1.0000	0.0075	32.00	1.00	0.1140
PARAFAC	38.34	0.9997	1.3304	0.031	0.96	120.00
PARAFAC2	N.R	N.R	N.R	N.R	N.R	N.R
System (*iii*)						
MCR-ALS	2.74	1.0000	0.0040	41.00	1.00	0.0887
PARAFAC	22.91	0.9983	3.4536	0.120	0.63	31.430
PARAFAC2	N.R	N.R	N.R	N.R	N.R	N.R

[a] Lack of fit,[b]Regression coefficient,[c]Relative error of calibration,[d] Not resolved

In general, when there is no shift in the systemsthe results of PARAFAC and MCR-ALS are comparable and similar to the true solutions in terms of LOF and multivariate AFOM (analytical sensitivity, selectivity and limit of detection). However, incorporation of noise and shift in data cause deviation from trilinear model assumption needed for PARAFAC and therefore the calibration performance get worse. In other words, the results of bilinear MCR-ALS model are still good in different levels of nise and shifts. It is important to note that the PARAFAC2 only converged in noiseless data and until 5% shift levels. Also, the resolved elution profiles for PARAFAC2 were negative and/or bimodalin test samples (in presence of uncalibrated species). Finally, the effects of uncalibrated species on the performance of PARAFAC and MCR-ALS were investigated. In general, the multivariate AFOM for PARAFAC got worse in the presence of uncalibrated species whrease MCR-ALS results were still acceptable.

It can be concluded that bilinear MCR-ALS has better performance in the presence of different artifacts compared to PARAFAC and PARAFAC2. Also, the multivariate AFOMs are approperite criteria to evaluate the performance of second-order calibration algorithms in the presence of different artifacts and interferences.

References

[1] Olivieri AC, Faber K. New Developments for the Sensitivity Estimation in Four-Way Calibration with the Quadrilinear Parallel Factor Model. *Anal. Chem*. 2012; **84**: 186–193.

[2] BauzaMC, Ibanez GA, Tauler R, Olivieri AC. Sensitivity Equation for Quantitative Analysis with Multivariate Curve Resolution-Alternating Least-Squares: Theoretical and Experimental Approach. *Anal. Chem.* 2012; **84**: 8697-8706.

[3] Allegrini F, Olivieri AC. IUPAC-Consistent Approach to the Limit of Detection in Partial Least Squares Calibration. *Anal. Chem*. 2014; **86**: 7858-7866.

[4] Olivieri AC. Analytical figures of merit: from univariate to multiway calibration, *Chem. Rev.* 2014; **114**: 5358-5378.

[5] Kalambet Y, Kozmin Y, Tikhonov P. Reconstruction of chromatographic peaks using the exponentially modified Gaussian function. *J. Chemom*. 2011; **25**: 352–356.

پوسترها

Poster Presentations

صفحه ۶۶ تا ۳۲۳

Investigating the effect of constraints on the accuracy of the area of feasible solutions for the real datasets with an improved cost function of the polygon inflation algorithm

Nahal Rahimdoust[a], Mathias Sawall[b], Klaus Neymeyr[b], [c], Hamid Abdollahi[a*]

[a]*Department of Chemistry, Institute for Advanced Studies in Basic Sciences, P.O. Box 45195-1159, Zanjan, Iran*

[b]*Universitat Rostock, Institut f¨ur Mathematik, Ulmenstrasse 69, 18057 Rostock, Germany*

[c]*Leibniz-Institut for Katalyse e.V. an der Universit¨at Rostock, Albert-Einstein-Strasse 29a, 18059 Rostock*

E-mail address: mozhdehi@iasbs. ac.ir

Abstract

Self-modeling curve resolution methods have continuously been improved during recent years.Choosing proper constraints and cost functions is critically important for the reduction of the rotational ambiguity because the constraints have a direct influence on the accuracy of the Area of Feasible Solution (AFS). In this work we introduce new improved cost function which serves to apply nonnegativity, unimodality, equality and monotony constraints. We investigate the reduction of the AFS under hard and soft constraints. Another point of this work is to evaluate the accuracy and precision of the reduced AFS in the presence of noise and perturbations, under hard and soft implementation of constraints. A comparison is given between the reduced AFS with soft constraints (small deviations from constraints are accepted) and the reduced AFS under hard constraints (restrictedly forced constraints). Graphical visualization of this comparison is presented for various model problems. The results show that an AFS computation with soft constraints provides more reliable results, especially in the presence of noise.Using tunable parameters ε, γ, ω, δ is one of the advantages of soft constrained cost function.Ultimately soft constraints can help to reduce the lack-of-fit and they are proper instrument to handle the effect of noise on the AFS.

Keywords*:* multivariate curve resolution, nonnegative matrix factorization, area of feasible solutions, soft constraints, polygon inflation.

1. Introduction

Multivariate curve resolution methods aim at decomposing sequences of spectra taken from a multicomponent chemical reaction system into the underlying contributions from the pure components. If these spectra are collected row-wise in a matrix D, then the Lambert-Beer law [1] says that D can approximately be factored into a product of a matrix C containing column wise the concentration profiles of the pure components and a matrix A containing row-wise the associated pure component spectra, that is $D = CA$. (1)

In general, the factorization (1) is not unique and contain all of possible nonnegative solutions exist [1]. In order to determine only one solution, the chemist wants the “true” or “chemically correct” one, various techniques based on soft models or hard models have been developed. Sometimes such methods can result in poor pure component factors. A fundamentally different approach is to compute the full set of all possible nonnegative solutions and afterwards to select a solution. The set of all solutions that are nonnegative and represent the data are called the Area of Feasible Solutions (AFS) [2]. The aim of this work is to demonstrate the impact of soft constraints on the solutions represented by the AFS and to present a hybrid approach which combines the conceptual rigor of an AFS computation with the successful

regularization techniques underlying soft constraints. The resulting method allows extracting chemically meaningful solutions from the set of feasible nonnegative factors.

2. Results and discussion

2.1. Model examples

In order to evaluate the effect of the soft and hard constraints on the AFS, we use some simulated model examples. All data matrix have 95 rows and 200 columns.

2.2. Real data sets

Next soft constraints are applied to a three component model problem in the form of the consecutive reaction.

$$X \rightarrow Y \rightarrow Z$$

The kinetic parameters without units are $k_1 = 1$ and $k_2 = 0.5$, and the time interval without unit is [0, 15]. The pure component spectra are assumed to be simple Gaussian curves on the wavenumber interval [0, 100]. Equidistant grids are used with k = 101 points along the time axis and n = 201 points along the frequency axis.

2.3. Discussion

Every SMCR application has an implied research hypothesis as Gemperline mentioned [4] ”There exists an unconstrained bilinear model with unimodal, nonnegative pure component concentration profiles and pure component nonnegative spectra that fits the data matrix This hypothesis is tested for any pair of concentration profiles and spectral profiles. If the hypothesis is true, then the borders of the AFS can be calculated with no active constraints. Due to noise and non-ideal response (chemical and physical), the AFS may calculate with final models that frequently require active constraints. In the practical data sets active constraints during the AFS computations may increase the models lack-of-fit and affect the accuracy of the AFS. In this paper, we present a cost function that employs “soft” constraints and works with polygon inflation algorithm. We also compare the soft AFS with hard AFS.

Accuracy of the AFS under nonnegativity constraints

In the case of practical data, profiles may have small negative entries due to the noise[3]. We illustrate that inan AFS with hard nonnegativity constraints no negative deviations (strictly enforced) are accepted. This meansthat the feasible bands of hard nonnegativity constraints may not contain the real noisy profiles, because for thehard nonegativity constraints ($\varepsilon = 0$) the defined abstract space is not the real space of the noiseless dataso for the real data sets the accuracy of the feasible bands is reduced due to the noise effect.

Unimodality constraints

The AFS can be strongly affected by implementing (hard or soft) constraints. Among the constraints unimodality (only one maximum in each profile is allowed) is commonly applied for concentration profiles.

Unreal shapes of the profiles under hard unimodality constraints

These spikes are omitted or changed by the implementation of hard unimodality constraints.

Thischange lead to a wrong suppression of some parts of the profiles. By using soft unimodality constraints, the obtained profiles have shapes much closer to the original noisy peaks.

The effect of the soft equality constraints on the accuracy of the AFS in the dual space

Equality constraints have special effects on both subspaces in three component systems. So implementing right known profiles as auxiliary data is important. Some practical data sets dueto noise, distortion or temperature effects suffer from shifts or small changes in the shape of the profiles thatcause small deviations from the bilinear structure of the data. Sometimes different noise levels on auxiliary andmixture data may cause problems in implementing the equality constraints because separate measured knownprofiles normally have different noise distortion from the profiles which exist in the mixture data sets. Use of the hard constraints in the practical data is problematic because the hard equality constraints reproduce profiles just like the auxiliary data, which leads to poorly fitted results and also changes the subspace of the dual profiles. In the case of soft constraints some near neighbors of the auxiliary data (known spectral profile) are automatically accepted to be within the AFS, so the lack- of-fit reduces and the dual space is calculated properly.

Conclusion

Using soft constraints can prevent from over-fitting, because some near neighbors of the known profile are automatically accepted to be within the AFS of soft constraints. In contrast, hard equality constraints may cause over-fitting problems because measured known profiles normally have different noise distortion than the profiles which exist in the mixture data sets, so it may lead to over-fitting in the results. From the other point of view all soft cost functions (nonnegativity, unimodality, and equality) are minmized with a new fast numerical scheme of the adaptive approximation. The AFS for three-component systems is calculated by the polygon inflation algorithm. The polygon inflation approximates the boundary [5,6] of the AFS in a smooth way. The numerical calculations are time efficient and fast enough to calculate and draw the abstract spaces and feasible bands in less than one second. It is possible to use the introduced method on which systems that have moderate noise. The reductionof the rotational ambiguity by using soft constraints simultaneously in both concentration and spectral space in the presence of moderate noise and deviations under various constraints is one of the remarkable abilities of this method. Moreover the polygon inflation technique can principally be generalized to a polyhedron inflation scheme in order to approximate the AFS in case of an n-component system.

References

[1] A. Beer. Bestimmung der Absorption des rothen Lichts in farbigen Flussigkeiten. *Annalen der Physik*, 162(5):78–88, 1852.

[2] H. Abdollahi and R. Tauler. Uniqueness and rotation ambiguities in Multivariate Curve Resolution methods. *Chemometrics and Intelligent Laboratory Systems*, 108(2):100–111, 2011.

[3] P.J. Gemperline. Computation of the Range of Feasible Solutions in Self-Modeling Curve Resolution Algorithms. *Analytical Chemistry*, 71(23):5398–5404, 1999.

[4] P.J. Gemperline and E. Cash. Advantages of soft versus hard constraints in self-modeling curve resolution problems. Alternating least squares with penalty functions. *Anal. Chem.*,75:4236–4243, 2003.

[5] M. Sawall, C. Kubis, D. Selent, A. B¨orner, and K. Neymeyr. A fast polygon inflation algorithm to compute the area of feasible solutions for three-component systems. i: concepts and applications. *Journal of Chemometrics*, 27(5):106–116, 2013.

[6] M. Sawall and K. Neymeyr. A fast polygon inflation algorithm to compute the area of feasible solutions for three-component systems. II: Theoretical foundation, inverse polygon inflation and FAC-PACK implementation. Technical Report, Universit¨at Rostock and LIKAT Rostock, 2014.

Simultaneous determination of cationic dyes from aqueous solution using first-order derivative spectrophotometry

Maryam Bagtash, Javad Zolgharnein*

Department of Chemistry, Faculty of Science, Arak University, Arak, Iran

E-mail address: bagtash@yahoo.com

Abstract

Derivative spectrophotometry has been introduced for the resolution of overlapping peaks. Derivative method has been widely used to enhance the signal and resolve the overlapped peak-signals due to its advantages in differentiating closely adjacent peaks, and identifying weak peaks obscured by sharp peaks. Therefore, a first-order derivative spectrophotometric method is developed for the simultaneous determination of Brilliant Green and Crystal Violet in binary solutions. The Savitzky-Golay convolution method is used for calculating the first derivative absorption spectra of binary mixtures. The linear concentration range and limits of detection for the simultaneous determination of BG and CV are found to be: 1–20, 1–15 mg/L, 0.3 and 0.5 mg/L, respectively.

Keywords: Brilliant Green, Crystal Violet, Derivative spectrophotometry

1. Introduction

Brilliant Green (BG) and Crystal Violet (CV) are among the cationic dyes which have been shown to have harmful effects on human beings. BG dye is used for various purposes, such as in the production of cover paper in the paper industry; for biological staining; as a dermatological agent; in veterinary medicine; and also as an additive, to inhibit the propagation of mold, intestinal parasites and fungus [1]. CV is a well-known dye used for dyeing cotton, silk, paper, bamboo, weed, straw and leather [2]. Therefore, it is essential to monitor these dyes in water. To date, different analytical methods for the determination of dyes have been published, which need the intensive sample pretreatment. However, these separation processes are not simple and the simultaneous determination of some dyes by traditional spectrophotometry techniques is difficult because of a serious overlapping of their absorption spectra. Therefore, the derivative spectrophotometric technique has been used, due to its facile and simple approach for simultaneous determination of dyes in mixtures [3].

2. Materials and methods

2.1. Apparatus and Reagents

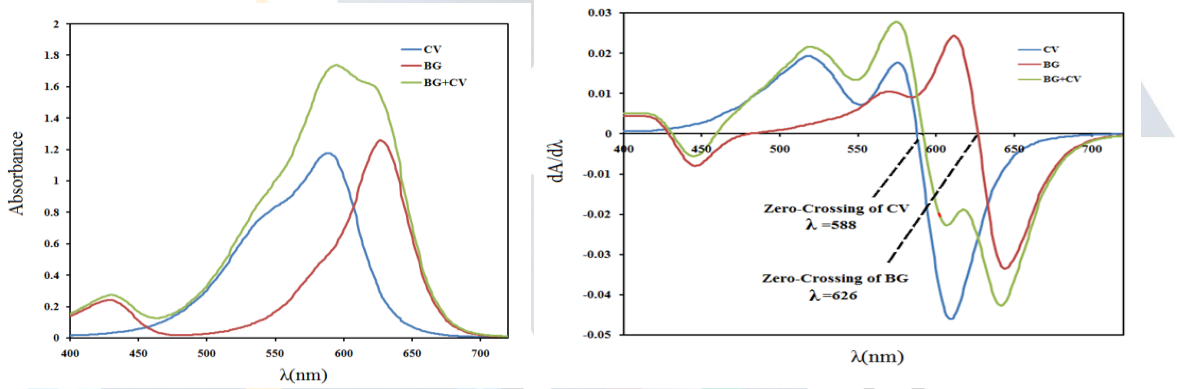
A UV–Vis spectrophotometer (Analytikjena SPECORD250) was used for analysis of BG and CV dyes. All chemicals were used with analytical-grade from Merck (Darmstadt, Germany) and doubly distilled water was used throughout the work.

3. Results and discussion

The simultaneous determination of BG and CV in binary solutions

For the simultaneous analysis of BG and CV dyes in a mixture, the binary mixture of these dyes were prepared and zero order absorption spectra were recorded (Fig. 1A). In binary mixtures, the absorption spectra of BG and CV dyes overlapped and showed interference between the zero order spectra of BG and the CV dyes, so their concentrations could not be simultaneously determined by direct absorbance measurements (Fig. 1A). The concentration of BG and CV in binary mixtures can be rightly determined by a first order derivative spectrophotometry (Fig. 1B). The first

derivative is the rate of change of absorbance against wavelength. It starts and finishes at zero, and passes through zero at the same wavelength as kmax of the absorbance band, with first a positive and then a negative band, and with the maximum and minimum at the same wavelengths as the inflection points in the absorbance band. The derivation of zero order spectra can lead to a separation of overlapped signals, and reduce the effect of spectral background interference caused by the presence of other compounds in a sample [4]. Hence, BG and CV dye concentrations in binary mixtures were determined by measuring the absorbance signal at the first order derivative wavelength. According to the zero-crossing derivative method it is necessary that zero-crossing wavelengths do not change with the varied concentrations of related species [4]. To evaluate the condition, changes in the pre-mentioned zero-crossing wavelengths for BG and CV were tested in the presence of different concentrations of another species. The calibration graphs for the determination of BG in the presence of CV were constructed by measuring derivative amplitudes at the zero crossing points of CV (588 nm). Similarly, calibration graphs were constructed by the measurement of derivative amplitudes at the zero crossing point of BG (626 nm) (Fig. 1B).



(A) (B)

Fig 4: (A) Absorption spectra of: CV (7 mg/L); BG (14 mg/L); and their mixture spectrum and (B) First derivative spectra of: CV, BG and their mixture spectrum.

In order to test the mutual independence of the analytical signals of BG and CV, calibration graphs were constructed for standard solutions containing various amounts of BG in the presence of 5 mg/L of CV. A similar procedure for standard solutions containing various amounts of CV in the presence of 10 mg/L of BG was performed. The similarity observed between regression equations of pure dye and the mixed solution suggested no interferences in the estimation of one dye in the presence of the other. The regression equations and coefficients of determinations of obtained calibration graphs are given in Table 1. The derivative amplitudes measured at 588 nm were found to be independent of the concentration of CV. Similarly, derivative amplitudes measured at 626 nm were found to be independent of the concentration of BG.

Table 1. Statistical analysis of the determination BG and CV in standard solutions by first-derivative spectrophotometry

Sample	Composition of solutions (mg/L)	Composition of solutions (mg/L)	Regression equations (at 588 nm for BG; at 626 nm for CV)	R^2	LOD (mg/L)
	BG	CV			
۱	1-20	0	D = 0.001C + 0.0012	0.997	0.3
۲	1-20	5	D = 0.001C + 0.001	0.998	0.3
۳	0	1-15	D = -0.002 + 0.001	0.999	0.5
۴	10	1-15	D = -0.002 + 0.007	0.999	0.5

The validity of the proposed method was determined in several synthetic binary mixtures containing BG and CV.

Table 2. Simultaneous determination of BG and CV in synthetic binary mixtures

Taken (mg/L)	Taken (mg/L)	Found (mg/L)	Found (mg/L)	Recovery %	Recovery %
BG	CV	BG	CV	BG	CV
8	10	7.8	10.3	103.0	97.5
4	15	4.1	15.4	102.7	102.5
6	9	5.9	9.3	103.3	98.3
14	5	14.3	5.1	102.0	102.1
12	3	11.7	3.1	103.3	97.5
10	7	10.2	6.8	97.1	102.0
5	17	4.9	17.4	102.3	98.0
Mean				99.71	101.97

However, recovery studies confirmed that the simultaneous determination of BG and CV contents in binary solutions could be accurately accomplished by taking the first-order derivatives of the spectrophotometric method. The results obtained are given in Table 2.

References

[1] M. Ghaedi, G. Negintaji, H. karimi , F. Marahel, Solid phase extraction and removal of brilliant green dye on zinc oxide nanoparticles loaded on activated carbon: New kinetic model and thermodynamic evaluation, J. Ind. Eng. Chem. 2014; **20;** 1444–1452.
(2013)

[2] Ofomaja AE, Ukpebor EE, Uzoekwe SA, Biosorption of Methyl violet onto palm kernel fiber: Diffusion studies and multistage process design to minimize biosorbent mass and contact time, *biomass bioenerg*. 2011; **35**; 4112-4123.

[3] Gao J-F, Zhang Q, Su K, Wang J-H, Competitive biosorption of Yellow 2G and Reactive Brilliant Red K-2G onto inactive aerobic granules: Simultaneous determination of two dyes by first-order derivative spectrophotometry and isotherm studies, *Bioresour. Technol.*, 2010; **101**; 5793–5801.

[4] Turabik M, Adsorption of basic dyes from single and binary component systems onto bentonite: Simultaneous analysis of Basic Red 46 and Basic Yellow 28 by first order derivative spectrophotometric analysis method, *J. Hazard. Mater*. 2008;**158**; 52–64.

Design of potent aromatase inhibitors through pharmacophore hybridization: QM/MM and docking study

Parvin Asadi [a,*] and Ghadamali Khodarahmi [a]

[a] *Department of Medicinal Chemistry, School of Pharmacy and Pharmaceutical Sciences, Isfahan University of Medical Sciences, Isfahan, 81746-73461, I.R.Iran*

E-mail address:asadi@pharm.mui.ac.ir

Abstract

Novel hybrid derivatives incorporating benzofuran, imidazole and quinazolinone pharmacophores were designed by molecular hybridization approach in such a way to arrange themselves in a symmetrical shape, similar to aromatase inhibitors. Subsequently, the binding of these novel hybrid compounds to the aromatase enzyme have been investigated in a docking procedure applying a combined quantum mechanical/molecular mechanical (QM/MM) method. The QM/MM calculation was performed on the reference structures, to obtain more accurate atomic charges on the ligand atoms in the active site. The results indicated that hybrid compounds adopted properly within the aromatase binding site, suggesting that they could be potential inhibitors of aromatase.

Keywords: QM/MM Docking; Pharmacophore hybridization; Benzofuran; Imidazole; Quinazolinone

1. Introduction

Breast cancer is the most frequently diagnosed cancer in women. Approximately two-thirds of breast cancer tumors are hormone-dependent and require estrogens to grow. One approach in treating hormone-dependent cancer involves interfering with endogenous hormone production. Aromatase, also known as estrogen synthase, has always been considered the most promising target for the endocrine treatment of breast cancer [1].

Computational docking of small molecules to a macromolecular target is a widely method used in rational drug design and drug discovery. Given the biological and pharmaceutical significance of molecular docking, considerable efforts have been directed towards spiriting use of this method [2]. Although molecular docking simulation of proteins, is a fast and inexpensive method for descriptions of the ligand-protein interactions, but poses some difficulties. Two of the most important limitations of conventional docking are assuming non protein flexibility upon ligand binding and using force field based fixed dielectric charges for both protein and ligand atoms [3]. Therefore, to increase accuracy in docking result, it is reasonable to expend additional effort to improve the quality of the charge.

2. Results and discussion

2.1. Molecular docking

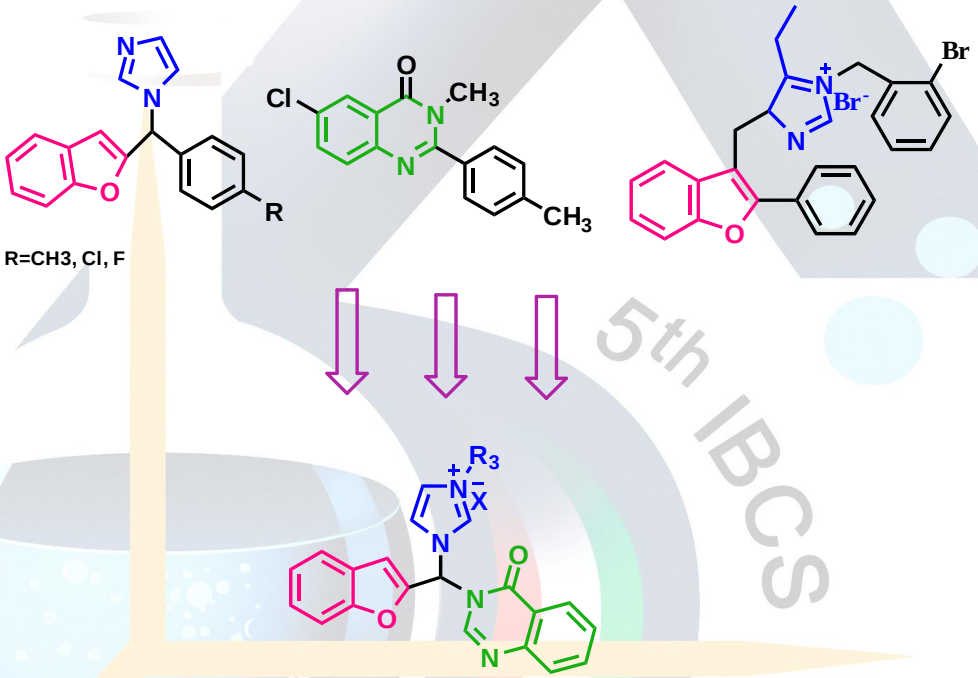
Molecular docking was performed by AutoDock4 software to elucidate the binding mode of aromatase with novel hybrid compounds and the best conformers were selected according to the lower docked free energy and top-ranked cluster to perform docking analysis with AutoDock Tools and PyMOL.

2.2. QM/MM Methodology

For QM/MM calculations, we employed the Gaussian 09 quantum chemistry package. "Our own N-layered integrated molecular orbital and molecular mechanic" abbreviated as ONIOM method implemented in Gaussian 09 was used for QM/MM calculations.

2.3. Design of new inhibitors

Considering potent cytotoxic activities of hybrid benzofuran-imidazolium and quinazolinone derivatives on breast cancer cell line (MCF-7), novel hybrid derivatives incorporating benzofuran, imidazole and quinazolinone pharmacophores were designed by molecular hybridization approach.



2.4. Evaluation of the designed inhibitors

Energetic and structural properties of desiened compounds were investigated as aromatase inhibitor using molecular docking and QM/MM calculations. Due to the positive charge on the imidazole ring of the designed ligands and also the presence of Heme iron in the active site of the enzyme, QM/MM calculations was performed, in which partial charges of the ligand were re-fitted according to the polarized active site environment of the enzyme to increase the accuracy of the docking results. According to the above results, the designed ligands formed a three branched structures in the protein environment which accommodated well into the active site. Compound 1 which yielded the highest ΔG_b (-11.47) and the best performing *K*i (48nM) values was assuming to be the best ligand (Fig. 1).

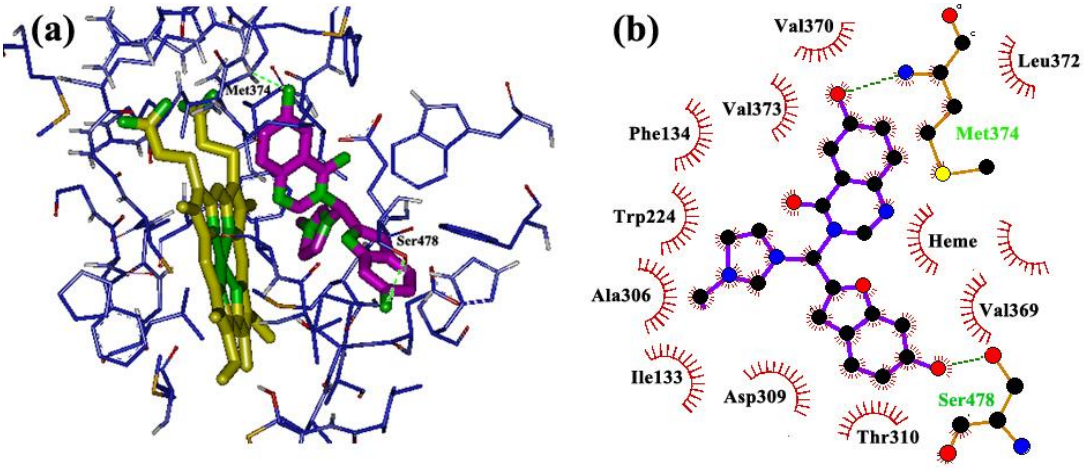


Fig. 1. Binding modes and hydrogen bonds interactions of compound 1 in aromatase active site respectively (a) 3D structure, (b) 2D structure

References

[1] Briest S, Davidson NE. Aromatase inhibitors for breast cancer. *Rev. Endocr. Metab. Disord*. 2007; **8**;215-228.

[2] Taylor RD, Jewsbury PJ, Essex JW. A review of protein-small molecule docking methods. *J. Comput-Aid. Molecul. Design*. 2002;**16**:151–166.

[3] CHO A, Guallar V, Berne B, Friesner R. Importance of Accurate Charges in Molecular Docking: Quantum Mechanical/Molecular Mechanical (QM/MM) Approach. *J. Comput. Chem*. 2005;**26**:915–931..

Multivariate calibration methods for simultaneous multicomponent determination of of two synthetic dyes in fruit juice samples

Rouhollah Khani[a*], Reza Rahmanian[b]

[a] *Department of Chemistry, College of Sciences, University of Birjand, Birjand, 97179–414 Iran.*

[b] *Department of Chemical Technologies, Iranian Research Organization for Science and Technology (IROST), 33535111, Tehran, Iran*

E-mail address: Rouhollah.khani@birjand.ac.ir

Abstract

In the present study a method for the simultaneous determination of sunset yellow (SY) and tartrazine (TT) in fruit juice samples has been developed. The method uses of β-cyclodextrin (β-CD) complexes as a new method for generation of three way data, combined with second-order calibration methods for quantification of a binary mixture of target synthetic dyes. At first, the basic experimental parameters affecting the formation of inclusion complexes between target analytes and β-CD were investigated and optimized. Then under the optimum conditions, parallel factor analysis (PARAFAC) and bilinear least squares/residual bilinearization (BLLS/RBL) were applied for deconvolution of trilinear data to get spectral and concentration profiles of SY and TT as a function of β-CD concentrations. The proposed method was validated by comparison with a reference method based on highperformance liquid chromatography -ultra violet detection (HPLC-UV), and no significant differences were found between the reference values and the ones obtained with the proposed method.

Keywords: Bilinear least squares/residual bilinearization; Beta–cyclodextrin; Parallel factor analysis; synthetic dyes; Fruit juices samples

1. Introduction

Food dyes are often added to foodstuffs and drinks in order to supply, intensify or restore their colour to create the desired coloured appearance [1]. Synthetic dyes are widely used as they show several advantages compared with natural dyes such as high stability to light, oxygen and pH, colour uniformity, low microbiological contamination and relatively lower production costs. However, in certain quantities they are harmful to human health [2], hence supervising of synthetic dyes in high consumption products such as fruit juice becomes an indispensable task. βcyclodextrin (β-CD) has the peculiar 'interior hydrophobic, exterior hydrophilic' structure forming a 1:1 or 1:2 inclusion complex with guest molecules, thus the physical, chemical and biochemical characters of guest molecules are modified [3]. Properties of β-CD affect spectroscopic properties, absorption and fluorescence spectra and especially molar absorptivities of SY and TT. The addition of β-CD to the solution mixture of target analytes and the subsequent change of the absorbance in the wavelength range help to have a series of spectra, which can make a bilinear data matrix.

By changing the initial concentration of binary mixtures of target dyes and addition of β-CD, a three dimensional data matrix can be generated.

2. Results and discussion

2.1. Effects of temperature and time on β-CD including SY and TT

Effect of temperature on β-CD including SY and TT was investigated. The results showed that in the inclusion process of SY and TT, the absorbance intensities gradually decrease, as the temperature increases. The maximum absorbance intensities were observed at room temperature (25 ± 1 °C).

For inclusion of SY and TT on β-CD, sonication was applied in the present study. The sonication time required for attaining the inclusion equilibrium depended on the system. Based on the obtained results, the room temperature (25 °C) and the time of 40 min were chosen for further experiments.

2.2. Prediction of SY and TT in validation samples

The performance of second-order calibration methods based on BLLS/RBL and PARAFAC were initially tested for determination of SY and TT in validation samples [4]. For BLLS/RBL model, the number of calibrated solutes and the number of expected interferents were set as 2 and 0, respectively. Number of interferents was set as 0, because validation samples were free from any interferents. For PARAFAC analysis, core consistency criterion [5], was adopted to find the optimum number of factors leading to optimum data deconvolution and the optimum number of factors was 2. The final results are shown in Figure 1.

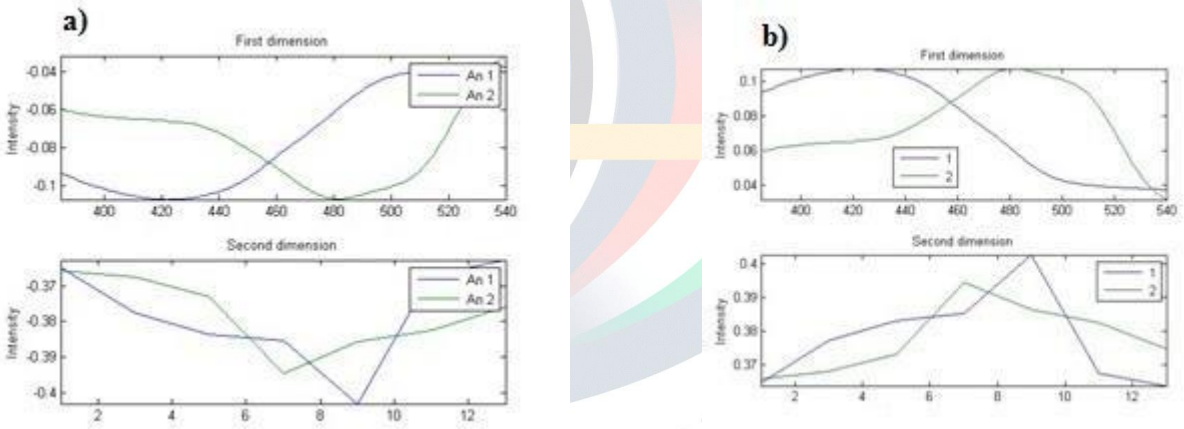


Figure 1. a) BLLS/RBL and b) PARAFAC outputs for spectral deconvolution of SY and TT in one of the calibration samples.

Validation samples show 3-D plots similar to those of calibration samples, therefore satisfactory results in the prediction step by applying BLLS/RBL and PARAFAC were reasonably expected (Figure 2). The prediction results for the validation set were reasonably good, leading to a mean recovery of 98% and

100.45% for SY and TT, respectively. These parameters indicate that the proposed method is a feasible methodology for achieving the secondorder advantage in cases of sample components with similar spectra.

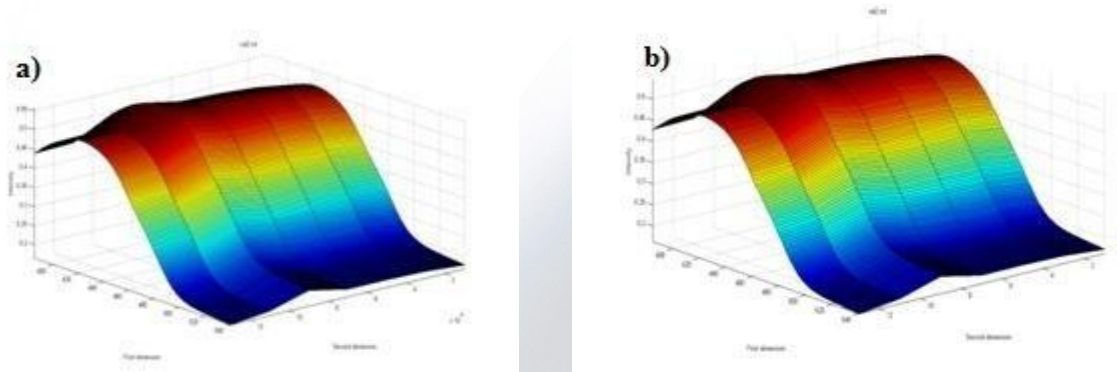


Figure 2. 3-D plots of: (a) calibration sample and (b) validation sample.

2.3. Simultaneous quantification of SY and TT in real samples

Determination of SY and TT in fruit juices samples was carried out using BLLS/RBL and PARAFAC. Even in the presence of interferents, both of two algorithms were satisfactory for delivering the spectral profiles of phenolic acids, and this would explain their high prediction power.

In order to validate the performance of the proposed method, the samples were also analyzed by the HPLC-UV detection method. In the proposed method, determination of SY and TT in the real samples (fruit juices samples) was carried out in the presence of β-CD under the optimum conditions, but in the HPLC method, the real samples were directly analyzed, so then some observed partial differences in the results are quite reasonable. The results of the t-test at appropriate confidence level (95%) revealed no significant differences between the reference method (HPLC) and the strategy described in the present report.

References

[1] Schenone AV, Culzoni MJ, Marsili NR, Goicoechea HC. Determination of tartrazine in beverage samples by stopped-flow analysis and threeway multivariate calibration of non-linear kinetic-spectrophotometric data. *Food Chem*. 2013; **138**: 1928–1935.

[2] Silva MLS, García BQ, Lima JLFC, Barrado E. Voltammetric determination of food colorants using a polyallylamine modified tubular electrode in a multicommutated flow system. *Talanta* 2007; **72**: 282–288.

[3] Zhang G, Shuang S, Dong C, Pan J. Study on the interaction of methylene blue with cyclode derivatives by absorption and fluorescence spectroscopy. *Spectrochim. Acta Part A* 2003; **59**: 2935–2941.

[4] Olivieri AC, Escandar GM, Mun˜oz de la Pen˜a A. Second-order and higher-order multivariate calibration methods applied to non-multilinear data using different algorithms. *Trends Anal. Chem.* 2011; **30**: 607–617.

[5] Bro R, Kiers HAL. A new efficient method for determining the number of components in PARAFAC models. *J. Chemom.* 2003; **17**: 274–286.

Estimating complicated baselines in analytical signals by iterative training of Bayesian regularized artificial neural networks

Ahmad Mani-Varnosfaderani[*1], Atefeh Kanginejad [1]

[1] Department of Chemistry, TarbiatModaresUniversity, Tehran, Iran.

E-mail address: a.mani@modares.ac.ir

Abstract

A new baseline correction method based on Bayesian regularized artificial neural networks (BRANN) was developed. Herein, an iterative subtracting algorithm based on BRANN technique has been developed for the first time for removing different types of baselines in analytical chemistry. In BRANN approach, the density function for the network's weights will be updated according to Bayes' rule. The regularization methodology inside the BRANN algorithm made it an interesting technique for unbiased estimation of complicated baselines. For comparison, some standard methods were applied to remove baseline drift such as adaptive iteratively reweighted penalized least squares (airPLS) and iterative polynomial fitting (iPF). Both airPLS and iPF techniques have some drawbacks, for example these methods are sensitive to outliers and have some parameters need to be optimized before implementation on data. The suggested algorithm in this work does not need to adjust parameters before implementation and is less sensitive to outliers. Calculation of both projection difference resolution (PDR) and signal to noise ratio (SNR) for simulated data revealed the superiority of BRANN over airPLS and iPF algorithms. Experimental results on real data such as GC-MS demonstrate the efficiency of the proposed algorithm in this work.

Keywords: baseline correction**,** Bayesian regularized artificial neural networks (BRANN), iterative polynomial fitting, adaptive iteratively reweighted penalized least squares

Introduction

The baseline drift is usually one of the main issues in chromatography. The signals of analytical instruments commonly consist of chemical information, baseline and random noises. However, the existence of the baseline can negatively affect qualitative or quantitative analytical results. The simplest method to remove baseline drift is fitting a straight line using the first and last points of the collected data. The fitted line is then subtracted from whole data points to produce baseline free signals. Feng Gan et al [1] implemented iterative polynomial fitting strategy with automatic threshold. But the performance of this method is poor in low signal-to-background ratio. Zhi-Min et al [2] presented adaptive iteratively reweighted penalized least squares (airPLS). This method is based on the penalized least squares methodology and need adjustment of penalty ratio (ƛ). This means that ƛ should be specified by the user and this needs peak detection before background correction. This makes airPLS methodology not as robust as expected. Eilers and Boelens [3] proposed an AsLS (Asymmetric Least Squares) method which does not require peak finding, but a new parameter named "P" was introduced to set weights asymmetrically. Generally in both airPLS and AsLS techniques finding ƛ and P is time consuming and controversy. In the present contribution the Bayesian regularized artificial neural networks (BRANN) has been used for removing background from one and two dimensional data such as GC, LC, GC-MS and HPLC-DAD data.

2-Bayesian regularized artificial neural network (BRANN) for background correction

In this section we report how a BRANN model can be used for removing background from simulated and real signals. Some different signals were simulated for benchmarking purposes. For a triangular baseline, Figure 1 (a) shows the estimated baseline based on iPF technique. As can be seen, the baseline for the middle part of the signal is not thoroughly estimated using iPF algorithm. Figure 1(b) shows baseline estimation based on airPLS algorithm. The estimated baseline using this algorithm is partially better than iPF but problem of the middle part of the signal still remains. Figure 1(c) illustrates the estimated baseline using BRANN algorithm. As can be seen a partial triangular baseline has been estimated using this technique.

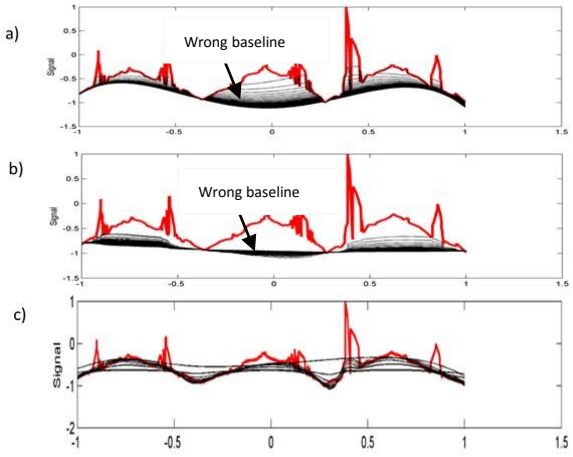


Figure 6. Comparison of different baseline correction methods for removing a triangular baseline (a) iterative polynomial fitting; (b) airPLS;(c) BRANN

Table1 summarizes the statistical parameters of airPLS, iPF and BRANN algorithms for baseline correction in a real GC-MS data. As can be seen in this table the performance of BRANN is better than airPLS and iPF algorithms in terms of PDR and signal to noise ratio.

Table1.Comparison of BRANN with AirPLS and Iterative polynomial fitting.

	PDR	Signal to Noise
Data	5.34	254.7087
Baseline correction using BRANN	5.47	2.1533×10^{1}
Baseline correction using AirPLS	5.30	1.0964×10^{4}
Baseline correction using iPF	3.81	3.7322×10^{3}

As illustrated above, after comparing with several popular baseline correction methods such as airPLS and iterative polynomial fitting, the proposed algorithm in this work, BRANN, reveals consistency and robustness for baseline correction and peak detection in analytical signals. This algorithm increases signal to noise and PDR values compare to iterative polynomial fitting and airPLS algorithms.

3. References

[1] Gan F, Ruan G, Mo J., Chemom.Intell.Lab.sys. 2006 ; **82**:59-65.

[2] Zhang ZM, Chen S, Liang YZ., Analyst. 2010 ;**135**:1138-1146.

[3] Eilers PHC., Anal. Chem. 2003 ;**75**:3631-3636.

QSAR modeling of Sphingosine kinase inhibitors using multivariate adaptive regression spline and projection based regression techniques

Ahmad Mani-Varnosfaderani*[1], Marzieh Sadat Neiband[2], Ali Benvidi[2]
[1] Department of Chemistry, Tarbiat Modares University, Tehran, Iran.
[2] Department of Chemistry, Yazd University, Yazd, Iran.
E-mail address: a.mani@modares.ac.ir

Abstract

In this work, the inhibitory activity of Sphingosine kinase (SphK) is predicted in terms of quantitative structure activity relationship (QSAR) models. Different algorithms were utilized to choose the best variables among large numbers of descriptors. The present contribution paves the way for comparing the performances of projection based modeling (PLS,PCR) and local modeling (MARS) techniques. The QSAR models in this work were developed with the aid of multivariate adaptive regression spline (MARS), stepwise multiple linear regression (Stepwise-MLR), genetic algorithm-partial least squares (GA-PLS), partial least squares (PLS), principle component regression (PCR) and successive projection algorithm (SPA-MLR). The models were validated and tested through the use of the external prediction set of compounds and leave-five-out cross validation methods. The statistical parameters of R^2 and root mean square error (RMSE) indicated that MARS is superior for modeling the inhibitory activity of SphK inhibitors over the other methods. The robustness and accuracy of the QSAR models were confirmed by the satisfactory statistical parameters ($R^2_{training}$ = 0.95365 and R^2_{L5O} = 0.9213) and low standard error values ($RMSE_{training}$ = 0.1525 and $RMSE_{L5O}$ =0.2483).

Key words: Sphingosine kinase inhibitor, Quantitative structure activity relationship, Multivariate adaptive regression spline

Introduction

The sphingolipids are a group of amino or amido di-hydroxylated lipids [1]. In addition to their structural roles in the cell membrane, it has become increasingly clear that certain sphingolipids can act as signaling molecules. Among the sphingolipids, ceramide, sphingosine (Sph) and sphingosine 1-phosphate (S1P) have received considerable attention due to the roles they are reputed to play in a broad array of biological processes including apoptosis, mitogenesis, lymphocyte migration, radio and chemo-sensitization and angiogenesis [2]. QSAR searches information relating chemical structure to biological activity by developing a mathematical model. Building of a QSAR model begins with calculating theoretical parameters or selecting structural features for the compounds involved. Variable selection techniques have become important for producing useful predictive models. A suitable feature selection method ensures the model stability and the consistency of relationship between the descriptors and biological activity [3].In order to make sure that the most important descriptors have been selected; cross-validation technique was used in this work. In this method, the data set was divided into several subsets, and variable selection process was performed for different combinations of these subsets. Then the most frequent descriptors in models were selected as most important variables for describing the inhibitory effect.

Results and discussion

The data set in this work consists of 44 derivatives of Sphingosine kinase (SphK) inhibitor and was taken from the article published by Yihong Li et al. [4]. In the present contribution, six different variable selection techniques have been used for describing and predicting the SphK inhibitory activities of a series of Sphingosine kinase amide derivatives. The results of the stepwise MLR, GA_PLS, PLS, PCR, SPA-MLR and MARS [2] models are listed in Table 1.

Table1. Statistical parameters of different models obtained for the training and L5O-CV procedures

Method	**Training**		**L5O**		**No. Comp/Var.**
	R_{2a}	RMSE	R^2	RMSE	
Stepwise-MLR	0.9319	0.5719	0.8794	0.7675	10
PLSR	0.6998	1.2408	0.4780	1.2691	5
GA-PLS	0.7770	1.0494	0.6731	1.0149	5
PCR	0.6975	1.2043	0.4534	1.6365	7
SPA-MLR	0.9283	0.6043	0.7923	0.9044	12
MARS	0.95365	0.1525	0.9213	0.2483	10

[a] All R^2 are adjusted coefficient regression.

Investigation of the results reveals that the MARS algorithm selects the best variables for predicting the inhibitory effect of SphK inhibitors. Generally, this contribution presents a simple and interpretable QSAR model based on multivariate splines which can help the medicinal chemists for designing novel Sphingosine kinase inhibitors. The results revealed that the best variables for describing the inhibition mechanism of Sphingosine kinase were ATS7e, SHP2, Ae, VEA2, DISPm, Qmean, MATS5e and Nn descriptors. A complete description of these molecular descriptors can be found in literature [6].

References

[1] Hannun YA, Obeid LM. Principles of bioactive lipid signalling: lessons from sphingolipids. *Nat. Rev. Mol. Cell. Biol*. 2008, **9**: 139-150.
[2] Spiegel S; Milstien S. The outs and the ins of sphingosine-1-phosphate in immunity.*Nat.* Rev. *Immunol*. 2011, **11** : 403-501.
[3] Ghafourian T, Cronin MTD, The impact of variable selection on the modeling of oestrogenicity. *SAR, QSAR. Environ. Res.* 2005, **16** :171–190.
[4] Darin JG, Yihong Li, Matthew LB, Xiaoshan M, Mike JS, Malgorzata W, Xiaodong W, Richard C, Sheere J, Mario C ,Alan CC, Shawn J, Brendon F, Shyun L, Shanling S, Mariwil W, Holger W, Guifen X, Timothy JC, Matthew P, Kurt M, Karen R ,Joanna S, Angela C, Nigel W, Frank K, Zhulun W. Structure guided design of a series of sphingosine kinase (SphK) *inhibitors Bioorganic & Medicinal Chemistry Letters*.2013, **23** : 4608– 4616.
[5] Friedman JH. Multivariate adaptive regression splines. *Annal. Stat* .1991, **19**: 1–141.
[6] Todeschini R., Consonni V.,Handbook of molecular descriptors, Wiley-VCH, Weinheim, 2008.

Study of anti-cancer activity of parthenin derivatives using support vector regression and molecular docking

Zahra Garkani-Nejad

Department of Chemistry, Faculty of Science, Shahid Bahonar University of Kerman, Kerman, Iran

E-mail address: z_garkani@uk.ac.ir

Abstract

A quantitative structure–activity relationship (QSAR) study has been done on the anti-cancer activity of parthenin derivatives against three human cancer cell lines, SW-620, DU-145, and PC-3. QSAR models were based on multiple linear regression (MLR) and support vector regression (SVR). First, stepwise MLR was employed as a descriptor selection procedure. Then selected descriptors were used as inputs for SVR model. Comparison of the results indicates that the SVR method shows better predictive power than other method. In order to show the effect of hydration energy (HE) on anti-cancer activity, docking study of parthenin derivatives with Nf-jB transcription factor has been done.

Key words: Quantitative structure–activity relationship (QSAR); Anti-cancer activity; Docking; Parthenin derivatives

Introduction

In recent years, the anti-cancer property of various sesquiterpenes has attracted a great deal of interest and extensive research works have been carried out to characterize the anti-cancer activity and the molecular mechanisms of sesquiterpenoids [1]. Sesquiterpene lactones (SLs) are the active constituents of many medicinal plants from the Asteraceae family. Parthenin is a sesquiterpene lactone (SL) that isolated from *Parthenium hysterophorus* L [2]. has found interest due to its medicinal properties like anticancer activity [3]. Several novel derivatives of parthenin have been synthesized by the dipolar cycloaddition using various dipoles such as benzonitrile oxides, nitrones, and azides. Majority of the compounds exhibited improved anti-cancer activity compared to the parthenin, when screened for their in vitro cytotoxicity against three human cancer cell lines including SW-620, DU-145, and PC-3. Sesquiterpene lactones (SLs) are potent anti-inflammatory substances. The anti-inflammatory effect of these compounds could be partly explained by the inhibition of the transcription factor of NF-jB [4]. At the present work, relationship between the structure of parthenin derivatives and their anti-cancer activities against three human cancer cell lines, SW-620, DU-145, and PC-3 has been considered using some chemometrics methods.

Results and discussion

The half maximal inhibitory concentration (IC50) values of parthenin derivatives against three human cancer cell lines, SW-620, DU-145, and PC-3 are taken from the literature. Molecular descriptors were generated using DRAGON software. A stepwise MLR procedure was used for model development. For regression analysis, data set as divided into two groups of training and test sets for each cell lines. G2v, H3u and Hydration energy (HE) are three common descriptors that entered in the best models for three cell lines.

SVR is the most common application form of SVM that is a powerful technique for predictive data analysis with many applications to varied areas of study. In this work, the SVR evaluations were carried out using the SVM toolbox in CLEMENTINE software. Selected descriptors using MLR models were employed as inputs. After that, the kernel function should be determined, which represents the sample distribution in the mapping space. In this work, the RBF (radial basis function) kernel was chosen.

Molecular docking was carried out by AutoDock 4.3 to understand the detailed binding model for the active site of the receptor with its ligands. For determining the appropriate binding conformations of studied compounds and check the main factors affecting the activity, docking study was performed for parthenin derivatives with the most anti-cancer activity. In order to show the effect of HE descriptor, the hydrogen bond between Transcription factor Nf-jB with derivatives of parthenin has been investigated. One of the most active derivatives of parthenin has been docked with Nf-jB factor. Nf-jB expression is completely inhibited by this compound. According to Figure 1, this molecule has three hydrogen bonds with Nf-jB transcription factor. This molecule has two hydrogen bonds with Lys37 that one bond is through oxygen atom of nitrile oxide ring and another bond is through of oxygen atom of cyclopentenone ring. The other hydrogen bond of this molecule is through hydroxyl group of cyclopentenone ring with Glu39. Docking of this compound with Lys37 and Glu39 leads to blocking and alkylation Cys38. Thus, docked compound inhibits Nf-jB factor completely.

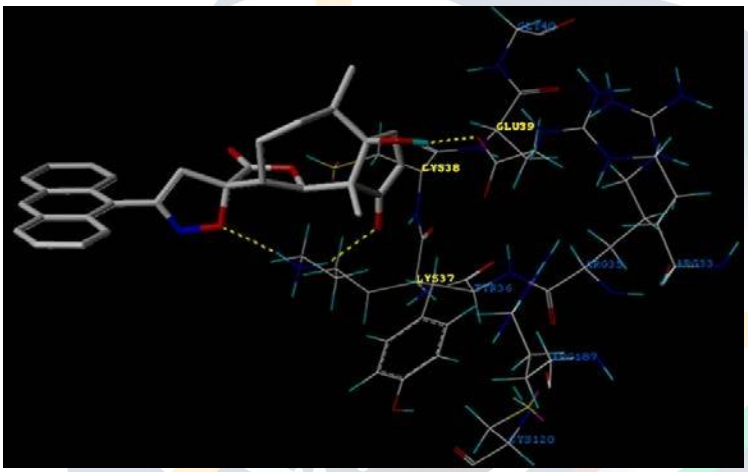


Figure 1

References

[1] Gershenzon J, Dudareva N (2007) The function of terpene natural products in the natural world. Nat Chem Biol 3:408–414

[2] Picman AK (1986) Biological activities of sesquiterpene lactones. Biochem Syst Ecol 14:255–281

[3] Fraga BM (2006) Natural sesquiterpenoids. Nat Prod Rep 23:943–972

[4] Garcı´a-Pin˜eres AJ, Lindenmeyer MT, Merfort I (2004) Role of cysteine residues of p65/NF-jB on the inhibition by the sesquiterpene lactone parthenolide and N-ethyl maleimide, and on its transactivating potential. Life Sci 75:841–856

In silico prediction of cutaneous penetration rate of some chemicals from their molecular structural descriptors

Seyedeh Mozhgan Behgozin*, Mohammad Hosein Fatemi and Kobra Samghani
Chemometics Laboratory, Faculty of Chemistry, University of Mazandaran, Babolsar, Iran
∗ Corresponding author. Tel.: +98 1125242931; fax: +98 11125342350
E-mail addresses:
m.behgozin@stu.umz.ac.ir

Abstract

The cutaneous penetration rate of some hormones and organic solvents to the stratum corneum was modeled and estimated by using interpretable molecular descriptors. Ten geometric and electronic descriptors were selected by stepwise multiple linear regression (MLR) as independent variables. To develop quantitative structure-activity relationship (QSAR) models, the methods of MLR, and artificial neural network (ANN), and random forest (RF), and support vector machine (SVM), and CORAL were used. Among these methods, the statistical parameters of ANN model were superior to the others. Robustness and reliability of the constructed ANN model were evaluated by using the leave-two-out cross-validation method, which produces the statistics of Q2=0.960 and SPRESS=0.059 for ANN model. The results of this study indicated the ability of developed QSAR models in the prediction of dermal penetration rate of various chemicals from their calculated molecular descriptors.

Key words: Quantitative-structure activity relationship; Dermal penetration rate; Artificial neural network; Multiple linear regressions; Molecular descriptor.

Introduction

The percutaneous/dermal absorption process is a global term, which describes the passage of molecules across the skin [1]. The penetration of molecules through the skin is affected mainly from its chemical structure and referred to as flux (fl) that can be determined experimentally either in vitro by measuring the diffusion rate across an excised layer of stratum corneum (SC), or in vivo, by measuring losses of the chemicals that had been hermetically applied on a defined area of skin over a certain period of time. Recently there has been much interest in the possibilities available to predict the dermal absorption to avoid unnecessary and costly in vivo and in vitro testings. One theoretical method to estimate chemical activities is quantitative structure-activity relationship (QSAR) methodology. In this method, activities or properties of chemicals were correlated to their relevant molecular structural features (descriptors). In the present work, we try to develop some QSAR models to predict the in vitro flux of some chemicals through human skin, which measured by flow-through or static cells methods .In these studies (multiple linear regression) MLR was applied as linear and (artificial neural network) ANN, and random forest (RF), and support vector machine (SVM), and CORAL and non-linear feature mapping techniques.

Results and discussion

The structures of 87 chemicals taken from literature were drawn and optimized with HyperChem software (version 7) using the AM1 semi-empirical method. Multiple linear regression and artificial neural network, RF, SVM, and CORAL were applied for modeling the flux of some chemicals. In these methods, the stepwise MLR method was used selected descriptors as inputs. Statistical parameters of these models are shown in Table 1.

Table 1- statistical results of models.

Model	Set	R	F	SE	Q^2	SPRESS
MLR[1]	-	0.572	74.1	0.936	-	-
MLR	Training	0.847	15.7	0.118	0.553	0.082
	Test	0.795	6.4	0.120		
SVM	Training	0.881	100.7	0.080	0.665	0.079
	Test	0.750	10.0	0.117		
RF	Training	0.817	120.2	0.045	0.678	0.070
	Test	0.646	25.3	0.067		
CORAL	Training	0.708		0.112	0.725	0.068
	Internal test	0.547		0.209		
	External test	0.845		0.081		
ANN	Training	0.902	292.2	0.080	0.960	0.059
	Internal test	0.894	27.8	0.110		
	External test	0.831	19.8	0.119		

1 [1]

In the present work, QSAR methodology was used to construct a quantitative relation between the stratum corneum logarithmic of flux of some chemicals and their calculated molecular descriptors. All descriptors that are used to develop these QSAR models encode features of molecules that influences on these interactions and govern the skin penetration of chemicals. Descriptors that were used in this work, can encode topological and electronical aspects of interested molecules. The superiority of the ANN model [2], over other models, indicate that there are some nonlinear relation between selected molecular descriptors and stratum corneum partition coefficient of studied chemicals.

References

[1] Eleftherios G, Samaras E, Riviere, and Taravat Ghafourian. The effect of formulations and experimental conditions on in vitro human skin permeation-Data from updated EDETOX database. *Int. J. Pharm.* 2012; **434**: 280-291.

[2] Fujii, M., Tekeda, Y., Yoshida, M., Utoguchi, N., Matsumoto, M., and Watanable, y.Comparation of skin permeation enhancement by 3lmenthoxypropane-1,2-diol and 1-menthol. *Int. J. pharm.* 2003; **258**: 217-223.

Prediction of flux through polydimethylsiloxane membranes using CoMFA and CoMSIA 3D QSAR.

Seyedeh Mozhgan Behgozin*, Mohammad H. Fatemi

Chemometrics Laboratory, Faculty of Chemistry, University of Mazandaran, Babolsar 47416-95447, Iran
E-mail address: *m.behgozin@stu.umz.ac.ir*

Abstract

The maximum steady state flux of 79 compounds through PDMS membranes was predicted using a refinement of a previously developed empirical quantitative structure-transportability relationship (QSTR) produced using the reported flux data for compounds. A three-dimensional quantitative structure activity relationship (3D-QSAR) study was performed on a variety of substituted benzenes, quinolines, isoquinolines and lesser numbers of other heterocyclic classes of compounds over a wide range of polarity. We have employing comparative molecular field analysis (CoMFA) and comparative molecular similarity indices (CoMSIA) techniques to investigate the structural requirements for substrates and derive a predictive model that may be used for the design of novel compound. Molecular models of all the compounds were simulated using SYBYL 6.0 molecular modeling software. The atomic charge of each individual compound was computed with the Gast-Hfick method using the same software. 3D-QSAR models were derived for benzene derivatives (training set N = 63, test set N = 16). The CoMFA model includes steric and electrostatic fields for the training set with the cross-validated q2value of 0.160 and the non-cross-validated r2 value of 0.250. The cross-validated q2 value of CoMSIA Model is 0.595 and the non-cross-vaildated r2 value is 0.780. From the cross-validated results, it can be seen that the CoMSIA model has a better predictive ability than CoMFA model. Based on the above results, the CoMFA and CoMSIA analyses can be used in the design of more potent compounds from skin.

Key words: CoMFA, CoMSIA, PDMS, QSTR, Flux, PLS, Counter maps.

Introduction

It is well known that, steady state flux through a membrane can be occurring when diffusion is membrane controlled and the concentration of diffusant in the receiver solution is kept at a negligible level (Daynes, 1920; Crank and Park, 1968). However, permeability is not easy to determine. As a result, prediction of flux using is difficult when the values of the required parameters are not available. Permeability and hence flux can be effectively estimated using simple quantitative structure-transportability relationships since the solute-solvent-membrane interaction energy is determined by the structural characteristics of the interacting species. In recent years, extensive studies have correlated the maximum steady state flux of different classes of compounds through a polydimethyl-siloxane (PDMS) membrane to some easily accessible physico-chemical properties. The purpose of this work is to validate the capability of a previously developed empirical quantitative structure-transportability relationship (QSTR) model for predicting the diffusional properties of an extended data set. Maximum steady state flux through a PDMS membrane of 79 new compounds with a wide range of polarity was predicted using a comparative molecular field analysis (CoMFA) and comparative molecular similarity indices (CoMSIA) techniques. The contribution of partial atomic charge to mass transport phenomena was further verified by the correlation of atomic charge to apparent permeability through PDMS membranes.

Results and discussion

A data set of 79 substituted benzenes, and quinolines, and isoquinolines derivatives [15] reported to have effects on the maximum steady state flux through a PDMS membrane, was used for the present 3D-QSAR (CoMFA and CoMSIA) studies. All molecular modeling and calculations were performed using the SYBYL-X 1.1 program package [30]. To choose the optimal result, we systemically altered the combination of fields and chose that value which gave the best cross-validated q^2 and the smallest errors, and the largest F value. CoMSIA and CoMFA analysis results are also summarized in Table 1.

Table 1. Summary of the results obtained from the CoMFA and ComSIA analysis.

Statistical parameters	CoMFA model	CoMSIA model
q^{2a}	0.200	0.595
ONC[b]	5	6
r^{2c}	0.890	0.940
SEE[d]	0.215	0.178
F_{ratio}[e]	140.155	225.210
r^2_{pred}	0.680	0.900
Contribution		
Stric	0.087	0.410
Electrostatic	0.198	1.186
Hydrophobic		0.750
H-bond acceptor		1.315
H-bond donor		0.955

a Cross-validation correlation coefficient after the leave-one-out procedure.

b Optimum number of component.

c Non-cross-validated correlation coefficient.

d Standard error of estimate.

e F-test value

f Correlation coefficient for test set predictions.

From the cross-validation results, it can be seen that the CoMSIA model has a better predictive ability than CoMFA model, suggesting that a reliable CoMSIA model is successfully constructed. The CoMSIA model gives better predictive r2 with less residual values of the test set compared to the CoMFA model. To view the field effect on the target property, CoMFA and CoMSIA contour maps were generated. The best CoMSIA model contain the steric, electrostatic, hydrophobic and hydrogen bond donor and hydrogen bond acceptor fields.

References

[1] Knust H, Achermann G, Ballard T, Buettelmann B, Gasser R, Fischer H, Hernandez MC, Knoflach F, Koblet A, Stadler H, Thomas AW, Trube G, Waldmeier P. The discovery and unique pharmacological profile of RO4938581 and RO4882224 as potent and selective $GABA_A$ α5 inverse agonists for the treatment of cognitive dysfunction. *Bioorg. Med. Chem. Lett.* 2009; **19** :5940–5944.

[2] Youdim MB, Buccafusco JJ. Multi-functional drugs for various CNS targets in the treatment of neurodegenerative disorders. *Trends Pharmacol. Sci.* 2005; **26** :27–35.

Developing a support vector machine based QSPR model for prediction of half-life of some herbicides

Kobra Samghani*, Mohammad Hossein Fatemi and Seyedeh Mozhgan Behgozin
Chemometrics Laboratory, Faculty of Chemistry, University of Mazandaran, Babolsar, Iran
E-mail address: *k.samghani@gmail.com*

Abstract

The half-life ($t_{1/2}$) of 61 herbicides were modeled by quantitative structure–activity relationship (QSAR) based molecular structure descriptors. The most relevant descriptors were selected by stepwise multiple linear regression for developing linear and nonlinear models, multiple linear regression (MLR) and support vector machine (SVM), respectively. Comparison between statistical parameters of these models indicates the suitability of SVM over MLR model for predicting the half-life of herbicides. By using the SVM model, the statistical parameters of R^2 and standard error for training set were; 0.98 and 0.087, respectively. This model was evaluated by leave one out cross validation test, which its result indicate the suitability of model. More over the parameters of R^2 and standard error were 0.96 and 0.094 for test set, respectively. This model was used for prediction the halt-life of other herbicides that are located in the applicability domain of model that were determined via leverage approach.

The non-linear relationship among selected molecular descriptors and herbicide's half-life emphasis that the process of degradation of herbicides in environment is very complex and can be affected by various environmental and structural features, therefore simple linear model can't able to successfully predict it.

Key words: Quantitative structure–activity relationship, Half-life, Support vector machine, Multiple linear regression, Leave one out cross validation, Applicability domain.

Introduction

Herbicides are chemicals which utilized in agricultural, aquatic, forest, and wild-land ecosystems to reduce the density of weeds to permit the growth of desirable species. Decomposition of herbicides occurs in soil, air, water, plants, animals, and microorganisms by photochemical, chemical, or microbiological means. Half-life ($t_{1/2}$) **of** herbicide can be considered as a criterion of its persistance in the environment [1]. Since carrying the experiments for determination of half-life is difficult and time consuming, therefore application of quantitative structure properties relationship (QSPR) methodology is very important. One of these methods is quantitative structure properties relationship (QSPR) methodology, which developed based on theoretical derived molecular descriptors. In the peresent work, quantitative structure–proprety relationship models were developed for prediction the half-life of some herbicides, that multiple linear regression (MLR) and support vector machine (SVM) were applied as linear and non-linear feature mapping techniques [2].

Results and discussion

The structures of 61 herbicides were drawn and optimized with HyperChem software (version 7) using the AM1 semi-empirical method. The theoretical descriptors for these chemicals were calculated by Dragon (version 3), CODESSA (version 2.7.2) and Accelrys Materials Studio (version 4.3) softwares. After removing the redundant descriptors, the stepwise multiple linear regression method was used for the selection of the most relevant descriptors.

Then data set was divided into the training and test sets. The result of diversity analysis[14] indicated that the data in the training and test sets are representative of the whole data set.

Multiple linear regression and support vector machine were applied as linear and nonlinear methods for modeling the halt-life of some herbicides. For developing the support vector machine model, radial bias function (RBF) was used as Kernel function. The optimum values of capacity (C), epsilon (ε), and the Kernel parameter (γ) for this model were achieved 10, 0.02 and 0.7. The results of MLR and SVM models are summarized in Table 1.

Table 1. Statistical Parameters of MLR and SVM Models

Model	Data set	R^2	SE
MLR	Training set	0.87	0.26
	Test set	0.87	0.26
SVM	Training set	0.98	0.087
	Test set	0.96	0.094

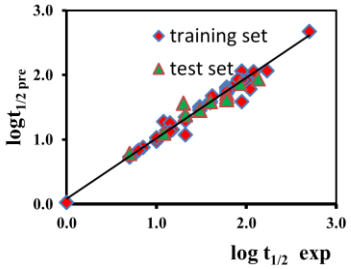


Figure 1. Plot of predicted versus experimental

Comparison between statistical parameters of these models indicates the suitability of SVM over MLR model for predicting the half-life of herbicides. In figure 1 the SVM predicted values of log $t_{1/2}$ for training and test sets are plotted against their experimental values that indicats the high correlation between experimental and predicted values. The parameters of leave one out cross validated correlation coefficient (q^2_{cv}) and standardized predicted residual error sum of squares (SPRESS) were achieved 0.52 and 0.37 respectively, which indicated the stability and robustness of the developed support vector machine model. By using this model, the half-life of 30 herbicides was predicted while their half-life was not reported and also they were placed in the applicability domain of the model.

Sensitivity analysis was performed for characterizing more important descriptors in developing the SVM model. According to the results of sensitivity analysis, the geometric and topologic aspects of molecules affect on their half-life in soil and these categories of descriptors are the most important variables for developing the reliable QSAR model for prediction the half-life of herbicides in soil.

References

[1]Emanuela B, Minarini A, Tumiatti V, Milelli A, Lunelli M, Pegoraro M, Rizzoli V, Luisa M, Paolo D. Synthetic polyamines as potential amine oxidase inhibitors: a preliminary study. *Amino Acids.* 2012; **42** :913–928.

[2] Li Y, Yuan J. Prediction of Key State Variables using Support Vector Machines in Bioprocesses. *Chem. Eng. Technol.* 2006; ***29***: 313-319.

Modeling and prediction of adsorption behavior of nanotubes

Afsane Heidari*, Mohammad H. Fatemi

Chemometrics Laboratory, Faculty of Chemistry, University of Mazandaran, Babolsar 47416-95447, Iran
E-mail address: afsane_heidari@yahoo.com

Abstract

A quantitative structure property relationship (QSPR) model is built up to predict adsorption coefficients of some small aromatic compounds on nanomaterials by CORAL software. CORAL (CORrealtions And Logic) is a freeware to assist QSAR/QSPR modeling by application of descriptors calculated with SMILES (simplified molecular input line entry system). The developed models by this software are one-variable models based on the optimal descriptors calculated with the Monte Carlo method. The adsorption behavior has been modeled with the representation of the molecular structure by hydrogen-filled graph (HFG) and by the hybrid representation of the molecular structure using combination of SMILES and HFG. A new method included three-dimensional response surface was used to optimize Monte Carlo parameters. The statistical results of both two models are acceptable, although model based HFG representation was better than model based hybrid representation. Moreover, the reliability of models was evaluated using the leave-one-out cross-validation method that led to Q^2 of 0.899 and 0.697, respectively. The results of the present study show that application of CORAL software along with presented optimization method can yield predictive and robust models.

Key words: Adsorption- molecular graph- Monte Carlo - QSPR- MWCNT- CORAL

Introduction

The adsorption energy of the nanoparticles is the primary driving force behind all of complicated nanomaterials interactions within biological systems [1, 2]. Competitive adsorption of small aromatic compounds onto the nanotube can consider as mimicking the molecular interactions of the nanotube with the residues of the proteins. Therefore, the evaluation of adsorption behavior of nanomaterials has drawn more attention in medicinal applications. In this study, experimental adsorption coefficients of forty organic compounds such as benzene derivatives, on multiwall carbon nanotube (MWCNT) were used to model the adsorption behavior of nanoparticles [3]. The CORAL software with method of balance of correlations with ideal slopes was applied to develop QSPR models [4]. In this method, training set is separated into subtraining set and validation set and the estimation of predictability is based on test set (structures, which are not used in building up of the model).

Results and discussion

Parameters of the Monte Carlo optimization are Threshold and N_{epoch} and we used subtraining, validation and test sets to obtain best values of them by considering the maximum value of squared correlation coefficient (R^2). The obtained response surfaces are indicated in Fig. 1. and optimum parameters are shown on it.

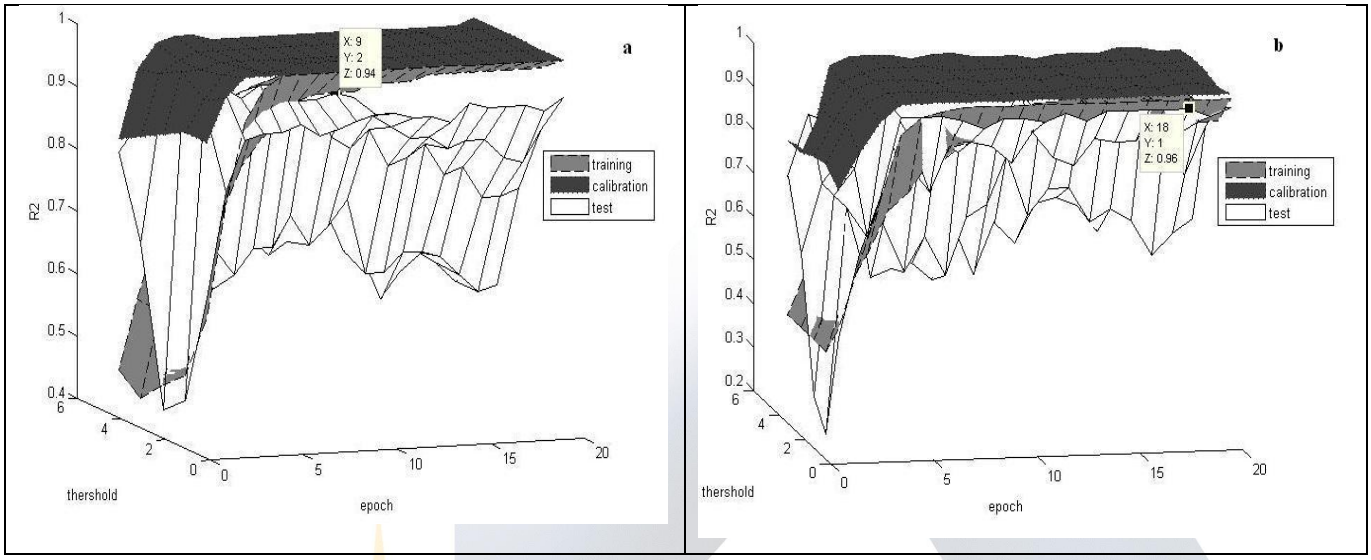


Fig. 3. Optimization of Monte Carlo Parameters: a) Hybrid model and b) HFG model

Results of two models are comparable according to statistical parameters of Table 1. Two 3-member external test sets were used as other evidence for acceptable prediction power of models, which identified in Table as external set 1 and 2. It concluded that CORAL software is able to build up predictive models.

Table 1. Statistical parameters of developed models.

	Subtraining set		Validation set		Test set			
	R^2	SE	R^2	SE	R^2	SE	$R^2_{external\ set\ 1}$	$R^2_{external\ set\ 2}$
Hybrid model	0.940	0.173	0.991	0.239	0.941	0.383	0.977	0.962
HFG model	0.985	0.086	0.997	0.258	0.961	0.235	0.983	0.871

References

[1] Nel, A. E.; Mädler, L.; Velegol, D.; Xia, T.; Hoek, E. M.; Somasundaran, P.; Klaessig, F.; Castranova, V.; Thompson, M. Understanding Biophysicochemical Interactions at the Nano-Bio Interface. Nat. Mater. 2009, 8, 543–557.

[2] Patil, S.; Sandberg, A.; Heckert, E.; Self, W.; Seal, S. Protein Adsorption and Cellular Uptake of Cerium Oxide Nanoparticles as a Function of Zeta Potential. Biomaterials 2007, 28, 4600–4607.

[3] Xia X-R, Monteiro-Riviere NA, Riviere JE: An index for characterization of nanomaterials in biological systems. Nat. Nanotechnol. 2010, 5:671-675.

[4] http://www.insilico.eu/CORAL/Assessed 07.07.2011

Novel application of the CORAL software to model Cellular Uptake of Magnetofluorescent Nanoparticles in Pancreatic Cancer Cells

Hanieh Malekzadeh [a,*], Mohammad H. Fatemi [a] and Setareh Gorji [b]

[a] *Laboratory of Chemometrics, Faculty of Chemistry, University of Mazandaran, Babolsar, Iran*

[b] *Department of Chemistry, Faculty of Sciences, Urmia University, Urmia, Iran*

E-mail address: H.Malekzadeh@stu.umz.ac.ir

Abstract *A quantitative structure-retention relationship model is built up by CORAL software based on representation of the molecular structure by the simplified molecular input line entry system and graphical representation. This model was used for the prediction of the cellular uptake of 109 magnetofluorescent nanoparticles (NPs) in pancreatic cancer cells. The hybrid version of the molecular structure representation by a combination of SMILES and the molecular graph provided the best correlation of the prediction for the cellular uptake. The ranges of statistical characteristics of Hybrid models are as follows: n=88, R^2=0.970, SE=0.163, F=3041 (sub-training set) ; n=11, R^2= 0.953, SE= 0.193, F=327 (calibration set) and n=10, R^2= 0.942, SE= 0.238, F=219 (test set). The reliability of established model was evaluated by using the leave-many-out cross-validation method (Q^2=0.932) as well as by y-scrambling that reveals the suitability of the developed model. The results of the present study have shown that CORAL software can be used for accurate prediction of Cellular uptake of 109 magnetofluorescent NPs in PaCa2.*

Keywords: Molecular descriptor; Structure-retention relationship; Linear retention indices; Nanoparticles

1. Introduction

Nanoparticles are microscopic particles having one or more dimensions of 100 nm or less. They are used in a variety of fields due to their unique physical and chemical properties. To find the effect of the nanoparticles (NPs) structure of its activity and also to accelerate the process of designing safe and efficient NPs, developing Chemoinformatics methods such as quantitative structure activity relationship (QSAR) that seems to be useful. CORAL software has been suggested as an efficient tool for the QSAR analysis. The CORAL models represent one–variable correlations between an endpoint and optimal descriptors. The optimal descriptors are calculated with special coefficients related to the presence of various molecular features. These coefficients (correlation weights) are obtained by the Monte Carlo method [1]. One can use as the representation of the molecular structure for the optimal graph-based descriptors [2], simplified molecular input line entry system (SMILES) [3-5], or a hybrid representation, which includes both the molecular graph and SMILES. The SMILES-based descriptors calculated with correlation weights of SMILES attributes, which are obtained by the Monte Carlo method. The building up of QSAR model for an arbitrary split into the sub-training, calibration and test sets should be qualified as a random event. The statistical quality of each model is a mathematical function of these splits. The comparison of aforementioned three representations of the molecular structure in the development of QSAR

approaches devoted to Cellular Uptake of Magnetofluorescent Nanoparticles in Pancreatic Cancer Cells is the aim of the present study.

2. Methodology

2.1 Data set

The experimental values of cellular uptake of 109 magnetofluorescent NPs in PaCa2 were taken from elsewhere. All NPs in the data set have exactly the same metal core decorated with different synthetic small molecules. Each NP is represented by the structure of organic surface modifier, which in turn is characterized by conventional molecular descriptors. Cellular uptake is expressed as the logarithm of the concentration (PM) of NP per cell, which varies from 2.23 to 4.44. Seven random splits into the training and test sets were examined.

3. Results and discussion

The main aim of the CORAL software may be formulated as the search for the optimum values of threshold (T*) and Nepoch (N*), which can give satisfactory statistical parameters for the test set and produce the maximum value of R^2 test. Theoretically, the correlation coefficients between experimental and calculated values of the endpoint for sub-training, calibration, and test sets depend on the threshold and number of epochs. After optimization the optimum values of R^2 against the number of epochs were achieved at Nepoch=N*=7 and the optimum value of the threshold was selected as T=T*=0.

Three types of models were developed based on SMILES, GRAPH and both SMILES and GRAPH descriptors for the cellular uptake of 109 magnetofluorescent nanoparticles in pancreatic cancer cells. Table 1 contains correlation between DCW and retention indices (R^2) of these models for all obtained data, which were obtained by CORAL software. Upon the analysis of all obtained models in table 1 one can see that the statistical parameters of the hybrid version model are better than the others.

Table 1 Comparison of predictability of optimal descriptors calculated with GRAPH, SMILES and both GRAPH and SMILES.

		R2(N)				
	Threshold	Probe 1	Probe 2	Probe 3	Average	Dispersion
GRAPH	0	0.9173(4)	0.9146(5)	0.9220(4)	**0.9180(4.33)**	0.0031(0.47)
	1	0.9003(2)	0.9079(4)	0.8999(3)	0.9027(3.00)	0.0037(0.82)
	2	0.9055(4)	0.9141(3)	0.9072(3)	0.9089(3.33)	0.0037(0.47)
	3	0.8997(3)	0.9201(4)	0.9119(3)	0.9106(3.33)	0.0084(0.47)
SMILES	0	0.9073(7)	0.9266(7)	0.8995(7)	0.9111(7.00)	0.0114(0.00)
	1	0.9291(7)	0.9219(7)	0.8913(7)	**0.9141(7.00)**	0.0164(0.00)
	2	0.9190(7)	0.8936(7)	0.8943(6)	0.9023(6.67)	0.0118(0.47)
	3	0.8684(7)	0.9252(7)	0.9086(7)	0.9007(7.00)	0.0238(0.00)
GRAPH and SMILES	0	0.95071(10)	0.9612(10)	0.9740(8)	0.9620(9.33)	0.0095(0.94)
	1	0.9625(7)	0.9644(6)	0.9659(5)	0.9643(6.00)	0.0014(0.82)
	2	0.9601(8)	0.9762(7)	0.9777(8)	**0.9714(7.67)**	0.0080(0.47)
	3	0.9616(8)	0.9713(7)	0.9628(8)	0.9652(7.67)	0.0043(0.47)

* In the above table N is the number of epoch and probe means run number. The best models are indicated by bold.

Fig. 1 represents graphically the best model of seven random splitting of data into the sub-training, calibration, and test sets for hybrid descriptors, which indicate that random split 3 is quite good.

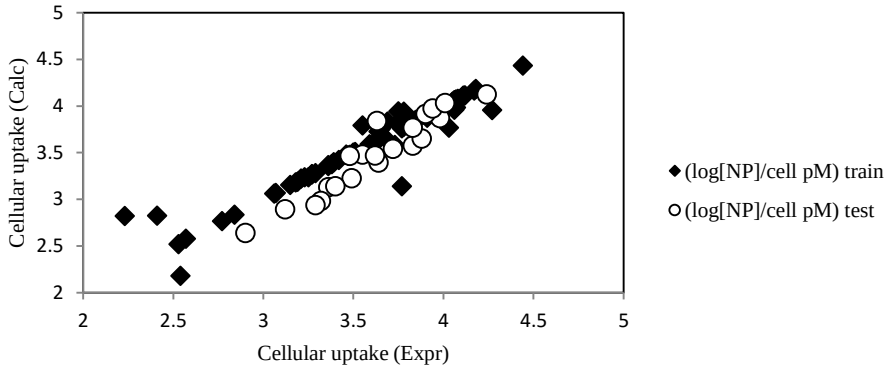


Fig. 1. Graphical representation of the performance of CORAL models for random split 3 for Hybrid model

This model has the following specifications:

Log [NP]/cell pM = 287.7875000 (± 4.3250663) + 43.4729000 (± 0.1340725) * DCW(0,13) (1)

n=88, R^2=0.970, SE=0.163, F=3041 (sub-training set)

n=11, R^2= 0.953, SE= 0.193, F=327 (calibration set)

n=10, R^2= 0.942, SE= 0.238, F=219 (test set)

where n is the number of compounds in a set; R is correlation coefficient; R^2 is correlation coefficients; SE is the standard error of estimation and F is a Fischer F-ratio.

To evaluate the prediction ability and generalization of this model (split 3), leave-one-out cross validation procedure was used. The high values of Q^2 (0.932) reveal the robustness of the model. In the next step Y-scrambling test was performed to examine that the resulted model has not been obtained by chance correlation. To do so, the Cellular uptake values vector in data matrix was randomly shuffled and the model coefficients were recalculated. This procedure was repeated many times (i.e. more than 30 times). The resulted correlation coefficients of the obtained models were ranging between 0.0003 and 0.2710, which were much lower than those established by model and revealed that there were not any chance correlations between selected molecular descriptors and dependent variable.

References

[1] Toropov AA,. Toropova AP, Martyanov SE, Benfenati E, Gini G, Leszczynska D, Leszczynski J. CORAL: QSAR modeling of toxicity of organic chemicals towards Daphnia magna. *Chemometrics and Intelligent Laboratory Systems* 2012; **110**: 177–181.

[2] Toropov AA, Toropova AP, Gutman I. Comparison of QSPR Models Based on Hydrogen-Filled Graphs and on Graphs of Atomic Orbital, *Croatica Chemica Acta* 2005; **78**: 503–509.

[3] Weininger D, SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules, *J. Chem. Inform. Comput. Sci.* 1988; **28**, 31–36.

[4] Weininger D, Weininger A, Weininger JL. SMILES. 2. Algorithm for generation of unique SMILES notation. *J. Chem. Inform. Comput. Sci.* 1989; **29**, 97–101.

[5] Weininger D. SMILES. 3. DEPICT. Graphical depiction of chemical structures, *J. Chem. Inform. Comput. Sci.* 1990; **30**, 237–243.

Selective extraction and determination of urinary nickel in foundry workers with the aid of a novel magnetic ion imprinted polymer nanoparticles and experimental design methodology

Niloofar Jalilian, Ali Akbar Asgharinezhad, Homeira Ebrahimzadeh*

E-mail address (H. Ebrahimzadeh): h-ebrahim@sbu.ac.ir

Abstract

In this work a new magnetic ion imprinted polymer nanoparticles was utilized for the rapid extraction, preconcentration and determination of Ni(II) ions in urine sample of foundry workers. The parameters affecting preconcentration were optimized by a Box-Behnken design through response surface methodology. Three variables (pH value, sorption time and magnetic imprinted polymer amount) were selected as the main factors affecting the sorption, while four variables (type, volume and concentration of the eluent; elution time) were selected for investigating the elution step in the optimization study. Following the sorption and elution steps, the Ni(II) ions were quantified by flame atomic absorption spectrometry. Equilibrium isotherms were studied, and two models were utilized to analyze the equilibrium sorption data. The results exhibited that the sorption process obeyed the Langmuir model. The maximum monolayer capacity and the Langmuir constant were 49.3 mg g^{-1} and 0.205 L mg^{-1}, respectively. Under the optimal condition the limit of detection and the relative standard deviations was 0.25 μg L^{-1} and were equal or less than 9.5%, respectively. Ultimately this nanosorbent was successfully applied to the selective determination of urinary nickel in foundry workers and satisfactory results were obtained.

