Interactions of H^+ and Cu(II) Ions with Poly(adenylic acid): Study by Factor Analysis

E. Casassas, R. Tauler,* and I. Marqués

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

Received August 5, 1993; Revised Manuscript Received November 17, 1993•

ABSTRACT: The interactions of H^+ and Cu(II) ions with poly(adenylic acid) (polyA) are investigated using different factor analysis techniques. From the changes in the ultraviolet and circular dichroism absorption spectra obtained in the course of spectroscopic acid–base titrations of poly(adenylic acid) solutions containing different amounts of Cu(II) ion, the unitary spectra, the concentrations, and the stabilities of the macromolecular species detected are estimated. Four spectroscopically different species are detected for poly(adenylic acid) in the pH interval 1.5–6.8. Three of them are different conformations of the protonated form of polyA, and the other is the unprotonated form of polyA. One new species is detected in the presence of Cu(II) ion.

Introduction

The study of equilibria between metal ions and macromolecular ligands is a field of great interest owing to its environmental and biological importance.[1,2] This study, however, is hindered by the fact that the law of mass action ruling the complexation equilibria is valid only separately for each one of the reaction sites of the macromolecule and several additional or secondary effects must be considered. These secondary effects have been classified[3] into three types: (a) polyfunctional effects assigned to differences in chemical nature and in electrostatic and steric environments of the coordination sites in the macromolecule, (b) conformational changes caused by the changes in pH and ionic strength of the medium or by the content of complexed ion, and (c) polyelectrolytic effects caused by the ionization of major sites of the macromolecule yielding changes in the local electric field at the surface of the macromolecule. All these effects contribute to the stability of the formed species, and their relative importance is difficult to define since it varies with the degree of site occupation (complexation or protonation).

The interpretation of the experimental data using traditional least-squares curve fitting approaches is consequently rather cumbersome and unsafe,[3–5] and there is a demand for the development of new approaches free from the constraints of the law of mass action and free from the prior postulation of a chemical model.

The study of conformation and structural transitions in synthetic polynucleotides such as poly(adenylic acid) (polyA) is important for a better understanding of the structures and interactions in naturally occurring nucleic acids.[6] These homopolynucleotides allow evaluation of constants of ion binding to predetermined types of bases.[7]

The reason for the significant interest in the studies of ion binding to nucleic acid components is that metal ions have profound effects on the structure and function of nucleic acids and mono- and polynucleotides. They alter the coding specificity of polynucleotides acting as templates for protein synthesis and are required for stabilizing the structure of tRNA. Their role in unwinding and rewinding of the double helix and in degradation of polynucleotides has been demonstrated.[8] The impact of copper ions on the conformation of nucleic acids is equally dramatic and of potential significance in disturbing the function of the genetic material.[9]

It has been well established that poly(adenylic acid) exhibits two conformations: helical, single stranded with partially stacked bases at neutral pH, and helical, double stranded with parallel chains and stacked protonated bases at acidic pH. However, there is evidence that poly(adenylic acid) can assume more than one conformation at acidic pH,[10–12] depending on the extent of protonation of the molecule and that a form (A) can exist associated with the fully protonated state while another form (B) would be associated with the partly protonated state.[11] Results based on circular dichroism (CD), electron paramagnetic resonance, and voltammetry yield further evidence that the transition of form B into form A occurs through a third form.[10]

In the present work, a new approach based on factor analysis techniques[13] is proposed for the study of the interaction of H^+ and Cu(II) ions with poly(adenylic acid). The method, valid in general for the study of equilibria involving macromolecular ligands, is based on our previously developed SPFAC procedure.[14–17] It is developed for the detection of the number of species present in the system, for the estimation of the concentration profile and the individual spectra of the species in equilibrium along a spectroscopic titration, and, in this particular case, for the deduction of the equilibrium properties of the macromolecule. Two related spectrometric techniques were used, UV normal absorption and UV circular dichroism absorption. The latter is a convenient method to follow conformational changes of macromolecules and although it is very suitable for the study of multiequilibria systems, it has been scarcely used for this purpose.[18] Factor analysis is a very appropriate method to deal with the two kinds of data since the instrumental response in both cases is linear with respect to the concentration of the active constituents. The proposed procedure does not need the initial proposal of a chemical model nor the fulfillment of the law of mass action, and although it was initially designed as a complementary tool to least-squares curve fitting approaches in nonpolyelectrolyte systems, it can be used as an independent model-free and self-resolving tool in the study of the spectroscopic changes produced by the interactions of metal ions with macromolecular systems.[19]

Experimental Section

Reagents. Poly(adenylic acid) (potassium salt) (polyA, Sigma) is a synthetic water-soluble polymer with the empirical formula of the monomer ($C_{10}H_{11}O_6N_5PK$). All other reagents and materials were of analytical grade quality. All solutions were

• Abstract published in *Advance ACS Abstracts*, February 15, 1994.

