Anal. Chem. **1997,** *69,* 2329−2336

Multivariate Curve Resolution and Trilinear Decomposition Methods in the Analysis of Stopped-Flow Kinetic Data for Binary Amino Acid Mixtures

J. Saurina, S. Hernández-Cassou,* and R. Tauler

Department of Analytical Chemistry, University of Barcelona, Diagonal 647, Barcelona 08028, Spain

A stopped-flow method is proposed to carry out the kinetic development of the reaction between amino acids and 1,2-naphthoquinone-4-sulfonate by mixing analytes and reagent in a three-channel continuous-flow system. The process is monitored using a diode array spectrophotometer. Thus, every sample produces a data matrix built up from the spectra registered at regular steps of time. As the reaction is faster for secondary amino acids than for primary ones, it is possible to distinguish between the kinetic formation of their corresponding derivatives. The method is applied to the simultaneous determination of phenylalanine and proline by using second-order multivariate curve resolution. The derivatives of these two amino acids present some differences in both orders of measure, i.e., their spectra and kinetic profiles, which can be exploited advantageously to quantify one of the analytes in the presence of the other as interference, without including any information about this interference in the modeling of the system.

The simultaneous determination of several analytes on the basis of their kinetic properties is usually performed by differential kinetic methods[1,2] based on the difference between reaction rate constants. These methods allow the quantification of mixtures of analytes with very similar or identical spectra which cannot be solved by equilibrium reaction-based methods. When univariate data are handled (i.e., for each sample, a single scalar measure is performed), the experimental and system conditions should be able to provide suitable selectivity. However, the analysis of data vectors[3] containing the evolution of the response over time is a much more powerful approach with lower selectivity requirements.

Curve-fitting methods for multicomponent kinetic analysis are based on the preliminary postulation of a chemical model (hard-modeling methods), for instance, through the order and the stoichiometry of the reaction. The experimental data (a data vector or a data matrix) are then fitted to the proposed model, and the best set of parameters is obtained by using a least-squares optimization.[4−7] The Kalman filter procedure has been also widely used in the analysis of kinetic data, allowing the estimation of a set of optimized parameters that describe the system under the assumption of a certain kinetic model.[8−10] Depending on the model to be fitted, either linear or nonlinear (extended) Kalman filters can be used. The Kalman filter approach has been successfully applied to the simultaneous determination of several analytes.[10] Both curve-fitting and the Kalman filter approach are hard modeling techniques.

A different group of multivariate chemometric methods consists of the so-called soft modeling methods based on factor analysis techniques. These methods do not require the assumptions of a chemical model defined by the kinetic order and the stoichiometry of the reaction. The only assumption postulated is that the experimental data should obey the Beer's law linear aditivity model. The partial least-squares method (PLS) is one of the most widely used factor analysis techniques for quantitative determinations.[11−13] This method utilizes a data vector as a response for each sample, and the standard samples should have the same interferences and matrix effects as those present in the unknown samples.

A different approach recently proposed to carry out multicomponent kinetic determinations is based on artificial neural networks (ANNs).[14,15] The input information, in the form of single values such as rate constants and concentrations[14] or in the form of scores of a principal component model,[15] is provided to the ANN.

A different type of factor analysis techniques has been developed with the main goals of resolution and recovery of the kinetic profiles of the species involved in the different steps of the kinetic process.[16] Data generated in each kinetic run consist of series of spectra taken at preselected times during the evolution of the reaction ordered in a data matrix. The individual analysis

(1) Mottola, H. A. *Kinetic Aspects of Analytical Chemistry*; Wiley: New York, 1988.
(2) Pérez-Bendito, D.; Silva, M. *Kinetic Methods in Analytical Chemistry*; Ellis Horwood: Chichester, 1988.
(3) Wentzell, P. D.; Karayannis, M. I.; Crouch, S. R. *Anal. Chim. Acta* **1989,** *224*, 263−274.
(4) Love, M. D.; Pardue, H. L. *Anal. Chim. Acta* **1994**, *299*, 195−208.
(5) Navas Díaz, A.; González García, J. A. *Anal. Chem.* **1994**, *66*, 988−993.
(6) Estela, J. M.; Cladera, A.; Cerdà, V. *Anal. Chim. Acta* **1995**, *310*, 307−318.
(7) Bugnon, P.; Chottard, J. C.; Jestin, J. L.; Jung, B.; Laurenczy, G.; Maeder, M.; Merbach, A. E.; Zuberbühler, A. D. *Anal. Chim. Acta* **1994**, *298*, 193−201.
(8) Wentzell, P. D.; Vanslyke, S. J. *Anal. Chim. Acta* **1992**, *257*, 173−181.
(9) Gui, M.; Rutan, S. C. *Anal. Chem.* **1994**, *66*, 1513−1519.
(10) Quencer, B. M.; Crouch, S. R. *Anal. Chem.* **1994**, *66*, 458−463.
(11) Blanco, M.; Coello, J.; Iturriaga, H.; Maspoch, S.; Riba, J. *Anal. Chem.* **1994**, *66*, 2905−2911.
(12) Xie, X. L.; Baeza-Baeza, J. J.; Ramis-Ramos, G. *Chemom. Intell. Lab. Syst.* **1995**, *27*, 211−220.
(13) García, J. M.; Jiménez, A. I.; Arias, J. J.; Khalaf, K. D.; Morales-Rubio, A.; de la Guardia, M. *Analyst* **1995**, *120*, 313−317.
(14) Ventura, S.; Silva, M.; Pérez-Bendito, D.; Hervás, C. *Anal. Chem.* **1995**, *67*, 4458−4461.
(15) Blanco, M.; Coello, J.; Iturriaga, H.; Maspoch, S.; Redón, M. *Anal. Chem.* **1995**, *67*, 4477−4483.
(16) Cladera, A.; Gómez, E.; Estela, J. M.; Cerdà, V. *Anal. Chem.* **1993**, *65*, 707−715.

