Resolution of temperature dependent conformational multiequilibria processes

Roma Tauler [*], Raimundo Gargallo, Montse Vives, Anna Izquierdo-Ridorsa

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

Abstract

Multivariate Curve Resolution (MCR) is applied to the study of temperature dependent conformational multiequilibria evolving processes. Experimental data sets are obtained by UV spectral monitoring of the melting behavior of the heteropolynucleotide poly(adenylic) acid–poly(uridylic) acid (poly(A)–poly(U)) and of the hompolynucleotides poly(adenylic) (poly(A)) acid and poly(urydilic) acid (poly(U)), i.e., recording UV spectra at different temperatures during the melting process of these polynucleotides. Separate study of every individual melting experiment by MCR did not give a satisfactory resolution of the heteropolynucleotide poly(A)–poly(U) melting process because of unresolved rotational ambiguities and rank deficiency problems. Conversely, the simultaneous MCR analysis of the melting process of poly(A)–poly(U) heteropolynucleotide together with the separate melting processes of the poly(A) and poly(U) homopolynucleotides, allowed the resolution of the species profiles and the elimination of the rank deficiency problems present in the individual analysis of the melting behavior of poly(A)–poly(U). © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Multivariate curve resolution (MCR); Temperature; Polynucleotides

Contents

[*] Corresponding author. Tel.: +34-93-4021545; Fax: +34-93-4021233; E-mail: roma@quimio.qui.ub.es

1. Introduction

In a broad sense, chemical processes can be defined as chemical systems which may evolve with the change of a system variable. With this definition in mind, the study of evolutionary chemical processes is of interest in many different chemical fields, like Chemical Engineering, Analytical Chemistry or Biochemistry. Application of chemometric techniques to chemical processes has been denominated Process Chemometrics [1]. Spectrometric monitoring of evolving chemical processes like those occurred in multiequilibria and kinetic reactions when a system variable like pH or time is changed provide very rich multivariate data describing the chemical changes occurring during the process. The spectral data obtained in one run of the process is organized in an ordered data matrix, with as many rows as spectra acquired along the process, and with as many columns as wavelengths or spectral channels measured. Each one of these data matrices contains information about two vector spaces, defined by the rows and the columns of the matrices. Row vector space gives the description of the evolution of the process, about how the concentrations of the different chemical components change during the process. Column vector space describes the spectral responses of these components.

Hard-modeling methods, i.e., methods based on the fulfillment of a chemical model like mass-action or rate laws, can only be applied when the system under study follows the proposed model, at least to the extent that unexplained variance be small. There are many cases where physical models cannot be applied in practice. One obvious case is when no physical model is known explaining the observed data variance. A second case is when the model variables cannot be fixed at the conditions where the model is fulfilled. A third case would be when only a small proportion of observed data variance can be described correctly by the physical model, leaving a major part of it, unexplained. In all these circumstances, soft-modeling methods provide a complementary and very useful tool of data analysis. Soft modeling methods do not use and do not postulate any a priori physical model to describe data variation; instead they rely on relatively some 'soft' assumptions about data, like linearity, non-negativity and others. Factor Analysis based methods, FA [2], are among the soft modeling methods which are more useful to model and analyze chemical data. Specially suitable for the study of evolutionary processes are methods like Evolving Factor Analysis, EFA [3]. Application of EFA to evolutionary processes, provides information about the number of evolving components, their appearance and disappearance, and about the local rank of the system [4]. One steep further in multivariate process modeling is Multivariate Curve Resolution, MCR, methods [5–8]. MCR methods try to obtain information about the identity of the components (spectra profiles) of the system, and about their concentration changes (species concentration profiles). There are different variants of Multivariate Curve Resolution algorithms which can be divided roughly as iterative and non-iterative curve resolution methods. Non-iterative methods try to recover the concentration and/or spectra profiles of the components of the system in one single step, using only the information about the regions of existence of each of them (local rank), whereas iterative methods try to recover the profiles in an iterative way, using an optimization algorithm as alternating least squares [7,8] or iterative target factor analysis [9]. Iterative methods can easily include several constraints during the optimization process, whereas in non-iterative curve resolution methods such an inclusion becomes rather difficult. Under selectivity or local rank favorable conditions, mathematically formulated as resolution theorems [10], the recovery of the correct profiles is theoretically possible. If such resolution conditions

are not fulfilled, the recovery of the species profiles is more problematic and ambiguous and the extent of this ambiguity depends on multiple factors, like data complexity (number of components overlapped) and data constraints. Although no general rules can be formulated to recover the correct solutions in the absence of local rank conditions, the results obtained via an alternating least squares iterative method [7,8] using natural constraints like non-negativity, unimodality, closure and others do fulfill most of the requirements for a semi-quantitative resolution of an unknown system allowing its chemical interpretation. In the case of monitoring an evolving process, such information would allow possible decisions about the process to be taken in future runs [11].

Additionally, MCR is easily extended to the simultaneous analysis of different correlated runs of a chemical process [7,8,11,12]. Combining data from different runs may contribute to improve data resolution because of different reasons. First, because the simultaneous analysis of multiple runs of a chemical process improves the stability and reliability of the obtained solutions obtained; since the system of equations to be solved is much more overdetermined than the system of equations to be solved in case of the analysis of a single run. Second, with an appropriate experimental design of data, resolution conditions not achieved in the analysis of a single run of the process, can be achieved instead, in the simultaneous analysis of different runs of the process. And third, the simultaneous analysis of several correlated data sets arranged in different data matrices, gives a three-way data structure, which under trilinearity conditions [13], can be resolved without the rotational ambiguities often present in resolution of two-way (one data matrix) data sets. Trilinearity, however, is a condition difficult to be fulfilled for experimental data in the general case, and three-way non-trilinear data is a much more common situation in practice [6,14].

An additional difficulty commonly present in process analysis and in chemical reaction based systems is rank deficiency [15]. In this case, the number of linearly independent contributions to data variance is less than the number of chemical components; this situation is often present in reaction based systems and in closed systems [15–17] when the number of independent reactions is lower than the number of

chemical species involved in these reactions. Rank deficient systems are obviously more difficult to resolve than full rank systems. However, simultaneous analysis of different runs of a process under independent conditions, may contribute to rank augmentation and to break rank deficiencies, allowing the resolution of components which could not be resolved in the individual analysis of each data set separately.

All these ideas have been used recently in the species resolution of biochemical processes related with the study of the solution properties of polynucleotides [18–20] and proteins [21,22] when the pH, the temperature or other variable of the system is changed. Modeling these pH or temperature dependent processes involving biomolecules is not usually feasible at present by means of hard modeling methods based on physical laws. Soft modeling description of biomacromolecular processes like acid–base macromolecular reactions, metal complexation reactions by macromolecules or any other macromolecular interaction may be achieved easily by means of MCR methods. Evolving processes like conformational changes or protein folding caused by different physical or chemical agents like temperature, pH, ionic strength or concentration of particular substances can be easily studied using MCR. MCR is relatively simple to implement compared with hard modeling methods, where all the chemical interactions and sources of data variance should be precisely defined a priori (model postulation) by means of a physical model. Soft modeling MCR is a much more analytical way to model experimental data using only very weak and soft assumptions. From the information gathered from MCR, the explanation of what occurs in the process in terms of analytical properties (concentration changes and spectra of the contributions) allows the understanding of the system from a chemical point of view, i.e., a chemical model is deduced a posteriori.

In this paper, UV spectrometric process monitoring of a melting (temperature dependent change of conformation) experiment of the heteropolynucleotide poly(adenylic)–poly(uridylic) acid (poly(A)–poly(U)) is studied and compared with the melting behavior of the homopolynucleotide constituents, polyadenylic acid, poly(A), and polyuridylic acid (poly(U)). Results obtained will be compared with those previously obtained in the monitoring of poly-

inosinic–polycytidylic acid (poly(I)–poly(C)) [22,23]. In both cases, rank deficiency problems are present. Solution to this problem by means of matrix augmentation is studied in detail. The ultimate goal of the study is the modeling and description of the process under study and the resolution of the chemical species responsible of the observed data variance of the system, i.e., the extraction of the number of species, the calculation of their melting concentration profiles and of their pure spectra.

2. Experimental

2.1. Reagents and solutions

Sodium chloride (Merck), potassium monohydrogenphosphate (Carlo Erba), sodium dihydrogenphosphate (Panreac) and poly(uridylic) acid, poly(adenylic) acid, poly(adenylic)–poly(uridylic) acid (SIGMA) were used without further purification. Solutions of polynucleotides were prepared from a known amount of the solid reagent and dissolution in the ionic medium used in this study. A pH value of 6.9 was adjusted by means of the buffer solution monohydrogenphosphate/dihydrogenphosphate. The concentration of the synthetic polynucleotide solutions was referred to the concentration of the cyclic monophosphate nucleotides cAMP and cUMP, which are the monomeric units in the polynucleotide chains.

2.2. Apparatus

UV absorption spectra were recorded on a Perkin-Elmer lambda-19 spectrophotometer. Temperature was adjusted with a Perkin-Elmer Digital Controller, based on a Peltier device, model C570-0701 with a rate of 1°C/min. Instrumental control,

data acquisition and spectra pre-processing were carried out via personal computers. pH measurements were performed with an Orion model 701A pHmeter (with a precision of ± 0.1 mV) and a combined Ross pH electrode (Orion 81–02).

2.3. Procedure

The melting experiments were performed by recording the UV spectrum (230–310 nm, 81 absorbance values read) at varying temperatures in the range 15–90°C. A quartz cell (Hellma, 1 cm optical path) with a Teflon® stopper was used in these experiments. The experimental conditions of the melting experiments performed are given in Table 1. Four melting experiments were analyzed at the conditions given in Table 1. Ionic strength of the experiments was 0.26 M and pH was set at 6.9 (37°C) According to previous studies of the acid–base properties of the polynucleotides (see Refs. [21,24,25]) at this pH a single acid–base species is present for these three polynucleotides.