0024-9297/94/2227-1729$04.50/0

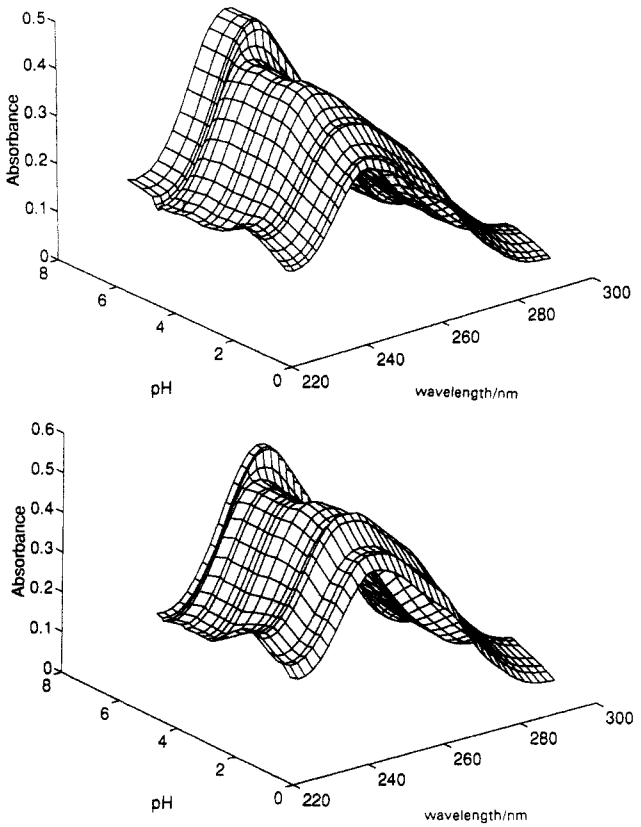

Figure 1. 3D plots of UV spectrometric titration data in the study of the systems poly(adenylic acid)–H^+ (a, top) and poly(adenylic acid)–Cu(II) ion (b, bottom).

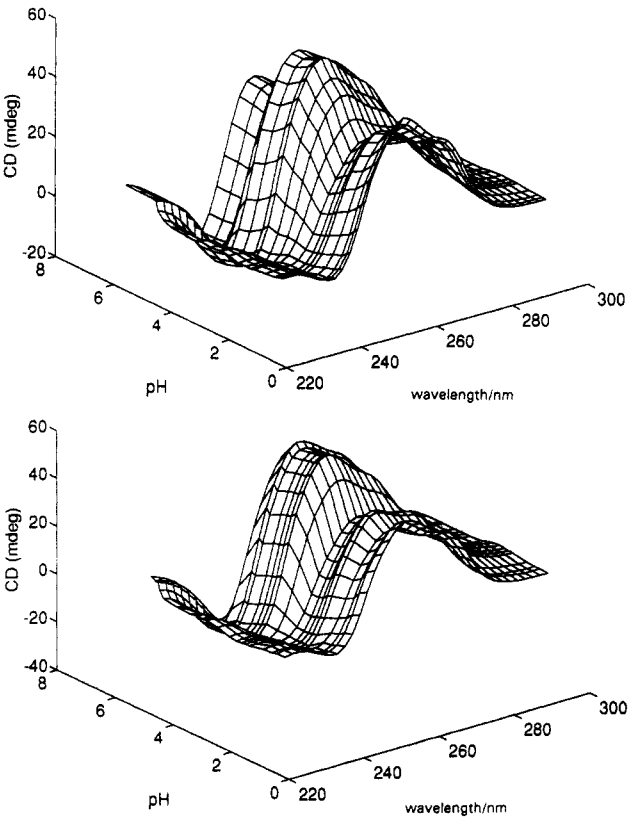

Figure 2. 3D plots of CD spectrometric titration data in the study of the systems poly(adenylic acid)–H^+ (a, top) and poly(adenylic acid)–Cu(II) ion (b, bottom).

made with deionized and CO_2-free water, adjusted to a ionic strength of 0.1 mol·L^{-1} with NaCl, and stored at 4 °C.

Apparatus. The UV–vis spectra were recorded at each pH using a Beckman DU-7 spectrophotometer controlled by a personal computer.[20] The circular dichroism spectra were recorded at the same pH values using a Jacso J-720 spectropolarimeter also controlled by a personal computer. In each case, the emf readings leading to pH values were made through an Orion 720 pH meter.