S0003-2700(96)00898-0 CCC: $14.00

of this single data matrix using factor analysis techniques can provide rich information about the number of species present in the system and about their kinetic behavior. These methods can take advantage of the possible differences in the two orders of measurement, the spectral order and the kinetic order. Among the several factor analysis-based methods proposed for resolution of mixtures of species, multivariate curve resolution based on alternating least-squares (ALS) optimization has been shown to be a powerful resolution method which can be easily adapted to different types of systems using different constraints.[16,17] The conditions under which the recovery of the true underlying profiles can be achieved have been examined by different authors.[18,19] Apart from rotational ambiguities, the factor analysis decomposition of a data matrix is associated with intensity ambiguities. There are some ways to break up the intensity ambiguities in the analysis of an individual data matrix. A first case involves using external information such as molar absorptivities of analytes. Another way is possible when the system is closed (i.e., the sum of concentrations of all species containing a certain compound, taken into account their stoichiometries, is constant in the process) and rotational ambiguities have been solved.[19] Finally, in the simultaneous analysis of several matrices (described below), the intensity ambiguities can be solved in relative terms.

Three-way data analysis methods[20] deal with the simultaneous analysis of two or more correlated data matrices (third-order data tensor). Examples of such methods are the generalized rank annihilation method (GRAM)[20,21] and its extension, the trilinear decomposition method (TLD),[22−24] the restricted Tucker model three-way analysis-based methods,[25,26] and the extension of the ALS multivariate curve resolution method to three-way data.[27,28] Under special conditions (trilinear data[20]), these methods can recover the true underlying profiles and allow the quantitative determination of analytes in the presence of unknown interferences by using pure analyte standards (with the only additional input that of the response matrix of the desired analyte or of a mixture containing it at a known concentration). This is a consequence of both the information provided in the two orders of the data matrix and the trilinear structure.[29] Trilinear structure is described by equal response profiles of the analyte in the two orders in the different data matrices simultaneously analyzed. When the system lacks a complete trilinear structure, some ambiguities in the recovery of qualitative and quantitative information can arise. However, even in these systems, an appropriate use of a set of constraints can partially compensate for the lack of complete trilinear structure.[28]

In the present work, the kinetic development of the reaction between amino acids and 1,2-naphthoquinone-4-sulfonate (NQS) is carried out by a stopped-flow methodology. The reaction was developed in a basic medium to give the corresponding derivatives. Each kinetic run was monitored spectrophotometrically by using a diode array detector, which produced a set of spectra registered at different times during the process. Previous studies showed slight differences in the shape of the UV−visible spectra of amino acid derivatives.[30] Moreover, some differences were also observed in the reaction rates of the amino acids and NQS. In particular, the reaction is faster in the case of secondary amino acids than for primary ones, allowing their kinetic discrimination. Phenylalanine (Phe) and proline (Pro) are chosen as models for primary and secondary amino acids, respectively. Phe and Pro derivatives present different behavior in the two orders of measurement (i.e., the spectral order and the kinetic order).

In particular, the goals of the study are (a) the investigation of the multivariate data structure provided by the kinetics of the reaction by rank analysis, (b) the resolution of the derivative species formed during the kinetic reaction by multivariate curve resolution (i.e., the estimation of their kinetic profiles and pure spectra), and (c) the quantitative determination of amino acids in unknown samples in the presence of interferences. Alternately, points b and c were also studied by applying TLD[23] in order to compare the ability of both methods in the resolution and quantification of such kinetic data.

EXPERIMENTAL SECTION

Reagents. All chemicals were of analytical grade. 1,2-Naphthoquinone-4-sulfonate (NQS) was supplied by Aldrich. The reagent solution contained 10^{-3} M NQS in 0.1 M HCl (Merck). The buffer solution contained 0.10 M sodium carbonate (Merck) plus 0.02 M plus sodium hydroxide (Merck). Phenylalanine and proline were supplied by Merck.

Apparatus. Detection was performed using a Hewlett-Packard HP8452A diode array spectrophotometer with a Hellma flow cell of 10 mm path length and 18 μL volume. Spectra were recorded from 290 to 590 nm in steps of 2 nm. Data acquisition and generation of ASCII files were done with a Hewlett-Packard Vectra N2 4/50 computer using the standard HP kinetic software.

Continuous-Flow Manifold. The reaction between amino acids and NQS was developed in a three-channel flow manifold (Figure 1). Solutions were pumped by means of a peristaltic pump (Scharlau HP4) using standard Tygon tubing. All connections, T-pieces, and coils were made of Teflon. NQS and buffer solutions were mixed in a TPFE coil of 2 m × 0.5 mm i.d., giving the reagent solution at the proper pH (pH = 9.7). Next, this reagent solution was mixed with the analyte solution in a coil of 10 m × 0.5 mm i.d. The flow rate of each channel was 0.5 m/min.

Kinetic Development. The kinetic development of the reaction was performed in the flow system using a stopped-flow methodology. The flow was stopped after achieving the steady state, starting at this point to register spectra at regular intervals of 30 s for 25.5 min. In this way, four kinetic data sets were obtained by pumping the following samples through the amino acid channel: (A) water; (B) Phe solution, 4×10^{-4} M; (C) Pro solution, 4×10^{-4} M; and (D) a mixture solution of Phe, 2×10^{-4}

(17) Tauler, R.; Izquierdo-Ridorsa, A.; Casassas, E. *Chemom. Intell. Lab. Syst.* **1993**, *18*, 293−300.
(18) Tauler, R.; Smilde, A.; Kowalski, B. R. *J. Chemom.* **1995**, *9*, 31−58.
(19) Manne, R. *Chemom. Intell. Lab. Syst.* **1995**, *27*, 89−94.
(20) Sanchez, E.; Kowalski, B. R. *J. Chemom.* **1988**, *2*, 265.
(21) Henshaw, H. M.; Burgess, L. W.; Booksh, K. S.; Kowalski, B. R. *Anal. Chem.* **1994,** *66*, 3328−3336.
(22) Sanchez, E.; Kowalski, B. R. *J. Chemom.* **1990**, *4*, 29−45.
(23) Booksh, K. S.; Lin, Z.; Wang, Z.; Kowalski, B. R. *Anal. Chem.* **1994**, *66*, 2561−2569.
(24) Gui, M.; Rutan, S. C.; Agbodjan, A. *Anal. Chem.* **1995**, *67*, 3293−3299.
(25) Smilde, A. K.; Wang, Y.; Kowalski, B. R. *J. Chemom.* **1994**, *8*, 21−36.
(26) Smilde, A. K.; Tauler, R.; Henshaw, J. M.; Burgess, L. W.; Kowalski, B. R. *Anal. Chem.* **1994**, *66*, 3345−3351.
(27) Tauler, R.; Smilde, A. K.; Henshaw, H. M.; Burgess, L. W.; Kowalski, B. R. *Anal. Chem.* **1994**, *66*, 3337−3344.
(28) Saurina, J.; Hernández-Cassou, S.; Tauler, R. *Anal. Chem.* **1995**, *67*, 3722−3726.
(29) Booksh, K. S.; Kowalski, B. R. *Anal. Chem.* **1994**, *66*, 782A−791A.
(30) Saurina, J.; Hernández-Cassou, S. *Analyst* **1995**, *120*, 305−312.