Key words: Magnetic ion imprinted polymer nanoparticles; Ni(II) ions; Response surface methodology; Selective extraction; Urine samples.

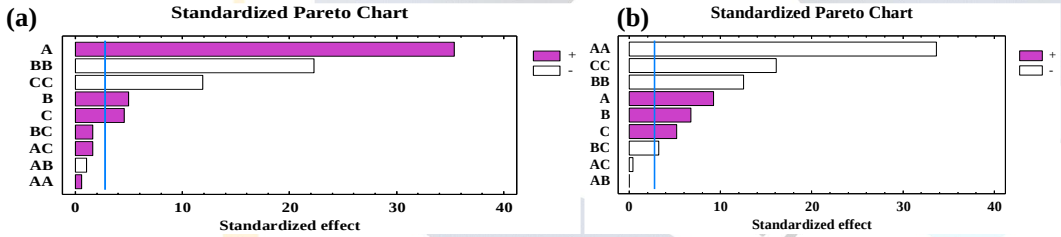
Introduction

In the recent decades, the ion-imprinting technique has gained research interest as a convenient and powerful extraction approach in which co-polymerization has been performed between functional and cross-linking monomers in the presence of desired target ions (imprint ions) [1]. In order to enhance the sensitivity and selectivity of solid phase extraction procedure (SPE), new sorbent based on ion imprinted polymers (IIPs) which have synthetic specific sites with excellent memory effect and high selectivity to the imprinted ions were developed [1]. However, low surface area, poor binding site accessibility, and difficulty to separate are the failures of bulk ion-imprinted polymers as sorbents [2]. The biological monitoring of toxic metals in urine samples has become a matter of wide interest owing to the high toxicity of these metals and their influence in controlling the course of biological activities. A Box-Behnken design was applied in order to investigate the optimum conditions of this method through response surface methodology. Ultimately, the nanosorbent was used for urinary monitoring of Ni(II) ions.

Results and discussion

The synthesis of Fe_3O_4@IIP NPs was confirmed by FT-IR spectroscopy, TGA/DTA, high-angle XRD, SEM, and EDX analysis. Box-Behnken design (BBD) was used to optimize the significant parameters affecting the separation/preconcentration of Ni(II) ions by performing as few experiments as possible. Sorption factors including pH of solution, amount of nanosorbent and uptake time are selected as effective parameter and other factors were kept constant and a BBD was applied to optimize the three factors. Effect of three factors (eluent concentration, eluent volume and elution time) was explored in elution step using BBD. The studied range of the factors were in the range of 5-8 for pH (A), 5-15 min for uptake time (B), 3-10 mg for sorbent amount (C), 0.1-1.5 mol L^{-1} for eluent concentration (A), 2-5 mL for eluent volume (B) and 2-7 min for elution time (C).

According to the overall results of the optimization study; pH, 8.0; sorption time, 10 min; and MIIP nanosorbent amount 7.0 mg were selected for sorption step and elution volume, 3.8 mL; elution time, 5.0 min; eluent concentration, 0.9 mol L^{-1} HCl were chosen as opted conditions for elution step. The Pareto charts (sorption and elution) of main and interaction effects are depicted in Fig. 1 ((a) sorption step and (b) elution step).



The effect of sample volume should be explored since it affects the preconcentration factor. As the sample volume was increased up to 1000 mL. Hence, 1000 mL was selected as the breakthrough volume in this work. The recovery of trace metal ions may be affected by other components of the samples too. All the studied metal ions have no remarkable effect on the preconcentration of Ni(II) ions at the opted conditions due to the high selectivity of the ion imprinted polymer toward the target ions. Quality features of the current method were evaluated under the final opted conditions. Under the opted conditions, the proposed method exhibited a linear dynamic range within the concentration range of 1.0-300 µg L^{-1} with a correlation of determination (r^2) of 0.996 and the extraction recovery was 99.0%. The limit of detection based on the ratio of three standard deviation of the blank signal to slope of the calibration curve ($C_{LOD}= 3S_b/m$) was 0.25 µg L^{-1}. A certified reference material (Seronorm LOT NO2525) with a certified Ni(II) content of 41.5±0.05 µg L^{-1} was used for method validation. The average concentration of the SRM sample and its standard deviation for three replicate measurements was 42.0±1.3 µg L^{-1}. The obtained *t* value for the analysis of Ni(II) in SRM was smaller than 4.30 ($t_{0.05,2}$ = 4.30), which confirms that there was no significant difference between the measured value and the certified value reported for the SRM and the method is reliable. The optimized method was applied for the determination of Ni(II) ion in the urine samples taken from the male workers from one of the Iranian automobile industries whom supposed to be exposed to Ni(II) through foundry process and satisfactory results were obtained.

References

[1] M. Liu, C. Chen, T. Wen, X. Wang, Synthesis of magnetic ion-imprinted composites and selective separation and preconcentration of U(VI), Dalton. Trans. 2014; 43: 7050-7056.

[2] L. Yan, L. Zhanchao, W. Yun, D. Jiangdong, G. Jie, X. Jimin, Y. Yongsheng, A surface ion-imprinted mesoporous sorbent for separation and determination of pb(II) ion by flame atomic absorption spectrometry, Microchim. Acta. 2011; 172: 309-317.

Characterization of Teucrium Polium Essential Oil component by GC-MS analysis using MCR and PARAFAC methods

Hamidreza Nikpour[1], Mehdi Mousavi[2,*], Hameedeh Asadollahzadeh[1]

[1] *Department of Chemistry, faculty of Sciences, Azad University, Kerman, Iran.*

[2] *Department of Chemistry, faculty of Sciences, Shahid Bahonar University of Kerman, Kerman, Iran.*

*E-mail address (corresponding author): mmousavi@uk.ac.ir

Abstract

Essential oil of Teucrium Polium was extracted by hydro distillation. Because of matrix complexity some fundamental problems may happen during direct gas chromatography – mass spectroscopy analysis which causes low similarity index in library search and uncertainty in components identification. In the present work, GC-MS is coupled to multivariate curve resolution alternative least square (MCR-ALS) and parallel factor analysis (PARAFAC) to extract more information from experimental data. The number of identified components with similarity index greater than 80 and calculated and reported retention index differences less than 40 increases from 76 in direct to 108 in MCR and 177 in PARAFAC methods.

Key words: Teucrium Polium; Essential oil; GC-MS analysis; Multivariate curve resolution; Parallel factor analysis.

Introduction

Teucrium polium is a native herb to the western Mediterranean and Irano-Turanian sphere with small pink to white flowers. In traditional persian medicine, Teucrium polium which locally called 'Kalpooreh' is used as an anti-hypertensive, anti-bacterial, carminative, nociceptor and anti-diabetes agent [1-2]. Gas chromatography–mass spectrometry (GC–MS) is one of the most applied techniques for essential oil analyses and components determination. For natural samples such as essential oils the sample is highly complicated so baseline drift, background, and some overlapping/embedded peaks cause high uncertainty in MS library search and decreases accuracy of components determination. On the other hand, one of the main goals of chemometric techniques is extraction of more information from experimental data. Therefore, in the recent decades resolution techniques such as evolving factor analysis (EFA), heuristic evolving latent projections (HELP), orthogonal projection resolution (OPR), multivariate curve resolution-alternating least squares (MCR-ALS) [3], and parallel factor analysis (PARAFAC) [4] have been developed to fulfill this goal. In this study, Teucrium polium essential oil constituents were identified using GC-MS coupled with MCR-ALS and PARAFAC techniques and the results were compared.

Method

Teucrium polium was collected from Koohbanan in north-western of Kerman province, Iran in April 2014. The areal parts of collected plant were separated and dried in shadow at room temperature for one week. Then 200gr of dried Teucrium polium weighed, semi-grounded and immersed in 500ml water. The mixture was hydro-distilled in a full glass Clevenger-type for 4 hours. Under optimum condition, GC–MS analyses of the essential oil was performed by using an Agilent Technologies (USA) GC-MS system equipped with a HP-5MS 5% Phenylmethyl Siloxan capillary

column coupled with quadrupole mass analyzer. In the temperature programmed analysis, carrier gas was helium (99.999%) with a flow rate of 1mLmin^{-1}; injector temperature, 260 °C; split ratio, 1:50. Enhanced Agilent MSD ChemStation software was used for data processing of the total ion chromatogram (TIC) of the essential oil.

Results and discussion

Careful inspection of the TIC chromatogram shows that there are several overlapped peaks which are due to complexity of T. Polium essential oil. This problem lead to uncertainty in direct library searche based on NIST-MS database, so that many components cannot be identified. In direct MS analysis, only 77 components were identified which had similarity index more than 80 and calculated and reported retention index differences (ΔRI) less than 40. To solve the problem MCR and PARAFACT resolution techniques were applied on the chromatogram. At the beginning, the TIC is divided into 177 individual peak clusters for the analysis and information of each peak cluster is converted to an ASCII format file by using the Agilent MSD Chem-Station. In MCR technique, several procedures including baseline correction, denoising, smoothing, chemical rank determination and MCR-ALS calculation[9] were carried on each peak cluster. At final step of MCR-ALS calculation, specteral and concentration profiles of each peak cluster is obtained which help to assign the components of the peak cluster. In MCR analysis, 108 components of the essential oil were identified which had similarity index more than 80 and ΔRI less than 40.

The PARAFAC algorithm is performed on raw ASCII format information of each peak cluster. To carry out PARAFAC, new software called CAIS is generated in our laboratory to process the data. The number of components of each peak cluster is determined by CORCONDIA and initializing is fitted by one of the DTLD/GRAM or SVD methods[10]. There is ability in CAIS software to find the best fitting model of initializing methods and this option is used for all PARAFAC calculations. In this part of study, 177 compounds of the essential oil were identified which had similarity index more than 80 and ΔRI less than 40.

Conclusion

The obtaind results indicate that the use of resolution techniques is highly recommended to exteract more information from raw experimental data. Also, comparisons of the results indicate that PARAFAC method is more reliable than MCR method and convergence rate and precesion is higher in PARAFAC.

References

1. Afifi, F.U., B. Al-Khalidi, and E. Khalil, *Studies on the in vivo hypoglycemic activities of two medicinal plants used in the treatment of diabetes in Jordanian traditional medicine following intranasal administration.* Journal of Ethnopharmacology, 2005*;100;*314-318.
2. Suleiman, M.S., et al., *Effect of Teucrium polium boiled leaf extract on intestinal motility and blood pressure.* Journal of Ethnopharmacology, 1988;22;111-116.
3. Gonçalves, D., et al., *Fractionation of orange essential oil using liquid–liquid extraction: Equilibrium data for model and real systems at 298.2 K.* Fluid Phase Equilibria, 2015**;**399;87-97.
4. Bro, R., *PARAFAC. Tutorial and applications.* Chemometrics and Intelligent Laboratory Systems, 1997;38;149-171.

Development of New Interfaces for Performing Chemometric Methods in Portable Instruments and Smartphones

Hamidreza Shaye, Hadi Parastar*

Chemometrics Research Lab., Department of Chemistry, Sharif University of Technology, Tehran, Iran

E-mail address: h.parastar@sharif.edu

Abstract

In the present contribution, a novel Java library entitled "MVC" is developed to perform different multivariate calibration methods. The developed library can implement in any Java-supported device as an environment for running. Different multivariate calibration methods, such as multiple linear regression (MLR), principal component regression (PCR) and partial least squares (PLS) are included in this library. Special versions of application were developed for Android and Linux operating systems (OS) based on MVC library. Android version named "MVC App" is capable to run in any smartphone and tablet based on Android OS (Version 4.0 and higher). Also, Linux versions specially designed to execute in Linux based portable platforms, such as Raspberry Pi and other ARM architecture based devices supported Java included in most of the portable analytical instruments (UV-Vis, NIR and Raman). These interfaces are designed for chemists who are not expert in programming or in advanced statistics.

Key words: Multivariate calibration; MLR; PCR; PLS; Android; Linux

Introduction

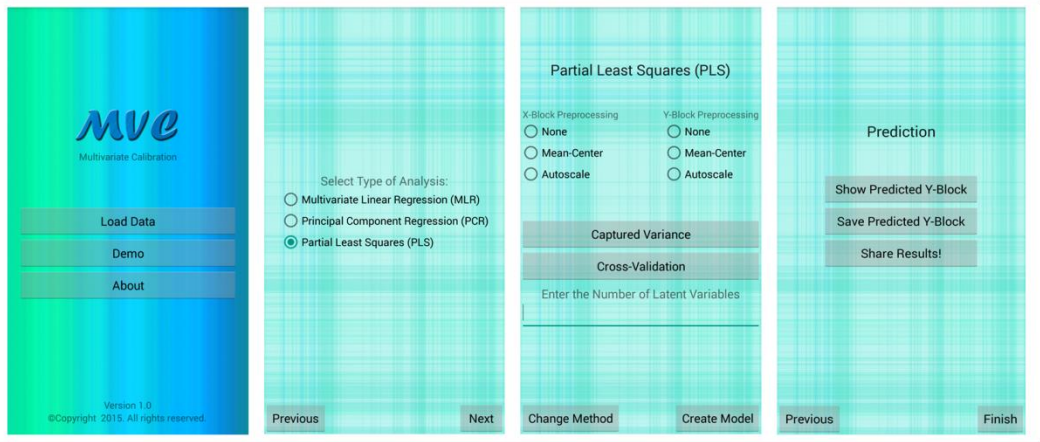
Owing to the fact that analytical chemistry can involve samples that are far from simple and often contain many components, therefore, multivariate calibration (MVC) methods have been proposed in recent decades to overcome fundamental mathematical challenges occurred during analysis of these complex mixtures [1]. Among different MVC algorithms, multiple linear regression (MLR) [2], principal component regression (PCR) [2] and partial least squares regression (PLSR) [3] has attracted great attention in chemistry in recent years. However, one the most important features of chemometric methods is development of user-friendly software in order to develop the range of users of these methods. In this regard, development of software for chemists who are not expert in programming or in advanced statistics will be an important mission for chemometricians.

The aim of this work was development of new interfaces (app) to perform different types of chemometric methods including multivariate calibration ones. These interfaces are based on Java language programming and it can be run on any hardware supported Java, such as Android, Linux and other ARM architecture based devices supported Java. On the other side, these types of apps can be coupled to portable instruments which are highly useful for in situ and field studies. In these cases, data can be collected with instruments and then, data can be transmitted to a smartphone or any supported device through a Wi-Fi or USB connection and finally, they can be easily manipulated and processed with no extra effort. Developed apps in this study include different MVC methods, such as MLR, PCR and PLSR. Also, several important tools for data preprocessing (e.g., mean-centring and auto-scaling), rank determination and model evaluation (e.g., Cross-Validation (CV) [4]) and also for android version plotting and sharing facilities are included. Running MVC app does not require a serious experience; however, a basic knowledge of the underlying methods is helpful to successfully interpret the results.

Results and Discussion

As an instance, for performing PLS modelling on Android based smartphones, first calibration and validation data sets are imported (via USB or Wi-Fi). Then, the number of latent variables (LVs) is chosen using leave-one-out cross-validation (LOO-CV). Afterwards, PLS model is built and the user can review the modelling results. In this regard, figures of merit (FOMs) of models, such as root-mean square error of prediction (RMSEP), standard error of prediction (SEP), bias and relative error (RE) and other parameters can be viewed for each analyte. Furthermore, various plotting options are included for each model. In this manner, there are two options: (1) showing data as a table which it is possible to scroll display horizontally and/or vertically to scan all matrix data and (2) showing data as a plot. In this way, horizontal scrolling in bottom-one-third of screen will change across the rows of data matrix and tapping the screen by three fingers will share current plot as JPEG-picture. Also, there is an option to share logs' file by smartphone via Bluetooth, message, Email and other supported methods depends on user's device. All of these options are available just by touching the screen, with no complexity that almost every chemist can use.

Figure 1 shows some of the pages of the developed interface for smartphones.



References

[1] Beebe K.R. and Kowalski B.R., An introduction to multivariate calibration and analysis. *Anal. Chem., 1987: **59**:1007A-1017A.*

[2] Brereton R.G., Chemometrics: data analysis for the laboratory and chemical plant. 2003: John Wiley & Sons, NY, USA.

[3] Wold S., Sjöström M., and Eriksson L., PLS-regression: a basic tool of chemometrics. *Chemom. Intell. Lab. Syst., 2001: **58**:109-130.*

[4] Kohavi R., A study of cross-validation and bootstrap for accuracy estimation and model selection. *Int. Joint Conf. Artif. Intell., 1995, **14**, 1137–1145.*

Partial least squares calibration method for the simultaneous determination of montelukast, fexofenadine and cetirizine in pharmaceutical tablets

Seyed Karim Hassaninejad-Darzi[1,*], Zarrin Eshaghi[2], Seyed Mohammad Nikou[3]

[1]*Research Laboratory of Analytical & Organic Chemistry, Department of Chemistry, Faculty of Science, Babol University of Technology, Shariati Av., Babol, Iran. P.O. Box: 484, Postal Code: 47148-71167, Tel/Fax: +98-1132334203.*

[2]*Department of Chemistry, Faculty of Science, Payame Noor University, P.O. Box 19395-3697, Tehran, I.R. Iran: eshaghi@pnu.ac.ir*

[3]*Department of Chemistry, Payame Noor University, P.O. Box 19395-3697, Tehran, I.R. Iran: smnikou@yahoo.com*

E-mail address (corresponding author): hassaninejad@nit.ac.ir

Abstract

The UV-Vis absorption spectra of sodium montelukast (MONT), fexofenadine hydrochloride (FEXO) and cetirizine hydrochloride (CET) are strongly overlapped and do not allow direct determination without previous separation by conventional methods. A simple, fast and precise spectrophotometric method combined with partial least-squares (PLS1) has been developed for the simultaneous determination of MONT, FEXO and CET. Data of analysis were obtained from UV-Vis spectra of three compounds. The method of central composite design was used for calibration and validation sets. A simple and fast method for wavelength selection in the calibration step is presented, based on the minimization of the predicted error sum of squares (PRESS) calculated as a function of a moving spectral window. The limit of detection was obtained 0.014, 0.062 and 0.112 mgL^{-1} for MONT, FEXO and CET, respectively. The procedure was successfully applied for simultaneous determination of the above compounds in pharmaceutical tablets.

Key words: Partial Least Squares 1; Montelukast; Fexofenadine; Cetirizine

Introduction

Asthma is a common chronic inflammatory disease of the airways characterized by variable and recurring symptoms, reversible airflow obstruction and bronchospasm. The control of asthma symptoms is a realistic goal and studies have shown that this can be achieved in most asthma patients leading to a higher quality of life [1]. Montelukast (MONT) is a potent and selective antagonist of the cysteinyl leukotriene receptor utilized for the treatment of asthma. Their discovery has had a significant impact on treatment strategies for the management of asthma [2]. Fexofenadine (FEXO) is a non-cardiotoxic and non-sedative terfenadine metabolite, which acts as a selective second-generation histamine H_1 receptor antagonist, relieving the uncomfortable manifestations of rhinitis [3]. Cetirizine hydrochloride (CET), a piperazine derivative and metabolite of hydroxyzine, described as a long-acting non-sedating antihistamine with some mast-cell stabilizing activity [4]. In the present work, we developed and validated a simple, fast and sensitive method for the simultaneous quantification of MONT, FEXO and CET with UVVis spectroscopy with the help of partial least squares multivariate calibration techniques in pharmaceutical tablets.

Results and discussion

The calibration set of 15 samples was built according to a central composite design (CCD) (three factors at twolevel in the cubic vertex, six experiment in the cubic face and one central point) in the concentrations range of 2.0– 12.0 mg L^{-1} for all drugs. The PLS1 model was developed using the calibration/prediction dataset that demand a suitable experimental design of the standards belonging to the calibration set in order to have good predictions. The full cross validation method suggested by Haaland and Thomas [5] was used for the selection of the optimum number of factors. Table 1 is shown data of test set composition, predicted values and relative error percentage for MONT, FEXO and CET .The minimum mean values of relative errors was calculated to be 3.93, 5.33 and 2.72 for MONT, FEXO and CET, respectively.

Table 1. Composition of test set and predicted values for MONT, FEXO and CET by PLS1 regression. Concentration values are expressed as mgL^{-1}.

	MONT			FEXO			CET		
Sample No.	Actual	Predicted	RE	Actual	Predicted	RE	Actual	Predicted	RE
Test 1	11.00	10.21	-7.18	11.00	10.68	-2.91	11.00	11.00	0.00
Test 2	3.00	2.80	-3.67	11.00	10.90	-0.91	11.00	10.38	-5.64
Test 3	11.00	10.03	-8.82	3.00	2.84	-5.33	11.00	10.50	-4.55
Test 4	11.00	10.20	-7.27	3.00	2.55	-15.00	3.00	2.91	-0.82
Test [3]	3.00	3.00	0.00	11.00	10.02	-9.02	3.00	2.85	-1.36
Test 6	3.00	2.88	-4.00	3.00	2.83	5.67	3.00	3.13	4.33
Test 7	7.00	6.80	-2.86	7.00	6.65	6.43	7.00	6.80	-2.86
Test 8	7.00	6.72	-4.00	7.00	6.85	2.14	3.00	2.93	-2.33
Test 9	7.00	6.88	-1.71	11.00	11.36	3.27	7.00	7.20	2.86
Test 10	3.00	2.97	-0.43	7.00	6.83	2.43	7.00	6.83	-2.43
R.E.			3.93			5.33			2.72

References

[1] Masoli M, Fabian D, Holt S, Beasley R. The global burden of asthma: executive summary of the gina dissemination committee report. Allergy 2004; **59**:469–78.

[2] Ibrahim MA, Amin EF, Ibrahim SA, Abdelzaher WY, Abdelrahman AM. Montelukast and irbesartan ameliorate metabolic and hepatic disorders in fructose-induced metabolic syndrome in rats. Eur J Pharmacol 2014; **724**:204–10.

[3] Simpson K, Jarvis B. Fexofenadine. Drugs 2000; **59**:301–21.

[4] Sweetman SC. Martindale-the complete drug reference. London: Pharmaceutical Press. 2002.

[5] Haaland D, Thomas E. Comparison of multivariate calibration methods for quantitative spectral analysis. Anal Chem 1990; **62**:1091–9.

Application of experimental design as a chemometric tool for optimization of new coupling of ultrasound assisted and reverse phase- dispersive liquid-liquid microextraction before high performance liquid chromatography for the sensitive determination of vitamin A in oil samples

Sahar Farsimadan*, Nasser Goudarzi, Mansour Arab Chamjangali and Ghadam Ali Bagherian

Faculty of chemistry, University of Shahrood, Shahrood, Iran

E-mail: Saharfarsimadan@yahoo.com

Abstract

Optimization of parameters using experimental design for the ultrasound assisted-reverse phase-dispersive liquid liquid microextraction technique is developed for the extraction and determination of vitamin A from oil matrices before a high performance liquid chromatography analysis. A methodology based on the full factorial design is carried out for choosing the significant parameters. Then the significant factors affecting the extraction efficiency including pH, volume of extraction solvent, and volume of disperser solvent are optimized using the Box-Behnken design. After analyzing the results obtained, the optimum conditions were: pH 4.5, 80-20 µL of the ethanol-water solvent mixture as extraction solvent, 110 µL of 1, 4- dioxane as the disperser solvent, and a sonication time of 10 min. For validation of the developed method, the linear dynamic range, repeatability, limit of detection, and recoveries were obtained under the optimum conditions. The detection limits of the method were 1.6 ng mL^{-1} for vitamin A. The extraction recovery percentages was above 91%, with acceptable relative standard deviation (RSD). The proposed methodology was successfully applied for the determination of the vitamin A in different oil samples.

Keywords: Experimental design; response surface methodology; Vitamin A; Olive oil samples

1. Introduction

Vitamins are organic compounds and the vital nutrients that an organism requires in limited amounts to stay healthy and avoid nutritional deficiencies. Vitamins can be classified as the two classes of water-soluble and fat-soluble [1, 2].

In this work, the chemometrics tool was used to study the effects of the experimental parameters on the efficiency of the microextraction method. Firstly, in order to obtain the significant parameters that affect the microextraction efficiency, a full factorial design (FFD) was applied as the screening design. Then the optimum experimental conditions for the pre-concentration procedure were obtained using the Box-Behnken design (BBD).

2. Results and discussion

The effects of four experimental factors including the pH (A), volumes of extraction solvent (B) and disperser solvent (C), and sonication time (D) on the extraction recovery were evaluated using a two-level FFD. The experimental response (Y) was represented by the following linear polynomial model:

Response = -532.75 + 157.0 A - 0.3625 B + 21.55 C - 1.805 D + 2.65 AB - 3.4 A+ 0.74 AD - 0.2 BC + 0.03 BD + 0.04 CD

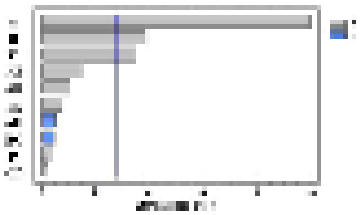


Figure 1. Standardized (P = 0.05) Pareto chart in FFD.

The experiments were conducted with three independent variables by applying a BBD. By applying the multiple regression analysis, a second-order polynomial equation was obtained, as follows:

Response = -414.805 + 633.047 A - 52.05 B + 23.2502 D - 202.853 A2 + 19.8 AB - 1.588 AD - 0.0217708 B2+ 0.1325 BD - 0.119783 D2

ANOVA test was applied to evaluate the statistical significance of the model. A summary of the ANOVA test is shown in Table 1. 3D response surface for the main factors (pH and volumes of the extraction and disperser solvents) vs. sum of the peak areas for vitamin A are shown in Fig. 2. Based upon the results obtained from BBD the pH has the largest influence on the extraction efficiency.

Table 4. ANOVA test results for BBD.

Source	Sum of Squares	Df*	Mean Square	F-Ratio	P-Value
A	1.48×10^6	1	1.48×10^6	90.79	0.0002
B	1.68×10^6	1	1.68×10^6	103.52	0.0002
C	86736.10	1	86736.1	5.33	0.0690
AA	370939.0	1	370939.0	22.81	0.0050
AB	980100.0	1	980100.0	60.26	0.0006
AC	39402.3	1	39402.3	2.42	0.1803
BB	280.01	1	280.006	0.02	0.9007
BC	70225.0	1	70225.0	4.32	0.0923
CC	331109.0	1	331109.0	20.36	0.0063
Total error	81317.4	5	16263.5		
Total (corr.)	5.07×10^6	14			

* DF: degree of freedom

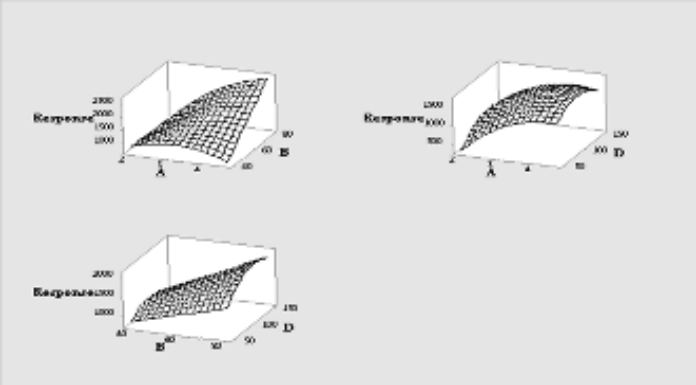


Figure 2. Estimated response surface plot.

4 Conclusion

In comparison with the conventional DLLME, the extraction phase in this method is a micro-volume of a mixture of ethanol and water with high safety that can be directly injected to HPLC. The proposed method is fast, of low cost, and does not require any toxic organic solvent. This method provided low LODs, a good repeatability, and an acceptable enrichment factor in a short analysis time.

References

[1] Santos J, Mendiola J, Oliveira M *et al.* Sequential determination of fat-and water-soluble vitamins in green leafy vegetables during storage. *J. Chromatogr. A* 2012; **1261**:179-188

[2] Plozza T, Trenerry VC, Caridi D. The simultaneous determination of vitamins A, E and β-carotene in bovine milk by high performance liquid chromatography–ion trap mass spectrometry (HPLC–MSn). *Food chem* 2012; **134**:559-563.

Chemometric optimization of ultrasound-assisted dispersive liquid-liquid microextraction based on solidification of floating organic droplets for determination of sertraline and citalopram in human serum by high performance liquid chromatography

Sahar Farsimadan*, Nasser Goudarzi, Mansour Arab Chamjangali and Ghadam Ali Bagherian

Faculty of chemistry, University of Shahrood, Shahrood, Iran

E-mail: Saharfarsimadan@yahoo.com

Abstract

A novel coupling of ultrasound-assisted dispersive liquid-liquid microextraction based on solidification of floating organic droplets (UA-DLLME-SFODs) and experimental design as chemometrics tool was developed for the simultaneous pre-concentration and determination of citalopram (CIT) and sertraline (SER) in the human serum samples. In this proposed procedure, 1-undecanol and ethanol were used as the extraction and disperser solvents, respectively. Several parameters affecting the extraction efficiency were investigated by a Plackett-Burman factorial design as a screening design. The variables showing significant effects on the analytical response were pH and volume of the extraction solvent; they were considered using central composite face centered design (CCFD). The method detection limits were 0.39 and 0.36 ng mL^{-1} for CIT and SER, respectively. Extraction recovery percentages for all the drugs were above 92.9 with acceptable relative standard deviations (RSDs).

Keywords: Anti-anxiety drugs; Experimental design; Human serum samples;

1 Introduction

Anxiety is the most common type of psychiatric illness that is widely spread around the world in the recent century. Variant classes of drugs have been introduced to control and treat this type of illness and the related psychiatric illnesses. Sertraline (SER) and citalopram (CIT) belong to the selective serotonin reuptake inhibitor (SSRI) class, which are clinically effective for the treatment of anxiety. [1, 2]. In this work, in order to use the advantages of modification in DLLME, we introduced a new ultrasound-assisted dispersive liquid-liquid microextraction based on solidification of organic droplets (UA-DLLME-SFOD) for the simultaneous determination of CIT and SER in biological matrices using HPLC.

2. Results and discussion

Since the extraction procedure depends on numerous factors, the one-factor-at-a-time (OFAT) classical optimization approach for all the factors is tedious and time-consuming. Based on the number of factors, a Plackett-Burman factorial design was applied to determine the main factors. In the proposed procedure, pH, type and volume of the extraction and disperser solvents, time, and salt effect were considered. The experimental response (Y) was represented by the following linear polynomial model:

$$Y = 49075.4 + 269.128\ A + 559.302\ B - 4.78366\ C + 272.594\ D + 42.0324\ F - 66.0037\ G + 0.392402\ E$$

In this study, a central composite faced-center design (CCFD) was carried out on twelve randomized runs consisting of four factorial points, four star points, and three center points with two replicates. Based on the results obtained from the Plackett-Burman factorial design, only the three factors pH, type, and volume of the extraction solvent were

significant, and had to be considered. The four variables that were not considered significant during the experimental design were set at their level based on their sign in the Pareto chart. Analysis of variance of the results obtained for CCFD is given in Table 1. The final model in terms of actual factors is as follows:

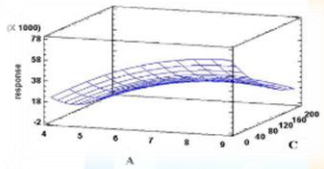
$$Y = -86814.9 + 35049.5\ A - 166.834\ C - 1991.51 \times A^2 - 34.951\ AC + 1.03771\ C^2$$

Table 1. ANOVA results for CCF.

Source	Sum of squares	Df	Mean square	F-ratio	P-value
A	1.1646×10^7	1	1.1646×10^7	90.39	0.0002
C	6.24403×10^6	1	6.24403×10^6	48.46	0.0009
AA	2.71611×10^6	1	2.71611×10^6	21.08	0.0059
AC	890683.0	1	890683.0	6.91	0.0466
CC	255785.0	1	255785.0	1.99	0.2179
Lack of fit	2999853.0	3	999951.0	2.18	0.329
Total error	644239.0	5	128848.0		
Total (corr)	2.21458×10^7	10			

The estimated response surface for pH and volume of the extraction solvent is shown in Figure 1. Based on the results obtained, pH and volume of the extraction solvent had a significant effect on the extraction recovery. The extraction recovery increased with decrease in the volume of the extraction solvent and increase in the pH.

Figure 1. Estimated response surface plotting volume of extraction solvent *vs.* pH.



4. Conclusion

The application of statistical experimental design for the selection of the UA-DLLME-SFOD conditions reduced the number of experiments required for the optimization and true optimum sets of conditions. Finally, this method was successfully applied to the analysis of these drugs in the human serum samples.

References

[1] Bosch ME, Sánchez AR, Rojas FS, Ojeda CB. Analytical methodologies for the determination of sertraline. *J. Pharm. Biomed. Anal* 2005 2008; **48**:1290-1302.
[2] Dejaegher B, Vander Heyden Y. Experimental designs and their recent advances in set-up, data interpretation, and analytical applications. *J. Pharm. Biomed. Anal* 2005 2011; **56**:141-158.

Potential of modified ionic liquid cold-induced aggregation dispersive liquid-liquid microextraction and central composite design for simultaneous preconcentration and determination of lead and cadmium in food and water samples

Nasibeh Saeedzadeh Amiri, Farzaneh Shemirani*, Rouhollah Khani

School of Chemistry, University College of Science, University of Tehran, P.O.Box 14155-6455 Tehran, Iran

Corresponding author: shemiran@khayam.ut.ac.ir

Abstract

A reliable and efficient method, termed ionic liquid cold-induced aggregation dispersive liquid–liquid microextraction (IL-CIA-DLLME), combined with flame atomic absorption spectrometry was applied for the simultaneous determination of lead (II) (Pb^{+2}) and cadmium (II) (Cd^{+2}) in food and water samples. The ionic liquid [Hmim][PF_6] is dispersed into a heated sample solution containing sodium hexafluorophosphate [$NaPF_6$] as a common ion source. The solution is then placed in an ice-water bath upon which a cloudy solution forms due to the decrease of the solubility of the IL. Lead and cadmium are complexed with ammonium pyrrolidine dithiocarbamate (APDC) and extracted into the IL. The enriched phase is dissolved in a diluting agent and introduced to the FAAS. The method is not influenced by variations in the ionic strength of the sample solution. The effective parameters, such as pH, amount of chelating agent, amount of IL and amount of common ion were optimized by a full factorial design to identify the most important parameters and their interactions, and central composite methodology was used to achieve the optimum point of effective parameters to the response.

Keywords: Modified ionic liquid cold-induced aggregation dispersive liquid-liquid microextraction; Flame atomic absorption spectrometry; Central composite design; Trace metals; Food samples.

Introduction

The increasing use of heavy metals over the past few decades has inevitably led to an increased pollution of metallic substances in natural water, and posed serious ecological and health risks. Cd^{+2} and Pb^{2+} are two elements that are among the most hazardous to human health. However, direct determination of Pb^{2+} and Cd^{2+} in water and food samples is often difficult, not only because of low concentration, but also because of matrix effects. To solve this problem, separation–preconcentration procedures are often involved

prior to analysis [1]. In 2010, a robust and practical sample preparation method termed ionic liquid cold-induced aggregation dispersive liquid-liquid microextraction was developed by Zhang et al. [2] to reduce the extraction time and required amount of IL. We applied this method for preconcentraton of Lead and Cadmium in real samples. Multivariate

techniques have been widely applied in optimization of methods in analytical chemistry. Recently, factorial design was used several times for optimization of preconcentration procedures for metal determination [3-4]. In this work, (M-IL-CIA-DLLME) was combined with FAAS by using microsample introduction system for determination of Pb and Cd in foods and water samples. The optimization step was performed by a central composite full factorial design.

Results and discussion

To evaluate the significance of the model equation and the related terms, analysis of variance (ANOVA) was considered. The F-values indicate that the model is significant and the lack of fit is not significantly relative to the pure error, hence confirm the validity of the model. This model that is shown in Eq. (1) consists of four main effects, two two-factor interaction effects, and four curvature effects as follows:

$$R = 0.34 + 0.014A + 0.00915B + 0.005583C + 0.006383D - 0.014AB - 0.002212AC - 0.001837AD + 0.009938BC + 0.006063BD - 0.010CD - 0.017A^2 - 0.009367B^2 - 0.022C^2 - 0.019D^2 \quad (1)$$

Under optimum conditions, the calibration curves were observed as linear in the concentration range of 5-100 µg L-1 Pb and 3-70 µg L-1 Cd by using 10 mL of the solution. The correlation coefficient of the calibration curve equations were 0.998 for Pb and 0.999 for Cd which indicate that a good linear regression were established between the absorbances and the concentrations. The detection limits, calculated according to six times the standard deviation of the blank signals with the preconcentration step, were 0.38 and 0.12 µg L-1 for Pb and Cd, respectively. The precision expressed as a relative standard deviation (RSD) for five replicate measurements of 50 µg L-1 Pb(II) was 2.3% and 1.7% for 20 µg L-1 of Cd(II). The enrichment factor was calculated by the ratio of slope of preconcentrated samples to those obtained without preconcentration and it was 98 for Pb and 77 for Cd.

References

1. Pourreza N, Rastegarzadeh S, Larki A. Simultaneous preconcentration of Cd(II), Cu(II) and Pb(II) on Nano-TiO2 modified with 2mercaptobenzothiazole prior to flame atomic absorption spectrometric determination. *J. Iind. Eng. Chem*. 2014; 20: 2680-2686.
2. Zhang H, Cheng M, Jiang X. Determination of benzoic acid in water samples by ionic liquid cold-Induced aggregation dispersive LLME coupling with LC. *Chromatographia*. 2010;72:1195–1119.
3. Pourbasheer E, Sadafi S, Ganjali MR, Abbasghorbani M. Dispersive liquid–liquid microextraction for preconcentration and determination of phenytoin in real samples using response surface methodology-high performance liquid chromatography. *RSC. Adv.* 2014; 4: 62190-62196.
4. Parbodi B, Londonio A, Griselda P, Smichowski P. On-line solid phase extraction of bismuth by FI–HG–AAS using l-proline immobilised on carbon nanotubes combined with factorial design. *Michrochem J*. 2015; 122: 70-75.
5. Medjahed B, Didi MA, Villemin D.Factorial design in optimization of extraction procedure for copper (II) using Aliquat 336 and Tri-nbutylphosphate based supported liquid membrane.Desalin Water Treat. 2014; 52: 3237-3245.

Investigation of the effect of various gravitational-constant functions on the feature selection performance of GSA: A QSAR study of antiprotozoal activity of benzyl phenyl ether diamidine derivatives

Tayebeh Baghgoli, Mehdi Mousavi*

Department of Chemistry, Faculty of Sciences, ShahidBahonar University of Kerman, Kerman, Iran

*E-mail address (corresponding author): mmousavi@mail.uk.ac.ir

Abstract

Binary gravitational search algorithm (BGSA) as a novel feature selection method is a time_dependent algorithm. In the present work the effect of one of the time-decreasing parameters, i.e. gravitational constant was evaluated on the performance of the algorithm in the selection of the best subset of descriptors. Evaluation functions including square root, linear, exponential and logarithmic were examined. The obtained results from QSAR study of benzyl phenyl ether diamidine derivatives, which were evaluated against acute infection of Trypanosoma brucei rhodesiense (T.b. rhodesiense) indicate that the exponential function is superior and more efficient than the other gravitational constant functions for applying in the binary gravitation search algorithm.

Key words: BGSA; Feature selection; Gravitational constant function; QSAR study; antiprotozoal activity.

Introduction

Feature selection process is an important step in QSAR studies. Since the selection of appropriate descriptors increases reliability of QSAR model. One of the most recent physics-based heuristic algorithms that are inspired by the law of gravity and mass interactions is Gravitational Search Algorithm (GSA) which was first introduced as a new feature selection method in chemical systems by Mohseni-Bababdani et al. [1]. In this algorithm the searcher agents are a collection of masses which evolve toward better regions of the search. All these masses attract each other by gravity force. Gravitational constant (*G)* is a decreased function of time, which can control the magnitude of the gravitational force by lapse of time. These forces cause a piecemeal movement of all objects towards the object with the heaviest mass. In order to find the global optimum and prevent from trapping in local optima, the gravity force should be great at the beginning (increase of exploration) and by lapse of iteration it should be small (increase of exploitation). This goal can be attained by using the appropriate gravitational constant function.

Method

In the present work, the effect of different gravitational constant functions on the exploration and exploitation ability of the gravitational search algorithm was studied. All assessments were performed by QSAR study of antiprotozoal activity of sixty benzyl phenyl ether diamidine derivatives [2, 3]. In order to generate the most stable conformer, all structures were optimized by PM3 semiempirical method. Then, 1497 descriptors were calculated for each molecule by implementing Dragon software. After evaluation of the descriptors for pairwise correlation and

constant and zero values, 175 descriptors remained for further investigation. In order to find the best subset of descriptors, binary version of the gravitational search algorithm (BGSA) was used. In BGSA agents were considered as objects and their performances were measured by Bayesian regularized artificial neural networks (BRANNs).

Results and discussion

In order to show the effect of *G* value on selected features and accuracy of the search result, four time-decreasing functions were studied in BGSA and all parameters of the algorithm (i.e. G_0, power of R, V_{max} and the number of iterations) were optimized. The selected descriptors were considered as input of BRANNs for modeling. Effectiveness of the selected descriptors was evaluated by statistical parameters including root mean square errors of training ($RMSE_t$), validation ($RMSE_v$), leave-one-out ($RMSE_{LOO}$), leave-five-out ($RMSE_{L5O}$), leave-ten-out ($RMSE_{L10O}$), and coefficient of multiple determinations of training (R_t^2), validation (R_v^2), leave-one-out (Q_{LOO}^2), leave-five-out (R_{L5O}^2), leave-ten-out (R_{L10O}^2) and Y-randomization test (R_r^2). These statistics are given in Table 1.

Table 1. Training, internal and external validation statistical parameters of generated models

G(t)	No.[a]	training		External		Internal validation					Y-randomization	
		R_t^2	$RMSE_t$	R_v^2	$RMSE_v$	Q_{LOO}^2	$RMSE_{LOO}$	R_{L5O}^2	$RMSE_{L5O}$	R_{L10O}^2	$RMSE_{L10O}$	R_r^2
Square root	7	0.994	0.119	0.928	0.424	0.881	0.527	0.877	0.537	0.863	0.565	0.283
Linear	7	0.992	0.138	0.800	0.749	0.966	0.281	0.950	0.341	0.934	0.394	0.215
Exponential	6	0.993	0.118	0.907	0.488	0.950	0.342	0.960	0.306	0.954	0.326	0.179
Logarithmic	7	0.996	0.087	0.877	0.564	0.922	0.426	0.930	0.401	0.931	0.399	0.250

[a]Number of descriptors in the model.

Conclusion

Comparison of the Table 1 results show unsatisfactory behavior of square root function due to high values of $RMSE_{LOO}$, $RMSE_{L5O}$, and $RMSE_{L10O}$ and low values of Q_{LOO}^2, , R_{L5O}^2 and R_{L10O}^2 relative to other functions. In the cases of other functions the obtained results are good and satisfactory for all of them, but the number of selected descriptors by exponential function is lower than the other two functions. In addition the lower value of Y-randomization test (R_r^2) for exponential function indicate superiority of the exponential function for using in BGSA.

References

[1] Behnam MB, Mehdi M. Gravitational search algorithm: A new feature selection method for QSAR study of anticancer potency of imidazo [4,5-b] Pyridine derivatives. Chemometrics and Intelligent Laboratory Systems.2013; 122: 1-11.

[2] Donald A P, Stanislav A B, Svetlana M B, Susan KJ, TanjaW, Todd B, Arvind K, David W B, Karl AW, RetoB, Richard RT. Synthesis and antiprotozoal activities of benzyl phenyl ether diamidine derivatives. European Journal of Medicinal Chemistry.2013; 67: 310-324.

[3] Jean. R.Human African trypanosomiasis, chemotherapy and CNS disease..Journal of Neuroimmunology. 2009; 211: 16–22.

Implementing partial knowledge based on Duality

Hamid Abdollahi[a*], Róbert Rajkó[b] and Elnaz Tavakoli[a]

[a] *Faculty of Chemistry, Institute for Advanced Studies in Basic Sciences, Zanjan, Iran*
[b] *Department of Unit Operations and Environmental Engineering, College Faculty of Food Engineering, University of Szeged, Hungary*

E-mail address: abd@iasbs.ac.ir

Abstract

Self-modeling curve resolution (SMCR) methods resolve data set to a range of feasible solutions assuming minimal constraints (non-negativity) analytically. Such a direct method for a two and three component system have been introduced by Lawton-Sylvestre in 1971 [1] and Borgen in 1985 [2], respectively. How to incorporate additional knowledge during SMCR and How to reduce the extent of feasible regions are among important questions in the SMCR field of study. Recently it has been highlighted that in order to elucidate the principles of existing theory in SMCR methods, geometrical view is required too [3-4]. In this study the well-known duality concept that first has been used by Henry [5] in chemometrics and generalized by Rajko for SMCR [6] is considered. The aim of research is to clarify the similarity between duality and algebraic Complementarity theorem introduced by Sawall et al [7].

Keywords: Self-Modeling Curve Resolution (SMCR); Duality; Complementarity theorem.

1. Introduction

The duality concept provides a very rigorous mathematical relationship between two spaces of a given data matrix. For the First time in the literature of chemometrics Henry [5] has introduced duality relationship with the analytical formulae for a three component data set containing multivariate receptor modeling of compositional data of airborne pollution. Considering that singular value decomposition (SVD) of the data leads to two sets of eigenvectors it has been shown that one set of eigenvectors spans a space in which source compositions are points and source contributions are hyperplanes and this space is dual to the space spanned by the second set of eigenvectors of the data in which source compositions are hyperplanes and source contributions are points. Later the concept of duality has been generalized by Rajko for SMCR which is based on singular value decomposition or principal component analysis (PCA) [6].In the presence of the partial knowledge, duality helps to extract more information of the data set. Another approach which uses partial knowledge of the pure components in order to reduce the feasible solutions is the complementarity theorem which has been introduced by Sawall et al [7]. They investigated that the knowledge of spectrum leads to linear restrictions on the concentration profiles of the other components. The aim of this study is to clarify the similarity between duality and algebraic Complementarity theorem.

2. Results and discussion

For any spectroscopic process that is measured as a bilinear two-way data (**M**) ,decomposition of the response matrix to concentration profile matrix (**C**) and spectral profile matrix (**A**) ,based on Bouguer–Lambert– Beer law is:

$$\mathbf{M}_{\mathbf{I}\times\mathbf{J}} = \mathbf{C}_{\mathbf{I}\times\mathbf{N}}\mathbf{A}^{\mathbf{T}}_{\mathbf{N}\times\mathbf{J}} \qquad (1)$$

Where **M** has dimension (I×J), **C** has dimension (I×N) and $\mathbf{A}^{\mathbf{T}}$ has dimension (N×J). I is the number of measured spectra, J is the number of wavelengths and N is the number of components.

According to SVD:

$$\mathbf{M}_{\mathbf{I\times J}} = \mathbf{U}_{\mathbf{I\times N}}\mathbf{D}_{\mathbf{N\times N}}\mathbf{V}^{\mathbf{T}}_{\mathbf{N\times J}} = \mathbf{X}_{\mathbf{I\times N}}\mathbf{V}^{\mathbf{T}}_{\mathbf{N\times J}} = \mathbf{U}_{\mathbf{I\times N}}\mathbf{Y}^{\mathbf{T}}_{\mathbf{N\times J}} = (\mathbf{X}_{\mathbf{I\times N}}\mathbf{Z}_{\mathbf{N\times N}})(\mathbf{T}_{\mathbf{N\times N}}\mathbf{V}^{\mathbf{T}}_{\mathbf{N\times J}}) = \mathbf{C}_{\mathbf{I\times N}}\mathbf{A}^{\mathbf{T}}_{\mathbf{N\times J}} \quad (2)$$

where $\mathbf{C}_{\mathbf{I\times N}} = \mathbf{X}_{\mathbf{I\times N}}\mathbf{Z}_{\mathbf{N\times N}}$ and $\mathbf{A}^{\mathbf{T}}_{\mathbf{N\times J}} = \mathbf{T}_{\mathbf{N\times N}}\mathbf{V}^{\mathbf{T}}_{\mathbf{N\times J}}$ (3)

So transformation of the **X** and **V** by any invertible transformation matrix like **T** and **Z** to **C** and **A** can be based on Eq. (2), note that $\mathbf{TZ} = \mathbf{I}$ and $\mathbf{Z} = \mathbf{T}^{-1}$.

Generating duality concept for a two component system with a known spectrum of the analyte ($\mathbf{A_i}$) leads to the following equations:

$$\mathbf{t} = \mathbf{A_i}\mathbf{V} \quad (4) \qquad \mathbf{tz} = \mathbf{0} \quad (5)$$

where **t** is a vector of the coordination of the known spectrum in the **V**-space and **z** is the general variable in the **U**-space. Both vectors, **t** and **z** have dimension of 1×2. This equation defines a line in the **U**-space corresponding to the known spectrum in the **V**-space. Knowledge about the known spectrum can be applied as the equality constraint too. Therefor unique solution of the concentration profile for the second component would be obtained [4]. The complementarity theorem represents the mathematical relationship between **t** and it's complementary column of **Z**, or **z**:

$$\mathbf{tz} = 0 \quad (6)$$

This theorem resembles the duality concept. Based on the knowledge in one abstract space, both of them provide information of the remained components on the other abstract space.

References

[1] Lawton W.A, Sylvester E.A. Self-Modeling Curve Resolution, Technometrics. 1971; **13** : 617–633.

[2] Borgen O.S, Kowalski B.R. An extension of the multivariate component-resolusion method to three components. Anal. Chim. Acta. 1985; **174** :1–26.

[3] Akbari M, Abdollahi H. Investigation and visualization of resolution theorems in self modeling curve resolution (SMCR) methods, J. Chemometrics (2013) DOI: 10.1002/cem.2519.

[4] Beyramysoltan S, Abdollahi H, Rajkó R. Newer developments on self-modeling curve resolution implementing equality and unimodality constraints. Anal Chim Acta. 2014; **827**: 1–14.

[5] Henry RC. Duality in multivariate receptor models. Chemom. Intell. Lab. Syst. 2005; **77**: 59–63.

[6] Rajko R. Natural duality in minimal constrained self modeling curve resolution. J. Chemometrics 2006; **20**: 164–169.

[7] Sawall M, Fischer C, Heller D, Neymeyr K. Reduction of the rotational ambiguity of curve resolution techniques under partial knowledge of the factors. Complementarity and coupling theorems. J. Chemometrics. 2012; **26**: 526–537.

Gaussian apodization evolving factor analysis as a novele method for curve resolution of gas chromatography–mass spectrometry datasets and precise local rank detection

Saeed Masoum*, Amir Hossein Alinoori

Department of Analytical Chemistry, Faculty of Chemistry, University of Kashan, Kashan, I.R. Iran

E-mail address: masoum@kashanu.ac.ir

Abstract

Gaussian-weighted moving window evolving factor analysis (GWMW-EFA) has been developed as a Gaussian apodization functions for factor analysis to assess the peak purity of the two-dimensional data. In GWMW-EFA method, rows (spectra) of submatrices which are extracted by fixed-size moving window, have been weighted by a Gaussian window, with the center of evaluated row (spectrum) and full width at half of the maximum (FWHM) equal to only one row width. Therefore, each submatrix mainly (89%) characterizes a single row and by performing factor analysis on this weighted submatrix, the number of principal components for each evaluated row, instead of all rows of a submatrix, is determined. Hence, the problem of time shifting in fixed-size moving window evolving factor analysis (FSMW-EFA) can be solved by this method. The results revealed that the GWMW-EFA algorithm could obtain proper initial estimation of elution profiles for curve resolution purposes in combination with alternating least-squares method.

Keywords: Gaussian-weighted moving window evolving factor analysis, Gaussian apodization functions, Local rank map, Curve resolution, Gas chromatography–mass spectrometry

Introduction

In chromatographic technique, factor analysis has been firstly used as a method for determining a number of pure components and separating their mass spectra [1]. Nowadays, variety of methods based on factor analysis has been successfully used to solve the problems arising from GC-MS analysis of real samples with complex matrices [2-4]. All of these methods contain one similar stage that involves applying PCA to a section of data to determine the eigenvalues of it. Through these techniques, the number of components can be mathematically estimated without any presumptions.

Theory

GWMW-EFA is as a Gaussian apodization functions for factor analysis. In GWMW-EFA, submatrices were extracted from a data matrix along of the elution time direction by moving a fixed size window with a default window size (*ws*). Submatrices in GWMW-EFA with any number of rows (*ws*) are weighted by Gaussian window with FWHM equal to only one row width and then they are factor analyzed, whereas all rows of the submatrices with the same weight are analyzed in FSMW-EFA. Based on this difference, GWMW-EFA exhibits higher accuracy in the determination of the start or end points of each component because the weighted moving window characterizes almost one row and the other rows have negligible weights. Thus, GWMW-EFA performance can be independent from window size (*ws*), and the problem of time shift and reduction of temporal resolution in FSMW-EFA can be solved completely in GWMW-EFA (Figure 1).

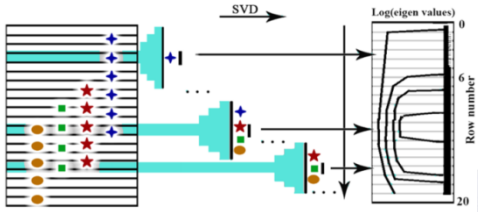


Figure 1. Diagram of GWMW-EFA algorithm. The submatrices are weighted by a Gaussian window with FWHM equal to only one row width

Results and discussion

The peak purity of three simulated GC-MS peak clusters were investigated which results in the local rank analysis map. In this regard, both GWMW-EFA and FSMW-EFA methods were performed and their results were compared. The elution profiles of three simulated GC-MS data and their corresponding factor analysis plots from FSMW-EFA with the window size of five and GWMW-EFA with sigma (σ) equal to 0.4246 are shown in Figure 2. In FSMW-EFA plots (Figure 2 b, e and h), regions of the achieved local rank map are greater than actual size and time shift in start point of each region of the rank map can be detected. FSMW-EFA gives same weight to all selected rows in a submatrix, and can easily cause inaccurate determinations of start and end points of each component. In GWMW-EFA plots (Figure 2 a, d and g), each region of the rank map was detected correctly and there is no difference between real rank maps and achieved local rank maps by GWMW-EFA. By GWMW-EFA algorithm, the rank of each row of a data matrix can be estimated more precisely. Therefore, the proposed algorithm could be more reliable for the peak purity assessment or the local rank analysis of complicated peak clusters.

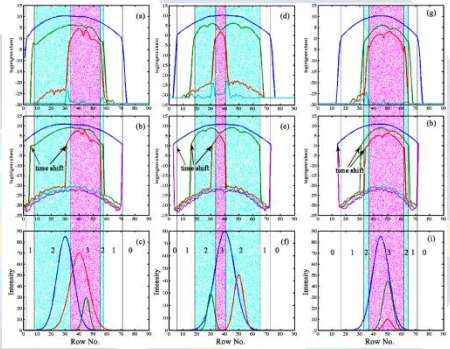


Figure 2. Obtained eigenvalues by GWMW-EFA and FSMW-EFA

References

[1] Sharaf MA, Kowalski BR. Extraction of individual mass spectra from gas chromatography-mass spectrometry data of unseparated mixtures. *Anal. Chem.* 1981; **53**: 518-522.

[2] Masoum S, Ghasemi-Estarki H, Seifi H, Ebrahimabadi EH, Parastar H. Analysis of the volatile chemical constituents in Mindium laevigatum by gas chromatography-Mass spectrometry and correlative chemometric resolution methods. *Microchem J.* 2013; **106**: 276-281.

[3] Seifi H, Masoum S, Seifi S, Performance assessment of chemometric resolution methods utilized for extraction of pure components from overlapped signals in gas chromatography-mass spectrometry, *J. Chromatogr. A* 2014; **1365**: 173-182

[4] Jalali-Heravi M, Moazeni-Pourasil RS, Sereshti H. Elimination of chromatographic and mass spectrometric problems in GC-MS analysis of Lavender essential oil by multivariate curve resolution techniques: Improving the peak purity assessment by variable size moving window-evolving factor analysis. *J. Chromatogr. B* 2015; **983-984**: 83-89.

Identification of potential antioxidant compounds in Walnut (Juglansregia L.) leaves extract by linear multivariate calibration techniques and HPLC-DAD-MS

Saeed Masoum [a]*, Naser Samadi [b], Bahare Mehrara [b]

[a] *Department of Analytical Chemistry, Faculty of Chemistry, University of Kashan, Kashan, Iran*

[b] *Department of Chemistry, Faculty of Science, Urmia University, Urmia, Iran*

E-mail address: masoum@kashanu.ac.ir

Abstract

The leaves of Juglans regia L. (Walnut) are used in traditional medicine in world due to its antifungal, antibacterial and antioxidant properties. In this study, the antioxidant activity in Walnut extracts was measured using 1,1-diphenyl-2-picrylhydrazyl (DDPH) assay and qualitative analysis was performed by HPLC-DAD-MS. The peaks potentially responsible for the antioxidant activity in Walnut samples were indicated by some linear multivariate calibration techniques with different preprocessing methods. From the studied techniques, orthogonal projection to latent structures (OPLS) was preferred to exhibit the potential antioxidant active compounds in Walnut leaves extract because of its high repeatability, simplicity and improved interpretability of the regression coefficients.

Keywords: Chromatographic fingerprints, Juglans regia L., Potentially antioxidant compounds, Linear multivariate calibration

Introduction

Walnuts (Juglans regia L.) have anti-oxidant and anti-inflammatory properties. Walnut leaves have been widely used in pharmaceutical industries for their antidiarrheic, antihelmintic, depurative and astringent properties [1, 2]. Several natural antioxidant compounds such as flavonoids, coumarins, curcuminoids, tannins, xanthons and terpenoids are found in the fruits, leaves and seeds of various plants. Phenolic compounds are the most active natural antioxidants in Walnuts.

In this study, the antioxidant activity of Walnut leaves extract was determined using a 1,1-diphenyl-2-picrylhydrazyl (DPPH) radical scavenging activity test. By high performance liquid chromatography-mass spectrometry (HPLC-MS) and linear multivariate calibration techniques, important compounds that contribute in the antioxidant properties of Walnut samples could be determined.

Results and discussion

Sixteen Walnuts samples were collected from sixteen regions of Iran at the end of July 2013. A data matrix with dimensions of 16 × 5525 was obtained from the chromatogram of Walnut. Response vector indicates the DPPH scavenging activities of 16 Walnut samples. Several multivariate calibration models for the antioxidant activity were developed using PCR, PLS, OSC-PLS and OPLS techniques, on the data matrix **X** containing 16 fingerprints and the response vector **y** (DPPH radical scavenging test results), representing the IC_{50} test results [3-5]. Since the DPPH scavenging test value decreases by increasing antioxidant activity of the sample, the negative peaks are corresponding

to the components with antioxidant activity, therefore these negative coefficients play greater roles compared to the other coefficients in the antioxidant activity. The peaks, which were potentially responsible for the antioxidant activity, could clearly be recognized in the fingerprints of the active samples. Based on the OPLS plots, compounds that elute at 4.95, 13.20, 14.30, 14.53, 14.72, 15.27 and 15.95 min are responsible for the Walnut antioxidant activity. Identification of these compounds has been done by HPLC–MS analysis (Figure 1).

In this work, PCR, PLS and OSC-PLS created some difficulties in interpreting the regression coefficients due to the presence of small irrelative positive and negative peaks. It was found that OPLS was the better performing technique in indicating the potential antioxidant active compounds in the Walnut leaves extract because of its simplicity and repeatability, and in order to remove the orthogonal information in the original data. OPLS caused the model complexity to be reduced, which in turn resulted in improved interpretability of the regression coefficients.

Seven compounds were indicated as being potentially antioxidant using HPLC–MS. In this study, regarding to the retention times, Monogalloyl-glucose**,** Digalloyl, 1,2,3,4,6 penta galloyl-glucose, Tellimagrandin II, Valoneic acide dilactone, Strictinin and (+)–Catechin are suggested as the Walnut potentially antioxidant compounds.

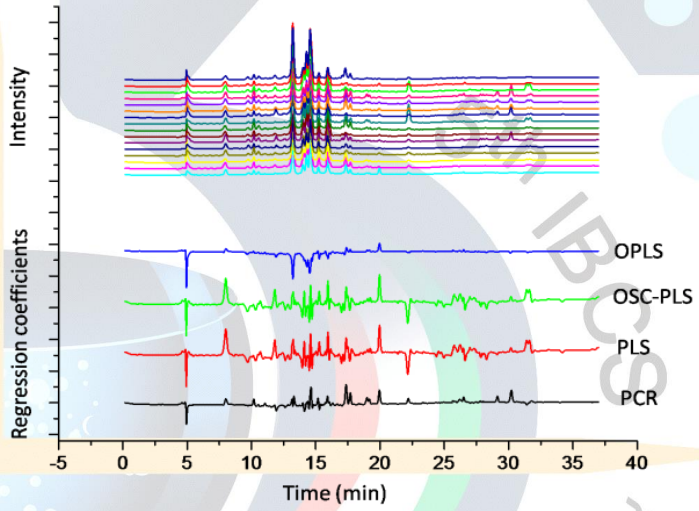


Figure 1. Fingerprints plot along with the obtained regression coefficients from some linear multivariate calibration methods

References

[1] Oliveira I, Sousa A, Ferreira IC, Total phenols, antioxidant potential and antimicrobial activity of walnut (Juglans regiaL.) green husks, *Food Chem. Toxicol.* 2008; **46**: 2326–2331.

[2] Pereira JA, Oliveira I, Sousa A, Bioactive properties and chemical composition of six walnut (Juglans regia L.) cultivars, *Food Chem. Toxicol*. 2008; **46**: 2103–2111

[3] Martens H, Martens M, Multivariate Analysis of Quality, An Introduction, Wiley, Chichester, 2001.

[4] Trygg J, Wold S, Orthogonal projections to latent structures (O-PLS), *J. Chemometr*. 2002; **16**: 119-128.

[5] Tistaert C, Dejaegher B, Nguyen Hoai N, Chataigné G, Rivière C., Nguyen Thi Hon V, Chau Van M, Quetin-Leclercq J, Vander Heyden Y, Potential antioxidant compounds in Mallotusspecies fingerprints. Part I: indication, using linear multivariate calibration techniques, *Anal. Chim. Acta* 2009; **649**: 24-32.

Potentiality of independent component regression in assessment of the peaks responsible for antioxidant activity of Thyme using gas chromatography-mass spectrometry technique

Saeed Masoum*, Narjes Ghasemi

Department of Analytical Chemistry, Faculty of Chemistry, University of Kashan, I. R. Iran

E-mail address: masoum@kashanu.ac.ir

Abstract

Thyme species are medicinal herbs that are famous throughout the world and are used as antioxidant, antimicrobial, antispasmodic and antiseptic. The purpose of this study is identification of potential antioxidant compounds in the essential oils of Thyme. For this purpose, the obtained chromatogram and antioxidant activity of essential oils were used. The antioxidant activity of essential oils was measured by DPPH radical scavenging test. By coupling of chromatographic data and antioxidant activity assisted by independent component regression (ICR) can identify the peaks potentially responsible for antioxidant activity. The results of this technique were displayed as a regression graph, which peaks with negative regression coefficients cause antioxidant properties. Finally, GC–MS was used to indicate the structure of chemical components, which are responsible for antioxidant activity.

Key words: Thyme, Antioxidant activity, Independent component regression, Gas chromatography-mass spectrometry

Introduction

Today, herbal therapies are used widely throughout the world. To ensure the productivity and safety of medicinal plants in the treatment of diseases, it is necessary to develop standard methods for assessment of their quality. Obtained results by fingerprint methods along with chemometric analyses lead to obtain additional information such as classification of samples and evaluation of the peaks responsible for their desired activity [1, 2]. Since independent component analysis (ICA) is commonly considered to be a further development of principal component analysis (PCA), a similar regression method based on ICA, independent component regression (ICR), is proposed by Chen and Wang [3]. In this work, ICR method is introduced and applied to determine important compounds that contribute in the antioxidant property of Thyme.

Results and discussion

Fingerprints of thirty samples of thyme were obtained according to GC temperature programming and are shown in Figure 1. Data matrix **X** with dimension of 30×4190 was obtained from the chromatograms of thyme samples. Response vector (**y**) indicates the DPPH scavenging activities of this medicinal herb. Different preprocessing methods such as normalization, mean centering and scaling were performed for evaluation of antioxidant activity. Normalization followed by mean centering shows the best result for applied multivariate calibration technique.

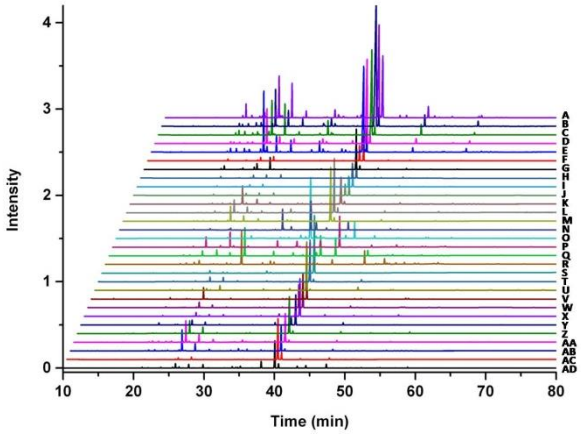


Figure 1. Fingerprints of the essential oils of thyme obtained by GC

Assessment of the peaks responsible for antioxidant activity of this medicinal herb was investigated. By increasing the antioxidant activity of a sample the DPPH scavenging test value decreases, so, the components with antioxidant activity show negative peaks. The corresponding regression coefficients are positive for the compounds with opposite behaviors. The regression coefficients that are obtained by ICR have been plotted in Figure 2.

To identify these potential antioxidant compounds, the GC-MS analyses have been applied. Based on the GC–MS analeses and according to ICR plot seven components such as α-Pinene, β-myrcene, p-Cymene, γ-Terpinen, Thymol Carvacrol and Caryophyllene are located in the areas that are corresponding to negative peaks in the regression plot of thyme. Some researces reported these components as strong antioxidant.

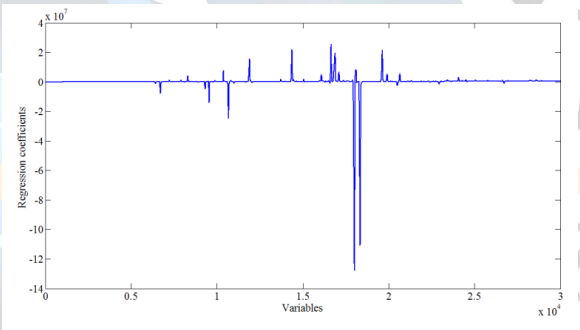


Figure 2. Obtained regression coefficients by ICR

References

[1] Tistaert C, Dejaegher B, Nguyen Hoai N, Chataigné G, Rivière C., Nguyen Thi Hon V, Chau Van M, Quetin-Leclercq J, Vander Heyden Y, Potential antioxidant compounds in Mallotusspecies fingerprints. Part I: indication, using linear multivariate calibration techniques, *Anal. Chim. Acta* 2009; **649**: 24-32.

[2] Masoum S, Mehran M, Ghaheri S, Identification of potential antioxidant compounds in the essential oil of thyme by gas chromatography with mass spectrometry and multivariate calibration techniques, *J. Sep. Sci.* 2015; **38**: 410-417.

[3] Chen J, Wang XZ, A new approach to near-infrared spectral data analysis using independent component analysis, *J. Chem. Inf. Comput. Sci.* 2001; **41**: 992-1001.

Advantage of multivariate calibration methods to introduce carbon dots as a novel pH sensor

Ali Barati*[1], Hamid Abdollahi[1], Mojtaba Shamsipur[2]

[1]*Faculty of Chemistry, Institute for Advanced Studies in Basic Sciences, Zanjan, Iran*
[2]*Department of Chemistry, Razi University, Kermanshah, Iran*
ali_barati@iasbs.ac.ir

Abstract

In this study, we prepared N-doped carbon dots (CDs) by the facile hydrothermal preparation using glucos as the carbon source. The prepared CDs show excitation-emission dependency that is typical for many reported CDs. we investigate the pH sensitivity of CDs and realized that emission spectra at different excitation wavelengths show different behavior toward the pH value. Beside the different changes in their emission intensity, different shift in their maximum toward lower or higher wavelengths were also observed respect to the pH. Neither emission intensity nor spectra shifts were not suitable signals to calibrate the related pH values in a broad range. In addition, ratio signals were proper only for a limited pH range.

However, recording the excitation –emissiom (Ex-Em) matrix of CDs at each pH value and utilizing multivariate calibration methods such as PLS and ANN helped to calibrate the pH values in the broad range of 2-14 with a high sensitivity and acuaracy. Using multivariate calibration not only take the advantage of ratio signal methods (which are usually free from the interfering quenchers) but also extend the range of the calibrated pH and the accuracy of the mesurements. This study would be the firt report on application of multivariate calibration methods to calibrate the pH dependency of nanoparticles signals.

Key words: Carbon dots, pH sensing, PLS, ANN

Introduction

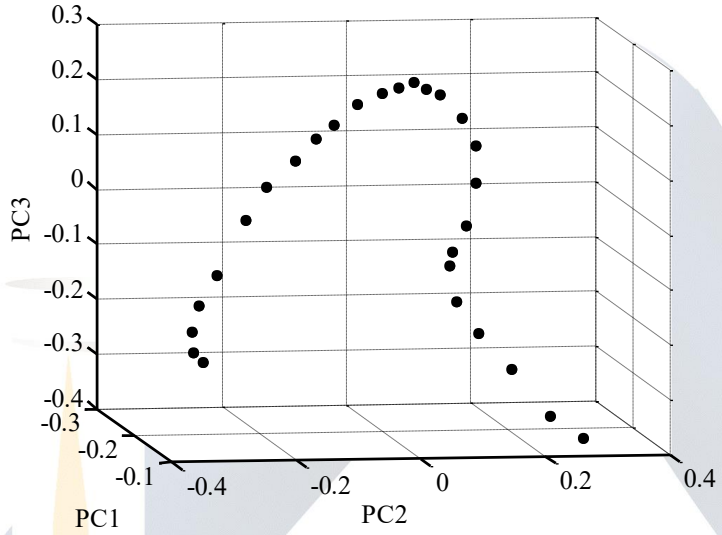
The pH sensors are found frequently in a wide range of applications required in various fields of science and technology such as chemical process control, medical diagnosis, and environmental analysis, as well as industrial applications. There are mainly two kinds of pH sensors such as electrochemical sensor and lately fiber optical sensor.

Although electrochemical pH sensors are well-established and can be used as reliable tools for a large number of analytical tasks, a number of disadvantages of the pH electrode including frequent calibration, the susceptibility to electrical interference, and corrosion by alkaline solutions or fluoride ions limit its usefulness. In the last 3 decades, many efforts have been directed toward the development of optical pH sensors, especially fiber optic pH sensors [1-2]. Optical pH sensors are based on reversible changes in the indicator's structures induced by pH and translated into changes in spectroscopic phenomena such as absorption, reflectance, and luminescence [3,4]. Among these techniques, the fluorescent pH sensor has gained considerable attention due to its particularly high sensitivity.

Many sensitive probes such as organic dyes, transition metal complexes and nanoparticles have been prepared and used to monitor pH values. Although they are sensitive, their fluorescence intensities may vary due to slight changes in the surrounding matrix, therefore ratiometric fluorescent pH sensors have been used. Therefore, searching for fluorophores with good properties including excellent photostability, and high quantum yield which respond to a wide pH range for optical sensing is a challenge for the research efforts in analytical chemistry.

Results and discussion

The score plot for the first three scores of the vectorized Ex-Em data matrix of CDs at different pH values are presented below. The pH values from left to right are 2, 3, 4, 5, 5.5, 6, 6.2, 6.4, 6.6, 6.8, 7, 7.2, 7.4, 7.6, 7.8, 8, 8.5, 9, 9.5, 10, 10.5, 11, 11.5, 12, 12.5, 13, and 14. As con be seen the point are weel separated, which indicate that a proper multivariate method can calibrate them to their related pH values.



References

[1] Wang, X. D.; Wolfbeis, O. S. Fiber-Optic Chemical Sensors and Biosensors (2008–2012) *Anal. Chem.* 2013, **85**, 487−508.

[2] Aigner, D.; Borisov, S. M.; Petritsch, P.; Klimant, I. Novel near infra-red fluorescent pH sensors based on 1-aminoperylene bisimides covalently grafted onto poly(acryloylmorpholine) *Chem. Commun.* 2013, **49**, 2139−2141.

[3] Cajlakovic, M.; Lobnik, A.; Werner, T. Stability of new optical pH sensing material based on cross-linked poly(vinyl alcohol) copolymer *Anal. Chim.* Acta 2002, **455**, 207−213.

[4] Aigner, D.; Ungerböck, B.; Mayr, T.; Saf, R.; Klimant, I.; Borisov, S. M. Fluorescent materials for pH sensing and imaging based on novel 1,4-diketopyrrolo-[3,4-c]pyrrole dyes *J. Mater. Chem. C* 2013, **1**, 5685−5693.

An Approach to Determine Accurate Model Parameters of Buffered Chemical Processes

Hamid Abdollahi*[a], Fereshteh Emami[a], Mohammad Javad Masroor[a]

[a]*Department of Chemistry, Institute for Advanced Studies in Basic Sciences, Zanjan, Iran*.
email address:abd@iasbs.ac.ir

Abstract

The pH plays an important role in many reactions in aqueous solution. In order to be able to quantitatively analyses reactions that involve shifting protonation equilibria, reaction solutions traditionally needed to be buffered to maintain constant pH. However, in some cases buffered solutions cannot be stayed at the expected pH, and pH is changed during the reaction. So assuming constant pH and neglecting the corresponding buffer equilibria make results inaccurate. We propose a general computational method, which allows analysis of reactions at non-expectable non-constant pH in the presence of buffer. In this work, the new concept of generalized equilibrium and kinetic study of thermodynamic parameters is introduced for dynamic buffered data. The examples include complexaction reaction at a assumed constant pH by the buffer, pH-dependent degradation of chemical compound, Ribpflavinin constant pH solutions.

Keywords: dynamic buffered data, Ribpflavinin, non-constant pH, non-ideal buffer behavior

1. Introduction

Spectrophotometric data that describes a chemical process can be mathematically analysed by using two different approaches: 1) hard_modeling based methods, which require a chemical model describing the process under investigation; 2) soft modeling based methods for which no the kinetic or equilibrium model link to the reaction studied is required [1]. Hard_modeling analysis provides basic parameters of process under study [2]. The pH plays a crucial role in most chemical processes in aqueous solution. In particular, many reaction rates in aqueous solution show a strong pH dependence [3,4]. Problems arise if some of the equilibrium components or the species formed from the components react in secondary, slow chemical reactions. Then it is impossible to reach equilibrium between the equilibrium species. The idea of this contribution is analysing the process considering both the kinetics and equilibrium processes.Software for the data analysis of kinetic processes with coupled instantaneous equilibria, usually protonation equilibria, has been developed and applied extensively [5]. We propose to utilize these analysis tools for a novel application, namely the determination of equilibrium and kinetic constants for processes that are affected by buffer non-ideal behavior.

2. Results and discussion

In this study, a more complete model than previous models has been used to determine the thermodynamic and kinetic parameters (different equilibrium and kinetic constants) in which there is a buffer non-ideal behavior. To provide a more complete model, the behaviors of the buffer and water with other equilibria and kinetics in the system have been considered leading to more complete and accurate results.

For the investigated example the task is: determining complex formation constant and kinetic rate constants in the photodegradation of Riboflavin in acetate buffer.

The mechanism of the photodegradation of Riboflavin in acetate buffer can be considered as follows:

$$RF \overset{k_1}{\rightarrow} FMF \overset{k_2}{\rightarrow} LC$$
$$\updownarrow K_3$$
$$FMF^- + H^+$$

K_3, The equilibrium constant of complex formation , $k_1$and k_2 the kinetic rate constants, RF (Riboflavin) , FMF (Formyl methyl flavin) and LC (Lumichrome)

The first model: equilibrium chemical reactions linked with kinetic consecutive procedures is considered to be, but it is assumed that the concentration of $[H^+]$ is a constant because the buffer is added to the solution.

The data was simulated by real models and is fitted by first model. As it is illustrated in Fig. 1, concentration profiles extracted from the fitting using first model with the concentration profiles used in the model is totally different. Real and estimated parameters of the model are shown in (Table 1).

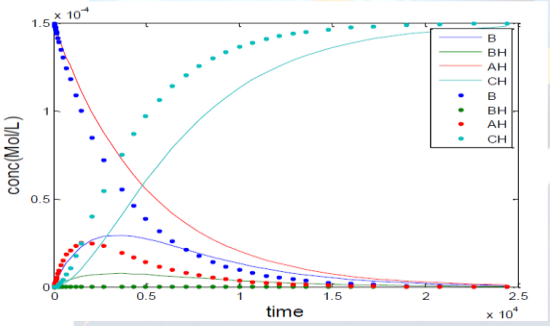


Figure 1: Concentration profiles obtained using first model (Point) and Real (Line)

The complete model: in the complete model the behavior of the buffer and water with other equilibria and kinetics in the system are considered leading to more complete and accurate results.

The data was simulated by real models and is fitted by complete model. As it is shown in fig. 2, concentration profiles extracted from the fitting using complete model with the concentration profiles used in the model covary the same. Results of the analysis are summarized in (Table 1).

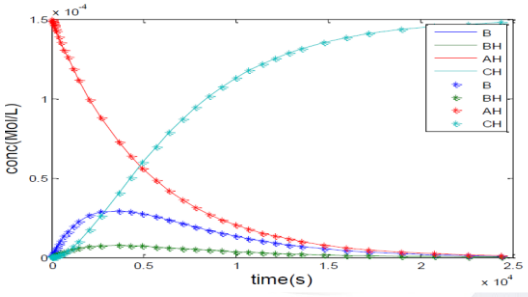


Figure 2: Composition of results concentration profiles obtained using complete model (Point) and Real (Line)

Table 1: Real and estimated parameters for the first model and complete model

Real values	pH	first model RF→FMF→LC ; ↕ FMF+H	Relative error % In first model	complete model RF→FMF→LC ; ↕ FMF+H H+OH↔H_2O; B+H↔BH	Relative error % In complete model
$K_۲$ =۳/۱۶×۱۰^{۳}	۴/۷۶	$K_۲$ =۴/۴۳×۱۰^{۳}	۲۸	$K_۲$ =۳/۱۶×۱۰^{۳}	۰/۰۰۳۲
$k_۱$ =۲×۱۰$^{-۴}$		$k_۱$ =۲/۷۴×۱۰$^{-۴}$	۳۷	$k_۱$ =۱/۹۹×۱۰$^{-۴}$	۰/۰۳
$k_۲$ =۲×۱۰$^{-۳}$		$k_۲$ =۱/۲×۱۰$^{-۳}$	۴۰	$k_۲$ =۱/۹۷×۱۰$^{-۳}$	۱

References

[1] Juan A. de, Casassas E., Tauler R., Soft modeling of analytical data, Encyclopedia of Analytical Chemistry: Instrumentation and Applications,Wiley, New York,2000.

[2] Norman S.,Maeder M., Crit.Rev.Anal.Chem,2006,36,199-209

[3] R. G. Wilkins, Kinetics and Mechanism of Reactions of Transition Metal Complexes, VCH, Weinheim, 1991.

[4] J. H. Espenson, Chemical Kinetics and Reaction Mechanisms, McGraw-Hill, New York, 1995.

[5] F. Emami, H. Abdollahi and M. Maeder, Model-based analysis of coupled equilibrium-kinetic processes Analyst, 2015, 140, 3121.

Chemometrical study of spectral curve fitting constraint on self-modeling curve resolution methods

Hamid Abdollahi*[a], Somayeh khodadadi[a]

[a] *Department of chemistry, Institute for Advanced Studies in Basic Sciences(IASBS), Zanjan, Iran.*

E-mail address: abd@iasbs.ac.ir

Abstract

In the current work the effect of apply model for spectral profiles have been investigated. Hard modeling techniques are understood as methods that exploit the physical structure of spectra by modeling the data with peak shaped functions, i.e. by peak fitting. For the resolve spectral data by hard modeling, different methods such as Complemental Hard Modeling (CHM) and Hard Modeling Factor Analysis (HMFA) have been suggested. Although using the chemical model for concentration profile lead to unique solution, the mathematical model for spectral profile can fit all feasible bands. We found that the curve fitting constraint is not a real constraint that has not a unique solution.

Key words: Hard modeling, Curve fitting, Mathematical model, Real constraint.

Introduction

Today, spectroscopy is the powerful measurement technique utilized in diverse industrial applications. For the analysis of the spectral raw data, different methods such as curve fitting are applied. The idea of spectral hard modeling is to represent peaks by certain mathematical function with some parameters and optimize these parameters to approximate the actual spectrum. Spectral line profiles can be represented as sum of distribution function like the Voigt function, which is the mathematical convolution of Gaussian and Lorentzian functions [1].

Hard modeling methods are, however, only applicable if the characteristic bands for quantitation can unambiguously be assigned to the different species in the mixture. For strongly overlapping bands, this may not be possible. By the way, there are certain situations where all existing methods fail, namely for reactive mixtures with molecular interactions and strongly overlapping bands, i.e., bands that may not readily be assigned to the components of interest. To overcome these limitations Marquardt and his coworkers have introduced a novel analysis method, called indirect hard modeling (IHM), which combines the advantages of multivariate linear methods with those of hard modeling techniques. IHM represents a mixture spectrum as a weighted sum of parameterized pure component models built by hard modeling [2]. In 2008 Marquardt and his coworkers introduced two methods for identification of unknown pure component spectra based on indirect hard modeling. The first method is Complemental hard modeling (CHM). CHM is a novel method for the identification of a single unknown pure component spectrum in an otherwise completely known mixture. It only requires a single mixture spectrum and the remaining pure component spectra as inputs. The second method is Hard

modeling factor analysis (HMFA). HMFA is a method for the identification of all pure spectra in unknown chemical systems [1].

Results and discussion

For The CHM method in the first step, an initial nonlinear IHM fit is performed. In contrast to CLS, a parameterized flexible spectral model is fitted to the data.

$$\mathcal{X}(v, w_A, \boldsymbol{\theta}) = \mathcal{B}(v, \boldsymbol{\theta}_B) + w_A \mathcal{S}_A(v, \boldsymbol{\theta}_P, \theta_S)$$

This model corresponds to the known part of the full IHM model. By solution of the least-squares problem the model is fitted by adjusting the pure component weight w_A, the baseline parameters θ_B and the shift parameter θ_S. The peak parameters of the known model θ_P are fixed. The fitting error is plotted, at the position of the highest fitting error; a new peak function is added in the next step. The complemented model is then fitted by solution of the nonlinear least-squares problem.

$$\min_{w_A, \boldsymbol{\theta}_B, \boldsymbol{\theta}_P^c} \sum_{i=1}^{n_v} \left[x_i - \mathcal{X}(v_i, w_A, \boldsymbol{\theta}_B, \boldsymbol{\theta}_P, \theta_S, \boldsymbol{\theta}_P^c)\right]^2$$

The weights of the known pure spectra w, the baseline parameters θ_B and the peak parameters θ_P of all complemental peak functions are adjusted at the same time. We observed that even in the wrong value of W_A, when residual has the contribution of known profile, we have a perfect fit.

The HMFA method starts with the estimation of peak functions by hard modeling of a representative mixture spectrum. A representative mixture spectrum reflecting the contributions of all components is generated out of data for the hard modeling step. Beacuse the representative mixture spectrum of all feasible band solutions are the same, and we can fit all feasible bands, the solution of HMFA resolving data is not unique.

We found that the curve fitting constraint is not a real constraint that has not a unique solution. It means that apply mathematical model for spectral profiles did not guarantee the unique real solution.

References

[1] E. Kriesten , D. Mayer , F. Alsmeyer , C.B. Minnich , L. Greiner , W. Marquardt . Identification of unknown pure component spectra by indirect hard modeling. *Chemometr. Intell. Lab. Syst*.. 2008; **93** :108–119.

[2] FRANK ALSMEYER, HANS-JU¨ RGEN KOß, and WOLFGANG MARQUARDT. Indirect Spectral Hard Modeling for the Analysis of Reactive and Interacting Mixtures. *Appl. Spectrosc.* 2004; **58** :975–985.