Procedure. The experimental data come from (a) four spectroscopic titrations of solutions containing poly(adenylic acid) at the concentrations (in monomer units) 9.34×10^{-5}, 1.08×10^{-4}, 1.15×10^{-4}, and 5.97×10^{-5} M and (b) four additional spectroscopic titrations of solutions containing poly(adenylic acid) and Cu(II) ion at the concentration ratios (polyA/Cu(II) ion) 9.16×10^{-5} M/9.07×10^{-5} M, 1.06×10^{-4} M/9.96×10^{-5} M, 1.14×10^{-4} M/9.90×10^{-5} M, and 6.05×10^{-5} M/4.95×10^{-5} M. All of solutions were 0.1 M in NaCl and were titrated by adding small amounts of acid or base to change the pH in the pH range 1.34–6.88. The experiments were thermostated at 25.0 ± 0.1 °C. At each pH, UV and/or CD spectra were recorded, with measurements taken every 0.2 nm within the wavelength range 230–300 nm, where polyA absorbs strongly. This wavelength range was used not only in the acid–base titrations but also in the complex-forming ones, because at the Cu(II) concentration needed (about 10^{-3} M) to follow the Cu(II) absorption band, precipitation phenomena were observed.

UV Experiments. The spectrometric titration experiments were conducted automatically under PC control: after each titrant addition and stabilization of the emf readings, a new spectrum was acquired and stored; then a new titrant addition was performed. A 3D plot of the whole set of spectra for the spectrometric titration of a Cu(II)-free poly(adenylic acid) solution (5.97×10^{-5} M in monomer units) is shown in Figure 1a; the plot for a solution containing poly(adenylic acid) (6.05×10^{-5} M in monomer units) and Cu(II) ion (4.95×10^{-5} M) is shown in Figure 1b.

CD Experiments. The titration cannot be done under PC control because of the limitations of the particular CD spectrometer used. Solutions of different pHs were prepared previously and kept in volumetric flasks until their serial measurement. Figure 2 shows 3D plots of CD spectra of a 5.97×10^{-5} M solution of poly(adenylic acid) (Figure 2a) and of a solution containing 6.05×10^{-5} M poly(adenylic acid) and 4.95×10^{-5} M Cu(II) ion (Figure 2b). For the joint treatment of UV and CD results (Figure 6), the intensity readings from CD experiments are divided by 100 to have similar intensities for both series.

Data Treatment

There are intrinsic difficulties in the analysis and interpretation of the data obtained from the study of acid–base and metal complexation properties of macromolecules if the study is performed using traditional least-squares approaches[21–23] because of the presence of polyelectrolytic, polyfunctional, and conformational effects.[3,4] In systems involving macromolecular ligands, mass action law can only be applied at the site level and equilibrium properties cannot be well defined for the overall molecule. In the present paper, a new model-free method is proposed for the study of acid–base and metal-complexing properties of macromolecules and, in particular, for the study of the acid–base and Cu(II)-complexing properties of poly(adenylic acid). The method, based on factor analysis and multivariate curve resolution techniques,[17,19,24,25] is as follows.

Data Arrangements and Linear Model (Figure 3). Factor analysis has been shown to be a powerful tool for solving multidimensional problems in chemistry.[13] The bilinear structure of the data matrices acquired in spectrometric titrations of macromolecules is very suitable for study using multivariate factor analysis techniques.

Suppose K titrations of samples at different concentrations of polynucleotide and metal are analyzed using UV and CD absorption spectrometric techniques. For each spectrometric titration using any one of these two techniques, a data matrix $\mathbf{D}_k$ is obtained. Under the assumption of linear response conditions, the data matrix