However, a single acid–base species can adopt different molecular conformations. In biochemical studies, the conformation existing at physiological conditions is often defined as the native conformation. In the melting experiments, we will assume that only one conformation is present at the beginning of the experiment, before temperature is raised and conformational changes induced. In the case of pure homopolynucleotides, poly(A) and poly(U), the conformations at high temperatures are also considered unique (selective). On the contrary for the heteropolynucleotide, poly(A)–poly(U), this last constraint is not considered, since a mixture of individual separated homopolynucleotides poly(A) and poly(U) can evolve when the melting denaturation of the heteropolynucleotide poly(A)–poly(U) occurs

Table 1
Experimental conditions of the melting experiments

Experiment	Polynucleotide	Concentration (M)	pH	Temperature range (°C)	No. of spectra
A	poly(A)	4.92×10^{-5}	6.8	15–90	26
B	poly(U)	4.87×10^{-5}	6.8	15–84	24
C	poly(A)–poly(U)	5.25×10^{-5}	6.8	15–90	32
D	poly(A)–poly(U)	4.73×10^{-5}	6.8	15–90	36

[23]. These initial assumptions about the initial and final conformations are specially relevant for the elimination of the resolution ambiguities (see below) and for the interpretation of results. All the detected conformational changes will be relative to the initial conformations present at the beginning of the experiment. Low temperatures facilitates nucleic base stacking and nucleic base pairing, the two main forces in ordered conformations. In contrast, high temperatures, breaks base stacking and base pairing inducing disordered structures, and in the case of heteropolynucleotides, inducing separation of the constituent hompolynucleotide chains [26].

3. Method

3.1. Data pretreatment

When a single melting experiment is analyzed, no data pretreatment is performed. In melting experiments, the total amount of absorbing material is constant, i.e., no addition nor subtraction of material occurs during the experiment, and the system is considered 'closed'. When the simultaneous analysis of several melting experiments is carried out, the experimental spectra were row-normalized to equal length (equal norm). This normalization is accomplished by dividing each experimental absorbance value by the length or norm of the corresponding spectrum (square root of the sum of all the squared absorbance values of the corresponding spectrum). This normalization eliminates the difficulties to know accurately the true initial concentration of the absorbing bases as constituents of the polynucleotide (see Section 2). Due to the closure constrain, changes in intensity of experimental spectra caused by normalization will not affect the relative amounts of the resolved components and will only affect the relative intensities of the (pure) component spectra. The relative amounts of the resolved components will only depend on the changes of the shape of the experimental spectra. As a conclusion, when closure constrain is applied, the normalization of the experimental spectra will not convolute the relative amounts of the components within an experiment but the relative intensities of the pure component spectra. When different melting experiments are simultaneously analyzed, row-normalization causes the loss of the quantitative information

between experiments, which could have been obtained from the comparison of the spectra intensity variations between different experiments. Since the spectra are equally normalized for all the experiments simultaneously analyzed, the same closure constant is applied to all the measured spectra. However, the information concerning the relative concentrations of the different species within the same melting experiment is not lost. As the main goal of the present study is the resolution of the system (estimation of species spectra and of melting profiles), the lost of quantitative information between experiments is not relevant. Different tests performed with simulated data showed that such an approximation is correct and that it will facilitate the analysis and interpretation of results of the systems under study (see Section 4).

3.2. Data treatment

Data treatment starts with the estimation of the number of independent contributions to data variance. In evolving processes this can be initially accomplished by visual inspection of eigenvalue, singular value or variance explained (or residual) principal component analysis plots. These plots give a first approximation of the major contributions to data variance, commonly referred as 'chemical rank', meaning by this, the number of chemical contributions evolving during the process. Evolving Factor Analysis [3] provides a more accurate way for visual inspection of intrinsically ordered data such as that occurring in the evolution of a chemical process. As process analysis data is logically ordered by time or by another variable which changes with time, like pH in a spectrometric acid–base titration or temperature in a melting experiment, EFA provides a powerful way to distinguish the evolution of the independent contributions or changes in concentration of the different components of the system along the experiment. EFA provides a way to estimate the appearance and disappearance of the different components in the system, i.e., the range or windows of existence of these compounds. EFA was initially proposed in 1985 [27] and its use has been extended during the recent years [3,28–30]. One of the more interesting features that EFA facilitates, is the estimation of how the local rank evolves during the process, which is extremely helpful for the assessment of resolution

conditions for an unambiguous resolution according to the Resolution Theorems [10] (see below). It should be noted, however, that the contributions detected by EFA do not correspond always with the real chemical contributions present in the system. On one hand there is the possibility to underestimate the number of chemical contributions because their concentrations do not change independently or because the spectral responses of two or more components are practically equal. And, on the other hand, there is the possibility to overestimate the number of chemical contributions because other undesired sources of data variance are present in the process, like detector and baseline shifts, background or solvent contributions, etc. Moreover, an additional aspect to be taken into account is the linearity of the spectral response. Non-linearities will be accompanied with upsurging of additional components in EFA plots. In conclusion, EFA plots are a very useful tool to investigate the evolution of a particular system, although they should be used with caution.

Once the global and local rank analysis of an evolving process has been performed, process modeling can be accomplished using Multivariate Curve Resolution [7,8]. The initial assumption is that the data is bilinear, which can be described mathematically as:

$$\mathbf{D} = \mathbf{C}\mathbf{S}^{\mathrm{T}} + \mathbf{N} \qquad (1)$$

where $\mathbf{D}(ns,nw)$ is the experimental data matrix with as many rows, ns, as spectra collected in the evolution of the process run and with as many columns, nw, as wavelengths or spectral channels are measured; $\mathbf{C}(ns,n)$ is the matrix describing the evolution of the n chemical contributions detected to be present during the process; $\mathbf{S}^{\mathrm{T}}(n,nw)$ is the matrix of pure spectra of these detected contributions. Matrix $\mathbf{N}(ns,nw)$ is the residual or noise matrix with the unexplained variance using the contributions in $\mathbf{C}$ and $\mathbf{S}^{\mathrm{T}}$. Self modeling curve resolution attempts to recover the $\mathbf{C}$ and $\mathbf{S}^{\mathrm{T}}$ matrices from the analysis of the original data matrix, $\mathbf{D}$. In well defined evolving systems, EFA provides a good initial abstract estimation of the matrix $\mathbf{C}$, from which an alternating least squares (ALS) non-linear optimization is started solving iteratively the two following equations:

$$\mathbf{S}_i^{\mathrm{T}} = \mathbf{C}_{i-1}^{+}\mathbf{D}_{\mathrm{PCA}} \qquad (2)$$

and

$$\mathbf{C}_i = \mathbf{D}_{\mathrm{PCA}}\big(\mathbf{S}_i^{\mathrm{T}}\big)^{+} \qquad (3)$$

where $\mathbf{S}_i^{\mathrm{T}}$ and $\mathbf{C}_i$ are the least squares estimations of these two matrices in the iteration i, and $\mathbf{C}_{i-1}^{+}$ is the best least squares pseudoinverse estimation of the matrix $\mathbf{C}_{i-1}$ obtained in the previous iteration $i-1$. In the first iteration, $\mathbf{C}_{i-1}$ is the matrix initially estimated by EFA [3,27,28]. $(\mathbf{S}_i^{\mathrm{T}})^{+}$ is the best least squares estimation of the pseudoinverse of matrix $\mathbf{S}_i^{\mathrm{T}}$. $\mathbf{D}_{\mathrm{PCA}}$ is the principal component analysis reproduced data matrix $\mathbf{D}$ for the selected number of components. This matrix has been noise filtered and provides more stability to the calculations. Instead of starting with EFA estimations of $\mathbf{C}$, in cases where the evolution of the concentrations follow complex patterns not changing in a stepwise evolving way during the process, it can be better starting the alternating least squares optimization using an initial estimation of the spectra, matrix $\mathbf{S}^{\mathrm{T}}$, specially if the spectra of the components have distinct features. In these cases, methods based on the detection of pure variables, like those proposed in SIMPLISMA [31], provide better and more reliable initial estimations to start the iterative optimization [32].

As a general rule, the unconstrained iterative optimization using Eqs. (2) and (3) will not necessarily converge to the true solution because of the so called rotational and intensity factor analysis ambiguities [5,7,33]. Manne [10] has stated mathematically the conditions under which the components present in a two-way data matrix can be resolved without ambiguities. These conditions have been postulated as resolution theorems in a chromatographic resolution context, but they can be easily extended to any evolutionary chemical process. The conditions under which resolution is possible can be tested using local rank detection techniques, such as those derived from evolving factor analysis techniques. If the local rank conditions are fulfilled, in principle the resolution can be obtained directly without imposing any additional constraint using an appropriate algorithm, such as that proposed in window factor analysis of Malinowski [34]. However, in practice and due to noise, it is better to use an iterative algorithm where local rank conditions are used as a constraints during an alternating least squares optimization. In fact, EFA initial

estimations do already contain the local rank information necessary (when present in the data) for an appropriate resolution, and this is the reason why their use has given good results in practice [3,7,27,28,30].

Apart from local rank conditions, the sought solutions are known to obey some other natural constraints like non-negativity, unimodality and closure. An improved resolution algorithm should incorporate the implementation of such constraints when the data obey them. This is quite simple to implement externally in an alternating least squares algorithm such as that previously described in Eqs. (2) and (3). At each new iterative recalculation of $\mathbf{C}_i$ and $\mathbf{S}_i^T$ these constraints can be applied. Non-negative solutions of Eqs. (2) and (3) can be found by non-negative least squares [35,36]; recently, a similar approach has been proposed for unimodality and non-negativity [37]. It is found however that in practice, the way how these constraints are applied affect mostly the number of iterations in the non-linear optimization, but it affects little the results; this is however a matter of discussion at present and deserves a deeper study. When local rank conditions for resolving rotational ambiguities are not present, the effect of using these natural constraints in the resolution of a two-way data matrix will decrease the number of possible solutions of Eqs. (2) and (3), inside a narrower band of possible solutions [5,6]. The important thing to recognize in such situations is that the found solutions although having still some ambiguity, they fit and explain the experimental data in a plausible way from a physical point of view and fulfill the constraints imposed. It should be recognised that in many contexts having these still somewhat ambiguous solutions is much better than having nothing.