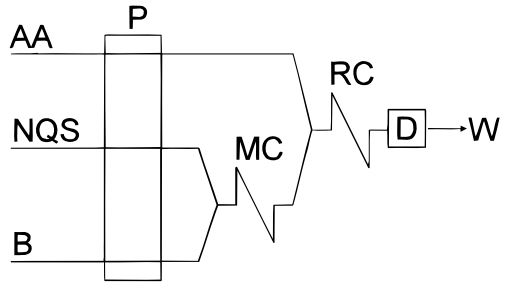

Figure 1. Continuous-flow manifold. P, peristaltic pump; D, diode array spectrophotometer; MC, mixing coil (2 m × 0.5 mm i.d.); RC, reaction coil (10 m × 0.5 mm i.d.); AA, amino acid sample solution; B, buffer solution; NQS, reagent solution; W, waste. Flow rate of channels AA, B, and NQS is 0.5 mL/min.

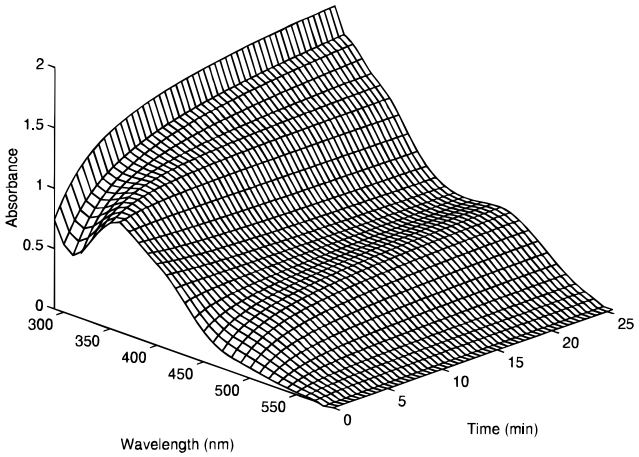

Figure 2. Tridimensional kinetic data obtained for mixture solution D of 2×10^{-4} M Pro and 2×10^{-4} M Phe by using the stopped-flow method.

M, and Pro, 2×10^{-2} M (Figure 2). Alternatively, a 10^{-4} M Phe solution and a 10^{-4} M Pro solution were used as standards.

The continuous-flow manifold allowed the control of the kinetic development under the same experimental conditions in all runs. The importance of this fact for the data analysis is that it assures a time synchronization, which can lead to equal shapes for the time profiles under appropriate experimental conditions. The potential error in the analysis due to the inertia in the flow after stopping the pump can be avoided, in the present case, because the reaction kinetics are slow and the kinetic process is developed when the completely continuous-flow system has achieved a steady state. At the same time, a long reaction coil (10 m) is used to improve the mixing.

Spectral data obtained in every run were subjected to data reduction, taking the absorbance in steps of 10 nm. In all, 31 working wavelengths were finally used between 290 and 590 nm. The reduction of the amount of data to be analyzed allowed us to decrease the computation time without losing significant information. Data treatment was carried out using Matlab for Windows, Version 4.2.

DATA ANALYSIS

Data Arrangement. Each of the four kinetic data sets A, B, C, and D previously described were ordered in data matrices [A], [B], [C], and [D], all of them of dimensions 52 × 31, where the number of rows (52) is equal to the number of spectra registered at the preselected regular time intervals and the number of columns (31) is equal to the number of wavelengths taken for each spectrum.

The four individual data matrices are arranged in the following augmented data matrices:

(1) Augmented columnwise data matrix [D;C;B;A] of dimensions 208 × 31, which is obtained by setting each of the data matrices on top of the other and keeping in common their column order (spectral wavelengths). This situation corresponds to the analysis of Phe and Pro in sample D, which is considered as unknown, while Phe, Pro, and the NQS reagent in the [B], [C], and [A] matrices are used as known standards. The [A] matrix provides information about the NQS reagent and matrices [B] and [C] about Phe and Pro, respectively.

(2) Augmented columnwise data matrix [D;B;A] of dimensions 156 × 31, which is obtained as in case 1, but with sample C absent. In this case, no standard information is provided for Pro, which acts as an interferent in the determination of Phe in sample D. Matrices [B] and [A] are standards which give information about Phe and the NQS reagent, respectively.

(3) Augmented columnwise data matrix [D;C;A] of dimensions 156 × 31, which is obtained as in case 1, but with sample B absent. Phe, considered as an interference, is present only in the unknown sample D, so there is no information about this species in the standard samples C and A. Matrices [C] and [A] provide standard information with respect to Pro and the NQS reagent.

(4) Additionally, three row-wise augmented data matrices are built up from individual data matrices [A], [B], [C] and [D], setting one beside the other and keeping in common their row order (kinetic time): [D,C,B,A], [D,B,A], and [D,C,A]. Although, the following steps of the method are mostly performed with the columnwise data matrices, row-wise matrices are used to check the data structure and chemical rank.