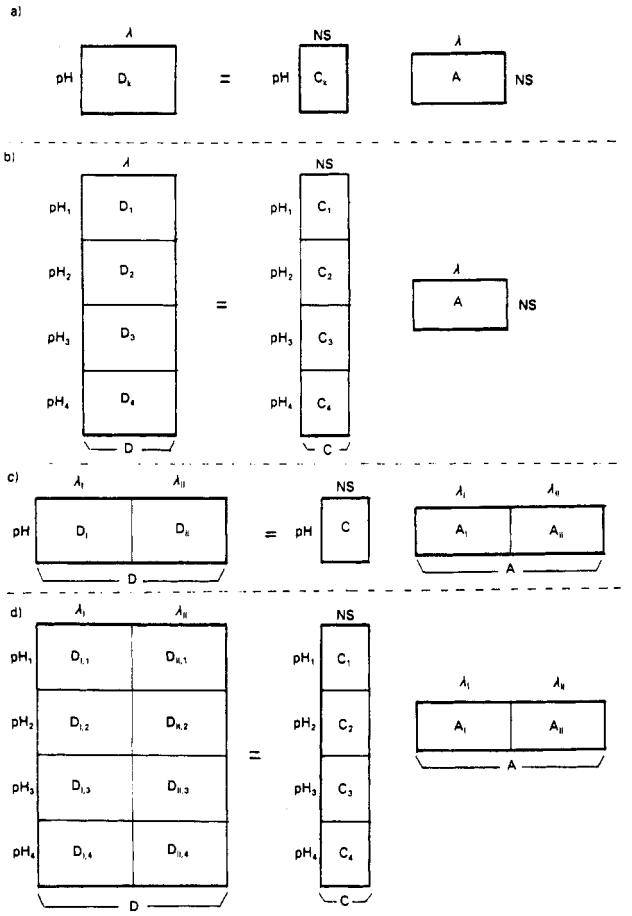

Figure 3. Schemes of the different data arrangements used in this work: (a) individual treatment of each data matrix; (b) simultaneous treatment of data matrices (acid–base and complexation) for each technique (UV or CD); (c) simultaneous treatment of data matrices (UV and CD) for each kind of titration (acid–base or complexation); (d) simultaneous treatment of all data matrices (UV, CD, acid–base, and complexation).

can be decomposed in the following way:

$$\mathbf{D}_k = \mathbf{C}_k\mathbf{A} + \mathbf{D}_{k0}, \quad k = 1, 2, ..., K \tag{1}$$

where $\mathbf{C}_k$ is the matrix of the concentration profiles of the chemical species (spectroscopically active) present during a particular titration k, $\mathbf{A}$ is the matrix of the unit or pure spectra of these species, and $\mathbf{D}_{k0}$ is the background or error absorption not caused by the considered species (Figure 3a).

When several titrations are performed over different samples using the same spectrometric technique, the same set of wavelengths being used for the measurements, the numerical analysis of the experimental data can be extended over all these titrations simultaneously. Each titration gives a data matrix $\mathbf{D}_k$. A new augmented data matrix $\mathbf{D}$ is obtained by setting the different data matrices $\mathbf{D}_k$ one on top of each other, keeping the column space (wavelength space) in common.

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_1 \\ \mathbf{D}_2 \\ \cdots \\ \cdots \\ \cdots \\ \mathbf{D}_K \end{bmatrix} = \begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \\ \cdots \\ \\ \cdots \\ \mathbf{C}_K \end{bmatrix} \mathbf{A} + \mathbf{D}_0 \tag{2}$$

$$\mathbf{D} = \mathbf{CA} + \mathbf{D}_0 \tag{3}$$

The new augmented data matrix $\mathbf{D}$ (Figure 3b) has a number of rows equal to the total number of acquired spectra in the different spectrometric titrations (measured pH values) and it is the product of two matrices, an augmented concentration matrix times the unit spectra matrix. The augmented concentration matrix will contain the different submatrices $\mathbf{C}_k$ explaining the concentration changes of the species present in each of the data matrices $\mathbf{D}_k$ analyzed. The shapes and intensities of the concentration profiles (species distribution) change from titration to titration and should be described by a "skinny" $\mathbf{C}$ matrix with the number of rows equal to the total number of rows of $\mathbf{D}$ (all the spectra acquired, all pH values measured) and with the number of columns equal to the number of species present in the titrations. $\mathbf{A}$ is a reduced size matrix of the unit pure spectra of the absorbing components, which are common (equal) in the different titrations, with the number of rows equal to the number of species present in the titrations and the number of columns equal to the number of spectrometric channels or wavelengths (which are common for all the titrations). $\mathbf{D}_0$ is the augmented background or noise absorption matrix composed by the different $\mathbf{D}_{k0}$ noise-background matrices at every titration k.

This study can be adapted easily to the simultaneous analysis of spectrometric titrations where some of the species are common and others not. This is the case for titrations of the macromolecule in the presence and/or absence of a metal ion. In the analysis of the individual titrations the number and nature (unit spectrum) of each species are initially determined. For those titrations that are known not to have a certain species, a zero is put in the corresponding column of the $\mathbf{C}$ matrix. When, at every point of the titration, spectra are measured with two different spectrometric techniques (UV and CD), two different data matrices are recorded for each titration. These two matrices have the rows (pH values) in common but the columns (spectrometric channels) are different. In this case, a new augmented data arrangement can be set keeping the row space in common and expanding the column space:

$$\mathbf{D} = [\mathbf{D}_I, \mathbf{D}_{II}] = \mathbf{C}[\mathbf{A}_I, \mathbf{A}_{II}] = \mathbf{CA} \tag{4}$$

Now the concentration $\mathbf{C}$ is a small size data matrix having the concentration profiles of the absorbing species which are common to both matrices $\mathbf{D}_I$ and $\mathbf{D}_{II}$. $\mathbf{A}$ is the new augmented data matrix containing the new unit or pure species spectra spanning the two spectrometric methods. Each of these spectra is defined as the composite of the unit spectrum of the species using the two spectrometric detection methods (Figure 3c).