Sargassum Glaucescens for the removal of Methylen blue, Crystal Violet and Safranin in ternary system by artificial neural network modeling, Plackett–Burman design and response surface analysis

Reza Tabaraki*, Ashraf Nateghi

Department of Chemistry, Ilam University, Ilam, Iran

rezatabaraki@yahoo.com

Abstract

Biosorption properties of three cationic dyes with different molecular sizes (Methylene blue (MB), Crystal violet (CV) and Safranin (SF)) on Sargassum glaucescens were studied. Principal component-wavelet neural network (PC-WNN) was used for the simultaneous determination of MB, CV and Safranin concentrations in ternary solutions. Plackett–Burman design was applied to identify the most significant factors in the removal of dyes in ternary system. The interaction between the factors and their optimum levels for maximum removal of dyes were determined using Box–Behnken design. The optimum biosorption conditions were determined as initial dye concentration 10^{-4}mol/L, biosorbant dosage 0.1 g/L and biosorbant particle size 0.1875 µm. At this condition 0.819 mmol/g biosorption capacity was predicted by the RSM model.

Keywords: Dye removal, artificial neural network, Plackett–Burman design, Box–Behnken design, Ternary system

Introduction

Dyes are synthetic chemical compounds having complex aromatic structures which are extensively used in the textile, cosmetic, plastic, food and pharmaceutical industries [1]. The dye-containing wastewater discharged from the industries can adversely affect the aquatic environment by impeding light penetration. Moreover, most of the dyes are toxic, carcinogenic and harmful to human health [2]. Biosorption has been found to be one of the prominent techniques for dye wastewater treatment in terms of cost and operation [3]. In this study, adsorption performance of *S. glaucescens* in the removal of basic dyes, namely, MB, CV and Safranin in ternary dye solutions were studied. The interaction between the factors was studied and optimized using Box–Behnken design under response surface methodology.

Results and discussion

The simplest and the most common method is a top-down variable selection where the PCs are ranked in the order of decreasing eigenvalue. The PC with highest eigenvalue is considered as the most significant one and subsequently, the PCs are introduced into the WNN model one after the other. WNN models were developed using different number of the PCs in input layer. The WNN variables consisted of the number of

PCs as an input layer, the number of nodes in the hidden layer, the learning rate, the momentum and the number of epochs, which optimized for each dye separately. The optimized variable parameters for each dye are given in Table 1.

Analysis of variance (ANOVA) is an essential tool for determining the significance of an effect or of a mathematical model. The most significant factors can be determined by using a statistical parameter, which is the P value (Table 2). From these results, it was found that the effects of initial concentration of dye, biosorbant dosage and biosorbant particle size are the most important and significant factors for dyes uptake. The other factors are not important and can be considered negligible. The calculated regression equation for the optimization of medium constituents showed that dyes maximum uptake (Y) was related with the function of biosorbant dosage (X_1), initial dye concentration (X_2) and biosorbant particle size (X_3). For the three factors which were studied, the Box–Behnken model efficiently designed a second order response fit for the surface. Second-order polynomial model for q (mmol/g) was: $q\ (mmol/g) = 0.388 - 0.211\ X_1 + 0.263\ X_2 - 0.111\ X_3 - 0.074\ X_{12} + 0.129\ X_{13} - 0.045\ X_{23} + 0.086\ X_1^2 - 0.072\ X_2^2 + 0.025\ X_3^2$ $R^2 = 95.2\%$

The optimized conditions were initial dye concentration 10^{-4}mol/L, biosorbant dosage 0.1 g/L and biosorbant particle size 0.1875 µm. At this condition 0.819 mmol/g biosorption capacity was predicted by the RSM model. This result was validated experimentally (0.798 mmol/g), which is close to the predicted value.

Table 1. The optimized parameters of PC-WNN models.

Parameters	Safranin	MB	CV
Input neurons	3	3	3
Hidden neurons	2	2	2
Output neurons	1	1	1
Learning rate	0.058	0.067	0.051
Momentum	0.87	0.22	0.2
Number of iterations	1000	3000	2000
Hidden transfer function	Morlet	Morlet	Morlet
Output transfer function	Linear	Linear	Linear

Table 2. Effects and coefficients for maximum uptake (q) of dyes.

Factor	Effect	Coefficient	SE Coefficient	P
Constant		0.25522	0.01815	0.000
biosorbant dosage	-0.10958	-0.05479	0.02029	0.024
dye concentration	0.34251	0.17126	0.02029	0.000
pH	0.00632	0.00316	0.02029	0.880
time	0.06358	0.03179	0.02029	0.152

particle size	-0.08805	-0.04403	0.02029	0.058

References

[1] E. Forgacs, T. Cserhati, G. Oros, Environ. Int. 2004, 30, 953-971.

[2] V. Vimonses, B. Jin, C.W.K. Chow, J. Hazard. Mater. 2010, 177, 420-427.

[3] X.J. Xiong, X.J. Meng, T.L. Zheng, J. Hazard. Mater. 2010, 175, 241-246.

Comparison of discriminatory power from soft-modeling methods coupled with multivariate pattern recognition technique for the identification of similar compounds

Hamid Abdollahi* and Fereshteh Emami and Davod Kazemi

Department of Chemistry, Institute for Advanced Studies in Basic Sciences, P.O. Box 45195-1159, Zanjan, Iran
E-mail address: abd@iasbs.ac.ir

Abstract

Pattern recognition techniques are employed to abstract multidimensional patterns (for a $n \times d$ matrix, n points in d dimensions) into known or possibly unknown categories. Soft modeling methods are based on the linear relationship between response and independent parameter like concentration. Coupling of soft-modeling methods with multivariate pattern recognition technique was proposed as an efficient approach that takes advantage of both soft-modeling and pattern recognition methods. Considering the fact that rotational ambiguity is the permanent part of soft-modeling methods, in this contribution discrimination ability of different possible solution in feasible regions is assessed. Finally, interaction between L&D-Cysteine and Au-NP monitored spectroscopically and for the recorded data set proposed strategy is checked in case of the L&D-Cysteine discrimination using different possible solution in feasible regions.

Keywords: Self modeling / multivariate curve resolution; pattern recognition; Gold Nanoparticle; Cysteine.

1. Introduction

There are many possible techniques for classification of data arrays. Principal component analysis (PCA) and linear discriminant analysis (LDA) are two commonly used techniques for data classification and dimensionality reduction. Coupling of soft-modeling and pattern recognition is recommended by Emami and Hasani which bear both advantages of soft-modeling and pattern recognition methods[1]. Considering the fact that soft-modeling results are always complicated with rotational ambiguity, investigation of discrimination ability for all possible solution in a feasible region is important issue. Cysteine (abbreviated as Cys or C) is an α-amino acid and is a semi-essential amino acid. Cysteine, mainly the L-enantiomer, is a precursor in the food, pharmaceutical, and personal-care industries. One of the largest applications is the production of flavors. Unmodified Au-NPs is incorporated as a chiral selector for D- and L- Cysteine (Cys).

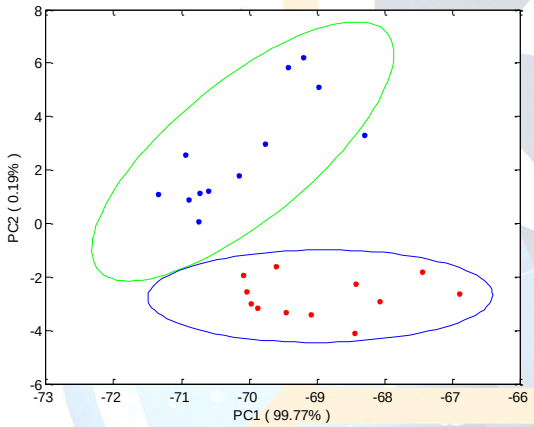
Finally, in order to determine discrimination of classes for different given solutions, confidence ellipse is used as a criterion [2].

Data collection

For each sample, 36 spectra were sequentially recorded every1 min in the wavelength ranges 200–800 nm during 36 min. The size of each spectral-time data matrix was 36×617. For each sample of cyst isomer with the concentrations of 24.3, 34.3, and 44.3 micro molar, spectra was recorded four times.

2. Results and discussion

L&D-Cys interact with Gold Nanoparticles in the consecutive reaction with different rate constant and resulted in identical products. Difference in the rate constant results in different concentration profiles. In the first step, recorded data matrices are vectorized and it is followed by application application of principal component analysis for the visualization in the abstract space. As it is shown in fig.1, L&D-cys groups are well-separated from each other in the V-space. Using confidence ellipce, discrimination value, (α), is calculated for considered data sets as α=0.87 which is satisfactory.



Alpha=87%

Fig. 1: Coordination of L-cys (red points) and D-syc (blue points) in the V-space. Ellipces are drawn based on confidence ellipce procedure [2].

In the following cases, feasible bands are calculated and different possible solutions are selected for pattern recognition purposes.

a : pattern is obtained using real concentration profile.

b, c, d, e, f, and g : different possible solution in feasible bands is considered for pattern recognition.

h: Points are selected in a manner which surrounds feasible region and corresponding profiles encompass all of possible profiles.

I, j, k, and l: using different initial estimates in MCR-ALS, resulted concentration profiles was used for pattern recognition.

Table 1. discrimination value, (α), for each designed cases.

Mode	A	b	c	d	e	F	g	h	I	j	k	l
Alpha	88%	67%	84%	85%	87%	87%	67%	91%	99.99%	69%	75%	0%

As it is tabulated in table 1, using different possible solutions resulted in different discrimination value. Althougth it is shown tha t incorporation of soft-modeling and pattern recognition is dramatically advantageous, rotational ambiguity dilemma should be taken into account.

References

1. M. Hasani and F. Emami, *Journal of Chemometrics*, 2013, **27**, 341-352.
2. N. A. Sinkov and J. J. Harynuk, *Talanta*, 2011, **83**, 1079-1087.

Simultaneous optimization of dispersive micro-solid-phase extraction of acidic and basic polloutants with the aid of desirability function

Ali Akbar Asgharinezhad, Homeira Ebrahimzadeh*

Faculty of Chemistry, Shahid Beheshti University, G.C., Evin, Tehran, Iran

E-mail address (H. Ebrahimzadeh): h-ebrahim@sbu.ac.ir

Abstract

In this study, a simple and fast coextraction of acidic and basic pollutants with the aid of magnetic dispersive micro-solid phase extraction based on mixed hemimicelles assembly was introduced. Cetyltrimethyl ammonium bromide (CTAB)-coated Fe_3O_4 nanoparticles as an efficient sorbent was successfully applied to adsorb 4-nitrophenol and 4-chlorophenol as acidic and chlorinated aromatic amines as basic model compounds, respectively. A central composite design methodology combined with desirability function approach was applied to find out the optimal experimental conditions. The opted conditions were: pH = 10, CTAB concentration = 0.86 mmol L^{-1}, sorbent amount = 55.5 mg, sorption time = 11.0 min, no salt addition to the sample, type and volume of the eluent = 120 µL methanol containing 5% acetic acid and 0.01 mol L^{-1} HCl, elution time = 1.0 min. Under the optimum conditions detection limits and linear dynamic ranges were achieved in the range of 0.05-0.1 µg L^{-1} and 0.25-500 µg L^{-1}, respectively. The extraction recovery and relative standard deviations (n = 5) were in the range of 71.4-98.0% and 4.5-6.5%, respectively. Ultimately, the applicability of the method was successfully confirmed by the extraction and determination of the target analytes in various water samples and satisfactory results were obtained.

Keywords: Desirability function; Central composite design; Dispersive micro-solid phase extraction; Mixed hemimicelles assembly, Simultaneous optimization; Acidic and basic pollutants.

Introduction

The simultaneous extraction of acidic and basic pollutants in a single step from various matrixes is a considerable and disputable concept in analytical chemistry. It has been reported that aniline, phenol and their derivatives are serious environmental pollutants and they are classified as the hazardous wastes and priority toxic pollutants by Environmental Protection Agency of America [1]. Anilines and phenols can easily permeate through soil and contaminate ground water due to their high solubility in water [1]. Herein, coextraction of these pollutants is in a point of view.

A fundamental motivation for developing a new separation and quantification method is reducing the required time and number of trials that ends in total required costs, and investigating the interactive effects among variables [2]. In the case of multiple response optimizations, a unique function named Global Desirability function (D or DF) or geometric mean (Geo mean) can be obtained when n variables (factor and responses levels) are combined after the transformation to desirability functions.

In this context, the aim is to develop a magnetic D-µ-SPE method based on mixed hemimicelles assembly for the coextraction and determination of some priority acidic and basic model pollutants. To the best of our knowledge, there is no report on the coextraction of acidic and basic pollutants using surfactant coated Fe_3O_4 NPs based D-µ-SPE

method. For simultaneous optimization, central composite design (CCD) methodology combined with desirability function approach was applied.

Results and discussion

For simultaneous optimization, the affecting factors were selected based on preliminary experiments and opted by a CCD experiment. In other words, CCD was applied to optimize the effect of five factors (pH of sample, CTAB concentration, sorption time, MNPs amount, and eluent volume). To get desired extraction efficiency as objective function, Geo mean as an indicator of extraction efficiency was maximized. It's important to note that an initial data preprocessing, i.e., normalizing the related response of each analytes is necessary before the analyzing of data. Subsequently, the obtained DF would be an input value for CCD. The studied ranges of the factors were: pH = 6.25-10.75, sorption time = 2.5-17.5 min, MNPs amount = 22.5-67.4 mg, CTAB concentration = 0.38-1.12 mol L^{-1} and eluent volume = 88-163 μL.

The experimental data presented a good agreement with the quadratic polynomial equation (Table 1S, Supplementary Information). The model F-value of 42.89 (p-value < 0.0001) implies that the model is significant and there is only a 0.01% likelihood that a model F-value this large could occur due to noise. The p-value for lack of fit (LOF) in the ANOVA table was higher than 0.05 that implies the LOF is not significant relative to the pure error. The goodness-of-fit of the model to the experimental values is depicted in Fig. 1S (Supplementary Information) which has an adjusted R^2-value of 0.9423. R^2 value greater than 0.80 implies that the obtained second-order polynomial model is sufficient for correlating the experimental results [3]. Moreover, this plot showed a good accordance between the predicted values and the actual experimental values. Two dimensional (2-D) contour plots and the response surfaces are depicted in Fig. 2S (Supplementary Information), exhibiting high desirability with warm ‘‘red’’ and low desirability with cold ‘‘green’’ colors.

Quality features of the current method were evaluated under the final opted conditions (Table 2S, Supplementary Information). To evaluate the accuracy and also applicability of the mentioned procedure for complicated samples, the coextraction of the aforementioned model compounds in real water samples (rain water, snow water and waste water) was performed and satisfactory results were obtained.

References

[1] Moradi M, Yamini Y, Kakehmam J, Esrafili A, Ghambarian M. A new strategy to simultaneous microextraction of acidic and basic compounds. *J. Chromatogr. A* 2011; **1218**: 3945-3951.

[2] Zolgharnein J, Shahmoradi A, Ghasemi JB. Comparative study of Box-Behnken, central composite, and Doehlert matrix for multivariate optimization of Pb(II) adsorption onto Robinia tree leaves. *J. Chemometr.* 2013; **27**: 12-20.

[3] Asgharinezhad AA, Ebrahimzadeh H. Coextraction of acidic, basic and amphiprotic pollutants using multiwalled carbon nanotubes/Fe_3O_4@polypyrrole composite. *J. Chromatogr. A* doi:10.1016/j.chroma.2015.07.087.

Is There Possible to Compare Biological Potency and Features of Colloidal Silver Nanoparticles Applied in Different Culturing Media?

Fateme Mirzajani [a,b,*]

[a] *Department of Biotechnology, The Faculty of New Technologies Engineering (NET), ShahidBeheshti University, G.C., Tehran, Iran.* [b] *Departmrnt of Nanobiotechnology, Protein Research Center (PRC), ShahidBeheshti University, G.C., Tehran, Iran.*

Abstract

In this study seven routine culturing media in biology and agriculture including; Gamborg's B-5, Murashige & Skoog Medium, Chu salt mixture, Nutrient, Mueller-hinton broth, Youshida and Carbon source/Vitamin mixture were subjected for their influences on NPs charachteristics. In addition to culturing media the influences of exlerimental conditions including; temperature, shaking and light/dark conditions at various times and silver nanoparticles were studied. Over 160 experiments were designed based on four varied parameters. The results and the model suggest that the increment in the temperature and duration of cultivation in the exposure of light and shaking lead to the severely decrease on the AgNPs stability and biological potency. In addition, the more complex culturing media; the more rapid lose of stability. Moreover, the additives including different carbon source or vitamin accelerate the AgNPs instability rate. The models are shown the stability, function and potency of AgNPs and provide the information for researchers using them.

Keywords: Silver Nanoparticle; Culturing Media; Central Composite Design; Stability Prediction.

1. Introduction

Silver nanoparticles (AgNPs) are one of the most wild spread nanomaterial and are utilized due to their antibacterial properties, [1] in variety of goods including agricultural pesticides and fungicides, [2], cosmetics and personal care products and food packaging, [3, 4]. Along with their applications, there are different reports on their toxic properties to human, animals, [4-7] and plants [8-10]. Nanoparticles have broad application in agriculture industry and plant sciences composing of a range of organic and inorganic compounds. While AgNPs have the ability to removal of metallic ions; they subjected for growth/aggregation/agglomeration which strongly influence on their biological potency or even deactivated them, [11-13]. The scope of this study is to treat different routine culturing media in the field of biology, agriculture and microbiology with a spherical AgNPs colloides under the varaied experimental conditions and evaluate the most influential conditions.

2. Results and discussion

The AgNPs colloid is produced and charachterized based on previous methods. They have maximum absorbance (lmax) at 426 nm, with homogenise 18.34 nm (X99) size. In order to study the influence of culturing media on the AgNPs stability and features seven types of routine seven routine biological culturing medid (listed in table 1) and glucose were prepared.

Table 1: Culturing media and their components, categorical factor in CCD.

Culture Media	Components
Muller Hinton Broth (MHB)	Beef infusion solids 4.0, Starch 1.5, Casein hydrolysate 17.5; g/L, Final pH 7.4 ±0.2 at 37°C.
Nutrient Broth (NB)	Peptone 15.0 Yeast extract 3.0 Sodium chloride 6.0 D(+)-Glucose 1.0; g/L, Final pH 7.5 ±0.2 at 25°C
Chu Medium (N6)	Ammonium Sulfate 463.0, Boric Acid 1.60, Calcium Chloride Anhydrous 125.330, Disodium EDTA Dihydrate 37.250, Ferrous Sulfate Heptahydrate 27.850, Magnesium Sulfate Anhydrous 90.370, Manganese Sulfate Monohydrate 3.330, Potassium Iodide 0.80, Potassium Nitrate 2830.0, Potassium Phosphate Monobasic 400.0, Zinc Sulfate Heptahydrate 1.50; mg/L, adjust to 1L, Final pH 5.7 ±0.1 at 25°C, autoclove at 121°C for at least 15 min.
Hydrophonic Medium (Youshida)	Ammonium nitrate 1142.50, Sodium dihydrogen phosphate.2H2O 503.70, Potassium sulphate 892.50, Calcium chloride.2H2O 1107.50, Magnesium sulphate.7H2O 4050.00, Manganese chloride.4H2O 18.75, Ammonium molybdate.4H2O 0.925, Zinc sulphate.7H2O 11.67, Boric acid 0.438 Copper sulphate.5H2O 0.300, Ferric Chloride.6H2O 96.25, Citric acid 148.75; mg/1L, Final pH 4.5 ±0.5 at 24°C.
Minimal Medium (M9)	Sodium hydrogen phosphate 25, Monopotassium phosphate 6, Sodium chloride 1, Ammonium chloride 1 g/L, autoclove at 121°C for at least 15 min, Glucose (20%) 20, Magnesium sulfate (1 M) 2, Calcium chloride (1 M) 0.1, Thiamine (0.5% w/v) 0.1; ml.
Gamborg Medium (B5)	Potassium nitrate 2500.00, Ammonium sulphate 134.00, Calcium chloride.2H2O 150.00, Magnesium sulphate 122.09, Sodium phosphate monobasic 130.42, Manganese sulphate.H2O 10.00, Boric acid 3.00, Potassium iodide 0.75, Molybdic acid (sodium salt).2H2O 0.25, Zinc sulphate.7H2O 2.00, Copper sulphate.5H2O 0.025, Cobalt chloride.6H2O 0.025, Ferrous sulphate.7H2O 27.80, EDTA disodium salt.2H2O 37.30, myo-Inositol 100.00, Thiamine hydrochloride 10.00, Pyridoxine hydrochloride 1.00, Nicotinic acid 1.00, Sucrose 20000.00; mg/L, Final pH 4.0 ±0.5 at 24°C, autoclove at 121°C for at least 15 min.
Murashige/Skoog Medium (MS)	Ammonium nitrate 1,650, Calcium chloride 440, Magnesium sulphate 370, Potassium phosphate 170, Potassium nitrate 1,900, Boric acid 6.2 , Cobalt chloride 0.025, Cupric sulphate 0.025, Ferrous sulphate 27.8, Manganese sulphate 22.3, Potassium iodide 0.83, Sodium molybdate 0.25, Zinc sulphate 8.6, Na2EDTA.2H2O 37.2, i-Inositol 100, Niacin 0.5, Pyridoxine.HCl 0.5, Thiamine · HCl 0.1, Glycine 2; mg/L, Final pH 5.7 ±0.1 at 25°C.

In this study a total of 160 experiments were performed to determine the effect of the factors on the AgNPs stability and features. The F-value in this table is the ratio of mean-squared error (MS) for each treatment to the residuals, which were obtained by repeating the experiment at the design centre point. The implication of the F-value depends on the degree of freedom of the model. The effects with the F-probability lower than 0.0001 are considered to be statistically significant. The values of R^2 and adjusted R^2 were 88.22 and 84.13% in the case of size and 84.95 and 74.33% in the case of the Shift of UV-Vis λ_{max}, respectively. The results demonstrate that with the increments of AgNps concentration the size increased and the UV-Vis λ_{max} showed a red shieft. In addition the increments of treatment time have the same influences. It seems that incubation time was capable of increasing size, causing aggregation and agglomeration of AgNPs, Figure 1 and Table 2. There are significant differences among MHB, a microbiological medium) and other medias. As can see in table 1 MHB composed of beef infusion solids, starch and casein hydrolysate with the final pH of 7.4. these media contains less amounts of salts and high amounts of proteins. It was previously mentioned that, nanoparticles/metal ions interactions have proposed that metal cations can be used as selective agents for AgNPs aggregation and assembly [9]. Furthermore, metal nanoparticles were reportedly aggregated in the presence of divalent metal ions by an ion-templated chelation process [12]. The most worst culturing media is hydrophonic media composing of high concentrations of salt in addition to glucose in the low pH. The results and the model suggest that the increment in the temperature and duration of cultivation in the exposure of light and shaking lead to the severely decrease on the AgNPs stability and biological potency. In addition, the more complex culturing media; the more rapid lose of stability.

Table 2. Some results of analysis of variances (ANOVA)

	Size (nm)	UV-Vis absorbtion (nm)
	Prob > F	Prob > F
A-AgNPS (ppm)	0.2896	0.0013
B-Time (h)	0.8477	0.9013
C-Temp (˚C)	0.0834	0.0016
D-Media (Type)	<0.0001	0.0014
AB	0.0927	0.0003
AC	-	0.0082
AD	0.0089	-
CD	0.0111	0.9945

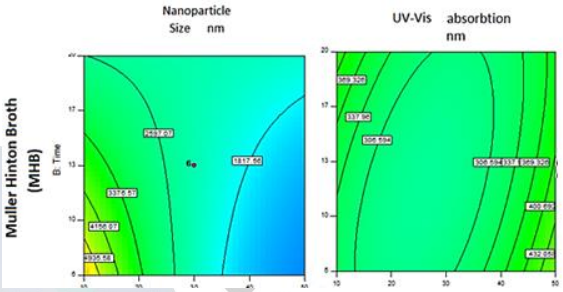


Figure 1: Response surface results of different types of culturing media based on time and AgNPs concentratios.

References

[1] Jegatheeswaran S, Sundrarajan M. PEGylation of novel hydroxyapatite/PEG/Ag nanocomposite particles to improve its antibacterial efficacy. Materials Science and Engineering: C. 2015;51:174-81.
[2] Taghiyari HR, Moradi-Malek B, Kookandeh MG, Bibalan OF. Effects of silver and copper nanoparticles in particleboard to control Trametes versicolor fungus. International Biodeterioration & Biodegradation. 2014;94:69-72.
[3] McShan D, Ray PC, Yu H. Molecular toxicity mechanism of nanosilver. Journal of food and drug analysis. 2014;22:116-27.
[4] Gaillet S, Rouanet J-M. Silver nanoparticles: Their potential toxic effects after oral exposure and underlying mechanisms–A review. Food and Chemical Toxicology. 2015;77:58-63.
[5] Aueviriyavit S, Phummiratch D, Maniratanachote R. Mechanistic study on the biological effects of silver and gold nanoparticles in Caco-2 cells–induction of the Nrf2/HO-1 pathway by high concentrations of silver nanoparticles. Toxicology letters. 2014;224:73-83.
[6] Munger MA, Radwanski P, Hadlock GC, Stoddard G, Shaaban A, Falconer J, et al. In vivo human time-exposure study of orally dosed commercial silver nanoparticles. Nanomedicine: Nanotechnology, Biology and Medicine. 2014;10:1-9.
[7] Austin LA, Ahmad S, Kang B, Rommel KR, Mahmoud M, Peek ME, et al. Cytotoxic effects of cytoplasmic-targeted and nuclear-targeted gold and silver nanoparticles in HSC-3 cells–A mechanistic study. Toxicology in Vitro. 2015;29:694-705.
[8] Mirzajani F, Askari H, Hamzelou S, Schober Y, Römpp A, Ghassempour A, et al. Proteomics study of silver nanoparticles toxicity on Oryza sativa L. Ecotoxicology and environmental safety. 2014;108:335-9.
[9] Mirzajani F, Askari H, Hamzelou S, Farzaneh M, Ghassempour A. Effect of silver nanoparticles on Oryza sativa L. and its rhizosphere bacteria. Ecotoxicology and environmental safety. 2013;88:48-54.
[10] Mustafa G, Sakata K, Hossain Z, Komatsu S. Proteomic study on the effects of silver nanoparticles on soybean under flooding stress. Journal of proteomics. 2015;122:100-18.
[11] Kuppusamy P, Yusoff MM, Maniam GP, Govindan N. Biosynthesis of metallic nanoparticles using plant derivatives and their new avenues in pharmacological applications–An updated report. Saudi Pharmaceutical Journal. 2014.
[12] Wang Y, Westerhoff P, Hristovski KD. Fate and biological effects of silver, titanium dioxide, and C 60 (fullerene) nanomaterials during simulated wastewater treatment processes. Journal of hazardous materials. 2012;201:16-22.
[13] Balakumaran M, Ramachandran R, Kalaichelvan P. Exploitation of endophytic fungus, Guignardia mangiferae for extracellular synthesis of silver nanoparticles and their in vitro biological activities. Microbiological Research. 2015;178:9-17.

Application of random forest (RF) for the modeling of Anti-HIV activity of some dibenzyle pyridinon derivatives, as HIV-1 integrase inhibitor, using group contribution descriptors

Nasser Goudarzi, Mansour Arab Chamjangali, and Atiyeh Asali*

Faculty of Chemistry, Shahrood University, Shahrood, Iran

E-mail address: atiyehasali88@gmail.com

Abstract

A QSAR study based the random forest (RF) was developed for the predict of anti-HIV activity of the some dibenzyle pyridinon derivatives. We performed studies on some dibenzyle pyridinon dervatives, as HIV-1 integrase inhibitor with anti-HIV biological activity, using quantitative structure–activity relationship (QSAR) methods that imply analysis of correlations and representation of models. In this work, the structural descriptors were generated by group contribution method. According to the group contribution concept a total of 8 descriptors were defined to account the overall structural features of compounds. The optimum RF model parameters of m= 8 and nt=100 were obtained. The mean square error (MSE) and determination coefficient (R^2) for the prediction of the pEC50 of test set were 0.0211 and 0.9298, respectively.

Key words: QSAR; Random forest; AIDS; Anti-HIV activity; Group contributions; Dibenzyle pyridinon dervatives.

Introduction

Acquired immune deficiency syndrome (AIDS) is a collection of symptoms and infections resulting from the specific damage to the immune system caused by the human immunodeficiency virus (HIV) [1]. In the HIV life cycle, reverse transcriptase (RT), protease (PR) and integrase (IN), are required. So, all of them are considered to be promising targets for the development of anti-HIV drugs. Recently QSAR study is widely developed for the designing of anti-HIV drugs. A basic assumption in QSAR studies relies on the fact that the structure of a compound entirely determines its properties [2]. Random forests are a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. Random Forest offers features which make it very attractive for QSAR studies.

Results and discussion

A data set containing the anti-HIV activity (pEC50) of 53 dibenzyl pyridinon derivatives, as HIV-1 integrase inhibitor was selected [3]. The main structure of this compound is shown in the figure 1. The structures and functional groups of pharmaceutical compounds can have an impact on the activity of these compounds. In order to consider the characteristics of this part of the molecule and its effect on drug activity, descriptors for compounds of interest were defined as contribution group. So, the structural descriptors were generated is called group contribution method.

According to the group contribution concept a total of 8 descriptors were defined to account the overall structural features of compounds.

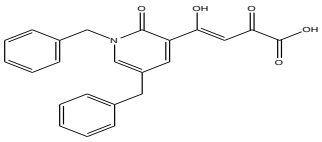


Figure 1. The main structure of the studied compound.

Finally, 8 descriptors were used for the RF model construction. The data set was consecutive divided into the training (43 molecules) and the test (10 molecules) sets. Two important parameters including the number of trees (n_{tree}) and the number of descriptors selected (mtry) were optimized to achieve the best RF model. The optimum RF model parameters including m= 8 and nt=100 were obtained. The validation study of the RF model was performed using test set and leave-group-out technique. pEC50 values of the molecules as test set were predicted for validation of RF model. Table 1 show the experimental and predicted values for the test set.

Table 1. *The results of the evaluation of the RF model using the test set*

	pEC50						
Compound No.	Exp.value	predicted value	% Relative error	Compound No	Exp.value	predicted value	% Relative error
2	6.0400	5.9717	-1.1302	33	6.1500	5.9601	-3.0875
10	6.5200	6.0573	-7.0967	39	5.7200	5.8115	1.5990
18	6.2200	6.0160	-3.2792	44	6.0900	5.9708	-1.9571
20	5.8500	5.8382	-0.2019	45	6.0000	5.9082	-1.5296
23	5.6000	5.8009	3.5880	52	6.3000	6.0209	-4.4303

The mean square error (MSE) and determination coefficient (R^2) for the prediction of the pEC50 of the test set were 0.0211 and 0.9298, respectively. These parameters in the leave-group-out method for all compounds were also calculated. The proposed RF model can used for the prediction of anti-HIV activity of the studied molecules. The statistical parameters such as R^2 and MSE show ability of the RF technique to predict the activity of drug compounds. Also, the RF method dose not suffer from the over fitting problem.

References

[1] R.J. Hu, J.P. Doucet, M. Delamar, R.S. Zhang, Eur. *J. Med. Chem*. 2009; **44**: 2158-2171.

[2] Todeschini R, Consonni V, Mannhold R, Kubinyi H, Timmerman H, Handbook of Molecular Descriptors, Wiley, 2008.

[3] I. Seo B, R. Uchil V, Okello M, Mishra S, Ma X, Nishonov M, Shu Q, Chi G, Nair V. Discovery of a Potent HIV Integrase Inhibitor That Leads to a Prodrug with Significant anti-HIV Activity. *ACS Med. Chem. Lett*. 2011; **2**: 877−881.

Chemometrics-assisted excitation-emission fluorescence analytical data for rapid and selective determination of optical brighteners in the presence of uncalibrated interferences

Atefeh Mohsenikia, Ali Gholami*, Saeed Masoum, Saleheh Abbasi

Department of Analytical Chemistry, Faculty of Chemistry, University of Kashan, Kashan, I.R. Iran
E-mail address: agholami@kashanu.ac.ir

Abstract

This study describes a novel approach for the simultaneous determination of CBS-X and CXT as widely used optical brighteners in household detergent, by combining the advantage of the high sensitivity of molecular fluorescence, and the selectivity of second-order chemometric methods. The proposed method is assisted by second-order chemometric analyses employing the parallel factor analysis (PARAFAC), self-weighted alternating trilinear decomposition (SWATLD) and Alternating penalty trilinear decomposition (APTLD) that help us to determine CBS-X and CXT in laundry powders and environmental samples. The proposed algorithms achieve the second-order advantage and in principle could be able to overcome the spectral uncalibrated interference problems in the determination of CBS-X and CXT at the ng g^{-1} level. High recoveries (90.00%-113.33%) for the spiked laundry powders and real environmental samples indicate the present method successfully faces this complex challenge without the necessity of applying separation and preconcentration steps in environmental contaminations.

Key words: Spectrofluorimetry, Second-order calibration methods, Optical brighteners, Real environmental samples.

Introduction

Optical brighteners (OBs) are synthetic chemical compounds that stick to the surface of the clothing or fabrics. OBs absorbing ultraviolet light (usually 340-370 nm), and re-emitting it as blue light (typically 420-470 nm), thus they enhance the intensity of visual reflection and strength the optical impression of whiteness and brightness [1]. Around 80% of fluorescent whitening agents (FWAs in the detergent industry) that are used detergent formulations and paper and textile industries are based on the stilbene derivatives. Benzenesulfonic acid, 2,2'([1,1'-biphenyl]-4,4'-diyldi-2,1-ethenediyl) bis-disodium salt or Tinopal CBS-X and disodium 4,4'-bis[(4-anilino-6-morpholion-1,3,5-trizin-2-yl)amino]stilbene-2,2'-disulphonate or CXT are most widely used in detergent industries [2]. Laundry wastewater is the largest supplier of optical brighteners to wastewater systems. They are removed from surface and ground water by adsorption into the soil and tend to accumulate on sewage sludge. High performance liquid chromatography (HPLC) and hyphenated method such as combining HPLC with mass spectroscopy has been applied to determine FWAs in detergents and environmental samples [3, 4]. The present work attempts to demonstrate the combination of EEM array in a photoluminescence system assisted by three-way calibration method for simultaneous determination of optical brighteners with strongly overlapping spectra in the presence of uncalibrated interferences.

Results and discussion

To generate fluorescence landscapes, excitation wavelengths between 250 and 380 nm with 5 nm intervals (27 excitation wavelengths) and emission wavelength between 300 and 550 nm with 0.5 nm intervals (502 emission wavelengths) were used. Thus, a three-way 27× 502 × 10 data array was collected. The scan rate was 1500 nm min^{-1}

excitation and emission slit width were set at 11 and 8 nm, respectively. Simultaneous quantitative analyses were provided by three-way methods in real samples. Table 1 indicates the resolved concentrations of unknown samples using APTLD, SWATLD and PARAFAC. For evaluation of accuracy of the proposed methods to analysis of the real samples, known concentration of two OBs were spiked in the real samples and the data array is obtained by stacking the standard set and spike samples, then recoveries for each spiked sample were calculated. Comparison of the predicted concentrations and recoveries provided by proposed algorithms shows a good predictive ability towards the spiked laundry powder and real samples by three algorithms.

Table 1. Quantitative analysis of CBS-X and CXT in the real samples

		Methods	CBS-X%	CXT%
Laundry powders	Sample 1	PARAFAC	5.24×10^{-2}	1.13×10^{-1}
		APTLD	5.37×10^{-2}	1.14×10^{-1}
		SWATLD	5.37×10^{-2}	1.14×10^{-1}
	Sample 2	PARAFAC	0.00	1.09×10^{-1}
		APTLD	0.00	1.02×10^{-1}
		SWATLD	0.00	1.02×10^{-1}
	Sample 3	PARAFAC	5.02×10^{-2}	0.00
		APTLD	5.10×10^{-2}	0.00
		SWATLD	5.10×10^{-2}	0.00
Agriculture soils	Sample 1	PARAFAC	1.31×10^{-4}	2.78×10^{-4}
		APTLD	1.21×10^{-4}	2.86×10^{-4}
		SWATLD	1.21×10^{-4}	2.86×10^{-4}
	Sample 2	PARAFAC	3.31×10^{-5}	7.88×10^{-5}
		APTLD	3.40×10^{-5}	5.73×10^{-5}
		SWATLD	3.40×10^{-5}	5.73×10^{-5}
Wastewater	Sample 1	PARAFAC	NA	NA
		APTLD	1.08×10^{-5}	3.07×10^{-5}
		SWATLD	1.08×10^{-5}	3.07×10^{-5}

References

[1] Field JA, Field TM, Poiger T, Siegrist H, Giger W, Behavior of fluorescent whitening agents during sewage treatment, *Water Res.* 1998; **32**: 1939-1947.

[2] http://www.wiley-vch.de/contents/ullmann/ull_10518.htm. (Ullmans Encyclopedia of Industrial Chemistry. 2001 Electronic release, 6th Edition. Wiley-VCH Interscience).

[3] Shu WC, Ding WH, Determination of fluorescent whitening agents in laundry detergents and surface waters by solid-phase extraction and ion-pair high-performance liquid chromatography, *J. Chromatogr. A* 2005; **1088**: 218-223.

[4] DuVal DL, Miyajima K, Ogura I, Characterization of brighteners in detergents by high-performance liquid chromatography/mass spectrometry/ultraviolet/fluorescence and three-dimensional high-performance liquid chromatography, *J. Am. Oil Chem. Soc.* 1995; **72**: 827-833.

Evaluation of physical properties and antioxidant activity of some Iranian herbhonies and classified by PCA and HCA chemometric method

Reza Tabaraki, Negar Sadeghinejad*
Department of Chemistry, Faculty of Science, Ilam University, Ilam, Iran.
E-mail address: negar_sadeghi.sh2007@yahoo.com

Abstract

until now there has been no research to determine physical and antioxidant properties of Iranian honeys.in the present study we investigated physical and antioxidant properties of fifty different types of herbhoney: Thymus Kotschyanus-Boiss ; EchinopsPersicum ; Petroselinum ; Nigella Sativa-sibth ;Eucalyptus globules ; Salvia officinal ; CoriandrumSativum ; QuerousBrantiikink ; Medicago sativa ; Astragalusgossypinusfinsch ; AstragalusadscendensBoiss ; Citrus SP ; Anethumgraveolensl; Cedrusatlantica ; TamarixManiferaehr.wecalssified these herbhoney by chemometric method such as PCA and HCA.

Key words: *antioxidant,herbhoney, chemometric method, PCA, HCA.*

Introduction

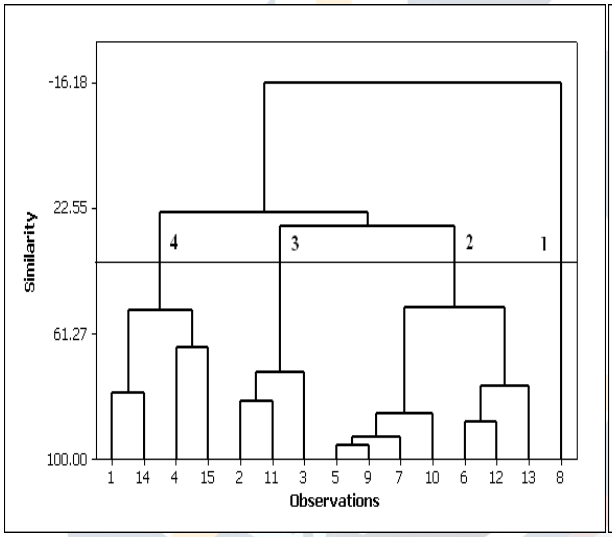
Honey is a natural complex product produced by honeybees from the nectar of blossoms or from exudates of trees and plants to produce nectar honeys or honeydews, respectively. Honey composition depends on the plants visited by the bees and on the climatic and environmental conditions. The strong sweetening capacity of honey is due to the presence of the monosaccharides fructose and glucose as majority components (60-85%) and phenolic compounds, minerals, proteins, free amino acids, enzymes, and vitamins as minor components. Honey is known to be rich in enzymatic and non-enzymatic antioxidants, including glucose-oxidase, catalase, flavonoids, ascorbic acid, phenolic acids and carotenoids. Research indicates that honey have functional properties in human health promotion which depend largely on the floral source of the honey. These properties could be associated to honey high osmolarity and antibacterial properties.

Results and discussion

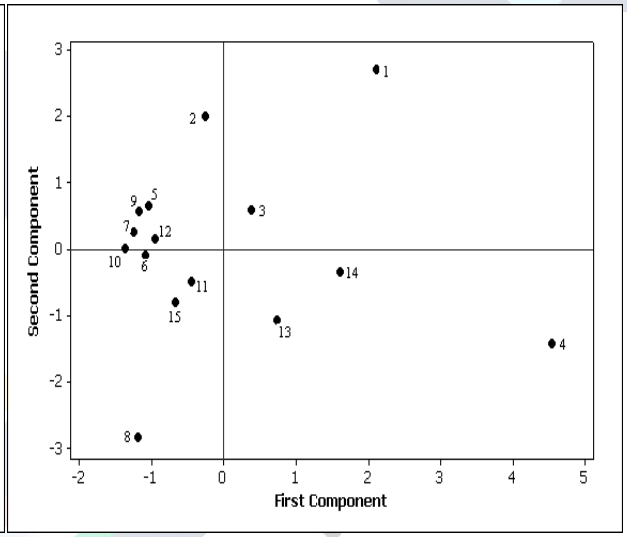
Physical properties of honeys

The pH was measured using a pH meter for a 10% (w/v) solution of honey prepared in milli Q water.*Ash content:* The ash content was determined by placing 2-3 g of honey samples in a crucible in a muffle furnace and heating at 640 °C for 6 h. Measurements of ash were done in triplicate and the mean was expressed in g% *Electrical conductivity:* The measurement of electrical conductivity is based on the determination of the electrical resistance. The electrical conductivity was measured for a 20% (w/v) solution of honey suspended in milli Q water (Bogdanovet al., 1997). The electrical conductivity of the milli Q water was < 10 µS/cm. Each sample was analysed in triplicate and the mean was

expressed in µS.*Color intensity: ABS*$_{450}$The net absorbance of the honey samples was determined by the method of Beretta, Granata, Ferrero, Orioli, and Facino (2005).The honey samples were diluted to 50% (w/v) with warm (45–50 °C) milli Q water and the solution was filtered through a 0.45 µm filter. There was a complete absence of coarse particlesin the honey solutions as all the commercial samples were noncrystalline liquid honeys. The absorbance was measured using a spectrophotometer at 450 and 720 nm and the difference in absorbance was expressed as mAU (Beretta et al., 2005).Total phenolic content, total flavonoid content, and amino acid of fifteen Iranian herbhonies were determined. Furthermore antioxidant avtivity for these honies were evaluated by five different methods: Free radical scavenging using 2, 2-diphenyl -1picryl hydrazyl (DPPH), Troloxequivalent antioxidant capacity (TEAC), Ferric reducing antioxidant power (FRAP), reducing power and total caretnoid content.By using of experiment resultsthesehoniesclassificated by chemometric methods such as PCA and HCA.Medicago sativa is in first group. Second group involve Anethumgraveolensl, CitrusSPP, Salvia officinal, Astragalusgossypinusfinsch, CoriandrumSativum, Medicago sativa and Eucalyptus globules. Petroselinum, AstragalusadscendensBoiss and EchinopsPersicum are in the third group. Thymus Kotschyanus-Boiss, Cedrusatlantica, Tamarix –Maniferaehr and Nigella Sativa-sibth are in the fourth group.



Classification of herbhonies by HCA method.



Classification of herbhonies by PCA method.

References

[1] Ortiz Valbuena, A.; Fernandez Maeso, M. C.; Subra Mun oz dela Torre, E. Principales caracter´sticas de la miel de La Alcarria. *InVestigacio´n Agraria en Castilla la Mancha*; Consejer ´a de Agricultura y Medio Ambiente de Castilla la Mancha: Castillala Mancha, Spain, 1996; 94 pp.
[2] Aljadi, A. M. ,&Kamaruddin, M. Y. (2004). Evaluation of the phenolic contents and antioxidant capacities of two Malaysian floral honeys. Food Chemistry, 85(4), 513 518.
[3] Effem, S.E.E., 1988. Clinical observations on the wound healing properties of honey. British Journal of Surgery 75, 679 681
[4] The National Honey Board (2003). Honey Health and therapeuticqualities. <www.nhb.org/download/factsht/compen-dium.pdf>.
[5] Aljadi, A. M. ,&Kamaruddin, M. Y. (2004). Evaluation of the phenolic contents and antioxidant capacities of two Malaysian floral honeys. Food Chemistry, 85(4), 513 518.
[6] Saxena,S., Gautam,S., Sharma,A.(2010) Physical, biochemical and antioxidant properties of some Indian honeys. Food Chemistry ,118 ,391 397

Simultaneous determination of paracetamol, diphenhydramine and dextromethorphane in pharmaceutical preparations using multivariate calibration 1

Abdolraouf Samadi-Maybodi[1,*], Seyed Mohsen Nikou[2], Seyed Karim Hassaninejad-Darzi[3]

[1]*Analytical Division, Faculty of Chemistry, University of Mazandaran, Babolsar, Iran, P.O.Box: 47416-95447. Tel/Fax: +98(112)5342350.*
[2]*Analytical Division, Faculty of Chemistry, University of Mazandaran, Babolsar, Iran, P.O.Box: 47416-95447.*
[3]*Department of Chemistry, Faculty of Science, Babol University of Technology, Shariati Av., Babol, Iran. P.O. Box: 484, Postal Code: 47148-71167, Tel/Fax: +98-11-32334203.*
E-mail address (corresponding author): samadi@umz.ac.ir

Abstract

Resolution of binary mixtures of paracetamol (PAR), diphenhydramine (DPH) and dextromethorphan (DEX) with minimum sample pre-treatment and without analyte separation has been successfully achieved by methods of partial least squares algorithm with one dependent variable, principal component regression and hybrid linear analysis. Data of analysis were obtained from UV– vis spectra of the above compounds. The method of central composite design was used in the ranges of 2–12 and 3-11mgL^{-1} for calibration and validation sets. The models refinement procedure and their validation were performed by cross-validation. Figures of merit such as selectivity, sensitivity, analytical sensitivity and limit of detection were determined for all three compounds. The procedure was successfully applied to simultaneous determination of the above compounds in pharmaceutical tablets.

Key words: Multivariate Calibration; Paracetamol; Diphenhydramine; Dextromethorphan.

Introduction

A mixture of paracetamol (PAR), diphenhydramine (DPH) and dextromethorphan (DEX) is widely used in diseases accompanied by cough, pain and fever such as the common cold and other viral infections as an analgesic, antipyretic, decongestant, antihistamine and antitussive [1]. Direct determination of an analyte is difficult due to the presence of one or several other constituents, instead of eliminating the interfering species (e.g. by a separation procedure). It is possible to use multivariate calibration for quantification of the analyte in the presence of the other compounds. A survey of literature showed that there are several spectrophotometric [2], HPLC [3,4] or LC-MS-MS [5] methods for the determination of these drugs alone or in combination dosage forms. In the present work, we developed and validated a simple, fast and sensitive method for the simultaneous quantification of PAR, DPH and DEX with UV-Vis spectroscopy with the help of multivariate calibration techniques in pharmaceutical tablets.

Results and discussion

Multivariate calibration methods allows extracting analytical information from the full-spectra, providing simultaneous determination of several components in the sample. Moreover, this techniquess permit a rapid analytical response with minimum sample preparation, reasonable accuracy and precision without separation procedures. For these reasons, these methods can be considered for routine analysis of the drugs in their formulations and human plasma. Correlation coefficient value for prediction (R^2_{PRED}) are 0.996, 0.995, 0.993 for PAR, DPH and DEX respectively.

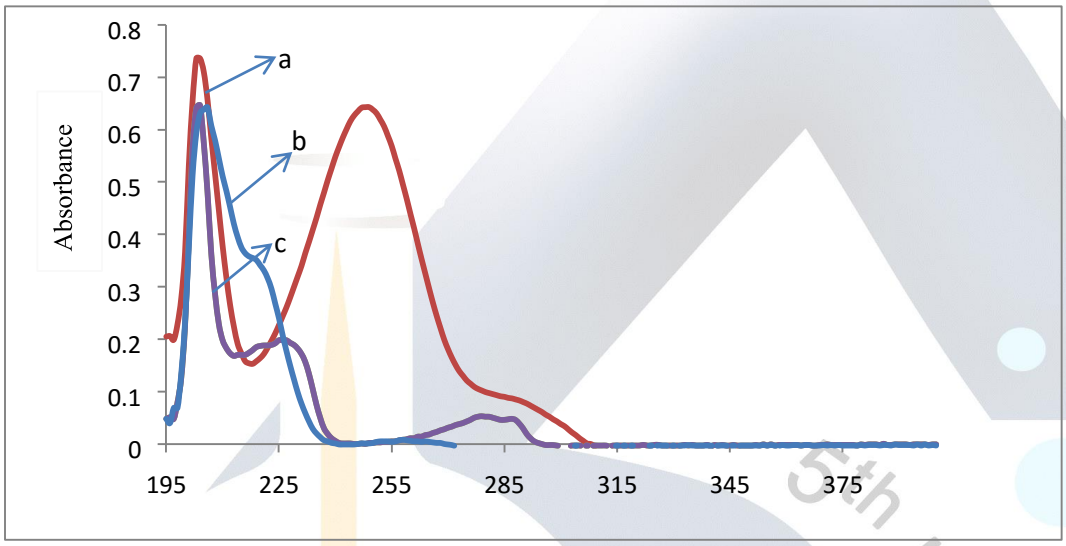


Fig. 1. Electronic absorbance spectra of three compounds in methanol: (a) paracetamol, (b) diphenhydramine and (c) dextromethorphan.

References

[1] P.P. Gerbino, Remington: *The Science and Practice of Pharmacy*, 20th ed., University of the Science, Philadelphia, 2000.

[2] Goicoechea HC, Olivieri AC. Simultaneous multivariate spectrophotometric analysis of paracetamol and minor components (diphenhydramine or phenylpropanolamine) in tablet preparations. *J. Pharm. Biomed. Anal*.1999; **20**:255–261.

[3] Carnevale L. Simultaneous determination of acetaminophen, guaifenesin, pseudoephedrine, pholcodine and paraben preservatives in cough mixture by high-performance liquid chromatography. *J. Pharm. Sci*. 1983;**72**:196–198.

[4] Ali MS, Ghori M, Rafiudin S, Khatri AR. A new hydrophilic interaction liquid chromatographic (HILIC) procedure for the simultaneous determination of pseudoephedrine hydrochloride (PSH), diphenhydramine hydrochloride (DPH) and dextromethorphan hydrobromide (DXH) in cough-cold formulations. *J. Pharm. Biomed. Anal*. 2007;**43**:158–167.

[5] Lou H, Yuan H, Ruan Z, Jiang B. Simultaneous determination of paracetamol, pseudoephedrine, dextrophan and chlorpheniramine in human plasma by liquid chromatography-tandem mass spectrometry. *J. Chromatogr. B.* 2010;**878**:682–688.

Spectrophotometric multicomponent analysis of a ternary and a quaternary mixture in human urine samples by analyzing first order data using MCR-ALS

Abdolhossein Naseri*, Bahar Ghasemzadeh and Saheleh Sheykhizadeh

Department of Analytical Chemistry, Faculty of Chemistry, University of Tabriz, Tabriz, Iran

E-mail address (corresponding author): a_naseri@tabrizu.ac.ir

Abstract

A novel method was developed for spectral resolution and further determination of three- and four-components mixture of drugs in urine samples. The method is simple, sensitive, and precise and could efficiently determine the components by a complementary application of technique Multivariate Curve Resolution using the Alternating Least Squares (MCR-ALS). In this work MCR-ALS with the proposed correlation and selectivity constraint along with other constraints has been applied to resolution and quantification of mixtures of drugs with matrix effects and unknown calibrated interferences in human urine using first-order data with a smaller number of calibration samples needs in first order multivariate calibration methods. The effect of different constraints was studied and the correlation of > 0.997 between the true spectra and the estimated spectral profiles were found for analytes. The results indicate good analytical performance towards the analytes, despite the intense spectral overlapping and the presence of unexpected constituents in the test samples. The maximum and minimum band boundaries of feasible solutions corresponding to the species profiles were estimated by multivariate curve resolution.

Key words: MCR-ALS, ternary mixture of drugs, quaternary mixture of drugs, first-order data, human urine.

Introduction

The problem of the appearance of unknown interference(s) is common in chemical analysis. In most cases, analysts have to deal with natural samples such as biological matrices, pharmaceuticals and environmental specimens, which are far from simplicity. Univariate calibration, which employs a single response per sample (known as zero order data) or multivariate calibration e.g. collecting a vector data for a sample (known as 1st order data), are not usually able to quantitate the analyte of interest in the presence of unknown and non-calibrated component(s)[1]. This means that the first-order calibration may compensate for interferences only if they are included in the calibration set. This explains why a large number of samples are needed in first-order calibration in comparison with second-order calibration, which can be performed using a few standards (in an extreme case, with only a single calibration sample). To cope with these issues, many sophisticated instrumentations which provide multidimensional (multi-way) data have been developed. A calibration model constructed using multi-way measurements makes the quantitation of the analyte of interest possible in new samples containing unknown component(s) which do not take part in the calibration data set[2]. This property

is known as the second-order advantage. Multivariate Curve Resolution (MCR) methods are reaching a mature state in Chemometrics and they have evolved as a powerful tool for the investigation of many types of chemical systems. Although MCR solutions have more physical meaning and an easier interpretation than those obtained by First-order calibration, they are not unique in the general Case, and they have an unknown amount of ambiguity. To overcome this problem, it has been shown that the correlation-constrained MCR-ALS version facilitates the analyte quantitation using first-order data[3]. In this work because of the presence of unknown interferences in calibration samples, the proposed multivariate calibration approach including the correlation and selectivity constraint is successful in determining the analyte of interest in the presence of unknown interference, considering and processing first-order multivariate.

Results and discussion

For training MCR-ALS model in resolution and quantification of mixtures of drugs with matrix effects and unknown calibrated interferences in human urine, the urine samples were collected two times from volunteer (in first time the urine samples without any dose of drugs are collected and the other time the urine samples containing of a specific brand drugs are collected). A calibration and validation were randomly prepared with spiking the different concentration of analyte to the first teme collected urine samples. The absorption spectra of data set show sever spectral overlap in the region from 200-400 nm. The spectral data matrix was subjected to the analysis by MCR-ALS algorithm using the spectral profiles of the components as initial estimates. For optimization, different constraints were applied to drive the final solution towards a chemical meaning. Non-negativity, selectivity, correspondence among species and correlation constraint are the constraints that are used. So the obtained results for MCR-ALS were free from the rotational ambiguity in ternary and quaternary mixtures of drugs despite the intense spectral overlapping and the presence of unknown constituents in the calibration and test samples. Results for the analysis of some examples of quaternary mixtures in the unknown sets are listed in **Table 1**.

Table 1. Results for the analysis of ibuprofen (IBU), aspirin (ASA), paracetamol (PAR), caffeine (CAF) in the unknown sets.

Sample	No	Taken (µg mL^{-1})				Found (µg mL^{-1})				Recovery (%)				RSD%			
		IBU	ASA	PAR	CAF	IBU	ASA	PAR	CAF	IBU	ASA	PAR	CAF	IBU	ASA	PAR	CAF
Urine samples containing of a specific brand of Acsar tablets	1	0.00	0.00	0.00	0.00	ND	0.47 ±0.01	4.11 ±0.03	1.17 ±0.02	-	-	-	-	-	3.00	1.03	2.42
	2	14.00	0.00	0.00	0.00	13.78 ±0.03	0.48 ±0.01	4.13 ±0.02	1.18 ±0.02	98.43	-	-	-	-	2.94	0.68	2.99
	3	0.00	7.40	6.50	0.00	ND	7.84 ±0.02	10.46 ±0.06	ND	-	99.59	97.69	-	-	0.36	0.81	-
	4	0.00	5.35	5.00	5.20	ND	5.92 ±0.05	9.01 ±0.04	6.39 ±0.03	-	101.87	98.00	100.38	-	1.19	0.62	0.66
	5	9.60	2.10	4.30	6.10	9.49 ±0.02	2.49 ±0.02	8.39 ±0.04	7.3 ±0.06	98.85	96.19	99.53	100.49	0.29	1.13	0.67	1.45

References

[1] Mohseni N, Bahram M, Olivieri A.C. Second-order advantage obtained from standard addition first-order instrumental data and multivariate curve resolution-alternating least squares. Calculation of the feasible bands of results. *Spectrochimica Acta Part A.* 2014; **122**: 721-730.

[2] Hurtado-Sánchez M. Comparison of the predictive ability of several second-order multivariate methods in the simultaneous determination of two therapeutic drugs in human urine. *Talanta*. 2012; **88**: 609-616.

[3] Goicoechea H.C, Olivieri A.C, Tauler R. Application of the correlation constrained multivariate curve resolution alternating least-squares method for analyte quantitation in the presence of unexpected interferences using first-order instrumental data. *Analyst*. 2010; **135**: 636-642.

Determination of acidity constants of cinnamic acid derivaties using two rank annihilation factor analysis (TRAFA)

A. Benvidi*, T. Zarei, S. Jahanbani

Department of Chemistry, Yazd University, Yazd, Iran

E-mail address (corresponding author): benvidi89@gmail.com

Abstract

In this work, we introduce a simple, selective, sensitive and low cost procedure for determination of acidity constants of cinnamic acid and cinnamic acid derivaties in binary mixtures of solvent-water at 25° C by applying RAFA and TRAFA methods to pH gradual change-UV-Vis spectral data.. In RAFA and TRAFA methods, the pKa values were obtained from relative standard deviation (RSD) versus hypothetical pKa values. Acidity constants of cinnamic acid derivaties obtained by RAFA and TRAFA methods are very close to each other. The effect of cinnamic acid derivaties struchter on the pKa was examined and a logical relationship was obtaind. By these methods and without any prior knowledge about the system, concentration profiles and pure spectra can be obtained from the experimental data. Obtained data by TRAFA model have more agreement with real acidity constants of cinnamic acid and cinnamic acid derivaties than RAFA model.

Key words: Rank annihilation factor analysis, TRAFA, TAR, Acidity constant, Spectrophotometry, Cinnamic acid derivaties.

Introduction

Cinnamic acid is an organic compound with the formula $C_6H^6CHCHCO_2H$. It is a white crystalline compound that is slightly soluble in water, and freely soluble in many organic solvents [1]. Classified as an unsaturated carboxylic acid, it occurs naturally in a number of plants. It exists as both a cis and a trans isomer, although the latter is more common [2]. Spectrophotometry is one of the most powerful techniques for the investigation of solution equilibria, although potentiometric / pH metric titrations are more convenient and more commonly used because of the simplicity of the equipment and minimal time requirements. The accurate determination of acidity constant is so important in various chemical and biochemical courses [3]. There are several problems in determination of acidity constants, including low solubility in aqueous solutions and the low values of acidity constants. Therefore, mixed solvents had been chosen to overcome these problems. Rank annihilation factor analysis (RAFA) is an efficient chemometrics technique based on rank analysis of two-way spectral data and this technique can be employed for quantitative analysis of some systems with unknown background. When RAFA combined by a chemical model it is comparable with hard-

soft modeling approach in solving some chemical problems. RAFA was originally developed by Ho et. al. as an iterative procedure [4]. To the best of our knowledge, there is no report for determination of acidity constants of cinnamic acid derivaties in gray mixture spectrophotometrically using RAFA [5] thus, in this work a simple, sensitive, selective and cheap procedure for the determination of acidity constants of cinnamic acid derivaties by applying RAFA to pH gradual change-UV-Vis spectral data (pH-spectra) is introduced. In addition RAFA and TRAFA methods are compared for determination of acidity constants of cinnamic acid derivaties.

Results and discussion

RAFA is a powerful chemometrics method for obtaining the acidity constant, especially when there is severe spectral overlap. Also, this method makes it possible to obtain pure absorption spectra and concentration profiles of species in several organic solvent-water binary mixture systems. As explained in the introduction section, RAFA is usually performed by annihilating one rank from the original data matrix. If the contributions of both acidic and basic forms of H_2A are simultaneously annihilated from the original data, the determination of the successive acid dissociation constants of diprotic acids can be feasible by RAFA, which this method was called two-rank annihilation factor analysis (TRAFA). Therefore, in this study TRAFA was used for the determination of the protolytic constants of cinnamic acid derivaties by applying pH gradual change-UV-Vis spectral data (pH-spectra). The acidity constants of cinnamic acid derivaties were calculated in the ethanol-water by applying different methods of RAFAand TRAFA.

The obtained pKa values from RAFA and TRAFA are summarized in Table 1.

Table 1. Comparison of the calculated acidity constants of cinnamic acid and cinnamic acid derivaties by RAFA and TRAFA in ethanol: water 10 %(w/w).

Material	Method	pK_{a1}	pK_{a2}
Cinnamic acid	RAFA	4.60	-
Ortho hydroxyl Cinnamic acid	RAFA	4.14	9.61
	TRAFA	4.12	9.58
Meta hydroxyl cinnamic acid	RAFA	4.45	8.55
	TRAFA	4.56	10

Para hydroxyl cinnamic acid	RAFA TRAFA	4.72 4.65	9.96 9.95

References

[1] Bradley JC, Neylon C, Williams A, Guha R, Open Notebook Science Challenge: Solubilities of Organic Compounds in Organic Solvents. Nature Precedings. 2010.

[2] Dorothea G, Cinnamic Acid in Ullmann's Encyclopedia of Industrial Chemistry, 2000, Wiley-VCH, Weinheim.

[3] Baggesgaard-Rasmussen H. Titration schwacher Säuren und Basen in Wasser-Alkohol-Gemischen. Zeitschrift für analytische Chemie 1936; 105: 269-282.

[4] Liang YZ, Kvalheim OM, Manne R. White, grey and black multicomponent systems: A classification of mixture problems and methods for their quantitative analysis. Chemom. Intell. Lab. Syst 1993; 18: 235-250.

[7] Abdollahi H, Safavi A, Zeinali S. Model-based rank annihilation factor analysis for quantitative analysis of mixtures of monoprotic acids using multivariate spectrophotometric acid-base titrations. Chemom. Intell. Lab. Syst 2008; 94: 112-117.

Fourier transform infrared spectroscopy coupled with multivariate classification methods to identify different adulterated infant formula powder with Melamine and Cyanuric Acid

Abdolhossein Naseri*, saheleh sheykhizadeh

Department of Analytical Chemistry, Faculty of Chemistry, University of Tabriz, Tabriz, Iran

E-mail address: a_naseri@tabrizu.ac.ir

Abstract

In the paper, Fourier transform infrared spectroscopy (FTIR) in association with multivariate chemometrics classification techniques was employed to predicte the quality of infant formula and to discriminate between the different adulterated infant formula samples including melamine and cyanuric acid without the need for using chromatographic separation or other expensive instruments. The ability of four different multivariate classification methods: principle component analysis- linear discriminant analysis (PCA-LDA), partial least squares-discriminant analysis (PLS-DA), K-nearest neighbor (KNN) and classification and regression tree (CART) were compared for infant formula classification. PCA-LDA technique for classification was found to be among the most effective ones with correctly prediction > 97% and KNN has shown worst result with correctly prediction 87.2%. Based on the results, it was concluded that FTIR coupled with multivariate classification could be used as a rapid and powerful technique for detecting melamine and cyanuric acid in infant formula powder, so Quality-Control could employ this method for rapid screening analysis to discourage adulteration practices.

Key words: Fourier transform infrared spectroscopy; classification; Melamine; Cyanuric acid; infant formula; adultration.

Introduction

Classification or supervised pattern recognition is a name given to a set of numerical techniques developed to solve the class membership problem. Many classification methods have been discussed and tried in the health sciences including food quality control. Milk products' safety has a close relationship with people's daily life and there are many methods to measure the milk safety such as for measure the adulteration contained melamine and related compounds were provided [1, 2]. Current analytical methods for testing melamine and related compounds are mainly chromatography based methods (HPLC and GC-MS). These methods exhibit high sensitivity, but for the most part are labor intensive, costly and time consuming [3]. Thus, there is still a need for a rapid, widely available, cost effective method for detection of melamine and its analogues in food ingredients. Vibrational spectroscopy techniques are good alternative for them in this field [4].

IR spectroscopy and chemometrics have been shown a great potential for monitoring the authenticity of dairy products, paving the way for their use by regulatory agencies in an effort to protect consumer health [4].

Results and discussion

Three powder blends were prepared that including a) infant formula: melamine, b) infant formula: cyanoric acid and c) infant formula: melamine: cyanoric acid. These blends had high and low w/w% ratio. The spectra of these samples were collected with FTIR (400-4000 cm^{-1}). The appropriate pre-treatment methods were done on the spectra and then these spectra were analyzed with PCA-LDA, KNN, PLA-DA and CART to identify whether infant formula has been adulterated or not and the type of adulterant used. The results of these calculations that obtained in optimized conditions are listed in Table 1.

Table 1. Result of multivariate classification method

Method	Optimum conditions	Correct of classification	Correct of prediction
PCA-LDA	Pretreatment: autoascale Principle components= 8	96.1%	97.5%
KNN	Pretreatment: autoascale and smoothing K nearest neighbors= 7	91.5%	87.2%
PLS-DA	Pretreatment: autoascale Latent variables= 12	93.6%	92.4%
CART	Pretreatment: autoascale and smoothing Leaf size= 10	90%	89.25%

As shown in Table 1, we can say that all these methods are satisfactory results. It can be concluded that, using the method described here, it is possible to identify whether infant formula has been adulterated or not and the type of adulterant used. It was concluded this method could be used as a rapid, nondestructive, capable and powerful technique for detecting melamine and cyanuric acid in infant formula powder.

References

[1]Tay M, Fang G, Chia PL, LiSFY. Rapid screening for detection and differentiation of detergent powderadulteration in infant milk formula by LC–MS.*Forensic Sci. Int.* 2013; **232** :32–39.

[2]Lin M. A review of traditional and novel detection techniques formelamine and its analogues in foods and animal feed. *Front. Chem. Eng. China* 2009; **3** :427–435

[3]Balabin RM, Smirnov SV. Melamine detection by mid- and near-infrared (MIR/NIR) spectroscopy: A quickand sensitive method for dairy products analysis including liquid milk, infantformula, and milk powder. *Talanta* 2011; **85** : 562–568.

[4]Souza SS, Cruz AG, Walter EH, Faria JA, Celeghini R, Ferreira M. Monitoring the authenticity of Brazilian UHT milk: A chemometric approach. *Food Chemistry* 2011; **124** :692-695.

Standard Addition Method Coupled with Principal Component Analysis for Quantitative Study of Real Samples Using Spectroscopic Data

Abdolhossein Naseri *, Farnaz Farahmand, Mahdieh Ghaffari

E-mail address: a_naseri@tabrizu.ac.ir

Abstract

Univariate calibration method is a simple, cheap and easy to use procedure in analytical chemistry. A univariate analysis will be successful if a selective signal can be found for the analyte(s). In this work, two simple methods are introduced to find such a selective signal, namely spectral ratio plot (SRP) and loading plot (LP) methods, both of which are able to perform a point-by-point search within whole set of variables. Electronic absorption data was used as a model in this work. Needed for the first method, SRP, were absorption spectra of unknown and standard samples. However, in the second method, LP, there is a need to augment first order data which was achieved via standard addition method using recorded spectra at multiple wavelengths. Once the selective wavelength range was found, univariate standard addition curve was plotted to get the analyte concentration without any contributions from the matrix effect. The standard addition curve was interpolated to reduce any bias error. To demonstrate the proposed methods, several synthetic and real datasets were analyzed and the results were reported. The proposed algorithms were used to determine some additives in food and hygienic real samples using spectrophotometric data.

Key words: principal component analysis, spectral ratio method, loading plot, spectral selective region, preservatives, standard addition.

Introduction

Analysis of unknown mixtures by simple spectrophotometric methods, among other analytical methods, has gain large deal of interest [1]. Many new instrumentation and instrumental techniques have been developed to deal with increasing complexity encountered with the samples. Quantitative analyses often involve the spectrophotometric resolution of mixtures of two components with partially overlapped spectra. The greater the extent of overlapping, the more difficult the resolution is rendered. Not surprisingly, this topic has been the subject of a number of chemometric studies originally intended to resolve binary mixtures while then extended to mixtures of three or more components [2]. In some cases, there may be a selective spectral zone for the analyte within the spectrum of multicomponent real samples. Components' concentration window can be found via such methods as Evolving Factor Analysis (EFA), Window Factor Analysis (WFA) and Heuristic Evolving Latent Projection (HELP); however, to the best of our knowledge, there is no obvious method in the literature to find selective zones within a spectral window. Getting information about these segments will be very useful in quantification and qualification analysis. The method of standard addition is a well-known analytical technique. This is used where sample matrix also contributes to the analytical signal (a situation known as the matrix effect), thus making it impossible to compare the analytical signal between sample and standard matrix using the traditional calibration curve approach [3]. In this work, we have used standard addition method along with interpolation to quantify food preservatives. As production of processed and convenience products increases, application of chemical food additives is becoming an increasingly important practice in modern food and herbal pharmaceutical industries.

Results and discussion

Loading Plot (LP) and Spectral Ratio Plot (SRP) has been used to quantify of food additives in real samples. Benzoic acid, sorbic acid, methylparaben and saccharin were selected as model analytes. Fig. 1A shows electronic absorption spectra of Benzoic acid, sorbic acid, methylparaben and saccharin. Different real samples were selected and checked to find selective zones within their absorption spectra using spectral ratio plot and loading plot methods. Fig. 1B shows the absorption spectra for the yoghourt drink sample in 200-400 nm before and after addition of different amounts of standard solution of benzoic acid. Fig. 1C and 1D show spectral ratio plot and loading plot of this dataset, respectively. The linear region of the loading plot with zero intercept as well as the linear region of spectral ratio plot at zero slope represent wavelength range (227 – 233 nm) of the selective region of benzoic acid in the yoghourt drink sample. In other words, benzoic acid is the only absorptive component in this region. After finding the selective region, standard addition curve was plotted at 230 nm and the concentration of benzoic acid was calculated using interpolation. In this work, several other real samples were analyzed by the proposed method: toothpaste, eye drop, shampoo, cola, moisturizing cream, juice and yoghourt drink. SRP and LP curves are plotted for each sample with the linear segment determined. Then, standard addition method within the selected wavelength range was used to perform quantification step. Spiked method was used to validate the obtained results.

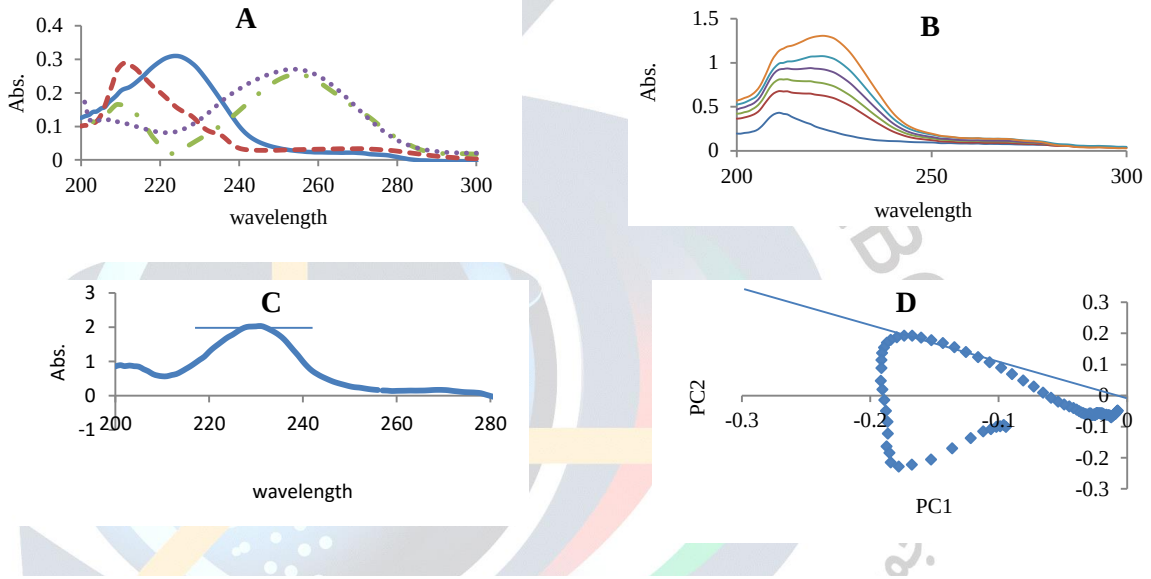


Fig1. Electronic absorption spectra of components (blue: benzoic acid, red: saccharin, green: methylparaben and purple: sorbic acid)(A), Absorption spectra data of yogurt drink after standard addition of benzoic acid (B), loading plot (C), and associated spectral ratio plot (D).

References

[1] A. Livaska, Analysis of a mixture of a known and an unknown weak acid by titration to a preset pH, Talanta, 22 (1975) 995-999.
[2] L. Hargis, J. Howell, R. Sutton, Ultraviolet and light absorption spectrometry, Analytical Chemistry, 68 (1996) 169-184.
[3] B.E. Saxberg, B.R. Kowalski, Generalized standard addition method, Analytical Chemistry, 51 (1979) 1031-1038.

Useing net analyte signal for determination of acidity constants caffiec acid in multivariate spectrophotometric analysis systems

A. Benvidi*, S. Jahanbani, T. Zarei

Department of Chemistry, Yazd University, Yazd, Iran

E-mail address (corresponding author): benvidi89@gmail.com

Abstract

A comparative study about advantages and limitations of net analyte signal (NAS) and DATa Analysis (DATAN) program in constant acid analysis has been performed. Net analyte signal (NAS) concept, which is a part of total signal that is directly related to the concentration of the component of interest. It monitors the concentration changes of any chemical species involved in the evolutionary process without requiring any pure spectra or having previous knowledge about the presence of the interferences. NAS method has some advantages such as the use of a full spectrum realization, therefore it does not require calibration and prediction steps and only a few measurements are required for the determination. By using these methods and without any prior knowledge about the system, concentration profiles and pure spectra can be obtained from the experimental data.

Key words: net analyte signal, DATa Analysis, Acidity constant, Spectrophotometry, caffiec acid.

Introduction

Caffeic acid (CA), 3,4-dihydroxycinamic acid is the major representative of hydroxycinamic acid in wines [1]. Caffeic acid is an organic compound that is classified as hydroxycinnamic acid. This yellow solid consists of both phenolic and acrylic functional groups.Caffiec acid is present in many fruits, vegetables, seasonings, beverages (coffee, wine) and olive oil. It is found in all plants because it is a key intermediate in the biosynthesis of lignin, one of the principal components of plant biomass and its residues [2]. The acidity and basicity form of dyes play a very fundamental role in many analytical procedures such as acid–base titration, solvent extraction and complex formation. Different methodologies have been proposed for the experimental determination of the acid dissociation constants including1H NMR spectroscopy, capillary electrophoresis, FT-IR spectrometry, UV–Vis absorption and fluorescence spectrophotometry and potentiometry. The net analyte signal (NAS) was defined by Lorberbased on spectroscopic methods, as the part of the spectrum of a mixture that is unique for the analyte of interest, i.e., it is orthogonal to the spectra of the interferences. The NAS is the part of the signal, which is directly related to the concentration predicted by the calibration model. In mathematical terms, it is the part of a spectrum which is orthogonal to the space spanned by the spectra of all analytes except one [3]. In this work, to obtain the NAS, the two methods Lorber et al and Goicoechea and Olivieri was used, followed by the two methods were compared. One of these

methods (Goicoechea and Olivieri) is hybrid linear analysis (HLA) which can be applied provided a very accurately measured pure spectrum of the analyte is available [4, 5]. The DATAN program, proposed by Kubista and coworkers calculates spectral profiles, concentrations and equilibrium constants by utilizing equilibrium expressions that are related to the components. However, to the best of our knowledge, there is not any report in the literature so far acidity constants the determination of Caffeic acid in grey mixture using the difference of absorption spectra of the analyte at different pH values.

Results and discussion

In this study NAS were used for determination of the protolytic constants of caffeic acid by applying pH gradual change-UV-Vis spectral data (pH-spectra).The first step in NAS is choosing the analyte and calculating the rank-annihilated data matrix R-m. Here we will discuss the steps of NAS analysis for a diprotic acid, for which the R_m matrix is obtained by either annihilating the contributions of the H_2A, HA or A^{2-} species. When the contribution of one species is annihilated from the total signal, R_{-m} contains spectral information for the remaining species and the spectral contribution from other sources such as interfering species). In this article, to obtain the NAS, the two methods Lorber et al and Goicoechea and Olivieri was used, followed by the two methods were compared. Comparison between methods of NAS and DATAN represents data close to each other. The acidity constants of caffeic acid were calculated in the AN-water by applying different methods of NAS and DATAN. The obtained pK_a values from NAS and DTAN are summarized in Table 1. This comparison indicates that the NAS method is an efficient method to obtain the constants of acid.

Table 1. The calculated acidity constants of caffeic acid in various Solvent percent (AN) by various methods.

Solvent percent	Method	pK_{a1}	pK_{a2}
AN: 20%(w/w)	NAS(Lorber)	5.95	9.70
	NAS(HLA)	5.92	9.65
	DATAN	6.05	9.81

AN: 60%(w/w)	NAS(Lorber) NAS(HLA) DATAN	6.00 5.98 6.09	9.90 9.95 10.05
AN: 80%(w/w)	NAS(Lorber) NAS(HLA) DATAN	6.15 6.20 6.25	10.25 10.30 10.42

References

[1] Franauet-Griell H, Checa A, Saurina J, Detrmination of polyphenols in spanish wines by capillary zone electrophoresis:application to wine characterization by using chemomtrics, J.Agric.Food chem., 2012;60:8340-8349.

[2] Boerjan W., Ralph J., Baucher M., Ligninbiosynthesis, Annual Review of Plant Biology, 2003; 54: 519-546.

[3] Lorber A., Error propagation and figures of merit for quantification by solving matrix equations, Anal. Chem., 1986; 58: 1167-1172.

[4] Bro R., Andersen C.M., Theory of net analyte signal vectors in inverse regression, J. Chemometrics, 2003; 17: 646–652.

[8] Mirmohseni A., Abdollahi H., Rostamizadeh K., Net analyte signal-based simultaneous determination of ethanol and water by quartz crystal nanobalance sensor, Anal. Chim. Acta, 2007; 585: 179-184.

Geographical origin identification of Iranain soil samples by ICP-MS and ICP-OES elemental analysis by chemometrics methods

Somayeh Veyseh [a], Ali Niazi *[,a], Jahan B. Ghasemi [b]

[a] *Department of Chemistry, Faculty of Science, Arak Branch, Islamic Azad University, Arak, Iran*
[b] *School of Chemistry, University College of Science, University of Tehran, P.O. Box 14155-6455, Tehran, Iran.*

E-mail address (corresponding author): ali.niazi@gmail.com & a-niazi@iau-arak.ac.ir

Abstract

The main outcome of this work is an elaboration of classification models for 156 soil samples representing the most widespread of Behabad soil samples. Analytical measurements are performed by ICP-MS and ICP-OES allowing the determination of individual elements. To achieved geographical classification of the soils by their elemental composition; cluster analysis (CA) and principal component analysis (PCA) were used for multivariate statistical modeling of the input data. CA displays the object similarity. The search for natural grouping in the samples was one of the main ways to study the structure of data. PCA was used as a descriptive tool to visualize the data in two dimensions, findings relationship between the variables and soil samples. Prior the chemometrics processing the data matrix was autoscaled.

Key words: Soil samples, ICP-MS, ICP-OES, Cluster analysis, Principal component analysis

Introduction

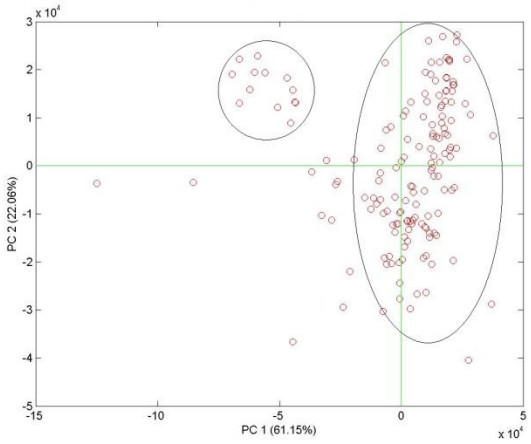
The application of chemometrics techniques to ICP-MS and ICP-OES data is investigated as a means differentiating soil samples on the basis of trace element content. There are several reports in the literature determining the content of inorganic elements in soil samples, measured by ICP-MS and ICP-OES techniques, with the goal of classifying samples according to their origin [1]. Various multivariate methods and chemometrics are used to interpret and extract information from complex data obtained by instrumental techniques in the analysis of soils. Results from other studies show that the correct discrimination and classification of soils depend not only on the smaple variability but also on the variables used for discrimination/classification.

Results and discussion

156 soil sampls from three different regions of Iran orginate, namely Behabad (128 samples), Bonestan (10 samples) and Asfij (12 samples). 70 trace elements (Ag, Al, As, Au, B, Ba, Be, Bi, Ca, Cd, Ce, Co, Cr, Cs, Cu, Dy, Er, Eu, F, Fe, Ga, Gd, Ge, Hf, Hg, Ho, I, In, Ir, K, La, Li, Lu, Mg, Mn, Mo, Na, Nb, Nd, Ni, Os, P, Pb, Pd, Pr, Pt, Rb, Re, Rh, Ru, Sb, Sc, Se, Sm, Sn, Sr, Ta, Tb, Te, Th, Ti, Tl, Tm, U, V, W, Y, Yb, Zn, Zr) were determined by ICP with mass spectrometry (MS) and optical emission spectrometry (OES) detections. The Analysis of variance is the initial step in identifying factors that are influencing a given data set. After the ANOVA test is performed, the analyst is able to

perform further analysis on the systematic factors that are statistically contributing to the data set's variability. The ANOVA test can provide information on the differences between soil samples.

Basic chemometrical characterization of the investigated soil samples is made by principal component analysis (PCA). PCA [2] allows the determination of the major sources of variability in the data set. It is an unsupervised technique, which can display a natural grouping of the studied objects (soil samples) in the plane or 3D space of the most important principal components, which are created by an uncorrelated linear combination of all original ariables (e.g. element concentrations of contained elements), optimized with respect to preserving as much as possible the total variance of the data. The search for natural groupings in the samples is one of the main ways to study the structure of the data. PCA is used as a descriptive tool to visualize the data in 2D or 3D space, finding the relationship between the variables and soil samples. Principal components can be, therefore, useful descriptors or even treated as predictors of some properties of data, so that the resulting grouping of variables can be used not only to distinguish between studied objects, according to the varieties or the region, but also to define the causes of variability.



In conlusion, this study achieved satisfactory discrimination among soil samples from Behabad region of Iran by multi-elemental anlysis using ICP-MS and ICP-OES. The data obtained and related to concentrations of elements in 156 different soil samples was used for their classification by means of the pattern recognition.

References

[1] Kruzlicova D, Fiket Z, Kniewald G. Classification of Croation wine varieties using multivariate analysis of data obtained by high resolution ICP-MS analysis. *Food Res. International* 2013; **54**: 621–626.

[2] Selih V.S, Sala M, Drgan V. Multi-elemental analysis of wines by ICP-MS and ICP-OES and their classification according to geographical origin in Slovenia. *Food. Chem.* 2014; **153**: 414-423.

Flotation-assisted homogeneous liquid-liquid microextraction and inductively coupled plasma-mass spectrometry for determination of thorium based on Box-Behnken Design

Somayeh Veyseh, Ali Niazi *

Department of Chemistry, Faculty of Science, Arak Branch, Islamic Azad University, Arak, Iran

E-mail address (corresponding author): ali.niazi@gmail.com & a-niazi@iau-arak.ac.ir

Abstract

In study, a fast, simple and efficient flotation-assisted homogenous liquid-liquid microextraction (FA-HLLME) combined with ICP-MS was successfully used to preconcentration and determination of thorium in different water samples. Response surface methodology based on a Box–Behnken Design (BBD) was used to optimize the extraction conditions (concentration of dimethyl vinyl phosphonate as chelating agent, sample pH, volume of extraction solvent and volume of homogenous solvent). The analysis of variance and some statistical tests such as lack-of-fit and coefficient of determination (R^2) showed good fit of the experimental data with the second-order polynomial model. Under the optimum conditions (concentration of dimethyl vinyl phosphonate = 2.4×10^{-4} mol L^{-1}, pH = 6.5, n-heptane = 150 μL and acetonitrile = 0.5 mL) the calibration graphs were linear in the range of 5.0–400.0 ng L^{-1} with detection limits of 2.62 ng L^{-1} ($3S_b/m$) for thorium. The preconcentration factor of this method for thorium determination reached 200. Repeatability (intra-day) and reproducibility (inter-day) of 0.83 and 1.03% for were obtained. The results indicated that simple, practical, and environmentally friendly proposed method could be successfully used to determination of thorium in different environmental water samples by using ICP-MS with the recoveries of 96.9-100.5%.