Chemical Rank. The chemical rank is defined as the number of chemical species distinguishable from noise. It is first estimated by singular value decomposition of the individual data matrices [A], [B], [C], and [D]. It is assumed that chemical species have associated singular values much larger than those associated with other possible contributions such as instrumental drift or experimental error. In this case, simple visual inspection of the magnitude of the singular values will provide an idea of the chemical rank. However, for experimental data, it is very difficult to distinguish unambiguously between singular values associated with significant factors and those associated with the experimental error. From the singular value decomposition, a reduced set of possible numbers of chemical species is deduced, which are then tested during the ALS optimization. The number of species finally chosen will be that which gives a low lack of fit as well as reliable kinetic and spectral profiles from a chemical point of view.

Singular value decomposition is also performed on augmented data matrices. The comparison of the significant singular values of the columnwise augmented matrix [D;C;B;A] (keeping in common wavelengths) and the row-wise augmented matrix [D,C,B,A] (keeping in common times) can be used to check whether data have a trilinear second-order structure. In the case of complete trilinear data structure, the number of significant singular values of both columnwise and row-wise augmented matrices must be the same (necessary but nonsufficient condition). The trilinear structure involves equal shape profiles in both orders, spectral and kinetic, for common species in different data matrices. Intermediate situations where trilinear structure is present for only some of the species are also possible, as has been shown in previous works.[28] In that case, chemical ranks of the

columnwise and row-wise augmented matrices are different. Therefore, the comparison of chemical ranks of individual and augmented data matrices gives a powerful tool to investigate the data structure and rank deficiency problems.[31,32]

Resolution of Kinetic Runs by Multivariate Curve Resolution. The analysis of the individual matrices [A], [B], [C], and [D] or of the augmented matrices [D;C;B;A], [D;B;A], and [D;C;A] by multivariate curve resolution allows the estimation of the kinetic profiles and of the pure spectra of the species present in each kinetic run. The resolution of an individual process is accomplished by an iterative constrained alternating least-squares (ALS) optimization based on the compliance of linear Beer's law:

$$\mathbf{D} = \mathbf{C}\mathbf{S}^{\mathrm{T}} + \mathbf{E} \tag{1}$$

where **D** is the experimental data matrix, **C** is the matrix of kinetic profiles of species, $\mathbf{S}^{\mathrm{T}}$ is the transposed matrix of pure spectra of species, and **E** is the matrix of error. During the optimization, the ALS algorithm updates kinetic and spectral matrices as follows:

$$\mathbf{C} = \mathbf{D}(\mathbf{S}^{\mathrm{T}})^{+} \tag{2}$$

and

$$\mathbf{S}^{\mathrm{T}} = \mathbf{C}^{+}\mathbf{D} \tag{3}$$

where $\mathbf{C}^{+}$ and $(\mathbf{S}^{\mathrm{T}})^{+}$ are the pseudoinverses[33] of the matrices **C** and $\mathbf{S}^{\mathrm{T}}$, respectively. The extended model for simultaneous analysis of several data matrices can be written by using the columnwise augmented data matrices and the augmented concentration matrix keeping species in common (instead of the individual matrices **D** and **C**, respectively). A more detailed description of the complete ALS optimization procedure has been given elsewhere.[17,18,27,28]

Constraints applied during the ALS optimization in order to have physically meaningful solutions from kinetic data analysis are as follows: (i) Kinetic profiles and pure spectra must be nonnegative. (ii) The system is closed with respect to the total NQS analytical concentration, which includes the initial NQS species, the NQS decomposition product, and the Phe and Pro derivatives. (iii) The pure spectrum of each species is the same in all kinetic runs where present. (iv) The kinetic profile of each species has the same shape in all kinetic runs where present.

However, this last condition is not strictly fulfilled for all the species since, depending on the amino acid concentration, the NQS consumption during the reaction modifies the shape of the kinetic profiles of unreacted NQS species (the initial NQS species and its decomposition product) in the different runs. It is important to realize that this implies a noncomplete trilinear data structure. Under these circumstances, the analysis of augmented columnwise data matrices is more convenient than the analysis of the corresponding row-wise matrices, since only the spectral (column) order is kept strictly constant between experiments.

Application of Multivariate Curve Resolution to the Quantification of Analytes in Unknown Mixtures. Once the resolution of the kinetic profiles and pure spectra has been achieved, the quantitative estimations are obtained by comparison of the areas below the kinetic profiles of the analytes in the unknown sample D with those from the standard samples B and/or C, respectively for Phe and Pro, as follows:

$$C_{\mathrm{unknown}} = (A_{\mathrm{unknown}}/A_{\mathrm{standard}})C_{\mathrm{standard}} \tag{4}$$

where C_{unknown} and C_{standard} are the concentrations of the analyte in the unknown and standard samples, respectively; A_{unknown} and A_{standard} are the areas below the kinetic profiles in the unknown and in the standard, respectively.

Resolution of Simultaneous Kinetic Runs by TLD. To compare the ability of ALS with respect to the TLD method in the simultaneous resolution of several kinetic runs, the kinetic systems described above were also analyzed using the TLD method. In the TLD method,[22−24] the set of correlated data matrices simultaneously analyzed is ordered in three-dimensional data cube structure. TLD is based on the QZ[33] algorithm for solving the generalized eigenvalue−eigenvector problem. This method assumes a trilinear model, which can be written as

$$\mathbf{R}_{twr} = \sum \mathbf{C}_{ti}\mathbf{S}_{wi}\mathbf{Q}_{ri} + \mathbf{E}_{twr} \tag{5}$$

where $\mathbf{R}_{twr}$ is the experimental data cube response, *i* refers to each component (chemical species), *t* is the time measures, *w* wavelength, and *r* the kinetic run. **C**, **S**, and **Q** are the matrices containing the kinetic concentration profiles, the spectral profiles, and the quantitative (relative) information, respectively. $\mathbf{E}_{twr}$ is the error cube.

Thus, the method assumes a strict trilinear data structure (i.e. there is an intrinsic kinetic and spectral profile unique for each analyte in every run). When this condition is far from being fulfilled, results are not reliable, and complex (imaginary) solutions can be obtained.

Validation of the Spectral Shapes Recovered. The validation of the spectra of Phe and Pro derivatives recovered by ALS and TLD was carried out by comparison with the pure spectra of Phe and Pro derivatives obtained by using an independent method. In this case, these pure spectra of derivatives in the absence of the NQS were obtained by liquid chromatography after their precolumn derivatization as described elsewhere.[34] Each amino acid derivative was recovered by collecting its corresponding fraction from the aqueous eluate solution emerging from the chromatographic system. These fractions were used to record the pure spectra of the amino acid derivatives.