An even more complex data arrangement is obtained when the different data matrices obtained in the different titrations using the two spectrometric techniques are set up together. That is

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_{I,1}\mathbf{D}_{II,1} \\ \mathbf{D}_{I,2}\mathbf{D}_{II,2} \\ \cdots \\ \cdots \\ \cdots \\ \mathbf{D}_{I,K}\mathbf{D}_{II,k} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \\ \cdots \\ \\ \cdots \\ \mathbf{C}_K \end{bmatrix} [\mathbf{A}_I\mathbf{A}_{II}] + \mathbf{D}_0 =$$

$$\mathbf{CA} + \mathbf{D}_0 \tag{5}$$

The gross new augmented data matrix will now be the product of a "skinny" concentration matrix $\mathbf{C}$ (containing the submatrices $\mathbf{C}_k$ with the concentration of the species in each titration) and a "fat" species spectral matrix $\mathbf{A}$

(which has the augmented unit spectra of each of the species using the two spectrometric methods). The amount of information about every species is now much higher, and consequently the resolving power of the method is considerably increased. In all cases, for a certain spectrometric method, the unit spectra of the common species in the different titrations are equal (see below for the case of noncommon components). The latter assumption is true whenever physical external conditions such as temperature and solvent composition are held constant (Figure 3d).

Factor Analysis. The data matrix **D**, augmented or not, is decomposed using factor analysis[13]

$$\mathbf{D} = \mathbf{U}\mathbf{V}^{\mathrm{T}} + \mathbf{E} = \mathbf{D}^* + \mathbf{E} \tag{6}$$

where now **U** and $\mathbf{V}^{\mathrm{T}}$ are respectively the score and loading matrices of **D** for the preselected number of components, **E** is the residual error matrix containing the variance not explained by **U** and $\mathbf{V}^{\mathrm{T}}$, and $\mathbf{D}^*$ is the reproduced data matrix based on **U** and $\mathbf{V}^{\mathrm{T}}$. Under the assumption of linearity, when the correct number of components is chosen, the residual error matrix **E** is close to the noise or experimental error. An intial estimation of the noise and background absorption level can be obtained from the spectral regions where considered species do not absorb.

The transformation of the **U** and $\mathbf{V}^{\mathrm{T}}$ matrices in the **C** and **A** matrices cannot be achieved directly if no additional information is provided. As **C** and **A** are unknown, the equation

$$\mathbf{D}^* = \mathbf{U}\mathbf{V}^{\mathrm{T}} = \mathbf{U}\mathbf{T}\mathbf{T}^{-1}\mathbf{V}^{\mathrm{T}} = \mathbf{C}\mathbf{A} \tag{7}$$

has an infinite number of solutions for any arbitrary transformation–rotation matrix **T**. There is an intrinsic ambiguity in the factor analysis solutions to solve for **C** and **A** if no more information is provided. The general task of the curve resolution methods and in particular of the proposed method is to constrain the number of possible solutions which eventually can give the physically and chemically meaningful actual solutions for **C** and **A**.

The number of components to be considered is that needed to reproduce the original data matrix within the experimental error. This number of principal factors or components should be intrinsically related to the number of species in the system. Several methods based on factor analysis have been proposed for the determination of the number of components of a data matrix, and they can be used in this case for the determination of the number of species in equilibrium.[13,26-28] The simpler approach is to look for the number of eigenvalues that explain most of the data variance up to a certain residual variance associated with background and noise contributions. This residual contribution is previously estimated from spectral regions with no absorption from the investigated species. Owing to experimental uncertainties and baseline contributions, especially in the CD method, the residual variance can be quite high, especially at lower wavelength values, approaching 2 or 3% of the maximum intensity of the CD bands.

The singular value analysis of the different data matrices arranged in augmented form (described previously) provides important information concerning the correspondence of species among titrations and among spectrometric techniques. If some new species appears in one titration and is not present in the others, the rank of the augmented data matrix should increase with respect to the rank of individual matrices. The number of components initially estimated in this way is checked later using evolving factor analysis and during the alternating least-squares regression optimization (see below), looking always for those solutions which better fit the data and which have a physical meaning, i.e., yield recovered profiles which can be accepted as real concentration and spectra profiles (see below).

There are two classes of ambiguities associated with factor analysis methods: the intensity and the rotational ambiguities. Because of the intensity ambiguity, the estimated concentrations and spectra will be scaled by some unknown factor. This is not a serious problem in qualitation (spectral identification, fingerprinting) but it is a serious problem in quantitation. This intensity ambiguity can be solved when data from several samples are treated simultaneously.[29]

Rotational ambiguity is also inherent to factor analysis decompositions[30] when there are two or more linearly independent overlapped components. The estimated unit spectrum for any of the species will be an unknown linear combination of the unit spectra of the true species.[31] However, for those spectra measured where only one component is present, selectivity is achieved and there is no rotational ambiguity.[32] This means that in this particular situation, the principal component analysis[13,25] solution yields the correct true (in shape) unit spectrum.