Key words: Box-Behnken Design, FA-HLLME, Determination, Thorium, Water sample

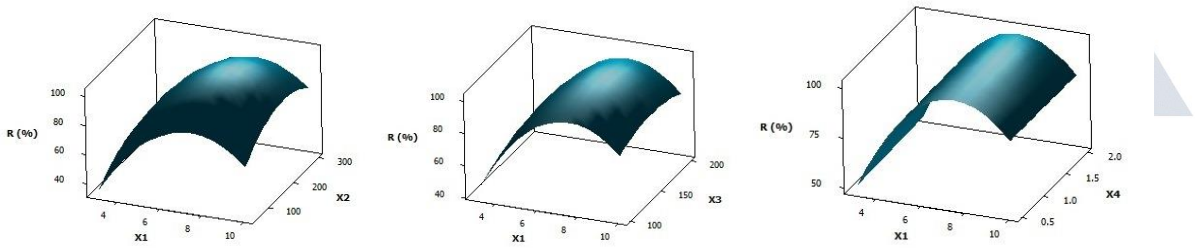
Introduction

The response surface methodology (RSM) has been extensively used for modeling and optimization in various analytical procedures [1]. Different factors can affect the extraction yield in the FA-HLLME procedure, therefore, a multivariate approach is recommended for their optimization. In the current work, Box–Behnken designs was used to optimize and study the influence of independent variables (concentration of chelating agent, sample pH, volume of extraction and homogeneous solvents) on the maximum extraction yield of thorium from water samples. Box–Behnken design is a second order multivariate technique based on three level partial factorial designs. Box–Behnken is a spherical, rotatable or nearly rotatable that consists of a central point and with the midpoints of the edges of the variable space. The main advantage of Box–Behnken design is that it requires a small number of runs and, therefore, is an important alternative avoiding time-consuming experiment [2].

Results and discussion

Box–Behnken experimental design was used to statistically study and optimize the significant variables. According to Box–Behnken matrix, a total of 27 tests containing 3 replicates at the center point were performed in random order. Experimental data were fitted to a second-order polynomial mathematical equation in order to express the relationship between input variables and responses. The analysis of variance (ANOVA) was applied to determine the significance of the model and the model terms so that the significance of each term was evaluated by their corresponding *p*-values. A statistically significant model only with significant terms can be written as follows eqution: $Y=1.2887+37.9668X_1+0.3182X_2+0.7293X_3+1.0007X_4-2.5205X_1\times X_1-0.0023X_3\times X_3-0.0933X_4\times X_4+0.0028X_1\times X_2-0.0040X_1\times X_3-0.1705X_1\times X_4-0.0084X_2\times X_4+0.0151X_3\times X_4$

To study and better understanding of the influences of the independent variables and their interactions on the response, three dimensional (3D) response surface plots were employed (sample figures are shown below, (X_1) sample pH; (X_2) volume of DVMPA; (X_3) volume of n-heptane; and (d) volume of acetonitrile).



In order to validate the proposed method, the linearity, precisions, LOD and enhancement factor were investigated under optimal conditions. The calibration curve for thorium determination was linear in the range of 5.0–400.0 ng L^{-1} (R^2 0.9997). LOD was calculated to be 2.62 ng L^{-1}. Precision, accuracy and stability were evaluated by intra- and inter-day analyses. The repeatability (intra-day) and reproducibility (inter-day) were determined by analyzing 5 spiked samples during a day and five replicates on 4 consecutive days, respectively. Intra- and inter-day precisions of the method were satisfactory with a relative standard deviation (RSD) of 0.83 and 1.03% for thorium determination. Finally, the preconcentration factor of 200 for thorium was determined.

References

[1] Govender S, Pillay V, Chetty D.J, Essack S.Y, Dangor C.M, Govender T. Optimisation and characterisation of bioadhesive controlled release tetracycline microspheres. *Int. J. Pharm.* 2005; **306**: 24–40.

[2] Zhang Z, Zheng H. Optimization for decolorization of azo dye acid green 20 by ultrasound and H_2O_2 using response surface methodology. *J. Hazard. Mater*. 2009; **172**: 1388–1393.

Visualization of error propagation and one-factorat-a-time local sensitivity analysis concepts

Akram Rostami[a] , Hamid Abdollahi[a]*, Marcel Maeder[b]

[a] *Department of Chemistry, Institute for Advanced Studies in Basic Sciences, P.O. Box 45195-1159, Zanjan, Iran,* [b] *Department of Chemistry, University of Newcastle, Callaghan, New South Wales 2308, Australia*

Email: abd@iasbs.ac.ir

Abstract

Until know everything about error propagation and sensitivity analysis was formulated. Sometime just using the formula cannot make a sense and it may be difficult for understand. In this study we will try to make a sense by visualization the error propagation and local sensitivity analysis, for ease to understand these concepts.

Key words: Error propagation, one-factorat-a-time (OAT) Lockal Sensitivity Analysis, Visualization

Introduction

Error propagation is simply the process of determining the uncertainty of an answer obtained from a calculation. Every time data are measured, there is an uncertainty associated with that measurement. If these measurements used in the calculation have some uncertainty associated with them, then the final answer will, of course, have some level of uncertainty. Determining the strength of the relation between a given uncertain input and the output is the job of sensitivity analysis [1,2]. A straightforward implementation of the "sensitivity" concept is provided by model output derivatives. If the model output of interest is Y, its sensitivity to an input factor X_i is simply $Y_{X_i} = \partial Y / \partial X_i$. This measure tells how sensitive the output is to a perturbation of the input. If a measure independent from the units used for Y and X_i is needed, $S^r_{X_i} = (X^0_i / Y^0)(\partial Y / \partial X_i)$. In this latter case an alternative measure is $S^{\sigma}_{X_i} = (\sigma_{X_i} / \sigma_Y)(\partial Y / \partial X_i)$, where the standard deviations σ_{X_i}, σ_Y are uncertainty analysis' input and output, respectively, in the sense that σ_{X_i} comes from the available knowledge on X_i, while δ_Y must be inferred using the model [3].

Results and discussion

In this study we used a reversible chemical reaction A↔B. with reaction rates k_1 and k_{-1} for the direct and inverse reactions, respectively, whose solution, for the initial conditions (ICs) :

(1) [[AB]]((tt==00))==0[A]$_0$

(2) $[A]= \frac{k_1[A+]k_{0\,-1}}{}\left(k_1 e^{-(k_1+k_{-1})t} + k_{-1}\right)$ $[B]= [A]_0 -[A]$

(3) $,kk_{1-1}\sim\sim NN((3,0.33,1))$

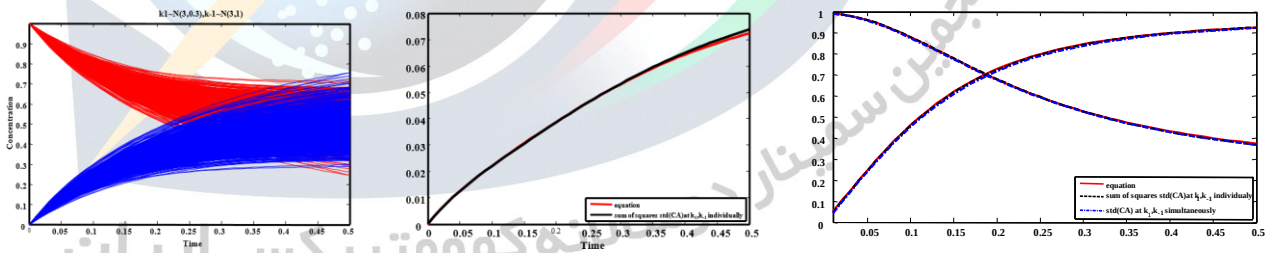
Here we assumed that the reaction rates are uncertain and describe by continuous random variables with known probability density functions (pdf's) that characterize their uncertainty as above (ref to eq 3). Based on this assumption existed a range for k_1 and k_{-1}.

In the first step k_1 or k_{-1} was fixed in the mean value "3" and another one was changed in the defined range. In the each set of k_1 and k_{-1} the concentration profile for A and B were calculated. When k_{-1} was fixed and k_1 was changed, the effect of uncertainty of k_1 on the concentration of A and B has been studied and vice versa. It was individual contribution of k_1 and k_{-1} in the error propagation equation.

In figure 1a we have demonstrated that uncertainty in both of k_1 and k_{-1} how changes the [A] and [B]. In the figure 1b error propagating base on this equation $\sigma^2_{[A]} \approx \sigma^2_{k_1}\left(\partial[A]/\partial k_1\right)^2 + \sigma^2_{k_{-1}}\left(\partial[A]/\partial k_{-1}\right)^2$ and numerical determination

has been shown. As seen in this figure, the results of numerical and theoretical calculation of error propagation approximately are matching.

a b c



Time

Figure 3: a) Time evolution of [A] and [B] when both of k_1 and k_{-1} simultanusly are uncertain parameters. B) Theoritical and numerical solution of error propagation in each time point. C) Theoritical and numerical solution of local sensitivity analysis.

Another concept that we have visualized was local sensitivity. This concept was calculated by $S^{\sigma}_{k1} = \frac{\sigma_{k_1}}{} \frac{\partial [A]}{}$, $S^{\sigma k\text{-}1} = \frac{}{} \frac{\sigma_{k\text{-}_1}}{} \partial [A]$. In addition to error propagation, local sensitivity could be numerically calculated $\sigma_{[A]} \partial k_1 \sigma_{[A]} \partial k^{-1}$ correctly (figure 1 c).

References

[1]Saltelli, A.; Chan, K; Scott, M. Sensitivity Analysis; John Wiley & Sons: Chichester, 2000.

[2]Saltelli, A.; Ratto, M.; Andres, T.; Campolongo, F.; Cariboni, J.; Gatelli, D.; Saisana, M.; Tarantola, S. Global Sensitivity Analysis. The Primer; John Wiley & Sons: Chichester, 2008.

[3] Saltelli, A.; Ratto, M.; Tarantola, S.; Campolongo, F. Chem. Rev. 2005, 105, 2811.

Pharmacophore-based virtual screening and toxicity risk analysis for identifying novel MPS1 inhibitors

Elmira Nazarshodeh[a] and Sajjad Gharaghani[a,*]

[a] *Department of Bioinformatics, Institute of Biophysics and Biochemistry, University of Tehran, Tehran, Iran*

E-mail address: s.gharaghani@ut.ac.ir

Abstract

Monopolar spindle 1 (MPS1), was identified as a novel target for cancer treatment. The significance of MPS1 kinase activity in spindle checkpoint activation has guided the search for MPS1 potent inhibitors. The availability of selective inhibitors with diverse structures has accelerated the suggestion of novel inhibitors by in silico methods. The identification of chemical features of MPS1 inhibitors will be useful to propose the potent candidates. We generated ligand-based pharmacophore model by LigandScout. The model was evaluated by ROC plot. Afterward, it was utilized for virtual screening of ZINC database. The retrieved molecules with high score values were filtered by Lipinski's rule of five and were then subjected to toxicity risk assessment.

Keywords: Monopolar spindle 1; Pharmacophore modeling; ROC; Virtual screening; Toxicity

1. Introduction

Monopolar spindle 1, a dual specificity protein kinase, controls the mitotic spindle checkpoint by monitoring correct chromosome attachment to spindle microtubules. Since inhibition of MPS1 kinase function weakens mitotic checkpoint signaling and accelerates apoptosis in cancer cells it is a promising target for cancer therapy [1]. Until now, no MPS1 inhibitor has entered clinical testing. So, the discovery of novel and non-toxic MPS1 inhibitors is an important issue. Pharmacophore approaches are efficient to identify novel molecule scaffolds. Hence, we developed the pharmacophore model to discover novel and promising MPS1 inhibitors. The model validated was used to screen chemical ZINC database. The retrieved molecules were then subjected to Lipinski's RO5 and toxicity risk assessment.

2. Results and discussion

2.1. Generation of pharmacophore using LigandScout

The pharmacophore model which was generated by LigandScout 3.12 is on the basis of the 3 structural classes of MPS1 selective and potent inhibitors from literatures [2-4]. The pharmacophore was validated by accessing the predictive ability on a decoy set to realize how correctly it can identify active and inactive compounds. The features of the model are two H-bond acceptor, one H-bond donor, one hydrophobic group, two aromatic ring and ten excluded volumes. Mapping the compound with the highest Pharmacophore-Fit score (59.08) on the model is shown in Figure 1. All features are consistent with the interactions between ligand and the binding site of MPS1 [2, 3].

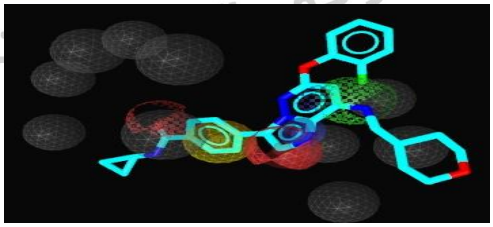


Figure 1. The best pharmacophore model

2.2. Pharmacophore performance and validation

A decoy set containing 10000 compounds from DUD-E and 349 active compounds from BindingDB were used for validating the reliability of pharmacophore model. These compounds were converted into LigandScout format, 258 out of 349 actives and 9962 out of 10000 decoys were obtained. A maximum of 25 conformations per ligand was calculated by the FAST settings of OMEGA. The statistical parameters were calculated to analyze the performance of the pharmacophore model. The results are the TP % = 83%, the FP % =22%, EF = 3.6. The pharmacophore retrieved 2344 hits and a total AUC value of 0.86% indicates the good specificity. They proposed the model has the ability for using in virtual screening.

2.3. Pharmacophore-based virtual screening

The ligand-based pharmacophore has used as a 3D query to search the chemical lead-like ZINC database using LigandScout. The 99 molecules were screened and satisfied all chemical features present in the model. The range of Pharmacophore-Fit score of hits and the hit rate % are 59.90-57.68 and 0.03%, respectively. These evidences confirm the accuracy of the model in discovery of potential hits. We removed the hits on the basis of the Lipinski's RO5 used as filter to assess ability of medicinal of compounds. The remained hits were used for toxicity risk analysis.

2.4. Toxicity risk assessment test

Hit molecules were subjected to the toxicity risk assessment test. The hits with no toxicity risk of mutagenicity, tumorigenicity, irritation and reproductive toxicity (Table 2) are proposed. The structures of these 5 top-scored and non-toxic hits associated with their Pharmacophore-Fit score values are shown in Figure 2.

Table 2: Results of toxicity risk assessment test

Hits	MUT	TUM	IRR	REP
ZINC33358003	No risk	No risk	No risk	No risk
ZINC69739078	No risk	No risk	No risk	No risk
ZINC91018779	No risk	No risk	No risk	No risk
ZINC09318833	No risk	No risk	No risk	No risk
ZINC40267234	No risk	No risk	No risk	No risk

MUT: mutagenicity, *TUM*: tumorigenicity, *IRR*: irritating effects, *REP*: reproductive effects (calculated by OSIRIS)

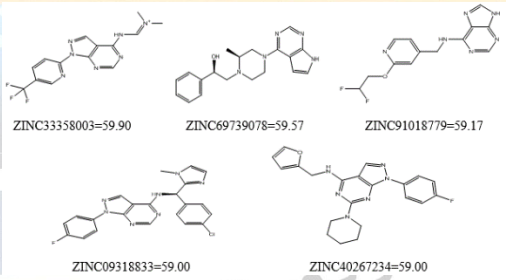


Figure 2. The non-toxic top-scored hits

References

[1] Maria K and Victor M. Bolanos-Garcia. Spindle Assembly Checkpoint (SAC): More New Targets for Anti-Cancer Drug Therapies. Advances in Cancer Drug Targets, 2014, 2, 54-79.

[2] Ken-ichi K, Nobuyuki I, Yataro D, Takeshi I, Takahiko Y, Hiroshi H, Kohei N, Hiroshi Y, Genta T, Sachie T, Kenichi H, Yousuke O, Yuji S, Makiko I, Motofumi I, Takayuki K, Yukichi I, Keiji D, Yasuto K, Shingo S, Shigeru A, Masahiro M, Masayo H, Yoshiyasu B, and Yusuke N. Discovery of Imidazo[1,2-b]pyridazine Derivatives: Selective and Orally Available Mps1 (TTK) Kinase Inhibitors Exhibiting Remarkable Antiproliferative Activity. J. Med. Chem. 2015, 58, 1760−1775.

[3] Ken-ichi K, Nobuyuki I, Yataro D, Takeshi, I, Takahiko Y, Hiroshi H, Yoshio H, Kenichi H, Yousuke O, Yuji S, Makiko I, Motofumi I, Takayuki K, Yukichi I, Keiji D, Yasuto K, Shingo S, Kazuya Y, Masahiro M, Masayo H, Kazuo U, Hidenori Y, Yoshiyasu B, Takeshi S, Hitoshi M, and Yusuke N. Indazole-Based Potent and Cell-Active Mps1 Kinase Inhibitors: Rational Design from Pan-Kinase Inhibitor Anthrapyrazolone (SP600125). J. Med. Chem. 2013, 56, 4343−4356.

[4] Ken-ichi K, Nobuyuki I, Yataro D, Takeshi I, Kenichi H,Yousuke O, Genta T, Yuki T, Yuji S, Makiko I, Tooru W, Motofumi I, Takayuki K, Yukichi I, Keiji D, Sachie T, Yasuto K, Shingo S, Kazuya Y, Masahiro M, Takahiko Y, Masayo H, Takeshi E, Kazuo U, Takeshi S, Hitoshi M, and Yusuke N. Diaminopyridine-Based Potent and Selective Mps1 Kinase Inhibitors Binding to an Unusual Flipped-Peptide Conformation. Med. Chem. Lett. 2012, 3, 560−564.

Evaluation of electrochemical process variables for removal of basic yellow 28 using a central composite design

Masoud Rohani Moghadam*, Saeideh Piroozi

Department of Chemistry, Faculty of Science, Vali-e-Asr University, 77188-97111 Rafsanjan, Iran.

E-mail address: m.rohani@vru.ac.ir

Abstract

In this study, electrochemical treatment was employed to remove color of basic yellow 28 dye using stainless steel electrodes in a batch-mode operation. The effects and relationships of the process variables were evaluated, and optimizations were carried out using response surface methodology to maximize removal efficiencies and minimize operating costs in relation to energy consumptionspotential= 5-20V, pH= 2-12, temperute= 25-40°C and [KCl]= 0.01-0.1mol L^{-1} The proposed models fit very well with experimental data, evidenced by high R2 correlation coefficients (>90%). At optimum conditions (applied potential= 11.99V; C_0= 20 mg/L; pH_i=7.81; temperate=20°C; [KCL]= 0.1 and t=4 min) maximum color removal and energy efficiency were 91.8% and 27.79mmol kWh^{-1}, respectively.

Key words: Experimental design, Desirability function, Basic Yellow **28,** Electrochemical decolorization

Introduction

Dyes and pigments represent one of the problematic groups;they are emittedintowastewaters fromvarious industrial branches, mainly from the dye manufacturing and textile finishing and also from food coloring, cosmetics, paper and carpet industries.Discharge of dye effluents into the natural streams may be toxic to the aquatic lives. Cationic dyes, commonly known as basic dyes, are a complex family and used in acrylic, nylon and wool dyeing.Advanced oxidation processes can be consideredfor the removal of recalcitrant pollutants fromwastewater. Owing to the efficiency of free radicals for thedegradation of recalcitrant compounds, most of the availablestudies deal with advanced oxidative processes [1]. Inthe case of an electroactive target compound, its oxidationor reduction can be implemented to achieve its degradation.An electrochemical reaction is based on a heterogeneouselectronic transfer, and hence, there is no need for the addition of supplementary chemical additives.Optimization is a critical stage to find the value that each factor must have to produce the best possible response. In this context, the multivariate design of experiments (DOE) is an important issue because it takes less time, effort and resources than the univariate procedures, and facilitates the gathering of large quantities of information while minimizing the number of experiments [2] When the optimization procedure involves more than one response, it is not possible to optimize each one in a separate way, because a number of solutions equal to the variables under study would be gathered. In 1980, Derringer and Suich found one of the solutions to optimize multiple responses by developing the Desirability function, which has been widely used since then in industry [3]. This function is based on the idea that the quality of a product or process that has many features is completely unacceptable if one of them is outside of a “desirable” limit. Its aim is to find operating conditions that ensure compliance with the criteria of all the involved responses and, at the same time, to provide the best value of compromise in the desirable joint response [4].

Results and discussion

The aim of this study was to use the experimental design (ED) for the modeling and the optimization of the yield of the electrochemical degradation of Basic Yellow 28 (BY28) dye in batch reactor by using central composite design (CCD). The analysis of variance of precentage of dechlorization (Y1) and energy efficiency (Y2) are shown inTables 1and 2.

Table1. Analysis of variance of precentage of dechlorization

Response: Degradation

ANOVA for Response Surface Reduced Quadratic Model

Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	22054.92	6	3675.82	32.21	< 0.0001	significant
A	*387.05*	*1*	*387.05*	*3.39*	*0.0785*	
B	*10731.51*	*1*	*10731.51*	*94.04*	*< 0.0001*	
C	*8263.65*	*1*	*8263.65*	*72.41*	*< 0.0001*	
D	*178.11*	*1*	*178.11*	*1.56*	*0.2241*	
D²	*1293.64*	*1*	*1293.64*	*11.34*	*0.0027*	
BC	*1200.97*	*1*	*1200.97*	*10.52*	*0.0036*	
Residual	2624.70	23	114.12			
Lack of Fit	*1383.36*	*18*	*76.85*	*0.31*	*0.9704*	*not significant*
Pure Error	*1241.34*	*5*	*248.27*			
Cor Total	24679.63	29				

Table 2. Analysis of variance of energy efficiency

Response: Efficiency

ANOVA for Response Surface Reduced Linear Model

Analysis of variance table [Partial sum of squares]

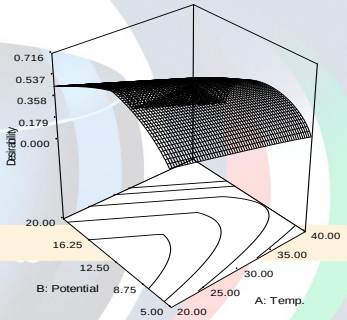
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	1432.95	2	716.47	11.58	0.0002	significant
A	*421.93*	*1*	*421.93*	*6.82*	*0.0145*	
B	*1011.01*	*1*	*1011.01*	*16.34*	*0.0004*	
Residual	1670.34	27	61.86			
Lack of Fit	*858.97*	*22*	*39.04*	*0.24*	*0.9917*	*not significant*
Pure Error	*811.37*	*5*	*162.27*			
Cor Total	3103.28	29				

With the five variables, data analysis gavemathematicalmodelsaccording to following equations :

Precentage of dechlorization=-22.17830-0.80317* Temp.-8.59259E-003* Potential-458.81481* Salt+16.10233* pH-1.07233* pH^2+102.68148* Potential * Salt

Energy efficiency =+69.50589-0.83858* Temp.-1.73078 * Potential

After studying the effects of the factors on each responses, the factors were simultaneously optimized for both of the responses by using the desirability function.The global desirability value was deternined as the geometric mean of the individual desirability functions by a feasibility grid search over the domain by the Design-Expert software (Stat-Ease Inc.).The Figure shows the response surface for the desirability function in various amounts of potential and temperute.



The optimized formulation was achieved at potential= 11.99V, pH= 7.81, temperute= 20°C and [KCl]=0.1mol $L^{-1}$with the corresponding desirability(D) value of 0.996. This factor level combination predicted the responses PD%= 91.8%, EE= 27.79mmol kWh^{-1}. Finally, to confirm the validity of the optimal parameters and predicted responses calculated, four batches of the optimized formulations were examined.

References

[1].IdrisYahiaoui,Farida Aissani-Benissad,Florence Fourcade,and AbdeltifAmrane. Combination of an electrochemical pretreatment with a biological oxidation for the mineralization of nonbiodegradable organic dyes: Basic yellow 28 dye. *Environ Progr Sustain Energ*. 2014;33: 160-169.

[2] Richard G. Brereton.Chemometrics: Data Analysis for the Laboratory and Chemical Plant, John Wiley & Sons, Chichester, 2003.

[3]. G. Derringer, R. Suich. Simultaneous Optimization of Several Response Variables. *J. Qual. Technol*. 1980; 12: 214–219.

[4].Raymond H. Myers, Douglas C. Montgomery, Christine M. Anderson-Cook. Response Surface Methodology: Process and Product Optimization Using Designed Experiments (Wiley Series in Probability and Statistics), Wiley, New York, 2009.

Experimental design and multiple response optimization for decolorization of acid blue 25 using electrochemical treatment

Masoud Rohani Moghadam*, Ali Reza Talebizadeh Rafsanjani, Leila Salehi

Department of Chemistry, Faculty of Science, Vali-e-Asr University, 77188-97111 Rafsanjan, Iran.

E-mail address: m.rohani@vru.ac.ir

Abstract

The dye acid blue 25 (AB 25) has found extensive use all over the world in the textile. In this context, electrochemical oxidation of acid blue 25 (AB 25) has been studied on a one-compartment batch reactoremploying using stainless steel electrodes. Central composite design(CCD) technique was used to investigate the effects of several experimental parameters, such as initial concentration, pH initial and supporting electrolyte on color removal and energy efficiency in multiple response optimization procedures using the desirability function. The optimal conditions were determined such as strongly acidic pH (pH =2), and KCl as a good supporting electrolyte.Under these conditions, a decolorization of 100% and 175.18mmol kVh^{-1} *energy efficiency of was obtained after only 1 min of electrolysis.*

Key words: Electrochemical decolorization, Experimental design, Desirability function, Acid blue 25

Introduction

Textile industry constitutes a major chunk of the industries in developing countries. The effluent from textile industries contains large variety of unused dye such as high organic materials, colors and surface active materials. As a consequence, dyes is becoming a major source of environmental contamination. As the international environmental standards are becoming more stringent, many research studies have been focussed on the treatments of colored wastewater. The chemical structures of dyes vary enormously, and some have complicated aromatic structures that resist degradation in conventional wastewater treatment processes because of their stability to sunlight, oxidizing agents, and microorganisms [1] The electrochemical oxidation technique has gained attention in recent years for textile industrial effluent treatment because of its efficiency in removing organic and inorganic pollutants from effluents. In addition, no additional chemical is required for electrochemical oxidation, because the electron is the main reagent in the electrochemical reaction.[2] In analytical chemistry, optimization is a critical stage to find the value that each factor must have to produce the best possible response. In this context, the multivariate design of experiments (DOE) is an important issue because it takes less time, effort and resources than the univariate procedures (which are surprisingly still being used in routine method development), and facilitates the gathering of large quantities of information while minimizing the number of experiments [3] When the optimization procedure involves more than one response, it is not possible to optimize each one in a separate way, because a number of solutions equal to the variables under study would be gathered. In 1980, Derringer and Suich found one of the solutions to optimize multiple responses by developing the Desirability function, which has been widely used since then in industry [17]. This function is based on the idea that the quality of a product or process that has many features is completely unacceptable if one of them is outside of a "desirable" limit. Its aim is to find operating conditions that ensure compliance with the criteria of all the involved responses and, at the same time, to provide the best value of compromise in the desirable joint response. This is achieved by converting the multiple responses into a single one, combining the individual responses into a composite function followed by its optimization [4].

Results and discussion

The analysis of variance of precentage of dechlorization (Y1) and energy efficiency (Y2) are shown inTables 1and 2.

Table1. Analysis of variance of precentage of dechlorization

Response: Degradiation

ANOVA for Response Surface Reduced Quadratic Model

Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	3928.99	3	1309.66	15.24	< 0.0001	significant
A	*2148.93*	*1*	*2148.93*	*25.01*	*< 0.0001*	
C	*1183.57*	*1*	*1183.57*	*13.78*	*0.0010*	
A²	*596.49*	*1*	*596.49*	*6.94*	*0.0140*	
Residual	2233.77	26	85.91			
Lack of Fit	*1927.75*	*21*	*91.80*	*1.50*	*0.3472*	*not significant*
Pure Error	*306.02*	*5*	*61.20*			
Cor Total	6162.76	29				

Table 2.Analysis of variance of energy efficiency

ANOVA for Response Surface Reduced Quadratic Model

Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	7.456E+005	5	1.491E+005	385.54	< 0.0001	significant
A	*4.819E+005*	*1*	*4.819E+005*	*1245.91*	*< 0.0001*	
B	*5248.38*	*1*	*5248.38*	*13.57*	*0.0012*	
C	*1.103E+005*	*1*	*1.103E+005*	*285.05*	*< 0.0001*	
A²	*41423.27*	*1*	*41423.27*	*107.10*	*< 0.0001*	
C²	*1.207E+005*	*1*	*1.207E+005*	*311.94*	*< 0.0001*	
Residual	9282.76	24	386.78			
Lack of Fit	*7318.12*	*19*	*385.16*	*0.98*	*0.5665*	*not significant*
Pure Error	*1964.65*	*5*	*392.93*			
Cor Total	7.549E+005	29				

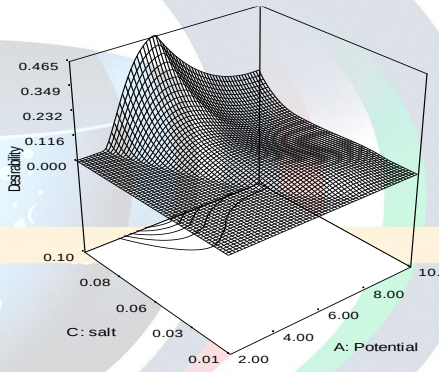
With the five variables, data analysis gavemathematical modelsaccording to following equations :

Precentage of dechlorization=+2.44708+18.38417* Potential+312.11111 * salt-1.13774* Potential^2

Energy efficiency = +1566.22845-185.33419* Potential-2.95758* Temp.-17163.80093* salt+9.54033* Potential^2

+1.28650E+005* salt^2

After studying the effects of the factors on each responses, the factors were simultaneously optimized for both of the responses by using the desirability function. Responses Y1and Y2 were transformed into individual desirability scales d1 and d2, respectively. All the responseswere set to be maximized, and 5 and 3 was selected as the importance values for decolorization and energy efficiency, respectively.Consequently, the global desirability value was determined as the geometric mean of the individual desirability functions by a feasibility grid search over the domain by the Design-Expert software (Stat-Ease Inc.) The Figure shows the response surface for the desirability function in various amounts of potential and salt concentration.



The optimized formulation was achieved at applied potential= 8V, pH=2, temp.= 20˚C and [KCl]= 0.1mol L^{-1} with the corresponding desirability(D) value of 0.996. This factor level combination predicted the responses PD100%, EE= 175.18 mmol kVh^{-1}.

References

[1]CHU W. Dye removal from textile dye wastewater using recycled alum sludge. Wat. Res. 2001; **35** :3147–3152.
[2]Martinez-Huitle CA, Brillas E. Decontamination of wastewaters containing synthetic organic dyes by electrochemical methods. A general review. *Appl. Catal. B-Environ.* 2009; **87** :105–145.
[3]Brereton RG. Chemometrics: Data Analysis for the Laboratory and Chemical Plant. 2003.
[4] Candioti LV, De Zan MM, Camara MS, Goicoechea HC. Experimental design and multiple response optimization. Using the desirability function in analytical methods development. Talanta. 2014; **124** :123-138.

Study of anti-cancer activity of parthenin derivatives using support vector regression and molecular docking

Zahra Garkani-Nejad

Department of Chemistry, Faculty of Science, Shahid Bahonar University of Kerman, Kerman, Iran

E-mail address: z_garkani@uk.ac.ir

Abstract

A quantitative structure–activity relationship (QSAR) study has been done on the anti-cancer activity of parthenin derivatives against three human cancer cell lines, SW-620, DU-145, and PC-3. QSAR models were based on multiple linear regression (MLR) and support vector regression (SVR). First, stepwise MLR was employed as a descriptor selection procedure. Then selected descriptors were used as inputs for SVR model. Comparison of the results indicates that the SVR method shows better predictive power than other method. In order to show the effect of hydration energy (HE) on anti-cancer activity, docking study of parthenin derivatives with Nf-jB transcription factor has been done.

Key words: Quantitative structure–activity relationship (QSAR); Anti-cancer activity; Docking; Parthenin derivatives

Introduction

In recent years, the anti-cancer property of various sesquiterpenes has attracted a great deal of interest and extensive research works have been carried out to characterize the anti-cancer activity and the molecular mechanisms of sesquiterpenoids [1]. Sesquiterpene lactones (SLs) are the active constituents of many medicinal plants from the Asteraceae family. Parthenin is a sesquiterpene lactone (SL) that isolated from *Parthenium hysterophorus* L [2]. has found interest due to its medicinal properties like anticancer activity [3]. Several novel derivatives of parthenin have been synthesized by the dipolar cycloaddition using various dipoles such as benzonitrile oxides, nitrones, and azides. Majority of the compounds exhibited improved anti-cancer activity compared to the parthenin, when screened for their in vitro cytotoxicity against three human cancer cell lines including SW-620, DU-145, and PC-3. Sesquiterpene lactones (SLs) are potent anti-inflammatory substances. The anti-inflammatory effect of these compounds could be partly explained by the inhibition of the transcription factor of NF-jB [4]. At the present work, relationship between the structure of parthenin derivatives and their anti-cancer activities against three human cancer cell lines, SW-620, DU-145, and PC-3 has been considered using some chemometrics methods.

Results and discussion

The half maximal inhibitory concentration (IC50) values of parthenin derivatives against three human cancer cell lines, SW-620, DU-145, and PC-3 are taken from the literature. Molecular descriptors were generated using DRAGON software. A stepwise MLR procedure was used for model development. For regression analysis, data set was

divided into two groups of training and test sets for each cell lines. G2v, H3u and Hydration energy (HE) are three common descriptors that entered in the best models for three cell lines.

SVR is the most common application form of SVM that is a powerful technique for predictive data analysis with many applications to varied areas of study. In this work, the SVR evaluations were carried out using the SVM toolbox in CLEMENTINE software. Selected descriptors using MLR models were employed as inputs. After that, the kernel function should be determined, which represents the sample distribution in the mapping space. In this work, the RBF (radial basis function) kernel was chosen.

Molecular docking was carried out by AutoDock 4.3 to understand the detailed binding model for the active site of the receptor with its ligands. For determining the appropriate binding conformations of studied compounds and check the main factors affecting the activity, docking study was performed for parthenin derivatives with the most anti-cancer activity. In order to show the effect of HE descriptor, the hydrogen bond between Transcription factor Nf-jB with derivatives of parthenin has been investigated. One of the most active derivatives of parthenin has been docked with Nf-jB factor. Nf-jB expression is completely inhibited by this compound. According to Figure 1, this molecule has three hydrogen bonds with Nf-jB transcription factor. This molecule has two hydrogen bonds with Lys37 that one bond is through oxygen atom of nitrile oxide ring and another bond is through of oxygen atom of cyclopentenone ring. The other hydrogen bond of this molecule is through hydroxyl group of cyclopentenone ring with Glu39. Docking of this compound with Lys37 and Glu39 leads to blocking and alkylation Cys38. Thus, docked compound inhibits Nf-jB factor completely.

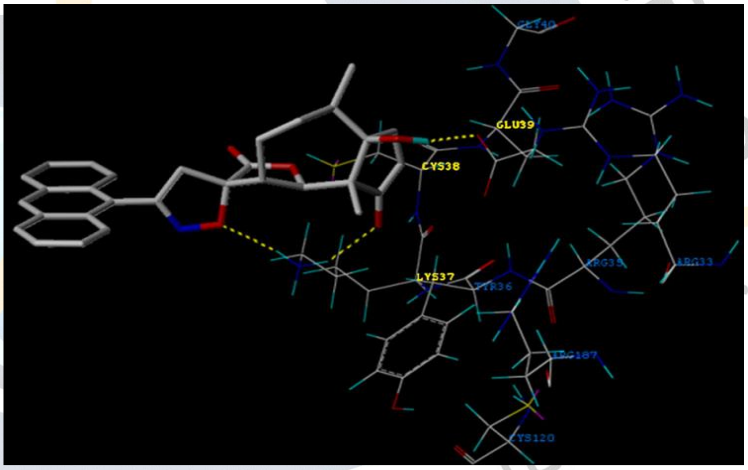


Figure 1

References

[1] Gershenzon J, Dudareva N (2007) The function of terpene natural products in the natural world. Nat Chem Biol 3:408–414

[2] Picman AK (1986) Biological activities of sesquiterpene lactones. Biochem Syst Ecol 14:255–281

[3] Fraga BM (2006) Natural sesquiterpenoids. Nat Prod Rep 23:943–972

[4] Garcı´a-Pin˜eres AJ, Lindenmeyer MT, Merfort I (2004) Role of cysteine residues of p65/NF-jB on the inhibition by the sesquiterpene lactone parthenolide and N-ethyl maleimide, and on its transactivating potential. Life Sci 75:841–856

Partial least squares multicalibration method for the simultaneous determination of montelukast, fexofenadine and cetirizine in the pharmaceutical tablets

Seyed Karim Hassaninejad-Darzi[1,*], Zarrin Es'haghi[2], Seyed Mohammad nikou[3]

[1]*Research Laboratory of Analytical & Organic Chemistry, Department of Chemistry, Faculty of Science, Babol University of Technology, Shariati Av., Babol, Iran. P.O. Box: 484, Postal Code: 47148-71167, Tel/Fax: +98-11-32334203.*
[2]*Department of Chemistry, Faculty of Science, Payame Noor University, P.O. Box 19395-3697, Tehran, Iran: eshaghi@pnu.ac.ir*
[3]*Department of Chemistry, Payame Noor University, P.O. Box 19395-3697, Tehran, Iran: smnikou@yahoo.com*
E-mail address (corresponding author): hassaninejad@nit.ac.ir

Abstract

The UV-Vis absorption spectra of sodium montelukast (MONT), fexofenadine hydrochloride (FEXO) and cetirizine hydrochloride (CET) are strongly overlapping and do not allow direct determination without previous separation by conventional methods. A simple, fast and precise spectrophotometric method combined with partial least-squares (PLS1) has been developed for the simultaneous determination of MONT, FEXO and CET. Data of analysis were obtained from UV-Vis spectra of three compounds. The method of central composite design was used for calibration and validation sets. A simple and fast method for wavelength selection in the calibration step is presented, based on the minimization of the predicted error sum of squares (PRESS) calculated as a function of a moving spectral window. The limit of detection was obtained 0.014, 0.062 and 0.112 mgL^{-1} for MONT, FEXO and CET, respectively. The procedure was successfully applied for simultaneous determination of the above compounds in pharmaceutical tablets.

Key words: Partial Least Squares 1; Montelukast; Fexofenadine; Cetirizine

Introduction

Asthma is a common chronic inflammatory disease of the airways characterized by variable and recurring symptoms, reversible airflow obstruction and bronchospasm. The control of asthma symptoms is a realistic goal and studies have shown that this can be achieved in most asthma patients leading to a higher quality of life [1]. Montelukast (MONT) is a potent and selective antagonist of the cysteinyl leukotriene receptor utilized for the treatment of asthma. Their discovery has had a significant impact on treatment strategies for the management of asthma [2]. Fexofenadine (FEXO) is a non-cardiotoxic and non-sedative terfenadine metabolite, which acts as a selective second-generation histamine H_1 receptor antagonist, relieving the uncomfortable manifestations of rhinitis [3]. Cetirizine hydrochloride (CET), a piperazine derivative and metabolite of hydroxyzine, described as a long-acting non-sedating antihistamine with some mast-cell stabilizing activity [4]. In the present work, we developed and validated a simple, fast and sensitive method for the simultaneous quantification of MONT, FEXO and CET with UV-Vis spectroscopy with the help of partial least squares multivariate calibration techniques in pharmaceutical tablets.

Results and discussion

The calibration set of 15 samples was built according to a central composite design (CCD) (three factors at two-level in the cubic vertex, six experiment in the cubic face and one central point) in the concentrations range of 2.0–12.0 mg L^{-1} for all drugs. The PLS1 model was developed using the calibration/prediction dataset that demand a suitable experimental design of the standards belonging to the calibration set in order to have good predictions. The full cross validation method suggested by Haaland and Thomas [5] was used for the selection of the optimum number of factors. Table 1 is shown data of test set composition, predicted values and relative error percentage for MONT, FEXO and CET .The minimum mean values of relative errors was calculated to be 3.93, 5.33 and 2.72 for MONT, FEXO and CET, respectively.

Table 1. Composition of test set and predicted values for MONT, FEXO and CET by PLS1 regression. Concentration values are expressed as mgL^{-1}.

	MONT			FEXO			CET		
Sample No.	Actual	Predicted	RE	Actual	Predicted	RE	Actual	Predicted	RE
Test 1	11.00	10.21	-7.18	11.00	10.68	-2.91	11.00	11.00	0.00
Test 2	3.00	2.80	-3.67	11.00	10.90	-0.91	11.00	10.38	-5.64
Test 3	11.00	10.03	-8.82	3.00	2.84	-5.33	11.00	10.50	-4.55
Test 4	11.00	10.20	-7.27	3.00	2.55	-15.00	3.00	2.91	-0.82
Test 5	3.00	3.00	0.00	11.00	10.02	-9.02	3.00	2.85	-1.36
Test 6	3.00	2.88	-4.00	3.00	2.83	5.67	3.00	3.13	4.33
Test 7	7.00	6.80	-2.86	7.00	6.65	6.43	7.00	6.80	-2.86
Test 8	7.00	6.72	-4.00	7.00	6.85	2.14	3.00	2.93	-2.33
Test 9	7.00	6.88	-1.71	11.00	11.36	3.27	7.00	7.20	2.86
Test 10	3.00	2.97	-0.43	7.00	6.83	2.43	7.00	6.83	-2.43
$\overline{R.E.}$			3.93			5.33			2.72

References

[1] Masoli M, Fabian D, Holt S, Beasley R. The global burden of asthma: executive summary of the gina dissemination committee report. Allergy 2004; **59**:469–78.

[2] Ibrahim MA, Amin EF, Ibrahim SA, Abdelzaher WY, Abdelrahman AM. Montelukast and irbesartan ameliorate metabolic and hepatic disorders in fructose-induced metabolic syndrome in rats. Eur J Pharmacol 2014; **724**:204–10.

[3] Simpson K, Jarvis B. Fexofenadine. Drugs 2000; **59**:301–21.

[4] Sweetman SC. Martindale-the complete drug reference. London: Pharmaceutical Press. 2002.

[5] Haaland D, Thomas E. Comparison of multivariate calibration methods for quantitative spectral analysis. Anal Chem 1990; **62**:1091–9.

Characterization and classification of Iranian honeys based on physicochemical properties and antioxidant activities, with chemometrics approach

Hamid Moloudian [1,2], Sepideh Abbasian [2], Nasser Nassiri-Koopaei [2], Mohammad Reza Tahmasbi [2], Golnaz alsadat Afzal [1], Maryam Sadat Ahosseini [2], Masud Yunesian [3], Mohammad Reza Khoshayand [1, 2,*]

[1] *Pharmaceutical Quality Assurance Research Center, Tehran University of Medical Sciences, Tehran 1417614411, Iran*

[2] *Department of Drug and Food Control, Faculty of Pharmacy, Tehran University of Medical Sciences, Tehran, Iran*

[3] *Department of Environmental Health Engineering, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran*

E-mail address (corresponding author): khoshayand@tums.ac.ir

Abstract

In the present study, 48 Iranian honey samples were tested for 15 physicochemical and antioxidant parameters.The parameters for which the samples were tested included color intensity, moisture, electrical conductivity, pH, free acidity, diastase activity, hydroxymethylfurfural content, proline level, total phenolic content, antioxidant activity, and radical scavenging activity. The study attempted todifferentiate honeys based on origin and composition. In the study, the Iranianhoney samples were classified according to their respective physiochemical properties and antioxidant activities using principal component analysis and hierarchical cluster analysis. Furthermore, the relationships between the geographical and botanical origins were determined for the samples used in the study.

Key words: Antioxidant activity, Chemometrics, Classification, Physicochemical properties

Introduction

Honey is a widely-used sweet food, with a long history in human consumption. Honey is naturally-occurring, complex, and multi-functional, and is a good source of antioxidants. It is used in the preparation of foods and beverages, as well as medical and pharmaceutical purposes [1]. At the present time, honey is required to meet specific quality criteria in order to be supplied in the market in most countries. The importance of honey sample authentication becomes even more apparent when considering countries or regions producing honeys from a diverse array of botanical and geographical origins. Multivariate techniques have shown some factors in honey to be more representative in quality control and more discriminating in origin than others [2]. In the present study, the physicochemical properties and antioxidant activities of different Iranian honey samples are investigated using various multivariate techniques in order to develop a quality management model. In this study, two different multivariate statistical techniques including cluster analysis (CA), factor analysis (FA)/principal component analysis (PCA) were applied to data pertaining to physiochemical and chemical composition of honey samples that was obtained in the laboratory in order to evaluate Iranian honeys from different botanical and geographical origins.

Materials and methods

For this study, 48 honey samples of various botanical origins (17 different species) and from a range of different geographical regions (48 regions from 19/31 provinces) throughout Iran were donated by the Animal Sciences Research Institute (ASRI) in the summer of 2014. Botanical origin of honeys was first ascertained by melissopalynological analysis carried out by ASRI. Fifteen parameters including physiochemical and antioxidant content indicators were measured/calculated. All measurements were triplicated. Data were organized into a matrix of 48 rows (honey samples) × 15 columns (composition). The results of the measurements were analyzed using various chemometric methods. The chemometrics evaluation was performed using two multivariate techniques: FA, including PCA, and CA, including hierarchical cluster analysis (HCA) using MATLAB 6.5 (TheMathworks, Natick, USA) and SPSS 11.5 for WINDOWS (SPSS Inc., Chicago, IL, USA, 2002).

Results and discussion

Approximately all ranges of parameters are subsets of the acceptable range (refer to national (according to recommendations of ASRI) and international standards [3]), with very few deviations. However, these results are derived from descriptive statistics, and applying statistical inference is necessary to achieve valid conclusions. PCA analysis showed that the first 5 PCs contain about 80% of the variance in data matrix. In other words, PCA can easily reduce dimension and complexity of a data matrix, while resulting in a loss of only about 20% of data. The results of the PCA show that correlating factors generally verify each other: the greater the color intensity by ABS_{450}, the greater the antioxidants properties by folin, FRAP, and proline, but the smaller IC_{50}. High DN and low HMF content are the indicators of honey freshness, and according to the loading plot, these two variables show a negative correlation.

In the present study, HCA is used to show the natural groupings within the dataset. This method is especially effective when the number of the cases examined is limited.

Results of chemometric classification methods carried out on physicochemical and antioxidant properties are highly matched to melissipalynological investigations performed by ASRI. In other word, honeys of similar botanical or geographical origins were clustered together or near each other in most cases. Chemometric combination of different factors including W%, proline, AC, EC, acidity, pH, HMF, diastase, and sugars leads to a strong differentiation of honeys from different geographical regions of Spain [3].

References

[1] Bogdanov S, Jurendic T, Sieber R, Gallmann P. Honey for nutrition and health: a review. *J. Am. Coll. Nutr.* 2008; **27** :677–689.
[2] Devillers J, Morlot M, Pham-Delegue MH, Dore JC. Classification of monofloral honeys based on their quality control data. *Food Chem.* 2004; **86** :305–312.
[3] Bogdanov S, Martin P. Honey authenticity: a review. *Mitt. Lebensm. Hyg.* 2002; **93** :232–254.

Sargassum Glaucescens for the removal of Methylen blue, Crystal Violet and Safranin in ternary system by artificial neural network modeling, Plackett–Burman design and response surface analysis

Reza Tabaraki*, Ashraf Nateghi

Department of Chemistry, Ilam University, Ilam, Iran

rezatabaraki@yahoo.com

Abstract

Biosorption properties of three cationic dyes with different molecular sizes (Methylene blue (MB), Crystal violet (CV) and Safranin (SF)) on Sargassum glaucescens were studied. Principal component-wavelet neural network (PC-WNN) was used for the simultaneous determination of MB, CV and Safranin concentrations in ternary solutions. Plackett–Burman design was applied to identify the most significant factors in the removal of dyes in ternary system. The interaction between the factors and their optimum levels for maximum removal of dyes were determined using Box–Behnken design. The optimum biosorption conditions were determined as initial dye concentration 10^{-4}mol/L, biosorbant dosage 0.1 g/L and biosorbant particle size 0.1875 µm. At this condition 0.819 mmol/g biosorption capacity was predicted by the RSM model.

Keywords: Dye removal, artificial neural network, Plackett–Burman design, Box–Behnken design, Ternary system

Introduction

Dyes are synthetic chemical compounds having complex aromatic structures which are extensively used in the textile, cosmetic, plastic, food and pharmaceutical industries [1]. The dye-containing wastewater discharged from the industries can adversely affect the aquatic environment by impeding light penetration. Moreover, most of the dyes are toxic, carcinogenic and harmful to human health [2]. Biosorption has been found to be one of the prominent techniques for dye wastewater treatment in terms of cost and operation [3]. In this study, adsorption performance of *S. glaucescens* in the removal of basic dyes, namely, MB, CV and Safranin in ternary dye solutions were studied. The interaction between the factors was studied and optimized using Box–Behnken design under response surface methodology.

Results and discussion

The simplest and the most common method is a top-down variable selection where the PCs are ranked in the order of decreasing eigenvalue. The PC with highest eigenvalue is considered as the most significant one and subsequently, the PCs are introduced into the WNN model one after the other. WNN models were developed using different number of the PCs in input layer. The WNN variables consisted of the number of PCs as an input layer, the number of nodes in the hidden layer, the learning rate, the momentum and the number of epochs, which optimized for each dye separately. The optimized variable parameters for each dye are given in Table 1.

Analysis of variance (ANOVA) is an essential tool for determining the significance of an effect or of a mathematical model. The most significant factors can be determined by using a statistical parameter, which is the P value (Table 2).

From these results, it was found that the effects of initial concentration of dye, biosorbant dosage and biosorbant particle size are the most important and significant factors for dyes uptake. The other factors are not important and can be considered negligible. The calculated regression equation for the optimization of medium constituents showed that dyes maximum uptake (Y) was related with the function of biosorbant dosage (X_1), initial dye concentration (X_2) and biosorbant particle size (X_3). For the three factors which were studied, the Box–Behnken model efficiently designed a second order response fit for the surface. Second-order polynomial model for q (mmol/g) was: $q\ (mmol/g) = 0.388 - 0.211\ X_1 + 0.263\ X_2 - 0.111\ X_3 - 0.074\ X_{12} + 0.129\ X_{13} - 0.045\ X_{23} + 0.086\ X_1^2 - 0.072\ X_2^2 + 0.025\ X_3^2$ $R^2 = 95.2\%$

The optimized conditions were initial dye concentration 10^{-4}mol/L, biosorbant dosage 0.1 g/L and biosorbant particle size 0.1875 µm. At this condition 0.819 mmol/g biosorption capacity was predicted by the RSM model. This result was validated experimentally (0.798 mmol/g), which is close to the predicted value.

Table 1. The optimized parameters of PC-WNN models.

Parameters	**Safranin**	**MB**	**CV**
Input neurons	3	3	3
Hidden neurons	2	2	2
Output neurons	1	1	1
Learning rate	0.058	0.067	0.051
Momentum	0.87	0.22	0.2
Number of iterations	1000	3000	2000
Hidden transfer function	Morlet	Morlet	Morlet
Output transfer function	Linear	Linear	Linear

Table 2. Effects and coefficients for maximum uptake (q) of dyes.

Factor	Effect	Coefficient	SE Coefficient	P
Constant		0.25522	0.01815	0.000
biosorbant dosage	-0.10958	-0.05479	0.02029	0.024
dye concentration	0.34251	0.17126	0.02029	0.000
pH	0.00632	0.00316	0.02029	0.880
time	0.06358	0.03179	0.02029	0.152
particle size	-0.08805	-0.04403	0.02029	0.058

References

[1] E. Forgacs, T. Cserhati, G. Oros, Environ. Int. 2004, 30, 953-971.

[2]V. Vimonses, B. Jin, C.W.K. Chow, J. Hazard. Mater. 2010, 177, 420-427.

[3] X.J. Xiong, X.J. Meng, T.L. Zheng, J. Hazard. Mater. 2010, 175, 241-246.

Simultaneous determination of paracetamol, diphenhydramine and dextromethorphane in pharmaceutical preparations using multivariate calibration 1

Abdolraouf Samadi-Maybodi[1,*], Seyed Mohsen Nikou[2], Seyed Karim Hassaninejad-Darzi[3]

[1]*Analytical Division, Faculty of Chemistry, University of Mazandaran, Babolsar, Iran, P.O.Box: 47416-95447. Tel/Fax: +98(112)5342350.*

[2]*Analytical Division, Faculty of Chemistry, University of Mazandaran, Babolsar, Iran, P.O.Box: 47416-95447.*

[3]*Department of Chemistry, Faculty of Science, Babol University of Technology, Shariati Av., Babol, Iran. P.O. Box: 484, Postal Code: 47148-71167, Tel/Fax: +98-11-32334203.*

E-mail address (corresponding author): samadi@umz.ac.ir

Abstract

Resolution of binary mixtures of paracetamol (PAR), diphenhydramine (DPH) and dextromethorphan (DEX) with minimum sample pre-treatment and without analyte separation has been successfully achieved by methods of partial least squares algorithm with one dependent variable, principal component regression and hybrid linear analysis. Data of analysis were obtained from UV–vis spectra of the above compounds. The method of central composite design was used in the ranges of 2–12 and 3-11mgL^{-1} for calibration and validation sets. The models refinement procedure and their validation were performed by cross-validation. Figures of merit such as selectivity, sensitivity, analytical sensitivity and limit of detection were determined for all three compounds. The procedure was successfully applied to simultaneous determination of the above compounds in pharmaceutical tablets.

Key words: Multivariate Calibration; Paracetamol; Diphenhydramine; Dextromethorphan.

Introduction

A mixture of paracetamol (PAR), diphenhydramine (DPH) and dextromethorphan (DEX) is widely used in diseases accompanied by cough, pain and fever such as the common cold and other viral infections as an analgesic, antipyretic, decongestant, antihistamine and antitussive [1]. Direct determination of an analyte is difficult due to the presence of one or several other constituents, instead of eliminating the interfering species (e.g. by a separation procedure). It is possible to use multivariate calibration for quantification of the analyte in the presence of the other compounds. A survey of literature showed that there are several spectrophotometric [2], HPLC [3,4] or LC-MS-MS [5] methods for the determination of these drugs alone or in combination dosage forms. In the present work, we developed and validated a simple, fast and sensitive method for the simultaneous quantification of PAR, DPH and DEX with UV-Vis spectroscopy with the help of multivariate calibration techniques in pharmaceutical tablets.

Results and discussion

Multivariate calibration methods allows extracting analytical information from the full-spectra, providing simultaneous determination of several components in the sample. Moreover, this techniquess permit a rapid analytical response with minimum sample preparation, reasonable accuracy and precision without separation procedures. For these reasons, these methods can be considered for routine analysis of the drugs in their formulations and human plasma. The best Correlation coefficient value for prediction (R^2_{PRED}) are 0.996, 0.995, 0.993 for PAR, DPH and DEX respectively.

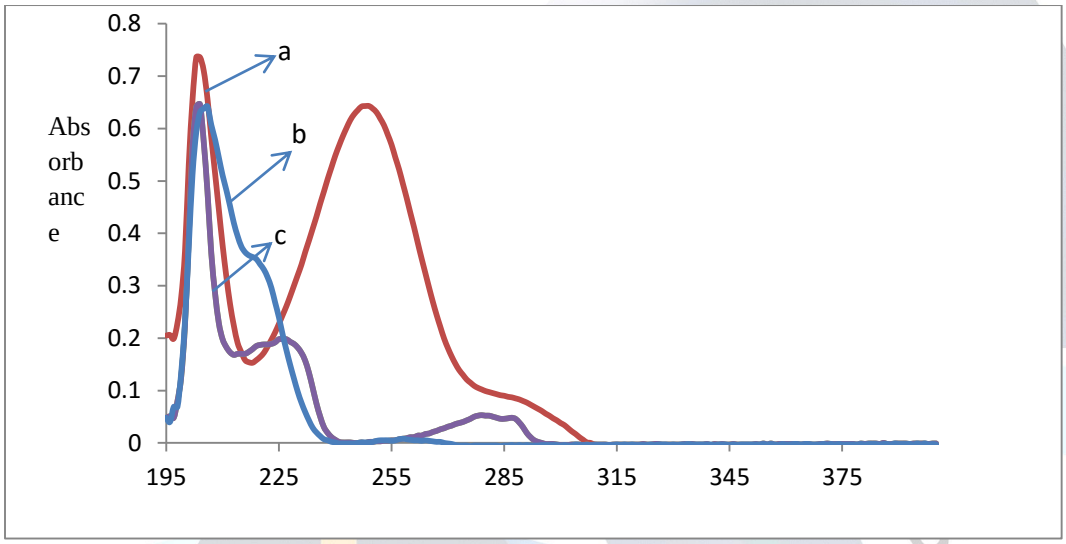


Fig. 1. Electronic absorbance spectra of three compounds in methanol: (a) paracetamol, (b) diphenhydramine and (c) dextromethorphan.

References

[1] P.P. Gerbino, Remington: *The Science and Practice of Pharmacy*, 20th ed., University of the Science, Philadelphia, 2000.

[2] Goicoechea HC, Olivieri AC. Simultaneous multivariate spectrophotometric analysis of paracetamol and minor components (diphenhydramine or phenylpropanolamine) in tablet preparations. *J. Pharm. Biomed. Anal.*1999; 20:255–261.

[3] Carnevale L. Simultaneous determination of acetaminophen, guaifenesin, pseudoephedrine, pholcodine and paraben preservatives in cough mixture by high-performance liquid chromatography. *J. Pharm. Sci.* 1983;72:196–198.

[4] Ali MS, Ghori M, Rafiudin S, Khatri AR. A new hydrophilic interaction liquid chromatographic (HILIC) procedure for the simultaneous determination of pseudoephedrine hydrochloride (PSH), diphenhydramine hydrochloride (DPH) and dextromethorphan hydrobromide (DXH) in cough-cold formulations. *J. Pharm. Biomed. Anal.* 2007;43:158–167.

[5] Lou H, Yuan H, Ruan Z, Jiang B. Simultaneous determination of paracetamol, pseudoephedrine, dextrophan and chlorpheniramine in human plasma by liquid chromatography-tandem mass spectrometry. *J. Chromatogr. B.* 2010;878:682–688.

QSAR modeling of binding affinity of some serotonin receptor antagonists using docking derived molecular descriptors

Mohammad H. Fatemi[a], Hengameh Bahrami[b]*, Sajjad Gharaghani[c]

[a]*Chemometrics Laboratory, Faculty of Chemistry, University of Mazandaran, Babolsar, Iran.*
[b]*Department of Chemistry, Shahid Bahonar University of Kerman, Kerman, Iran.*
[c]*Department of Bioinformatics, Laboratory of chemoinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran.*

mail address: atom_79@yahoo.com

Abstract

The binding affinity of some tetrahydrocarbazole derivatives as 5-HT6 receptor antagonists was predicted by protein structure-based QSAR and ligand- based QSAR modeling. In the first step, three- dimensional (3D) structure of the 5-HT6 was obtained by the homology modeling. Then the molecular docking was used to find out binding site and consequently types of the interactions between the most potent drug and receptor. Results showed that hydrogen bonds and hydrophobic interactions play an important role in binding affinity of studied chemicals. The obtained information relating to the binding site of the receptor utilized for docking of all drugs to the receptor and finding the optimized conformation for each drug candidate which were used to develop structure- based quantitative structure-activity relationship (QSAR) model. In the next step, more than 1900 descriptors were calculated for docked drugs. Then, selected descriptors were related to the binding affinity using MLR and LS-SVM regression. The best MLR model has the statistics of R_{train}=0.89, R_{test}=0.88, SE_{train}=0.792 and SE_{test}=8.49. Finally, results of protein based QSAR model were compared whit ligand based model and indicate the superiority of protein based over ligand based QSAR model.

Keywords: Homology modeling, Molecular docking, binding affinity, protein structure- based QSAR.

Introduction

Some brain disorders such as; alzheimer's disease, anxiety, depression, epilepsy, obesity, schizophrenia and bulimia are controlled in part by the neurotransmitter serotonin (5-HT) [1]. Serotonin, a small molecule is derived from tryptophan [2]. The most dramatic effects of serotonin are in the brain and help to control our emotions, moods and thoughts. Serotonin is released from vesicles in nerve. The activity of 5-HT is mediated through activation of members of a large family of 5-HT receptor proteins that can be grouped into seven subfamilies(5-HT1-7) on the basis of sequence homology and signaling mechanisms [3]. Many different drugs bind to serotonin receptors and modify the normal signaling process. Among these receptors, 5-HT6 acts selectively in the central nervous system [4].

The ability and the strength of a chemical to join to its receptor (for example to 5-HT6) is expressed by binding affinity (Ki), which is widely used to characterize the binding of molecules to receptors [5]. Modeling methods such as QSAR (quantitative structure–activity relationship) approaches have been found to be valuable in further optimization and development of novel inhibitors in drug design. Docking descriptors are calculated based on the major interactions between inhibitors and receptor. Thus, the obtained QSAR model from these descriptors could be very useful for rational drug design.

Results and discussion

In the present study, the interactions between a series of new synthesized N,N-Dimethyl-[9-(arylsulfonyl)-2,3,4,9-tetrahydro-1H-carbazol-3-yl]amines [6] and 5-HT6 receptor are identified. Then the binding affinity of these 5-HT6 receptor antagonists are predicted by protein structure- based QSAR and ligand- based QSAR modeling. The data set used in this study contains the values of the binding affinity of 5-HT6 receptor, reported in Ki (nM). Homology

modeling is used to obtain a 3D structure of the receptor. The amino acid sequence of 5-HT6 receptor was taken from the NCBI Web site. Then, it was loaded to the SWISS-MODEL server.

To realize the binding site of the 5-HT6 receptor, the best potent reported compound, was docked into the receptor. The minimum free energy of the receptor-drug was obtained (– 5.22 Kcal/mol). The binding sites were GLY 149, ALA 150, LEU115, CYS 118, LEU 199, LEU 122 and LEU 119. Docking study indicated that hydrogen bond and hydrophobic interaction of compounds with binding site of 5-HT6 receptor plays an important role in determining the potency of the drugs (fig 1).

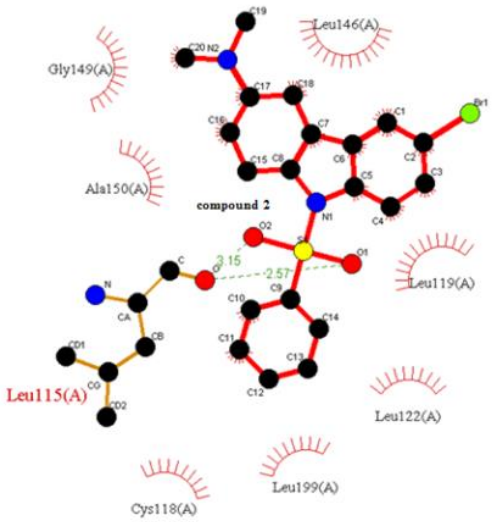


Fig 1. 2D scheme of interaction between compound 2 whit 5-HT6 receptor.

More than 1900 descriptors were calculated by using Dragon, Codessa, HyperChem and Autodock 4.2 software for docked drugs. Then, selected descriptors were related to the binding affinity using MLR and LS-SVM regression. The statistical parameters of the above models show that the MLR model was better than the LS-SVM model, and indicates that the relation between descriptors and experimental binding affinity is mainly linear. Finally, the results of protein based QSAR model were compared whit ligand based model and indicate the superiority of protein based over ligand based QSAR model (table 1).

Table 1. The statistical parameters of the protein based (I) and ligand based (II) QSAR model.

Model	R		SE		F		Q^2_{LOOCV}	SPRESS
	train	predict	train	predict	train	predict		
I	0.89	0.88	7.92	8.49	141.27	28.70	0.73	9.59
II	0.85	0.96	9.02	5.09	100.66	93.13	0.58	14.06

Reference

[1] Codony X, Vela J M, Ramírez M J. 5-HT6 receptor and cognition. *Curr.Opin. Pharm*. 2011;**11**: 94-100.
[2]Young S. N. How to increase serotonin in the human brain without drugs. J. Psychiatry Neurosci. 2007; **32** :394–399.
[3] Kroeze WK, Kristiansen K, Roth BL. Molecular biology of serotonin receptors structure and function at the molecular level. *Top. Med.Chem.* 2002 ; **6**: 507-28.
[4] Woolley M L, Marsden C A, Fone K C F. 5-ht6 Receptors. *CNS & Neurological Disorders - Drug Targets*.2004; **3**: 59-79.
[5] Kenakin T. Principles: Receptor theory in pharmacology. *Trends Pharmacol. Sci.* 2004; **25** : 186-192.
[6] Nirogi R V S, Konda J B, Kambhampati R, Shinde A, Bandyala T R, Gudla P, Kandukuri K K, Jayarajan P, Kandikere V and Dubey P K. *Bioorg. Med. Chem. Lett*. 2012; **22**: 6980-6985.

Determination of acidity constants of cinnamic acid derivaties using two rank annihilation factor analysis (TRAFA)

A.Benvidi*, T.Zarei, S.Jahanbani
Department of Chemistry, Yazd University, Yazd, Iran
E-mail address (corresponding author): benvidi89@gmail.com

Abstract

In this work, we introduce a simple, selective, sensitive and low cost procedure for determination of acidity constants of cinnamic acid and cinnamic acid derivaties in binary mixtures of solvent-water at 25° C by applying RAFA and TRAFA methods to pH gradual change-UV-Vis spectral data.. In RAFA and TRAFA methods, the pKa values were obtained from relative standard deviation (RSD) versus hypothetical pKa values. Acidity constants of cinnamic acid derivaties obtained by RAFA and TRAFA methods are very close to each other. The effect of cinnamic acid derivaties struchter on the pKa was examined and a logical relationship was obtaind. By these methods and without any prior knowledge about the system, concentration profiles and pure spectra can be obtained from the experimental data. Obtained data by TRAFA model have more agreement with real acidity constants of cinnamic acid and cinnamic acid derivaties than RAFA model.

Key words: Rank annihilation factor analysis, TRAFA, TAR, Acidity constant, Spectrophotometry, Cinnamic acid derivaties.

Introduction

Cinnamic acid is an organic compound with the formula $C_6H_5CHCHCO_2H$. It is a white crystalline compound that is slightly soluble in water, and freely soluble in many organic solvents [1]. Classified as an unsaturated carboxylic acid, it occurs naturally in a number of plants. It exists as both a cis and a trans isomer, although the latter is more common [2]. Spectrophotometry is one of the most powerful techniques for the investigation of solution equilibria, although potentiometric / pH metric titrations are more convenient and more commonly used because of the simplicity of the equipment and minimal time requirements. The accurate determination of acidity constant is so important in various chemical and biochemical courses [3]. There are several problems in determination of acidity constants, including low solubility in aqueous solutions and the low values of acidity constants. Therefore, mixed solvents had been chosen to overcome these problems. Rank annihilation factor analysis (RAFA) is an efficient chemometrics technique based on rank analysis of two-way spectral data and this technique can be employed for quantitative analysis of some systems with unknown background. When RAFA combined by a chemical model it is comparable with hard-soft modeling approach in solving some chemical problems. RAFA was originally developed by Ho et. al. as an iterative procedure [4]. To the best of our knowledge, there is no report for determination of acidity constants of cinnamic acid and cinnamic acid derivaties in gray mixture spectrophotometrically using RAFA [5] thus, in this work a simple, sensitive, selective and cheap procedure for the determination of acidity constants of cinnamic acid and cinnamic acid derivaties by applying RAFA to pH gradual change-UV-Vis spectral data (pH-spectra) is introduced. In addition RAFA and TRAFA methods are compared for determination of acidity constants of cinnamic acid and cinnamic acid derivaties.

Results and discussion

RAFA is a powerful chemometrics method for obtaining the acidity constant, especially when there is severe spectral overlap. Also, this method makes it possible to obtain pure absorption spectra and concentration profiles of species in several organic solvent-water binary mixture systems. As explained in the introduction section, RAFA is usually performed by annihilating one rank from the original data matrix. If the contributions of both acidic and basic forms of H_2A are simultaneously annihilated from the original data, the determination of the successive acid dissociation constants of diprotic acids can be feasible by RAFA, which this method was called two-rank annihilation factor analysis (TRAFA). Therefore, in this study TRAFA was used for the determination of the protolytic constants of TAR by applying pH gradual change-UV-Vis spectral data (pH-spectra). The acidity constants of cinnamic acid derivaties were calculated in the ethanol-water by applying different methods of RAFAand TRAFA. The obtained pKa values from RAFA and TRAFA are summarized in Table 1. The pKa values obtained by RAFA and TRAFA were compared with those obtained using famous chemometrics algorithm, DATAN. The results show that there is a very close agreement between the pKa values obtained by RAFA and TRAFA model and those obtained using DATAN (see Table 1).

Table 1. Comparison of the calculated acidity constants of cinnamic acid and cinnamic acid derivaties by RAFA and TRAFA in ethanol: water 10 %(w/w).

Material	Method	pK_{a1}	pK_{a2}
Cinnamic acid	RAFA	4.60	-
Ortho hydroxy Cinnamic acid	RAFA	4.14	9.61
	TRAFA	4.12	9.58
Meta hydroxy cinnamic acid	RAFA	4.45	10.02
	TRAFA	4.56	10
Para hydroxy cinnamic acid	RAFA	4.72	9.96
	TRAFA	4.65	9.95

References

[1] "Open Notebook Science Challenge: Solubilities of Organic Compounds in Organic Solvents". Nature Precedings. 2010. doi:10.1038/npre.2010.4243.3.

[2] Dorothea Garbe "Cinnamic Acid" in Ullmann's Encyclopedia of Industrial Chemistry, 2000, Wiley-VCH, Weinheim. doi:10.1002/14356007.a07_099.

[3] Baggesgaard-Rasmussen H. Titration schwacher Säuren und Basen in Wasser-Alkohol-Gemischen. Zeitschrift für analytische Chemie 1936; 105: 269-282.

[4] Liang YZ, Kvalheim OM, Manne R. White, grey and black multicomponent systems: A classification of mixture problems and methods for their quantitative analysis. Chemom. Intell. Lab. Syst 1993; 18: 235-250.

[5] Abdollahi H, Safavi A, Zeinali S. Model-based rank annihilation factor analysis for quantitative analysis of mixtures of monoprotic acids using multivariate spectrophotometric acid-base titrations. Chemom. Intell. Lab. Syst 2008; 94: 112-117.

Visualization of error propagation and one-factorat-a-time local sensitivity analysis concepts

Akram Rostami[a] , Hamid Abdollahi[a]*, Marcel Maeder[b]

[a] *Department of Chemistry, Institute for Advanced Studies in Basic Sciences, P.O. Box 45195-1159, Zanjan, Iran,*
[b] *Department of Chemistry, University of Newcastle, Callaghan, New South Wales 2308, Australia*

Email: abd@iasbs.ac.ir

Abstract

Until know everything about error propagation and sensitivity analysis was formulated. Sometime just using the formula cannot make a sense and it may be difficult for understand. In this study we will try to make a sense by visualization the error propagation and local sensitivity analysis, for ease to understand these concepts.

Key words: Error propagation, one-factorat-a-time (OAT) Lockal Sensitivity Analysis, Visualization

Introduction

Error propagation is simply the process of determining the uncertainty of an answer obtained from a calculation. Every time data are measured, there is an uncertainty associated with that measurement. If these measurements used in the calculation have some uncertainty associated with them, then the final answer will, of course, have some level of uncertainty. Determining the strength of the relation between a given uncertain input and the output is the job of sensitivity analysis [1,2]. A straightforward implementation of the "sensitivity" concept is provided by model output derivatives. If the model output of interest is Y, its sensitivity to an input factor X_i is simply $Y'_{X_i} = \partial Y/\partial X_i$. This measure tells how sensitive the output is to a perturbation of the input. If a measure independent from the units used for Y and X_i is needed, $S^r_{X_i} = (X^0{}_i/Y^0)(\partial Y/\partial X_i)$. In this latter case an alternative measure is $S^\sigma_{X_i} = (\sigma_{X_i}/\sigma_Y)(\partial Y/\partial X_i)$, where the standard deviations σ_{X_i}, σ_Y are uncertainty analysis' input and output, respectively, in the sense that σ_{X_i} comes from the available knowledge on X_i, while δ_Y must be inferred using the model [3].

Results and discussion

In this study we used a reversible chemical reaction A↔B. with reaction rates k_1 and k_{-1} for the direct and inverse reactions, respectively, whose solution, for the initial conditions (ICs) :

(1)
$$[A](t=0) = [A]_0$$
$$[B](t=0) = 0$$

(2)
$$[A] = \frac{[A]_0}{k_1 + k_{-1}}\left(k_1 e^{-(k_1+k_{-1})t} + k_{-1}\right) ,$$
$$[B] = [A]_0 - [A]$$

(3)
$$k_1 \sim N(3,0.3)$$
$$k_{-1} \sim N(3,1)$$

Here we assumed that the reaction rates are uncertain and describe by continuous random variables with known probability density functions (pdf's) that characterize their uncertainty as above (ref to eq 3). Based on this assumption existed a range for k_1 and k_{-1}.

In the first step k_1 or k_{-1} was fixed in the mean value "3" and another one was changed in the defined range. In the each set of k_1 and k_{-1} the concentration profile for A and B were calculated. When k_{-1} was fixed and k_1 was changed, the effect of uncertainty of k_1 on the concentration of A and B has been studied and vice versa. It was individual contribution of k_1 and k_{-1} in the error propagation equation.

In figure 1a we have demonstrated that uncertainty in both of k_1 and k_{-1} how changes the [A] and [B]. In the figure 1b error propagating base on this equation $\sigma^2_{[A]} \approx \sigma^2_{k_1}\left(\partial [A]/\partial k_1\right)^2 + \sigma^2_{k_{-1}}\left(\partial [A]/\partial k_{-1}\right)^2$ and numerical determination has been shown. As seen in this figure, the results of numerical and theoretical calculation of error propagation approximately are matching.

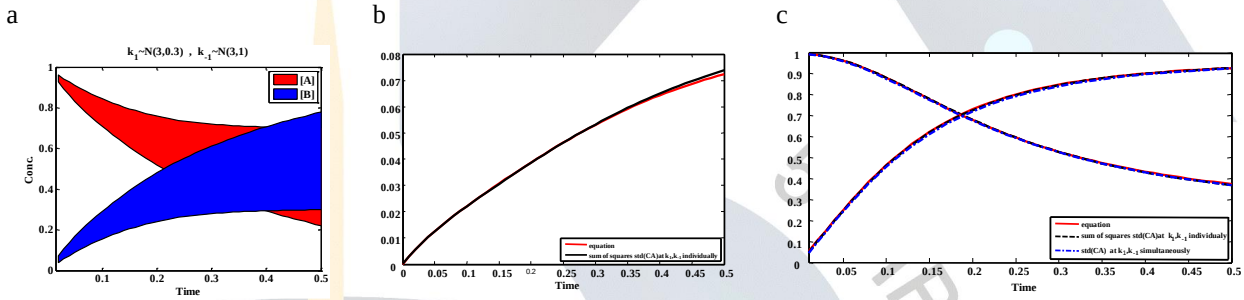


Figure 3: a) Time evolution of [A] and [B] when both of k_1 and k_{-1} simultanusly are uncertain parameters. B) Theoritical and numerical solution of error propagation in each time point. C) Theoritical and numerical solution of local sensitivity analysis.

Another concept that we have visualized was local sensitivity. This concept was calculated by $S^{\sigma}_{k1} = \frac{\sigma_{k_1}}{\sigma_{[A]}}\frac{\partial [A]}{\partial k_1}$, $S^{\sigma}_{k-1} = \frac{\sigma_{k_{-1}}}{\sigma_{[A]}}\frac{\partial [A]}{\partial k_{-1}}$. In addition to error propagation, local sensitivity could be numerically calculated correctly (figure 1c).

References

[1]Saltelli, A.; Chan, K; Scott, M. Sensitivity Analysis; John Wiley & Sons: Chichester, 2000.

[2]Saltelli, A.; Ratto, M.; Andres, T.; Campolongo, F.; Cariboni, J.; Gatelli, D.; Saisana, M.; Tarantola, S. Global Sensitivity Analysis. The Primer; John Wiley & Sons: Chichester, 2008.

[3] Saltelli, A.; Ratto, M.; Tarantola, S.; Campolongo, F. Chem. Rev. 2005, 105, 2811.

An efficient chemometric strategy based on wavelet transform for preprocessing of multi-capillary column – electronic nose data

Saeed Masoum*, Amir Hossein Alinoori

Department of Analytical Chemistry, Faculty of Chemistry, University of Kashan, Kashan, I.R. Iran

E-mail address: masoum@kashanu.ac.ir

Abstract

The multi-capillary column (MCC) that is hyphenated with electronic nose (E-nose) is a novel sensitive and strong technique with ability to separate components of odors to facilitate detection of them and it can be used in many applications. Small signal with high noise levels in MCC-Enose dataset demonstrates a necessity of noise removal technique before chemometric analyses. Signal averaging is the classical and the simplest smoothing method that traditionally have been used for noise reduction. Recently, many new algorithms based on transformation for smoothing, denoising and compression of analytical signals have been proposed. The use of the wavelet denosing technique improves the signal-to-noise ratio (SNR) along with amplifying the instrument resolution. The wavelet denoising method was applied to amplify significant information and suppress noise level and remove spike in the raw MCC-Enose dataset that represents a preference in comparing with another denoising method. The proposed methods have been performed on the real noisy raw MCC-Enose dataset with known peak positions as a typical noisy signal. The experimental results show that this method is robust denoising method and it also performs well in preserving signal features in MCC-Enose dataset.

Key words: Multi capillary column chromatography, Electronic nose, Wavelet transform, Signal denoising.

Introduction

An electronic nose (e-nose) is a device that identifies the specific components of an odor and analyzes its chemical to identify them [1]. Design and construction of new less expensive and more sensitive device with ability to separate components of an odor led to improvement of capability of detection of chemical components of unknown samples. This new version, MCC-Enose is a single unit containing both the multi-capillary column (MCC) for separation that is coupled with an array of metal oxide sensors and the processing units. Through a combination of Enose with multi-capillary column chromatography (MCC-Enose), a fast, robust, non-invasive and easy-to-use system for qualitative analyses of VOCs odour is now available. However, MCC-Enose spectra present chemometric challenges because of the high dimensionality (hyphenated system) and redundancy of information (several variables associated to a single VOC). This study is focused on preprocessing, denoising and smoothing of this MCC-Enose data by wavelet transform (WT), a mathematical technique that has been widely used in engineering. It was introduced to chemistry in 1990s and has now attracted the attention of many chemists. Most of the WT applications on processing chemical signals are based on the dual-localization property. Generally, the signal is composed of baseline, noise, and chemical signals. Since the frequencies of the three parts as mentioned above are significantly different from one another, it is not difficult to use the WT technique by removing the high-frequency part from the signal for denoising and smoothing and by removing the low-frequency part for baseline correction [2].

Results and discussion

When trying to denoise the signal by WT, the first problem encountered was the choice of wavelet. Too many wavelets could be employed to meet different requirements from various signal conditions. A possible way for optimization was that the noise structure (heteroscedasticity and correlation) of real signal was characterized at first. Simulated noise with the same structure was then added to an ideal signal that is simulated by a computer to form a simulated noisy signal, on which WT was performed with different wavelets [3]. Thus, the denoising efficiency could be evaluated in statistics by computing the root mean square error (RMSE) between the denoised signal and the ideal one. The wavelet producing the lowest RMSE was finally applied to the real signal as the optimum. The denoising efficiency was evaluated with other frequently applied denoising techniques such as Savitzky-Golay smoothing and moving average and verified to be superior to them in terms of noise removal and peak preservation.

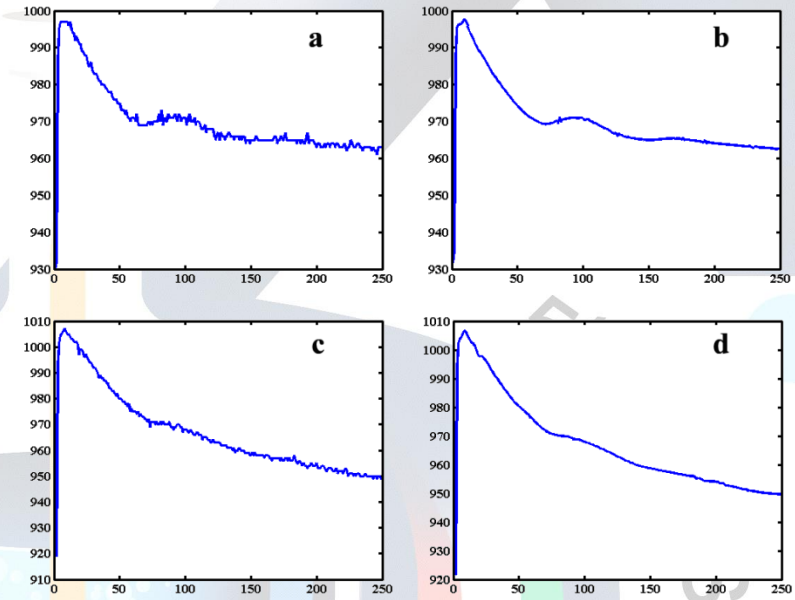


Figure 1. Signal denoising for MCC-Enose by WT. (a) real signal from MQ2, (b) denoised signal from MQ2, by db5 at decomposition level 6; (c) real signal from MQ3, (d) denoised signal from MQ3, by db5 at decomposition level 6

Figures 1.b and 1.d show the WT denoised signals by db5 at decomposition level 6. It was obvious that the noise was successfully removed and the shape of peak was well preserved. Even the slight variation of baseline was kept, comparing with original signal in Figure 1.a. and 1.c. No peak information was included in noise. It turned out that WT was a powerful tool of signal denoising for MCC-Enose preprocessing.

References

[1] Di Natale C. Electronic Noses. In: Moretto LM, Kalcher K, editors. Environmental Analysis by Electrochemical Sensors and Biosensors: Springer New York; 2014. p. 651-66.
[2] Liu B-F, Sera Y, Matsubara N, Otsuka K, Terabe S. Signal denoising and baseline correction by discrete wavelet transform for microchip capillary electrophoresis. *Electrophoresis*. 2003; **24**:3260-5.
[3] Pasti L, Walczak B, Massart DL, Reschiglian P. Optimization of signal denoising in discrete wavelet transform. *Chemometrics Intellig Lab Syst*. 1999; **48**:21-34.

Useing net analyte signal for determination of acidity constants caffiec acid in multivariate spectrophotometric analysis systems

A.Benvidi*, S.Jahanbani, T.Zarei

Department of Chemistry, Yazd University, Yazd, Iran

E-mail address (corresponding author): benvidi89@gmail.com

Abstract

A comparative study about advantages and limitations of net analyte signal (NAS) and DATa Analysis (DATAN) program in constant acid analysis has been performed. Net analyte signal (NAS) concept, which is a part of total signal that is directly related to the concentration of the component of interest. It monitors the concentration changes of any chemical species involved in the evolutionary process without requiring any pure spectra or having previous knowledge about the presence of the interferences. NAS method has some advantages such as the use of a full spectrum realization, therefore it does not require calibration and prediction steps and only a few measurements are required for the determination. By using these methods and without any prior knowledge about the system, concentration profiles and pure spectra can be obtained from the experimental data.

Key words: net analyte signal, DATa Analysis, Acidity constant, Spectrophotometry, caffiec acid.