The estimation of the similarity between the pure (chromatographic) and the recovered spectra of each amino acid derivative was performed by calculation of the correlation coefficient as the cosine of the angle between these two spectral vectors. Correlation values of 1 (or close to 1) indicate very similar shapes for both spectra, while lower correlation values mean that pure and recovered profiles are different.

RESULTS AND DISCUSSION

Rank Analysis of Individual Kinetic Runs. (a) Two chemical species were found in the rank analysis of the matrix [A]

(31) Amrhein, M.; Srinivasan, B.; Bonvin, D.; Schumacher, M. M. *Chemom. Intell. Lab. Syst.* **1996**, *33*, 17−33.

(32) Izquierdo-Ridorsa, A.; Saurina, J.; Hernández-Cassou, S.; Tauler, R. *Chemom. Intell. Lab. Syst.*, in press.

(33) Golub, G. H.; Van Loan, C. F. *Matrix Computations*; The Johns Hopkins University Press: Baltimore, MD, 1989.

(34) Saurina, J.; Hernández-Cassou, S. *J. Chromatogr. A* **1996**, *740*, 21−29.

Table 1. Resolution and Quantification of Phe and Pro Contents in Mixture Sample D by Second-Order Multivariate Curve Resolution of Augmented Data Matrix [D;C;B;A] under Different Applied Constraints

resolution option	constraints applied[a]	added Phe concn ($M \times 10^4$)	recovered Phe concn ($M \times 10^4$)	quantitation error (%)	Phe spectral recovery[b]	added Pro concn ($M \times 10^4$)	recovered Pro concn ($M \times 10^4$)	quantitation error (%)	Pro spectral recovery[b]	fitting error[c] (%)
a	i and iii	2	1.87	6.5	0.980	2	2.07	3.5	0.980	0.35
b	i, ii, and iii	2	1.78	11.0	0.964	2	1.98	1.0	0.955	2.17
c	i, ii, iii, and iv(t)	2	1.76	12.0	0.967	2	1.99	0.5	0.955	3.47
d	i, ii, iii, and iv(p)	2	1.82	9.0	0.963	2	2.04	2.0	0.955	2.20

[a] Constraints applied: i, kinetic and spectral profiles must be positive; ii, closure with respect to the NQS analytical concentration; iii, pure spectrum of each species is the same in every kinetic run; iv(t), kinetic profile of each species has the same shape in every kinetic run; iv(p) kinetic profile of each amino acid derivative has the same shape in every kinetic run. [b] Spectral recovery measured by the correlation (cosine) between the pure (chromatographic) spectrum and that calculated by ALS. [c] % fitting error = $(\sum(A_{ij\text{experimental}} - A_{ij\text{reproduced}})^2)^{1/2} / (\sum A_{ij\text{experimental}}^2)^{1/2}) \times 100$, where $A_{ij\text{experimental}}$ is the experimental absorbance at the wavelength i and time j, and $A_{ij\text{reproduced}}$ is the reproduced absorbance obtained by multivariate curve resolution.

corresponding to the kinetic behavior of the reagent in the absence of amino acid. This result was interpreted on the basis of the NQS decomposition process, where the initial NQS species (species I) suffered a degradation resulting in the NQS decomposition product (species II).

(b) The chemical rank of the matrix [B] was found to be 3, which indicated the presence of an additional species with respect to those observed in the matrix [A]. Apart from the two reagent species I and II, the new species must be a Phe derivative (species III) produced in the reaction between Pro and NQS (the unreacted Phe does not absorb in the wavelength range used in the study).

(c) The chemical rank of the matrix [C] was found to be 3. Similarly to case b, the new species must be due to the formation of a Pro derivative (species IV).

(d) The chemical rank of the matrix [D], obtained for the mixture solution of Phe and Pro, was found to be 4. This result agrees with the formation of the Phe and Pro derivatives, in addition to the two NQS species.

Rank Analysis of the Augmented Data Matrices. The singular value analysis of the augmented data matrix [D;C;B;A], keeping wavelengths in common, shows that the chemical rank of this system is 4, which shows the presence of four independent spectral contributions. Therefore, no chemical rank augmentation is observed with respect to the matrix [D]. Obviously, these four chemical species are the two NQS species and the Phe and Pro derivatives. For the augmented data matrix [D,C,B,A], keeping the time order in common, the chemical rank is close to 4. This fact suggests that the system could have a trilinear data structure. Unfortunately, the experimental error associated with the data can hinder the sensitivity of the singular value decomposition to estimate the chemical rank. For this reason, the verification of trilinearity required further study of the results obtained in the resolution, i.e., looking at the shapes of the kinetic profiles of each species recovered in the different runs. On the basis of previous works,[27] it is expected that the trilinearity condition is not completely fulfilled, since the total concentration of NQS was not in sufficient excess.

Initial Estimations of Pure Spectra. The following spectra were chosen as initial estimations to be used in the ALS optimization procedure: (1) spectrum of initial NQS species (species I), registered at the beginning of the data set A, which corresponds to the NQS species in absence of reaction with amino acids; (2) spectrum of the NQS decomposition product (species II), registered at the end of the kinetic data set A; (3) spectrum of Phe derivative (species III), obtained by subtracting the spectrum at the end of the kinetic set B, where the amount of Phe derivative is expected to be maximum, from the spectrum at the end of the kinetic set A; (4) spectrum of Pro derivative (species IV), obtained by subtracting the spectrum at the end of the kinetic set C, where the concentration of the Pro derivative is expected to be maximum, from the spectrum at the end of the kinetic set A.