The data acquired in spectrometric titrations are ordered data; the spectra are obtained in a determined order, as pH, time, concentration of a reagent, or any other parameter is systematically changed. Eigenanalysis can be performed for all the submatrices of one whole data matrix acquired in one spectrometric titration; for instance, if the titration is performed from acidic to basic pH range, all the submatrices originated when pH is increased can be successively analyzed (forward analysis). The same can be done with all the submatrices originated when pH is decreased (backward analysis). The number and magnitude of the emerging eigenvalues are related with the evolution of the real contributions and with how they affect the data variance structure; if the concentration of a certain species increases, the same will happen with the corresponding eigenvalue. With this idea in mind, evolving factor analysis[33,34] and submatrix analysis[35] were developed. From the evolving factor analysis plots,[14-17,19] the number of significant factors is recognized and a first estimation of the changes in concentrations of the significant species, i.e., the distribution of species, is obtained. The plot is called the abstract distribution plot and constitutes the starting point for obtaining the distribution of real concentrations of the species in the system.[14,15]

Constrained Alternating Least-Squares Analysis of Individual Spectrometric Titrations. Once the number of species present in each titration is known and an initial estimation of the concentration distributions (matrix) **C** is obtained from evolving factor analysis, an alternating least-squares regression procedure is started, consisting of two parts:

(a) In the first part an estimation of the unknown species spectra is performed from the linear equation (derived from the generalized Beer equation) by least squares

$$\mathbf{D} = \mathbf{C}\mathbf{A} \tag{8}$$

from which the matrix of unit spectra **A** is estimated by least squares

$$\mathbf{A} = \mathbf{C}^{+}\mathbf{D}^* \tag{9}$$

where $\mathbf{D}^*$ is the reproduced data matrix for the considered number of factors and $\mathbf{C}^{+}$ is the pseudoinverse of **C** (which

is estimated from $(C^TC)^{-1}C^T$). The matrix A gives the current least-squares estimation of the unit spectra. The UV absorptivities must be positive, whereas CD unit absorptivities can be positive or negative. This constraint is applied accordingly during the least-squares optimization.

(b) In a second stage, a new estimation of the concentration distributions is obtained by least squares using the equation

$$C = D^*A^+ \quad (10)$$

where now A^+ is the pseudoinverse of the A matrix (estimated from $A^T(AA^T)^{-1}$). In this case the concentrations derived from the equation are not only constrained to be positive but also to give unimodal profiles (a departure of the unimodal condition can happen in very special cases,[36] but it is quite rare in general). As the total concentration of the absorbing species is known, normalization of the concentration profiles by closure is also applied at this stage.

(c) Steps a and b are repeated until the data matrix D^* is well explained within experimental error. Convergence is achieved usually in a few iterations of the alternating least-squares regression method.

Concentration profiles and unit spectra obtained in the individual analysis of every spectrometric titration can be somewhat discordant between titrations because of the subsisting presence of intensity and rotational ambiguities, which are not solved in the individual analysis of each spectrometric titration.

Constrained Alternating Least-Squares Simultaneous Analysis of Several Spectrometric Titrations. The numerical solutions obtained in the treatment of individual titrations can be highly constrained in the simultaneous least-squares analysis of the different data matrices obtained in different spectroscopic titrations. Steps a–c of the previous constrained alternating least-squares procedure are applied to the new augmented data matrix arranged as proposed in Figure 3. For the analysis using simultaneously both UV and CD data matrices (matrix arrangement 3d), a set of species spectra and concentration profiles are obtained using eqs 9 and 10, where now D^* and C are the augmented matrices described in eqs 2, 4, and 5. The set of possible numerical solutions of these equations is now more constrained than in the analysis of individual spectroscopic titrations. The analysis of the same titration using two independent spectrometric methods enhances greatly the resolving power of the method.

In the simultaneous analysis of different titrations, the right correspondence between species is found from detailed analysis of single titration results. Species in different titrations with the same unit spectra correspond to the same species in the simultaneous analysis. This assignment is not difficult because the formation of the different species in the different titrations is stepwise and takes place in a reproducible order. For those species not found in a particular titration, the concentration values in the appropriate column of the concentration matrix C are constrained to be zero. Convergence of the alternating least-squares optimization using the constraints previously described is usually fast and gives fitting errors always below 3% of the higher intensity input spectral values.