Resolution of individual experiments or process runs can be improved if several correlated experiments or process runs are analyzed simultaneously. Correlated experiments are those which have some structure in common, for instance when some chemical species are common in the different experiments. In this case, several advantages are obtained in the simultaneous analysis. The first advantage is a signal averaging one, i.e., the system of equations to be solved is more overdetermined for the common structure, and therefore, the noise effect is reduced. Second, local rank conditions needed for resolution without ambiguities in a single data matrix are ex-

tended to the simultaneous analysis of several data matrices. If one species has a favorable local rank condition in one matrix, this species can be correctly resolved not only in this data matrix, but also in the others simultaneously analyzed, even if this species was not resolved in this matrix when it was analyzed individually. A third advantage is that in the new three way data structure obtained when several correlated data matrices are simultaneously analyzed, the rotational ambiguity present in two way data sets is drastically reduced, specially if the new data structure is trilinear [7,8,38]. In the proposed MCR-ALS method, the simultaneous analysis of NR different experiments or process runs is performed using an extension of linear Eqs. (1)–(3) to augmented column-wise data matrices:

$$\mathbf{D}_{\text{aug}} = \begin{bmatrix} \mathbf{D}_1 \\ \mathbf{D}_2 \\ \cdots \\ \mathbf{D}_{\text{NR}} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \\ \cdots \\ \mathbf{C}_{\text{NR}} \end{bmatrix} \mathbf{S}^T + \mathbf{N} = \mathbf{C}_{\text{aug}} \mathbf{S}^T + \mathbf{N}$$

(4)

$$\mathbf{S}^T = \begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \\ \cdots \\ \mathbf{C}_{\text{NR}} \end{bmatrix}^+ \begin{bmatrix} \mathbf{D}_1 \\ \mathbf{D}_2 \\ \cdots \\ \mathbf{D}_{\text{NR}} \end{bmatrix} = \mathbf{C}_{\text{aug}}^+ \mathbf{D}_{\text{aug}}$$

(5)

$$\begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \\ \cdots \\ \mathbf{C}_{\text{NR}} \end{bmatrix} = \begin{bmatrix} \mathbf{D}_1 \\ \mathbf{D}_2 \\ \cdots \\ \mathbf{D}_{\text{NR}} \end{bmatrix} (\mathbf{S}^T)^+ = \mathbf{D}_{\text{aug}} (\mathbf{S}^T)^+$$

(6)

where the number of rows in $\mathbf{D}_{\text{aug}}$ and in $\mathbf{C}_{\text{aug}}$ is equal to the total number of spectra collected in the different NR process runs, run by run. In this column-wise matrix arrangement, it is not necessary that all the experiments have the same number of spectra. On the contrary, the method allows that the number of spectra in different experiments be different. The number of columns in column-wise augmented data matrix $\mathbf{D}_{\text{aug}}$ and in spectral matrix $\mathbf{S}^T$ is the same and equal to the number of measured wavelengths, NW. The number of columns of matrix $\mathbf{C}_{\text{aug}}$ and of rows of matrix $\mathbf{S}^T$ is equal to the number of species. The augmented concentration matrix $\mathbf{C}_{\text{aug}}$ has in its columns the concentration profiles of the different detected species in the different runs. The spectral

data matrix $\mathbf{S}^T$ has in its rows the spectra of these detected species.

The iterative solution of Eqs. (5) and (6) by ALS optimization required initial estimates of the matrix $\mathbf{C}_{aug}$ if the first equation solved in the first iteration is Eq. (5), or of the matrix $\mathbf{S}^T$ if the first equation solved in the first iteration is Eq. (6). In both cases initial estimates should be provided for each of the species considered to be present. Usually initial estimates are obtained via the previous individual analysis of the different data matrices included in the augmented matrix. In case a new species is proposed in the simultaneous analysis which was not detected in the individual analysis (as in the case of rank deficiency problems in the analysis of individual matrices, see below), a new estimate of its profile has to be proposed. There are several possibilities to do this and the best approach is problem dependent. Apart from EFA or pure variables [31], an initial vector having ones where the species is supposed to be present and zeros where not, i.e., defining where exists and where not, is good enough as an initial estimation.

During the iterative ALS optimization using Eqs. (5) and (6), the same constraints as in the analysis of a single experiment can be applied to matrices $\mathbf{C}_{aug}$ and $\mathbf{S}^T$, selectivity (in $\mathbf{C}_{aug}$), non-negativity (in $\mathbf{C}_{aug}$ and $\mathbf{S}^T$), unimodality (in $\mathbf{C}_{aug}$), and closure (in $\mathbf{C}_{aug}$). Additionally, new constraints related with the new data structure can be now applied. First, owing to the column-wise data structure of Eq. (3), the species spectra of common species in different runs are equal and correspond to the same row of matrix $\mathbf{S}^T$; second, the concentration of those species known to be absent in a particular experiment are forced to be zero setting the corresponding rows in matrix $\mathbf{C}_{aug}$ equal to zero during the ALS optimization. Correspondence between species in different experiments is implemented in matrix $\mathbf{C}_{aug}$. Different correspondences between species in the different melting experiments will give different solutions and fitting values. That correspondence between species giving profiles with unreasonable shapes and unsatisfactory data fitting values, will be rejected. The main goal of the simultaneous analysis is to find a resolution of species which on one hand fits adequately the data (lack of fit values at the level of experimental error, i.e., below 1%), and which on the other hand, finds a

chemically 'sound' resolution of the system. In contrast to the species spectra in matrix $\mathbf{S}^T$, the concentration profiles of one species in different experiments or process runs ($\mathbf{D}_k$ in $\mathbf{D}_{aug}$) are allowed to be different, in shape and in intensity. This property is specially relevant in many chemical process systems, where the equality of shape of concentration profiles in the different runs cannot be accomplished, and therefore, the three-way data structure is not trilinear. In the particular case that the considered species fulfill this condition (equal shape of its concentration and spectrum profile in all the experiments simultaneously analyzed) an additional trilinearity can be applied as described in previous publications [7,8, 14,39]. In any case, calculation of pseudoinverses in the least squares estimation of matrices $\mathbf{C}_{aug}$ and $\mathbf{S}^T$ is not problematic if these matrices are full rank matrices.

Finally, an additional advantage of the simultaneous analysis of several experiments using column-wise augmented data matrices and Eqs. (4)–(6) compared to the analysis of individual data matrices and Eqs. (1)–(3), is that rank deficiency problems present for individual matrices can be solved by matrix augmentation [15–17] in the simultaneous experiments data analysis. As stated by Amrhein et al. [15], column-wise matrix augmentations helps solving rank deficiencies present in the matrix of concentrations, $\mathbf{C}$, for reaction based systems, where the number of independent concentration profiles can be lower than the number of real species present in the systems due to the reaction network constraints, mass balance, mass action law, and closure of these systems. The important thing to emphasize here is that, species hidden in the analysis of a single experiment or process run can be resolved in the simultaneous analysis of several experiments or process runs using the column-wise augmented data matrix approach. The example selected in this work shows clearly this situation.

4. Results and discussion

4.1. Analysis of the poly(adenylic acid) melting experiment

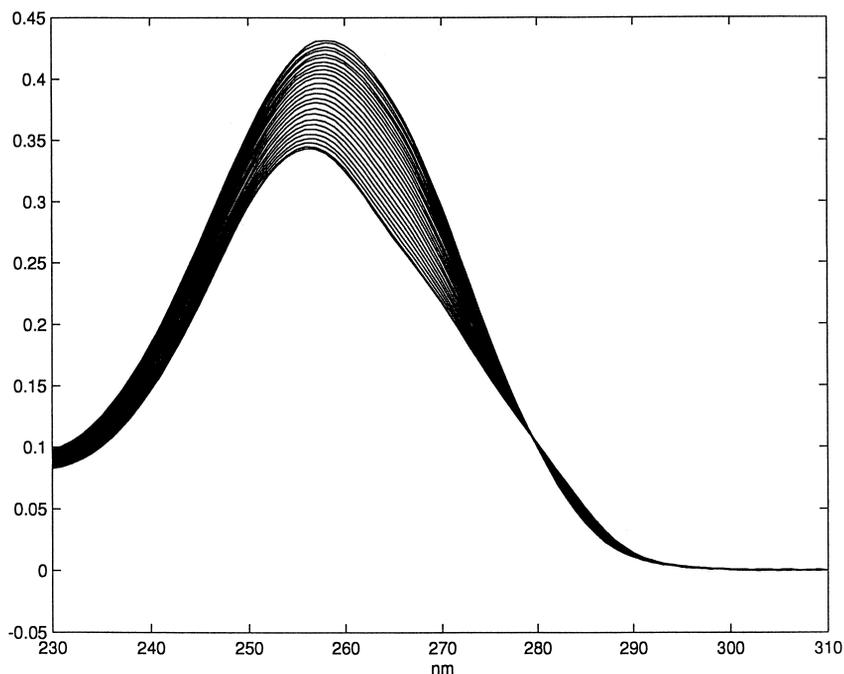In the poly(A) melting experiment A (see Fig. 1) some spectral changes were observed when the tem-

Fig. 1. Experimental UV spectra obtained for poly(adenylic) acid (poly(A)) at pH 6.8 and varying temperatures from 15°C to 90°C (experiment A).

perature was changed. An isosbestic point is present at 279.4 nm. Rank analysis of the corresponding data matrix (see Table 2) confirmed two major contributions with large singular values followed by much lower ones related with experimental noise. In Table 3, PCA lack of fit for two components is only a 0.12%. Fig. 2 shows EFA plots where two major contributions are observed, with log of singular values reaching values higher than -1. At the bottom of the plot, singular values related with pure noise

Table 2
Rank analysis of the experiments

Singular value no.[a]	Exp. A	Exp. B	Exp. C	Exp. D	Exp. A, B, C and D[b]
10	0.0018	0.0016	0.0024	0.0026	0.0051
9	0.0018	0.0016	0.0025	0.0027	0.0098
8	0.0019	0.0017	0.0026	0.0030	0.0126
7	0.0021	0.0018	0.0031	0.0036	0.0212
6	0.0023	0.0019	0.0057	0.0078	0.0391
5	0.0035	0.0020	0.0211	0.0228	0.0738
4	0.0048	0.0030	0.0286	0.0290	0.1719
3	0.0071	0.0055	0.1492	0.1925	0.5889
2	0.2155	0.0364	1.0854	1.0597	1.6169
1	9.6942	7.6901	20.6542	21.4058	32.2077

[a] Singular values at increasing size order.
[b] Singular values of the augmented column-wise data matrix built up with the different experiments (A, B, C and D).