Case 1. Resolution of the Augmented Data Matrix [D;C;B;A]. Simultaneous Determination of Pro and Phe. The resolution of the augmented data matrix [D;C;B;A] by ALS was performed under different constraints. Table 1 summarizes the effect of these constraints on the data fitting as well as on the simultaneous determination of Phe and Pro in the unknown sample D. Results shown in Table 1 indicate that the best option to carry out the resolution and the simultaneous quantification is (a), which is the least constrained optimization. This option gives the best recovery of the spectral shapes of Pro and Phe derivatives. This is because, although the experimental data obtained in the kinetic runs are close to fulfilling closure and trilinearity, none of them is strictly achieved. In the case of closure, the system is theoretically closed with respect to the total NQS analytical concentration, but, in practice, the total NQS concentration is slightly different from the theoretical value due to small variations in the experimental conditions (especially flow rates) between runs.

It is expected that unreacted NQS species (the initial NQS species and the NQS decomposition product) have slightly different kinetic profiles depending on the amino acid concentrations, due to the consumption of a certain amount of NQS by the reaction to produce the derivatives. Moreover, the results obtained in the ALS optimization indicate that the shapes of the kinetic profiles of the two derivatives are not completely equal in the different kinetic runs. The total maintenance of the shape of these profiles involves having first- or pseudo-first-order kinetics. However, in the present study, the amount of reagent with respect to the amino acid concentration was not enough in excess to provide pseudo-first-order conditions.

The kinetic profiles and spectra recovered using option a of Table 1 are shown in Figure 3. The pure chromatographic spectra of derivatives are also included (Figure 3b). The shape of the kinetic profiles of Pro and Phe derivatives indicates that the rate of formation of the Pro derivative is faster than the rate of formation of the Phe derivative. These two kinetic profiles are, however, very similar, as is apparent from the value of the correlation between the two profile vectors (cosine value 0.975),

Table 2. Resolution and Quantitation of Phe Contents in Mixture Sample D, in the Presence of Pro as Interference, by Second-Order Multivariate Curve Resolution of Augmented Data Matrix [D;B;A] under Different Applied Constraints[a]

resolution option	constraints applied	added Phe concn ($M \times 10^4$)	recovered Phe concn ($M \times 10^4$)	quantitation error (%)	Phe spectral recover	added Pro (interference) concn ($M \times 10^4$)	correlation interference-Pro	fitting error (%)
a	i and iii	2	1.87	6.5	0.984	2	0.961	0.27
b	i, ii, and iii	2	1.64	18.0	0.966	2	0.958	0.72
c	i, ii, iii, and iv(t)	2	1.84	8.0	0.971	2	0.952	0.60
d	i, ii, iii, and iv(p)	2	1.63	18.5	0.965	2	0.958	0.83

[a] Refer to Table 1 footnotes.

and there is also a lack of selectivity of one with respect to the other. Furthermore, the pure spectra of both amino acid derivatives are highly overlapped, as is apparent from the cosine value between them, which was 0.983, and with a total absence of spectral selectivity (specific wavelengths for a given species). In spite of these difficulties, the ALS method was able to provide a reliable resolution of the system.

Case 2. Analysis of the Augmented Data Matrix [D;B;A]. Determination of Phe in the Presence of Pro as Interference. The columnwise augmented data matrix [D;B;A] was analyzed considering the presence of four species, which agrees with the results deduced from rank analysis of this matrix. The same initial estimations of these species as in case 1 were used. However, there was no information in the standards about the fourth component (Pro derivative), which was present only in the unknown sample D. Thus, its initial estimation was arbitrary. The resolution of the augmented data matrix [D;B;A] was performed by ALS under constraints i, ii, iii, and iv, combined in different ways (see Table 2). The concentration of Phe in the unknown sample D was calculated from the areas below the kinetic profile of Phe derivative in matrix [B], which was considered as standard (eq 4).

Results shown in Table 2 indicated that the best option to carry out the resolution was (a), as occurred in case 1. The ALS optimization was powerful enough to allow a reasonably accurate resolution and quantitative determination of the analyte in the presence of interferences (Pro) in the unknown sample D, even in the absence of a complete trilinear data structure. Figure 4 shows the kinetic profiles of all kinetic runs and the pure spectra of the four species obtained by ALS under option a. The concentration of Phe calculated by ALS in the unknown sample was 1.87×10^{-4} M (6.5% error), and the ALS lack of fit between experimental and reproduced data was only 0.27%.

Figure 4 showed that, apart from the information recovered for NQS species and Phe derivative, the method was able to provide an estimation of the kinetic and spectral profiles associated with the interference, which in this case is the Pro derivative. The comparison of the spectrum recovered for the interferent species with the pure chromatographic Pro derivative spectrum (see Figure 4b) indicated a rather good correlation between their shapes (cosine value 0.984). The difference in intensity of both spectra comes from the absence of any standard information about the interference.

Case 3. Analysis of the Augmented Data Matrix [D;C;A]. Determination of Pro in the Presence of Phe as Interference. The matrix [D;C;A] was also analyzed considering the presence of four species. The kinetic and spectral profiles of the species were recovered by the ALS optimization using different combinations of constraints i, ii, iii, and iv, as shown in Table 3. In each case, the concentration of Pro in the unknown sample D was calculated using sample C as reference (eq 4). Taking into account the fit between experimental and reproduced data, the best resolution was attained under option a. The concentration of Pro was 2.07×10^{-4} M (3.5% error), and the ALS lack of fit error was only 0.35%.