The proposed method for the simultaneous analysis of several data matrices differs from other methods recently proposed by other authors.[37,38] (1) The method handles data matrices with linear correlation in one or in two of the orders of every analyzed data matrix (the unit spectral profiles and/or the unit concentration profiles of the same species are equal in the different data matrices); it does not require that both orders were totally correlated; i.e., it does not require that the data have a completely second-order structure.[38] (2) The method proposed here allows the simultaneous analysis of more than two data matrices, in constrast to other proposed methods based on eigenvalue–eigenvector decompositions such as GRAM[39] or in Procrustes rotation.[37] (3) The method is a least-squares iterative method, with its advantages (real solutions) and disadvantages (slow convergence and local minima).

The experimental data analyzed in the present work do not have a completely second-order structure and cannot be studied by other three-way data analysis methods such as GRAM[39] or by Procrustes-based methods such as DATAN.[37] Because of the lack of a complete second-order data structure, true solutions can only be obtained under a constrained least-squares optimization method such as that proposed in the present work.

Equilibrium Information. From the species concentrations finally obtained in the alternating least-squares procedure of SPFAC, it is possible in principle to determine the stability constants or better the concentration quotients for the involved equilibria. The possibility of obtaining equilibrium information, without the previous use of the mass action law, becomes especially interesting in the case of the study of equilibria with macromolecular ligands, where the mass action law is only valid locally for each bonding site, and not simultaneously to all the bonding sites of the same type in the macromolecule.[3]

The different factor analysis and regression techniques described so far have been implemented in a single FORTRAN computer program, SPFAC, and in a small set of MATLAB functions.[40] The aim in the development of the computer procedures is to provide a simple way for the self-curve resolution of concentration profiles and species spectra of the components present in spectrometric ordered data systems using factor analysis techniques.

Results

Singular value analysis of a single data matrix using UV detection showed that three absorbing species are present in the H^+–polyA system. When CD detection was used, another species was detected for the same system, suggesting a higher resolution power of the CD method to detect differences in conformation which are difficult to detect by UV absorption. In the analyses of data matrices where Cu(II) was also present an additional species was detected. Singular value analyses of the augmented data matrices built up from single titration data matrices confirmed that the total number of species does not increase when more matrices are added. When data matrices coming from titrations where Cu(II) was present were put together with data matrices where Cu(II) was not present, the number of species detected was still equal to the maximum number of species detected in the individual analysis of titrations having Cu(II) ion. There is no rank increase nor decrease when several titrations are analyzed together, and the system behaves following a linear model as expected. When only one titration is analyzed using both detection techniques, the number of species (singular values) is equal to the number of species detected by circular dichroism. Finally, when the singular value analysis is performed over the gross augmented data matrix built up using titrations with and without copper, using UV and CD detection techniques (Figure 4), the total number of species present is deduced equal to five. Five species are proposed to be present in the titrations

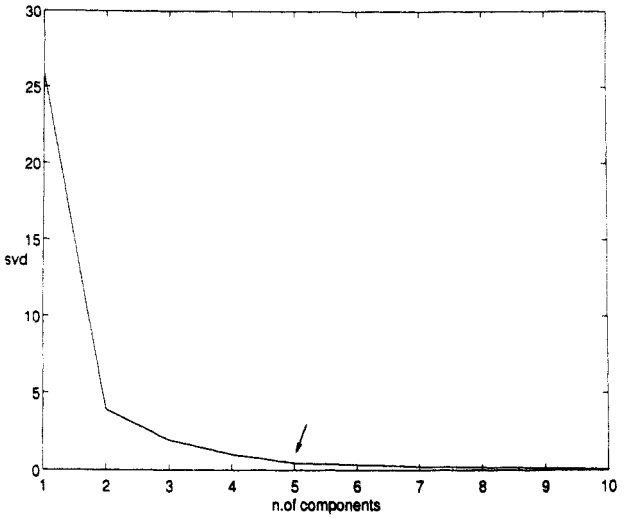

Figure 4. Singular value plot of the augmented matrix with all titrations (UV, CD, acid–base, and complexation) (see arrangement 3d).

which contain Cu(II) ion, whereas for those titrations without Cu(II) ion, a maximum of four species is proposed. These results are checked and eventually confirmed in the next steps of the data treatment.

In Figure 5, the evolving factor analysis plot of an acid–base UV titration of a polyA solution (Figure 5a) and the evolving factor analysis plot of a CD titration of a Cu(II)–polyA solution (Figure 5b) are shown. These are two examples of the results obtained by evolving factor analysis of individual data matrices. Similar plots are obtained for other data matrices. The plots show how complex the systems under study are and the strong overlap in the formation of the different species. Rank one spectral selective regions are very rare although they seem to be present at the beginning and the end of some of the titrations. Evolving factor analysis plots provide the initial estimates[33-36] of the species concentrations in each titration to be used in the alternating least-squares optimization.