Table 3
MCR-ALS data fitting results

	Matrix **A**	Matrix **B**	Matrix **C**	Matrix **D**	Augmented matrix [**A;B;C;D**]
PCA lof[a]	0.12 (2)[b]	0.48 (1)	0.18 (3)	0.18 (3)	0.17 (5)
ALS lof[c]	0.27 (2)	0.48 (1)	0.43 (3)	0.36 (3)	0.43 (5)

[a] PCA lack of fit calculated by means of the equation:

$$\text{PCA lof} = 100 \sqrt{\frac{\sum_{i,j}\left(d_{i,j} - d_{i,j}^{c}\right)^{2}}{\sum_{i,j} d_{i,j}^{2}}},$$

d_{ij} experimental values for spectrum i, wavelength j; d_{ij}^{c} PCA calculated values for spectrum i, wavelength j.
[b] In parenthesis, number of species considered in the calculations.
[c] ALS lack of fit (at convergence with a 0.01% of change of fit between iterations). The same formula as in [a], but with d_{ij}^{c} are the ALS calculated values for spectrum i, wavelength j.

with log values below -2.5 are shown. Three other small contributions with log of singular values between -2.5 and -2 appear close to the pure error ones. These contributions are related with non-random noise contributions due to experimental uncertainties of unknown nature (detector drifts, non-linearities, solvent and medium spectral effects, etc.). Unfortunately, these contributions are frequently and
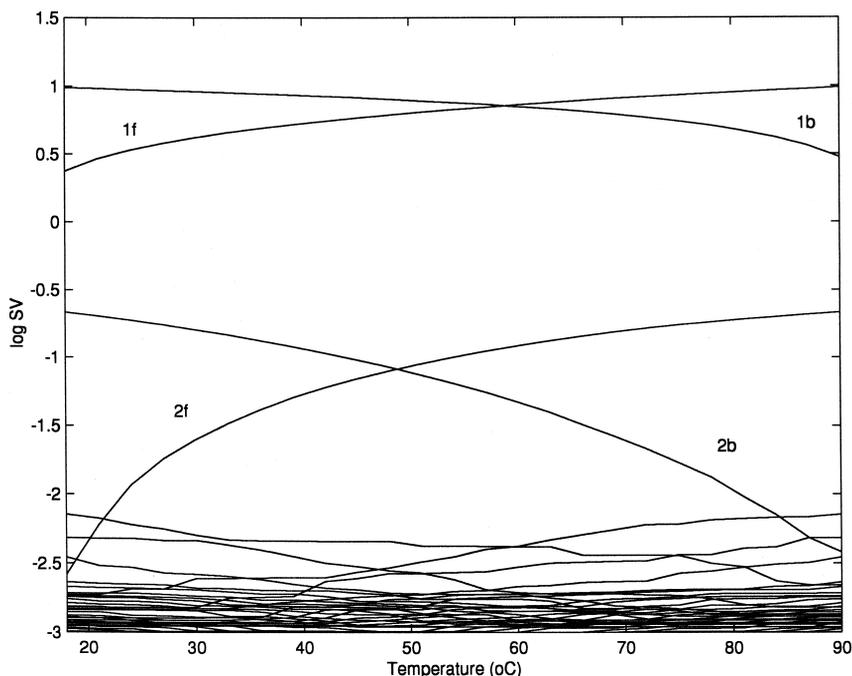

Fig. 2. Evolving Factor Analysis (EFA) plot of experiment A (Fig. 1). Y-axis gives the log of the singular values found by EFA. 1f and 2f are the lines corresponding to the first and second forward singular values found by EFA. 1b and 2b are the lines corresponding to the first and second backward singular values found by EFA See text and Refs. [3,25,26] for interpretation.

unavoidable present in experimental data making difficult the accurate assertion of the possible presence of minor chemical contributions from EFA plots. Considering the presence of two chemical contributions, an initial estimation of the two evolving profiles is obtained from EFA plots (Fig. 3). Resolution of the two conformations associated with these two contributions is achieved by MCR-ALS using the additional assumption that selectivity is present at the lowest and highest temperatures, where each conformation is considered as the only one present at these temperatures. Under these assumptions, ALS lack of fit (Table 3) was only of 0.27% and the resolution is unique [7,10]. The plot shown in Fig. 4 gives the relative concentrations evolution from the conformation present at the lowest temperature to that present at the highest temperature. The resolved concentration profiles describe the thermal denaturation. of the polynucleotide conformational structure. The species resolved at lower temperature is mostly related with a single-stranded helical (ordered) conformation whereas that resolved at higher temperatures is mostly related with the denatured (disordered) poly(A) conformation. The spectra of these two conformations are

shown in Fig. 5. As expected, the denaturation process increases the absorbance values and shifts slightly the absorbance band to higher wavelengths. Both spectra cross (isosbestic point) at 279 nm and have their maximum absorbance at 256 nm (native) and 258 nm (denatured).

4.2. Analysis of the poly(uridylic acid) melting experiment

In poly(U) melting experiment 2 (see Fig. 6), very small absorbance changes in the UV spectra were observed between 15 and 90°C, which indicates that poly(U) is always present in the same conformation, a single disordered random coil conformation [40]. Rank analysis of the data matrix corresponding to such experiment (see Table 2), gives a first singular value which is much larger than the second one (211 times). The second singular value is larger than expected for a system with a single species and probably describes small spectral changes and shifts due to temperature which are not assigned to significant conformational changes (see below). PCA lack of fit with a single component is, in this case, 0.49%, a lit-
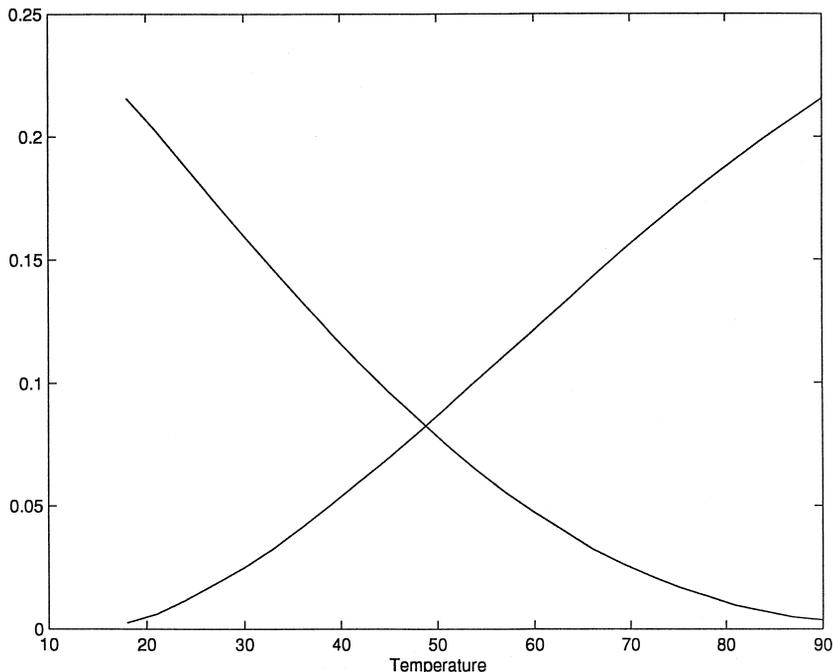