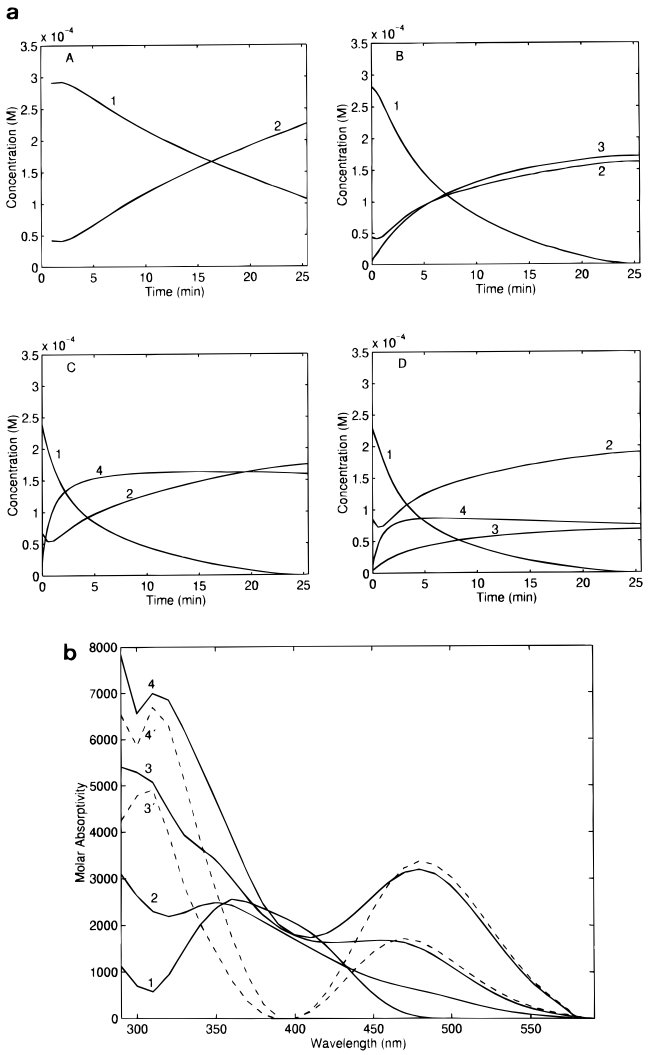

Figure 3. Kinetic profiles (a) and pure spectra (b) recovered by multivariate curve resolution of the simultaneous analysis of the augmented data matrix [D;C;B;A]. Species assignment: 1, initial NQS species; 2, NQS decomposition product; 3, Phe derivative; 4, Pro derivative; 3′, pure (chromatographic) Phe derivative; 4′, pure (chromatographic) Pro derivative.

Analysis of Cases 1, 2, and 3 by Using the TLD Method. The same kinetic systems studied by ALS were analyzed by the

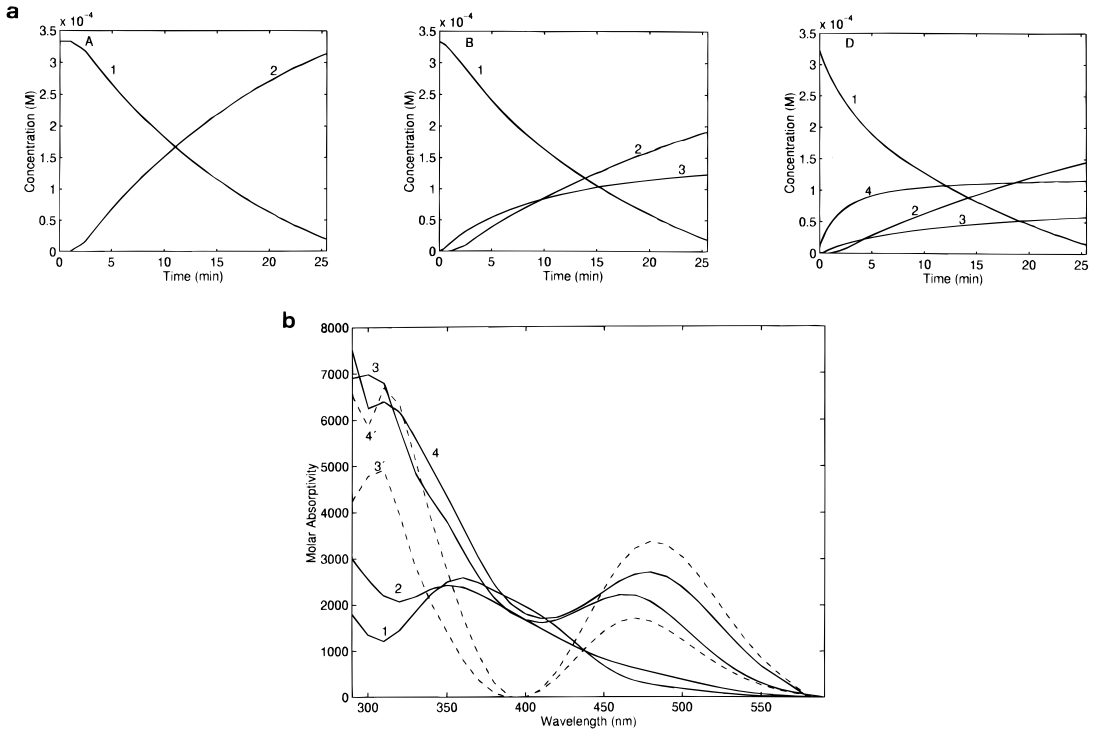

Figure 4. Kinetic profiles (a) and pure spectra (b) recovered by multivariate curve resolution of the simultaneous analysis of the augmented data matrix [D;B;A] under the four species model. Species assignment: see Figure 3.

Table 3. Resolution and Quantitation of Pro Contents in Mixture Sample D, in the Presence of Phe as Interference, by Second-Order Multivariate Curve Resolution of Augmented Data Matrix [D;C;A] under Different Applied Constraints[a]

resolution option	constraints applied	added Pro concn ($M \times 10^4$)	recovered Pro concn ($M \times 10^4$)	quantitation error (%)	Pro spectral recover	added Phe (interference) concn ($M \times 10^4$)	correlation interference-Phe	fitting error (%)
a	i and iii	2	2.07	3.5	0.951	2	0.806	0.35
b	i, ii, and iii	2	1.97	1.5	0.989	2	0.969	0.99
c	i, ii, iii, and iv(t)	2	2.01	0.5	0.960	2	0.971	0.72
d	i, ii, iii, and iv(p)	2	2.08	4.0	0.960	2	0.964	1.00

[a] Refer to Table 1 footnotes.