Introduction

Caffeic acid (CA), 3,4-dihydroxycinamic acid is the major representative of hydroxycinamic acid in wines [1]. Caffeic acid is an organic compound that is classified as hydroxycinnamic acid. This yellow solid consists of both phenolic and acrylic functional groups.Caffiec acid is present in many fruits, vegetables, seasonings, beverages (coffee, wine) and olive oil. It is found in all plants because it is a key intermediate in the biosynthesis of lignin, one of the principal components of plant biomass and its residues [2]. The acidity and basicity form of dyes play a very fundamental role in many analytical procedures such as acid–base titration, solvent extraction and complex formation. Different methodologies have been proposed for the experimental determination of the acid dissociation constants including1H NMR spectroscopy, capillary electrophoresis, FT-IR spectrometry, UV–Vis absorption and fluorescence spectrophotometry and potentiometry. The net analyte signal (NAS) was defined by Lorberbased on spectroscopic methods, as the part of the spectrum of a mixture that is unique for the analyte of interest, i.e., it is orthogonal to the spectra of the interferences. The NAS is the part of the signal, which is directly related to the concentration predicted by the calibration model. In mathematical terms, it is the part of a spectrum which is orthogonal to the space spanned by the spectra of all analytes except one [3]. In this work, to obtain the NAS, the two methods Lorber et al and Goicoechea and Olivieri was used, followed by the two methods were compared. One of these methods (Goicoechea and Olivieri) is hybrid linear analysis (HLA) which can be applied provided a very accurately measured pure spectrum of the analyte is available [4, 5]. The DATAN program, proposed by Kubista and coworkers calculates spectral profiles, concentrations and equilibrium constants by utilizing equilibrium expressions that are related to the components. However, to the best of our knowledge, there is not any report in the literature so far acidity constants the determination of Caffeic acid in grey mixture using the difference of absorption spectra of the analyte at different pH values.

Results and discussion

In this study NAS were used for determination of the protolytic constants of caffeic acid by applying pH gradual change-UV-Vis spectral data (pH-spectra).The first step in NAS is choosing the analyte and calculating the rank-annihilated data matrix R-m. Here we will discuss the steps of NAS analysis for a diprotic acid, for which the R_{-m} matrix is obtained by either annihilating the contributions of the H_2A, HA or A^{2-} species. When the contribution of one species is annihilated from the total signal, R_{-m} contains spectral information for the remaining species and the spectral contribution from other sources such as interfering species). In this article, to obtain the NAS, the two methods Lorber et al and Goicoechea and Olivieri was used, followed by the two methods were compared. Comparison between methods of NAS and DATAN represents data close to each other. The acidity constants of caffeic acid were calculated in the AN-water by applying different methods of NAS and DATAN. The obtained pK_a values from NAS and DTAN are summarized in Table 1. This comparison indicates that the NAS method is an efficient method to obtain the constants of acid.

Table 1. The calculated acidity constants of caffeic acid in various Solvent percent (AN) by various methods.

Solvent percent	Method	pK_{a1}	pK_{a2}
AN: 20%(w/w)	NAS(Lorber)	5.95	9.70
	NAS(HLA)	5.92	9.65
	DATAN	6.05	9.81
AN: 60%(w/w)	NAS(Lorber)	6.00	9.90
	NAS(HLA)	5.98	9.95
	DATAN	6.09	10.05
AN: 80%(w/w)	NAS(Lorber)	6.15	10.25
	NAS(HLA)	6.20	10.30
	DATAN	6.25	10.42

References

[1] Franauet-Griell H, Checa A, Saurina J, Detrmination of polyphenols in spanish wines by capillary zone electrophoresis:application to wine characterization by using chemomtrics, J.Agric.Food chem., 2012;60:8340-8349.

[2] Boerjan W., Ralph J., Baucher M., Ligninbiosynthesis, Annual Review of Plant Biology, 2003; 54: 519-546.

[3] Lorber A., Error propagation and figures of merit for quantification by solving matrix equations, Anal. Chem., 1986; 58: 1167-1172.

[4] Bro R., Andersen C.M., Theory of net analyte signal vectors in inverse regression, J. Chemometrics, 2003; 17: 646–652.

[5] Mirmohseni A., Abdollahi H., Rostamizadeh K., Net analyte signal-based simultaneous determination of ethanol and water by quartz crystal nanobalance sensor, Anal. Chim. Acta, 2007; 585: 179-184.

Quantitative Nanostructure–Activity Relationship (QNAR) modeling for the cellular uptake of magnetofluorescent engineered nanoparticles in pancreatic cancer cells

Zahra Pahlavan Yali, Mohammed Hossein Fatemi*
Department of Chemistry, University of Mazandaran, Babolsar, Iran

E-mail address: zpahlevanyali@yahoo.com

Abstract

In this study, quantitative nanostructure–activity relationship (QNAR) was used for modeling the cellular uptake of magnetofluorescent nanoparticles. The nine descriptors selected by multiple linear regressions (MLR) were used as inputs of stepwise-multiple linear regression (SW-MLR), for the construction of the linear model. The artificial neural network (ANN) and support vector machine (SVM) methods were employed as nonlinear feature mapping. Comparison between the values of statistical parameters it showed that the SVM performed better than SW-MLR and ANN methods. The relative standard error (Se) of the training and test set for the SVM model were 0.117 and 0.245, and the correlation coefficients (R) were 0.957 and 0.863, respectively. While the square correlation coefficient of the cross validation (Q^2) was 0.614 for the training set, revealing the reliability of SVM model.

Key words: quantitative nanostructure–activity relationship, cellular uptake, nanoparticle, magnetofluorescent, support vector machine.

Introduction

One of the goals of nanotechnology in biomedical field is to functionalize biocompatible and inert compound to transmit precise biological functions. For this purpose, Union between nanotechnology and small-molecule chemistry will necessitate large scale production of nanomaterial with new surface properties. Magnetofluorescent nanoparticles as new materials have recently been Expressed for therapeutic and diagnostic use [1]. recent studies have shown that quantitative structure-property relationship (QSPR) modeling can be employed in computational nano particle studies such as fullerenes[2],carbon nanotubes(CNT)[3] and MNPs [4].

Methodology

The experimental values of cellular uptake of 109 magnetofluorescent (MNPs) in PaCa2 were taken [1]. These nanoparticles have exactly the same metal core decorated with reactive amines, carboxylic acids, alcohols, sulfhydryls or anhydrides. The cellular uptake is defined as decadic logarithm of the concentration (pM) of nanoparticles per cell [5]. The molecular descriptors were calculated using CODESSA (ver. 2.72)[2], DRAGON (ver. 3.0)[3], PaDEL (ver. 2.11)[4] and MATERIAL STUDIO software (vers. 4.3), on the basis of the minimum energy molecular geometries that were optimized with Hyperchem program (ver. 5.0).

Results and discussion

Definitions and the statistical parameters of the final selected descriptors of MLR appearing in this model are given in Table1. The statistical parameters for the SW-MLR, ANN and SVM models of the training set and test set are tabulated in Table2. So SVM is of better generalization performance than SW-MLR and ANN. Fisher statistic (F) and SPRESS values for training set on the SVM model respectively were 81.215 and 0.296.

Table1. Selected descriptors of Multiple Linear Regression.

No.	Descriptor name	Notation	Coefficient	Se
		Constant	1.249	1.085
1	3D-MoRSE - signal 29 / unweighted	Mor29u	-1.789	0.215
2	Moran autocorrelation of lag 7 weighted by I-state	MATS7s	-1.067	0.21
3	d COMMA2 value/weighted by atomic van der Waals volumes	DISPv	-0.034	0.009
4	Hydrogen bonding acceptor ability of the molecule	HADSA1	-0.01	0.001
5	Radial Distribution Function - 085 / weighted by atomic mass	RDF085m	0.063	0.018
6	3D-MoRSE – signal 29/ weighted by atomic van der Waals volumes	Mor29v	2.283	0.427
7	Counts the number of connected sp^2hybridised Carbon	C3SP2	-0.057	0.021
8	average connectivity index of order 4	X4A	1.492	0.421
9	Moran autocorrelation-lag 4/weighted by atomic mass	MATS4m	2.93	1.042

Table 2. The statistical parameters of different constructed QNAR models.

Model	$Se_{training}$	Se_{test}	$R_{training}$	R_{test}
SW-MLR	0.191	0.254	0.875	0.839
SVM	0.117	0.245	0.957	0.863
ANN	0.118	0.294	0.961	0.930

References

[1]Weissleder, Ralph, et al. "Cell-specific targeting of nanoparticles by multivalent attachment of small molecules." Nature biotechnology 23.11 (2005): 1418-1423.

[2]Liu, Huanxiang, et al. "Accurate quantitative structure-property relationship model to predict the solubility of C60 in various solvents based on a novel approach using a least-squares support vector machine." *The Journal of Physical Chemistry B* 109.43 (2005): 20565-20571.

[3] Yilmaz, Hayriye, Bakhtiyor Rasulev, and Jerzy Leszczynski. "Modeling the Dispersibility of Single Walled Carbon Nanotubes in Organic Solvents by Quantitative Structure-Activity Relationship Approach." *Nanomaterials* 5.2 (2015): 778-791.

[4] Singh, Kunwar P., and Shikha Gupta. "Nano-QSAR modeling for predicting biological activity of diverse nanomaterials." *RSC Advances* 4.26 (2014): 13215-13230.

[5]Fourches, Denis, et al. "Quantitative nanostructure– activity relationship modeling." ACS nano 4.10 (2010): 5703-5712.

Evaluation of long-heating kinetic process of edible oils using ATR-FTIR and chemometrics tools

Marjan Mahboubifar [1, 3], Bahram Hemmateenejad [1, 2,*], Katayoun Javidnia [1], Saeed Yousefinejad [2]

[1] *Medicinal and Natural Products Chemistry Research Center, Shiraz University of Medical Sciences, Shiraz, Iran*

[2] *Chemistry Department, Shiraz University, Shiraz, Iran*

[3] *Student Research Committee, Shiraz University of Medical Sciences, Shiraz, Iran*

Abstract

Long thermal oxidative kinetic and stability of four different edible oils (Colza, Corn, Frying, Sunflower) from various brands were surveyed with the use of attenuated total reflectance-Fourier transform infrared spectroscopy (ATR-FTIR) combined with multivariate curve resolution-alternative least square (MCR-ALS). Sampling from the heated oils (at 170 ºC) was performed each 3 hours during a 36-hours period. Changes in the ATR–FTIR spectra of the oil samples in the range of 4000–550 cm^{-1} were followed as a function of heating time. MCR-ALS was utilized to resolve the concentration and spectral profiles of three detected kinetic components. Three variations in resolved concentration profiles were related to the thermal-deduction of triacylglycerol of unsaturated acid, appearance of hydroperoxides and generation of secondary oxidation products. The kinetic profiles of these species were dependent on the type of oil. The proposed method can define a new way to monitor the oils' quality.

Keywords: Edible oils; Heating process; Oxidative kinetic; ATR-FTIR; MCR-ALS

1. Introduction

Checking the oil quality during heating is an important concern in food industries. oxidation of oil induced by heating is an important degradation reaction and an undesirable chemical change with significant effect on the oil's flavour, aroma, and nutritional quality and also may cause side effects on human health [1-3]. Therefore, it is so important to make sure about the oils' quality after alterations accelerated by heating of edible oils. Investigation of oil's oxidation processes is an approach for evaluation of its quality but most of the official methods are time-consuming, laborious and are environmentally hazardous. In the current research, MCR-ALS was utilized to resolve spectral and concentration pattern of the main produced/degraded species during the long-time heating treatment of edible oils.

2. Results and discussion

GC-MS was utilized to determine different fatty acids composition in the oil samples. The major fatty acid compositions in all kinds of the tested oils were generally with 14, 16, and 18 carbon atoms.

IR spectra of the oil samples were collected in absorption mode. The spectra were collected in the range 4000–550 cm^{-1}. The maximum differences on the infrared spectra of all types of oils during 36 hours heating procedure are in the regions around 1050–800, 3450, 2700 and 1650 cm^{-1} due to changes in their composition. It seems that these

differences correspond to variation in the unsaturation degree and form of the acyl groups and their chain length [4], and also assign to oxidation products such as hydroperoxydes, acids, aldehydes or ketones [2].

Rank analysis for estimating the number of components in the data matrix is the first step before MCR-ALS optimization. Herein, the number of chemical components in each time-derived spectra set was determined using the singular value decomposition (SVD) and evolving factor analysis (EFA) methods. However the presence of three components in spectra of heating process was indispensable, but in some case the probability of existence of more components (4 or 5) could not be neglected. Thus for more carefulness, we have performed MCR-ALS for 2-5 compounds.

The MCR-ALS optimization program was applied on all the recorded data during 36 hours heating of all oil samples by considering 2- 5 components as the optimum number of factors. Three approaches were used for making initial estimate to use in ALS process. According to the statistics results and the logical MCR-ALS extracted profiles (spectral and concentration), the existence of 3 components was confirmed even with different initial estimates.

Description of results shows that these three compounds were related to the deduction of triacylglycerol of unsaturated acid, appearance of hydroperoxides and secondary oxidation products during the 36-hours oxidation process of oils respectively. Due to differences which existed in the oil samples, the changing in the concentration profile of the detected compounds was different.

References

[1] Gonçalves, R. P., Março, P. H., & Valderrama, P. Thermal edible oil evaluation by UV-Vis spectroscopy and chemometrics. *Food Chemistry*, 2014; **163**, 83–6.

[2] Le Dréau, Y., Dupuy, N., Artaud, J., Ollivier, D., & Kister, J. Infrared study of aging of edible oils by oxidative spectroscopic index and MCR-ALS chemometric method. *Talanta*, 2009; **77**, 1748–1756.

[3] Ng, C.-Y., Leong, X.-F., Masbah, N., Adam, S. K., Kamisah, Y., & Jaarin, K. Heated vegetable oils and cardiovascular disease risk factors. *Vascular Pharmacology*, 2014; **61**, 1–9.

[4] Moros, J., Roth, M., Garrigues, S., & Guardia, M. de la. Preliminary studies about thermal degradation of edible oils through attenuated total reflectance mid-infrared spectrometry. *Food Chemistry*, 2009; **114**, 1529–1536.

Application of Extended ratio subtraction and simultaneous ratio subtraction methods for determination of Toluidine Blue and Methyl violet in their binary mixture

Ghadamali Bagherian, Naser Goudarzi, Zahra Zahednia*

Faculty of Chemistry, Shahrood University, Shahrood, Iran
zzahednia@gmail.com

Abstract

Two spectrophotometric methods, the extended ratio subtraction method (EXRSM) and the simultaneous ratio subtraction method (SRSM), have been developed for the simultaneous determination of Toluidine Blue (TB) and Methyl violet (MV) without prior separation. Linear calibration curves were obtained for single solution of TB and MV in the range of 0.5-25 and 0.5-12.5 $\mu grmL^{-1}$ *respectively. Under the optimum conditions, the EXRSM method showed a linear range 0.5-10* $\mu grmL^{-1}$ *for TB. Also using SRSM method, the TB and MV can be determined in the range 0.5-5* $\mu grmL^{-1}$ *for both dyes. The developed methods have been succssefully used for the simultaneous determination TB and MV in water sample.*

Key words: Toluidine Blue, Methyl violet, Extended ratio subtraction, Simultaneous ratio subtraction.

Introduction

Dyes as organic compounds, are widely used in various industril fields such as textiles, paper, rubber, plastics, leather, cosmetics, pharmaceuticals and food productions. Such extensive use of dyes often poses problems in the form of colored water which can affect plant life, and thus the entire ecosystem can be destroyed by the contamination of various dyes in water [1, 2]. In the recent years, the presence of dyes in the wastewater is an important enviromental problem and many different processes have been employed for eliminataing them such as biological treatment. The TB and MV belong to dyes classification of thiazin and triphenylmethane respectively and they are widely used in the textiles, foodstuffs, pulp and paper [3]. The release of these complex dyes in the environment cause toxicity problems. Therefore, simultaneous determination of these dyes is important.

The extended ratio subtraction method (EXRSM) and simultaneous ratio subtraction method (SRSM) are two novel methods that were introduced for the first time by Lotfy et al. They are simple, sensitive, and precise and have lower cost compared to the official methods [4]. In the perasent stady EXRSM and SRSM as two simple and accurate methods were used for simultaneous determination of TB and MV. To the best of our knowledge no studies have been reported for simultaneous determination of TB and MV by these two methods.

Results and discussion

The effects of different parameters such as the pH, concentration of the divisor, and the interferences were investigated for both dyes in the two methods. Simultaneous determination of two substance with overlapping spectra is difficult by spectrophotometric methods. But they are measurable with simplicity and high selectivity by the proposed

methods. For the simultaneous analysis of TB and MV, a binary mixture of these dyes, containing 5 $\mu grmL^{-1}$ of TB and 2.5 $\mu grmL^{-1}$ of MV in water were prepared and the zero-order absorption spectrum were recorded (Fig. 1). As the figure shows, due to the spectral overlapping, determination of TB and MV dyes in binary solution was not possible by the direct absorption measurement accurately. To solve this problem the EXRSM and SRSM methods were applied.The SRSM method is described briefly:

Firstly the spectrum of drug mixture was divided by a divisor, wich is zero-ordered spectrum of TB. Then by subtraction of constant, the term of $\frac{MV}{TB'}$ was obtained. By inserting $\frac{MV}{TB'}$ in ratio calibration, the concentration of MV was obtained. Therfore, the concentration of TB was calculated by multiplying constant by TB'.

$$\frac{MV+TB}{TB'}=\frac{MV}{TB'}+\frac{TB}{TB'}=\frac{MV}{TB'}+constant$$

$$\frac{MV}{TB'}+constant-constant=\frac{MV}{TB'}$$

$$\frac{TB}{TB'}=constant$$

$$TB=constant*TB'$$

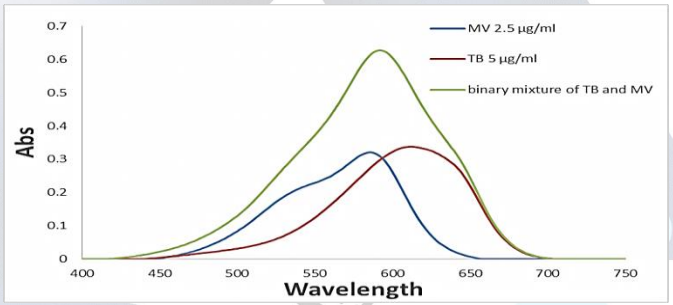


Fig. 1. Zero-order spectra of 5 µg/mL (TB), 2.5 $\mu grmL^{-1}$ (MV) in water, and the binary mixture of the same concentration.

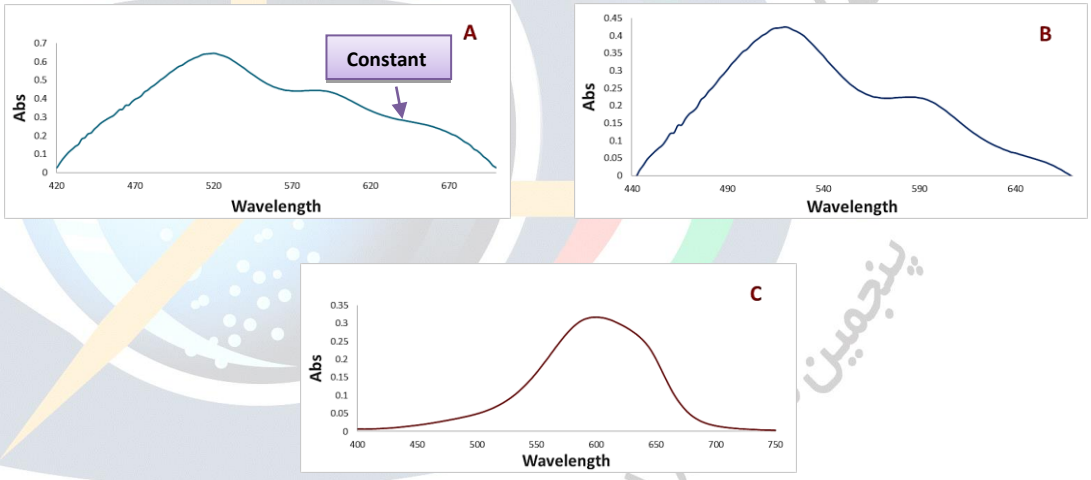


Fig. 2. (A) Absorption spectra mixture of TB 5 $\mu grmL^{-1}$and MV 2.5 $\mu grmL^{-1}$ using 25 $\mu grmL^{-1}$ of TB as a divisor; (B) The subtraction of constant from (A); (C) Multiply of the divisor (25 $\mu grmL^{-1}$) by the constant (Zero order absorption spectrum of TB) .

References

[1] Lafi R, Fradj A, Hafiane A, Hameed BH. Coffee waste as potential adsorbent for the removal of basic dyes from aqueous solution. *Korean J. Chem. Eng*. 2014; **31** ;2198–2206.

[2] Rauf MA, Qadri SM, Ashraf S, Al-Mansoori KM. Adsorption studies of Toluidine Blue from aqueous solutions onto gypsum. *J. Chem. Eng*. 2009; **150** ;90–95.
[3] Alhassani HA, Rauf MA, Ashraf SS. Efficient microbial degradation of Toluidine Blue dye by Brevibacillus sp. *Dyes and Pigments*. 2007; **75** ; 395–400.
[4] Lotfy HM, Hegazy MAM Simultaneous determination of some cholesterol-lowering drugs in their binary mixture by novel spectrophotometric methods. *Spectrochim. Acta Mol. Biomol. Spectrosc*. 2013; **113** ; 107–114.

Odor Visualization and Discrimination of Herbal Distillates Using a Colorimetric Sensor Array

Bahram Hemmateenejad*, Javad Tashkhourian, Mohammad Mahdi Bordbar, Nabiollah Mobaraki

Chemistry Department, Shiraz University, Shiraz, 71454, Iran

Email address: hemmatb@shirazu.ac.ir

Abstract

This work proposes a pattern-based recognition approach for the rapid discrimination of forty six Iranian herbal distillates using a low cost and sensitive colorimetric sensor array composed of 25 indicators. The color changes of the sensor exposed to the vapor of the herbal distillates can be monitored easily with an ordinary flatbed scanner. The digital representation of the array response was analyzed with Hierarchical Clustering Analysis (HCA). Using variable selection strategy, 6 indicators among the 25 employed indicators were selected as discriminant elements of the array. This colorimetric sensor array demonstrates excellent potential for quality assurance/control applications of herbal distillates.

Key words: Colorimetric sensor array, Discrimination, Herbal distillate, HCA, Variable selection

Introduction

Today, traditional systems of medicines (such as herbal distillates) become important resources for providing healthcare benefits because of inexpensive process economics, fewer side effects of plant–derived medicines and better compatibility of them with human body rather than modern medicines [1]. The ability to discriminate among closely similar herbal products is critical to ensure their efficacy [2]. To evaluate the quality control of herbal distillates, some chromatographic and spectroscopic techniques such as gas chromatography and Fourier transform infrared spectrometry can be available but they are destructive, expensive, and need time-consuming sample preparation [3]. To solve these limitation, electronic nose based on colorimetric detection have proven to be good alternatives for traditional techniques in nondestructive odor analysis of food [4]. The recorded patterns by sensor arrays need data analysis tools for extraction of useful information. Among the various chemometrics tools, HCA is the simplest and free model approach that provides a better representation of high dimensional data [5]. It classifies the analyte vectors according to their Euclidean distance and shows the degree of similarity of the array responses to different analytes [5].

Results and discussion

The responses of the designed sensor array to some studied herbal distillates are given in Fig. 1. Although the color responses reveal differences between the herbal distillates, the color change profiles are individual fingerprints for each specifics analytes and are distinguished excellent by eye. The array responses were obtained in the optimized conditions.

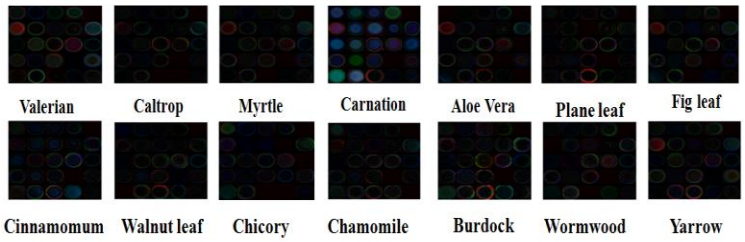


Fig. 1. Color difference maps of 14 herbal distillates.

As shown in Fig. 2, HCA resulted in clear discrimination of herbal distillates from each other with no misclassification. This clustering analysis created the library of array responses which tell us what the unknown sample is like to the existing library entries.

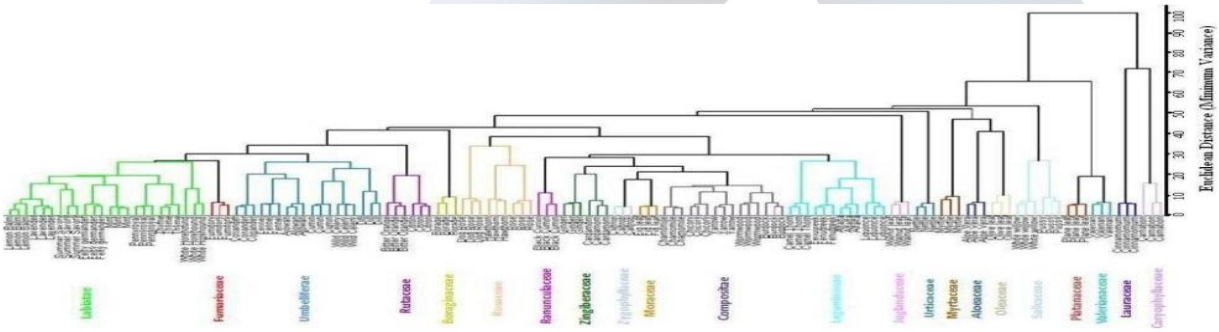


Fig. 2. HCA dendrogram of the colorimetric array responses to 46herbal distillates using minimum variance.

References

[1] Pan S, Neeraj A, Srivastava KS, Kishore P, A Proposal for a Quality System for Herbal Products. *J. Pharm. Sci.* 2013; **102** :4230-4241.

[2] Gao W, Yang H, Qi LW, Liu EH, Ren MT, Yan YT, Chen J, Li PJ, Unbiased metabolite profiling by liquid chromatography–quadrupole time-of-flight mass spectrometry and multivariate data analysis for herbal authentication: Classification of seven Lonicera species flower buds. *J. Chromatogr. A* 2012; **1245** :109-116.

[3] Cordell C, Moussa I, Martel AC, Sbirrazzuoli N, Cuvelier LL, Recent developments in food characterization and adulteration detection: technique-oriented perspectives. *J. Agric. Food Chem.* 2002; **50** : 1751−1764.

[4] Askim JR, Mahmoudi M, Suslick KS, Optical sensor arrays for chemical sensing: the optoelectronic nose. *Chem. Soc. Rev.* 2013; **42** :8649-8682.

[5] Ciosek P, WrÔblewski W, The analysis of sensor array data with various pattern recognition techniques, *Sens. Actutors B* 2006; **114** :85-93.

Evaluation of carbon nanotube sulfonation by chemometrics techniques

Mahsa Mohammadi, Amir Bagheri Garmarudi*, Mohammadreza Khanmohammadi, Mohammad Babaei Rouchi

Chemistry Department, Faculty of Science, Imam Khomeini Internatioal University, Qazvin, Iran

Corresponding Author: bagheri@sci.ikiu.ac.ir

Abstrac

The purpose of this study, was to evaluation of the structural and chemical changes in Carbon Nanotubes during the functionalization of CNT by spectroscopic – chemometrics method. A wet chemical functionalization process of carbon nano tubes (CNT) is done to study the process of sulfonation of CNT. Sulfonation was conducted by reflux condition. In order to evaluate the reaction evolution trend, diffuse reflectance mid IR spectroscopy was employed to obtain data along the process time and spectral data were processed by multivariate curve resolution alternating least squares (MCR-ALS) chemometrics techniques to monitoring the reaction.

Key words*: Chemometrics , IR Spectroscopy, Multiwall carbon nanotubes, Functionalization*

1. Introduction

The Multivariate curve resolution alternative least square (MCR-ALS) chemometric technique is, the methods of soft modeling analysis , this method based on factor analysis , can separation matrix data into two categories , concentration profile and spectra profile ,and not require to have information about the chemical model system , is a appropriate method for monitoring the process [1].FT-IR commonly used for Qualitative study of functional groups but in recent years quantitative analysis of organic and inorganic compounds in different matrix even in aqueous matrix have been developed in many application . In other ways FT-IR quantitative based methods provide fast signal and information about studied system. MCR-ALS associated to SIMPLISMA method resolves FT-IR data into pure spectra of species produced during the reaction and their associated concentration profiles, without a priori information.Therefor FT-IR combined with Multivariate Curve Resolution Alternative Least Square (MCR-ALS) have been used widely for monitoring of various chemical reaction.[2]

2. Results and discussion

2.1. Fourier Transformed Infrared Spectroscopy (FT-IR)

FT-IR spectroscopy is a well used qualitative technique for the evaluation of the chemical structures or characterization of chemically modified carbon nanotubes. In the IR spectrum of pristine CNT the most important band at 1600 cm^{-1} is related to stretching C=C , confirming SP^2 hybridation as the most common in the virgin CNT structure.Acid treated CNT shows signals at 1155 cm^{-1} due to C-O , 1583 cm^{-1} due to C=O , and 1634 cm^{-1} due to C=C . Also a peak at 2900 cm^{-1} is related to C-H, and the broad signal at 3500 cm^{-1} is due to O-H of –COOH groups. Comparison between pristine CNT with CNT-COOH spectrum reveals the success in functionalization by acid mixture, according to C=O and C-H peaks. The signal related to C-H in the spectral region below 3000 cm^{-1} confirms that carbon atoms in oxidized MWCNTs are in SP^3 hybridation , while in pristine CNT carbon atoms are in SP^2 hybridation. IR spectrum of S-CNT sample shows some spectral features at 1105 cm^{-1} and 1380 cm^{-1} assigned to symmetrical and asymmetrical sulphate group respectively. This result demonstrated sulphate group is successfully formed on the surface of CNTs [3].

2.2. Chemometrics data processing

In Figure 1 shows the 3D form of infrared spectral data obtained along the sulfonation reaction time. Optimizing the concentration and spectral profiles (Figures2) the PCA model explained 99.36 % of variance while residual standard deviation and PCA lacked fit were 0.013 and 1.058% respectively .They the model was reliable : As observed in concentration profile 2 components are reduced while an other one is increased solid line is related to oxidized CNT while is desending along the reaction . Related spectral profile demonstrates a wide peak at 3450 cm^{-1} due to OH of CH2OH functional group. The signal at 1634 cm^{-1} is due to C=O in kinone structure oxidized CNT which is reduced during sulfuration of CH2OH on CNT surface . An other peak at 1630 cm^{-1} due to sp2 hybrid c=c of CNT is also noticable .In case of dash line spectral profile a band at 1150 cm^{-1} is due to acylor alkoxy C-O is the intermediate structure during sulfunation .

2.3. Kinetics analysis

In case of solid line (ralated to CNT) , reduction rate is about twice of dash line (ralated to intermediate) and the formation rate of sulfure containing structure is much more less , which can be attrebouted to less rate of the second step of sulfuration process. This is reasonable , considering time and temperature of reaction (6h and 70 ºC) . On the other hand reducing trend of solid line (CNT) is semi –Logarithmic which looks like 1st order reactions . As the formation of sulfure structure on the surface of CNT is increasing during the reaction it can also reveal the 1st order kinetics.

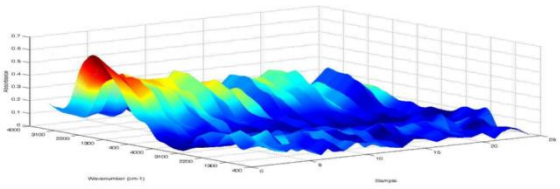


Figure 1- 3D form of infrared spectral data obtained along the sulfonation reaction time

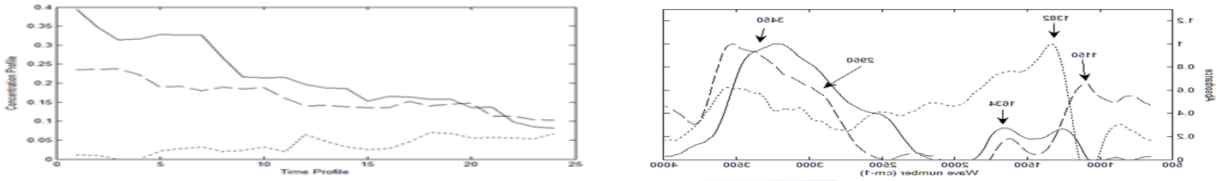


Figure 2- Concentration profile obtained from ALS optimization, Spectral profile obtained from AlS optimization

References

[1]. A. De Juan, S.C. Rutan, R. Tauler, D.L. Massart, Chemometr. Intell. Lab. Syst. 40 (1) (1998) 19–32.

[2]. JI, X., CHEN, Y., WANG, X. & LIU, W. 2011. Sulphate-functionalized multi-walled carbon nanotubes as catalysts for the esterification of glycerol with acetic acid. Kinetics and Catalysis, 52, 555-558.

[3]. Y Tian, Z. Hu, Y. Yang, X. Chen, W. Ji, Y. Chen, Thermal analysis–mass spectroscopy coupling as a powerful technique to study the growth of carbon nanotubes from benzene, Chem. Phys. Lett. 388 (2004) 259–262.

Study of human prostate anticancer activity of novel arylpiperazines using genetic algorithm and adaptive neuro-fuzzy inference system

Mahnaz Akrami , Zahra Garkani-Nejad*

Department of Chemistry, Faculty of Science, Shahid Bahonar University of Kerman, Kerman, Iran

E-mail address: z_garkani@uk.ac.ir

Abstract

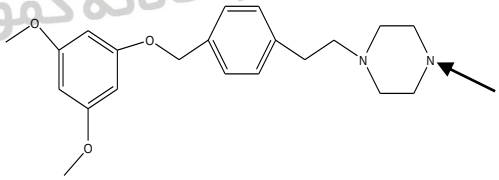
In this study, QSAR approach was carried out for the prediction of anticancer activity of novel arylpiperazine derivatives on two PC3 and DU145 human prostate cancer cell lines using adaptive neuro-fuzzy inference system. Genetic algorithm is utilized to select the most important variables and then these variables were used as inputs of ANFIS to predict anti cancer activity of arylpiperazine derivatives. The statistical parameter of R^2 for DU145 and PC3 cell lines are 0.968 and 0.963, respectively. Root mean square error (RMSE) for these two cell lines are 0.030 and 0.039, respectively. The ability and robustness of GA-ANFIS model in predicting anticancer activity of arylpiperazine derivatives are illustrated using cross-validation and Y-randomization techniques. Comparison of the results for GA-ANFIS, MLR-ANFIS and MLR-ANN methods indicated the proposed model is superior over the others.

Key words: Quantitative structure-activity relationship; Adaptive neuro-fuzzy inference system; Arylpiperazines

Introduction

Prostate cancer is the most common cancer in men and is the second-leading cause of cancer-related deaths [1]. Compounds with arylpiperazine moieties have a wide range of bioactivities including antiallergic, antidepressant, antipsychotic, antimalarial and antiplasmodial properties. Recent studies have shown that arylpiperazine compounds could possibly exert an anticancer effect and inhibit prostate cancer cell growth [2, 3]. These findings indicate that arylpiperazines might be useful as anti-cancer drugs. A series of novel arylpiperazinyl derivatives were synthesized based on piperazine to identify new anti-prostate cancer drug candidates. Obtained experimental data showed that some of these compounds exhibited strong anti-cancer activities against the tested cancer cells.

In the present study, relation between the biological activity of novel arylpiperazine derivatives on two types of cancer cells *(PC-*3 *and DU145)* and their molecular structures has been considered. The main aim of this study is to create mathematical models between measured biological properties of compounds and their structures. Then, created models would be used for predicting the anticancer activity of new compounds.



Results and discussion

In the present study, the data set consists of 26 molecules of novel arylpiperazine derivatives together with their anticancer activities (IC_{50}). The data set was divided into two groups of training and prediction sets. The 3D structures of the studied compounds were optimized using semi-empirical method of AM1. Structural descriptors were calculated for each compound using Dragon software. Then, we have examined stepwise multiple linear regression (MLR) and genetic algorithm (GA) techniques for selection the most important descriptors and artificial neural networks (ANN) and adaptive neuro-fuzzy inference system (ANFIS) for model developing. Finally, the validity of the generated models was investigated using cross-validation and Y-randomization techniques.

Table 1 shows the statistical parameters for different models for two cancer cells of PC-3 and DU145. It is shown in this work that GA as variable selection method combined with ANFIS as developing model can successfully predict anticancer activities (IC_{50}) of novel arylpiperazine derivatives better than the MLR-ANN and MLR-ANFIS modeling methods.

Table 1. Statistical parameters for different models

		PC-3		DU145	
Method	Set	R^2	RMSE	R^2	RMSE
MLR	Training	0.939	0.048	0.892	0.051
	Prediction	0.903	0.095	0.875	0.147
MLR-ANN	Training	0.941	0.050	0.915	0.046
	prediction	0.923	0.086	0.917	0.095
MLR-ANFIS	Training	0.948	0.033	0.964	0.037
	prediction	0.959	0.069	0.933	0.068
GA-ANFIS	Training	0.963	0.039	0.968	0.030
	prediction	0.958	0.063	0.944	0.061

References

[1] Greenlee, R.T.; Murray, T.; Hill-Harmon, M.B.; Thun, M.J. Cancer statistics, 2001. CA: Cancer J. Clin. **2001**, 51, 15–36.

[2] Kanda, H.; Ishii, K.; Ogura, Y.; Imamura, T.; Kanai, M.; Arima, K.; Sugimura, Y. Naftopidil, a selective alpha-1 adrenoceptor antagonist, inhibits growth of human prostate cancer cells by G1 cell cycle arrest. Int. J. Cancer 2008, 122, 444–451.

[3] Gotoh, A.; Nagaya, H.; Kanno, T.; Nishizaki, T. Antitumor action of α(1)-adrenoceptor blockers on human bladder, prostate and renal cancer cells. Pharmacology **2012**, 90, 242–246.

Simultaneous determination of drospirenone and ethinyl estradiol by spectrophotometric method using partial least square and support vector regression

Mahmoud Reza Sohrabi[1], Saeedeh Rahimi Farahani[1,*], Eslam pourbasheer[2]

1. Department of Chemistry, Faculty of Chemistry, Azad University, North Tehran Branch, Tehran, Iran
2. Department of Chemistry, Payame Noor University, Tehran, Iran

E-mail address: saeede.rahimi@ymail.com

Abstract

Partial least square regression (PLSR) and support vector regression (SVR) were used for simultaneous quantitative determination of drospirenone (DRSP) and ethinyl estradiol (EE) by spectrophotometric method in synthetic binary mixtures. For this purpose, a series of standard solutions were prepared for designing calibration set and 7 mixtures were prepared by choosing concentration of components randomly for test set in linear concentration range of each analyte and UV spectra of them were recorded from 200 to 300 nm. In PLSR method, the optimal number of factor was obtained 2 and 5 for DRSP and EE respectively with cross validation method and concentration of components were predicted. In SVR technique, the value for parameters ε (insensitive factor), γ (width of kernel function) and C (capacity factor) were optimized for training set using Leave- One- Out cross validation method based on minimum error in term of root mean square error (RMSE) and maximum value of Q^2_{Loo}. Root mean square error of prediction (RMSEP) for test set for DRSP and EE were obtained 0.7 and 1.6620 for PLS and 0.48 and 1.6579 for SVR and mean recovery was calculated 98.20% and 101.95% for PLS and 98.21% and 99.79% for SVR respectively.

Key words: PLSR, SVR, drospirenone, ethinyl estradiol, spectrophotometric method.

Introduction

Drospirenone (DRSP) is an analogue of spironolactone, its biochemical and pharmacological specifications resemble to endogenous progesterone, particularity concerning antimineralocorticoid and antiandrogenic activities. [1] Ethinyl estradiol (EE) is derivative of 17β – estradiol with having high estrogen potency which is the first orally active form of synthetic steroidal estrogen.Combination of DRSP and EE as a contraception is used to prevent pregnancy, it is also used to treat the premenstrual symptom (PMS), symptom of premenstrual dysphoric disorder (PMDD) [2, 3] and androgen dependent disorders entitled Hyperandrogenism [4]. literature survey revealed analytical methods just based on high performance liquid chromatography (HPLC) [5] and reversed phase high performance liquid chromatography (RP-HPLC) [6] for estimation of investigated drugs therefore, submitting simple technique for simultaneous quantitative determination of these drugs based on spectrophotometric methods like UV-visible along with chemometric methods such as partial least square as a basic tool and known method and support vector regression based on structural risk minimization [7, 8] can increase the speed of analysis and improve quality control of these components.

Results and discussion

The results indicated the ability of the mentioned multivariate calibration models to resolve the highly overlapped UV spectra of the 7 synthetic mixtures, yet using cheap and easy to handle instruments like the UV spectrophotometer. SVR model gives more accurate compared to PLSR model and show high generalization ability, however, PLSR still maintains the advantage of being fast to optimize and implement.

References

[1] Praveen C, Ranganath MK and Divakar P. Method Development and Validation for Simultaneous Estimation of Ethinyl Estradiol and Drospirenone and Forced Degradation Behavior by HPLC in combined dosage form, *Pharmaceut Anal Acta*, 2013; **4**: 231-235.

[2] Karla M, Carolyn W. The state of hormonal contraception today: established and emerging noncontraceptive health benefits, *American Journal of Obstetrics & Gynecology*, 2011.

[3] Biggs WS, Demuth RH. Premenstrual syndrome and premenstrual dysphoric disorder. *American Family Physician*, 2011; **84**: 918-24.

[4] M. guido, D. Romualdi, M. Giuliani, R. Suriano, L. Selvaggi, R. Apa, A. Lanzone. 2004, *The Journal of Clinical Endocrinology & Metabolism*, **89(6)**:2817–2823.

[5] K. Ahmed Shaikh, A. Tanaji Patil, S. katare. A Validated LC Method for the Simultaneous Determination of Drospirenone and Ethinylestradiol in Tablet Dosage form by Using Combine Fluorescence and UV Detectors, *Journal of Trace Analysis in Food and Drugs*, 2013; **1**: 30-37.

[6] S. Warkad, B. Santhakumari, A.V. Chandewar, Development and validation of a simple and sensitive RP-HPLC method for simultaoues estimation of drospirenone and ethinyl estradiol in combined tablet dosage form. *Int. J. of Pharm. Pharm. Sci.*, 2012; **4**: 452-457.

[7] R. G. Brereton, *Chemometrics; Data Analysis For The Laboratory and Chemical Plant*, John Wiley & Sons, 2003; 297.

[8] N. Chen, W. Lu, J. Vang, G. Li, *Support Vector Machine In Chemistry*, World Scientific Publishing, 2004.

Prediction of the acidity number of edible oils during long heating procedure using chemometrics tools based on FTIR-ATR results

Marjan Mahboubifar [1, 3], Bahram Hemmateenejad [1, 2,*], Katayoun Javidnia [1]

[1] *Medicinal and Natural Products Chemistry Research Center, Shiraz University of Medical Sciences, Shiraz, Iran*

[2] *Chemistry Department, Shiraz University, Shiraz, Iran*

[3] *Student Research Committee, Shiraz University of Medical Sciences, Shiraz, Iran*

Abstract

Determination of oil's acidity number is a method for evaluation of oil's quality, and is an important concern in food control laboratories. On the other hand, oxidation of oil induced by heating is an undesirable chemical change affect on the acidity number and quality of oils. Sampling of 14 oil samples from four types of edible oils (Canola, Corn, Frying, Sunflower), heated at 170 ºC, was performed each 6 hours during a 36-hours heating period. The ATR-FTIR spectra of the collected samples were recorded and then were used as input of PLS regression. The acidity number of the samples was determined by an standard procedure and were used as predicted variable. The acidity number of the oil samples was predicted using partial least square (PLS) and genetic algoritm-PLS (GA-PLS) as the chemometrics treatments on the ATR–FTIR spectra of the oil samples in the range of 4000–550 cm^{-1}. The effect of auto scaling and Multiplicative scatter correction (MSC) as data preprocessing methods was also investigated.

Keywords: Edible oils; acidity number; PLS; GA-PLS

1. Introduction

One of the most common methods for preparing of human kind foods in household kitchen and industries is Deep-fat frying at about 180 ◦C. Series of reactions such as hydrolysis, oxidation, isomerization, and polymerization take place during frying procedure witch influence on quality of the used cooking oil and final products [1]; Determination of acidity number of oils is a method for evaluation of oils quality and is an important concern in food control laboratories.

Public demand for quality and safety of oils and food products obviously requires appropriate analytical tools for fast and accurate analysis of them. Many analytical methods have been proposed for checking the oil's quality [2].

IR is a common and available apparatus in most quality control laboratories [3] which is a low costing, non-destruction, and environment protecting technique, whitch express a unique ''fingerprint'' spectrum for any compound; This feature enables this method to be used with chemometrics tools [4-5].

In the current research, the acidity number of oil samples from four different types (Canola, Corn, Frying, Sunflower) which was heated at 170 ºC for a 36-hours period, were estimated using partial least square (PLS) and genetic algoritm-PLS (GA-PLS) as the chemometrics treatments on the ATR–FTIR spectra of the oil samples. This method is proposed as a fast technique for direct acidity number quantification in edible oil samples, without any usage of organic solvents or previous treatment.

2. Results and discussion

The oil samples heated at 170 °C using a heater thermostated and sampling was done each 6 h for duration of 36 h. From each oil, 7 samples were collected during the heating process and 98 samples were totally collected for 14 oil samples which were from four different types of edible oils.

IR spectra of the oil samples were collected in absorption mode in the range of 4000–550 cm^{-1}. The maximum differences on the infrared spectra of all types of oils during 36 hours heating procedure were in the regions around 1050–800, 3450, 2700 and 1650 cm^{-1} due to changes in their composition.

The acidity number the oil's samples were measured by titration of each sample with a solution of potassium hydroxide.

The number of 98 oil samples divided to two different sets: calibration set (78 samples) and prediction set (20 samples); and the acidity number of the data sets determined using partial least square (PLS) and genetic algoritm-pls (GA-PLS) on the FTIR-ATR results. Under the best condition studied, the root mean square error (RMSE) values for calibration and prediction sets were respectively 0.09 and 0.13. The results show that these models have the capacity for determination the acidity number of the tested oils and could be proposed as a fast technique for acidity number quantification.

References

[1] Zhang, Q., Saleh, A. S. M., Chen, J., & Shen, Q. Chemical alterations taken place during deep-fat frying based on certain reaction products: A review. *Chemistry and Physics of Lipids*, 2012; **165**, 662-681.

[2] Giacomelli, L., Mattea, M., & Ceballos, C. Analysis and characterization of edible oils by chemometric methods. *Journal of the American Oil Chemists' Society*, 2006; **83**, 303-308.

[3] Moore, J. C., Spink, J., & Lipp, M. Development and application of a database of food ingredient fraud and economically motivated adulteration from 1980 to 2010. *Journal of Food Science*, 2012; **77**, 118-126.

[4] De Luca, M., Terouzi, W., Ioele, G., Kzaiber, F., Oussama, A., Oliverio, F., Tauler, R., & Ragno, G. Derivative FTIR spectroscopy for cluster analysis and classification of morocco olive oils. *Food Chemistry*, 2011; **124**, 1113-1118.

[5] Zhou, Y., Liu, T., & Li, J. Rapid identification between edible oil and swill-cooked dirty oil by using a semi-supervised support vector machine based on graph and near-infrared spectroscopy. *Chemometrics and Intelligent Laboratory Systems*, 2015; **143**, 1-6.

Metal-dependent activity of the 8-17 DNAzyme studied by resolution of rank deficient multi-way fluorescence data

Dr.Mohsen Kompany-Zareh*, Zohreh Shomali

Institute for Advanced Studies in Basic Sciences, Zanjan, Iran

kompanym@iasbs.ac.ir

Abstract

The recent advances in multi-way analysis provide new solutions to traditional enzyme activity assessment. In the present study enzyme activity has been determined by Hard-modeling and restricted-Tucker3. In Tucker3 analysis of three-way data array obtained from a chemical or biological system, it is sometimes possible to use a priori knowledge about the system to specify what is called a restricted Tucker3 model. Often, the restricted Tucker3 model is characterized by having some elements of the core forced to zero. The method relies on monitoring spectral changes of reaction mixture at specific time points during the course of the whole enzyme catalyzed reaction and employs multi-way analysis to detect the spectral changes. Parameters of chemical model (rate constants), interaction (for core of Tucker3) and initial concentrations (enzyme, substrate and metal ion) are estimated during optimization process by Replacement Method (RM). The work presented in this paper was undertaken to test the hypothesis that FRET and chemometric multiway analysis can be used as a universally applicable approach for rapid assessment of enzyme activity without using any external standards. A simulated chemical equilibrium data set is used to evaluate the applicability of this hypothesis. The obtained information from our study by using chemometrics methods will be helpful for prediction of DNAzyme behavior during processes and determination of substrate yield that is cleaved.

Key words: Three-way data, Restricted Tucker3, Rank deficiency, Hard modeling, Replacement method

Introduction

One of the most important discoveries in the last decades is that DNA molecules are not only materials for genetic information storage, but also catalysts for a variety of biological reactions, and therefore DNA molecules with catalytic properties are called catalytic DNAs or DNAzymes. DNAzymes that catalyze the cleavage of RNA are by far the largest class of catalytic DNA molecules [1] and almost is applied as quantitative analysis of cations[2] (Figure1). A wide variety of techniques have been developed to investigate metal-dependent activity of DNAzyme [3]. In this study, three-way data array was recorded by measuring excitation-emission fluorescence during the titration of DNAzyme-substrate complex with metal ion.

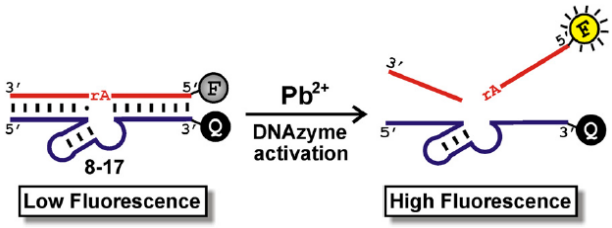


Figure 1. A Pb2+ Biosensor Based on 8-17.[Ref. 1]

Smilde and Kiers [4,5] have pioneered the work on restricted Tucker3 models. Based on the idea that PARAFAC may be too restricted and simple for many problems and that Tucker3 is too flexible, they developed an idea of using prior chemical/physical knowledge of the system studied to fix certain elements in the Tucker3 core array to zero and thereby in the end develop new models; restricted Tucker3 models, which are unique and chemically meaningful.

In kinetic hard-modelling concentration profiles are calculated by numerical integration of the rate laws describing the postulated kinetic hard-model. For some kinetic models and experimental conditions, however, the concentration matrix is rank deficient and the pure component spectra cannot be computed, as the linear regression step cannot be performed. Different solutions have been proposed in order to circumvent this rank deficiency problem and to allow the fitting of pure component spectra [6]. Here we used some independently known component spectra to the analysis by restricted Tucker3.

A (P, Q, R) component Tucker3 model on a data array $X \in RI \times J \times K$ can be represented in matrix notation as in Eq. (1), $X=AG(C \otimes B)T + E$ (1)

Where $A \in RI \times P$, $B \in RJ \times Q$ and $C \in RK \times R$ are the component matrices in the first, second and third mode, respectively, and G represents the core array ($G \in RP \times Q \times R$) matricized in the first mode. X is the metricized data array and E holds the residual variation not explained by the model. The symbol $\otimes$ denotes the Kronecker product.

In order to optimize the parameters (rate constants (k1:k5), interaction paramters at core of Tucker3 (g5:g13) and nitial concentrations (c1:c3)) replacement method (RM) was used. RM is a very simple method with prediction ability in a simple, rapid, high performance and reproducible way.

In the present case study,we apply this systematic method to sets of kinetic data, monitored by fluorescence spectroscopy, recorded during the course of the reaction of 8-17 DNAzyme with substrate catalysed by metal ion (see Fig. 1). FRET ocurrs between Enzyme and substrate strand. Therefore, rank deficiency of FRET is excitation and emission modes. Kinetic mechanism of this reaction will discussed here by hard modeling. First step of reaction is second order and rank deficiency because of closure will be possible. We applied independently known component spectra to the analysis by restricted Tucker3 to this catalytic reaction in order to resolve of chemical system with the rank deficiency in three mode (concentration, excitation and emission). Parameters optimized by replacement method.

The obtained information from our study by using chemometrics methods will be helpful for prediction of DNAzyme behavior during processes and determination of substrate yield that is cleaved.

Results and discussion

We supposed chemical model for cleavage activity of DNAzyme at Figure 2.

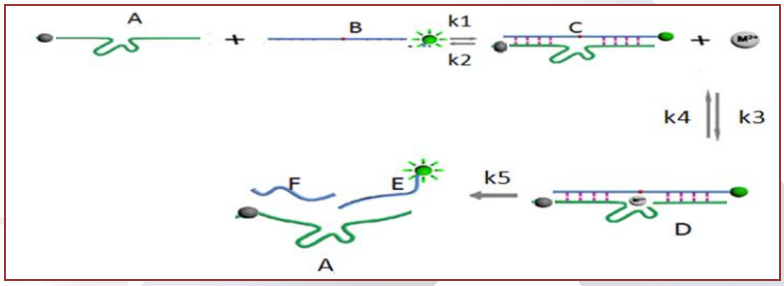


Figure 2. Mechanism of metal-dependent activity of DNAzyme

Simulated (81time×21em×31ex) fluorescence data array from complexometric titration of DNAzyme-substrate complex C with metal ion M (a cleavaging agent) is considered. The concentration profiles of each species of considered analyte (A, B, M, C, D, E) can be expressed as a function of five parameters (k1:k5) and initial analytical concentrations (c1:c3 for enzyme, substrate and metal ion respectively). F is non-absorbing species.Restricted-Tucker3 is applied for analysis of three-mode rank deficient chemical system. Dimentions of core array G is (6×3×3) for concentration, excitation and emission modes, respectively. Because we need six concentration profiles for explain chemical model and three independent excitation and emission profiles for A, B and M. Interaction parameters (g5:g13) at core of G are very informative, they estimated during optimization process. Other interaction prameters supposed are known. Replacement Method (RM) is applied to optimized all of parameres (3initial concentration+5rate constants+9interaction). Optimization algorithm is replicated until estimated parameters is converged and residual error for reconstruction of data be minima. Mentioned algorithm can estimate enzyme dosage and substrate concentration without needing calibration samples. In biological systems, RNA-cleaving DNAzymes can be used as therapeutic agents for treatment of cancer and a number of viral and microbial diseases. The obtained information from our study by using chemometrics methods will be helpful for prediction of DNAzyme behavior during processes and determination of substrate yield that is cleaved.

References

[1] K. Schlosser and Y. Li, Chemistry & Biology, 2009; **16**: 311-322.

[2] J. Li and Y. Lu, J. Am. Chem. Soc., 2000; **122**: 10466-10467

[3] K:. Schlosser and Y. Li, Chemistry & Biology, 2009; **16**: 311-322.

[4] A.K. Smilde, Y. Wang, B.R. Kowalski, J. Chemometr. 1994; **8**: 21–36.

[5] H.A.L. Kiers, A.K. Smilde, J. Chemometr. 1998; **12**: 125–147.

[6] J. Billeter, Y. Neuhold, K. Hungerbühler, Chemometrics and Intelligent Laboratory Systems, 2009; **98**: 213–226

A review on the application of chemometrics methods in the field of industrial lubricants

AydaYousefian*, Abolfazl Semnani, Gholamhossein Azimi[2], Hedayat Haddadi, Shima Ghanavati nasab

[1]*Department of Chemistry, Faculty of Sciences, University of Shahrekord, Shahrekord, P.O. Box 115, Iran*

[2]*Department of Chemistry, Faculty of Sciences, University of Isfahan, Isfahan, Iran*

.

aydayousefian@yahoo.com

Abstract

Current methods of analyzing the industrial lubricants involve the consumption of considerable amounts of these fluids. Also, these routes are time consuming. Therefore, developing new methods which are fast and can be performed by the least quantities of the mentioned materials can have a positive effect on a vast range of industries. In this paper different applications of chemometrics methods in the field of lubricants have been reviewd. The results indicate that classification, calibration and multiresponse optimization associated with spectroscopic techniques are the most employed methods. Also, the foundings show that little attention has been made by Iranian chemometrician to this subject. The authors hope this review would be a first step to join chemometrics with lubricanting industry in our country.

Key words: Classification, Calibration, Optimization, Spectroscopy, Lubricants.

Introduction

Lubricants are commercial products usually prepared through the mixture of base oils and additives; the former in a proportion ranging from 70 to 99%. Additives are used to strengthen the features of the base oil and/or to supply additional properties to the base oils [1]. Lubricants are substances introduced to reduce friction between mobile surfaces and hence prevent the wear of material surfaces as a consequence of relative mutual movement. They play other functions as tightness improvement , cooling action by removing the heat produced by the contact of moving parts, prevention of corrosion and rust against weathering and contaminative agents such as, oxygen, water, acids, glycol, fuel, sulphure and keeping the different elements clean as detergents. Therefore monitoring the lubricating oil condition is essential because of their key role in machine duration and performance.

The main properties of lubricants, which are usually indicated in the technical characteristics of the product, are: viscosity, viscosity index, pour point, flash point, demulsibility, Base number, acid number, etc. The mentioned properties of lubricating oils are usually determined by the standard methods such as ASTM, ISO and others [2].

These methods assessing lubricant properties are time consuming and expensive. The corresponding tests require specific equipments and large amounts of sample for the determination of each parameter of interest. Therefore it could be interesting to look for alternative or complementary methods like spectroscopic methods combined with chemometrics techniques.

Results and discussion

Specroscopic methods especially vibrational spectroscopy (IR, NIR and Raman) have been widely applied as an analytical technique for the determination of physical and chemical properties for petroleum products, because of the fact that most products of petroleum refining and petrochemicals consist of hydrocarbons and other compounds, which have spectra in range 4000-14000 cm^{-1} [1].

Chemometrics technique especially multivariate data analysis (MDA), have boosted the use of IR instruments. Only MDA methods are able to process enormous amounts of sophisticated experimental data that are provided by IR techniques [2].The correlation between NIR spectra and the carcinogenic potential of base oils was proved by employing principle component analysis (PCR). MIR spectroscopy and partial least squares regression (PLS) were used to determine the chemical composition (paraffines, Iso paraffines, naftens, aromatics and heteroaromatics) and its influence on the physic-chemical properties (viscosity and viscosity index) in mineral based lubricants. The potentiality of MIR spectroscopy for the prediction of viscosity in lubricating oils for Locomotives and diesel engines was assessed by PCR and interval-pls. MIR and NIR spectroscopy have been employed for the prediction of contaminants, degradation products and additives employing PCR, PLS and Interval-PLS [4]. The possible differentiation and classification between different types of commercial motor oils by using IR spectroscopy and mathematical data analysis have been shown. Many studies from Balabin et al have revealed the importance of quality control of motor oil adultration by means of NIR. A large number of chemometrics techniques like SIMCA, PLS; KNN, MLP and SVM were used to evaluate the origin of motor oils according to their base stock and to their kinematic viscosity at low and high tempratures [1]. An extreme vertices mixture design and multiple response optimizations through a desirability function (D) were applied for the first time to formulate and optimize hydrulic and turbine lubricant [5]. FTIR coupled to chemometrics techniques (PLS, CLS and PCR) were used to study of motor oil adultration and oil parameters prediction. (Viscosity index and base number) [3]. Chemometrics treatments conjunctions with spectroscopic data have had a vital role in development of classification, optimization, calibration, modeling, adultration, determination of lubricants origin, monitoring the service condition and regression in lubricants oils. Unfortunately, Iranian chemometricians do not have enough information about the application of novel and helpful methods in industry. In the presented short review, the authors have tried to represent versatile application of spectroscopic techniques associated to multivariate analysis for quality and quantity control of lubricants. The authors hope this review would be a first step to join chemometrics techniques to lubricant industry of Iran and introduce applicable aspects of chemometrics to interested Iranian reaserchers.

References

[1] Balabin R.M, Safieva Z, Lomakina I. Near- infrared (NIR) spectroscopy for motor oil classification from discriminant analysis to support vector machine. *Microchemical journal.* 2011; 121-128

[2] Gracia. N, Thomas .S, Bazin.P, Duponchel.L, Lerasle.O. Combination of mid-infrared spectroscopy and chemometrics factorization tools to study the oxidation of lubricating base oils. *Catalysys today.* 2010; **155**:255-260.

[3] Al-Ghouti M.A, Al-Degs Y.S, Amer.M. Application of chemometrics and FTIR for determination of viscosity index and base number of motor iols. *Talanta*. 2010; **81**:1096-1101.

[4] Bassbasi.M, Hafid.A, Tauler.R, Oussama.A.Study of motor oil adultration by infrared spectroscopy and chemometrics methods. Fuel. 2013; **104**:798-804

[5] Khalilian.H, A.Semnani, H.Haddadi, M.Nekoeinia Multiresponse Optimization of a Hydraulic Oil Formulation by Mixture Design and Response Surface Methods. *Journal of Tribology*.2015, **138**:1015-1024

.

Determination of Ofloxacine in human plasma using dispersive liquid-liquid microextraction combined with spectrofluorimetry with the aid of box-behnken design

Farah Assadian[a], Ali Niazi*[a], Jahan B. Ghasemi[b]

[a]*Department of Chemistry, Faculty of Science, Arak Branch, Islamic Azad University, Arak, Iran*
[b] *School of Chemistry, University College of Science, University of Tehran, P.O. Box 14155-6455, Tehran, Iran.*
E-mail address (corresponding author): ali.niazi@gmail.com

Abstract

In this research an improved method for the determination of Ofloxacine in human plasma has been developed using dispersive liquid–liquid micro extraction (DLLME) by fluorescence spectrometry analysis. The volume of extraction solvent (CHCl3), disperser solvent (acetonitrile), pH, and extraction time were selected as interested variables in DLLME process. Firstly, a four factor, three-level Box-Behnken experimental design employed to identify the important factors. Then the explanation and optimization relationship between the response and the important factors was investigated by response surface methodology (RSM). The optimum experimental conditions found from this statistical evaluation were included: sample volume: 10.0 mL, volume of extraction solvent (chloroform): 260 µL, volume of disperser solvent (acetonitril): 0.5 mL, pH: 4.5, and extraction time: 30.0 min.Under the optimum conditions, the preconcentration factor of 183 was achieved. Calibration graph was linear in the range of 0.5-8 mg L^{-1} with correlation coefficient of 0.982. The Enrichment factor (EF) 171.6 was obtained for Riboflavin determination.

Key words: Dispersive liquid-liquid microextraction; Ofloxacine; Box-Behnken design; fluorimetry; response surface methodology; human plasma.

Introduction

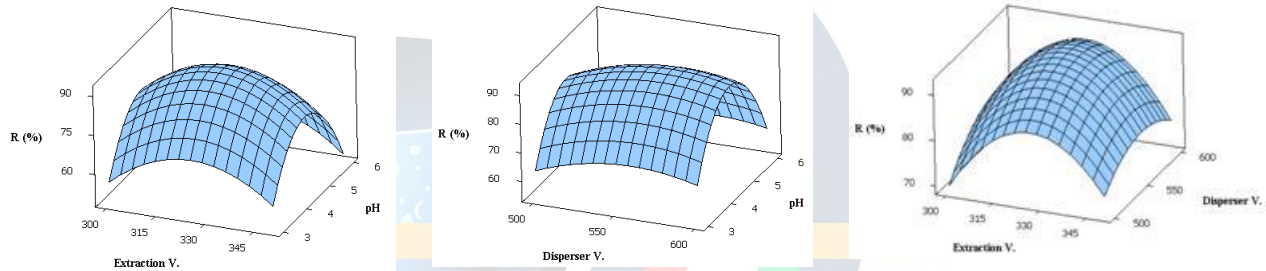
Ofloxacine (OFL), is a member of the third generation fluoroquinolones (FQ) with antibacterial activity. This antibacterial agent is widely used in the treatment of respiratory tract, urinary tract and tissue-based infections [1]. Several analytical methods have been developed for OFL determination in pharmaceutical formulations and biological samples. Dispersive liquid-liquid microextraction (DLLME) is one the practical methods that offers several advantages such as high recovery, simplicity of operation, low cost and rapidity. It has been widely applied to preconcentrate and separate organic and inorganic compounds in different samples [2]. In this research the effective factors in efficiency of DLLME are studied by Box-Behnken design that is a second-order multivariate design technique based on three-level factorial designs. Then, the effects of factors at different level and their influence on each other considered by response surface methodology (RSM) [3].

Results and discussion

In this study, a new method based on microextraction technique described as dispersive liquid–liquid microextraction (DLLME) has been developed to determine the Ofloxacine in human plasma. The optimization of the variables of DLLME was carried out by using response surface methodology and experimental design. There are four

factors, volume of extraction solvent, volume of disperser solvent, initial pH of the solution, and the extraction time, which can affect the extraction in DLLME process. Acetonitrile, acetone, and methanol were tested as the disperser solvents and chloroform, Carbon tetrachloride, 1,2 dichloroethane and chlorobenzene were studied as the extraction solvents. Most suitable disperser and extraction solvent pair was selected as acetonitrile–chloroform solvent.

Response surface methodology (RSM), involving Box-Behnken design matrix in four most important operating variables, were employed for the study and chose the best condition for extraction of OFL. The 27 experiments were required to investigate the effect of parameters on extraction process and also optimizing the experimental parameters and choosing the best condition by determination effective factors. The optimized conditions for this process were; volume of extraction solvent: 260 µL, volume of disperser solvent: 0.5 mL, pH: 4.5, and extraction time: 30.0 min. Under the optimum conditions, the preconcentration factor of 183 was achieved. The Enrichment factor (EF) 171.6 was obtained for OFL determination. The employed working range was set from 0.5-10 mg L^{-1} (number of calibration points, N=11). The experimental results showed good linear relationship between emission intensity and the concentration with a correlation coefficient of 0.982. The three-dimensional response surfaces plots were generated by Minitab software (Figure 1). Response surface plots when optimizing the following pair of factors, while maintaining constant the remaining one at its mean values. The influence of the different process variables on the response factor (% removal) are visualized in the figure. In extraction processes, the extraction solvent volume, disperser solvent volume, and solution pH play important role. These variables are considered to be most effective in influencing the extraction process.



(a) (b) ©

Figure 1. Effect of different pair of factors on the extraction process

These results indicated the proposed method had high sensitivity and stability and high potential to be a powerful and suitable preconcentration tool for trace analysis.

References

[1] Ichihara N, Tachizawa H, Tsumura M, Une T, Sato K. Phase I study on DL-8280, Chemotherapy, 1984; 32:118-149.

[3] Zeeb M, Ganjali MR, Norouzi P. Dispersive liquid-liquid microextraction followed by spectrofluorimetry as a simple and accurate technique for determination of thiamine (vitamin B1). Microchim. Acta, 2010; 168:317-324.

[5] Govender S, Pillay V, Chetty DJ, Essack SY, Dangor CM, Govender T. Optimisation and characterisation of bioadhesive controlled release tetracycline microspheres. Int J Pharm 2005;306:24–40.

Rapid chromatographic analysis of selected UV filters and parabens in sunscreen products exploiting second-order calibration methods

Nafise Shekari, Maryam Vosough* and Kourosh Tabar-Heydar

Chemistry and Chemical Engineering Research Center of Iran, P.O. Box 14335-186 Tehran, Iran

E-mail address: vosough@ccerci.ac.ir

Abstract

In the present study, a simple strategy based on liquid-liquid extraction (LLE) followed by fast high-performance liquid chromatography (HPLC) with diode array detection coupled with chemometrics tools has been proposed for the determination of five UV filters and two parabens in 31 sunscreen products in less than 4.5 min. At first, a full factorial design (FFD) as a screening step and then a Box-Behnken design (BBD) have been applied to optimize the factors influencing the chromatographic separation quality containing flow rate, methanol and acetonitrile volume fractions. The performances of MCR/ALS and ATLD, as second-order algorithms, were studied and comparable results were obtained from implication of these modeling methods. Acceptable qualification and quantification results were achieved in the presence of matrix interferences and the second-order advantage was fully exploited with superiority of MCR-ALS modeling. The average recoveries were in the rage of 81.1% - 119.6% and relative standard deviation values were lower than 9 %. The limits of detection were in the rage of 0.006-0.067 μg mL^{-1} and the limits of quantification were in the rage of 0.020-0.222 μg mL^{-1}.

Keywords: Rapid analysis, UV filters; parabens; FFD; BBD; Second-order calibration

1. Introduction

Monitoring some chemical compounds such as sunscreen agents and parabens in cosmetic products is necessary to satisfy the requirements of legislative frameworks and directives as they are prone to pose a threat to human. The bioaccumulation of sunscreen agents is considerable due to their high hydrophobicity (log K_{ow} = 5-8). Estrogenic activity is a common unwanted effect of these compounds [1]. Benzophenone-3 (BZ3), 4-methylbenzylidene camphor (MBC), octocrylene (OCR), ethylhexyl dimethyl PABA (EDP), butyl methoxydibenzoylmethane (BDM) are the UV filters and methyl paraben (MP) and propyl paraben (PP) are the parabens that were investigated in this study.

Alternating trilinear decomposition (ATLD) and multivariate curve resolution-alternating least squares (MCR-ALS) algorithms were applied for a fast analysis of the mentioned analytes in 31 miscellaneous sunscreen samples.

2. Results and discussion

2.1. Optimization of HPLC separation factors

In order to select suitable chromatographic conditions, volume fractions of the mobile phase (MeOH, ACN and water) and flow rate were optimized by full factorial design (FFD) as a screening procedure and subsequently Box-Behnken design (BBD) as an optimization strategy, keeping in mind that reducing the analysis time is a goal of the

utmost importance. The optimized condition was as mobile-phase composition of acetonitrile:H_2O:methanol (55:20:25, v/v) and a flow rate of 1.5 mL min^{-1}.

2.2. Preparation of calibration and validation samples

A set of 15 calibration samples were prepared considering the linear range of each analyte. Chromatographic profiles in multiple wavelengths belong to one of the calibration and validation samples is shown in Fig. 1a and Fig 1b, respectively. Three validation samples with different complexity were provided and spiked at three spiked levels. All samples were analysed in triplicates.

2.3. MCR/ALS and ATLD analysis of spiked sunscreen products samples

Nine global matrices and multiway arrays containing each of the validation samples and external calibration set were constructed. Considering the predefined number of components and initial estimates, the modeling process were accomplished by MCR/ALS and ATLD methods. The recovered spectral and successive chromatographic profiles for MP of one typical validation sample is depicted in Fig. 1c and Fig. 1d, respectively. The relative recovery and RSD values for one spike level of three validation samples is shown in Table 1. The obtained analytical merits obtained through both modeling consisting of mean recovery and RSD values confirmed the superiority of MCR/ALS relative to ATLD method.

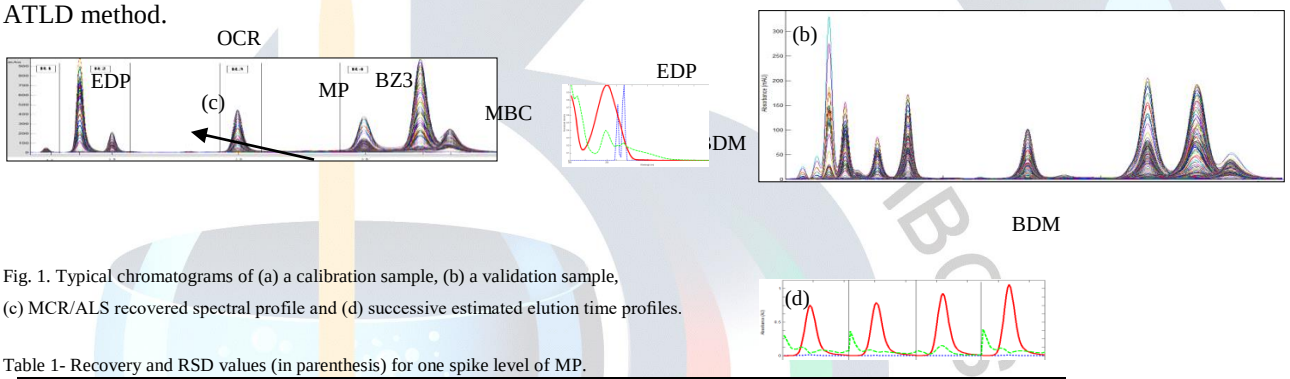


Fig. 1. Typical chromatograms of (a) a calibration sample, (b) a validation sample, (c) MCR/ALS recovered spectral profile and (d) successive estimated elution time profiles.

Table 1- Recovery and RSD values (in parenthesis) for one spike level of MP.

Analytes		MP	PP	BZ3	MBC	OCR	EDP	BDM
Spike Level 1 (μg mL^{-1})		30	30	75	75	75	75	75
Method	ATLD	79.1 (6.5)	79.3 (2.9)	87.2 (6.1)	80.7 (8.5)	83.6 (4.9)	104.8 (1.5)	95.0 (4.4)
	MCR-ALS	84.2 (6.4)	94.3 (3.4)	97.4 (7.1)	94.0 (8.5)	100.3 (5.6)	108.4 (3.0)	97.2 (5.3)

2.3. Conclusion

In the present work, simultaneous determination of UV filters and parabens in different sunscreen products in the presence of overlapped profiles was investigated with the aid of ATLD and MCR-ALS methods. A simple sample preparation was utilized to extract the analytes and no further clean-up was necessary. In spite of coelution of the analytes in the optimized HPLC method, an accurate quantitation was achieved by both methods with the superiority of MCR-ALS. The proposed method showed outstanding merits in accordance with green analytical methods.

References

[1] Salvador A, Chisvert A. An environmentally friendly ("green") reversed-phase liquid chromatography method for UV filters determination in cosmetics. Anal. Chim. Acta 2005; **537**:15–24.

Multi-element analysis of hair samples by ICP-MS and classification using chemometrics methods

Narjes Sadri, Somayeh Veyseh, Ali Niazi *

Department of Chemistry, Faculty of Science, Arak Branch, Islamic Azad University, Arak, Iran

E-mail address (corresponding author): ali.niazi@gmail.com & a-niazi@iau-arak.ac.ir

Abstract

The proposed methodology was applied to 50 donors aged 25–35 years old from various geographical areas of the (Arak, Ashtian, Tafresh, Saveh, Delijan, Shazand, Farmahin, Mahalat, Khomein) in Markazi province of Iran which were characterized by differing environmental conditions and dietary habits. 25 trace elements (Ag, Al, B, Ba, Bi, Ce, Cu, Fe, Li, Mn, Nb, Ni, Pb, Sb, Sc, Sn, Sr, Ta, Th, Ti, Tl, U, W, Zn and Zr) were determined by inductively coupled plasma techniques (ICP) with mass spectrometry (MS) detections. Analytical steps were deeply discussed including pre-treatment (washing) and acid treatment (microwave system). After Wilcoxon-Mann-Whitney test and correlation analysis, 6 useful elements (Bi, Cu, Fe, Li, Ni and Zn) and 4 toxic elements (Pb, Sb, Sn and Sr) were found to be evidently different between hair samples in Markazi Region. Based on the data set of these elemental variables in 50 hair sample, classification of origins was successfully achieved utilizing principal component analysis (PCA) and K-nearest neighbor (KNN) clustering.

Key words: Hair sample, ICP-MS, Principal component analysis, K-nearest neighbor clustering, Classification.

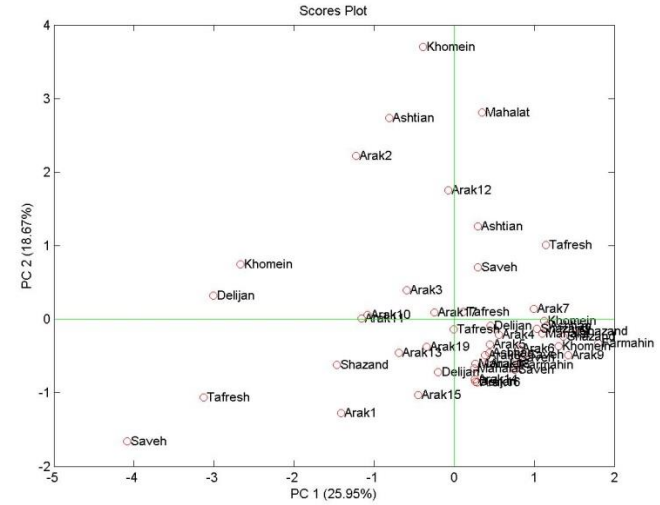
Introduction

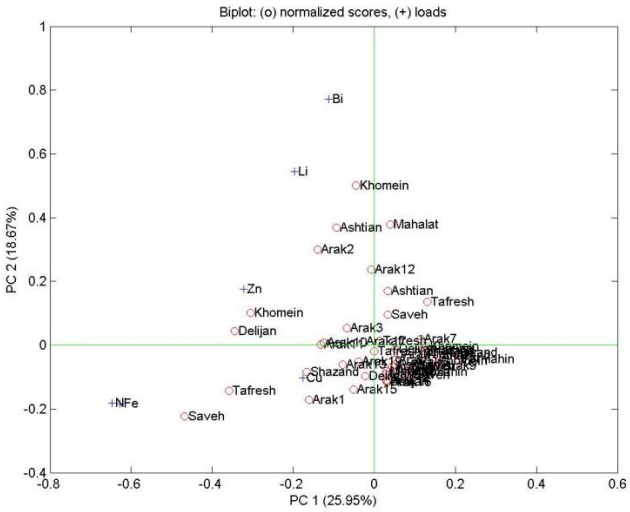
Concern about the effects of environmental exposure to tracemetals on human health has driven the scientific community to find reliable tools and methods for assessing the impact of emissions of toxic metals from anthropogenic activities. Biological monitoring has been extensively employed with this end in view, and blood, urine, feces, hair and nails are the biological materialsmost frequently analyzed to ascertain the levels of many metals [1]. According to this, human hair can provide a permanent record of trace and toxic elements associated with normal and abnormal metabolism as well as their assimilation from the environment than other biological matrices as blood and urine. In addition, hair is easily collected, conveniently stored, and easily treated.

Results and discussion

The elemental concentrations measured in human hair of the Markazi province population are collected in data matrix. Apart from correlation analysis, PCA and KNN were also performed using PLS-Toolbox. A cross validation setup was defined, which is a more time consuming validation method than leverage correction, but the estimate of the residual variance is more reliable. This residual variance decreased as the number of PCs in the model increased. Two different plots were obtained from PCA: a) scores which show the projected locations of the objects onto the PCs and define very useful sample patterns; and b) loadings that define the variable correlations. In below sample Figures shows

the score and biplot plots obtained from the experimental data ((Bi, Cu, Fe, Li, Ni and Zn). The classification of the subject groups has been discussed using the trace metal contents as discriminant variables.





The results of the present study might be useful in elaboration of reference values for the concentration of elements in human hair of the Markazi province with the consideration of age, gender, smoking and diet habits. The key to obtaining reliable, accurate, and precise hair element analysis is to ensure that the laboratory employs standardized and documented procedures such as those described in this article.

References

[1] Jing Ch. Hou J. SVM and PCA based fault classification approaches for complicated inductrial process. *Neurocomputing.* 2015; **167**: 636–642.

[2] Tie C, Hu T, Guo B, Zhang J. Novel strategy for herbal species classification based on UPLC-HRMS oligosaccharide profiling. *J. Pharm. Biomed. Anal.* 2015; **111**: 14-20.

Principal component analysis and K-nearest neighbor clustering for classification of soil samples via ICP-MS in Maygan salt lake

Mahbobeh Rabiee, Somayeh Veyseh, Ali Niazi *

Department of Chemistry, Faculty of Science, Arak Branch, Islamic Azad University, Arak, Iran

E-mail address (corresponding author): ali.niazi@gmail.com & a-niazi@iau-arak.ac.ir

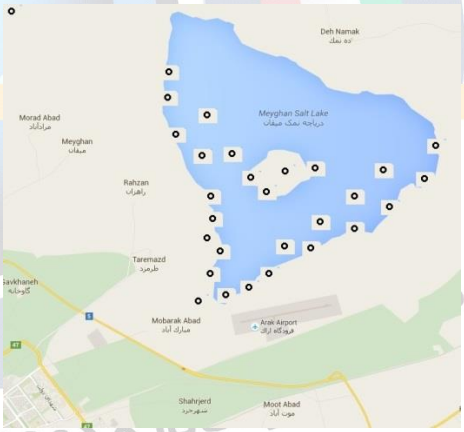
Abstract

Forty-Four inorganic elements (Ba, Be, Bi, Cd, Ce, Co, Cs, Cu, Dy, Er, Eu, Ga, Gd, Hf, Ho, In, La, Lu, Mn, Mo, Nb, Nd, Ni, Pb, Pr, Rb, Rh, Ru, Sb, Sc, Sm, Sn, Sr, Tb, Th, Ti, Tl, Tm, U, W, Y, Yb, Zn and Zr) in soil smaples from Maygan salt lake were determined using ICP-MS. After Wilcoxon-Mann-Whitney test and correlation analysis, 16 elements (Bi, Cd, Co, Cu, Fe, Li, Mn, Ni, Pb, Sb, Sn, Sr, Th, U, Zn and Zr) were found to be evidently different between soil samples in different region of Maygan salt lake. In this study, PCA and KNN are used for classification of soil samples. The results of this work suggest that PCA+KNN are a promising method to be considered for classification of soil samples datasets. The two regions were discriminated with 100% accuracy using 16 of these elements.

Key words: Soil sample, Maygan salt lake, ICP-MS, PCA, KNN, Classification.

Introduction

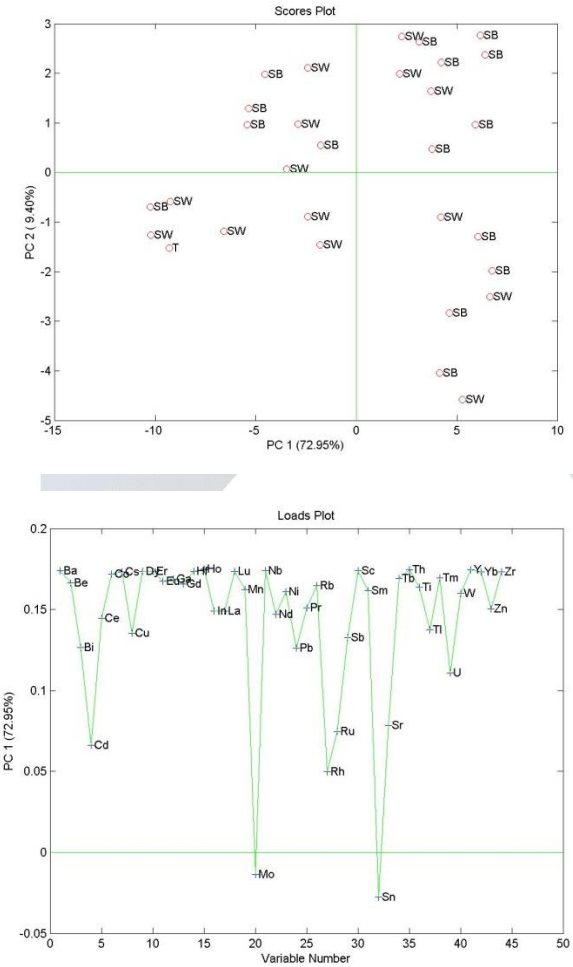
In the current work, intraregional classification of soil samples from one region (Naygan salt lake), was investigated, in order to establish the limits and reliability of the application in the relatively small geographically area of a single soil district. The distribuation of the soil samples are shown in below Figure.



Results and discussion

The elemental concentrations measured in soil samples of Maygan salt lake are collected in data matrix. Apart from correlation analysis, PCA and KNN were also performed using PLS-Toolbox. A cross validation setup was defined, which is a more time consuming validation method than leverage correction, but the estimate of the residual

variance is more reliable. Two different plots were obtained from PCA (Score and Loading plots). In below sample Figures shows the score and loading plots obtained from the experimental data (Bi, Cd, Co, Cu, Fe, Li, Mn, Ni, Pb, Sb, Sn, Sr, Th, U, Zn and Zr).



The classification of the subject groups has been discussed using the trace metal contents as discriminant variables. The results of the present study might be useful in elaboration of reference values for the concentration of elements in soil samples of Maygan salt lake.

References

[1] Coetzee P.P, Jaarsveld F.P, Vanhaecke F. Intraregional classification of wine via ICP-MS elemental fingerprint. *Food Chem.* 2014; **164**: 485-492.

[2] Tie C, Hu T, Guo B, Zhang J. Novel strategy for herbal species classification based on UPLC-HRMS oligosaccharide profiling. *J. Pharm. Biomed. Anal.* 2015; **111**: 14-20.