Fig. 3. Initial estimation of concentration profiles obtained from EFA plot of experiment A (Fig. 2).
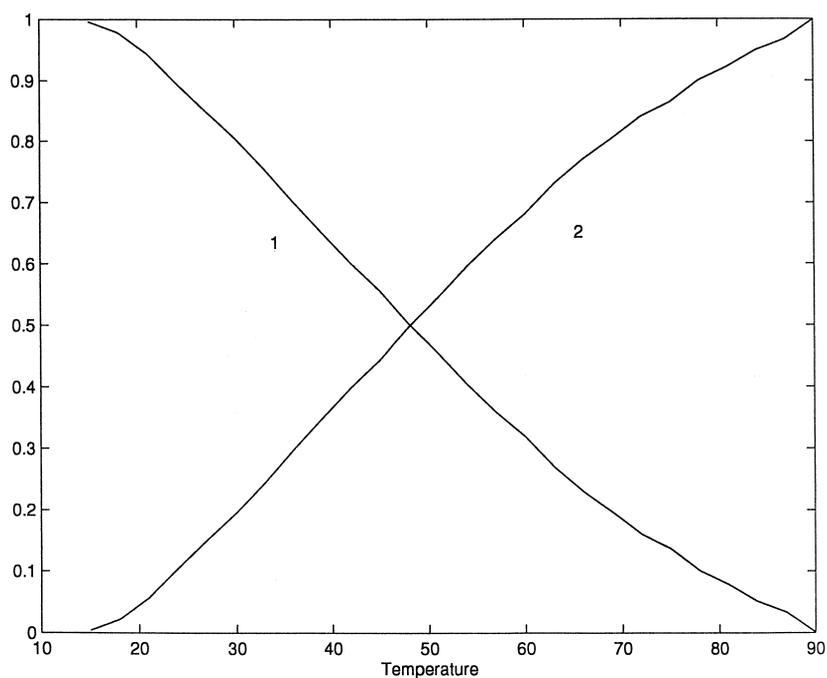
Fig. 4. MCR-ALS optimized concentration (melting) profiles for the two species detected in experiment A (poly(A)). Species 1 is the native species at lowest temperature and species 2 is the denatured species at highest temperatures.
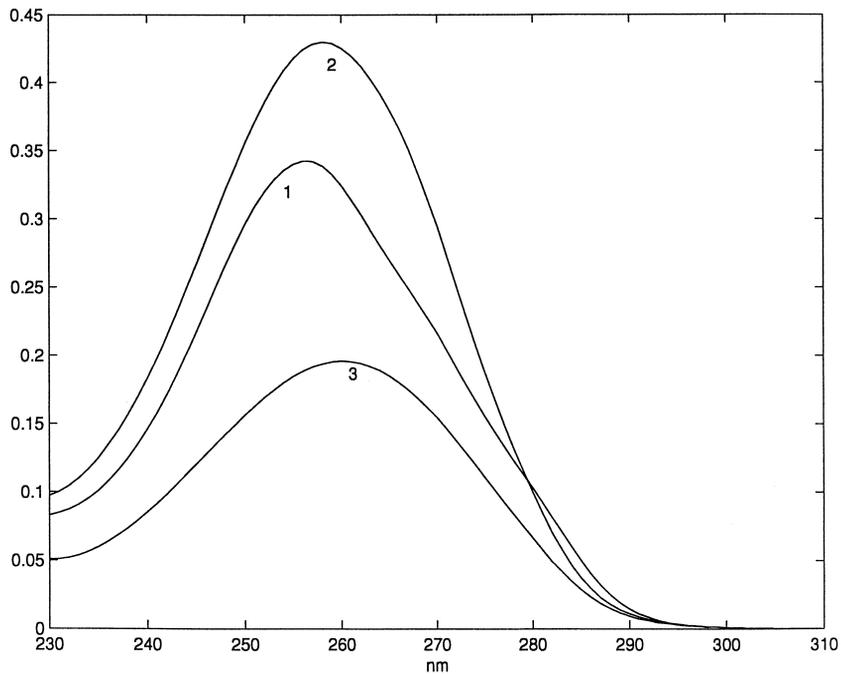

Fig. 5. MCR-ALS optimized spectra profiles for the species detected in experiment A. Species 1 and 2 (poly(A)) and species 3 (poly(U)).
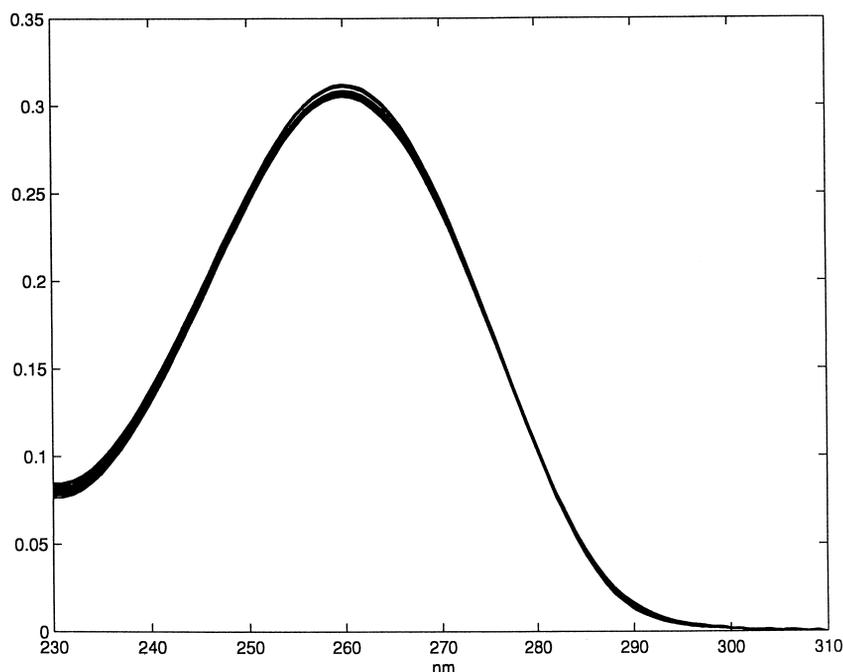
Fig. 6. Experimental UV spectra obtained for poly(uridylic acid) (poly(U) at pH 6.8 and varying temperatures from 15°C to 84°C (experiment B).

tle higher than in poly(A) for two components. In Fig. 7, the EFA plot of the poly(U) melting experiment is given. The second singular value reaches levels up to log values of −1.5, one log unit below that the second singular value in the poly(A) experiment. If the resolution of a second conformation is attempted by MCR-ALS, the species spectra of the two resolved species are practically identical, with a correlation value between them higher than 0.9999 and with only a small difference at the lowest wavelengths, close to 230 nm, where solvent and salt medium contributions can be more important. Therefore, only one component is considered for this system. No further data treatment is necessary in this case since the same single conformation is present along the whole experiment. Its spectrum is given together with those resolved for poly(A) in Fig. 5.

4.3. Analysis of the poly(adenylic acid)–poly(uridylic acid) melting experiment

Double stranded poly(A)–poly(U) melting experiments C and D (Fig. 8) showed significant spectral changes between 15°C and 90°C. A first spectral change is rather weak and occurs mostly between 40°C and 60°C, with an isosbestic point at approximately 276 nm and an small increase of absorbance at 257 nm. This spectral change is difficult to detect at the lower temperatures, below 35°C, but it becomes more apparent at temperatures around 45°C. A second stronger spectral change is produced very fast, in less than 10°C, between 58°C and 68°C, with a considerable increase of absorbance and a weak shift to higher wavelengths. These fast temperature dependent spectral changes are usually associated with melting processes [23,26,40]. Rank analysis of the two data matrices obtained in experiments C and D showed three main contributions with three large singular values (Table 2). In Fig. 9, the EFA plots with a non-fixed (9a) and a fixed (9b) size window are shown. In the first plot, Fig. 9a, three main contributions with log of singular values larger than −1.5 and four intermediate contributions with log of singular values between −2.5 and 1.5 are obtained. The first contributions are related with three main chemical changes, and the second contributions are related with
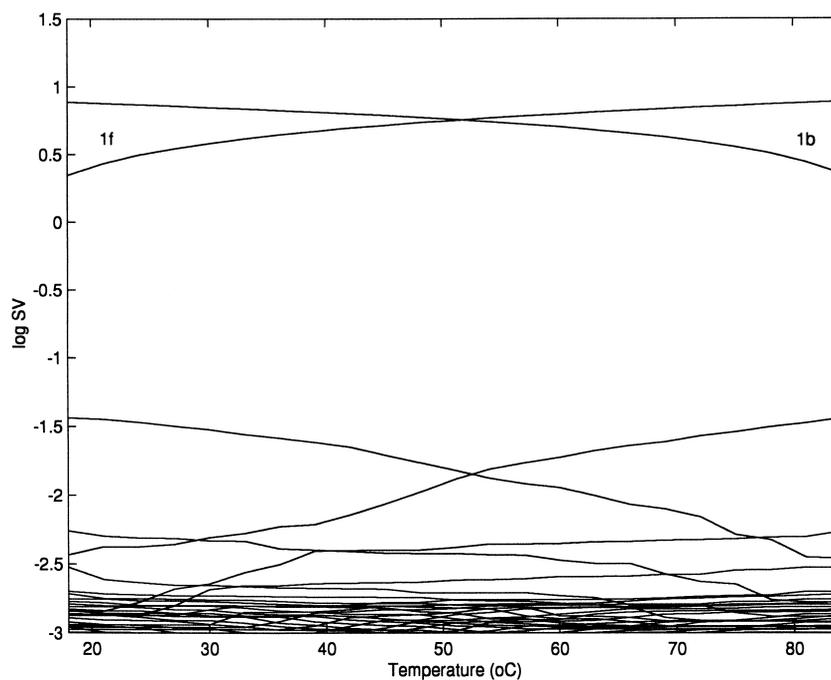
Fig. 7. Evolving Factor Analysis (EFA) plot of experiment B (Fig. 6). 1f is the line corresponding to the first forward singular value found by EFA. 1b is the line corresponding to the first backward singular value found by EFA. See text and Refs. [3,25,26] for interpretation.
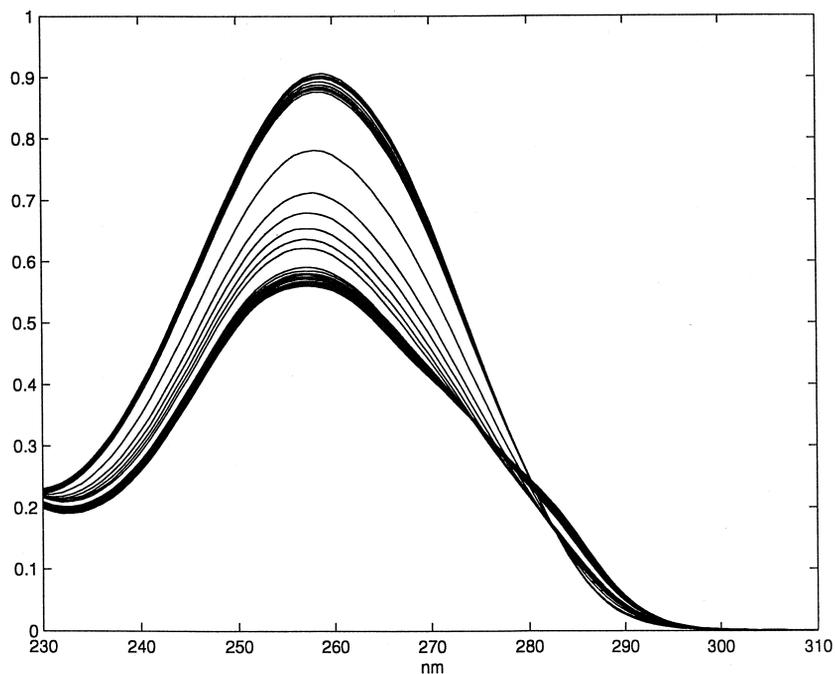