Table 4. Determination of Phe and Pro Concentrations in Sample Mixture D from the resolution of the Multiple Kinetic Systems [D;C;B;A], [D;B;A], and [D;C;A] by TLD[a]

kinetic system	added Phe concn ($M \times 10^4$)	recovered Phe concn ($M \times 10^4$)	quantitation error (%)	Phe spectral recovery	added Pro concn ($M \times 10^4$)	recovered Pro concn ($M \times 10^4$)	quantitation error (%)	Pro spectral recovery	fitting error (%)
[D;C;B;A]	2	2.88	44.0	0.793	2	2.11	5.5	0.928	3.05
[D;B;A]	2	2.35	17.5	0.899	2	interference		(0.700)	3.67
[D;C;A]	2	interference		(0.287)	2	3.05	52.5	0.966	2.05

[a] Refer to Table 1 footnotes *b* and *c*.

trilinear decomposition method (TLD) in order to compare the capacity of the ALS method in the resolution and quantification of such systems with respect to TLD. Therefore, the simultaneous determination of Phe and Pro (system [D;C;B;A]), as well as the determination of each one of them in the presence of the other as interference (systems [D;B;A] and [D;C;A]), was considered.

Although the system has four chemical species, TLD only provided reasonable kinetic and spectral profiles by using three components. Quantitative results were also better for three components. As the system is not rank-deficient, this discrepancy was attributed to departures from the trilinearity. Table 4 summarizes the results obtained, where the accuracy in the quantification as well as in the spectral recovery of species is worse than those obtained by ALS. Figure 5 shows the kinetic profiles and the spectra of each species in the resolution of the system [D;C;B;A] by TLD. From this figure, it can be deduced that the species unmodeled by TLD is the NQS decomposition product, while the three components included in the model correspond to the initial NQS species and the two amino acid derivatives. The contribution attributed to the NQS decomposition product cannot be resolved; thus, it is mixed with that of the derivatives. The lack of a strict trilinear structure is responsible for the low

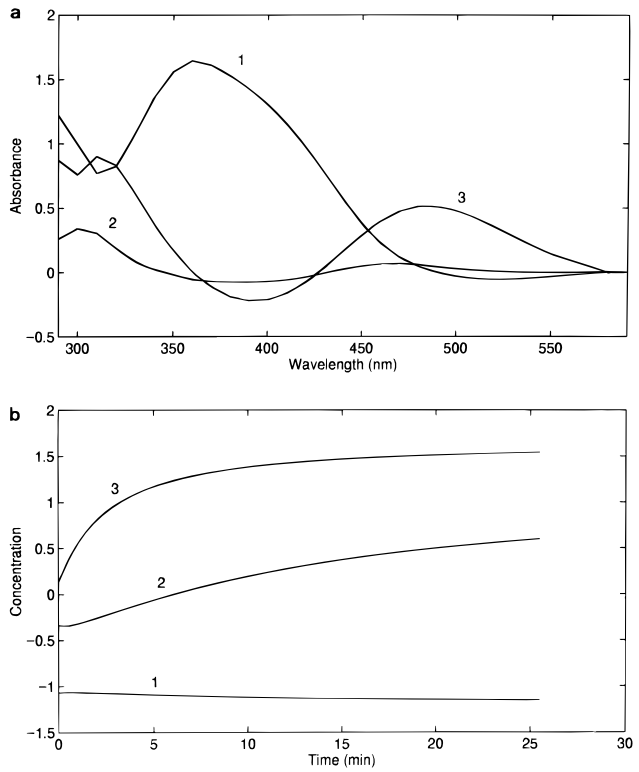

Figure 5. Kinetic profiles (a) and pure spectra (b) recovered by TLD in the simultaneous analysis of the augmented data matrix [D;C;B;A]. Species assignment: see Figure 3.

accuracy of the recovered profiles and quantitative estimations.

The TLD method was applied to the study of two data matrices (one standard and one unknown, i.e., equivalent to GRAM method). In this approach, the data matrix [A] containing the kinetic behavior of NQS was removed. Then, systems under study were [D;B] for the determination of Phe in the presence of Pro as interferent, in which the concentration of Phe found was 1.49 $\times 10^{-4}$ M, and [D;C] for the determination of Pro in the presence of Phe as interferent, in which the concentration of Pro found was 2.80 $\times 10^{-4}$ M. Quantification errors are of the same order as those obtained for [D;B;A] and [D;C;A] systems using TLD.

Other Cases. The analysis of the sample mixture D was also performed by using a 1 $\times 10^{-4}$ M Phe solution and a 1 $\times 10^{-4}$ M Pro solution as standards B and C, respectively, instead of the standards solutions of cases 1, 2, and 3. The resolution of [D;C;B;A] by ALS gave a concentration of Phe of 1.81 $\times 10^{-4}$ M (9.5% error) and a concentration of Pro of 1.84 $\times 10^{-4}$ M (8.0% of error). In the determination of Phe in the presence of an interferent ([D;B;A] system), the concentration of Phe found was 1.61 $\times 10^{-4}$ M (19.3% error), whereas in the resolution of the [D;C;A] system the concentration of Pro recovered was 1.88 $\times 10^{-4}$ M (6.4% error). In these cases, in spite of quantification errors slightly increased with respect to the cases 1, 2, and 3 listed above, the ALS method can be applied to these systems with still rather accurate results. On the other hand, the application of TLD to these new cases provided worse results, with quantification errors higher than 50%.

CONCLUSIONS

The ALS multivariate curve resolution of multiwavelength kinetic spectrophotometric data proved to be a powerful tool for mixture analysis of little selective systems. The application of the ALS method to the study of a kinetic reaction between NQS and two amino acids (Phe and Pro) allowed the resolution of all spectroscopically active species present. Spectral profiles recovered by ALS were compared with those obtained by an independent chromatographic method, and a good concordance between them was observed. Moreover, the quantification of Phe and Pro in the unknown mixture was successfully performed under different approaches, including the determination of one of them in the presence of the other as interferent. The proposed ALS method proved to be especially suited for the resolution of chemical systems where, as in the present work, the trilinear condition was not completely fulfilled. In this case, ALS represented an advantageous alternative to those methods based on the assumption of a strict trilinear structure of data such as TLD. The method can be extended to the kinetic determination of other amino acid mixtures whenever the amino acids have some differences in both kinetic and spectral orders.

ACKNOWLEDGMENT

This work has been supported by the CYCIT Project PB93-0744 and by the CIRIT Project 1995 SGR-00458.

Received for review September 9, 1996. Accepted March 20, 1997.[⊗]

AC9608984

[⊗] Abstract published in *Advance ACS Abstracts,* May 1, 1997.