A method for rapid screening of interactions of pharmacologically active compounds with bovine serum albumin by chemometrics methods

Firouzeh Manouchehri[1,2], Yahya Izadmanesh[1], Jahan B. Ghasemi*[,1,2]

[1] *Faculty of Chemistry, K.N. Toosi University of Technology, Tehran, Iran*

Department of Chemistry, Islamic Azad University, Central Tehran Branch (IAUCTB), Tehran, Iran

School of Chemistry, University College of Science, University of Tehran, P.O. Box 14155-6455, Tehran, Iran.

E-mail address: jahan.ghasemi@gmail.com

Abstract

In this research, the binding interactions of glycyrrhizin (GL) and glycyrrhetinic Acid (GA) with bovine serum albumin (BSA) have been investigated by the novel method of spectrophotometric-gradient flow injection technique. GL and GA exist in the aqueous extract of licorice root. A hard-modeling multivariate approach was used for calculation of binding constants and estimation of concentration-spectral profiles. The Stability Quotients from Absorbance Data (SQUAD) computer program was used for fitting the predefined complexation model to the spectral mole ratio data. The proper selection of the chemical model was verified by the determination of the number of absorbing species by using a singular value decomposition of each data set. The binding constants obtained for GL and GA with BSA are 4.58, 4.48, respectively.

Key words: Glycyrrhizin, Glycyrrhetinic Acid, Bovine Serum Albumin, Flow injection, SQUAD

Introduction

Glycyrrhizin is a well-known pharmacologically bioactive natural glycoside. Glycyrrhizin (GL) has been widely used as a therapeutic agent for chronic active liver diseases. Glycyrrhetinic acid (GA) is an aglycone and an active metabolite of Glycyrrhizin. Serum albumin is the most abundant protein present in the circulatory system of a wide variety of organisms and is the major macromolecule contributing to the osmotic blood pressure [1]. The most important property of this group of proteins is that they serve as transporters for a variety of compounds.

Specific noncovalent [2] binding of micromolecules to proteins is of great importance in pharmacology, biopharmaceutics, biochemistry, immunochemistry, and related fields. Flow injection analysis (FIA) is based on an injection of a sample solution into a flowing carrier, where controllable dispersion and generation of a reproducible signal at a detector occur [3-4]. In the present work, we wish to report the FIA in conjunction with a photodiodearray UV–Vis spectrophotometer and a whole-domain spectral processing computer program, SQUAD, in the study of the binding interactions GL and GA with BSA.

Results and discussion

In order to calculate the binding constants and estimate concentration- spectral profiles of protein-drug interaction; it is needed to define a exact chemical model of interaction. By assuming the 1:1 stoichiometry for protein: drug interaction, the mass balance equations and chemical model of interaction in any point along the FIA reaction tube can be written in the following way:

$$P_t + D_t \leftrightarrow PD_t \qquad (1) \qquad\qquad \beta = \frac{[PD_t]}{[P_t][D_t]} \qquad (2)$$

where, P_t, D_t, and PD_t are free protein (BSA), free drug and drug attached to BSA concentrations. β Indicates stability constant. The following mass balance equations can be written for the above equilibrium:

$$C_{t,D} = [D]_t + [PD]_t \qquad (3)$$

$$C_{t,P} = [P]_t + [PD]_t \qquad (4)$$

$C_{t,D}$ and $C_{t,P}$ are total concentrations (free and attached forms) of the drug and BSA concentrations. The combination of Eqs. (2) - (4) as a function of the free drug concentration $[D]_t$ yields the quadratic polynomial:

$$\beta[P]_t{}^2 + [P]_t\left(1 + \beta C_{t,D} - \beta C_{t,P}\right) - C_{t,P} = 0 \qquad (5)$$

$C_{t,P}$ is constant in any point along the FIA tube and is equal to the concentration of the BSA solution used as carrier. $C_{t,D}$ has a concentration gradient along the tube and is calculated by calibration of dispersion. By initial estimation of β and having spectral data, the concentration-spectral data and final value of β can be obtained. The β obtained for GL and GA are 4.58, 4.48, respectively.

References

[1] Carter D C, Ho J X. Structure of serum albumin. *Adv. Prot. Chem.* 1994; **45**; 153–203.

[2] *Sjöholm I, Ekman B, Kober* A, *Ljungstedt-Påhlman* I, *Seiving B, Sjödin T. Binding of drugs to human serum albumin: XI. The specificity of three binding sites as studied with albumin immobilized in microparticles*. Molecular Pharmacology. 1979; **16**; 767-777.

[3] Hansen E.H., Miro M., *Miniaturization of environmental chemical assays in flowing systems: The Lab-on-a-Valve approach vis-à-vis Lab-on-a-Chip microfluidic devices.* Trends Anal. Chem. 2007; **26**; 18-26

[4] Izadmanesh Y., Ghasemi J.B., *A rapid, automated gradient flow injection–spectrophotometric technique for study of metal complexation reactions*. Talanta, 2014; **128**; 511-517.

Application of vortex − assisted dispersive liquid–liquid microextraction coupled with spectrofluorimetry for determination of aspirin in human urine by response surface methodology

Zahra Tamiji, Ali Niazi *

Department of Chemistry, Faculty of Science, Arak Branch, Islamic Azad University, Arak, Iran

*E-mail: a-niazi@iau-arak.ac.ir & ali.niazi@gmail.com

Abstract

In this work an effective method based on vortex − assisted dispersive liquid – liquid microextraction (VA − DLLME) developed for extraction of aspirin (acetyl salicylic acid) prior to its determination by spectrofluorimetry in human urine. The factors affecting the extraction efficiency, such as volume of extraction solvent (chloroform), volume of disperser solvent (acetonitrile), extraction time, centrifuging time, sample pH and ionic strength, were investigated. Then significant variables were optimized by the response surface method using the Box–Behnken design. Under the optimum extraction conditions, a linear calibration curve in the range of 0.1 to 130 ng mL^{-1} with a correlation coefficient of r^2 = 0.998 was obtained. The limits of detection (LOD) and limit of quantification (LOQ) were 0.031 and 0.103 ng mL^{-1}, respectively. The relative standard deviations (RSD) were less than 4%. Enrichment factor and recoveries were obtained for the extraction of aspirin in human urine. The proposed method gives a very rapid, simple, sensitive and low–cost procedure for the determination of aspirin.

Key words: Spectrofluorimetry, Dispersive liquid – liquid microextraction, Box–Behnken design.

Introduction

Acetylsalicylic acid (ASA), known as aspirin, is a non-steroidal anti-inflammatory drug, which therapeutic action happens both at locally and peripheral inflammatory sites by inhibiting cyclooxygenase enzymes [1]. The DLLME method was developed in 2006 by Rezaee et al. [2]. It is generally based on a ternary component solvent system, in which extraction and disperser solvents are rapidly introduced into the aqueous sample to form a cloudy solution[3]. Box–Behnken is a rotatable second order design based on three level incomplete factorial design. It is a special 3level design because it does not contain any points at the vertices of the experiment region [4].

Results and discussion

In this work, for the first time, a new combined methodology of VA − DLLME , box – behnken design and spectrofluorimetric determination of aspirin has been developed. The excitation and emission spectra of aspirin are at 276±3 and 375±3 nm, respectively. The response function (Y), representing fluorescence intensity (Y)is a function of pH (X_1), Volume extraction solvent ($CHCl_3$) (X_2), Volume of disperser solvent (ethanol) (X_3) and Extraction time (X_4). Mathematical equation representing the fluorescence intensity is given in Eq. (1).

$$\text{Fluorescence intensity (Y)} = 286.00 - 35.33\,X_1 - 14.25\,X_2 - 34.42\,X_3 - 28.17\,X_4 - 28.00\,X_1^2 - 11.62\,X_2^2 - 9.87\,X_3^2 - 37.50\,X_4^2 + 16.75\,X_1X_2 + 28.00\,X_1X_3 - 12.75\,X_1X_4 + 13.25\,X_2X_3 - 19.75\,X_2X_4 + 23\,X_3X_4 \quad (1)$$

The adequacy of the developed model was analyzed by using the ANOVA technique. The ANOVA test results are presented in Table 1. An ''F − value'' equal to 47.18 shows that the model is significant. The ''lack of fit (LOF) F-value'' of 0.901 is not significant relative to the pure error and confirms the validity of the model.

Table 1
ANOVA for response surface quadratic model.

Source	Degree of freedom	Sum of Squares	Mean square	F value	p-Value
Model	14	59730.7	4266.5	47.18	0.000
X_1	1	14981.3	14981.3	165.65	0.000
X_2	1	2436.8	2436.8	26.94	0.000
X_3	1	14214.1	14214.1	157.17	0.000
X_4	1	9520.3	9520.3	105.27	0.000
X_1X_2	1	1122.2	1122.2	12.41	0.004
X_1X_3	1	3136.0	3136.0	34.68	0.000
X_1X_4	1	650.2	650.2	7.19	0.020
X_2X_3	1	702.2	702.2	7.77	0.016
X_2X_4	1	1560.3	1560.3	17.25	0.001
X_3X_4	1	2116.0	2116.0	23.40	0.000
Residual	12	1085.3	90.4		
Lack-of-Fit	10	683.3	68.3	0.34	0.901
Pure Error	2	402.0	201.0		

The accuracy of the present method was evaluated by determination of ASA in spiked human urine.The obtained results are shown in Table 2. As can be seen, calculated amounts of recoveries varied between 93.2-97.5% for human urine, indicating both accuracy and precision.

Table 2
Results of recoveries of spiked urine samples

Sample	ASA added (ng ml^{-1})	*ASA found (ng ml^{-1})	Recovery (%)
Human urine	5	4.79±0.16	95.8
	10	9.75±0.28	97.5
	100	93.22±3.15	93.2

References

[1] W. Wesemann, K. H. Sontag, and J. Maj, "Pharmacodynamics and pharmacokinetics of nonsteroidal anti-inflammatory drugs in species of veterinary interest," *Arzneimittelforschung.*, vol. 33, no. 8, pp. 1122–1134, 1983.

[2] M. Rezaee, Y. Assadi, M.-R. M. Hosseini, E. Aghaee, F. Ahmadi, and S. Berijani, "Determination of organic compounds in water using dispersive liquid–liquid microextraction," *J. Chromatogr. A*, vol. 1116, no. 1, pp. 1–9, 2006.

[3] D. Han and K. H. Row, "Trends in liquid-phase microextraction, and its application to environmental and biological samples," *Microchim. Acta*, vol. 176, no. 1–2, pp. 1–22, 2012.

[4] G. E. P. Box, W. G. Hunter, and J. S. Hunter, "Statistics for experimenters," 1978.

Box-behnken experimental design based optimization technique in headspace-solid-phase microextraction–gas chromatography/mass spectrometry for simultaneous determination of BTEX in Urine.

Mahdi Rahmani, Ali Niazi*, Mahmoudreza Sohrabi

Department of Chemistry, Faculty of Sciences, Arak Branch, Islamic Azad University, Arak, Iran.

E-mail: a-niazi@iau-arakac.ir & ali.niazi@gmail.com

Abstract

The paper shows Box-behnken experimental design (BBD) as a usefull technique to optimize an extraction reaction and visualize the effect that some experimental factors exert on several analytical responses. A five-factor, three-level BBD was used to optimizing. The entire proposed procedure has been applied in the optimization of extraction temperature, extraction time, desorb time, amount of NaCl and volatile organic compound (VOC) as internal standard (IS) in drog free Urine. Forty six batches were prepared and evaluated for responses as dependent variables. Mathematical equations and response surface plots used to relate the dependent and independent variables. The statistical validity of the polynomials was established, and optimized formulation factors were selected by feasibility. The use of BBD approach helped in identifying the critical formulation parameters in the head space micro extraction–gas chromatography/mass spectrometry (HS-SPME-GC/MS). In the optimal experimental conditions (0.32 gr NaCl,2.67 microliter VOC, 40 ◦C extraction temperature, 15 min extraction time and 60 second desorb time) the determination by means of a HS-SPME–GC/MS system is carried out. Validation of the optimization study with 46 confirmatory runs indicated high degree of prognostic ability of response surface methodology.

Keywords: Box–Behnken; optimization; HS-SPME-GC/MS; BTEX.

Introduction

BTEX (monocyclic aromatic chemicals), are mainly used as additives in gasolineto enhance octanerating, inreplacement of alkyl-lead compounds, and to improve the combustion process. And also present in main- and side-stream tobacco smoke, which provides one of the largests ources of personal exposure to air pollutants. Meanwhile, toluene, xylene and ethylben-zene are neurotoxic. Benzene is hematotoxic and has been classified as a known carcinogen to humans by the International Agency for Research on Cancer (group 1)[1]*. Solid-phase microextraction (SPME) is a comparatively new method usable for aromatic compounds extraction. It is a very easily and efficient, solvent-free sample preparation method for analysis of organic compounds. A coated silica fiber is directly injected into the sample or into the headspace above the sample (HSSPME). After ward organic analytes are adsorbed to the respective fiber coating. During the subsequent gas chromatography, the analytes are desorbed by the heated injection system and reach the GC column (HSSPME/GCMS). Box–Behnken experimental design (BBDs) has been used as a tool for multivariate optimization also are a class of rotatable or nearly rotatable second-order designs that are based on three-level incomplete factorial designs*[2]

Result and discussion

A five-factor three-level Box–Behnken design for response surface methodology was applied to study the combined effect of independent variables on dependent variables over three levels which resulted in 46 experimental runs in the present study. Factors as follows: NaCl (mg), VOC (micro liter), extraction time (min), extraction temperature(c) and desorbtion time (sec.) In optimizing condition of HSSPME (Table 1), calibration regession graph was made(Table 2) and real samples were test (Table 3).

Table1. Variables and *Optimized* Responses in Box–Behnken Design for BTEX *HSSPME*.

Variables	*NaCl(gr)*	*VOC(microliter)*	*Extraction time(min)*	*Extraction temp. (C)*	*Desorbtion time(sec.)*
Optimizing condition	*0.32 gr*	*2.76 microliter*	*15.1 min*	*40.6 C*	*60.4 sec*

Table2. Summary of Results of Regression Analysis for Responses.

Peak	*RT*	*0-1*	*0-2*	*0-3*	*SD*	*LOD*	*LOQ*	*Slope*
Benzene	*2.113*	*16084*	*17821*	*15022*	*1413*	*0.045*	*0.149*	*93685*
Toluene	*4.182*	*80390*	*77321*	*78541*	*1545.21*	*0.036*	*0.119*	*128375*
Ethyl benzene	*8.509*	*59615*	*51201*	*50210*	*5167.71*	*0.053*	*0.176*	*290638*
m,p-xylene	*8.935*	*180246*	*174302*	*197320*	*11949.07*	*0.085*	*0.281*	*421062*
o-xylene	*10.097*	*167532*	*152301*	*149321*	*9768.18*	*0.087*	*0.287*	*336605*

Table3. Real sample analyze.

Peak	*RT*	*Area Repeat 1th*	*Area Repeat 2th*	*Area Repeat 3th*	*SD*	*Mean*	*RSD*
Benzene	*2.113*	*684542*	*658221*	*675573*	*13381*	*672779*	*1.99*
Toluene	*4.182*	*842228*	*834969*	*887418*	*28419*	*858472*	*3.32*
Ethyl benzene	*8.509*	*1883082*	*2016157*	*1986302*	*69827*	*1961847*	*3.56*
m,p-xylene	*8.935*	*4442709*	*4841336*	*4832310*	*227587*	*4705452*	*4.84*
o-xylene	*10.097*	*3942943*	*4249021*	*4151533*	*156366*	*4114501*	*3.8*

Optimization of HSSPME/GCMSS is a complex process that requires one to consider a large number of variables and their interactions with each other. The present study conclusively demonstrates the use of a Box–Behnken statistical design is valid for optimizing BTEX concentration with HSSPME/GCMSS . The derived polynomial equations and contour plots aid in predicting the values of selected independent variables for preparation of optimum extraction with desired properties.

References

[1] S. Fustinoni, F. Rossella, L. Campo, R. Mercadante, and P. A. Bertazzi, “Urinary BTEX, MTBE and naphthalene as biomarkers to gain environmental exposure profiles of the general population.,” *The Science of the total environment*, vol. 408, no. 14, pp. 2840–9, Jun. 2010.

[2] M. J. Chaichi and M. Ehsani, “Determination of Glucose and Cholesterol Using a Novel Chemiluminescence Method by Box – Behnken Design,” 2015.

Box-behnken experimental design (BBD) for optimization instrumental parameters effects for powder coating particle size.

Mahdi Rahmani * , Alireza Sabouri Azad

Department of Chemistry, Faculty of Sciences, Arak Branch, Islamic Azad University, Arak, Iran.

Kian Rangi co.

E-mail: mahdirahmani_mp@yahoo.com

Abstract

In this study using a statistical Box-behnken experimental design (BBD) as a means of adequately optimizing the various parameters to evaluate the main and interaction variables that affect particle size, a three-factor three-level BBD was employed to schedule and perform the experiments. The effects of variable instrumental parameters, such as the classifier rate, the rotor rate and air flowing rate, were investigated systematically. In the optimal experimental conditions (0.34% air flow rate, 1062.62 r/m classifier rate and 2595.96 r/m rotor rate) the determination by means of a particle size analyzer is carried out. Validation of the optimization study with 15 confirmatory runs indicated high degree of prognostic ability of response surface methodology.

Keywords: Box–Behnken; optimization, particle size analyzer.

Introduction

Powder coatings are made by mixing polymer with charge agents and other additives in a melt. This is then extruded to form chips or pellets that are subsequently ground to the required particle size in a series of steps involving milling or grinding and separation/ classification. The maximum particle size is determined by the sharpness of the classification inside the impact classifier mill. Influence parameters on such an undefined top size are dimensions of the housing, speed of air and product inside the impact classifier mill, type of classifier wheel and design of the outlet. The width of the particle size distribution also depends on the design and the efficiency of the grinding system Important factors include the classifier rate (rpm), rotor rat(rpm) and airflow conditions. The optimum conditions for optimizing particle size were determined by means of a Box–Behnken experimental design (BBD) combining with response surface modeling (RSM) and quadratic programming. RSM is a statistical and graphical technique for developing, improving and optimizing process which can overcome the following disadvantages: the classical one factor-one-time design method in a time-consuming process; unrealistic number of experiments and difficulties in determination of optimal conditions. Box-Behnken design is an especially efficient response surface method for obtaining mathematical models[1] This design was suitable for constructing a second-order polynomial model. Only a small number of experimental runs (i.e.15 runs) was necessary for the optimization process. This design consisted of replicated center points and multicubes at each edge of the center of surface composition.

Result and discussion

This method may neglect important factors, which might lead to loss of the supreme values of these factors and of the final results. Therefore, we used the response surface-Box–Behnken method to further optimize the three most important factors in the medium.

A three-factor three-level Box–Behnken design for response surface methodology was applied to study the combined effect of independent variables on dependent variables over three levels which resulted in 46 experimental runs in the present study. The three factors examined were classifier rate, rotor rate and air flow. The three-dimensional response surface and the two-dimensional contour plots are shown in Figs. 1. These plots are graphical representations of the regression equation; they are plotted to show the interaction of the variables and to identify the optimum level of each variable for a maximum response. Each response surface plot for D50 production represents different combinations of two test variables at once.

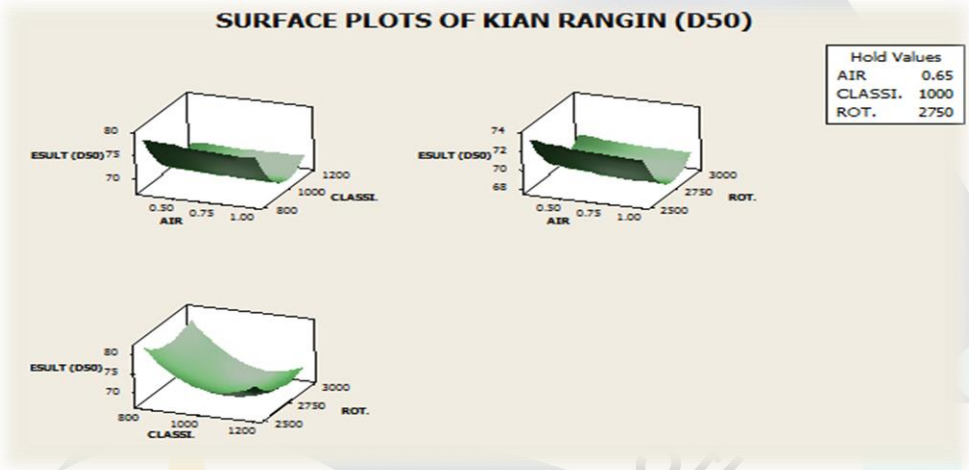


Fig1. Response surface plots of particle size of powder coating, showing the interaction of classifier, rotor and air flow**.**

As a result, significant effects of classifier(c) and rotor(r) and air flow(a) rate. A second- order regression formulation was derived by ANOVA as the following regression Eq (1),

$$Y=-0.278\ c + -0.257\ r- 10.076a +1.418aa+ 0.004\ ca+0.001ra \qquad (1)$$

Linear, quadratic, and linear interaction effects were calculated for the optimization process, and significant coefficients were preserved to create the corresponding response surfaces. According to the results obtained from the 15 experimental runs, the various independent factors resulted in particle size distribute of powder coating sample varying from 58 to 80micron and optimal experimental conditions are 0.34% air flow rate, 1062.62 r/m classifier rate and 2595.96 r/m rotor rate.

References

*[1] Box G, Behnken DW, Some new three level designs for the study of quantitative variables. Technometrics, 1960; **2**:455–475.*

Chemometrics-assisted of ultrasonic emulsification–microextraction for simultaneous spectrophotometric determination of Cobalt, Nickel and Copper

Zohreh Doroudi[1], Ali Niazi[2]*, Neda khorshidi[2]

[1]*Department of Chemistry, Faculty of Science, Yadegar-e-Imam Khomeini Branch, Islamic Azad University, Tehran, Iran*
[2]*Department of Chemistry, Faculty of Science, Arak Branch, Islamic Azad University, Arak, Iran*

E-mail address: a-niazi@iau-arak.ac.ir & ali.niazi@gmail.com

Abstract

A sensitive and simple method for rapid extraction, preconcentration and determination of cobalt, nickel and copper as their 1-(2-pyridilazo)-2-naphthol (PAN) complexes based on ultrasound-assisted emulsification–microextraction (USAEME) and multivariate calibration of spectrophotometric data is presented. Various parameters affecting the extraction efficiency were optimized by Box–Behnken design. The resolution of ternary mixtures of these metallic ions was accomplished by using partial least-squares regression (PLS) and orthogonal signal correction- partial least-squares regression (OSC-PLS). Under the optimized condition, limit of detection was found 0.15 ng mL^{-1} (Co^{2+}), 0.16 ng mL^{-1} (Ni^{2+}), and 0.19 ng mL^{-1} (Cu^{2+}) and the relative standard deviation was <3.75%. The method was successfully applied to the simultaneous determination of Co^{2+}, $Ni^{2+}$and Cu^{2+} in water samples.

Key words: USAEME, Box–Behnken, OSC, PAN, Water sample

Introduction

Water pollution is a worldwide problem affecting developing and developed countries alike. Heavy metal contaminants are one prevalent type of water pollutant. They are persistent in the environment once discharged and removal from source waters is necessary to ensure a clean drinking water supply. Among those cobalt, nickel and copper are metals, which appear together in many real samples. So, it is very important to determine their concentrations. Sample preparation represents a major challenge and a very important step in the development and application of an analytical method [1]. This study consists of two parts. The first part is the preconcentration and optimization by experimental design and second section focuses on the multivariate calibratin to relate the concentrations of the metal ions to the spectral properties.

Results and discussion

Various parameters that affected the extraction efficiency such as types and volumes of extractant, ultrasonic time, pH of the solution, ionic strength, concentration of ligand, centrifugation speed and time [2]. An optimization procedure was applied in order to find out the exact values of the most important factors. The design consisted of 27 sets of experiments carried out according to the Box–Behnken [3].

The first step in simultaneous determination of Co^{2+}, Ni^{2+} and Cu^{2+} in mixtures by multivariate methods involved constructing the calibration matrix. It is usual to use multivariate calibration, of which the partial least squares (PLS) algorithm is the most widespread. In order to select the mixtures for prediction set, their compositions were randomly designed. For the evaluation of the predictive ability of a different model, the root mean square error of prediction (RMSEP) and relative standard error of prediction (RSEP) can be used. OSC is a preprocessing method for PLS regression to find variation in a descriptor matrix that is unrelated to that in a response matrix and which therefore can be extracted and analyzed separately [4]. The results show that score plots have better results when OSC-PLS is used. To validate the applicability of the method for simultaneous extraction and determination of Co^{2+}, Ni^{2+} and Cu^{2+} in aqueous samples, waste water, well water and tap water were collected and analyzed with the proposed method. The results were shown in Table 1. Under the optimal conditions, calibration curves were the range of 2.0–120.0 ng mL^{-1} for Ni^{2+}, 2.0–150.0 ng mL^{-1} for Co^{2+} and 2.0–150.0 ng mL^{-1} for Cu^{2+} in initial solution. The correlation of determination (r^2) was 0.997 for Ni^{2+}, 0.996 for Co^{2+} and 0.997 for Cu^{2+} ions. Under the optimized condition, limit of detection was found 0.15 ng mL^{-1} (Co^{2+}), 0.16 ng mL^{-1} (Ni^{2+}), and 0.19 ng mL^{-1} (Cu^{2+}) and the relative standard deviation was < 3.70%. The preconcentration factor was obtained 166.

Table 1. Determination of Cu^{2+}, Ni^{2+} and Co^{2+} in real water sample

Sample	Analyte	Added	Found (n=3)	Recovery %	RSD
Waste water	Ni^{2+}	80.0	83.3	104.2	2.93
	Cu^{2+}	80.0	84.3	105.4	3.15
	Co^{2+}	80.0	82.3	102.8	2.81
Well water	Ni^{2+}	60.0	63.2	105.4	4.05
	Cu^{2+}	40.0	38.4	96.1	2.80
	Co^{2+}	60.0	58.7	99.8	2.67
Tap water	Ni^{2+}	60.0	62.5	104.2	3.69
	Cu^{2+}	100.0	104.3	104.3	2.42
	Co^{2+}	40.0	38.1	95.2	2.78

References

[1] Hejazi L., Mohammadi D. E., Yamini Y., Brereton R. G. Solid-phase extraction and simultaneous spectrophotometric determination of trace amounts of Co, Ni and Cu using partial least squares regression. Talanta, 2004; **62** :185.

[2] Regueiroa J., Llomparta M., Psillakis E., Garcia-Monteagudoc J. C., Garcia-Jaresa C. Ultrasound-assisted emulsification-microextraction of phenolic preservatives in water. Talanta, 2009; **79** :1387.

[3] Niazi A., Khorshidi N., Ghaemmaghami P. Microwave-assisted of dispersive liquid-liquid microextraction and spectrophotometric determination of uranium after optimization based on Box-Behnken design and chemometrics methods Spectrochim. Acta, Part A, 2015; 135 :69.

[4] Wold S., Antii H., Lindgren F., Ohman J. Orthogonal signal correction of near-infrared spectra, Chemom. Intell. Lab. Syst., 1998; 44 :175.

Simultaneous spectrophotometric determination of oxytetracycline, tetracycline and doxycycline in milk and honey samples. Comparison PLS and PARAFAC

Zohreh Doroudi[1], Ali Niazi[2]*

[1] *Department of Chemistry, Faculty of Science, Yadegar-e-Imam Khomeini Branch, Islamic Azad University, Tehran, Iran*
[2] *Department of Chemistry, Faculty of Science, Arak Branch, Islamic Azad University, Arak, Iran*

E-mail address: a-niazi@iau-arak.ac.ir & ali.niazi@gmail.com

Abstract

A simple, sensitive, rapid and precise spectrophotometric method has been developed and validated for simultaneous determination oxytetracycline (OTC), tetracycline (TC) and doxycycline (DXC) in milk and honey samples. Chemometrics methods suitable for overlapping spectra and have resolved successfully the overlapping bands. Two chemometric techniques parallel factor analysis (PARAFAC) and partial least square (PLS) were applied to the determination of TCs in their ternary mixture and the proposed calibration techniques were validated by analyzing synthetic mixtures consisting of these drugs. The procedure was repeated at different pH values. The best PLS model was obtained at pH= 8.0 The PARAFAC model was applied to a three-way array constructed using all the pH data sets and anable better results. The RMSEP for DXC, OTC and TC with PLS and PARAFAC were 0.123, 0.236, 0.167 and 0.0196, 0.0480, 0.0316 respectively

Key words: PLS, PARAFAC, Tetracyclines, Simultaneous spectrophotometric, Milk, Honey

Introduction

Tetracycline (TC), Oxytetracycline (OTC) and Doxycycline (DXC) are members of the tetracycline group of broad-spectrum antibiotics [1], widely are used in human and animals. The main applications of tetracyclines in animal husbandry are for preventative treatment of bacterial infections and to increase growth rates (Elmund et al. 1971). A more powerful way to treat the ternary and more complex mixtures of organic compounds, which exhibit similar spectral characteristics, is the application of multivariate analysis methods. The PARAFAC and N-PLS regression methods are well known chemometrics tools involving factor analysis and have successfully been applied to the spectral data analysis [2-4].

The aim of this work is to develop a very simple and sensitive method for the determination of tetracycline, oxytetracycline and doxycycline based on spectrophotometric methods and chemometrics approaches.

Results and discussion

Parallel factor analysis (PARAFAC)

The main advantage of three-way multivariate calibration is that it allows concentration information of an individual component to be extracted in the presence of any number of uncalibrated constituents. In this study, we

selected the pH = 2.0, 5.0, 6.0, 7.0, 8.0 and 11.0 for three-way data. The data were arranged in a three-way array 9×281×6, composed of 9 solutions, with different TCs concentrations in the rows, 281 wavelengths in the columns and 6 pH values in the slices. An important parameter to determine is the number of PARAFAC components, which are necessary to build the data. Several methods can be used to determine this parameter. In this work, the method used is core consistency diagnostic (CORCONDIA). Three factors give a CORCONDIA value of 100% (a perfect trilinear model) whilst, when using four or more factors, this value diminishes to values below to 1%.

PLS analysis

The multivariate calibration is a powerful tool for determinations, because it extracts more information from the data and allows building more robust models. Therefore, it was decided to perform a multivariate calibration using PLS models built for each pH value individually and compare it with PARAFAC model. Result show PLS model at PH=8.0 is the best model.

Determination of TC, DXC and OTC in milk and honey

In order to show the analytical applicability of the proposed methods, first calibration curve obtained from PARAFAC and PLS model at PH=8.0 were applied to determination of for TC, DXC and OTC in real samples (milk and honey).

Table 1. Determination of TC, DXC and OTC in honey using PARAFAC and PLS-PH=8 models (μgmL^{-1})

Type of samples		Added (ppm)	Amount Found (PARAFAC)	Recovery (%)	RSD%	Amount found (PLS-PH=8.0)	Recovery (%)	RSD%
Honey Sample	TC	2.00	1.91	95.5	0.06	1.74	87.0	0.21
	DXC	2.00	1.96	98.0	0.08	1.62	81.0	0.19
	OTC	2.00	2.05	102.5	0.07	2.29	114.5	0.16
Milk Sample	TC	2.00	1.97	98.5	0.11	1.82	91.0	0.16
	DXC	2.00	2.03	101.5	0.08	2.12	106.0	0.18
	OTC	2.00	1.95	97.5	0.09	1.76	88.0	0.14

References

[1] Goodman, L.S. and Gilman, A.G. The Pharmacological Basis of Therapeutics, ninth ed. McGraw-Hill Companies Inc., New York, USA. 1996.

[2] Elmund, G. K., Morrison, S. M., Grant, D. W. and Nevins, M. P. Role of excreted chlorotetracycline in modifying the decomposition process in feedlot waste. Bulletin of Environmental Contamination and. Bulletin of Environmental Contamination and Toxicology, 1971; **6** : 129-132.

[3] Valverde R. S., Garcia M. D. G., Galera, M. M., Goicoechea, H. C. Determination of tetracyclines in surface water by partial least squares using multivariate calibration transfer to correct the effect of solid phase preconcentration in photochemically induced fluorescence signals. Anal.a Chim. Acta 2006; **562** : 85–93.

[4] Niazi A., Yazdanipour A., PLS and PARAFAC applied to determination of noscapine in biological fluids by excitation-emission matrix fluorescence. Pharm. Chem. J. 2007; **41** :170-175.

Bee algorithm as a descriptor selection in QSAR study of β-site amyloid precursor protein cleaving enzyme1 (BACE1) inhibitors

Morteza Atabati*, Leili Akhoundi

School of Chemistry, Damghan University

E-mail address: atabati@du.ac.ir

Abstract

The β-site amyloid precursor protein (APP) cleaving enzyme 1 (BACE1) is one of the most hotly pursued targets for the treatment of Alzheimer's disease. In this work QSAR study based on bee algorithm (BA) for prediction of IC_{50} has been used to estimate the IC_{50} of 36 compounds including BACE1 inhibitors. A set of 1,439 descriptors were used for each molecule in the data set. A major problem of QSAR is the high dimensionality of the descriptor space; therefore, descriptor selection is one of the most important steps. In this paper, six descriptors were selected by BA. The multiple linear regression (MLR) method with and without leave-one-out cross validation (LOOCV) were applied. The mean square error (MSE), average absolute relative deviation and correlation coefficient (r) were 0.06, 2.6% and 0.9259, respectively. With LOOCV, the mean square error (MSE), average absolute relative deviation and correlation coefficient (r) were 0.10, 3.3% and 0.8656, respectively.

Keywords: Bee Algorithm, BACE1, Alzheimer's disease

Introduction

β and γ secretases are the two enzymes critically responsible for forming β amyloid. Scientist discovered β -site amyloid precursor protein cleaving enzyme 1 (BACE1) is the rate-limiting enzyme for the formation of Aβ peptides and represent a potential disease-modifying treatmentof AD [1].

In this paper we developed a model for prediction of performance inhibitors (IC_{50}) in a series of aminooxazoline xanthene β-Site amyloid precursor protein cleaving enzyme (BACE1) inhibitors in 36 variable structures, using bee algorithm for descriptor selection [2, 3].

Results and discussion

The known experimental IC_{50} in a series of aminooxazoline xanthene β-Site amyloid precursor protein cleaving enzyme (BACE1) inhibitors values for a diverse data set consisted of 36 compounds were taken from literature [1]. 3-D structures of these compounds were optimized using HyperChem software (version 7.0) with semi empirical AM1 optimization method. After optimization, a total of 1439 descriptors were generated using Dragon software (version 3.0). Bee algorithm (BA) program was written in Matlab (Ver. 7.0.4) and then was used in this work. BA was applied for finding the best descriptors with total number of initial bee of 20000, 40 bees recruited for neighborhood search, and the number of other bees in the population is assigned to search randomly other sites and the number of iteration, were 40 and 20000, respectivly.

Quantitative structure–activity relationship (QSAR) model with multiple linear regression (MLR) method (BA-MLR) for prediction of pIC_{50} for BACE1 inhibitors was build on the basis of six descriptors were selected as independent variables.

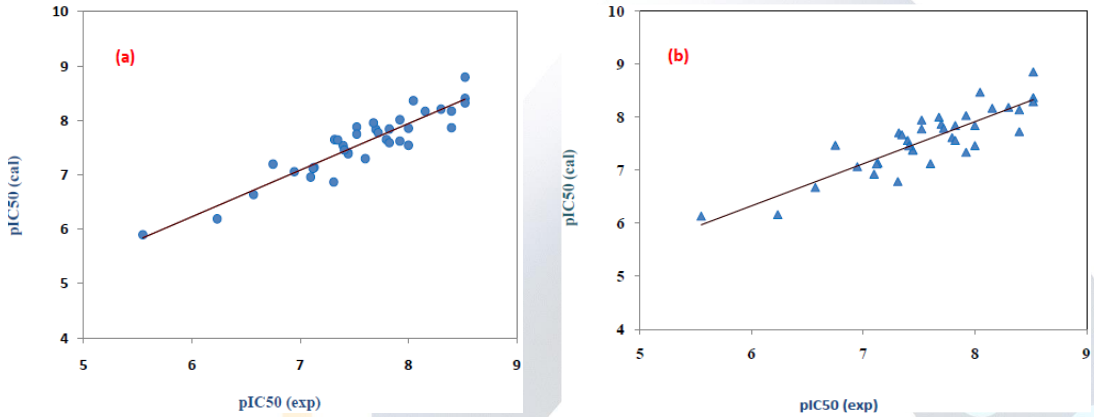


Figure 1. Relationship between the experimental and calculated pIC_{50} values of BACE1 inhibitors. a) with and b) without LOOCV.

For with and without LOOCV, the relationships between experimental and calculated values of pIC_{50} are shown in Fig. 1. The mean square error (MSE), average absolute relative deviation and correlation coefficient for this model without LOOCV were 0.06, 2.6% and 0.9259, respectively, while the mean square error (MSE), average absolute relative deviation and correlation coefficient for tis model with LOOCV were 0.10, 3.3% and 0.8656, respectively. The results show that the applied procedure was suitable for prediction of activity of BACE1 inhibitors.

References

[1] Epstein O, Bryan MC, Cheng AC, Derakhchan K, Dineen TA, Hickman D, Hua Z, Human JB, Kreiman C, Marx IE, Weiss MM, Wahl RC§ Wen PH, Whittington DA, Wood S, Zheng XM, Fremeau RT, White RD, Patel VF. Lead Optimization and Modulation of hERG Activity in a Series of aminooxazoline xanthene β-Site Amyloid precursor protein cleaving enzyme (BACE1) inhibitors. *Journal of Medicinal Chemistry*. 2014; **57**: 9796–9810.

[2] Zarei K, Atabati M, Borhani A. Ant colony optimization as a descriptor selection in QSPR modeling: Estimation of the λ_{max} of anthraquinones-based dyes. *Journal of Saudi chemical society*. 2015; In Press.

[3] Pham DT, Ghanbarzadeh A, Koç E, Otri S, Rahim S, Zaidi M. The bees algorithm-a novel tool for complex optimization problems. *In proceedings of the 2nd International virtual conference on intelligent production machines and systems*. 2006; 454–459.

Spectrophotometric determination of acidity constant of Cinnamic acid using rank annihilation factor analysis (RAFA)

Ali Benvidi*, Abbas Dadras
Department of Chemistry, Yazd University, Yazd, Iran
Corresponding author: E-mail addresses: abenvidi@yazd.ac.ir, benvidi89@gmail.com ;

Abstract

Rank annihilation factor analysis (RAFA) was a means employed in order to determine spectrophotometrically the acidity constant of Cinnamic Acid in water at 25 °C and an ionic strength of 0.1 M. A general method in rank annihilation factor analysis (RAFA) is annihilation of the contribution of one chemical component from the original data matrix. The contributions HA were annihilated from the absorbance data After recording the electronic absorbance spectra of the acid at different pH, making possible the determination of successive acidity constant. It was obtained that there was a very close and exact agreement between the resulted values by RAFA(pKa=4.44)and the declared values(pKa=4.44).

Key words: Spectrophotometric; acidity constant ; Cinnamic Acid ; RAFA

Introduction

A significant parameter for understanding and quantifying chemical phenomena reaction rates, biological activity, biological uptake, biological transport and environmental fate [1] can be dissociation constants. In order to the determination of dissociation constants, several procedures such as potentiometric titration, spectrophotometric determination and conductometry, have been reported [2,3].

Spectrophotometric is one of the methods used for the determination of acidity constants and also appropriate for determination of acidity constants in solution under diverse experimental conditions. One of the important problems is overlapping of spectra of different chemical species involved in the equilibria. Chemometric methods can easily resolve the overlapped spectra, too. RAFA was exploited to investigate the experimental data in this work[4-9].

Results and discussion

The changes in the absorbance spectra of Cinnamic acid was illustrated in Fig. 1a respectively, as a function of pH in a Water. The plotted section in Fig. 1b is related to the variation of R.S.D, as a function of hypothetical pKa values for the reaction systems of Fig.1a. As it is detected, a twodimensional plot has been obtained for Cinnamic acid. The acidity constant of Cinnamic acid in water is 4.44, based on the plot displayed in Fig.1b. Also, the same outcomes were obtained by annihilating the contribution of A^- instead of HA.

The calculated pure spectra and concentration profile of the two protogenic species of Cinnamic acid in water are represented in Fig. 1c and 1d, respectively. It was found that there is a very close and exact agreement between the resulted values by RAFA(pKa=4.44)and the declared values(pKa=4.44).

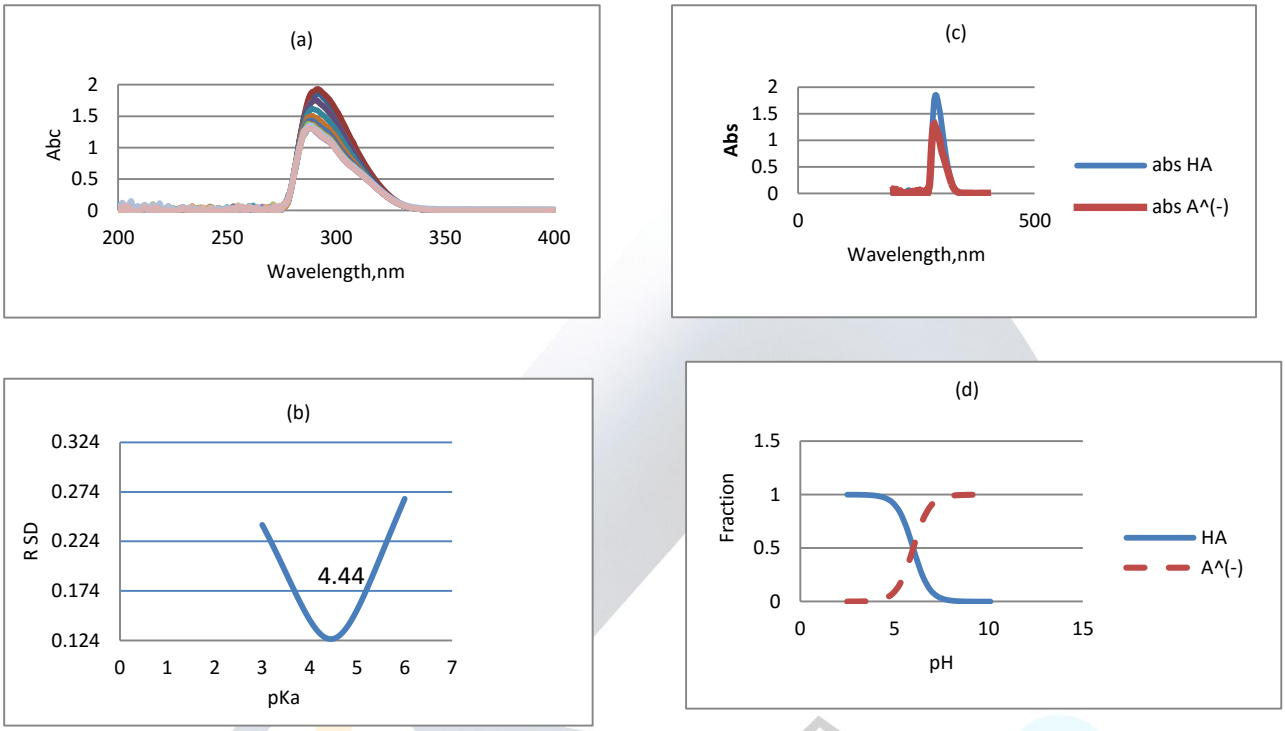


Fig. 1 – Plots of pH-metric titration absorbance spectra of Cinnamic acid in Water(a)R.S.D. (b)purespectra(c)concentrationprofiles(d)

References

[1] S.J. Gluck, K.P. Steele, M.H. Benko, J. Choromotogr. A 745 (1996)117.

[2] J.L. Beltran, R. Codony, M. Granados, A. Izquierdo, M.D. Part, Talanta40 (1993) 157.

[3] A.M. Gerakis, M.A. Koupparis, C.E. Efstathiou, J. Pharm. Biomed.Anal. 11 (1993) 33.

[4] A. Safavi, H. Abdollahi, Talanta 53 (2001) 1001.

[5] K.S. Booksh, B.R. Kowalski, Anal. Chem. 66 (1994) 782A.

[6] M. Kubista, R. Sjoback, B. Albinsson, Anal. Chem. 65 (1993) 994.

[7] S.M. Hoenigman, C.E. Evans, Anal Chem. 68 (1996) 3274.

[8] A. Afkhami, L. Khalafi, Anal. Chim. Acta 569 (2006) 267.

[9]B.Hemmateenejad, A.Abbaspour,H.Maghami ,A. Anal. Chim. Acta 2008; 607: 142-152.

QSPR Approaches to Elucidate the Stability Constants between β-Cyclodextrin and some Organic Compounds: Docking Based 3D Conformer

F. Mirrahimi[1], M. Salahinejad[2], Jahan B. Ghasemi[1, 3]

[1] *Drug Design in Silico Lab., Chemistry Faculty, K N Toosi University of Technology, Tehran, Iran*

[2] Nuclear Science and Technology Research Institute, Tehran, Iran

[3] Chemistry Faculty University of Tehran, Tehran, Iran

Abstract

In the present investigation, a combination of docking and quantitative structure-property relationship (QSPR) approaches was applied to elucidate the host-guest interactions in β-cyclodextrin complexation with diverse set of organic compounds. Molecular docking was performed to find correct conformations of organic molecules in the cavity of β-cyclodextrin. The conformation with the lowest binding free energy was chosen to calculate molecular descriptors. Some additional descriptors relevant to characterizing the structural properties of inclusion complexes were also calculated and used in QSPR model building. Genetic function approximation technique was applied to choose the best subset of descriptors. The selected descriptors explain that the hydrophobicity, surface area and shape of guest molecules play important roles in the β-cyclodextrin complexation. The final QSPR model, based on *multiple linear regression* (MLR) method, was characterized by satisfactory statistical performance; calibration (R^2_c) and prediction (R^2_p) correlation coefficient of 0.83±0.02 and 0.78±0.07 respectively.

Keywords: QSAR, Docking, Cyclodextrins, Complexation, Stability constant

1. Introduction

Cyclodextrins (CD), cyclic oligomers are a group of structurally related natural products *with a hydrophilic outer surface and a lipophilic inner/central cavity* which can encapsulate a guest molecule to form an inclusion complex [2-3-4]. The CD's are water soluble, nonreducing macrocyclic polymers containing glucose molecules linked by α- 1,4-glycosidic bonds [1]. As a consequence of these structural features, CDs have been utilized in many different fields [5]. Due to difficulty in experimental determination of stability constant of inclusion complexes, especially for low solubility of guest molecules, computational chemistry has recently been used as a tool for studying the stability of CD inclusion complexes. The objective of the present study is to build a multiple linear regression QSPR model, able to correlate and predict the complex stability constant between diverse guest molecules and β-CD. Special emphasis will be given to elucidate the driving forces and *conformational* reorganization *effects* leading to the complexation of the set of molecules under study. Hence, docking procedure was used to select the best pose for each ligand and docking descriptors are calculated based on the major interactions between ligand and cyclodextrin. A docking based QSAR

model is developed to represent a relationship between descriptors originating from docking and complexation of β-CD with different organic compounds.

2. Results and discussion

2.1. Docking procedure

AUTODOCK 4.2 [٦] with Lamarckian Genetic Algorithm (LGA) method was used to generate the inclusion complex of β-CD with the guest molecules. The optimized conformation of each inclusion complexes, with the lowest binding free energy was used to calculate molecular descriptors.

2.2. Descriptor calculation

Two kinds of descriptors were computed to consider the conformational changes in the complexation process of CD with different organic compounds. The ligand descriptors were calculated based on the conformation of guest molecule with the lowest energy in inclusion complex exported from docking procedure. Molecular Operating Environment (MOE, 2011) package was used to calculate various physicochemical, structural, topological and geometrical ligand descriptors. Complex descriptors were calculated based on ligand and receptor interactions using AutoDockTools and BINANA (BINdingANAlyzer) programs [٧].

2.3. Docking based QSPR

The results of statistical performance of QSPR models with five descriptors explained Q^2 value of 0.82±0.02, R^2_c of 0.83±0.02 and R^2_p of 0.78±0.07 with SEP=0.39±0.03 and SEC= 0.38±0.03.

2.4. Description of the descriptors

The definitions of the descriptors involved in the QSPR models provided in Table 1. The contribution of descriptors after 120 *runs* using different *random divisions* of the data set into training and prediction sets. The logP coefficient is well-known as one of the principal parameters for the estimation of hydrophobicity of chemical compounds and the effects of solvation/desolvation for ligands. The obtained QSPR model confirmed the importance hydrogen bonding and hydrophobic interactions on the complexation of β-cyclodextrin with diverse set of organic compounds.

Table 1 Description of the descriptors contributed in the models

Descriptor	Descriptions
LogP ($_{o/w}$)	Log of the octanol/water partition coefficient (including implicit hydrogens).
vdw_area	Area of van der Waals surface calculated using a connection table approximation.
pmi 1	First diagonal element of diagonalized moment of inertia tensor
std_dim3	Standard dimension 3: the square root of the third largest eigenvalue of the covariance matrix of the atomic coordinates.
vsurf_CP	Critical packing parameter
vsurf_ID7	Hydrophobic integy moment7
Binding energy	Sum of the intermolecular energy
C-H	Atom-type pair counts within 2.5 Å C-H
HD-HD	Atom-type pair counts within 2.5 Å HD-HD
C-C	Summed electrostatic energy by C-C (atom-type pair in J/mol)

References

[1] R. Agrawal, V. Gupta, Cyclodextrins–a review on pharmaceutical application for drug delivery, Int. J. Pharm. Front., 2 (2012) 95-112.
[2] S. Li, W.C. Purdy, Cyclodextrins and their applications in analytical chemistry, Chem. Rev., 92 (1992) 1457-1470.
[3] J.B. Ghasemi, M. Salahinejad, M.K. Rofouei, Review of the quantitative structure–activity relationship modelling methods on estimation of formation constants of macrocyclic compounds with different guest molecules, Supramol. Chem., 23 (2011) 614-629.
[4] Connors, K.A., 1997. The stability of cyclodextrin complexes in solution. Chem. Rev. 97, 1325-1358.
[5] Del Valle, E.M., 2004. Cyclodextrins and their uses: a review. Process biochem. 39, 1033-1046.
[۱] G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A.J. Olson, AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility, J. Comput. Chem., 30 (2009) 2785-2791.
[۲] J.D. Durrant, J.A. McCammon, BINANA: a novel algorithm for ligand-binding characterization, J. Mol. Graph. Model., 29 (2011) 888-893.

Probing and identification of the solute effects on the methanol–water association by chemometrics methods

B. Hemmateenejad[a*] Z. Shojaeifard[a], M. Shamsipur[b], K. Neymeyr[c, d], M.Sawall[c, d] and A. Mohajeri[a]

[a] *Department of Chemistry, Shiraz University, Shiraz, Iran*

[b] *Department of Chemistry, Razi University, Kermanshah, Iran*

[c] *Institute of Mathematics, Rostock University, Rostock, Germany*

[d] *Leibniz Institute for Catalysis, Rostock, Germany*

E-mail address (corresponding author):hemmatb@shirazu.ac.ir (B. Hemmateenejad)

Abstract

In this work, solute effects on the changing in the equilibrium constant of methanol-water association have been investigated. In this regard, the behavior of four solutes (solvatochromic dyes) in methanol-water binary mixtures are probed using augmentation of infrared and UV-Vis spectroscopy methods followed by multivariate analysis. In order to get a more reliable concentration and spectral profiles, with fewer ambiguities, complementarity and coupling theorems has been utilized. A ternary system was resolved by multivariate analysis on the noted augmented data for each chemical probe. Concentration profiles showed unique behavior of each solute in the binary-solvent system. Therefore, individual association constant for the formation of the methanol-water association was achieved for each solvatochromic probe. Indeed, different solutes can change solvent-solvent interaction based on their hydrogen binding affinity.

Key words: Complementarity and coupling theorems, Methanol, Water, Chemometrics

Introduction

Solvent-solute and solvent-solvent interactions are of vital importance in studying chemical, environmental, industrial, and biological processes in liquid phase.[1] Methanol-water binary solvents, as the widely used solvents in liquid chromatography, have been studied in the past using both theoretical and experimental methods. It has now been accepted that, in the binary mixtures of methanol and water, a methanol-water association (or cluster) is formed.[2] However, all of the previous studies ignored the solute effects on the structure of this association. This led to the inconsistency between the obtained results by different authors. There is no agreement between the stoichiometric ratio(s) of the associate(s) [3] and also the reported equilibrium constants for even the same association model. Here, we demonstrate the effect of different solutes on the methanol-water association in their binary mixtures. In the past, both native FTIR spectra of the pure solvents as well as the mixed solvents and the spectra of solvatochromic dyes in the pure and mixed solvents have been used to monitor the methanol-water association. However, the changes in absorbance spectra of the solvatochromic dyes can also provide some indirect measure of the solvent-solvent association. Thus, in this work, we used both techniques simultaneously for following the solute effects in dynamic parameters of binary solvent.

Results and discussion

Four standard solvatochromic indicators including N,N-dimethyl-4-nitroaniline, 4-nitroanisol, 4-nitroaniline and indigo carmine were used to probe methanol-water association. A fixed amount of each dye was dissolved in the methanol-water binary solvents and then the ATR-FTIR as well as UV-Vis absorbance spectra of the resultant solutions were recorded. Both singular value decomposition (SVD) and evolving factor analysis (EFA) were employed to determine the number of the significant principal components, attributing to the number of the solvated complex in the systems. Principal component analysis (PCA) of the data matrices of all dyes revealed the presence of three significant chemical species attributed to the pure methanol, pure water, and a methanol-water association. The absorbance data matrix of each dye can be discomposed into the contribution of each solvated state of the dye:

$$\mathbf{D = D_m + D_w + D_{mw} = c_m s_m + c_w s_w + c_{mw} s_{mw} = [c_m c_w c_{mw}]\,[s_m;\, s_w;\, s_{mw}] = CS} \qquad \mathbf{(1)}$$

The effect of solutes on solvent structure is very small, since there might be an overlap between the ranges of solutions of different dyes and, hence, the comparison might not be meaningful. So, for obtaining a unique solution that is a key in comparing the effects of solutes, complementary and coupling theorem have been used. [4], [5]

Four dyes has the similar ternary contribution in the methanol-water mixed solvents. However, the concentration of the associated solvent varies as dyes vary. The observed differences can be attributed to the changing in the equilibrium constant of the association model with the form: $\mathbf{M + W \leftrightarrows MW}$ **(2)**

The association constant can be estimated by nonlinear fitting of the concentration profiles to association model of Eq. (2). As it is observed from Table 1, there is a large difference in the estimated association constants for different solutes. It can be attributed to the affinity of variant electron donor and electron acceptor functional groups toward hydrogen bonding. Therefore, attributing a fix and invariant association constant to binary solvent association is not acceptable and even the type of solute should be considered.

Table1 Association constant of methanol- water complex by using various dyes

Dye	Association constant	ssq
4-nitroanisol	2.32 ± 0.0003	1.60
4-nitroaniline	0.64 ± 0.0007	0.35
N,N dimethyl 4-nitroaniline	0.54 ± 0.0013	0.82
Indigo carmine (ICM)	0.10± 0.0043	0.34

References

[1] Y. Marcus, *The Properties of Solvents*. Wiley-Vch, 1998.

[2] S. Nigam, Juan A, R. Stubbs, and S. Rutan, Characterization of methanol-water and acetonitrile-water association using multivariate curve resolution methods, *Anal. Chem.*, 2000, **72**, 1956–1963.

[3] Z. Zhao and E. R. Malinowski, Detection and identification of a methanol-water complex by factor analysis of infrared spectra, *Anal. chem.*, 1999, **71**, 602–608.

[4] M. Sawall, C. Fischer, D. Heller, and K. Neymeyr, Reduction of the rotational ambiguity of curve resolution techniques under partial knowledge of the factors. Complementarity and coupling theorems, *J. Chemom.*, 2012, **26**, 526–537.

[5] M. Sawall and K. Neymeyr, "On the area of feasible solutions and its reduction by the complementarity theorem., *Anal. Chim. Acta*, 2014, **828**, 17–26.

The study of pre-concentration and determination of some esterphthalates in plastic containers by using QuEChERS method and high performance liquid chromatography

Banafsheh Aghvami*, Hakim Faraji

E-mail address (corresponding author): Banafshehaghvami@yahoo.com

Abstract

Plasticizers are chemical compounds which are usually used to improve flexibility, efficiency and workability of plastics particularly polymer and some particular color compounds. These properties have spurred increasing application of plasticizers during the recent years. Phthalates are the main chemical compound being used in plasticizers. Some of these compounds such as phthalate ester are regarded not optimal due to having potential toxicity and the ability of transference from materials into environment. So in this case they can pollute water, soil, air and food substance putting human health subject to jeopardy. In this research, four main phthalate ester compounds including DMP, BBP, DBP and DEHP were used for the purpose of pre-condensation, separation and quantitative evaluation. These processes were carried out in disposable vessels samples made of polystyrene and polyethylene using QuEChERS method coupled with liquid chromatography (HPLC) and DAD. To ensure optimality of experiment, chemo-metrics method was particularly devised using FFD method. Having considered the effects of different parameters including type and quantity of absorbents, the time of extraction and volume of extractors were selected.

In this study, homogenized samples (1gr samples of plastic vessels) were put in centrifuge tubes containing 50mL of MgSO4 and NaCl (1gr). Having improved the samples based on relevant standards, 10mL acetonitrile was added to them and were put to be shook for 1 min. in this step, 150mgC18 and 300 PSAmg were added to the tube content. Then the content was put subject to ultrasound instrument for 15 minute and then they were centrifuged for 10 minutes by a speed of 5000 rpm. At the end, 20ml extracted solution was used for decomposition of HPLC-DAD.

Results obtained from standard addition method indicated that linear range (LR) of rating curve was 20 for the samples in 10-500 mg,kg-1 and for BBP in range of 20-500 mg kg^{-1}. Quantitative coefficient (R2) for rating curves was 0.9926 and 0.9913 respectively for polystyrene and polyethylene compounds. Limit of discernment was in the range of 0.6 -2.0 mg kg -1 for analyte and 0.6-1.4 mgkg -1 for polyethylene compounds. Also limit of quantity for the said samples was 2.0-6.6 mg kg -1 and 2.0-4.6 mg kg^{-1} respectively. Repeatability of this method was assessed and the intermediator precision for analyte was found to be less than 6.6 and 7.4. also, relative recovery of the compounds was assessed to be 94/41 – 102/00 for polystyrene and 95/26-102/05.

Key words: *QuEChERs, experimental design, chemo-metrics, liquid chromatography, plasticizer*

Classification of *Mentha piperita* accessions according to their antioxidant activity

Parisa Izadiyan[1,3], Bahram Hemmateenejad[1, 2*]

[1] *Medicinal and Natural Products Chemistry Research Center, Shiraz University of Medical Sciences, Shiraz, Iran*
[2] *Chemistry Department, Shiraz University. Shiraz, Iran*
[3]*Student Research Committee, Shiraz University of Medical Sciences, Shiraz, Iran*

E-mail address (corresponding author): hemmatb@sums.ac.ir

Abstract

In this work, phenolic fingerprints of twenty two accessions of Mentha piperita (peppermint) collected from different provinces of Iran were investigated by HPLC-DAD analysis and chemometric methods. Different supervised and unsupervised pattern recognition methods were evaluated to gategorize the samples based on their antioxidant activity level. Evaluation of the phenolic fingerprints of peppermint samples revealed that it is possible to classify the samples using chemometrics methods. Using their HPLC fingerprints the samples were categorized into three classes with high, medium and low antioxidant activity. The proposed classification model could also be applied in food industry for quality control and for commercial marketing in the field of herbal medicine.

Key words: Fingerprinting, HPLC-DAD, PLS-DA, *Mentha piperita*, peak alignment

Introduction

Mentha piperita (peppermint) is a member of the *Labiatae* family with a high antioxidant activity (1). It is widely used not only as a culinary but also as a medicinal herb in Iran and other countries all around the world (2). The chemistry of peppermint extract is very complex and highly variable according to different parameters including; growth season, geographical origin, and collection time (3). The chemical alterations resulting from the mentioned factors will probably affect antioxidant activity of the plants too.

In this work, high performance liquid chromatography coupled with diode array detector (HPLC-DAD) fingerprinting and antioxidant activity analysis in terms of 1,1-diphenyl-2-picrylhydrazyl (DPPH) scavenging activity have been carried out to investigate firstly the chemical composition of Iranian peppermint and secondly to introduce a chemometric classification model for clustering the peppermint samples based on their antioxidant activity level.

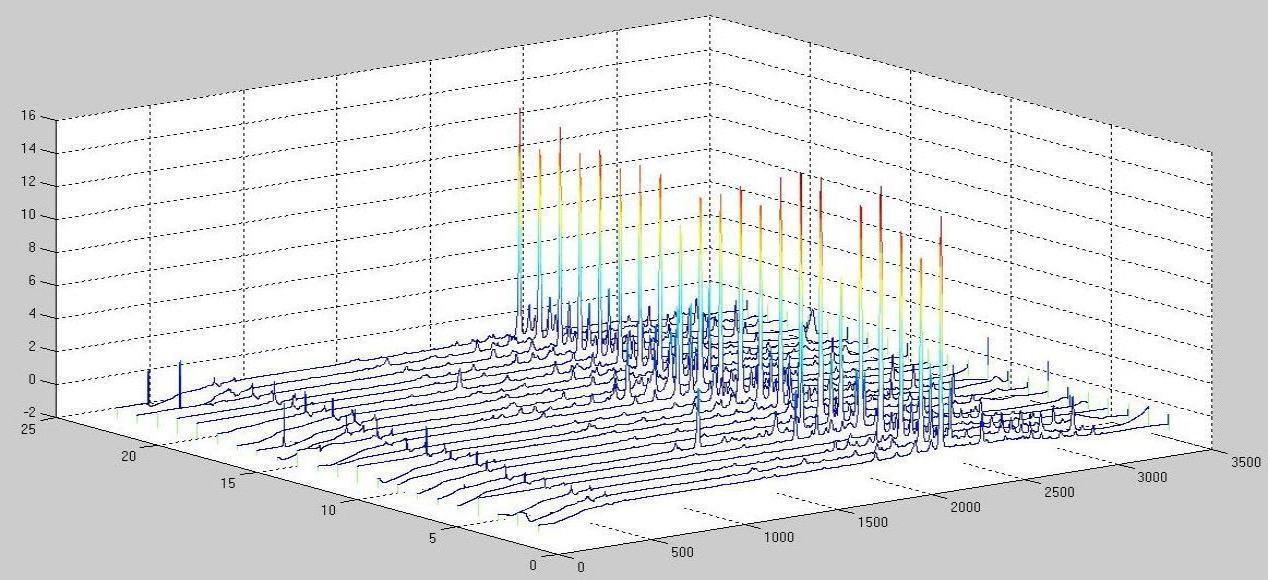
Results and discussion

Mentha piperita accessions were collected during different growth seasons from varied geographical sources including northern, southern, eastern and western cities of Iran. Fingerprints of 22 peppermint samples were achieved by high-performance liquid chromatography coupled with a diode array detector (HPLC-DAD). DPPH antioxidant assays were also performed on all of the samples and the antioxidant activities were expressed in the form of half maximal inhibitory concentration (IC_{50}). Afterwards, peppermint samples were grouped into three levels of high, medium and low antioxidant activity based on the corresponding logarithms of IC50 values.

Chromatographic fingerprints were extracted in text format from the instrument and were prepared for chemometrics data analysis. According to some chromatographic conditions peak shifts occur from sample to sample that should be

considered and revised. In order to correct the peak shifts, various peak alignment algorithms were examined such as: the correlation optimised shifting algorithm (Coshift), dynamic manifold warping (DMW) and correlation optimized warping (COW), among which, the last algorithm performed the best (Figure 1). As the next step, preprocessed data were subjected to two exploratory data analysis using Principal Component Analysis (PCA) and hierarchical cluster analysis (HCA). The results were not satisfactory. Thereafter, different algorithms were also assessed for constructing the classification model such as; soft independent modeling of class analogy (SIMCA) and partial least square discriminant analysis (PLS-DA). Cross validation was also used for model validation. The best discrimination among samples with three different antioxidant levels was obtained using PLS-DA classification model with model and cross validated non-error rate of 0.92 and 0.90, respectively. The obtained results once again showed up the superiority of PLS-DA algorithm over other tested classification methods. The proposed classification model could also be applied in food industry for quality control and for commercial marketing in the field of herbal medicine.

Figure 1. Cow aligned peppermint samples



References

1. Naghibi f, mosaddegh m, mohammadi motamed m, ghorbani a. labiatae family in folk medicine in iran: from ethnobotany to pharmacology. iranian journal of pharmaceutical research. 2010:63-79.
2. Singh r, shushni ma, belkheir a. antibacterial and antioxidant activities of mentha piperita l. arabian journal of chemistry. 2011.
3. Karray-bouraoui n, ksouri r, falleh h, rabhi m, jaleel ca, grignon c, et al. effects of environment and development stage on phenolic content and antioxidant activities of mentha pulegium l. J Food Biochem. 2010;34:79-89.

3D-QSAR Studies for Prediction of Graphene Dispersibility in Diverse Organic Solvents

F. Mirrahimi[1], M. Salahinejad[2], Jahan B. Ghasemi[1, 3]

[1] *Drug Design in Silico Lab., Chemistry Faculty, K N Toosi University of Technology, Tehran, Iran*

[2] Nuclear Science and Technology Research Institute, Tehran, Iran

[3] Chemistry Faculty University of Tehran, Tehran, Iran

Abstract

This work focuses on the dispersibility of graphene in different solvent phases. A chemometrics study was performed on the dispersibility of graphene in liquid phases. Two models were constructed using theoretical and empirical parameters of the solvents. Molecular descriptors were computed using GRid Independent Descriptors (GRIND). After variable selection via genetic algorithm method, Partial Least Squares (PLS) analysis was carried out. A simple and fairly good predictive linear model based on VolSurf descriptors was also developed that showed an adequate prediction power of the the dispersibility of graphene. The results confirmed that dispersive interactions, in comparison with H-bonding and polar interactions, might have a more significant role in increasing the graphene dispersibility.

Keywords: 3D-QSAR, Graphene, Dispersibility, GRID,VolSurf descriptors

1. Introduction

Graphene, a two-dimensional sheet of sp_2-bonded carbon atoms arranged in a honeycomb lattice [1] , is the thinnest known material and the strongest ever measured [2]. Potential applications of graphene for nano-electronics [3,4], sensors [5], and nanocomposites [6] have been actively pursued [7]. Research on its potential for biological applications has also been reported, e.g. in drug loading and delivery. Among different computational methods, quantitative structure–activity/property relationship (QSAR/QSPR) methods seem to provide a compromise between the speed and accuracy of the predictions and the QSAR models can become progressive tools to help prioritizing the experimental work. In our recent study, we developed a valid and predictive 3D-QSAR model using GRid INdependent Descriptors (GRIND) [8] and VolSurf descriptors that it is able to correlate and predict the dispersibility of graphene in different organic solvents.

2. Results and discussion

Alignment free, three dimensional quantitative structure activity relationships (3D-QSAR) of dispersibility of graphene in different organic solvents was reported. GRIND methodology, where descriptors are derived from GRID molecular interaction fields (MIF), was used [8]. After variable selection via genetic algorithm (GA), Partial Least Squares (PLS) analysis was carried out, and a highly descriptive and predictive model was obtained. The statistical parameter of GIRIND model obtained with Q2 value of 0.76, R2c of 0.91 and R2p of 0.83 , and Sep of 0.13,and Sec of 0.13.

VolSurf descriptors, as the other known GRID-based method, were applied to develop a simple and predictive multiple linear regression (MLR) model. The statistical parameter of VolSurf model obtained with Q2 value of 0.83, R2c of 0.88 and R2p of 0.72 ,and Sep of 0.15 ,and Sec of 0.15. Based on the molecular descriptors in these models, it was suggested that some weak van der Waals interaction would help in the dispersibility of graphene. Among these interactions, dispersive interactions, in comparison with H-bonding and polar interactions, might have a more significant role in increasing the graphene dispersibility in various organic solvents.

References:

[1] T. Seyller, A. Bostwick, K. V. Emtsev, K. Horn, L. Ley, J. L. McChesney, T. Ohta, J. D. Riley, E. Rotenberg, F. Speck, Phys. Status Solidi B 2008, 245,1436 –1446.
[2] A. K. Geim, Science 2009, 324, 1530 –1534.
[3] Y. Kopelevich, P. Esquinazi, Adv. Mater. 19 (2007) 4559.
[4] X.L. Li, X.R. Wang, L. Zhang, S.W. Lee, H.J. Dai, Science 319 (2008) 1229.
[5] J. Zhan, G.F. Chen, L. Zhu, G.X. Li, Electrochem. Commun. 13 (2011) 31.
[6] S. Stankovich, D.A. Dikin, G.H.B. Dommett, K.M. Kohlhaas, E.J. Zimney, E.A. Stach,R.D. Piner, S.T. Nguyen, Nature 442 (2006) 282.
[7] D. Li, M.B. Müller, S. Gilge, R.B. Kaner, G.G. Wallace, Nat. Nanotechnol. 3 (2008)101.
[8] A´ngel Dura´n, Ismael Zamora, and Manuel Pastor, *J. Chem. Inf. Model.* 2009, *49*, 2129–2138

MIA for a QSAR study of inhibitory activity of CCB with using WT and GA for pixel processing

Ahmad Akrami [a] Ali Niazi *[,a], Jahan B. Ghasemi [b]

[a]*Department of Chemistry, Arak Branch, Islamic Azad University, Arak, Iran*

[b]*Faculty of Chemistry, K.N. Toosi University of Technology, Tehran, Iran*

E-mail address: a-niazi@iau-arak.ac.ir & ali.niazi@gmail.com (A. Niazi)

Abstract

The quantitative structure-activity relationship (QSAR) studies have been carried out on a series of dihydropyridine (DHP) derivatives known as calcium channel blocker (CCB) drugs. The partial least squares (PLS) algorithm was used for prediction of inhibitory activity of calcium channel antagonists as a function of the bidimensional images. In this paper is investigated the effect of pixel selection by application of genetic algorithms (GAs) for PLS model, because of the GAs is very useful in the variable selection in modeling. On the other hand, pre-processing methods such as wavelet transform (WT) were used, in order to improve the predictive power of multivariate calibration methods. The subset of pixels, which resulted in the low prediction error, was selected by genetic algorithm. In order to evaluate of models in this study (PLS, GA-PLS, WT-PLS and WT-GA-PLS), the inhibitory activity of several compounds, which were not in the modeling procedure were predicted. The results of models showed high prediction ability with root mean square error of prediction of 0.779, 0.288, 0.159 and 0.070 for PLS, GA-PLS, WT-PLS and WT-GA-PLS, respectively. The WT-GA-PLS method was used for prediction of inhibitory activity of the new antagonists.

Key words: CCB, QSAR, PLS, GA, WT, MIA.

Introduction

Calcium channel blockers (CCBs) are a group of medicines that affect the way calcium passes into certain muscle cells. The DHP derivatives are an important class of drugs known as calcium channel antagonists [1]. An important step in constructing QSAR/QSPR models is to find one or more molecular descriptors that represent variation in the structural property of the molecules by a number [2]. The most applied computational methods for QSAR/QSPR are based on 3D or multidimensional approaches, which provide information about the relation between physicochemical parameters, such as steric and electrostatic fields, and bioactivity. Esbensen and Geladi [3] have demonstrated that image analysis may provide useful information in chemistry; the descriptors do not have a direct physicochemical meaning, since they are binaries. In multivariate image analysis-quantitative structure activity relationship (MIA-QSAR), descriptors are pixels (binaries) of 2D images, which are the chemical structures of the active compounds. In MIA-QSAR, images (2D chemical structure) have shown to contain chemical information, allowing the correlation between chemical structures and properties [4].

Results and discussion

The inhibitory activity data of DHP derivatives were taken from literature [1].

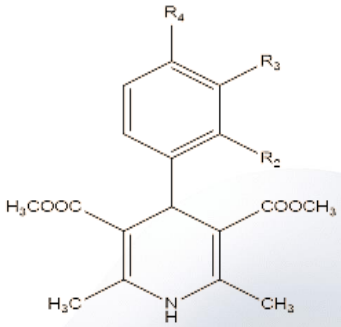


Figure 1. Chemical structure of DHP derivatives

In order to enhance the predictive power of multivariate calibration methods, molecular (pixels) descriptors are often corrected prior to the data analysis. In order to evaluate the WT-GA-PLS method, four methods (PLS, GA-PLS, WT-PLS and WT-GA-PLS) were tested with set of tests. The statistical parameters obtained by these methods are listed in Table 1.

Table 1. Observation and calculation values of $\log(1/IC_{50})$ using PLS, GA-PLS, WT-PLS and WT-GA-PLS models

Number of compounds (**Ref. 1**)	Observation $\log(1/IC_{50})$	PLS		GA-PLS		WT-PLS		WT-GA-PLS	
		Predicted	Error (%)	Predicted	Error (%)	Predicted	Error (%)	Predicted	Error (%)
3	8.66	9.40	8.54	9.00	3.92	8.85	2.19	8.73	0.80
11	7.80	8.64	10.76	8.10	3.84	7.95	1.92	7.88	1.02
12	7.68	6.98	-9.11	7.42	-3.38	7.51	-2.21	7.59	-1.17
14	7.46	8.1	8.57	7.60	1.87	7.52	0.80	7.51	0.67
16	7.37	6.63	-10.04	7.03	-4.61	7.26	-1.49	7.33	-0.54
19	7.13	6.27	-12.06	6.89	-3.36	6.97	-2.24	7.05	-1.12
29	6.89	7.79	13.06	7.23	4.00	7.11	3.19	6.96	1.01
LVs		10		8		7		6	
R^2		0.496		0.840		0.934		0.985	
RSEP (%)		10.268		3.798		2.096		0.929	

The above results indicate that WT-GA-PLS is a very promising tool for the prediction of inhibitory activity of calcium channel antagonists for 35 compounds based multivariate image analysis alone. The results well illustrate the power of pixels in prediction of inhibitory activity of DHP derivatives. The model could predict the inhibitory activity of DHP derivatives not existed in the modeling procedure accurately.

References

[1] Hasegawa K, Miyashita Y, Funatsu K. GA strategy for variable selection in QSAR studies: GA-based PLS analysis of calcium channel antagonists. *J. Chem. Inf. Comput. Sci*. 1997; **37**: 306-310.

[2] Isayev O, Rasulev B, Gorb L, Leszczynski J. Structure-toxicity relationships of nitroaromatic compounds. *Mol Diversity*. 2006; **10**: 233-245.

[3] Geladi, P, Esbensen K. Can image análisis provide information useful in chemistry?. *J. Chemom.* 1989; **3:** 419-429.

[4] Nunes CA, Freitas MP. Introducing new dimensions in MIA-QSAR: a case for chemokine receptor inhibitors. *Euro. J. Med. Chem.* 2013; **62**: 297-300.

Determination of Enantiomer Excess of Some Amino Acids Using Fluorescence Spectroscopy and Multivariate Curve Resolution Analysis

B. Hemmateenejad*[1], M. Shamsipuor[2], A. Naghashian Haghighi[1]

(1) Department of Chemistry, Shiraz University, Shiraz, Iran.
(2) Department of Chemistry, Razi University, Kermanshah, Iran.

E-mail address: hemmatb@sums.ac.ir

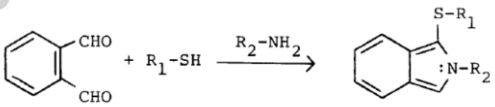
Abstract

In this work a non-separative spectroflourimetric method based on the time-dependent fluorescence emission was proposed for the detection of amino acid (AA) enantiomers. An optically active chiral thiol compound, 1-mercapto-2-propanol (MP), is used as a chiral derivatization reagent. This compound and o-phthaldialdehyde (OPA) react with amino acid enantiomers to produce fluorescent diastereomers that are detectable in fluorescence spectroscopy. At room temperature and in the borate buffer solution of pH 9.6, d the formed fluorescent diasteriomers of D and L-amino acid exhibited difference in time-dependent fluorescence intensity at about 450 nm. To gain information from the small spectral changes, the kinetic-flourscence data matrices were analyzed by Multivariate Curve Resolution Alternating Least Squares (MCR-ALS) method. A linear calibration cureve is achived to distinct D and L-Lysine in different mole ratios by appling appropriate constraints in MCR-ALS proceedures. In the present investigation, detection of enantiomeric excess can benefit from direct fluorescence spectra analysis without appling any separating methods.

Key words: amino acid, enantiomer, *o*- phthaldialdehyde, Chiral derivatization, fluorimetric detection, MCR-ALS

Introduction

Chiral compounds exhibit different properties in biochemical systems. Two molecules that are mirror images of each other are called anenantiomeric pair. The key step in chiral recognition is the formation of diastereoisomeric complexes between the enantiomers and a chiral selector. Molecular recognition results because of the differences in Gibbs free energy between the two diastereoisomeric enantiomer–selector complexes.[1] Particular attention has been paid to the chiral separation of amino acids (AAs) because of the different biological activities of their stereoisomers.[2]AAs in their native form are generally weak chromophores and do not possess electrochemical activity, so they must first be chemically modified (derivatized). OPA readily reacts with primary amines in alkaline media in the presence of thiol compounds, to give intensely fluorescent N-alkyl-2-alkylthio-substitutedisoindole derivatives.[3]OPA rapidly forms fluorescent derivatives (λex= 340 nm, λem= 455 nm) at room temperature, is nonfluorescent itself, and when present in excess does not break down or react to form fluorescent by-products. However, the derivatives formed are somewhat unstable and this severely limits the general utility of the reaction.[4]



Almost in all of the researches have been done up to now, OPA-thiol adduct has been used to separate AAs by the help of chromatographic methods. In this research we are going to apply OPA-MP as a chiral selector compound to

determine AAs enantiomers by analyzing their fluorescence spectra without using chromatographic approaches by appling Multivariate Curve Resolution Alternating Least Squares (MCR-ALS) method. MCR-ALS has become a popular chemometric method used for the resolution of multiple component responses in totally or partly unknown unresolved mixtures. This recognition is due to the great variety of data sets that can be analyzed by curve resolution methods. Other reasons for the acceptance of MCR-ALS are its ability to deal with multiple data matrices simultaneously[5,6] and the diversity and flexible application of constraints to help and improve the resolution results.

Results and discussion

Appropriate amounts of reagents and AA enantiomers solutions were added to fluorescence quartz cuvette respectively in dark conditions. The emission spectra exhibited similar peak maxima at 450 nm. However, they differ in kinetic profile (Fig. 1). By recoding flourscence spectra in 25.5 min time duration, a kinetic-flourescence data was obtained for each enatiomeric composition. Then simultaneous analysis of data matrices of 11 studied enatiomeric ration was achieved by MCR-ALS. The area under concentration propfile of one resolved component showed linear dependency on enantiomeric excess (Fig. 2).

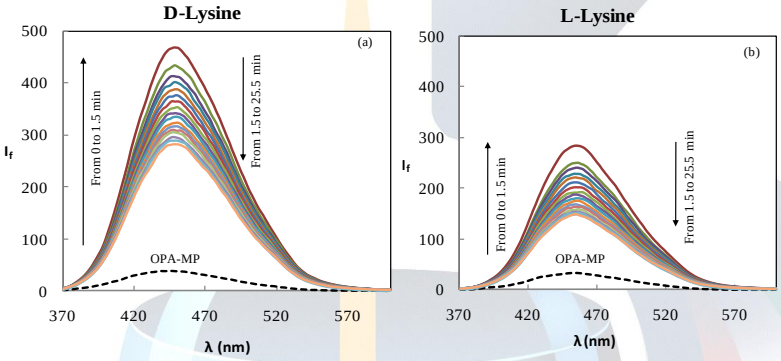


Fig. 1. Emission spectra of (a) D-Lysine and (b) L-Lysine which increases in 1.5 min after injecting AA to the quartz cuvette containing OPA-MP adduct and then decreases in 25.5 min.

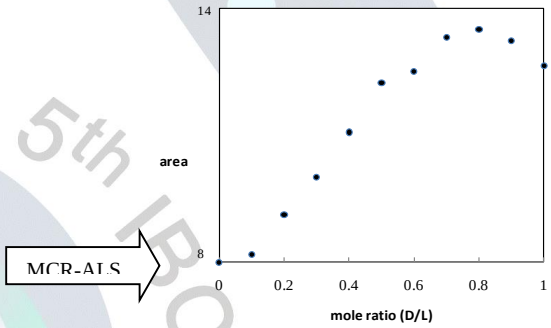


Fig. 2. Linear calibration curve obtained for D and L-Lysine mole ratios by calculating the areas of resulting concentration profile for each ratio (Conc. Vs. time) from MCR-ALS.

References

[1] Berthod A.; Chiral Recognition Mechanisms, *Anal. Chem.* 2006, 1, 2093-2099.

[2] Wan H. and BlombergL.G., Optimized Chiral Separation of 20 Amino Acids Derivatized with 9-Fluorenylmethyl Chloroformate Using Cyclodextrins as Chiral Selectors in Capillary Electrophoresis, *Journal of Chromatographic Science, 1996, 34, 540-546.*

[3] Fujiwara M., Ishida Y., Nimura N., Toyawa A., Kinoshita T., Postcolumn fluorometric detection system for liquid chromatographic analysis of amino and imino acids using o-phthalaldehyde/N-acetyl-l-cysteine reagent *Anal. Biochem*, 1987, **166**, 72-78.

[4] García Alvarez-CoqueM. C., Medina HernándezM. J., Villanueva CamañasR. M.,Mongay Fernández C., Formation and instability of o-phthalaldehyde derivatives of amino acids, *Anal. Biochem*., 1989, **178**, 1-7.

[5] Ahmadi G, Tauler R, Abdollahi H, Multivariate calibration of first-order data with the correlation constrained MCR-ALS method, C*hemometrics and Intelligent Laboratory Systems*, 2015, **142**, 143–150

[6] Jaumot J., Gargallo R., de Juan A., Tauler R., A graphical user-friendly interface for MCR-ALS: a new tool for multivariate curve resolution in MATLAB, *Chemometrics and Intelligent Laboratory Systems*,2005, **76**, 101–110.

Investigation of the interaction between adenosine and DNA G-quadruplex by fluorescence resonance energy transfer and molecular modeling

Zahra Hashemian[a] , Marziyeh Poshteh Shirani[a] and Taghi Khayamian[a,*]

[a] *Department of Chemistry, Isfahan University of Technology, Isfahan, Iran*

E-mail address: taghi@cc.iut.ac.ir

Abstract

A new biosensor was designed for analysis of adenosine with DNA G-quadruplex based on Fluorescence Resonance Energy Transfer (FRET). The proposed method was applied for analysis of adenosine in patients with cancer. Furthermore, the interaction of two adenosine molecules with DNA G-quadruplex was explored by using docking and Molecular Dynamic Simulation (MDS). The molecular docking results showed that two adenosine molecules are placed at the center of G-quartet to stabilize the DNA G-quadruplex structure and two adenosine molecules have two hydrogen bonds and seven hydrophobic bonds with DNA G-quadruplex. The MDS results revealed that the conformations of DNA G-quadruplex with two adenosine molecules were changed in different time scale of MDS. At in time scale of 9 ns two adenosine molecules have two hydrogen bonds and four hydrophobic bonds with DNA G-quadruplex. In addition, a hydrophobic bond was created between two adenosine molecules.

Keywords: Fluorescence Resonance Energy Transfer; Adenosine; DNA G-quadraplex; Molecular docking; Molecular dynamics simulation

1. Introduction

Adenosine plays a crucial role in the regulation of physiological activity in various tissues and organs. As adenosine is a possible biomarker for cancer, the determination of its level presents a demanding task for deeply monitoring progress of diseases**.** Therefore, adenosine can act as a tumor marker and its concentration in the urine can be determined to monitor the progress of the diseases [1]. Huizenga et al. indicated that the structure of the complex between ssDNA-aptamer and adenosine in solution consists of two small Watson–Crick helices and two G-quartets. Four guanosines assembled via hydrogen bonds and two G-quartets stack on each other to form a stable structure in the presence of metal cations. Aptamer catches two adenosine molecules between the top G-quartet and the two short stems form a pocket [2].

2. Results and discussion

2.1. Chemical structures of adenosine and DNA G-quadruplex

The chemical structure of adenosine was optimized using quantum mechanics at the density functional theory with B3LYP method 6-311G ** basis set. The crystallographic structure of DNA G-quadruplex was taken from the Protein Data Bank.

2.2. Molecular docking

The docking studies were performed by the Auto dock Vina program [3]. At first, one adenosine molecule was docked with DNA G-quadruplex. Then, the lowest binding energy conformation from the docking of the first adenosine was loaded into the DNA G-quadruplex. This complex was used for docking the second adenosine molecule.