Fig. 8. Experimental UV spectra obtained for poly(adenylic acid)–poly(uridylic acid) (poly(A)–poly(U)) at pH 6.8 and varying temperatures from 15°C to 90°C (experiment D).
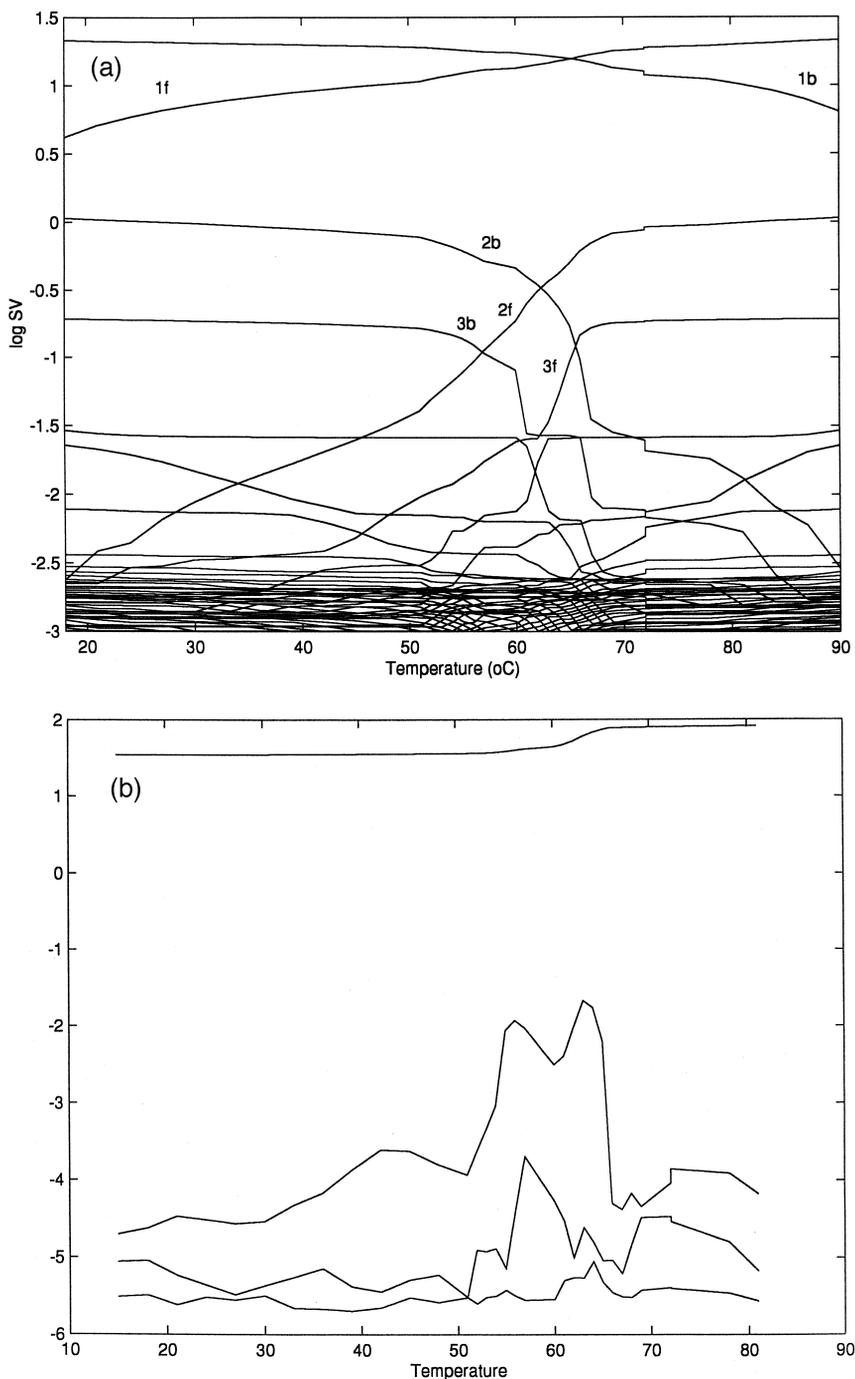
Fig. 9. (a) Evolving Factor Analysis (EFA) plot of experiment D (Fig. 8). 1f, 2f and 3f are the lines corresponding to the first, second and third forward singular values found by EFA. 1b, 2b and 3b are the lines corresponding to the first, second and third backward singular values found by EFA. See Refs. [3,25,26] for interpretation. (b) Evolving Factor Analysis With a Fixed Size Moving Window (FSMWEFA) plot of experiment D (Fig. 8). See text and Ref. [27] for interpretation.

other sources of non-random error contributions. At the bottom, with log values below -2.5, the pure error contributions appear. In the second plot, Fig. 9b, local rank analysis with evolving factor analysis with a fixed size moving window, showed two rank one regions at the beginning and at the end of the experiment (no important changes of speciation in this regions) and one rank two region or even rank three region at intermediate temperatures, between 50–70°C. This confirms that at this intermediate region, significative changes in speciation occur, related with a fast change of conformations.

Resolution of these three contributions using the MCR-ALS procedure gave the three species melting profiles given in Fig. 10 and the species spectra given in Fig. 11. ALS lack of fit values for the two experiments C and D, gave respectively (Table 3), 0.43% and 0.36%. In both cases, these values are considered very low and close to the values found by PCA analysis for the same number of components (PCA lack of fit 0.18% in both cases). The nature of the re-

solved species is difficult to be elucidated from the analysis of this single data matrix. As it was previously indicated, fast increases of absorbance values such as those detected around 65°C at 257–259 nm are usually associated related with conformational changes giving disordered denatured molecular forms of the polynucleotide. The two species spectra resolved for the two first species (I and II in Fig. 11) have a band maximum at approximately the same wavelength (257 nm); the first species spectrum (I in Fig. 11) has a weak shoulder around 278 nm that has disappeared in the second species spectrum. The species spectrum of the third resolved species (III in Fig. 11) has a band maximum with a weak shift of nearly 2 nm (maximum at 259 nm) and a higher intensity. These resolved band shapes and maxima are not coincident with those resolved in the melting experiments of the homopolynucleotides poly(A) and poly(U). Experiments C and D, only differ in the total amount of polynucleotide initially present in the melting process. When the MCR-ALS resolution re-
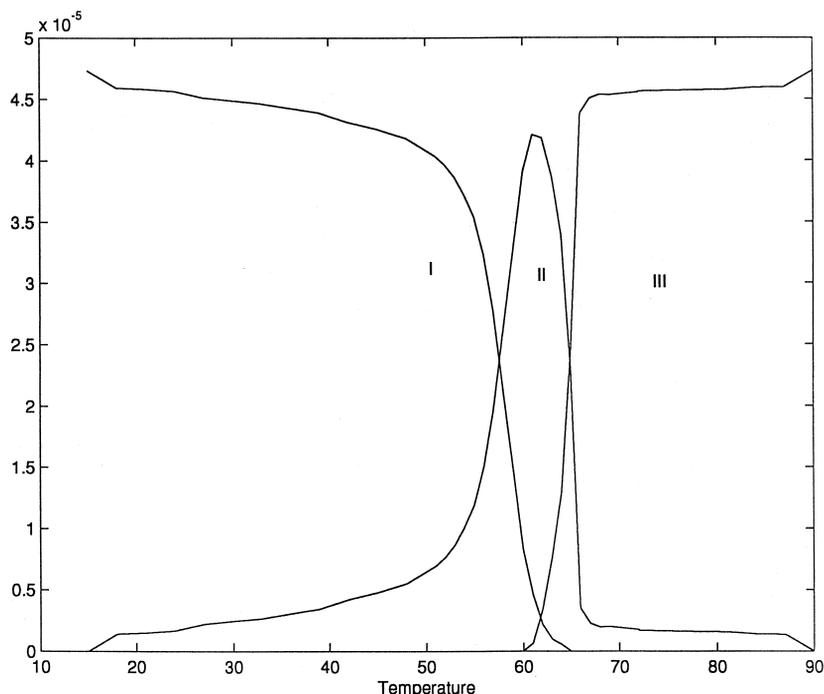

Fig. 10. MCR-ALS optimized concentration (melting) profiles for the three species detected in experiment D (rank deficient data matrix) when temperature is changed. Species I is the native species at low temperatures, species II is an intermediate species and species III is the denatured species at higher temperatures.
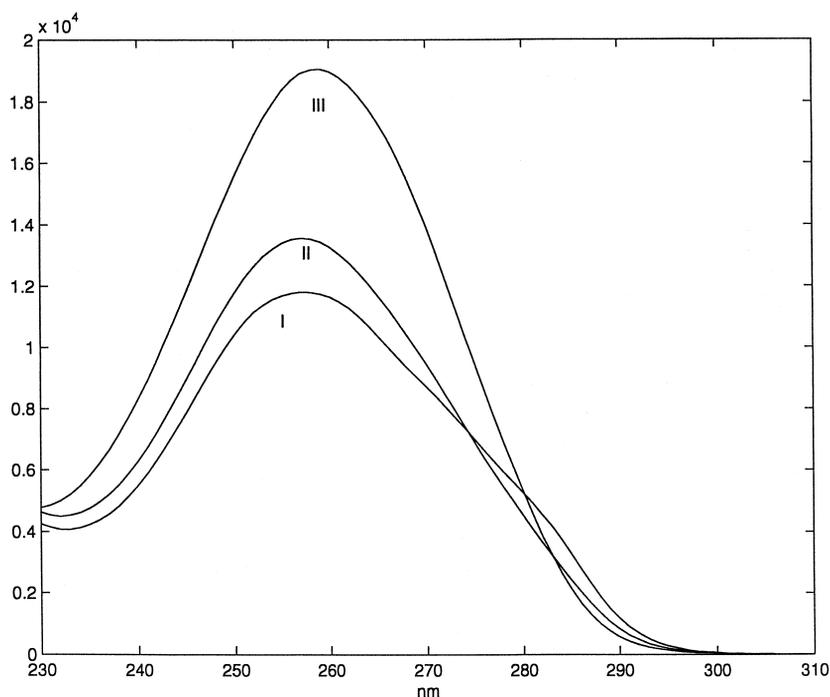
Fig. 11. MCR-ALS optimized spectra profiles for the three species detected when temperature is changed in experiment D (rank deficient data matrix). Species I, II and III the same as in Fig. 10.