2.3 Molecular dynamic simulation of two adenosine molecules with DNA G-quadruplex

The lowest conformation of two adenosine molecules with DNA G-quadruplex, resulted from the docking was loaded for MDS. The topological parameters of two adenosine molecules were generated using the Dundee PRODRG2.5 server (beta) [4]. In order to obtain conformation of DNA G-quadruplex with two adenosine molecules in a water environment, 10 ns MDS were performed in a cubic water box.

We experimentally shown that each DNA G-quadruplex captures two adenosine molecules [5].The results of MDS at time scale t=9 ns are shown Fig 1(a) and at in this time scale, a hydrophobic bond was created between two adenosine molecules. The results of this time scale of MDS are in agreement with the statement of Huizenga and Szostak, [2] where two adenosine molecules may stack between the top G-quartet and two short stems.

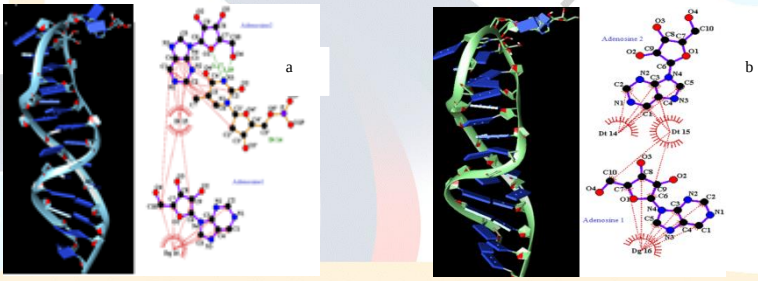


Figure 1. molecular dynamic simulation of two adenosine molecules with DNA G-quadruplex

References

[1] M. Mujoomdar, D. Hoskin, J. Blay, Adenosine stimulation of the proliferation ofcolorectal carcinoma cell lines: roles of cell density and adenosine metabolism; *Biochem. Pharmacol*; 2003; **66**; 1737–1747.

[2] D.E. Huizenga, J.W. Szostak; A DNA aptamer that binds adenosine and ATP; *Biochemistry*; 1995; **34**; 656–665.

[3] O Trott, A J. Olson; Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading; *J. Comput. Chem*. 2010, **31**, 455–461.

[4] A. W. Schuttelkopf and D.M.F. van Aalten, PRODRG: A tool for high-throughput crystallography of protein-ligand complexes, *Acta Crystallogr*; 2004; **60**; 1355–1363.

[5] M. E. Najafabadi, T. Khayamian, Z. Hashemian, Aptamer-conjugated magnetic nanoparticles for extraction ofadenosine from urine followed by electrospray ion mobilityspectrometry; *J. Pharm. Biomed. Analysis*; 2015; **107**; 244-250.

Tabu search as a descriptor selection in QSAR study of 1,5-diheteroarylpenta-1,4-dien-3-ones: a class of promising curcumin-based anti-cancer agents

Morteza Atabati*, Maryam Teimuori

School of Chemistry, Damghan University

E-mail address: atabati@du.ac.ir

Abstract

Quantitative structure–activity relationship (QSAR) studies based on tabu search (TS) were carried out for the prediction of inhibitory rate of 1,5-diheteroarylpenta-1,4-dien-3-ones derivatives. Ts is a meta-heuristic algorithm proposed to feature selection. A major problem of QSPR and QSAR is the high dimensionality of the descriptor space; therefore, descriptor selection is the most important step. Tabu search has important links to evolutionary and "genetic" methods, often overlooked, through its intimate connection with scatter search and path relinking — evolutionary procedures. In this work, a TS algorithm was used to select the best descriptors.Then selected descriptors were applied for model development using multiple linear regression (MLR).

Key words: Tabu search; Descriptor selection; 1,5-diheteroarylpenta-1,4-dien-3-ones; Anti cancer

Introduction

In a continuing study of 1,5-diheteroaryl-1,4-pentadien-3-ones mimics as potential drug candidates to treat prostate cancer. These 1,5-diheteroaryl-1,4-pentadien-3-ones mimics were tested for cytotoxicity against androgen-sensitive LNCaP and androgen-insensitive PC-3 and DU-145 human prostate cancer cell lines, as well as HeLa human cervical cancer cells [1]. Another interesting aspect of 1,5-diheteroaryl-1,4-pentadien-3-ones activity is the ability to exert both adioprotective effect in normal cells and radiosensitizing effects in cancer cells. More and more preclinical studies support the idea that this chemical could be a promising anti cancer drug for a variety of tumors. 1,5-diheteroaryl-1,4-pentadien-3-ones has been demonstrated to exert its anti tumor activity mainly through the inhibition of NF-kB, which regulates the expression of a number of genes involved in cancer development and progression [1,2].

The tabu search, proposed by Glover, is a meta-heuristic method that can be used to solve combinatorial optimization problem. Its flexible control framework and several spectacular successes in solving NP-hard problems caused rapid growth in its application. It differs from the local search technique in the sense that tabu search allows moving to a new solution which makes the objective function worse in the hope that it will not trap in local optimal solutions. Tabu search uses a short-term memory, called tabu list, to record and guide the process of the search [3].

Results and discussion

Experimental inhibitory rate values of 1,5- diheteroarylpenta-1,4-dien-3-ones derivatives were assembled from previous publications [1]. The data set was randomly divided in two groups, a calibration set and a prediction set consisting of 30 and 13 compounds, respectively. The calibration set was used for the model generation and the prediction set was used for the evaluation of the generated model. Structures of the molecules were drawn in the HyperChem. The optimization of the molecular structures was carried out by the semi-empirical AM1 method. After optimization of the molecular structures different descriptors were calculated by Dragon. After elimination of descriptors with zero value and descriptors with same value for all molecules, the remained descriptors were used.

Tabu search (TS) program was written in Matlab in our laboratory and used for descriptor selection in this work. After optimization of TS parameters, the best descriptors (9 descriptors) were selected. Multiple linear regression (MLR) method was build with these descriptors. The relationships between experimental and calculated values of inhibitory rate for the calibration and prediction sets are shown in Fig. 1.

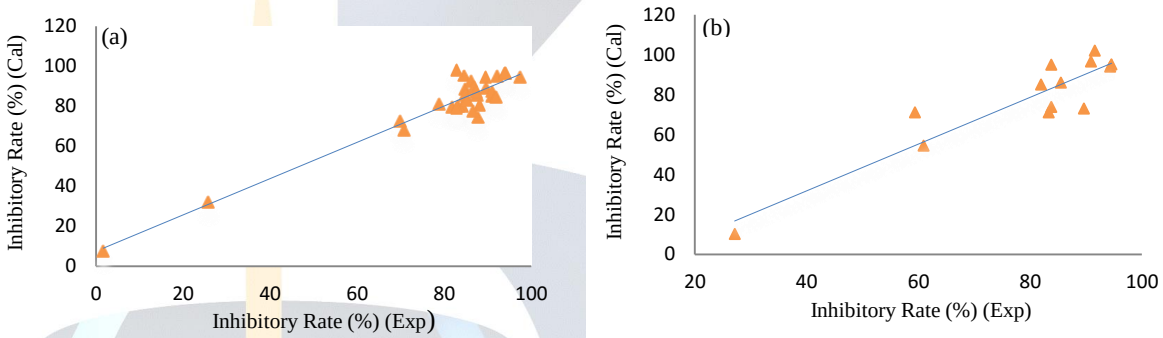


Figure 1. Relationship between the experimental and calculated pIC_{50} values of inhibitory rate for a) calibration set and b) prediction set.

The root mean square error (RMSE) and correlation coefficient (r) for the calibration set were 5.9 and 0.9521, respectively, while the root mean square error (RMSE) and correlation coefficient (r) for the prediction set were 9.9 and 0.9193, respectively.

References

[1] Rubing W, Chengsheng C, Xiaojie Z, Changde Z, Qiu Z, Guanglin C, Qiang Z, Shilong Z, Guangdi W, Chen QH. Structure activity relationship and pharmacokinetic studies of 1,5- Diheteroarylpenta-1,4-dien-3-ones: aclass of promising curcumin-based anti-cancer agents. *J. Med. Chem*. 2015; ۵۸: 4713-4726.

[2] Adams B, Ferstl E, Davis M, Herold M, Kurtkaya S, Camalier R, Hollingshead M, Kaur G, Sausville E, Rickles F. Synthesis and biological evalution of novel curcumin analogs as anti-cancer and anti-angiogenesis agents. *Bioorg. Med. Chem*. 2004; **12**: 3871-3883.

[3] Porto SCS, Rabeiro CC. Parallel tabu search message-passing synchronous strategies for task scheduling under precedence constraints. *J. Heurist*. 1995; **1**: 207-255.

The Effect of Acetylation on Lysine Residues on the Structure and Dynamics of TDP-43

Jahan B. Ghasemi* & Farahnaz Rezaei Makhuri

Drug Design in Silico Laboratory, Chemistry Faculty, KN Toosi University of Technology, 15875-4416, 470 Mirdamad Ave. West, 19697 Tehran, Iran.

E-mail address (corresponding author): : jahan.ghasemi@gmail.com

Abstract

TAR DNA binding Protein (TDP-43) pathology is a disease hallmark that characterizes amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD-TDP). Post-translational modifications (PTMs) of proteins play a key role in different cellular processes ranging from enzymatic activation to regulation of signal transduction to cell-cycle control. The effects of acetylation, one of the most common protein post-transcriptional modifications, on TDP-43 lysine residues within the RNA-binding domains (RRMs) are characterized herein using a series of computer simulations. Acetylation of lysine residues is thought to promote conformational changes and to control protein binding with or dissociation from other nucleotides. Indeed, our results suggested that a single acetylation event is sufficient to impair RNA binding, implicating Lys-145 as a critical determinant of TDP-43 function.

Key words: acetyl lysine, TDP-43, molecular dynamics simulation, Amyotrophic lateral sclerosis.

Introduction

Amyotrophic lateral sclerosis (ALS) is the most common adult motorneuron disease that affects ~2/100 000 individuals each year worldwide. Patients with ALS suffer from rapidly progressive degeneration of motorneurons ultimately leading to death. The major pathological features observed in post-mortem tissue from patients with ALS are motorneuron loss, cortical spinal tract degeneration, gliosis and cytoplasmic neuronal inclusions formed by TDP-43 or TAR DNA binding Protein with a molecular mass of 43 kDa, which are now recognized as the signature lesions of sporadic ALS [1]. TDP-43 is a highly conserved and ubiquitously expressed nuclear protein that contains two RNA-recognition motifs (RRMs) involved in RNA and DNA binding, as well as a glycine-rich C-terminal sequence, which harbours the majority of the ALS-linked mutations. TDP-43 has diverse cellular roles in regulating RNA splicing and RNA stability as well as other gene regulatory functions [2].

Proteins can undergo various post transcriptional modifications that function as switches or markers required for a large number of cellular events. Protein acetylation is a major class of PTM whose main target is the lysine residue. Acetyl groups are transferred from acetyl coenzyme A by acetyltransferase [3]. Lysine acetylation is thought to promote conformational changes and to control protein binding with or dissociation from other biomolecules, including proteins and nucleotides, by neutralizing positive charge. In this study, acetylation of lysine residues inside the RNA-binding domains (RRMs) of TDP-43 was evaluated in terms of its effect on the affinity of TDP-43 for RNA using computer simulation techniques. These results suggest that acetylation of lysine residues would have a large effect on association with RNA.

Results and discussion

The crystal structure of the TDP-43 (PDB code 4BS2) was taken from the protein data bank. MD simulations of the TDP-43 -RNA complex were performed using GROMACS with amber99SB force field. Each complex consisting of wild type or mutant TDP-43 was neutralized by the addition of the sodium counterions, Na, and soaked in solute molecules in a rectangular parallelepiped box filled with a periodic box of TIP3P water molecules. Potential energy was minimized to optimize the initial position of each atom. Long-range electrostatic interactions energy was calculated using the Particle Mesh Ewald method. The coordinates and energies of each atom were stored every picosecond (ps) for further analysis. At the end, the results suggested that a single acetylation event is sufficient to impair RNA binding, implicating Lys-145 as a critical determinant of TDP-43 function.

References

[1] Redler RL, Dokholyan NV. The Complex Molecular Biology of Amyotrophic Lateral Sclerosis (ALS). Prog Mol Biol Transl Sci. 2012; 107: 215-262.

[2] Hummel AV, Lee W, Ke Y. TDP-43 Mouse Models of Frontotemporal Dementia and Amyotrophic Lateral Sclerosis. *SBME*. 2015; 46 :12–15.

[3] Petrov1 D, Margreitter C, Grandits M, Oostenbrink C, Zagrovic B. A Systematic Framework for Molecular Dynamics Simulations of Protein Post-Translational Modifications. PLoS Comput Biol. 2013; 7:1-9.

QSAR Study of tri-substituted 1,2,4-triazoles as inhibitors of the annexin A2–S100A10 protein using the multiple linear regression model

Vahid Mohammadrezaei[2], Ahmad Mani-Varnosfaderani*[1] Mahmoud Ebrahimi*[2]

[1] *Department of Chemistry, Faculty of Sciences, Islamic Azad University, Mashhad Branch, Iran*

[2] *Department of Chemistry, Tarbiat Modares University, Tehran, Iran.*

E-mail address: a.mani@modares.ac.ir

Abstract

QSAR technique is used to relate one block of properties to another and is a tool for linking chemical activities with composition and molecular structures. Recently a set of tri-substituted1,2,4-triazoles as inhibitors of the annexin A2–S100A10 protein were synthesized. In this study predictive quantitative structure-activity relationships for a data set of 41 molecules of tri-substituted1,2,4-triazoles were developed. Multiple linear regression (MLR) model was employed to construct the QSAR model. For developing the model different descriptors were calculated using Dragon software. The obtained descriptors were programmed and filtered using MATLAB software to reach the least correlated descriptors. The results of cross validation revealed that the developed model is reliable and predictive and can be used for designing better and potent annexin A2–S100A10 protein inhibitors.

Keywords: QSAR , 1,2,4-triazoles derivatives , SPA_MLR , GA_MLR , Stepwise_MLR

Introduction

Annexin A2 (AnxA2) is a member of the larger annexin family of Ca2+ and phospholipid binding proteins [1–3]. In common with all annexins, AnxA2 contains a conserved C-terminal core domain, which confers Ca2+ and phospholipid binding properties, and a less conserved, smaller N-terminal domain. The N-terminal domain of AnxA2 consists of 30 amino acids of which the first 14 residues constitute the binding site for its typical binding partner S100A10, a member of the S100 protein family [4].

S100 proteins constitute a family of small globular adaptor proteins that regulate cell functions by virtue of their capacity to interact with protein binding partners. [5,6] S100 proteins are Annexins's important binding partners of [6] and they contain a short helical sequence feature at the N-terminus that allows binding to several members of the S100 family. The S100A10 protein and annexin A2 (AnxA2) form a classic pairing in this way. [7] The interaction between these proteins is very well characterised and shows similarities with the tractable protein interactions described above, in that a small helix docks into a deep well-defined binding crevice [8,9]. Both S100A10 and AnxA2 have been implicated in cell matrix invasion, cell movement and cell adhesion and as such playimportant roles in the regulation of therapeutically relevant processes such as vascular neo-angiogenesis and tumour cell metastastis.[10]

Results and discussion

As it is mentioned in the abstract 41 different molecules of tri-substituted1,2,4-triazoles were used for model development. The molecular descriptors for these molecules have been calculated and analyzed by three variable selection MLR techniques including stepwise-MLR, successive projection algorithm (SPA) and genetic algorithm (GA). As it is illustrated in the following table, for describing and predicting the tri-substituted1,2,4-triazoles inhibitory activities the best model is developed using GA MLR technique which yields the highest correlation coefficient .(table1).

Method	Training		L50		No. Comp/Var
	R^{2a}	RMSE	R^2	RMSE	
Stepwise_MLR	0.54875	0.3810	0.4691	0.5759	10
SPA_MLR	0.78336	0.2554	0.86154	0.3245	5
GA_MLR	0.9397	0.1413	0.9557	0.3067	10

[a] All R2 are adjusted coefficient regression.

Ten variables including nCL, X_4v, CSI, IDET ,HVopx ,T(F..Br) ,BEHm6 , ATS3m , ATS8e and MATS2e have been selected by GA_MLR and the optimized equation is :

$Y= -8.9139 + 58.6812\ nCL+ 5.3866X_4v – 0.0063CSI + 3.8732IDET + 1.6349HVopx +0.0958T(F..Br) + 0.2797BEHm6 – 9.5372ATS3m-2.1619ATS8e + 0.5254MATS2e$

References

[1] Gerke, V., Creutz, C.E. and Moss, S.E. (2005) Annexins: linking Ca2+ signaling to membrane dynamics. Nat. Rev. Mol. Cell Biol. 6, 449_461.

[2] Gerke, V and Weber, K. (1984) Identity of p36K phosphorylated upon Rous sarcoma virus transformation with a protein purified from brush borders; calcium-dependent binding to non-erythroid spectrin and F-actin. EMBO J. 3, 227–233.

[3] Crumpton, M.J. and Dedman, J.R. (1990) Protein terminology tangle. Nature 345, 212.

[4] Johnsson, N., Marriott, G. and Weber, K. (1988) P36, the major cytoplasmic substrate of src tyrosine protein kinase, binds to its p11regulatory subunit via a short amino-terminal amphiphatic helix. EMBO J. 7, 2435–2442.

[5] Salama, I. Malone, P. S. Mihaimeed, F. Jones, J. L. Eur. J. Surg. Oncol. 2008, 34,357.

[6] Santamaria-Kisiel, L.Rintala-Dempsey, A. C. Shaw, G. S. Biochem. J. 2006, 396,201.

[7] Gerke, V ,Weber, K. EMBO J. 1985, 4, 2917.

[8] Johnsson, N. Marriott, G. Weber, K. EMBO J. 1988, 7, 2435.

[9] Rety, S. Sopkova, J.; Renouard, M. Osterloh, D. Gerke, V. Tabaries, S. RussoMarie, F. Lewit-Bentley, A. Nat. Struct. Biol. 1999, 6, 89.

[10] Ling, Q. Jacovina, A. T. Deora, A .Febbraio, M.Simantov, R. Silverstein, R. L. Hempstead, B. Mark, W. H. Hajjar, K. A. J. Clin. Invest. 2004, 113, 38.

[11] Tummala R. K. Reddy , Chan Li, Xiaoxia Guo, Peter M. Fischer, Lodewijk V. Dekker . Design, synthesis and SAR exploration of tri-substituted 1,2,4-triazoles as inhibitors of the annexin A2–S100A10 protein interaction . Bioorganic & Medicinal Chemistry ;2012;49 ;5378–5391

Diagnosis of brain glioma using near-infrared micro-spectroscopy comparing partial least squares discriminant analysis and target factor analysis

M.Khorani, M.Khanmohammadi*, K.Ghasemi A. Bagheri Garmarudi, M. Babaei Rouchi

m.khanmohammadi@sci.ikiu.ac.ir

Abstract

In the present study it was tried to develope a rapid approach based on attenuated total reflectance – Fourier transform near infrared (ATR-FT-NIR) micro-spectrometry which can be useful for diagnosing of the healthy tissue smaples from malignant ones. A total of 101 brain samples, including 85 gliomas and 16 normal cases were studied in 8000–4000 cm^{-1} spectral region.The preprocessing methods such as multivariate wavelet denoising were utilized by the related software for all spectra. Two procedures including partial least squares-discrimination analysis (PLS-DA) as a hard modeling method and target factor analysis (TFA) as a soft modeling method were applied to classify the brain tissues.Several figure of merits were calculated to evaluate the performance of methods. The classification results shows that combination of FT-NIR spectroscopic data with chemometric techniques is a reliable route for classifying of the normal and glioma tumors.The spectra were classified into glioma and normal tissue classes with accuracy of 94% and 96% for PLS-DA and TFA models, respectively. Finally, the obtained results were validated with gold standard and showed good reliability for medical diagnosis purposes.

Key words: Glioblastoma, Near Infrared Microspectroscopy, Multivariate wavelet denoising, PLS-DA, TFA.

Introduction

Glioblastoma is one of the most aggressive types of human cancers which form in the brain owing to changes in the amount of water, protein, lipid, nucleic acid and carbohydrate [1].At the beginning of the tumor formation, many cancer cells are in the molecular levels which are not visually identifiable by traditional methods of pathological detection[2]. As a result, the cancer cells cannot be recognized; therefore, the survival rate of the patients will decrease [3]. To overcome the common limitations of pathological diagnosis, the near infrared micro-spectroscopy as an analytical technique is used based on the spectrochemical analysis of tissues [4]. However, the NIR spectra require suitable diagnostic models to discriminate between glioma and normal tissues [5]. In this research, a pattern recognition method using PLS-DA and also TFA that is a factor analysis method were applied to examine the combination of NIR spectroscopy and chemometrics methods as a potential tool for distinguishing between normal and malignant brain tissues.

Results and discussion

Multivariate wavelet denoising was applied for same areas in the range of 4000 to 8000 cm^{-1}. However, the spectra had overlapping, weak and broad peaks without distinct signature of individual components. Consequently, the biological origins of the spectral differences were difficult to be entirely addressed and the discrimination of normal and cancer tissues was so hard. Since this method did not yield a satisfactory outcome, PLS-DA and TFA methods as chemometrics methods were used to classify the samples. In this research, the original dataset was splitted as subsets of 50 training and 51 test set by using Kennard algorithm.For doing PLS-DA, threshold value were determined for classes after modeling the data.After that for all of the samples calculated responses were gained.Samples which had the calculated responses more than threshold value, were assigned to the first class and ones which had less value were assigned to second class.The results show that 3 out of 50 samples of training set and 5 out of 51 samples of test set were misclassified. To estimate a better classification model compared with PLS-DA technique, TFA as a soft modeling method was used based on the groups properties. In this method at first 85 normal samples considered as target set and subsequently PCA were done on it and its scores and loading were extracted .Then, cancerous samples were used as test set and were projected on the target set vector space. Distance and angle values of test set toward vector space of target set were calculated and demonstrate that cancerous and normal samples have significant differences in these values.The results of Accuracy, sensitivity and correct classified samples of PLS-DA and TFA methods are shown in table.1.

Table1. Figures of merit for PLS-DA and TFA methods

Parameters	PLS-DA	TFA
correct classified samples	92%	94%
Sensitivity	97%	100%
Accuracy	94%	96%

This is noticeable that Soft modelings methods focus on spectral vectors based on the characteristics of each groups. As a result, soft modeling method can discriminate features of the groups better than hard modeling metod during the survey of tomur spectral data that have many overlappings. The data in Table.1. Show that the use of TFA as a soft modeling method reports a better state compared to PLS-DA as a hard modeling method for separating healthy tissues from glioma ones.Also the results show that using of the FTNIR microspectroscopy and chemometrics methods are useful ways for assessing the data of the brain tissues spectra.

References

[1]J. R. Hands , P. Abel , K. Ashton, et al.Investigating the rapid diagnosis of gliomasfrom serumsamples using infrared spectroscopy and cytokineand angiogenesis factors. Anal Bioanal Chem ,2013;23:7347-7355

[2]F. M. Lyng,R. M. Ramos,O. Ibrahim , and H. J. Byrne.Vibrational Microspectroscopy for Cancer Screening.Appl. Sci.2015; 5, 23-35.

[3]M. Preusser, S. de Ribaupierre,A.Wohrer, Sara C,et al. Current Concepts and Management of Glioblastoma. Neurological Progress,2011;70:9-21.

[4]E. Gazi, J. Dwyer, P. Gardner, A. Ghanbari-Siahkali, AP. Wade,et al, Applications of Fourier transform infrared microspectroscopy in studies of benign prostate and prostate cancer. A pilot study, J Pathol,2003; 201: 99–108.

[5]A. A. Bunaciua, V. D. Hoangb , H. Y. Aboul-Eneinc,Applications of FTIR Spectrophotometry in Cancer Diagnostics, Critical Reviews in Analytical Chemistry (2015) 904733, 156-165.

Improvement of the patial least squares model for simultaneous determination of dyes by Orthogonal signal correction

Ali Niazi[a],Saba Ahmadi[b]

a Department of Chemistry, Faculty of Sciences, Azad University of Arak, Arak, Iran
b Yong Researchers Club, Azad University of Azad, Arak, Iran
E-mail address (saba ahmadi): s.ahmadi6586@yahoo.com

Abstract

The simultaneous determination of dyes mixtures by using spectrophotometric methods is a difficult problem in analytical chemistry, due to spectral interferences. By multivariate calibration methods, such as partial least squares (PLS), it is possible to obtain a model adjusted to the concentration values of the mixtures used in the calibration range. Orthogonal signal correction (OSC) is a preprocessing technique used for removes the information unrelated to the target variables based on constrained principal component analysis. OSC is a suitable preprocessing method for partial least squares calibration of mixtures without loss of prediction capacity using spectrophotometric method. In this study, the calibration model is based on absorption spectra in the 300–600 nm range for 21 different mixtures of dyes. Calibration matrices were containing 0,5–35, 0,5–15.5 and 0,5–30.5 μg ml−1 of annatttos, tartrazine and turmeric, respectively. The RMSEP for annatto, tartrazine and turmeric with OSC and without OSC were 0.6567, 0.2692, and 0.3134, and 1.3818, 1.2181, and 0.3953, respectively. This procedure allows the simultaneous determination of dyes in real matrix samples and good reliability of the determination was proved.

- **Key words:** Annatto,Tartrazine,Turmeric, Determination; Spectrophotometric; PLS; OSC

Introduction

C16H9N4Na3O9S2 a chemical compound with the formula with the chemical formula ID is 6321403 . The molar mass of 534. 3 g / mol is .Tartrazine Rngknndh a chemical substance , the orange or lemon yellow food color additives that are used.(Turmeric or turmeric) Scientific name : Curcuma longa) is a spice plant and that the plant is prepared . Category ginger turmeric walls (Zingiberidae), Cut ginger (Zingiberales), yellowe color Annatto The color is produced from the seeds of the annatto plant the seeds of the annatto plant is available .

The basic concept of PLS regression was originally developed byWold and application of PLS in spectrometry have been discussed by several workers. In addition, several multicomponent determinations based on the application ofthese methods to spectrophotometric data have been reported.Orthogonal signal correction (OSC) was introduced byWold et al. to remove systematic variation from the response matrix*X*that is unrelated, or orthogonal, to the property matrix *Y*.Therefore, one can be certain that important information regarding the analyte is retained. Since then, several

groups have published various OSC algorithms in an attempt to reducemodel complexity by removing orthogonal components from the signal.

Results and discussion

This procedure allows the simultaneous determination of dyes by PLS and OSC-PLS in real samples such as ice cream with good reliability of the determination was proved.

Determination of Annatto, Tartrazine and Turmeric by OSC-PLS

Sample	Add			Prediction			Recovery (%)		
	Annatoo	Tartrazine	Turmeric	Annatto	Tartrazine	Turmeric	Annatto	Tartrazine	Turmeric
M1	4.0	8.0	10.0	3.9	8.0	10.1	98.8	100.3	101.2
M2	8.0	10.0	5.0	7.9	9.8	5.0	99.0	98.6	101.2
M3	20.0	5.0	8.0	20.1	5.0	8.0	100.6	100.4	100.8
M4	3.0	3.0	3.0	3.0	2.9	2.9	100.7	98.7	99.3
M5	10.0	10.0	10.0	10.0	10.1	9.9	100.5	101.2	99.6
M6	7.0	6.0	15.0	7.0	6.0	14.9	100.7	100.3	99.8
NF				3.0	3.0	3.0			
PRESS				0.1001	0.0091	0.0008			
RMSEP				0.0367	0.0611	0.0530			
RSEP(%)				07806	0.8215	0.7299			

References

[1] A. Niazi, J. Ghasemi, A. Yazdanipour, Anal. Lett. 38 (2005) 2377.

[2] A. Niazi, J. Ghasemi, M. Zendehdel, Talanta 74 (2007) 247.

Application of genetic algorithms for pixel selection in multivariate image analysis for a QSAR study of Inhibition of the growth of tumor cell lines for novel antitubulin agents

Masoomeh Aghaei [1,*], Ali Niazi[1], Ali Yeganeh Faal[2],

[1]*Department of Chemistry, Faculty of Science, Arak Branch, Islamic Azad University, Arak, Iran*
[2] *Department of Chemistry, Payam Noor University, Qom, Iran*

M.Aghaei2003@gmail.com

Abstract

Quantitative structure-activity relationship (QSAR) analysis has been directed to a series of 40 compounds of 2,6-dinitro-4-(trifluoromethyl)phenoxysalicylaldoxime and novel chalcone derivatives as novel antitubulin agents with Inhibition of the growth of tumor cell lines that was performed by chemometrics methods.Bidimensional images were used to calculate some pixels. Multivariate image analysis was applied to QSAR modeling of the Inhibition of the growth of tumor cell lines of 2,6-dinitro-4-(trifluoromethyl)phenoxysalicylaldoxime and novel chalcone derivatives by means of multivariate calibration such as principal component regression (PCR) and partial least squares (PLS). In this study we investigate the effect of pixel selection by application of genetic algorithms (GAs) for PLS model. GAs is very useful in the variable selection in modeling and calibration because of the strong effect of the relationship between presence/absence of variables in a calibration model and the prediction ability of the model itself. The subset of pixels, which resulted in the low prediction error, was selected by genetic algorithm. The resulted model showed high prediction ability with RMSEP of 0.0783, 0.0462 and 0.0062 for PCR, PLS and GA-PLS models, respectively. Furthermore, the proposed QSARmodel with GA-PLS was used for modification of structure and their activity predicted.

Key words: antitubulin, QSAR, Multivariate Image Analysis, PCR, genetic algorithm, PLS

Introduction

Quantitative structure–activity relationship (QSAR) as one of the most important areas in chemometrics gives information useful for pharmaceutical chemistry, drug design, toxicology and eventually most facts of chemistry, and for this reason, several investigations have been carried out in order to improve the results [1–3]. Multivariate calibration such as PCR and PLS is a method that can be useful in dealingwith the problem of the unfavorable

more variable/object ratio and collinearity [4,5]. The PLS theory and its application in QSAR are reported by several of the workers [5-9]. A GA is a stochastic method to solve optimization problems defined by a fitness criteria applying the evolution hypothesis of Darwin and different genetic functions, i.e., crossover and mutation. Leardi [10] demonstrated that GA, after suitable modifications, produces more interpretable results since the selected variables are less dispersed than with other methods. Nowadays, image analysis is becoming more important because of its ability to perform fast and non-invasive low-cost analysis on different processes in chemistry. Image analysis is a wide denomination that encloses classical studies on gray scale or (red–green–blue) RGB images [11]. Esbensen and Geladi [12] have

demonstrated that image analysis may provide useful information in chemistry; the descriptors do not have a direct physicochemical meaning since they are binaries. In QSAR, images (2D chemical structure) have shown to contain chemical information [13,14], allowing the correlation between chemical structures and properties. The present study is focused on the application of 2D images, which are the proper structures of the compounds that can be drawn with aid of any appropriate program, as descriptors in QSAR. Then, multivariate image analysis-quantitative structure activity relationship study (MIA-QSAR) is proposed to model and predict the Inhibition of the growth of tumor cell lines as antitubulin activity of a series of 2,6-dinitro-4-(trifluoromethyl)phenoxysalicylaldoxime and novel chalcone derivatives by genetic algorithm-partial least squares (GA-PLS) modeling method.

Results and discussion

Using GA-PLS, a QSAR model for the prediction of Inhibition of the growth of tumor cell lines for 40 compounds based multivariate image analysis alone have been developed. The results well illustrate the power of pixels in prediction of Inhibition of the growth of tumor cell lines of a series of 2,6-dinitro-4-(trifluoromethyl)phenoxysalicylaldoxime and novel chalcone derivatives. The model could predict the Inhibition of the growth of tumor cell lines of 2,6-dinitro-4-(trifluoromethyl)phenoxysalicylaldoxime and novel chalcone derivatives not existed in the modeling procedure accurately. The results of this study clearly show the potential and versatility of GA-PLS modeling in QSAR study of Inhibition of the growth of tumor cell lines of 2,6-dinitro-4-(trifluoromethyl)phenoxysalicylaldoxime and novel chalcone derivatives using multivariate image analysis, which could be applied to prediction of Inhibition of the growth of tumor cell lines.

References

[1] B. Hemmateenejad, R. Miri, M. Akhond, M. Shamsipur, Chemom. Intell. Lab. Syst. 64 (2002) 91–99.

[2] J. Mon, M. Flury, J.B. Harsh, J. Hydrol. 316 (2006) 84–97.

[3] A. Niazi, S. Jameh-Bozorghi, D. Nori-Shargh, Turk. J. Chem. 30 (2006) 619–628.

[4] A. Niazi, S. Jameh-Bozorghi,D.Nori-Shargh, J.Hazard.Mater. 151 (2008) 6063–6606.

[5] P. Geladi, B.R. Kowalski, Anal. Chim. Acta. 185 (1986) 1–17.

[6] E.V. Thomas, D.M. Haaland, Anal. Chem. 62 (1990) 1091–1099.

[7] D.M. Haaland, E.V. Thomas, Anal. Chem. 60 (1988) 1193–1202.

[8] D.M. Haaland, E.V. Thomas, Anal. Chem. 60 (1988) 1202–1208.

[9] Y. Liang, D. Yuan, Q. Xu, J. Chemom. 22 (2008) 23–35.

[10] R. Leardi, Genetic algorithms in feature selection, in: J. Devillers (Ed.), In Genetic Algorithms in Molecular Modeling, Academic Press, London, 1996.

[11] J.M. Prats-Montalban, A. de Juan, A. Ferrer, Chemom. Intell. Lab. Syst. 107 (2011) 1–23.

[12] K. Esbensen, P. Geladi, J. Chemom. 3 (1989) 419–429.

[13] K. Esbensen, P. Geladi, H. Grahn, Chemom. Intell. Lab. Syst. 14 (1992) 357–374.

[14] M.P. Freitas, S.D. Brown, J. Martins, J. Mol. Struct. 738 (2005) 149–154.

Application of mean centering of ratio spectra for dye degradation study using Fenton reaction

Morteza Bahram*,1 , Mostafa Aran Ardabili[1], Masoumeh Hasani[2], Siavash Bahari[2]

[1]*Department of Chemistry, Faculty of Science, Urmia University, Urmia, Iran*

[2]*Department of Analytical Chemistry, Faculty of Chemistry, Bu Ali Sina University, Hamadan, Iran*

E-mail m.bahram@urmia.ac.ir

Abstract

In this work an application of mean centering of ratio spectra is proposed to the study of pollutant dyes degradation using Fenton reaction. Two-way kinetic– spectrophotometric data related to dye(s) mixtures degradation are divided to simulated kinetic profiles, mean centered and their ranks were monitored versus simulated dye degradation constant. For accurate dye degradation constant(s) the rank of resulted matrix will be reduced by one. By this search all kinetic constants related to all included dyes are acquired. In order to find out which k values is related to the degradation of dye of interest, a preliminary deflation of data from its contribution is done and the procedure is repeated. By this search the kinetic constant related to dye of interest will be omitted while others will be remained. The applicability of the proposed method was evaluated by using model data. Finally second order kinetic-spectrophotometric data in the degradation of malachite green in the presence of crystal violet and methylene blue were analyzed by the proposed method as a real system.

Key words: Mean centering of ratio spectra, two-way data, Fenton Reaction, degradation, Toxic Dyes

Introduction

Developing of an effective and reliable analytical method to monitor the dye concentrations is necessary [1, 2]. In dye degradation studies through advanced oxidation processes, spectrophotometric techniques are widely used as the monitoring technique because of the cheapness of the instrument, easy interpretation and handling of the spectral data. But, traditional univariate spectrophotometry can only be performed if there are no interferent or there is selective wavelength range for the analyte of interest [3].

In particular, mean centering of ratio spectra is used to remove the contribution of absorbing known components from data matrix exactly [4-8]. Mean centering of ratio spectra have been presented by Afkhami and Bahram [3] applied for the simultaneous analysis of binary and ternary mixture analysis [5–8]. Also recently mean centering of ratio spectra was applied to resolve two-way kinetic–spectral data of consecutive reactions and acquire the rate constants and the spectrum of intermediate component [9].

In this work an application of mean centering of ratio spectra is used to resolve the two-way kinetic-spectrophotometric data of simultaneous dyes degradation. In comparison to multivariate curve resolution (MCR) method unique results were acquired using this procedure.

Results and discussion

The results of applying the proposed method to estimate the k-values of degradation of dye one (arbitrarily) in the presence of unknown dyes 2-3 are shown in Fig. 1. As can be concluded from this fig the k-values eques 0.02 has been omitted for deflated data. Therefore the k value of degradation kinetic of dye one is estimated as 0.02.

To investigate the effects of overlapping degree of kinetic profiles, a series of kinetic profiles using different k-values were simulated. The simulated k- values and the predicted k-value for dye one as component of interest are shown in table 1.

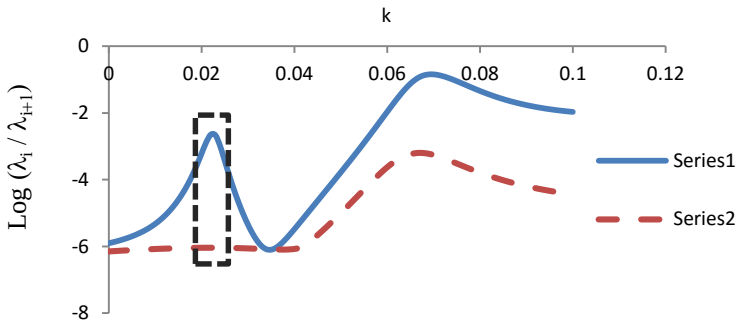


Fig. 1

Table 1. The simulated data and the results obtained by applying the proposed method

	Simulated			Predicted k_1	Predicted k_1 (noisy)			RSE%[a]		
No.	k_1	k_2	k_3	(without noise)	0.001	0.003	0.01	0.001	0.003	0.01
1	0.02	0.05	0.08	0.02	0.0200	0.0194	0.0187	0.00	-2.86	-6.50
2	0.01	0.05	0.08	0.01	0.0110	0.0101	0.0092	10.00	0.72	-8.00
3	0.03	0.05	0.07	0.03	0.0306	0.0280	0.0270	2.01	-6.61	-10.00
4	0.04	0.07	0.08	0.04	0.0403	0.0422	0.0418	0.68	5.54	4.38
5	0.01	0.07	0.9	0.01	0.0096	0.0097	0.0093	-4.00	-3.12	-7.00

References

[1] C. Fernández, M.S. Larrechi, M.P. Callao, Anal. Chem. 29 (2010) 1202.

[2] B. Maheri Kalejahi, M. Bahram, A. Naseri, S. Bahari, M. Hasani, J IRAN CHEM SOC, 11 (2014) 241.

[3] A.K. Gupta, A. Pal, C. Sahoo, Dyes Pigm. 69 (2006) 224.

[4] A. Afkhami, M. Bahram, Talanta 66 (2005) 712.

[5] A. Afkhami, M. Bahram, Anal. Chim. Acta 526 (2004) 211.

[6] A.Afkhami, M. Bahram, Talanta 68 (2006) 1148.

[7] A. Afkhami, T. Madrakian, M. Bahram, J. Haz. Mater. 123 (2005) 250.

[8] M. Bahram, Anal. Chim. Acta , 603 (2007) 13.

[9] M. Bahram, M. Mabhooti, Anal. Chim. Acta, 639 (2009) 19.

Target transformation fitting for simultaneous dye degradation using Fenton reaction

Morteza Bahram*,[1] , Rasol Gheytaran[1], Abdolhosein Naseri[2]

[1]*Department of Chemistry, Faculty of Science, Urmia University, Urmia, Iran*

[2]*Department of Analytical Chemistry, Faculty of Chemistry, Tabriz University, Tabriz, Iran*

E-mail m.bahram@urmia.ac.ir

Abstract

In this work, for the first time, Target Transformation Factor Analysis (TTFA) was applied to study the simultaneous removal of dyes by the use of Fenton reaction. The presented method was evaluated by studying the simulated data. The effect of three parameters in the simultaneous removal of three carcinogenic, mutagenic and toxic organic dyes were studied. The studied parameters were: concentration of Fe^{2+}, concentration of H_2O_2 and initial pH of the reaction media. Three studied dyes were: malachite green, crystal violet and methylene blue. First, by applying Central Composite Experimental Design 18 samples were designed which involve different conditions of the mentioned factors. The spectrophotometric- kinetic two way data were recorded and studied. The UV-Vis spectra of mixture of three dyes during degradation time (for 15 min with 30 seconds increments) were recorded. Using Target Transformation Factor Analysis (TTFA) the data were resolved and the concentration profile of three dyes despite of extreme spectral overlap were attained. Rate constant values of three dyes was computed and was used as responses of Central Composite Experimental Design to build a model. For each of three dyes, response surfaces was plotted separately and the effective factors of degradation and their optimum values were computed for each dye at the range of study.

Key words: Target Transformation Factor Analysis (TTFA). Fenton reaction. Central Composite Experimental Design. Response Surface

Introduction

Synthetic organic dyes are widely used in different industries such as textile dyeing, cosmetic and food [1]. Depletion of wastewaters of such industries at the environment is highly problematic. There are several ways to remove or treat these dangerous materials from wastewaters such as adsorption, biological degradation, electro chemical process, and advanced oxidation processes (AOP). AOPs are processes based on the generation of the hydroxyl radical (·OH), which

is a strong oxidant to destruct a wide variety of organic dyes. A special kind of AOP is the Fenton reaction, in which the hydroxyl radical is formed from the reaction between H_2O_2 and Fe (II) ions [2-4]. According to sources, degradation reaction of organic dyes by the Fenton reaction is from a first order reaction.

TTFA replaces the principal component axes by real factors with chemical meaning. This is done after testing individually a pool of potential real factors (targets) which may be useful to build a chemical model for the data set [5, 6].

A=CE → A=SL → $A=STT^{-1}L=CE$ → C=ST → $E=T^{-1}L$ A: Data Matrix S: Score Matrix L: Loading Matrix T: Transformation Matrix C: Concentration Profile E: Spectral Profile

Results and discussion

Application of TTFA for analyzing and resolving the Uv-Vis data is a reliable method for extracting of quantitative information of monitoring of simultaneous dyes degradation despite of drastic spectral overlap and presence of possible interferences. Important considerable instances in Fenton reaction are Fe^{2+} and H_2O_2 concentration and pH effect.

18 experiments (14 + 4 central point relicates) were designed to study the effect of 3 mentioned parameters on the simultaneous degradation of malachite green, crystal violet and methylene blue. The two-way kinetc spectrophotometric data was recorded for each sample. The results were analyzed by TTF. A case data and results of application of TTF to find out the degradation rate constants (k values) are represented here.

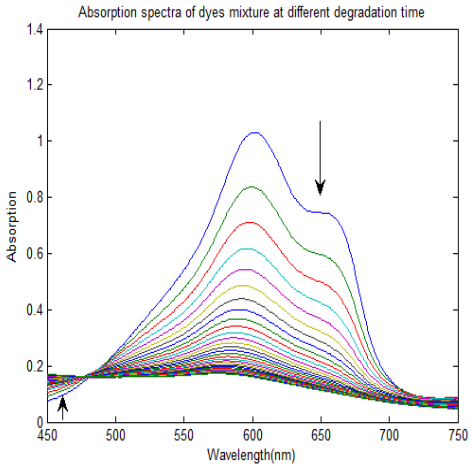


Fig. 1. Recorded spectra during Fenton reaction for an arb experiment,

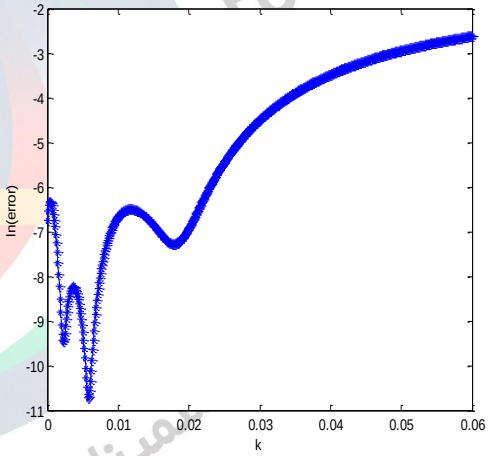


Fig. 2. Diagram of Ln(Error) against constant rate(k)

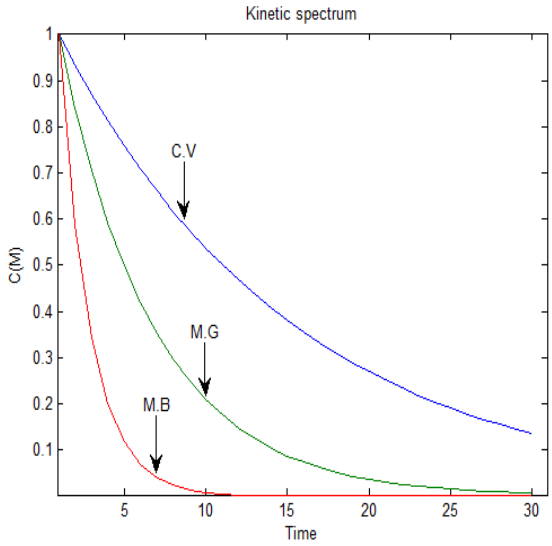


Fig. 3. Concentration profiles of each dye.

After collecting all degradation constants for each conditions, the regression model was build and using the response surface plot , the optimum conditions were acquired. In all cases there is an increase in degradation value with increase of Fe^{2+} and H_2O_2 and pH, but after an appointed limit degradation value decreases. This is maybe because of hydroxyl radical generation value and Fe^{2+} solubility: With increasing the pH to basic range, Fe^{2+} is converted to $Fe(OH)_3$ and deserts the reaction medium by sedimentation. Increasing the Fe^{2+} and H_2O_2 concentration both increase the hydroxyl radical generation, but after appointed limit, it seems these radicals involve into sideway Reactions. Anyway using Fenton Reaction to remove organic materials specially organic dyes from industrial wastewaters is an effective, simple, low cost and clean method. Using of Experimental Design, compared to one variable at the time optimization helps us to reduce the expense and time of experimentations by reducing the number of experiments and also clear the Important Factors and variables and optimum conditions of Reaction.

References

1. H. Zollinger, Color Chemistry. Synthesis, Properties and Applications of Organic Dyes and Pigments, 3rd edn (VCH, New York, 1991)
2. E.G. Garrido-Ramirez, B.K.G. Theng, M.L. Mora, Appl. Clay Sci. 47, 182 (2010)
3. J.J. Pignatello, E. Oliveros, A. MacKay, Environ. Sci. Technol. 36, 1 (2006)
4. E. Neyens, J. Baeyens, J. Hazard. Mater. 98, 33 (2003)
5. Target testing procedure for determining chemical kinetics from spectroscopic data , Chemometrics and intelligent lab. Syst. 50, 2000, Page 63-73.
6. Quantitation of spectroscopic systems using computer assisted target factor analysis, Freswneius J. Anal. Chem., 369,2001, Page 124-133.

Application of Central Composite Design for Optimization of Coacervative Extraction of Cu (II) Using Anionic Surfactant

Morteza Bahram,* Loghman Shokri, Naimeh Mohseni

Department of Chemistry, Faculty of Science, Urmia University, Urmia, Iran

E-mail address: m.bahram@urmia.ac.ir

ABSTRACT

The aim of this work was to develop a new and simple coacervative extraction method for the preconcentration and spectrophotometric determination of Cu (II) in water samples. Dithizone was used as the chelating agent while an anionic surfactant, namely sodium dodecyl sulfate (SDS), was used as extracting agent at room temperature. Central composite design (CCD) based response surface method (RSM) was employed to optimize main experimental parameters such as pH values of solutions, concentration of the surfactant and volume of the saturated salt. Analytical characteristics of the method such as limit of detection, linear range, relative standard deviation (R.S.D.) and relative standard error (R.S.E.) were calculated under the optimum conditions. The calibration graph was linear in the range of 10-100 ng mL^{-1} with the detection limit of 3.8 ng mL^{-1}. The interference effects of some anions and cations were also studied. The method was successfully applied for the determination of Cu (II) in water samples.

***Keywords*:** Coacervative extraction; Cu (II); Dithizone; SDS; Central composite design

Introduction

Copper is a vital nutrient mineral, essential for the health, growth and reproduction of all higher plants and animals. It is widely released into the environment by both natural sources and human activities [1]. Copper has been more considered by its applications in metallurgical and chemical industries; however, it is one of the essential microelements for human body [2,3], so that copper deficiency can cause impaired growth in children while it helps detecting a number of metabolic problems in adults. Nevertheless, once copper level exceeds a certain value, immune mechanism will be activated protecting the body against excess copper.

Nonionic surfactants, such as Triton X-100 and Triton X-114, have been widely used as the extracting agents for various compounds [4]. However, CPE has to be carried out at high temperatures (sometimes >70°C) and this may affect stability of the compounds of interest.

It should be pointed out that CPE involves the phase separation of neutrally charged (nonionic or zwitterionic) surfactants induced by the temperature while coacervative extraction involves phase separation of ionic amphiphiles induced by other parameters (e.g., addition of electrolytes, pH change, addition of organic co-solvents, or simple mixing of oppositely charged amphiphiles). Coacervative extraction can be performed at a low temperature while providing a higher rate of recovery and lower limit of detection compared to the previous methods [5].

In the present work, an anionic surfactant, namely sodium dodecyl sulfate (SDS), was used as an extracting agent at room temperature.

Results and discussion

CCD was used to optimize effective experimental parameters on preconcentration and determination of Cu (II). Three independent factors, namely pH (F1), SDS concentration (F2) and added volume of saturated KCl (F3) were studied at five levels with four repeats at the central point, using a circumscribed central

composite design. For each of the three studied variables, high (coded value: +1) and low (coded value: -1) set points were selected as shown in Table 1**.**

Table 1. Variables used for central composite design along with their values.

		Coded factor levels				
Variable name		**-1 (low)**	**-0.59**	**0**	**+0.59**	**+1 (high)**
F1	pH	1	1.81	4.2		5
F2	SDS (5% w/v) (mL)	0	0.1	0.25	0.4	0. 5
F3	Saturated KCl (mL)	0	0.24	0.6	0.96	1.2

Response Surface Methodology (RSM) is useful for modeling while the response of interest is inuenced by several variables and the objective is to optimize this response. Three-dimensional (3D) graphs provide insight about the effect of each variable in order to achieve peak performance. Plots for the predicted responses were prepared based on the model function to analyze the change in response surface. Some of the response surface plots are represented in Figure 1, which illustrates the relationship between two variables and the absorbance (510 nm) of samples, while two other variables were fixed at the central point.

Figure 1. Response surface plots.

References

[1] Şatıroğlu N., Arpa Ç., Cloud point extraction for the determination of trace copper in water samples by flame atomic absorption spectrometry, *Microchim Acta*., 2008; **162**: 107-112..

[2] Ulusoy H.I., Gürkan R., Akçay M., Kinetic spectrophotometric determination of trace copper(II) ions by their catalytic effect on the reduction of brilliant cresyl blue by ascorbic acid, *Turk J Chem*., 2011; **35**: 599-612. Shokrollahi A., Ghaedi M., Gharaghani S., Fathi M.R., Soylak M., Cloud point extraction for the determination of copper in environmental samples by flame atomic absorption spectrometry, *Quim. Nova.*, 2008; **31**: 70-74..

[3] Yurtman-Gündüz Z., Aydın-Urucu O., Kök-Yetimoğlu E., Filik H., Cloud point extraction for the determination of trace copper(II) in environmental samples using a new synthesized reagent and flame atomic absorption spectrometry, Maced. J. Chem. Chem. Eng., 2013; **32**: 69-78.

[4] Purkait M.K., Banerjee S., Mewara S., Das Gupta S., De S., Cloud point extraction of toxic eosin dye using Triton X-100 as nonionic surfactant, *Water Res.*, 2005; **39**: 3885-3890.

[5] Bahram M., Khezri S., Multivariate optimization of cloud point extraction for the simultaneous spectrophotometric determination of cobalt and nickel in water samples, *Anal. Methods*, 2012; **4**: 384-393..

Elicitation of the most important structural properties of volatiles affecting electroantennogram responses in *Aphis fabae*; a QSPR approach

Davoud Beiknejad, Mohammad Javad Chaichi*, Mohammad Hossein Fatemi

Faculty of Chemistry, University of Mazandaran, Babolsar 47416-95447, Iran

*E-mail: jchaichi@yahoo.com

Abstract

A quantitative structure-property relationships based approach to correlate the electroantennographic responses with molecular descriptors was developed using a newly developed software. A data set of 29 compounds with known electroantennographic response was used. Statistically validated quantitative structure-property relationships model has been developed using genetic algorithm-multiple linear regression technique for a given training data set. Descriptors which were used by this model are: BCUT (BELm6, BEHv6), 2D autocorrelations (GATS2v), and functional group counts (nOH). The statistical characteristics for the developed QSARINS model are the following: R2= 0.9181, Q2loo= 0.8628, Q2lmo= 0.8595, SE= 0.0788, F= 53.2685. An attempt was also made to interpret the physical meaning of the descriptors.

Keywords: Bioelectrochemical response; Electroantennogram; GA-MLR; QSPR; *Aphis fabae*; VOCs

Introduction

Volatile organic compounds (VOCs) play an important role in nature as messenger compounds to transmit selective information between species. Insects use VOCs for mating and to find host plants. In general, electroantennography (EAG) is the most widely used method reflecting the insect volatile reception [1]. *An electroantennograph is a bioelectrochemical device* that exploits metabolic activities of insect antennae for generation of electric potential. EAG is, however, tedious, expensive, labor-intensive and relatively time-consuming. An efficient alternative way of obtaining information to understand the mechanisms behind the studied systems without the need to perform expensive laboratory experiments is application of quantitative structure–property relationship (QSPR) techniques. Black bean aphid (*Aphis fabae*) is one of the most polyphagous and widely distributed aphid species. It attacks nearly 120 plant families. The aim of the present work was to study the potential of QSARINS, a newly developed QSAR software based on GA-MLR, to construct a model and to predict the EAG amplitudes of *Aphis fabae* from molecular structural descriptors.

Results and discussion

The volatile compounds and the corresponding normalized EAG data used in this study were selected from literature [2]. The data set (n=29) was divided randomly into two groups: training set (n=24) and test set (n=5). The QSPR models were built using the QSARINS (QSAR INSubria) software version 2.1[3]. With the selected descriptors, we built a linear model using the training set data, and the following equation was obtained:

$$EAG = -0.3426 + 0.735\ BELm6 - 0.5945\ BEHv6 + 1.2357\ GATS2v - 0.1349\ nOH \quad (1)$$

R^2_{train}= 0.9181, R^2_{adj}= 0.9009, $RMSE_{train}$= 0.0701, SE= 0.0788, F= 53.2685, Q^2_{loo}=0.8628, Q^2_{lmo}= 0.8595, R^2_{ext}= 0.8693, $RMSE_{ext}$= 0.184

The selected variables demonstrate that BCUT (*BELm6*, *BEHv6*), 2D autocorrelations (*GATS2v*), and functional group counts (*nOH*) descriptors affect the EAG amplitudes of the studied compounds. The model was also used to estimate the electroantennographic amplitudes of new 29 volatile compounds which their EAG values were not been reported experimentally. In summary the QSPR model allows us to refocus the role of *in-silico* methods in electroantennographic studies. The authors believe that developed QSPR model can be used to predict response of insect's olfactory receptors to new target chemicals. Such novel strategies based on a molecular understanding of insect chemosensation, may also contribute to the control of insects that cause enormous harm to human health and agriculture.

References

[1] Ockenfels, P, Van der Pers, J. Electroantennography a Practical Introduction, Syntech, Kirchzarten, 2004, http://www.syntech.nl/manuals/EAGpract.pdf.

[2] Park KC, Hardie J. An improved aphid electroantennogram. *J. Insect Physiol*. 1998; **44**: 919–928.

[3] Gramatica P, Chirico N, Papa E, Cassani S, Kovarich S. QSARINS: A new software for the development, analysis, and validation of QSAR MLR models. *J. Comput. Chem*. 2013; **34**:2121-2132.

Structural analysis of ZSM-5 nano zeolite by diffuse reflectance near-infrared spectrometry and chemometrics

Sh. Amani, M. khanmohammadi*, A. Bagheri Garmarudi

Department of Chemistry, Faculty of Science, Imam Khomeini International University, Qazvin, Iran.
E-mail address : m.khanmohammadi@sci.ikiu.ac.ir

Abstract

The main objective of this research is to use near-infrared diffuse reflectance method to evaluate and quantify specific surface area and pore volume of ZSM-5 zeolie samples. In order to achieve this goal, 61 different ZSM-5 zeolite samples by 3 different classes of surface area and degree of prosity were analyzed using FT-NIR spectrometry. By splitting the spectral data in training and independent test sets based by Kenard-Stone sampling algorithm, two classification modesl were built using least squares support vector machines (LS-SVM) and soft independent modeling of class analogy (SIMCA) chemometric techniques. The models were applied for the prediction of 21 samples from the independent test set. The percent of correctly classified samples using SVM and SIMCA methods were 95.2 % and 90.5 %, respectively.

Key words: ZSM-5 zeolite, Surface area, porosity, SVM and SIMCA techniques, DRIFT spectrometry.

Introduction

Aluminosilicate H-ZSM-5 zeolite has been used in wide range of chemical processes, such as xylene isomerization [1], fluid catalytic cracking [2], methanol to gasoline conversion [3], and is known as useful candidate catalyst for methanol to propylene (MTP) process [4-5]. It has a medium size pore system with two channels, both including 10-ring. Straight channels (5.1- 5.5 Å) are intersected by zigzag channels (5.3- 5.6 Å), thus creating a three-dimensional network [6]. DRIFT spectroscopy is a powerful tool for the study of materials such as catalysts and powders. This study investigates the relationship between specific surface area and pore volume of ZSM-5 crystals and DRIFT spectroscopy.

Results and discussion

Textural properties of the different ZSM-5 zeolite synthesized in different condition were measured using nitrogen adsorption and desorption isotherms at 77 K. The surface area was calculated by using the Brunauer–Emmett–Teller (BET) method based on the adsorption data in the partial pressure (P/P_0) range 0.05–0.20. Micropore volume was obtained by t-plot analysis. Total pore volume was obtained from the amount of nitrogen adsorbed at P/P_0 = ca. 0.99. The reference textural properties of the ZSM-5 zeolites have presented in Table 1. Three different classes were defined based on specific surface area and total pore volume obtained by N_2 adsorption and desorption isotherms at 77 K characterization tests (Table 2).

Table 1. Textural properties of the ZSM-5 zeolites were obtained by N_2 adsorption and desorption isotherms at 77 K.

Num.	S_{BET} (m^2g^{-1})	V_{micro} (t- plot)	V_{tot} (cm^3g^{-1})
1	451	۰,۱۹٦۲	0.225
2	433	0.1863	0.201
3	421	0.189	0.212
4	421	0.1833	0.218
5	264	0.0906	0.440
٦	۳۲۱	۰,۱۲٤	۰,۱۸۹

Table 2. Specific surface area and pore volume classes of ZSM-5 samples for classification approach.

class	S_{BET}	V_{tot} (cm^3g^{-1})
1	300-400	< 0.2
2	>400	0.2- 0.4
3	200-300	> 0.4

As results in Table3 point out high suitability of DRIFT spectrometry to be used in the field of catalyst characterization for quantitative investigations in catalyst science.

Table 3. Accuracy results of predicting the test samples by SIMCA and SVM methods.

	class	samples	Misclassified SIMCA	SVM
Training set	1	5	0	0
	2	26	0	0
	3	9	0	0
	Total	40	0	0
Test set	1	5	2	1
	2	14	0	0
	3	2	0	0
	Total	21	2	1

References

[1] T.C. Tsai, S.B. Liu and I. Wang, Disproportionation and Transalkylation of Alkylbenzenes over Zeolite Catalysts, Applied Catalysis A: General. 181 (1999) 355-398

[2] Y. Adewuyi, D. Klocke, J. Buchanan, Effects of high-level additions of ZSM-5 to a fluid catalytic cracking (FCC) RE-USY catalyst, Applied Catalysis A: General, 131 (1995) 121-133.

[3] H. Zaidi, K. Pant, Catalytic conversion of methanol to gasoline range hydrocarbons, Catalysis Today, 96 (2004) 155-160.

[4] J. Liu, C. Zhang, Z. Shen, W. Hua, Y. Tang, W. Shen, Y. Yue, H. Xu, Methanol to propylene: Effect of phosphorus on a high silica HZSM-5 catalyst, Catalysis Communications, 10 (2009) 1506-1509.

[5] J.F. Haw, W.G. Song, D.M. Marcus, J.B. Nicholas, The Mechanism of Methanol to Hydrocarbon Catalysis, Accounts of chemical research, 36 (2003) 317-326.

[6] H.G. Jang, H.K. Min, J. K. Lee, S. B. Hong, G. Seo, SAPO-34 and ZSM-5 nanocrystals' size effects on their catalysis of methanol-to-olefin reactions, Applied Catalysis A: General, 437 (2012) 120– 130.

Multivariate curve resolution-alternating least squares assisted voltammetry for simultaneous determination of non-steroidal anti-inflammatory drugs

Sayed Mehdi Ghoreishi*, Faezeh Saeidi Nejad, Saeed Masoum, Mohsen Behpour

Department of Analytical Chemistry, Faculty of Chemistry, University of Kashan, Kashan, Islamic Republic of Iran

E-mail address (corresponding author): s.m.ghoreishi@kashanu.ac.ir

Abstract

In the present work differential pulse voltammetry (DPV) coupled with multivariate curve resolution-alternating least squares (MCR-ALS) was applied for simultaneous determination of two non-steroidal anti-inflammatory drugs (mefenamic acid (MEF) and diclofenac (DCF) at the surface of a $Cu(OH)_2$ nanoparticles modified carbon paste electrode ($Cu(OH)_2$ nanoparticles/CPE). Characterization of the synthesized $Cu(OH)_2$ nanoparticles was carried by scanning electron microscopy (SEM). Multivariate optimization based on central composite rotatable design (CCRD) coupled with response surface methodology (RSM) was used to optimize all effective experimental variables on DPV responses of the two drugs. Multivariate curve resolution alternating least squares (MCR-ALS) was applied for the resolution and quantification of MEF and DCF with a high degree of overlapping voltammograms. The lack of fit (LoF) of 3.74% was obtained by MCR-ALS. The proposed method was successfully applied for simultaneous determination of two drugs in human serum samples.

Key words: Non-steroidal anti-inflammatory drugs, Experimental design, Multivariate curve resolution alternating least squares, $Cu(OH)_2$ nanoparticles modified carbon paste electrode, Voltammetric analysis

Introduction

Non-steroidal anti-inflammatory drugs such as mefenamic acid (MEF) and diclofenac (DCF) are the most commonly employed first line drugs for the control the pain and inflammation occurred due to any disease. [1]. These drugs are used to relieve the symptoms of many diseases, such as rheumatoid arthritis, nonarticular rheumatism and sports injuries [2]. Various analytical methods have been applied for the determination of these two drugs, including UV spectrophotometry, liquid chromatography, fluorescence spectrometry, and electrochemical methods. In differential pulse voltammetry (DPV) technique, MEF and DCF show peaks with high overlapping, thus this method cannot provide the necessary information to resolve this system with severe voltammograms overlap only. Chemometric methods is able to resolution of overlapping signals of analytes in multicomponent systems using of second-order data analysis methods [3] such as the multivariate curve resolution-alternating least squares (MCR-ALS) [4]. In this work, differential pulse voltammetry (DPV) coupled with multivariate curve resolution-alternating least squares (MCR-ALS) was applied for simultaneous determination of MEF and DCF at the surface of a $Cu(OH)_2$ nanoparticles modified carbon paste electrode ($Cu(OH)_2$ nanoparticles/CPE).

Results and discussion

Main purpose of this study is simultaneous determination of MEF and DCF. For achieving this aim, differential pulse voltammograms of a mixture containing 60.0 µM MEF and DCF in phosphate buffer solution (PBS) (0.1 M, pH 6.4) was recorded at the surface of $Cu(OH)_2$ nanoparticles/CPE (Fig. 1) in the optimized conditions. As can be seen, a high degree of overlapping was observed for the voltammograms of MEF and DCF. Thus, it is necessary to resolve these voltammograms by MCR-ALS method.

Thus, a calibration set of known mixtures was prepared and differential pulse voltammograms of the each mixture were carried out at a pulse height of 10.0–60.0 mV with a 10 mV interval. In next step, MCR-ALS as a powerful chemometrical method was applied on a column-wise augmented data matrix obtained from calibration set of MEF and DCF and a lack of fit (LoF) of 3.74% was obtained. Therefore, using MCR-ALS, decomposed voltammograms of the mixtures of MEF and DCF that applied in calibration curve are obtained that presented in **Error! Reference source not found.**.

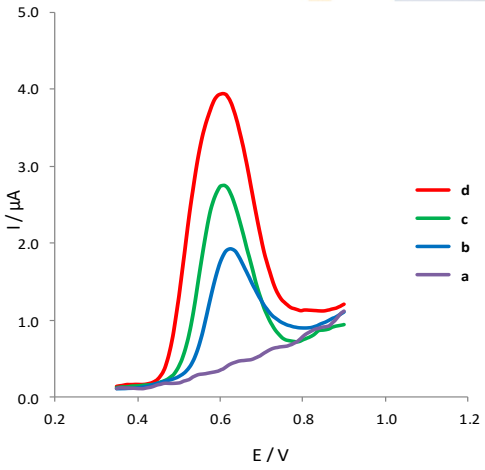


Fig. 1.. DPV of $Cu(OH)_2$ nanoparticles/CPE (a) in PBS (pH 6.4), (b) in the presence of 60.0 µM DCF, (c) in the presence of 60.0 µM MEF, and (d) in the presence of 60.0 µM DCF + 60.0 µM MEF.

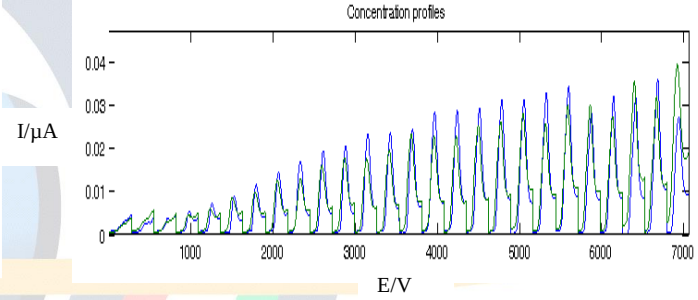


Fig. 8. Resolving of the augmented data by MCR-ALS for the data of standards of MEF and DCF.

References

[1] Kormosh Z, Hunka I, Bazel Y, Laganovsky A, Mazurenko I, Kormosh N, *Cent. Eur. J. Chem.* 2007; 5: 813.

[2] Muraoka S, Miura T, *Life Sci.* 2003; 72: 1897.

[3] Gomez V, Callao M.P, *Anal. Chim. Acta* 2008; 627: 169–183.

[4] Tauler R, Smilde A, Kowalski B.R, *J. Chemom.* 1995; 9: 31–58.

Spectrophotometric determination of oxazepam in drug and biological samples by parallel factor analysis

Ali Niazi*, Somayeh Yadegari

Department of Chemistry, Faculty of Sciences, Islamic Azad University, Arak, Iran

Abstract

In this work, direct determination of Oxazepam, an anxiolytic and sedative agent, in pharmaceutical formulations and biological fluids(plasma) was accomplished based on ultraviolet spectrophotometry(220-380 nm) using parallel factor analysis (PARAFAC)
and partial least squares (PLS). The study was carried out in the pH range from 1.0 to 12.0 and with a concentration range from 0.1to10 μgml−1 of oxazepam Multivariate calibration models using PLS at different pH and PARAFAC were elaborated for ultraviolet spectra deconvolution and oxazepam quantitation. The best models for the system were obtained with PARAFAC and PLS at pH=3,pH= 7 and pH= 12. The capabilities of the method for the analysis of real samples were evaluated by determination of oxazepam in pharmaceutical preparations and biological (plasma) fluids with satisfactory results. The accuracy of the method, evaluated through the root mean square error of prediction (RMSEP), was0.0233 for oxazepam with best calibration curveby PARAFAC and 0.0689 for oxazepam with PLS model at best pH. Protolytic equilibria of oxazepam were evaluated by DATAN program using the corresponding absorption spectra-pH data. The obtained p*K*a values of oxazepam are2.8 and11.38 for p*K*a1 and p*K*a2, respectively.

Keywords: Oxazepam; PARAFAC; PLS; DATAN; Pharmaceutical formulations; Biological fluids; Protolytic equilibria; Ultraviolet spectrophotometry.

Chemometric approach to optimize micellar liquid chromatographic separation of some natural anthraquinone dyes

Amir M. Ramezani, Saeed Yousefinejad, Mohammad Nazifi, Ghodratollah Absalan*

Professor Massoumi Laboratory, Department of Chemistry, College of Sciences, Shiraz University, Shiraz 71454, Iran

E-mail: gubsulun@yahoo.com

Abstract

A simple chemometrics approach has been developed for simultaneous optimization of resolution and analysis time in micellar liquid chromatography of some natural antraquinone dyes in isocratic mode. In this regard, the experimental parameters were designed according to a face-centered cube response surface methodology. To optimize the peak separation of four anthraquinone dyes, the effect of surfactant concentration, type and amount of the modifier, volume of the acetic acid on both retention time and resolution was investigated by means of multivariate analysis. The Pareto-Optimality Method was employed to select the best possible combinations of separation quality and analysis time.The applicability of the methodology was tested on madder root for separation and measurement of Alizarin, Purpurin, Danthron, and Quinizarin.

Key words: Experimental design; Micellar Liquid chromatography; antraquinone dyes; Multivariate analysis; central composite design; Isocratic elution

Introduction

Textile dying with antraquione dyes is characterized by high environmental pollution and high health risks to personnel handling harmful substances. This can be harmful and cause allergies in humans [1]. Hence, finding a reliable separation technique for natural dyes is a great analytical challenge. In this regard, micellar liquid chromatography (MLC) as a reversed-phase liquid chromatographic mode with a mobile phase consisting of an aqueous solution of surfactant above its critical micellar concentration is promising [2]. In MLC, usually several criteria such as retention factor, peak resolution, peak asymmetry, and retention time, should be considered to achieve an optimum experimental condition. One of the best systematic optimization methods for chromatographic separations is the central composite design (CCD), which is very efficient and potent in providing sufficient information on experimental variables with low overall experimental error and requires minimum number runs [3]. To the best of our knowledge, so far, this chemometrics approach has not been utilized to optimize MLC parameters for separation of antraquinone dyes which have different polarities.

Results and discussion

In order to achieve optimum experimental condition for suiatable separation of target compounds, effects of four operational factors including concentration of surfactant, alkyl chain length of the organic modifier, volume percentage of modifier and volume percentage of acetic acid (as acidic adjustment in micellar mobile phase) were studied. The experimental value range of the factors was selected based on chromatographic insight and physical limitations. The experiments were run in a random order and design using five levels for each factor in a face centered cube central composite design. The utilized approach provided enough information for calculating of the regression models containing linear interactions and non-linear effects. As it is shown in Fig. 1, the suggested approach was able to separate the chromatographic peaks of analalytes very well with adequate resolution in a reasonable analysis time. The results of the study suggested a suitable model with descriptive and predictive ability for the chromatographic response function, which allows finding the optimum conditions in the separation of natural dyes with different polarity.

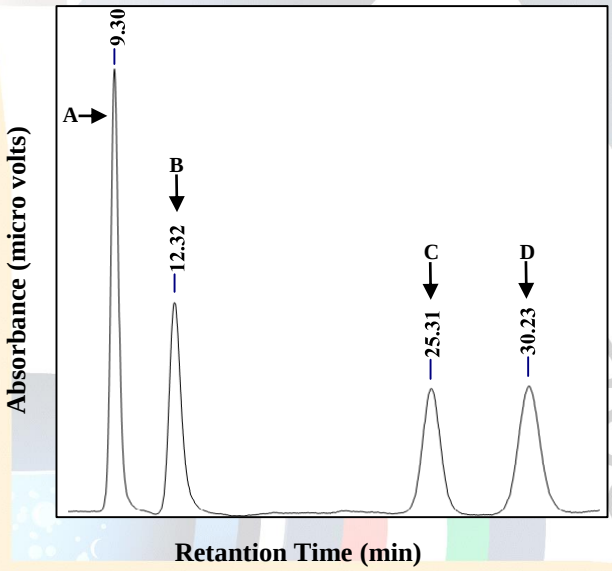


Figure 1. Chromatograms obtained for separation of Alizarin (A), Purpurin (B), Danthron (C), and Quinizarin (D) anthraquinones

References

[1] Dabiri M, Salimi S, Ghassempour A, Rassouli A, Talebi M. Optimization of microwave-assisted extraction for alizarin and purpurin in Rubiaceae plants and its comparison with conventional extraction methods. J. Sep. Sci. 2005; 28: 387-396.

[2] Hadjmohammadi MR, Nazari SSSJ. Chromatographic behavior of aromatic diamines in hydro-organic, micellar and submicellar reversed phase liquid chromatographic modes. Chromatographia 2013; 76: 23-31.

[3] Hadjmohammadi MR, Nazari SSSJ. Optimization of separation of flavonoids in micellar liquid chromatography using experimental design and derringer's desirability function. J. Liq. Chromatogr. Rel. Technol. 2012; 36: 943-957.

Evaluation of partial least squares-discriminant analysis for classification of gas chromatography fingerprints of saffron

Ghazaleh Aliakbarzadeh[a], Hadi Parastar*[,b], Hassan Sereshti[a]

[a]*Department of Chemistry, Faculty of Science, University of Tehran, Tehran, Iran*
[b]*Department of Chemistry, Sharif University of Technology, Tehran, Iran*

E-mail address: h.parastar@sharif.edu

Abstract

The present work aims at illustrating the relevance of gas chromatograpy-flame ionization detector (GC-FID) fingerprints of saffron for assessing the geographical origin of saffron samples from different regions of Iran. The volatile components of 72 saffron samples were extracted and preconcentrated using ultrasound-assisted dispersive liquid-liquid microextraction (US-DLLME) and then analysd by GC-FID. The obtained chromatograms were baseline corrected, aligned, and then subjected to multivariate data analysis by principal component analysis (PCA) and partial least squares discriminant analysis (PLS-DA). In order to eliminate the possibility of the model overfitting, variable importance in projection (VIP) and selectivity ratio (SR) were examined in PLS-DA analysis. Several major chemical components within the regions that are relevant to the discriminant problem were suggested as being the most influential.

Keywords: Saffron; Chromatography; Fingerprinting; Classification; Partial least squares discrimination analysis.

Introduction

Crocus sativus L. is an autumnal geophyte flower whose dried stigmas, known as saffron, is the most valuable spice in the world. Saffron is known for its aromatic and coloring power. Also, it has been used for medical purposes since immemorial time. Saffron is cultivated in a wide range of environments with mild to dry climates [1]. The Khorasan provinces of Iran with approximately 300 tones production in a year are considered as the main producers of saffron in the world. Since saffron is a valuable food and pharmaceutical product with high economic efficiency, its quality control based on its origin and also harvest and postharvesting procedures are of most importance [2]. The combination of gas chromatography (GC) fingerprinting with multivariate analysis provides a reasonable approach to study the profile of saffron's volatiles in relation to its geographical origin [3]. In the present study, the potential of partial least squares-discriminant analysis (PLS-DA) with different preprocessing and variable selection methods is evaluated for classification of GC fingerprints of saffron. In addition, it is tried to use the combination of fast and cheap microextraction method with GC-FID multivariate classification analysis to discriminate the Iranian saffron based on their cultivation area.

Results and discussion

Twenty-four saffron samples were collected from different resgion of the Razavi Khorasan and south Khoransan provinces of Iran. The saffron volatile components were extracted and preconcetrated using previously optimized ultrasound-assisted dispersive liqid-liquid microextraction (US-DLLME) [4]. The extracts were then analysed by GC-FID. The obtained chromatograms were baseline corrected by means of asymmetric least squares (AsLS) approach and aligned by correlation optimization warping (COW) method. The COW parametrs including the slack size and length were also optimized. In order to eliminate the artificial source of the variation, the chromatograms were normalized toward the maximum value of the internal standard. Also, data were mean-centered and Pareto scaled. For the classification and discrimination between samples, PLS-DA was carried out. The data were divided to two group of the training and test sets in order to assessing the validaty of the model. Based on variable selection methods including variable importance in projection (VIP) and selectivity ratio (SR), β-isophoron, dihydrooxophorone, safranal, HTCC, ketoisophorone and 2,4,4-trimethyl-3-carboxaldehyde-5-hydroxy-2,5-cyclohexadien-1-one were recognized as the main components which cause the differences among the two classes. The classification and prediction figures of merit were summarized in Table 1.

Table 1. Statistics for classification for five Latent variables (Total X variance 84%)

			Class 1 (South Khorasan)		Class 2 (Razavi Khorasan)	
	Error rate	Non-error rate	sensitivity	specificity	sensitivity	specificity
Fitting	0	1	1.00	1.00	1.00	1.00
Cross validation	0.04	0.96	1.00	0.93	0.93	1.00
Test set	0.06	0.94	0.86	0.81	0.81	0.86

It is concluded that PLS-DA can be considered as an efficient method for classification of chromatographic fingerprints of complex mixtures like saffron and for discrimination of most significat chemical components.

References

[1] T. K. Lim, in *Edible Medicinal and Non Medicinal Plants*, Springer Netherlands, 2014, vol. 8, ch. 3, pp. 77-136.

[2] L. Maggi, M. Carmona, S. D. Kelly, N. Marigheto and G. L. Alonso, Geographical origin differentiation of saffron spice (Crocus sativus L. stigmas) – Preliminary investigation using chemical and multi-element (H, C, N) stable isotope analysis, *Food Chem.*, 2011, **128**, 543-548.

[3] M. Jalali-Heravi, H. Parastar and H. Ebrahimi-Najafabadi, Self-modeling curve resolution techniques applied to comparative analysis of volatile components of Iranian saffron from different regions, *Anal. Chim. Acta*, 2010, **662**, 143-154.

[4] H. Sereshti, R. Heidari and S. Samadi, Determination of volatile components of saffron by optimised ultrasound-assisted extraction in tandem with dispersive liquid–liquid microextraction followed by gas chromatography–mass spectrometry, *Food Chem.*, 2014, **143**, 499-505.

Prediction of photolysis half-lives of dihydroindolizines by genetic algorithm-multiple linear regression (GA-MLR)

Davoud Beiknejad, Mohammad Javad Chaichi*, Mohammad Hossein Fatemi

Faculty of Chemistry, University of Mazandaran, Babolsar 47416-95447, Iran

*E-mail: jchaichi@yahoo.com

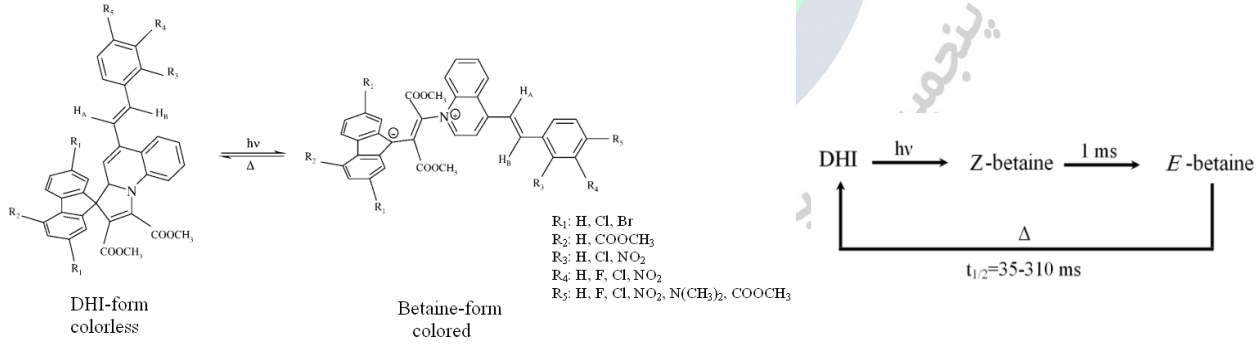
Abstract

Quantitative structure-property relationship (QSPR) study was carried out for the prediction of photolysis half-lives of dihydroindolizines (DHIs). Genetic algorithm (GA) variable selection was used in the multiple linear regression (MLR) modeling. Four descriptors including topological (AAC), BCUT (BEHm3), and RDF (RDF090m and RDF065v) descriptors were selected by the algorithm, revealing that electrostatic and steric interactions are the most important features affecting photolysis half-life of studied DHI photochromic compounds. The prediction of log $t_{1/2}$ for an external test set of molecules gave an R^2_{ext}= 0.9418, $RMSE_{ext}$= 0.0692 log $t_{1/2}$ units. The results showed that the developed model provides statistically significant predictions of photolysis half-lives of DHIs.

Keywords: Dihydroindolizines; Photochromism; Photolysis half-live; QSPR; GA-MLR

Introduction

Interest in photochromic reactions has increased dramatically in recent years [1]. Photochromism is defined as a reversible transformation of a chemical species between two forms having different absorption spectra induced in one or both directions by photoirradiation [2]. Fluorene-9'- styrylquinolinedihydroindolizines belong to an exciting class of DHIs which their synthesis and photophysical properties have been studied by Ahmed [3]. The photochromism of this class of DHIs is based on the 1,5-electrocyclization which can be reversed upon exposure to light (Scheme 1).



Scheme 1

Due to solubility of DHIs and related reagents in organic solvents, the reactions and experiments of DHIs including synthesis, separation, purification and photophysical study have been carried out mainly in *hazardous and/or toxic*

organic solvents. However, experimental determination of photokinetic parameters of DHIs is expensive, labor intensive, relatively time consuming and requires special equipments. Therefore development of a theoretical method to predict the photokinetic parameters will be of great importance in the study of DHIs photochromic systems.

Results and discussion

The values of $t_{1/2}$ for all studied DHIs were taken *from literature* [3]. The data set (n=28) was divided randomly into two groups: training set (n=23) and test set (n=5). QSARINS software (version 2.2), a newly introduced QSAR software, was used for GA-MLR analysis [4]. The relationship between log $t_{1/2}$ and four descriptors is shown as follows:

$$\log t_{1/2} = 8.2183 - 2.7005\ AAC - 0.4682\ BEHm3 + 0.0126\ RDF090m - 0.0167\ RDF065v \quad \text{Eq. (1)}$$

R^2_{train}= 0.9485, R^2_{adj}= 0.9371, $RMSE_{train}$= 0.0443, SE= 0.0501, F= 82.8812, Q^2_{loo}=0.9190, Q^2_{lmo}= 0.9091, R^2_{ext}= 0.9418, $RMSE_{ext}$= 0.0692

The appeared descriptors demonstrate that topological (*AAC*), BCUT (*BEHm3*), and RDF (*RDF090m* and *RDF065v*) descriptors affect the photolysis half-life of the studied DHIs (Eq. 1). The correlation matrix of the appeared descriptors was built and none of them were highly correlated to one another, meaning that each of the descriptors provides independent information to the QSPR model. The developed model was also used to estimate the photolysis half-lives of new 18 DHIs. Electrostatic and steric interactions are the most important features affecting photolysis half-life of DHI photochromic compounds. Based on these findings, the authors believe that this study allows material designers to predict the photokinetic properties of novel photochromic materials in DHI family prior to synthesis, and to understand the relationships between the microscopic properties of molecular components and the photochromic behavior of materials.

References

[1] Irie M, Yokoyama Y, Seki T. *New Fontiers in Potochromism*, Springer, Japan, 2013.

[2] Irie M, Fukaminato T, Matsuda K, Kobatake S. Photochromism of diarylethene molecules and crystals: memories, switches, and actuators. *Chem. Rev.* 2014; **114**:12174-12277.

[3] Ahmed SA. (2002) Photochromism of dihydroindolizines. Part II-Synthesis and photophysical properties of new photochromic IR-sensitive photoswitchable substituted fluorene-9′-styrylquinolinedihydroindolizines. *J. Phys. Org. Chem.* 2002; **15**:392–402.

[4] Gramatica P, Chirico N, Papa E, Cassani S, Kovarich S. QSARINS: A new software for the development, analysis, and validation of QSAR MLR models. *J. Comput. Chem*. 2013; **34**:2121-2132.