4.4. Simultaneous analysis of poly(adenylic acid), poly(uridylic acid) and poly(uridylic acid)–poly(adenylic acid) melting experiments (A, B, C and D)

As it was indicated in the data pretreatment section, in the simultaneous analysis of several experiments (Eq. (4)), a row data normalization is first applied. In this way, problems associated with polynucleotide concentration correspondences between different experiments and with the closure constraint are eliminated. Moreover, this row-normalization facilitates the ALS optimization and the interpretability of results.

When the individual melting experiments A, B, C and D are all together simultaneously analyzed using column-wise data matrix augmentation (see Eq. (4)), rank analysis of this new augmented data matrix gave

sults of these two matrices are compared, no significant differences were observed, with the same trends in the description of the melting process evolution.

singular values four and five larger than those found in the individual data analysis of matrices **C** and **D** (Table 2). This rank increase can be related with different hypothesis about the correspondence of the different species between experiments (common species) and also with the possible presence of rank deficiency problems in some of the data sets, specially in experiments C and D when analyzed individually. It should be noted first, that every common species in the different experiments, should have a rank one contribution in the column-wise augmented matrix, and second, that the total number of different species deduced from individual experiments is six: in experiment A two species, in experiment B one species and in experiments C and D three species. However rank analysis of the augmented matrix suggested a maximum number of five independent species. This could be explained if some of the species in experiments A and B were also present in experiments C and D. In a previous study of the melting behavior of poly(inosinic acid)–poly(citydilic acid) [23], it was found that the melting of this

heteropolynucleotide, gave an admixture of the denatured constituent homopolynucleotides, poly(I) and poly(C). Therefore, it is also reasonable to expect a similar situation in the melting experiments of poly(A)–poly(U), where at high temperatures the double stranded poly(A)–poly(U) is converted to an admixture of the homopolynucleotide denatured conformations of poly(A) and poly(U). Following this hypothesis, spectral initial estimates of the $\mathbf{S}^T$ matrix were available from the resolved spectra in the individual matrices of poly(A) and poly(U). A fourth species spectrum was obtained from the first experimental spectrum of the experiments C and D (poly(A)–poly(U) native form) and a fifth spectrum was obtained from the same experiments at the intermediate region where strong spectral changes were observed (melting), selecting the one which is more pure or selective [31]. This makes a total of five spectra in $\mathbf{S}^T$, from which an initial estimation of $\mathbf{C}_{aug}$ is immediately available using Eq. (6). Selectivity, non-negativity, closure, and unimodality constraints are applied to this augmented concentration matrix $\mathbf{C}_{aug}$, whereas only nonnegativity constraint is applied to $\mathbf{S}^T$ during the ALS optimization. Selectivity

at the beginning of experiments is justified from the point of view of defining the initial state or conformation of the system, which is postulated to be the one present at the lowest temperatures of the study. In all the experiments it was observed that the spectral changes observed between 15°C and 37°C (biological conditions) were small and that the more important changes, related with the denaturation of the native conformations present at biological condition, occurred at temperatures over 50°C. It is between 50°C and 70°C where the more important spectral changes are always observed (Figs. 4 and 11). Two additional constraints related with the correspondence of species between experiments and with the equality of species spectra of common species in different experiments were applied.

Fig. 12 gives the best set of five (concentration) melting profiles resolved by MCR-ALS for the heteropolynucleotide poly(A)–poly(U) in experiment D. Of these five resolved species, two were coincident with those found in the poly(A) experiment A and another one was coincident with the only one detected in the poly(U) experiment B. Lack of fit (lof) values using five species were very good (ALS lof
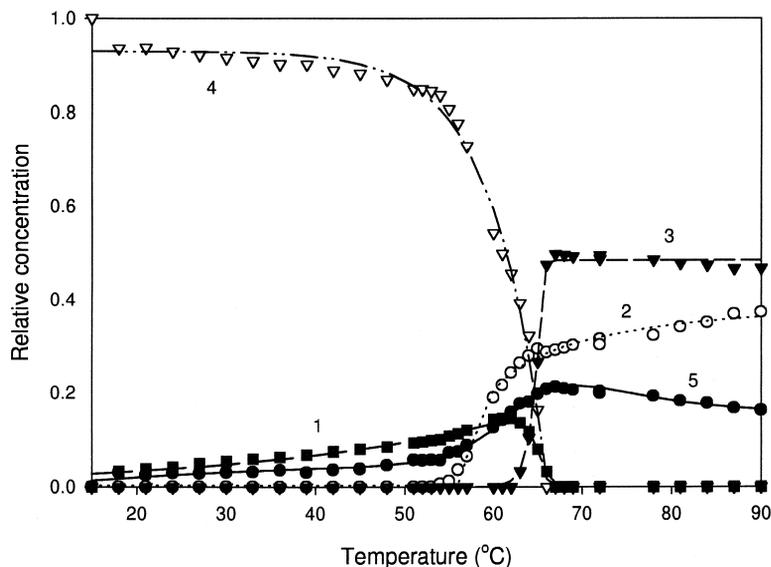

Fig. 12. MCR-ALS optimized concentration (melting) profiles for the five species detected when temperature is changed in experiment D when it is analyzed simultaneously with experiments A, B and C, using full rank column-wise matrix augmentation (Eqs. (4)–(6)). Species 1 and 2 refer to the same species as those resolved in the experiment A for poly(A) (Fig. 4 and Fig. 5); species 3 refers to the single species detected for poly(U) in experiment B; species 4 and 5 are two new species detected only for poly(A)–poly(U) in experiments C and D (see text for interpretation).

value of only 0.43% in Table 3) close to lof values previously found in the individual analysis of experiments and also close to that obtained by PCA (PCA lof value of 0.17%) for the same number of components. Both experiments C and D gave similar melting profiles. Postulation of other number of species or of another correspondence between species gave much worse data fittings and unreasonable shapes of the melting profiles and of the species spectra. Two new species not found in experiments A and B were detected in experiments C and D (Figs. 12 and 13). One of these new species (number 4 in Figs. 12 and 13) is the species present at the lowest temperatures of the study. The other species (number 5 in Figs. 12 and 13) is a new species not found in the previous studies. The other three species spectra corresponding to species 1, 2 and 3 are identical to those given in Fig. 3.

Rank deficiency found in the individual analysis of matrices **C** or **D** was solved in the simultaneous analysis of experiments A, B, C and D by column-wise matrix augmentation. This is in agreement with previous studies of reaction based rank deficient systems [15–17]. Comparison of resolved profiles in the simultaneous analysis of experiments A, B, C and D, with those obtained in the individual analysis of experiments C or D, confirmed that data matrices corresponding to the melting experiments of poly(A)–poly(U), C and D, were rank deficient. This rank deficiency was broken in the simultaneous study using the corresponding column-wise augmented matrix and MCR-ALS. This was possible because the different matrices included in the augmented matrices gave independent information about the system. Resolution of the constituent species was possible because much better resolution conditions [10] were present in the column-wise augmented matrix than in the individual matrices. For instance, the three species related with the individual homopolynucleotides poly(A) and poly(U) (species 1, 2 and 3) could not be resolved in the individual analysis of the data matrices **C** and **D** because these two matrices were rank deficient and also because the melting profiles of these three species were totally embedded inside the
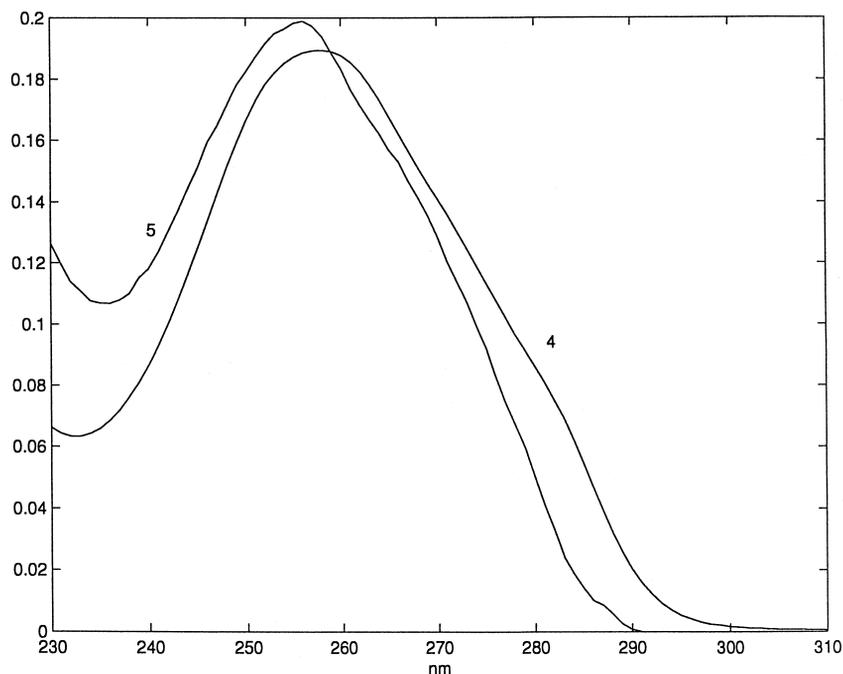

Fig. 13. MCR-ALS optimized spectra profiles for the two new species (4 and 5) detected when temperature is changed in experiments C and D, when they are analyzed simultaneously with experiments A and B using full rank column-wise matrix augmentation (Eqs. (4)–(6)). The melting profiles of these two species 4 and 5 are given in Fig. 12.

others (Fig. 12) and they do not have any selective or favorable local rank region needed for its resolution. In contrast to this, in the augmented data matrix, these three species fulfill the resolution conditions due to the inclusion of the experiments A and B, where the mentioned species have selectivity and favorable local rank regions (see Fig. 4) for poly(A) species, and there is only one poly(U) species. Possible rotational ambiguity in spectrum of species 4 is also solved by the postulation of a single native conformation at the beginning of the experiment. The only remaining rotational ambiguity would be reduced to species spectrum 5. In Table 4, information about the resolved profiles for each species is given. Whereas, the resolved species spectra are highly correlated, the resolved melting profiles are more different between them, although they are embedded, specially in experiments C and D, making more difficult their resolution, specially in the analysis of the individual experiments. Also in Table 4, the values of the absorption band maxima are given. Only seven nanometers is the highest separation between band maxima. In spite of all these difficulties, the proposed ALS-MCR method provided a reasonable semi-quantitative description of the whole process under study.