Qantitative Structure-Electrochemistry Relationship in Non-aqueous Solutions: On the Reduction Potential of Anthraquinone Derivatives in Organic phase

Fatemeh Honarasa [a], Saeed Yousefinejad [b] *

[a] *Department of Chemistry, Shiraz Branch, Islamic Azad University, Shiraz, Iran*
[b] *Department of Chemistry, Shiraz University, Shiraz, Iran*

E-mail address (corresponding author): yousefinejad.s@gmail.com

Abstract

A new strategy is proposed for the quantitative structure-electrochemistry relationship (QSER) in non-aqueous systems, based on cooperation of both theoretical properties of electrochemical analytes and empirical parameters of solvents. The modeling of the first reduction potential of some Anthraquinone derivatives in three organic solvents (acetonitrile, N,N-dimethyl formamide and dimethylsulfoxide) were done using the proposed approach. By using combination of properties of electrochemical solvent and solute in a single model, it was shown that the structural features of Anthraquinones compounds are important in predicting the electrochemical properties of these non-aqueous systems as well as some properties of involved solvents. The correlation coefficients of cross validation and external test set were 0.93 and 0.96 respectively which indicate stability and prediction ability of the model. To our knowledge, this is the first report on simultaneous including the features of analyte and solvent in a QSER. This wok obtains a way to decrease hazardous experiences in non-aqueous solvents.

Key words: Non-aqueous electrochemistry, Structure-electrochemistry relationship, Solvent empirical parameters, anthraquinones

Introduction

One of the earliest applications of non-aqueous solutions was in electrochemistry for exploring new chemical possibilities from one side and obtaining information about the dynamic and static solvent effects on different chemical processes from the other side [1].

In spite of these advantages, using organic solvents in electrochemical reactions (and other applications) could lead to problematic side effects on to environment and human health. So nowadays, the focus is on making “green” approaches, including low-risk organic solvents, ionic liquids, supercritical fluids, immobilized solvents, and even solventless processes. One of the other ways toward green approaches is decreasing trial-and-error experiences in non-aqueous solvents by estimating the analyte’s electrochemical property using quantitative structure-property relationships (QSPR) which is well-known as quantitative structure-property relationships (QSER) [2]. In the current work, for the first time a simple and informative approach was suggested for QSER in different non-aqueous solvents by considering both structural properties of analytes and some empirical scales of solvent [3]. The proposed approach

is away toward clarification of the involved solvent-electrochemical analyte interactions in addition to determination of important structural features of analytes. 30 electrochemical mixtures of antraquinones (Aq) in three organic solvents were chosen as the model to show the potential of proposed approach.

Results and discussion

A new approach was proposed based on the combination of structural descriptors of electrochemical analyte and empirical scales of electrochemical solvents for structure-electrochemistry modeling. The capability of this approach was shown by modeling the first reduction potential of 30 electrochemical mixtrure of anthraquinones (obtained by cyclic-voltametric) in different organic solvents. The constructed model covered more than 96% of variance in 24 training mixtures (See Fig. 1). Leave one-out and leave-many out cross validation and also prediction of the external test set, indicated the high stability and prediction ability of this model. The current study could conduct new avenue to predict different electrochemical properties of compounds (especially analog series) in non-aqueous solutions. Such predictions could be used to reach the optimized non-aqueous electrochemical systems with lower trial-and-error and risky experiments in organic solvents.

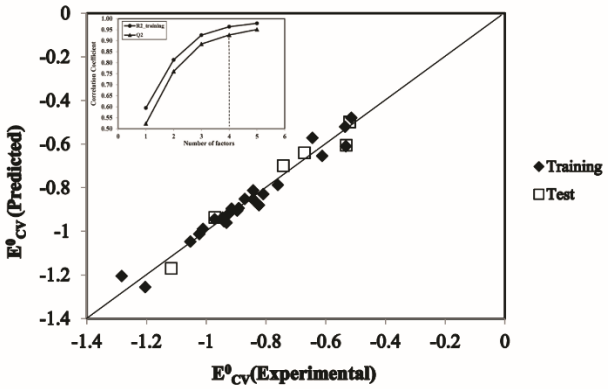


Fig. 1 Plot of predicted E^0_{CV} versus their experimental values for 30 electrochemical mixtures in training and test set using a four parametric model. Inset: Correlation coefficient of the training set (R^2_{train}) and cross-validation (Q^2) versus number of parameters to select the optimum number of parameters for model development.

References

[1] K. Izutsu, Electrochemistry in Nonaqueous Solutions, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2002.

[2] M. Shamsipur, A. Siroueinejad, B. Hemmateenejad, A. Abbaspour, H. Sharghi, K. Alizadeh, et al., Cyclic voltammetric, computational, and quantitative structure-electrochemistry relationship studies of the reduction of several 9,10-anthraquinone derivatives, J. Electroanal. Chem. (2007) **600**: 345–358.

[3] S. Yousefinejad, B. Hemmateenejad, A chemometrics approach to predict the dispersibility of graphene in various liquid phases using theoretical descriptors and solvent empirical parameters, Colloids Surfaces A Physicochem. Eng. Asp. (2014) **441**: 766–775.

Simultaneous determination of anionic dyes from aqueous solution using partial least square regression

Maryam Bagtash, Javad Zolgharnein*

Department of Chemistry, Faculty of Science, Arak University, Arak, Iran

E-mail address: bagtash@yahoo.com

Abstract

In order to determine the amount of anionic dyes such as Alizarin red and Congo red, spectrophotometry is used as a simple, rapid and economical method. The spectra of these compounds in their mixture show severe overlapping. Therefore, Partial least square regression (PLS) as a multivariate calibration method is developed for the simultaneous determination of Alizarin red and Congo red in binary solutions to overcome the severe spectral overlapping. Some statistical values such as the root mean square error of prediction, relative standard error of prediction, and the correlation coefficient were employed to evaluate the predictive ability of multivariate calibration model.

Keywords: Alizarin Red, Congo Red, Partial least square regression

1. Introduction

The synthetic dyes such as Alizarin Red S (AR) and Congo Red (CR) are well-studied anionic dyes. AR has been widely used for dyeing textile materials since ancient times as an important red dye, mainly for dyeing textile fabrics [1]. CR dye is benzedene based dye. This dye has been known to cause an allergic reaction and to be metabolized to Benzedrine. Its decomposition results in carcinogenic products [2]. Therefore, it is essential to monitor these dyes in water. To date, different analytical methods for the determination of dyes have been published, which need the intensive sample pretreatment. However, these separation processes are not simple and the simultaneous determination of target dyes by traditional spectrophotometry techniques is difficult because, generally, the absorption spectra overlap in a bright region and the superimposed curves are not suitable for quantitative evaluation. In recent years, attention has been directed toward chemometric methods for analysis of multi-component systems because of fast data collection using rapid scanning spectrophotometers [3]. In this work, a spectrophotometric method has been developed for simultaneous determination of binary mixtures of dyes (AR and CR) in water samples with applying partial least square regression (PLS).

2. Materials and Methods

2.1. Apparatus and Reagents

A UV–Vis spectrophotometer (Analytikjena SPECORD250) was used for analysis of AR and CR dyes. The data were processed on a Pentium IV computer with programs written in MATLAB 6.5 on Windows. All chemicals were used with analytical-grade from Merck (Darmstadt, Germany) and doubly distilled water was used throughout the work.

3. Results and discussion

3.1. The assay of AR and CR in a single solution

The absorbance spectra of single solution of AR (60 mg/L), CR (20 mg/L) and their mixture recorded between 370 and 700 nm are shown in Fig. 1. The dye concentration was calculated from a calibration curve of absorbance versus concentration at λ_{max} of each dye in single dye solutions. For constructing the individual calibration lines the absorbencies were measured at 520 and 500 nm for AR and CR, respectively. The linear regression equation for the calibration graph for AR for the concentration range of 1–120 mg/L was A = 0.0117 C_{AR} -0.007 (R^2 = 0.998) and for CR for the concentration range of 1–50 mg/L was A =0.036 C_{CR} +0.012 (R^2 = 0.998). The limits of detection were 0.5 and 0.3 mg/L for AR and CR, respectively, calculated according to calibration lines characteristics.

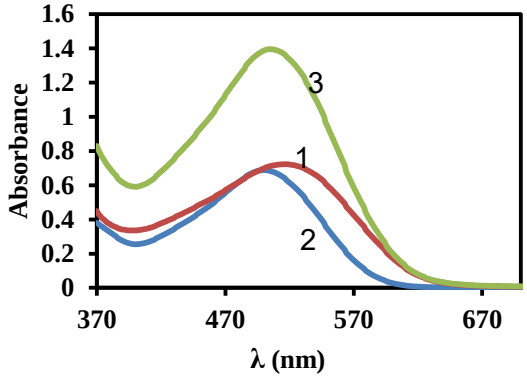


Fig 1: Absorption spectra of: 1) AR (60 mg/L); 2) CR (20 mg/L); and 3) their mixture spectrum.

3.2. The simultaneous determination of AR and CR in binary solutions

A number of 21 solutions were prepared to construct the models (calibration set (Table 1)) and another 18 solutions to validate them (prediction set) that these not included in the previous set were employed as an independent test (Table 2). All recorded data are just meancentered. The composition of samples was randomly designed to obtain enough information procedure. In order to determine the optimum number of factors (latent variables) in the PLS calibration algorithm, a leave-one-out cross validation method was used. The prediction residual error sum of squares (PRESS) was calculated for each dye for the prediction set, which was the samples not participating in the construction of the PLS model. PRESS is defined as follows:

$$PRESS = \sum_{j=1}^{n}\{C_j - \hat{C}_j\}^2 \quad (1)$$

where n, is the number of samples in the prediction set, C_j is the real concentration in the jth sample, and $\hat{C}_j$ is its estimated value. One logical choice for the optimum number of factors would be that number which yielded the minimum PRESS. The cross validation was repeated 40 times for each number of factors until each factor had been left out once. Concequently, the optimal number of factors for both AR and CR were obtained as two, with the lowest value of PRESS [4]. Some statistical values such as the root mean square error of prediction (RMSEP), relative standard error of prediction (REP), and the correlation coefficient (R^2) can be used to evaluate the predictive ability of a multivariate calibration model [3, 4] (Table 3).

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2}{n}} \quad (2) \qquad REP = 100\sqrt{\frac{\sum_{j=1}^{n}(x_i - \hat{x}_i)^2}{\sum_{j=1}^{n} x_i^2}} \quad (3) \qquad R^2 = \frac{\sum_{i=1}^{n}(\hat{x}_i - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \quad (4)$$

where x_i was the predicted value of model, $\hat{x}_i$ as the experimental value, n was the number of data and $\bar{x}_i$ was the average of the experimental value.

Table 1. Composition of calibration set.

Sample	C_{AR} (mg/L)	C_{CR} (mg/L)	Sample	C_{AR} (mg/L)	C_{CR} (mg/L)	Sample	C_{AR} (mg/L)	C_{CR} (mg/L)
1	18.91	12.19	8	52.50	25.00	15	30.00	5.00
2	86.08	12.19	9	52.50	17.50	16	30.00	40.00
3	18.91	22.80	10	15.86	10.12	17	55.00	30.00
4	86.08	22.80	11	44.14	10.12	18	79.74	11.22
5	5.00	17.50	12	15.86	34.87	19	٥٥.00	19.00
6	100.0	17.50	13	10.00	22.50	20	79.74	26.78
7	52.50	10.00	14	50.00	22.50	21	20.00	19.00

Table 2. Composition of prediction set.

Sample	Real C_{AR}	Real C_{CR}	PLS C_{AR}	PLS C_{CR}	Sample	Real C_{AR}	Real C_{CR}	PLS C_{AR}	PLS C_{CR}
1	6.19	15.25	5.51	15.02	10	8.02	13.49	7.42	13.18
2	16.80	15.25	13.92	15.21	11	12.97	13.49	12.14	13.23
3	6.19	33.75	4.74	34.29	12	8.02	25.51	7.31	25.41
4	16.80	33.75	15.59	33.96	13	12.97	25.51	11.62	24.85
5	4.00	24.50	3.31	24.66	14	7.00	19.50	6.54	19.42
6	19.00	24.50	17.22	24.58	15	14.00	19.50	13.10	19.43
7	11.50	11.42	9.85	11.33	16	10.50	11.00	10.21	10.76
8	11.50	37.58	9.85	38.14	17	10.50	28.00	9.44	27.90
9	11.50	24.50	10.63	24.40	18	10.50	19.50	9.83	19.33

Table 3. The calculated results of RSEPs and RMSEP for PLS.

method	RMSEP (%)	RMSEP(%)	REP (%)	REP (%)	PRESS	PRESS
	C_{AR}	C_{CR}	C_{AR}	C_{CR}	C_{AR}	C_{CR}
PLS	1.2	0.28	0.11	0.012	26.58	1.46

According to statistical parameters calculated for spectral data using PLS (Table 3) the proposed method could be applied to the determination of other dyes with severe overlapping spectra in their binary solutions.

References

[1] Pirillo S, Ferreira ML, Rueda EH, The effect of pH in the adsorption of Alizarin and Eriochrome Blue Black R onto iron oxides, *J. Hazard. Mater*. 2009; **168**;168–178.

[2] Demirbas A, Agricultural based activated carbons for the removal of dyes from aqueous solutions, *J. Hazard. Mater.* 2009; **167**; 1-9.

[3] An L, Deng J, Zhou L, Li H, Chen F, Wang H, Liu Y, Simultaneous spectrophotometric determination of trace amount of Malachite Green and Crystal Violet in water after cloud point extraction using partial least squares regression, *J. Hazard. Mater*. 2010; **175;** 883–888.

[4] Gao J-F, Wang J-H, Yang C, Wang S-Y, Peng Y-Z, Binary biosorption of Acid Red 14 and Reactive Red 15 onto acid treated okara: Simultaneous spectrophotometric determination of two dyes using partial least squares regression, *Chem. Eng. J*. 2011;**171;** 967–975.

Dispersive liquid-liquid microextraction based on solidification of organic drop with central composite design for the determination of 5-hydroxymethyl-2-furfural using high-performance liquid chromatography

Sajad Sobhani[a], Fereshteh shiri[*a] and Beshareh Hashemi[b]

[a]*Department of Chemistry, University of Zabol, P.O. Box 98615-538, Zabol, Iran*
[b]*Department of Chemistry, Razi University, 67149-67346 Kermanshah, Iran*
Fereshteh.shiri@gmail.com

Abstract

A simple, rapid and inexpensive dispersive liquid-liquid microextraction based on solidification of organic drop (DLLME-SFO) combined with high-performance liquid chromatography was developed for the extraction and determination of trace levels of 5-hydroxymethyl-2-furfural (HMF) in fruit juice. Effect of variables such as extracting and dispersive solvent volume and pH were investigated simultaneously using Central composite design. Under the optimum conditions, the calibration graph was linear in the range of 1 µg L^{-1} with the detection limit of 200 µg L^{-1}. The optimized method revealed a good precision with relative standard deviation (RSD) for 4 replicate determinations of 50 µg L^{-1} HMF was 4.1%. The performance of the method was evaluated for extraction and determination of HMF in fruit juice.

Key words: DLLME-SFO; 5-hydroxymethyl-2-furfural; central composite design; relative standard deviation.

Introduction

In recent years, Assadi and co-workers demonstrated a novel microextraction method called dispersive liquid–liquid microextraction (DLLME). DLLME is based on a ternary solvent system, in which a mixture of extracting and dispersive solvent is rapidly injected into an aqueous sample containing the analytes of interest, which causes the formation of a cloudy solution. In most common version of DLLME, solvents with the densities higher than water are required and further, they are not often compatible with reverse phase HPLC and in addition, the high-density extraction solvents, being mostly halogenated, are commonly hazardous to the environment. This feature is considered as the main drawback of this method.

Therefore, a simple dispersive liquid–liquid microextraction method based on the solidification of floating organic drop (DLLME-SFO) was introduced by Leong *et al*. [1] It is based on DLLME and the solidification of floating organic drop [2-5]. In this method solvents with densities lower than water are used and the floated extractant is solidified to be collected easily for analysis.

In the present work, DLLME-SFO followed by HPLC with UV detection was used for extraction and determination of HMF. The influence of main experimental parameters on the performance of this method were thoroughly investigated and discussed by experimental design. Finally, the applicability of the proposed method was

tested by the determination of HMF in fruit juice samples. 5-Hydroxymethylfurfural (5-hydroxymethyl-2-furaldehyde, HMF) is considered to be an intermediate product of the famous Maillard Reaction. When foods containing reducing sugars and amino acids are cooked and processed, we can observe the occurrence of this reaction. The process of HMF production begins with the acid-catalysed dehydratation of hexoses via 1,2-enolisation which is followed by glucosamine hydrolysis ; the result is a natural product in which water coexists with monosaccharides in acidic medium.

Results and discussion

In the present study, a DLLME-SFO method combined with HPLC-UV was developed and applied for pre-concentration and determination of HMF in orange juice. To obtain high efficiency, the effect of different parameters such as type of dispersive and extraction solvents, their volumes, pH and salt concentration were studied using both the one variable at a time and central composite design. 1-undecanol was selected as the extraction solvent for subsequent experiments. Methanol was chosen as a disperser solvent. The ionic strength and pH of the solution was adjusted to an appropriate amount (sodium chloride, 1% w/v; pH, 4). Considering confidence limits for selection of regression terms, the mathematical equation describing the relationship between three factors and responses can be expressed as following equations:

$$R= -153157 + 23375.9\ X_1 + 78110.1X_2 + 6045.97\ X_3 -2494.98\ X_1^2 -22829.4\ X_2^2 -76.3127\ X_3^2$$

Statistics ANOVA results for these models are presented in Table 1.

Table1 ANOVA results for the reduced cubic model of HMF

Source	SS	df	F-value	*p*-value
Model	35589	9	20.89	<0.0001
Residual	1883	8	-	-
Lack of fit	1583	5	15.2	0.052
Pure error	577	3		
R2: 0.9691, adjusted R2: 0.9322, predicted R2: 0.8102				

SS: sum of squares; df: degree of freedom

References

[1]. M.-I. Leong and S.-D. Huang, *J. Chromatogr. A*, 2008, **1211**, 8-12.

[2]. S. Berijani, Y. Assadi, M. Anbia, M.-R. Milani Hosseini and E. Aghaee, *J. Chromatogr. A*, 2006, **1123**, 1-9.

[3]. N. Fattahi, Y. Assadi, M. R. M. Hosseini and E. Z. Jahromi, *J. Chromatogr. A*, 2007, **1157**, 23-29.

[4]. M. Rezaee, Y. Yamini, S. Shariati, A. Esrafili and M. Shamsipur, *J. Chromatogr. A*, 2009, **1216**, 1511-1514.

[5]. M. Rezaee, Y. Assadi, M.-R. Milani Hosseini, E. Aghaee, F. Ahmadi and S. Berijani, *J. Chromatogr. A*, 2006, **1116**, 1-9.

Using QSAR calculation of Indol derivatives and 7-azaindoles for the treatment of Multiple Sclerosis(MS)

Seyedeh Maryam Moosavi[*], *Ghasem Ghasemi*

Department of Chemistry, Islamic Azad University, Rasht Branch, Rasht, Iran. Email: Maryammoosavi20@yahoo.com

Keywords:"Multiple Sclerosis's disease" ,"Quantitative structure activity relationship" ,"Indol derivatives" ,"Multiple linear regression","ICA".

Introduction

Multiple sclerosis, or MS, is a long-lasting disease that can affect brain, spinal cord, and the optic nerves in eyes. It can cause problems with vision, balance, muscle control, and other basic body functions. The goal of treatment for patients with symptoms but not eliminate the symptoms under control [1]. On the pattern, quantitative structure–activity relationship (QSAR) study has been done on aseries of Indol derivatives and 7-azaindoles for the treatment of Multiple Sclerosis's disease. The purpose of QSAR study is to find a relation between the composition or structure of a compound with its bio or chemical activity, in order to design a new compound with expected properties or predict the properties of an unknown compound. Up to now, a lot of successful applications have been reported in many different types of cases, e.g., medicine design, environmental chemistry exploration, pesticide searching, etc [2].The artificial neural networks (ANNs) are known as a good method in expressing highly non-linear relationship between the input and output variables, hence, greater interests were attracted in applying them to the pattern classification of complex compounds [3]Genetic algorithms (GAs) were introduced by Holland. They mimic nature's evolutionary method of adaptation to a changing environment. GAs are stochastic optimization methods hat provide powerful means to perform directed random searches in a large problem space as encountered in chemometrics and drug design .The Imperialist Competitive Algorithm (ICA), derived from the field of human social evolution, is a component of swarm intelligence theory. It was first introduced in 2007 to deal with continuous optimization problems, but recently has been extensively applied to solve discrete optimization problems. In multiple linear regression (MLR), for a given data set consisting of a target variable and M descriptors for n compounds, a model is made with good fitting to define the combination of m descriptors (m< M) on target variable. Running through all combinations usually is too time-consuming. Therefore, several approximate methods have been proposed for this reason, but none of them guarantied to find very best combination in all cases.

Methods

The 3D structures of the molecules were generated using the built optimum option of Hyperchem software (version 8.0), Then, the structures were fully optimized based on the ab initio method, usingDFT level of theory. Hyperchem, ChemOffice and Dragon (version 5.5) programs were employed to calculate the molecular descriptors . All calculations were performed using Gaussian 98W program series. Geometry optimization of compounds was carried out by B3LYP method employing 6–31G (d) basis set. In this study, the independent variables were molecular descriptors and the dependent variables were the actual half maximal inhibitory concentration (IC50) values. Overall, more than 3226 theoretical descriptors were

selected and calculated. Finally, Unscrambler (version 9.7) program was used for analysis of data and statistical methods.For each compound in the training sets, the correlation equation was derived with the same descriptors. Then, the obtained equation was used to predict log (1/IC50) values for the compounds from the corresponding test sets. In the present work, the method of stepwise multiple linear regression (stepwise, MLR) to select the most appropriate descriptor of all descriptor was used. Totally 3226 descriptors were generated that were too many to be fitted in our models. So, it was necessary to reduce the number of descriptors through an objective feature selection which was performed in three steps. The structures of the Indoland 7-azaindolesderivatives used in this study (Fig 1).

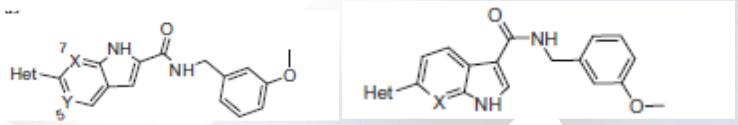


Fig. 1 The molecular structure ofIndol and 7-azaindoles derivatives.

Results and discussion

The selected descriptors through these methods were used to construct some linear and non-linear models by using MLR,ANN and ICA methods. The five most significant descriptors which were selected by ICA are as followsGATS4e,EEig06x.,*Mor6p ,G3m ,ISH* Based on the types of variable selection method and also the types of the feature mapping technique, these models can be shown as GA-MLR and GA-ANN. It revealed that the GA–ANN model was much better than other models is given in Table 2.

Table 1 Experimental and predicted values of log (1/IC50) using Jack–Knife.

Molecule	-log(1/IC50)	Jack–Knife	Δ
۱	-2.28	-2.61	0.34
۲	-0.30	-0.34	0.04
۳	-0.30	-0.71	0.41
۴	-0.30	-0.87	0.58
۵	-2.20	-2.38	0.18
۶	-1.56	-2.65	1.10
۷	-3.08	-2.11	-0.96
۸	-2.72	-2.47	-0.24
۹	-2.62	-1.12	-1.49
۱۰	-0.70	-0.95	0.25

Table 2 The statistical parameters of different constructed QSAR models.

Method	RMSE test	RMSE train	R^2
GA–ANN Jack–Knife	0.4218	0.1358	0.8555

Conclusion

In the present study, two linear and non-linear variable selection methods were used to select the most significant descriptors, and the MLR, ANN,ICAand GA were used to construct a quantitative relation between the activities heterocyclic Indol and7-azaindoles derivatives and their calculated descriptors. ANN has been successfully used for finding a QSAR model for heterocyclic Indol and7-azaindoles derivatives.

References

[1] C.T. Kresge, M.E. Leonowicz, W.J. Roth, J.C. Vartuli, J.S. Beck, *Nature* 359 (**1992**) 710.

[2] Hasegawa, K., Miyashita, Y., 1992. Chemom. Intell. Lab. Syst. 16, 69–75.

[3] RHuuskonen, J., 2000. J. Chem. Inf. Comput. Sci. 40, 773–777.

۱۱	-2.34	-1.79	-0.54
۱۲	-0.30	0.30	-0.60
۱۳	-0.85	-1.51	0.66
۱۴	-0.95	-0.55	-0.39
۱۵	-2.88	-2.73	-0.15
۱۶	-0.90	-0.47	-0.43
۱۷	-2.11	-2.86	0.76
۱۸	-2.23	-3.09	0.87
۱۹	-0.48	-0.29	-0.18
۲۰	-2.87	-2.72	-0.14
۲۱	-2.26	-2.76	0.50

Using QSAR calculation of amid derivatives for the treatment of Parkinson's disease

Molood Naziri* , Ghasem Ghasemi

Department of Chemistry, Islamic Azad University, Rasht Branch, Rasht, Iran. Email: Molood.Naziri@yahoo.com

Abstract

Parkinson's disease is one of the most common neurological disease is considered malicious. Parkinson's disease is a chronic condition and always in progress. The goal of treatment for patients with symptoms but not eliminate the symptoms under control [1]. On the pattern, quantitative structure–activity relationship (QSAR) study has been done on a series of Amid derivatives for the treatment of Parkinson's disease [2]. Multiple linear regression (MLR), partial least squares (PLS) and principal component regression (PCR) were used to create QSAR models. For this purpose, ab initio geometry optimization performed at B3LYP level with a known basis set (6–31G). Hyperchem, ChemOffice and Gaussian 03W softwares were used for geometry optimization of the molecules and calculation of the quantum chemical descriptors. Finally, Unscrambler program was used for analysis of data. In the present study, the root mean square error of the calibration and R^2 using MLR method, were obtained 0.898 and 0.806 According to the obtained results, we found that MLR model is the most favorable method toward the other statistical methods and is suitable for being used in QSAR models . All calculations were performed using Gaussian 09 program. Using Dragon (version 5.5) were employed to calculate the molecular descriptors, using MATAB for GA, ANN, MLR. In the work stepwise multiple linear regression (stepwise-MLR) and GA variable subset selection methods were used for the selection of the most relevant descriptors from all of discriptors. The root-mean-square errors of the training set and the test set for GA-ANN model using Jack-knife were 0.3354, 0.1405, R^2 = 0.8934, the training and test set for GA-ANN cross validation were 0.2319 , 0.1294.

Keywords: Parkinson's disease, Quantitative structure activity relationship, Amid derivatives, Multiple linear regression

References:

[1] F. V. Petegem, D. L. Jr. Minor. The structural biology of voltage-gated calcium channel function and regulation. Biochem Soc. Trans. 2006 (34) 887-893.

[2] Z. Xiang, A. D. Thompson, J. T. Brogan, M.l L. Schulte, B. J. Melancon, D. Mi, L. M. Lewis, B. Zou, L. Yang, R. Morrison, T. Santomango, F. Byers, K. Brewer, J. S. Aldrich, H. Yu, E. S. Dawson, M. Li, O. McManus, C. K. Jones, J. Scott Daniels, C. R. Hopkins, X. Simon Xie, P. Jeffrey Conn, C. David Weaver, C.W. Lindsley. The Discovery and Characterization of ML218: A Novel, Centrally Active T-Type Calcium Channel Inhibitor with Robust Effects in STN Neurons and in a Rodent Model of Parkinson's Disease. ACS Chem. Neurosci. (Article in press) 2011.

Comparison of Cadmium-induced toxicisity in ionic and nano forms applying ^{1}HNMR based Metabolomics analysis

Maryam Khoshkam[1,2]*, Ali Ramazani[2], Mehrdad Hamidi[2]

[1] *Department of Chemistry, University of Mohaghegh Ardabili, Ardabil 56199-11367, Iran*
[2] *School of Pharmacy, Zanjan University of Medical Sciences,45139-56184, Zanjan, Iran*

E-mail address (corresponding author): khoshkam@uma.com

Abstract

The study of metabolomics, as applied to toxicology, has applications that are unique, e.g. the identification of pathways of toxicity (PoT), signatures of toxicity (SoT) for regulatory evidence. In this study the toxicity of CdTe quantum dots (QDs) and bulk Cd^{2+} has been investigated and compared applying metabolomics methods. The datasets were ^{1}HNMR data from mice plasma which had been taken from four groups of mice in different time intervals. Then the data were analysed applying chemometrics methods and the metabolites were found from human metabolome data base (HMDB). The results show the significant change in level of some metabolites in diffedrent groups with different amounts of received Cd.The effect of applying different preprocessing methods in the obtained metabolites was studied, as well.

Key words: NMR based metabolomics; in vivo; PCA; CdTe; nanotoxicometabolomics.

Introduction

Recent advances in nanotechnology have led to increased production and use of many types of nanomaterials in many fields of science. Nanotoxicometabolomics methods are in fact the metabolomics study of toxicity of nano materials in life science [1-3]. These studies has application that is unique e.g. the identification of pathways of toxicity (PoT), signatures of toxicity (SoT) for regulatory evidence and applications feeding into systems toxicology approaches, which are to some extent distinct from other areas in the life sciences [4].

The study of semiconductor quantum dots (QDs) has received wide interest in the past two decades due to their unique size dependent optical properties, efficient emissions, and potential applications in biological and biomedical research, electronics, and light-emitting devices[5, 6]. However, these nano particles can be toxic to aquatic organisms and poses significant ecological risks. The sources of this toxicity for QDs like CdTe are first the presrnce of Cd^{2+} as a heavy metal and the last is the risk of potential toxicity of particles in nano scale which is general for all types of nano particles. However, the toxicological impacts of heavy metal constituted QDs on aquatic species are largely unknown, especially at the molecular level. Moreover, it is crucial to determine whether the toxicity of metallic QDs is quantitative or mechanistically different than that of soluble metal components. The aim of this study was to compare the toxicities of ionic cadmium (Cd) and and Cd- containing QDs at in vivo level. To our knowledge this study is the first toxicometabolomics study of CdTe QDs and ionic Cd^{2+}.

Results and discussion

In order to study the toxicity of CdTe QDs and ionic Cd^{2+}, four groups of mices were selected. There were 5 mices in each group; Group(I) was control group and group(II) and (III) were the mices which they received 5mg/Kg and 10 mg/Kg of CdTe QDs in PBS. Group (IV) were the mices which they received 5mg/Kg ionic Cd^{2+}. The blood sample of mices was taken in different time intervals and the plasma samples were kept in -80ºC until used. ^{1}HNMR spectra of plasma were recorded according to CPMG protocol and were analysed by PCA after proper pretreatment like mean centering, normalization, alignment and data bining. After applying PCA the groups were seprated and shown in score plot. Loading plot of data was used to show the chemical shifts of metabolites which were changed in the exposure of CdTe QD and Cd^{2+}. The results show the significant change in the level of some metabolites of mice and also show that Cd^{2+} is shown more toxic effects than equivalent amount of Cd from CdTe QDs which this refers to ionic nature of Cd^{2+}. After finding the changed metabolites, the biological pathways were investigated to show that which pathways will be changed.

References

[1] Ratnasekhar, C.; Sonane, M.; Satish, A.; Mudiam, M. K. R., Metabolomics reveals the perturbations in the metabolome of Caenorhabditis elegans exposed to titanium dioxide nanoparticles. *Nanotoxicol. 2015;* **9(8)**: 994-1004.

[2] Schnackenberg, L.; Sun, J.; Beger, R., Metabolomics Techniques in Nanotoxicology Studies. Reineke, J., Ed. Humana Press: 2012; Vol. 926, pp 141-156.

[3] Wei, L.; Liao, P.; Wu, H.; Li, X.; Pei, F.; Li, W.; Wu, Y., Metabolic profiling studies on the toxicological effects of realgar in rats by 1H NMR spectroscopy. *Toxicol. Appl. Pharm.* 2009; **234 (3)**: 314-325.

[4] Bouhifd, M.; Hartung, T.; Hogberg, H. T.; Kleensang, A.; Zhao, L., Review: Toxicometabolomics. *J. Appl. Toxicol.* 2013; **33 (12)**: 1365-1383.

[5] Abd El-sadek, M. S.; Ram Kumar, J.; Moorthy Babu, S., The role of potassium tellurite as tellurium source in mercaptoacetic acid-capped CdTe nanoparticles. *Curr. Appl. Phys.* 2010; **10 (1)**: 317-322.

[6] Yun, Z.; Zhengtao, D.; Jiachang, Y.; Fangqiong, T.; Qun, W., Using cadmium telluride quantum dots as a proton flux sensor and applying to detect H9 avian influenza virus. *Anal. Biochem.* 2007;**364 (2)**: 122-127.

Quality control of standard saffron samples from food colorant adulteration using chemometrics analysis of transmittance FT-IR spectra

Sadegh Karimi [1*], Fatemeh Mehrjo[1], Javad Feizi[2]

[1]*Department of Chemistry, College of Science, Persian Gulf University, Bushehr, Iran*

[2]*Testa quality control laboratory, Mashhad -Iran*

E-mail address (corresponding author): sakarimi@pgu.ac.ir

Abstract

The aim of the present study is to investigate one of the applications of transmittance FT-IR spectroscopy and pattern recognition to check the quality of standard saffron which is suffered from various types of food colorants. Transmittance FT-IR spectra have been obtained for reference saffron samples and those which have been mixed with six well known food colorants namely Tartarzine, Sunset yellow, Azorubine, Quinoline-yellow, Allura red and Sudan II. GA-LDA based clustering of variable concept has been applied to transmittance FT-IR spectra for classification of standard saffron from fraud ones. Analysis of the selected clusters of variables indicates that two regions band corresponding to 2400-2800 and 3795-3876cm^{-1} are responsible for differentiation of standard samples from their adulteration ones. Transmittance FT-IR combined with clustering of variable has been found to be efficient, promising, sensitive and rapid tool in the detection food colorant in trade saffron.

Key words: Saffron; Quality control; Food authenticity; Pattern recognition; clustering of variable concept.

Introduction

Food quality control (authenticity) is one of the increasingly important and vital subjects for consumers, regulatory agencies, and food industry. Meanwhile fraud detection by routine analytical methodologies is usually difficult and time consuming [1]. On the other hand, saffron is one the most expensive species in food industry. In addition, this product is just produced in a few countries such as Iran and Spain. These two factors cause that the saffron can be good candidate for adulteration conducted for economic gain and has been subjected to various types of adulteration over the centuries. Transmittance FT-IR spectroscopy is a simple analytical technique largely applied for its rapidity and reproducibility in food fraud detection. Although transmittance FT-IR spectra have high throughput out for each samples, this subject led to create small sample size problem. In this condition, the classification methods such as LDA have a tendency to show over-fitting result. This subject can be solved using clustering of variable concept for transmittance FT-IR spectra before LDA analysis. The current study presents an approach to discriminate adulterations in saffron by means of transmittance FT-IR and pattern recognition method.

.

Results and discussion

The transmittance FT-IR spectra of all studied saffron samples (standard and fraud ones) have been collected in a data matrix D of the dimension of ($n_s \times n_w$), where n_s and n_w are the number of sample and wavenumber respectively. Data matrix (D) has been divided into the calibration and prediction sets by the DUPLEX algorithm. In our case 42 samples have been included in the training and the remaining 28 samples have been selected as test. In the first step of GA-LDA based on SOM, Kohonen SOM is applied to cluster the wavenumber based on their similarity [2]. Each *n*-node Kohonen SOM model leads to ($n \times n$) cluster of variables. Therefore, the number of clusters (q) produced by each Kohonen map model is equal to n^2. Seven Kohonen SOM networks from the node sizes of 2×2 to 8×8 have been checked. The distribution pattern of wavenumbers in the (4 × 4) SOM network is presented in Figure 1. Each cluster is subjected to PCA analysis separately and the meaningful PCs and loading are extracted. Extracted PC of whole clusters builds a new data matrix (D_{new}) which their columns are significant principal components retained from these local clusters. This new data set have been used to constructions the classification model using linear relation by genetic algorithm PC selection. Table 1 lists the statistical classification parameters of the obtained models from different number of clusters through Kohonen SOM method. The Not Error Rate (NER) of calibration, validation and prediction reveals that GA-LDA model obtained from Kohonen nodes q=4, (16 cluster) is the optimum one for both calibration and prediction classification ability. Finally, the discriminant function plot (DF1) of Kohonen network size q = 4 is given in Figure 2. As it is evident, a clear separation between samples from the LDA plot of this cluster size is observed.

Fig1. Distribuition of wavenumber in 4×4 Kohonne network

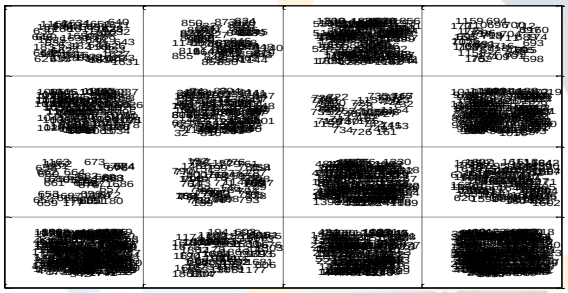


Table 1. Statistical parameter of the GA-LDA based SOM for different network size.

Number of segments (Kohonen nodes)	N_{EPC}	N_{SPC}	NER_{cal}	NER_{val}	NER_{pre}
4 (2×2)	24	2	0.9522	0.8529	0.8129
9 (3×3)	33	3	1.000	0.8771	0.8356
16 (4×4)	**87**	**4**	**1.000**	**1.000**	**1.000**
25 (5×5)	89	4	1.000	0.9532	0.9433
36 (6×6)	91	3	1.000	0.8923	0.9104
49 (7×7)	94	4	1.000	0.8661	0.8588
64 (8×8)	95	4	1.000	0.8433	0.8215

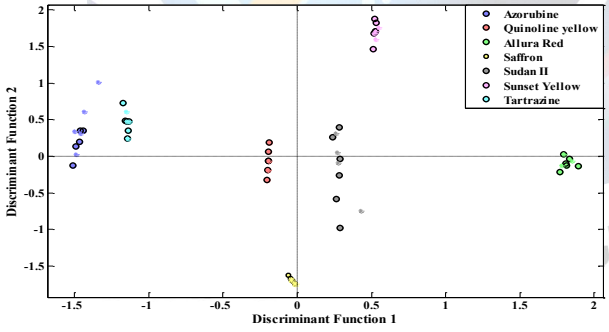


Fig 2. Classification using GA-LDA based on SOM for adulteration in saffron data set.

Refrencs:

1. Petrakis, E. A.; Cagliani, L. R.; Polissiou, M. G.; Consonni, R., Evaluation of saffron (Crocus sativus L.) adulteration with plant adulterants by 1 H NMR metabolite fingerprinting. *Food chemistry* 2015**, 173:** 890-896**.**

2. Karimi, S.; Farrokhnia, M., Leukemia and small round blue-cell tumor cancer detection using microarray gene expression data set: ombining data dimension reduction and variable selection technique. *Chemome. Intell. Lab.Syss* 2014**, 139**:6-14**.**

Gene expression classification (tumor or non-tumor) using wavelet-based feature extraction method

Zahra Heidari, Jahan B. Ghasemi*

Faculty of Chemistry, K. N. Toosi University of Technology, Tehran, Iran

School of Chemistry, University College of Science, University of Tehran, P.O. Box 14155-6455, Tehran, Iran.

E-mail address: jahan.ghasemi@gmail.com

Abstract

Microarray-based tumor gene expression profiles have made cancer diagnosis possible. However, due to ultra-high dimensionality, the biological interpretation of these large data sets is difficult. Therefore there is an urgent need to apply variable selection methods to discover the most informative genes. Here, in this study we propose the application of a wavelet based feature extraction method for selecting the most discriminant features in order to classification of tumor and normal tissues in pancreatic cancer samples. Our results confirm that the CWT-based feature extraction method is a very useful tool for introducing the most discriminant features to SVM and PLS-DA models for constructing models of diagnostic cancer using gene expression data.

Key words: Gene expression, Wavelet transform, feature extraction, classification, Cancer diagnosis, SVM, PLS-DA.

Introduction

Gene expression microarray technology generates an image of expression level of thousands of genes within a tissue or cell sample. This technology provides useful information that can be used for cancer diagnosis. These microarrays are characterized by many features (genes) on only a few samples. In spite of the large number of genes only a few underlying genes are differentially expressed across the samples. Therefore detecting and extracting the features that are differentially expressed across the samples is one of the challenges in microarray data analysis. Many statistical methods are used to extract informative features among existing genes in a microarray [1].

Wavelet transform is one of the methods that has been widely used for data mining purposes and its usefulness has also been proved in the domain of microarray data analysis. Wavelet transform is a multiresolution method that manifests the information content of a signal at different scales as wavelet coefficients [2]. In this study, we approach the feature extraction using a new developed wavelet-based method that extracts the most discriminative features among microarray features [3]. Extracted features are used to be introduced to classification models in ordet to classifying tomur and non-tomur tissues.

Results and discussion

The proposed feature extraction method is tested on a cancer microarray data set consisting of 16 normal (non-tumor) and 36 tumor-samples [4]. Applying this method in 5 scales using db2 could reduce the number of variables more that 91%. The results of performing the principal component analysis on the initial space of features and extracted features are shown in figure 1. It shows that using extracted features more efficient class separation is provided. Moreover, developing different classification models, Partial Least Squares Discriminant Analysis (PLS-DA) and Support Vector Machine (SVM) on selected features yields classifiers with a better performance than the models developed on original feature space.

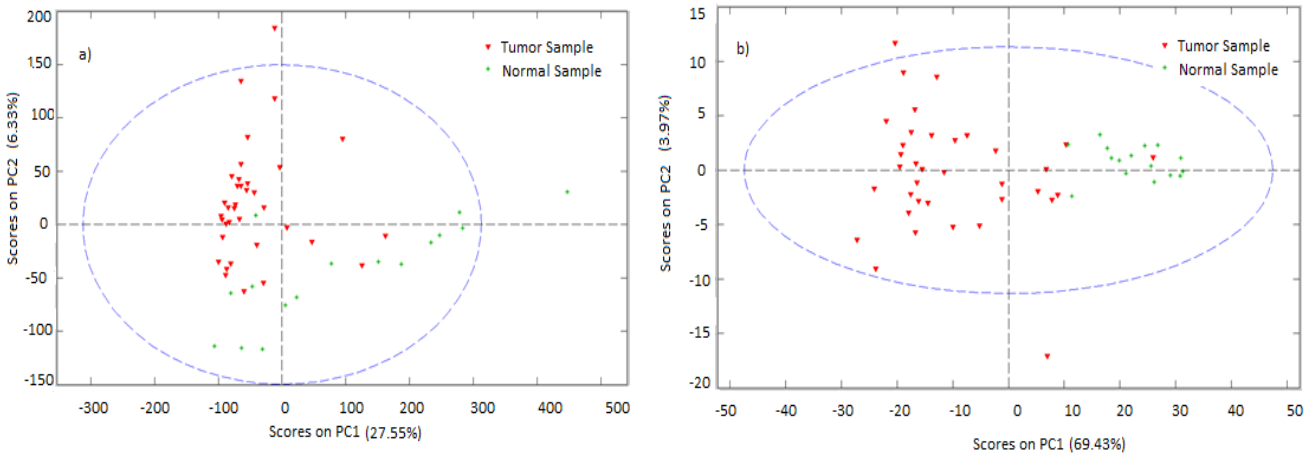


Fig. 1. Score plots of the first two PCs for tumor and normal (non-tumor) samples (a) in initial space of gene expression and (b) in feature space provided by wavelet based method.

References

[1] Lobenhofer EK, Bushel PR, Afshari CA, H K, Progress in the application of DNA microarrays, *Environ Health Prospect*, 2001; **109(9)**: 881–89.

[2] Mallat S.G, Multiresulotion Approximations and Wavelet Orthonormal Bases of L2(R), *Transactions of the American Mathematical Society*. 1989; **315(1)** :69-87.

[3] Ghasemi JB, Heidari Z, Jabbari A. Toward a continuous wavelet transform-based search method for feature selection for classification of spectroscopic data. *Chemometr Intell Lab Syst.* 2013; **127** :185-194.

[4] Pei H, Li L, Fridley BL, Jenkins GD, Kalari KR, Lingle W, Petersen G, Lou1 Z, Wang L. FKBP51 affects cancer cell response to chemotherapy by negatively regulating Akt. *Cancer cell.* 2009; **16(3)** :259-266.

Multivariate optimization and determination of Hg(II) content in aqueous samples after preconcentration with modified graphene oxide

Nayereh_Amiri[1,*], Mohammad K. Rofouei[1], Jahan B. Ghasemi[2]

[1]*Faculty of Chemistry, Kharazmi University, Tehran, Iran*

[2] *School of Chemistry, University College of Science, University of Tehran, P.O. Box 14155-6455, Tehran, Iran.*
Nayereh_amiri@yahoo.com

Abstract

A new modified graphene oxide (GO) was prepared as a solid adsorbent and used for preconcentration of trace amounts of Hg(II) in water samples prior to the determination by ICP-AES. The method was based on the sorption of Hg(II) ions on GO which is functionalized by a new ligand; (1-(p-acetyl phenyl)-3-(o-methyl benzoate) triazene (AMBT)). In the adsorption step, two-level full factorial design and central composite design were used to evaluate and determine the optimum level of important variables (pH, amount of adsorbent and adsorption time). In the desorption step, some factors such as type, concentration, amount and flow rate of eluting agent were investigated. Many common ions did not interfere. Under the optimal conditions, the adsorption capacity, limit of detection and relative standard deviation were found to be 80 mg g^{-1}, 0.1 µg L^{-1} and lower than 2.5% (n=3) respectively. The present method was successfully applied for determination of mercury ions in real water samples.

Key words: 1-(p-acetyl phenyl)-3-(o-methyl benzoate) triazene, Full factorial design, Central composite design

Introduction

Determination of mercury in nature is very important due to its highly toxic and bioaccumulative properties that cause severe damage to human health and living organisms. Mercury ions could be transformed into methyl mercury during bacterial activities in aqueous solutions and so it could be entered into human body through the food chain. The permissible mercury limits set by Word Health Organization (WHO) and US Environmental Protection Agency (EPA) for drinking water are 5 µg L^{-1} and 2 µg L^{-1} respectively [1]. Several quantitative analyzing techniques for determination of Hg have been reported based on ICP-AES [1, 2], ICP-MS [3], fluorescence spectroscopy [4] and atomic absorption spectrometry [5]. Due to the low permissible level of mercury and the presence of interfering matrix components, the analyte should be determined in environmental samples after using a preconcentration step. The solid phase extraction as preconcentration method has advantages such as consumption of small volume of organic solvents, a high enrichment factor, simplicity and rapidity. In this work, preconcentration of trace mercury (II) in aqueous media using GO-AMBT and subsequent determination by ICP-AES have been proposed. Full factorial design and central composite design were used to evaluate the three parameters of adsorption step.

Results and discussion

The AMBT compound was purified by repeated recrystallization in acetonitrile. X-ray crystallographic analysis, CHN analysis, FTIR and HNMR spectra were used to characterize the structure of AMBT. All the experimental results confirmed the structure of the triazene compound. FTIR spectrum, CHN elemental analysis and SEM images were used

to characterize GO-AMBT. All experimental results confirmed that GO was modified with AMBT. The interaction between Hg(II) and the triazene compound was studied by spectrophotometric experiments.

In the adsorption step, three factors e.g. pH, adsorbent mass, sonicating time were considered in the evaluation of adsorption of Hg(II) on the modified GO. A two-level full factorial design with 8 (2^3) runs was carried out to find the main factors and their interactions. Analysis of variance was used to evaluate the significance of the variables. The pH, ligand mass and sonicating time were effective factors while interactions among these factors were not considered significant. The value of R^2 was calculated as 0.99. A central composite design with 20 runs was carried out to optimize the values of the three significant variables and to obtain response surface model (RSM). According to ANOVA table, the model F-value of 19.7, R^2-value of o.97, low p-value of lack of fit (<0.0001) and low standard deviation (~1) establish the validity of the cubic model. According to the overall results of optimization study, the following optimal set points were chosen: pH 4.5, amount of sorbent 13 mg and sonicating time 13 min.

In desorption Step, combination of 0.5 mL H_2O_2 (5%) and 2 mL aqua regia (5%) was found to be the proper eluting reagent. complete extraction efficiency can be achieved with the use of medium and slow filter papers in the desorption step.

The break through volume was obtained 300 mL. The effects of some counter ions were investigated. The results showed that in excess of 1000-fold Na^+, K^+, Ca^{2+}, Mg^{2+}, Mn^{2+}, Cr^{3+}, Co^{+2}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, Pb^{2+}, Ag^+, Cl^-, NO_3^-, SO_4^{2-} and CH_3COO^-, 200-fold Cd^{2+} and 100-fold Fe^{2+} have no important interferences in the preconcentration and determination of the analyte. The maximum capacity of sorbent, extraction percentages, RSD, LOD and linear range was found to be 80 mg g^{-1}, >99%, 2.5 (n=3), 0.1 µg L^{-1} and 0.2-150 µg L^{-1} respectively. The application of the recommended method was investigated by determination of the analyte in the spiked tab water, river water and mineral water samples. The results showed the agreement between the actual and measured analyte amounts. Therefore, the proposed procedure can be appropriately used for analysis of different matrices.

References

[1] Vasimalai N, Abraham John S, Ultrasensitive and selective spectrofluorimetric determination of Hg(II) using a dimercaptothiadiazole fluorophore, *J. Lumin.* 2011; 131: 2636-2641.

[2] Rofouei M K, Hosseini S M, Khani H, Rahimi-Alangi S, Highly selective determination of trace quantities of Hg(II) in water samples by spectrophotometric and ICP-OES methods after cloud point extraction, *Anal. Method* 2012; 4: 759-765.

[3] Huang C, Hu B, Silica-coated magnetic nanoparticles modified with γ- mercaptopropyltrimethoxysilane for fast and selective solid phase extraction of trace amounts of Cd, Cu, Hg, and Pb in environmental and biological samples prior to their determination by ICP-MS, *Spectrochim. Acta B* 2008; 63: 437-444.

[4] Huang C C, Yang Z, Lee K H, Chang H T, Synthesis of highly fluorescent gold nano-particles for sensing mercury (II), *Angew. Chem. Int. Ed.* 2007; 119: 6948-6952.

[5] Adlnasab L, Ebrahimzadeh H, Asgharinezhad A A, Nasiri Aghdam M, Dehghani A, Esmaeilpour S, A preconcentration procedure for determination of ultra-trace mercury (II) in environmental samples employing continuous-flow CV-AAS, *Food Anal. Method* 2014; 7: 616-628.

Simultaneous quantification of triphenylmethane dyes using three-way data analysis

Mehrnoosh Bayat*, Jahan B. Ghasemi and Farzaneh Shemirani

School of Chemistry, University College of Science, University of Tehran, P.O. Box 14155-6455, Tehran, Iran.

E-mail address (corresponding author): Mehrnoosh_bayat@ut.ac.ir

Abstract

In this research a new method for generation of three way data, combined with second-order calibration methods for quantification of two important dyestuffs; malachite green and crystal violate, in real samples was reported. In this work we used two independent variables; wavelength and gradual changing in the hydrophobicity of media; to generate second order data. Several two component solutions for calibration set were made at constant temperature and resulting solutions were titrated by a sodium dodecyl sulfate (SDS) solution in a stepwise manner. The obtained matrix can be folded to form a three-way data matrix $\bar{X}$. Before the addition of the first increment the absorption spectrm of the solution was recorded in the range 380-780 nm and the subsequent spectra were obtained after each addition of SDS solution. The effects of several important variables were investigated and thoroughly discussed. The data processing was performed by MCR-ALS, GRAM and Bilinearization methods in MATLAB environment by PLS_Toolbox.

Key words: Second-order advantage, MCR-ALS, Dyes, Real sample.

Introduction

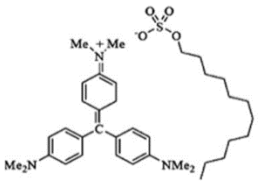
Triphenylmethane dyes, such as malachite green (MG) and crystal violet (CV) have been widely used to prevent and treat infections in aquaculture products due to high efficacy, availability and low price [1]. Due to the carcinogenic and mutagenic properties they are banned, in many countries all over the world, for use in fish for human consumption [2]. So they have to be analyzed simultaneously in the presence of many chemical interferents. In this case, sophisticated instrumentation and mathematical tools are available to deal with the complexity [3]. Second-order calibrations are particularly attractive for determinations in complex matrcies. There are several methods to generate three-way data. These types of data have a very efficient advantage known as second order advantages. In this approach a data matrix will be obtained for each single sample and grouping matrices for all calibration samples give three-dimensional data known as three-way array [4].

Results and discussion

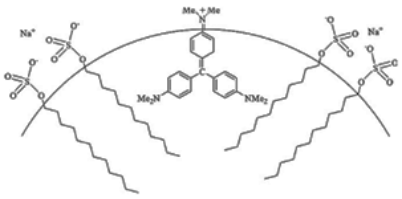
The CV and MG -SDS monomer ionic pair formation results in a decrease of the electrostatic repulsions between the $-N^{+}(CH_3)_2$ groups of dyes and simultaneously, formation of small premicellar aggregates which result in screening of dye molecules by the SDS monomers. This screening results in a destruction of dimer structures, because

$-N^{+}(CH_3)_2$ groups of the dye will interact simultaneously with various surfactant head groups $-OSO^{-}_3$ (Scheme 1). In the presence of large SDS concentrations, [SDS] $>8.0*10^{-3}$M, the micellization process begins, and the dye molecules will bind to the micellar surface where the $-N^{+}(CH_3)_2$ groups are solvated by water and screened by SDS head groups (Scheme 2). The aromatic rings of dyes will be incorporated to the micellar core to minimize their interactions with water [5]. It should be noted that the dye concentration used in our experimental conditions is lower enough to guarantee that no dimers or ion pairs are present in the absence of surfactant.

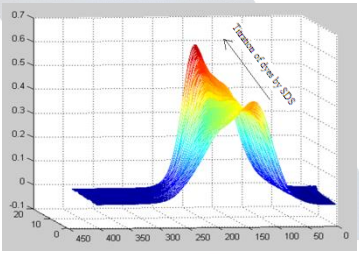
As a result, by titration of dyes solution by SDS solution in a stepwise manner, after the addition of the first increment, CV and MG -SDS ionic pairs were formed. By continuing addition of SDS, ionic pair structures were destructed and monomer spectra of dyes were observed (scheme 3).



Scheme 1



Scheme 2



Scheme 3

References

[1] L´ opez-Guti´ errez N, Romero-Gonz´ alez R, Luis Mart´ınez Vidal J, Garrido Frenich A. Analysis of triphenylmethane dyes in seafood products a review of extraction methods and determination by liquid chromatography coupled to mass spectrometry. *Anal. Methods* 201; **5**:3434.

[2] Kaplan M, Olgun EO, Karaoglu O. A rapid and simple method for simultaneous determination of triphenylmethane dye residues in rainbow trouts by liquid chromatography tandem mass spectrometry. *J. Chromatogr. A.* 2014;**1349**:37-43.

[3] Khani R, Ghasemi JB, Shemirani F. Second-order data obtained by beta-cyclodextrin complexes: A novel approach for multicomponent analysis with three-way multivariate calibration methods. *Talanta* 2014; **128**:254-262.

[4] Khani R, Ghasemi JB, Shemirani F, Rahmanian R. Application of bilinear least squares/residual bilinearization in bulk liquid membrane system for simultaneous multicomponent quantification of two synthetic dyes. *Chemometr. Intell. Lab*. 2015; **144**:48-55.

[5] Garcı´a-Rı´o L, Hervella P, Mejuto J.C, Parajo´ M. Spectroscopic and kinetic investigation of the interaction between crystal violet and sodium dodecylsulfate. *Chem. Phys.* 2007; **335**:164-176.

Multiset analysis of ATR-FT-IR and synchronous flourscence spectra by multivariate curve resolution analysis to study the hydrogen-bond disruption in Ethylene Glycol-water binary mixtures

Bahram Hemmateenejad*, Elaheh Talebanpour-Bayat

Department of Chemistry, Shiraz University, Shiraz, 71457, Iran

E-mail address: hemmatb@sums.ac.ir

Abstract

The disruption of hydrogen bonding of ethylene glycol (EG) and EG–water (W)binary mixtures in the liquid phase as a function of the molar fraction studied using augmentation of infrared and Constant energy synchronous fluorescence spectroscopy methods followed by multivariate analysis. The structure of the cluster was proposed as association of two molecules of water with a molecules of EG ($EG(W)_2$). Beacause of baseline drift, Multivariate Curve Resolution Method – Net Analyte Signal(MCR-NAS) was used for calculation of the equlibrium constant of this cluster. Overall, the results clearly illustrate that the presence of waterdisrupts the ethylene glycol−ethylene glycol hydrogen bonds.

Key words: Multivariate curve resolution, Net analyte signal, Ethylene glycol-water binary mixture, Baseline drift.

Introduction

Ethylene glycol (EG) is one of the most important cryoprotectants and it has a broad range of applications in industry (e.g. heat-transfer engineering, food addititives, paint formulations chemical intermediates and etc.), biology and medicine [1].Some of the major interesting properties of EG molecules are related to its competitive capability of forming inter-intra molecular hydrogen bonding. Hence, its properties were strongly influenced by hydrogen-bonding interactions.

Interaction of hydrogen bonded glass forming ethylene glycol with water is very interesting since due to the good solubility of ethylene glycol in water this molecule may form a variety of clusters due to intermolecular and intramolecular geometry of individual molecules. So an examination of the structure and stability of EG−W clusters could elucidate the structural features of this molecule in aqueous media and related properties of EG at the molecular level [2].

MCR-ALS that completely decompose a two-way data matrix into the product of concentration and spectral matrices are meant to describe processes without using explicitly the underlying chemical model linked to them[3].

The effect of concentration on molecular structure and hydrogen bonding of binary mixtures of EG–water are investigated using augmentation of infrared and Constant energy synchronous fluorescence spectroscopy methods followed by multivariate analysis.

Results and discussion

The Attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy (ATR-FTIR) and Constant energy synchronous fluorescence (CESF) spectra of various binary mixtures of water and EG bwere recorded at 25 ^{0}C. The data matrices of two spectroscopic methods were augmented and then were analyzed by MCR-ALS method. In the rank analysis of the row-wise augmented matrix by singular value decomposition (SVD) and evolving factor analysis (EFA), three siginificant chemical components were detected. Pure concentration profile of MCR-ALS has been shown in Fig. 1.The structure of the cluster was proposed consisting of two molecules of water that may interact both with a molecules of EG (EG(W)$_2$). L. C. Gomide Freitas et. Al. [4] propsed that the most stable isomer is one in which water molecules bridge the two hydroxyl groups of EG. The equilibrium constant of this cluster was calculated 1.85×10^{-5} based on MCR-NAS.

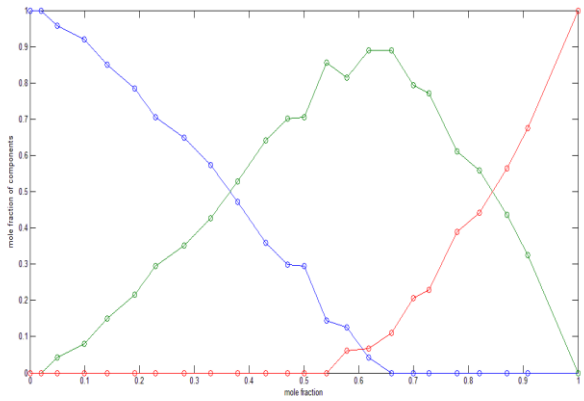


Fig.1. concentration profile of augmented spectra of ATR-FTIR andCESF at 25 ^{0}C.

References

[1] ChaudhariA,SL Lee,A computational study of microsolvation effect on ethylene glycol by density functional method. *J. Chem. Phys. 2004;* **120:**7464-7469.

[2]Kumar RM, Baskar P, Balamurugan K, Das S, Subramanian V,On the Perturbation of the H-Bonding Interaction in Ethylene GlycolClusters upon Hydration .*J. Phys. Chem. A* 2012; **116**: 4239−4247.

[3] Ruckebusch C, Blanchet L, Multivariate curve resolution: A review of advanced and tailoredapplications and challenges. *Anal.Chim.Acta* 2013; **765** : 28– 36.

[4] Oliveira OVd, Gomide Freitas LC, Molecular dynamics simulation of liquid ethylene glycoland its aqueous solution.*j. mol. struc –theochem. 2005*; **728** : 179–187.

Mean centering of ratio spectra for colorimetric recognition of morphine using melamine modified gold nanoprobe

Naimeh Mohseni*, Morteza Bahram

Department of Analytical Chemistry, Faculty of Chemistry, Urmia University, Urmia, Iran
E-mail address: n.mohseni@urmia.ac.ir

Abstract

The colorimetric assay for morphine presented here is based on melamine-modified gold nanoparticles (mel-AuNPs) and the finding that hydrogen bonding between melamine and morphine results in the aggregation of mel-AuNPs and a consequent color change of AuNPs from wine red to blue. We used mean centering of ratio spectra for blank bias error elimination. This procedure gives more accurate results than the traditional approach using the absorbance ratio of the aggregated to the dispersed AuNPs. The system exhibited a liner range, from 0.07 μM to 3.0 μM with a correlation coefficient of 0.9934 and a 17 nM detection limit (3σ/m). Comparison of linear range, detection limit, regression coefficient and reproducibility (R.S.D.) of traditionally method with those of mean centering ones slightly improves the figure of merit of the method.

Key words: Melamine . Gold nanoparticles . Morphine . Colorimetric detection . Mean centering of rartio spectra

Introduction

Morphine (MOR) as a major component in opium is frequently used to relieve severe pain for patients, but when overdosed or abused, it is toxic and can cause disruption in the central nervous system. It is recommended by the World Health Organization (WHO) for the relief of moderate cancer-related pain [1]. Thus, simple, sensitive and selective determination of MOR in biological samples is a vital of interest in control of drug abuse, clinical toxicology and forensic cases.

Gold nanoparticles (AuNPs) have received much attention and are widely utilized in nanoscience, owing to their excellent biocompatibility and unique optical and electrochemical properties [2,3]. Sensors based on the color change of AuNPs have been applied to determine many substances; however, as far as we are aware, this is the first demonstration for the simple direct analysis of MOR by colorimetric assay.

The reagent blank, the response obtained from an analyte-free sample solution, is considered as an integral part of measurement technique and its response usually is subtracted from the standard response. Mean centering of ratio spectra can be used for blank bias error elimination [4,5].

Results and discussion

Melamine molecule contains three exocyclic amino groups and a three nitrogen hybrid ring, thus it may strongly coordinate to AuNPs by the ligand exchange with weakly surface-bound citrate ions. First, mel-AuNPs were synthesized by the self-assembly of melamine onto the surface of AuNPs then the MOR was added to the mel-AuNPs solution. MOR has a high ability to link to melamine through the hydrogen bonding interaction, which results in the aggregation

of mel-AuNPs. Upon the gradual addition of different concentrations of MOR together with 0.9 μM melamine, the absorption at 520 nm decreased in intensity and a new absorption peak appeared at 690 nm, which were ascribed to the aggregation of AuNPs. The absorption ratio A_{690}/A_{520} was used to quantify the concentration of MOR. A typical plot of this absorption ratio versus the concentration of MOR was obtained. We also used mean centering of ratio spectra for the determination of the analyte concentration free from bias error (Figure 1 (b)). The analytical characteristics of the method are summarized in Table 1. The figure of merit of the method improved dramatically when using mean centering of ratio spectra in order to removing the blank absorbance.

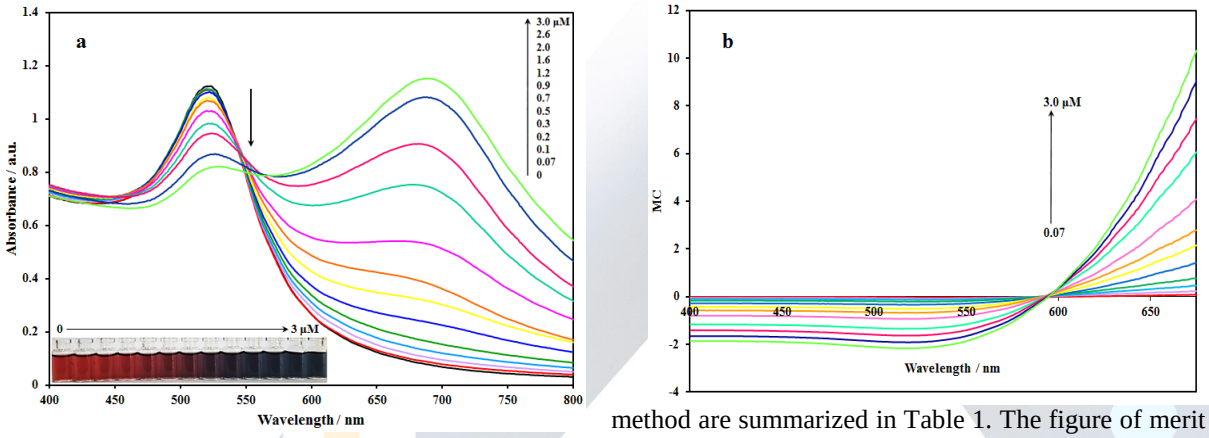


Figure 4. ***UV–Vis spectra and photographic images of mel–AuNPs solutions with different concentrations of MOR (a) and mean cetering of their ratio spectra (b).***

	Absorption ratio A_{690}/A_{520}	Mean centering of ratio spectra
Concentration range (μM)	0.1-3.0	0.07-3.0
Slope	0.4721	4.2297
Intercept	-0.0764	-0.2791
Correlation coefficient	0.9944	0.9934
Limit of detection (nM)	26	17
%R.S.D. (n = 8) (for 2 μM MOR)	2.70	1.20

Table 1. Analytical characteristics of the method.

References

[1] Dehdashtian S, Gholivand MB, Shamsipur M, Kariminia S. Construction of a sensitive and selective sensor for morphine using chitosan coated Fe_3O_4 magnetic nano particle as a modifier. *Matter. Sci. Eng.* C. 2015; **58** :53–59.
[2] Xiao L, Wildgoose GG, Compton RG, Sensitive electrochemical detection of arsenic (III) using gold nanoparticle modified carbon nanotubes via anodic stripping voltammetry. *Anal. Chim. Acta*. 2008; **620** :44-49.
[3] Chi H, Liu B, Guan G, Zhang Z, Han MY. A simple, reliable and sensitive colorimetric visualization of melamine in milk by unmodified gold nanoparticles. *Analyst.* 2010; **135** :1070-1075.
[4] Afkhami A, Madrakian T, Bozorgzadeh E, Bahram M, Spectrophotometric determination of beryllium in water samples after micelle-mediated extraction preconcentration. *Talanta*. 2007; **71** :1103-1109.
[5] Afkhami A, Bahram M. Mean centering of ratio spectra as a new spectrophotometric method for the analysis of binary and ternary mixtures. *Talanta*. 2005; **66** :712–720.

Principal Component Analysis of Molecular Dynamics Trajectories on Unbound and Agonist-Bound Dopamine D2

Shabnam Hooshmand[1*], Jahan B. Ghasemi[2]

[1]*Chemistry Department, Faculty of Sciences, Golestan University, Gorgan, Iran.*
[2]*Chemistry Department, Faculty of Sciences, K N Toosi University of Technology, Tehran, Iran*

E-mail address: shabnam_hooshmand@yahoo.com

Abstract.

Molecular dynamics (MD) simulations were performed for investigating the effect of an agonist on Dopamine D2 receptor as an anti-Parkinson agent. Employing the constructed 3D structure of Dopamine D2 receptor obtaining from homology modeling procedure as a starting structure, we carried out MD simulations of: (i) the free Dopamine D2; (ii) the complex of a potent agonist and D2 receptor. Producing the later ligand-protein complex, the ligand was docked in the active site of D2. The trajectory of the 10ns MD runs was analysis using PCA method. Comparing the results show that there are almost a similar trends in the major movement direction of both free and complex protein but with about 2ns delay for free Dopamine.

Key words: Dopamine, Parkinson, Molecular Dynamics Simulation, PCA

Introduction

Dopamine neurotransmitter and its receptors play a crucial role in the cell signaling process responsible for information transfer in neurons functioning in the nervous system. Development of improved therapeutics for some disorders like schizophrenia and Parkinson's disease would be considerably enhanced with the availability of the 3D structure for the Dopamine receptors and of the binding site for Dopamine and other agonists and antagonists. Dopamine receptors (DR) belong to the super family of GPCRs. Since GPCRs are bound to the membrane, making it difficult to express in sufficient quantities for crystallization. In the absence of high-resolution crystallographic data, homology modeling plays a significant role in prediction of the 3D-structure of the proteins. On the other hand, Molecular dynamics (MD) is the only technique available for obtaining dynamic protein data at atomic spatial resolution and picoseconds or finer temporal resolution. Recently, the use of MD in biological research, both to examine phenomena that cannot be resolved experimentally and to generate hypotheses that direct further experimental research, was significantly increased.

Results and discussion

According to the lack of tree-dimensional structure, in this study the construction of Dopamine D2 model was performed using the homology modeling module of MOE (Molecular Operating Environment) package based on the Dopamine D3 crystal structure (PDB codes 3PBL) as template. The best model of protein structure was minimized with MOE using the AMBER94 force field. The derived model that adapted the Ramachandran plot criteria was selected for more

refinement applying a molecular dynamics simulation (MDS). One of the most potent D2 agonist was chosen from literature [1]. To investigate the binding site interactions of Dopamine D2 against the agonists, a flexible molecular docking was performed using AutoDock Vina. The optimum docking conformation with a more negative docking score value which express more favorable interaction within the Dopamine D2 binding site was chosen for further assay employing molecular dynamics simulation. 10ns MD simulations of both free-protein and complex structure were performed in the NPT ensemble at 310K using Gromacs 4.5.3 package (http://www.gromacs.org/) with the AMBER03 force field. Principal components analysis (PCA) is a method that takes the MDS trajectory and extracts the dominant modes in the motion of the molecule. Here, the most important motion modes of the protein were extracted by PCA of the Cartesian coordinate covariance matrix, yielding eigenvectors and corresponding eigenvalues. The eigenvectors with the largest associated eigenvalues define the essential subspace in which most of the protein dynamics occurs. Such analysis showed that the first two eigenvectors account for 70.2% and 11.2% (i.e. 81.4%) of all the motion in free Dopamine receptor while 68.0% and 16.0% (i.e. 84.0%) in case of bound Dopamine D2 (Fig. 1). Time-development of the conformational changes associated with first and second eigenvectors was compared between Dopamine D2 bound and unbound form. The projection of trajectories on first eigenvector showed similarity with about 2ns delay for free Dopamine receptor while on second eigenvector was found opposite to each. It could be estimate that the agonist make the Dopamine movements faster but the main direction of movements was preserve.

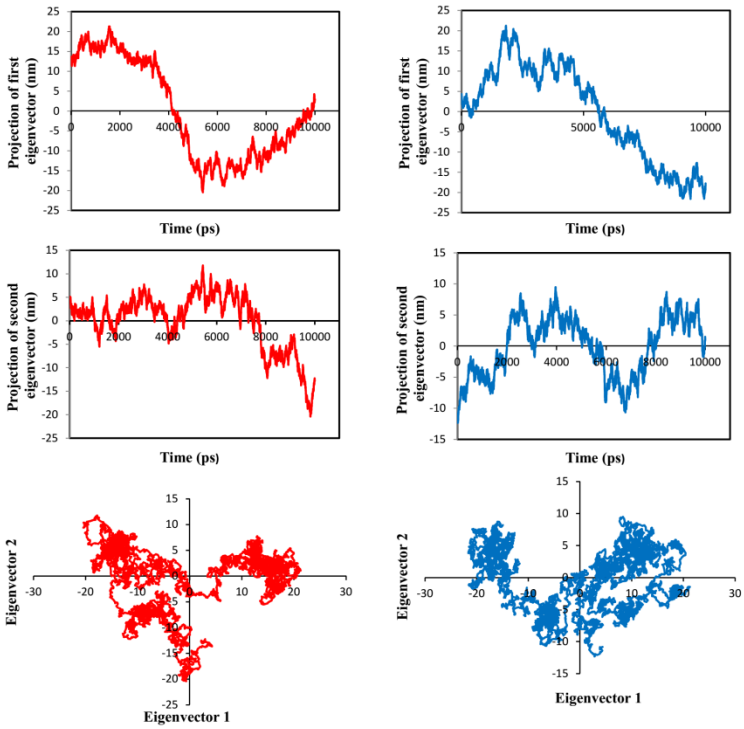


Figure 1. Projection of the first and second eigenvector to a trajectory of the unbound (blue) and bound (red) Dopamine and the correlation of first and second eigenvector

References

[1] Katritzky A. R, Dobchev D. A, Stoyanova-Slavova I. B, Kuanar M, Bespalov M. M, Karelson M, Saarma M. Novel computational models for predicting dopamine interactions, *Experimental Neurology*. 2008; **211**: 150–171.

Simultaneous Spectrophotometric determination of Cysteine and Ascorbic acid using $Cu(OH)_2$/CuO Nanowires and Parallel Factor Analysis

Maryam kuhsarian, Ali Mokhtari, Shabnam hooshmand*, Alireza Goudarzi

Facalty of Science, GolestanUniversity,Gorgan, Iran.
shabnam_hooshmand@yahoo.com
sh.hooshmand@gu.ac.ir

Abstract

We found that cysteine (Cys) and ascorbic acid (AA) could be autoxidized on the surface of $Cu(OH)_2$/CuO NWs. Oxidation of Cyson the NWs produced a new peak at 240-260 nm and UV absorption intensity of this bound increased with time. Moreover, AA could also be autoxidized on the surface of NWs so that its UV absorption peak at 265 nm disappeared slowly in presence of NWs. Based on these findings we proposed a novel spectrophotometric method for the simultaneous determination of Cys and AA using parallel factor analysis (PARAFAC).

Key words: Cysteine, Ascorbic Acid, PARAFAC, Spectrophotometry.

Introduction

Ascorbic acid (AA) and L-cysteine (Cys) are important substances for the metabolism of living cell and are added to pharmaceuticals food, cosmetic etc[1]. There are some reports on the simultaneous determination of AA and Cys with spectrophotometric detection [1-2], chemiluminescence (CL), electrochemistry, capillary electrophoresis and HPLC. A recent challenging research field is the use of nanomaterial for increasing selectivity. In the spectroscopic investigations of our new synthesized copper hydroxide/copper oxide nanowires ($Cu(OH)_2$/CuO NWs) we found that Cys and AA could be autoxidized on the surface of $Cu(OH)_2$/CuO NWs. The UV absorption spectra of a mixture including Cys and AA at different times after mixing with NWs are shown in Fig 1.

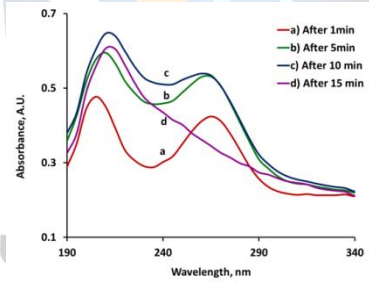


Fig. 1. typical UV absorption spectrum of a mixture including Cys and AA

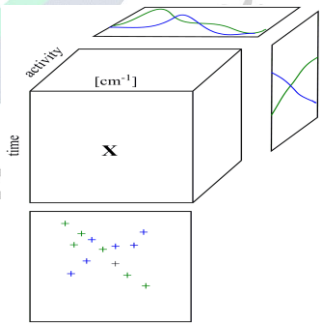


Fig. 2. Schematic illustration of a PARAFAC decomposition

Based on the difference effect of NWs on the UV spectrum of Cys and AA we proposed a novel spectrophotometric method for the simultaneous determination of Cys and AA using parallel factor analysis (PARAFAC). PARAFAC is a mathematical tensor decomposition method suitable for decomposing multilinear datasets into loadings or scores using an alternating least squares algorithm which results in a unique solution.

Results and discussion

After calibration data attainment the data were exported to Matlab (The Mathworks Inc., MA, USA) where they were treated. The three-way array, X, was arrenged which indicate the dimensions of 20 samples × 31 wavelengths × 4 time points (I × J × K). To decompose the resulted tensor PARAFAC analysis was performed using PLS Toolbox (Eigenvector Research Inc., WA, USA). In order to determine the optimal number of factors the value of the core consistency parameter was analyzed. The optimum model with core consistency of 100% suggested that two components were appropriate.

The three dimentional tensor decomposed by PARAFAC (Fig. 2) into three mode loadings which in this particular study refer to sample mode, fingerprint mode and kinetic mode. The PARAFAC loading of the wavelength-direction mode coincide with the pure spectrum of Cys and AA species (Fig. 3). As illustrated in Fig. 4 the PARAFAC kinetic mode loadings shows different tendency for Cys and AA. That could be explained by the fact that their reaction rates with NW were changing totally different. The loadings of the sample mode that referred to as concentration-dependent scores exhibited well relationships with the analytical concentrations of Cys and AA. This shows that these scores can be used as calibration curve for the simultaneous determination of Cys and AA in unknown samples, even in the presence of non-modeled interferences.We will use this unique property for assay of Cys and AA in biological and pharmaceutical samples.

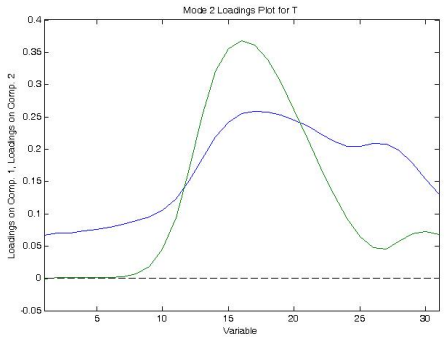


Fig. 3. PARAFAC fingerprint mode loadings for Cys (blue line) and AA (green line)

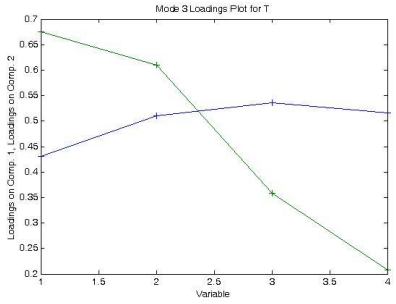


Fig. 4. PARAFAC kinetic mode loadings for Cys (blue line) and AA (green line)

References

1. Ghasemi, J., et al., *Differential kinetic spectrophotometric determinations of ascorbic acid and L-cysteine by partial least squares method.* Analytical letters, 2005. **37**(4): p. 725-737.

2. Ghasemi, J., et al., *Simultaneous Kinetic Spectrophotometric Determination of Ascorbic Acid and L-Cysteine by H-Point Standard Addition Method.* Microchimica Acta, 2004. **148**(3-4): p. 259-265.

5th IBCS
پنجمین سمینار دوسالانه کمومتریکس ایران

The complexicity of molecular docking and need for chemometrics: studying drug resistance in HIV-1 protease as an example

Bahram Hemmateenejad*[1], Nabiollah Mobaraki[1], Amirhossein Sakhteman[2]

[1]*Chemistry Department, Shiraz University, Shiraz, Iran*

[2]*Department of Medicinal Chemistry, School of Pharmacy, Shiraz University of Medical Sciences, Shiraz, Iran*

E-mail address: hemmatb@sums.ac.ir

Abstract

One of the requirements in the clinics to treat HIV is eliminating resistance mutations. One of the existing methods is designing the inhibitors with the capability to fit within the consensus volume of natural substrates, defined as substrate envelope. In this work we try to use different Chemometrics methods to study the differences between the effects of the substrates and the inhibitors on the protease cavity when their bounded to the proteases. Crystallographic structure of a set of 223 substrate/inhibitor bounded protease structures with different mutations and 202 protease inhibitors were selected. Cross docking simulations was performed on this large dataset. The resulted energy scores analyzed by different pattern recognition methods such as Principal Component Analysis (PCA), Hierarchical Clustering Analysis (HCA), Extended Canonical Variate Analysis (ECVA) and Kohonen Self-organizing map (SOM) methods. According to the obtained results it can be deduced that the substrates induced the different structure to the protease. This is in agreement with the substrate envelope hypothesis. By comparing these results with the effects of mutations on the protease selectivity between substrates and ligands, it was found that the substrate envelope plays more important role in this context. Therefore, it can be concluded that by designing the substrate-like inhibitors we expect to see fewer drug resistance.

Key words: HIV-1 Protease, Cross docking simulation, substrate envelope, Chemometrics.

Introduction

In antiviral treatment against AIDS, HIV-1 protease (PR) is an important drug target [1].The performance of the current active-site inhibitors is generally decreased by the growing of drug-resistant mutations. Different techniques have been adapted to design novel antiviral protease inhibitors (PI) therapies against drug-resistant HIV [2]. A new effective strategy is based on the utilizing the substrate envelope constraints in the structure-based drug design. This technique is according to the observation that substrates with large sequence diversity accept a conserved shape after bound to HIV-1 protease [3]. A volume that is used by the bounded substrates defined as substrate envelope.

Many research groups worked on the substrate envelope hypothesis. All of them have emphasized that using of the inhibitors similar to the substrates in terms of the size and structurecaused to occur less drug resistance [4].

Results and discussion

As mentioned in the abstract section, the various methods were applied on the energy scores data matrix. Figure 1A is shown the distribution of the data of each protease in the energy scores data matrix. As can be understood from this figure, histogram of the Substrate bounded proteases have a different pattern related to other groups. It is shown that substrates induced a specific structure to the protease when they bounds to the protease. These findings are in good agreement with the substrate envelope hypothesis. On the other hand, the protease selects the inhibitors on the basis of their ability to fit inside a consensus volume. SOM was applied on the row and column directions of energy scores data matrix. Data matrix reshaped based on the SOM outputs. As shown in the figure 1B, proteases could be categorized in a number of subclasses (fig. 1B).

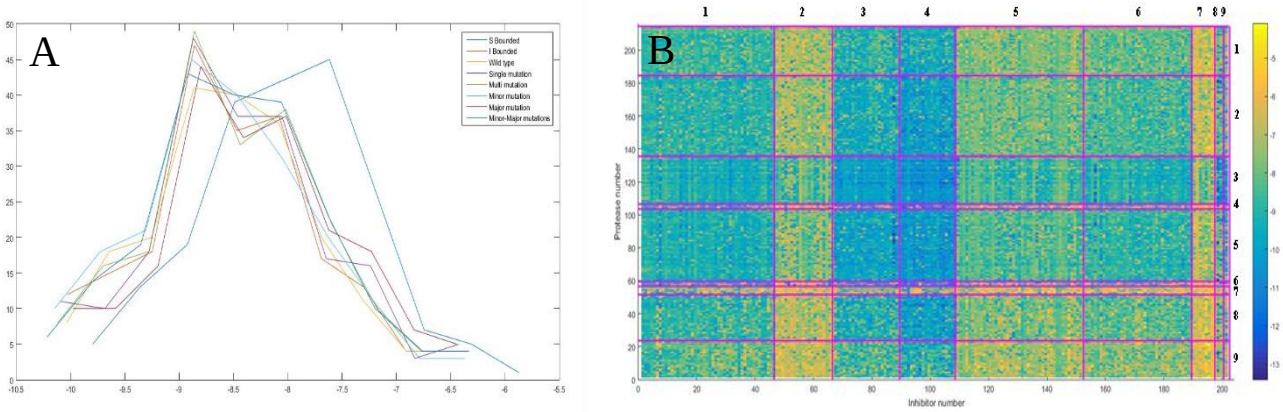


Figure 1. A) Distribution of the data of each protease in the energy scores data matrix. B) Energy scores data matrix reshaped based on the SOM outputs.

The effects of mutations on the classification were investigated and compared to the substrate envelope results. As is indicated, no relationship existed between the mutation types (major or minor) or number of mutations (single or multi) and the energy scores of ligands (substrate or inhibitor) bounded to the proteases. Following the arguments above, It can be concluded that for binding of ligand to the protease the role of ligand type is more important than the mutation types (major or minor) or number of mutations. Thus, this study indicates that for reduce the drug resistant effect, we must use the substrate-like inhibitors.

References

[1] Debouck C. The HIV-1 protease as a therapeutic target for AIDS.*AIDS Res. Hum. Retroviruses*. 1992; **8**: 153–164.

[2]Youle M. Overview of boosted protease inhibitors in treatment-experienced HIV-infected patients. *J. Antimicrob. Chemother*.2007; **60**: 1195–1205.

[3]Prabu-Jeyabalan M, Nalivaika E,Schiffer CA. Substrate shapedetermines specificity of recognition for HIV-1 protease: analysis of crystalstructures of six substrate complexes. *Structure*. 2002; **10**: 369–381.

[4]Prabu-Jeyabalan M, Nalivaika E, Schiffer CA. How doesa symmetric dimer recognize an asymmetric substrate? A substrate complexof HIV-1 protease. *J. Mol. Biol*. 2000; **301**: 1207–1220.

Author Index