In accordance with current biochemical literature [40,41], a possible interpretation of the process shown in Fig. 12 is given. At low temperatures, poly(A)–
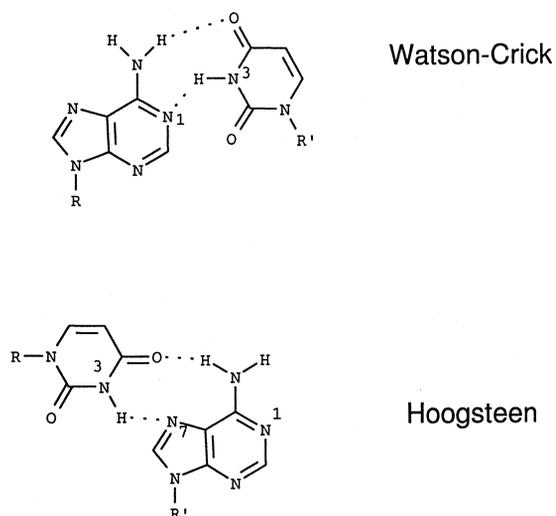

Fig. 14. Proposed scheme of base pairing explaining formation of species 5 in melting experiments C and D (see text and Refs. [37,38] for interpretation).

poly(U) is present as a double stranded structure via Watson–Crick base pairing (see Fig. 14), yielding a structure similar to that found in natural polynucleotides. This is the structure proposed for component 4 in Fig. 13. On the other hand, some authors have observed that it is also possible to find ordered triple stranded structures [40,41] between one poly(A) strand and two poly(U) strands. In this structure one of the poly(U) strands is linked to the poly(A) strand via Watson–Crick base pairs while the other poly(U) strand is linked to poly(A) via Hoogsteen-type base pairs (see Fig. 14). This structure is possible because the adenine heterocycle is simultaneously able to engage in both Watson–Crick and Hoogsteen base pairs since the functional groups in both processes are different. This triple helix conformation has been experimentally obtained from disproportionation of the double stranded Watson–Crick one as salt concentration is raised [40,41]. An increase of the temperature facilitates also the disproportionation process, yielding components 1 and 5 in Fig. 13, which can be related respectively with free single stranded poly(A) and with the postulated poly(U)–poly(A)–poly(U) helix. Thermal denaturation of free poly(A) gives species 2, with a fast increase of concentration close to the melting temperature of the poly(A)–poly(U)

Table 4
Correlation between spectra and melting profiles of the 5 MCR-ALS resolved species

(A) Species spectra

	1	2	3	4	5	Wavelength[a] (nm)
1	1					256
2	0.9972	1				258
3	0.9941	0.9958	1			260
4	0.9936	0.9881	0.9946	1		257.5
5	0.9826	0.9839	0.9792	0.9761	1	256.5

(B) Melting profiles

	1	2	3	4	5
1	1				
2	0.4567	1			
3	0.0577	0.2025	1		
4	0.1347	0.0580	0.0034	1	
5	0.2272	0.2754	0.1121	0.7758	1

[a]Wavelength of the maximum absorption of the considered species spectrum.

system. Simultaneously, the homopolynucleotide single poly(U) species appears at the melting temperatures (between 60–70°C). This description of the melting process of poly(A)–poly(U) would confirm that, as in the case of the poly(I)–poly(C) system [22], the melting of this heteropolynucleotide gave the two separated homopolynucleotides poly(A) and poly(U) in the same conformation as they were detected in the individual melting experiments of these systems (species 1, 2 and 3 in Fig. 3).

References

[1] B.M. Wise, B.R. Kowalski, Process chemometrics, in: F. McLennan, B.R. Kowalski (Eds.), Process Analytical Chemistry, Blackie Acad. and Profes., London (1995).

[2] E.R. Malinowski, Factor Analysis in Chemistry, 2nd edn., Wiley, New York (1991).

[3] M. Maeder, Anal. Chem. 59 (1987) 527.

[4] P. Geladi, S. Wold, Chemomet. Intell. Lab. Syst. 2 (1987) 273.

[5] W.H. Lawton, E.A. Sylvestre, Technometrics 13 (1971) 617.

[6] O.S. Borgen, B.R. Kowalski, Anal. Chim. Acta 174 (1985) 1.

[7] R. Tauler, A. Smilde, B.R. Kowalski, J. Chemometrics 9 (1995) 31.

[8] R. Tauler, Chemomet. Intell. Lab. Syst. 30 (1995) 133.

[9] B.G.M. Vandeginste, W. Derks, G. Kateman, Anal. Chim. Acta 173 (1985) 253.

[10] R. Manne, Chemomet. Intell. Lab. Syst. 27 (1995) 89.

[11] R. Tauler, B.R. Kowalski, S. Flemming, Anal. Chem. 65 (1993) 2040.

[12] R. Tauler, A. Izquierdo-Ridorsa, E. Casassas, Chemomet. Intell. Lab. Syst. 18 (1993) 293.

[13] E. Sanchez, B.R. Kowalski, J. Chemometrics 4 (1990) 29.

[14] A. de Juan, S.C. Rutan, R. Tauler, D.L. Massart, Chemomet. Intell. Lab. Syst. 40 (1998) 19.

[15] M. Amhrein, B. Sirnivassan, D. Bonvin, M. Schumacher, Chemomet. Intell. Lab. Syst. 33 (1996) 17.

[16] A. Izquierdo-Ridorsa, J. Saurina, S. Hernández-Cassou, R. Tauler, Chemomet. Intell. Lab. Syst. 38 (1997) 183.

[17] E. Casassas, R. Gargallo, A. Izquierdo-Ridorsa, R. Tauler, Reactive and Functional Polymers 27 (1995) 1.

[18] A. de Juan, A. Izquierdo-Ridorsa, R. Tauler, G. Fonrodona, E. Casassas, Biophysical J. 73 (1997) 2937.

[19] E. Casassas, R. Tauler, M.I. Marques, Macromolecules 27 (1994) 1729.

[20] J. Mendieta, M.S. Diaz-Cruz, M. Esteban, R. Tauler, Biophysical J. (1998), in press.

[21] J. Mendieta, M.S. Diaz-Cruz, R. Tauler, M. Esteban, Anal. Biochem. 240 (1996) 134.

[22] R. Gargallo, R. Tauler, A. Izquierdo-Ridorsa, Anal. Chem. 68 (1997) 1785–1792.

[23] E. Casassas, R. Gargallo, I. Gimenez, A. Izquierdo-Ridorsa, R. Tauler, Anal. Chim. Acta 283 (1993) 538.

[24] W. Saenger, Principles of Nucleic Acid Structure, Springer-Verlag, New York (1988).

[25] H. Gampp, M. Maeder, Ch. Mayer, A.D. Zuberbühler, Talanta 33 (1986) 943.

[26] R. Tauler, E. Casassas, J. Chemometrics 3 (1988) 151.

[27] H.R. Keller, D.L. Massart, Anal. Chim. Acta 246 (1991) 379.

[28] O.M. Kvalheim, Y.Z. Liang, Anal. Chem. 64 (1992) 936.

[29] R. Gargallo, F. Cuesta-Sànchez, D.L. Massart, R. Tauler, Trends in Analytical Chemistry 15 (1996) 279.

[30] E. Spjotwoll, H. Martens, R. Volden, Technometrics 3 (1982) 173.

[31] E.R. Malinowski, J. Chemometrics 6 (1992) 29.

[32] C.L. Lawson, R.J. Hanson, Solving Least Squares Problems, Prentice Hill, Englewood Cliffs, NJ (1974).

[33] R. Bro, S. de Jong, J. Chemometrics 11 (1997) 393.

[34] R. Bro, N.D. Sidiropoulos, submitted to J. Chemometrics.

[35] A. de Juan, A. Izquierdo-Ridorsa, R. Gargallo, R. Tauler, G. Fonrodona, E. Casassas, Anal. Biochem. 249 (1997) 174.

[36] R. Tauler, I. Marqués, E. Casassas, J. Chemometrics 12 (1998) 55.

[37] R.D. Blake, J. Massoulié, J.R. Fresco, J. Mol. Biol. 30 (1967) 291.

[38] J. Massoulié, Eur. J. Biochem. 3 (1968) 439.

[39] R. Tauler, D. Barceló, Trends in Analytical Chemistry 12 (1993) 319.

[40] R.S. Preisler, C. Mandal, S.W. Englander, N.R. Kallenbach, Biopolymers 23 (1984) 2099.

[41] C. Stevens, G. Felsenfeld, Biopolymers 2 (1964) 293.