Results obtained in the individual analysis of the different spectrometric titrations using the constrained alternating least-squares procedure described before become unreliable because of the lack of selectivity in both orders (unit spectra and concentration profiles for the different species are strongly overlapped). In particular, the results obtained from the UV titrations are less reliable than the results obtained from the CD titrations. When the analysis is performed over the augmented data matrices obtained by the grouping of the titrations with the same technique and type, the results improve since the number of possible solutions of the linear equation system decrease. Building up of the augmented data matrix containing all the titrations (both UV and CD titrations), including acid–base and complexation titrations, is very useful to discover the right species correspondence, i.e., which species are common and which are not among titrations in the presence or in the absence of Cu(II) ion. In particular, the best assignment is that where the three first species formed at more acidic media are common for both types of titrations, a fourth species is only present in the case of the presence of Cu(II) ion (probably corresponding to a metal complex of the unprotonated macromolecule), and a last species also common to all solutions is the deprotonated polynucleotide. This indicates that Cu(II) ions interact only weakly with unprotonated poly(adenylic acid) in a very similar way that Cu(II) ions interact with related compounds such as adenosine and 9-methyladenine.[41] Finally, when a unique data matrix formed with all titrations using

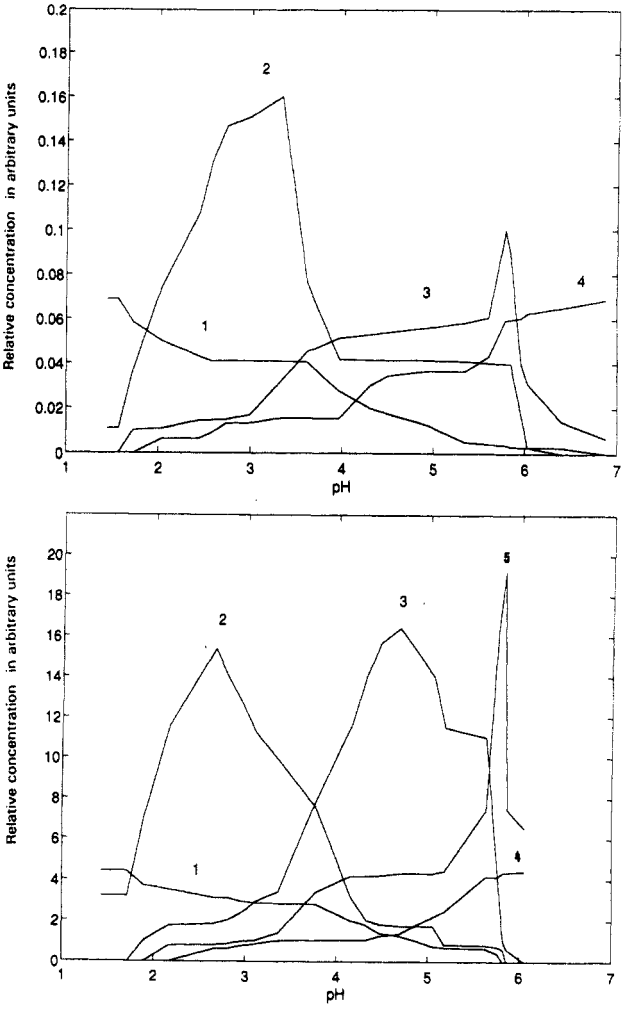

Figure 5. (a, Top) Evolving factor analysis plot for an acid–base UV spectrometric titration of a poly(adenylic acid) solution. (b, Bottom) Evolving factor analysis plot for a CD titration of a Cu(II)–polyA solution. In both plots, lines describe the initial estimation of the species distribution. Numbers identify the different species (see Figure 6).

both spectrometric techniques (acid–base, complexation, UV, and CD) (see Figure 3d) is analyzed by the proposed constrained alternating least-squares analysis, the number of components (or species) is defined precisely and the residual fitting error is similar to that expected as experimental error (around 2% of the maximum experimental signal intensity). The mixed unit species spectra (UV and CD) obtained in this treatment are shown in Figure 6, and the species distribution is shown in Figure 7. The titration chosen for this joint analysis are those of solutions containing 5.97×10^{-5} and 1.15×10^{-4} M poly(adenylic acid) and this acid plus Cu(II) ion at the concentration ratios 1.14×10^{-4} M/9.90×10^{-5} M and 6.05 $\times 10^{-5}$ M/4.95×10^{-5} M, respectively. These spectra are formed by two parts: one, on the left, is the UV spectrum, while the part on the right is the CD spectrum of each species.

Four spectroscopically different species (species 1–4) are found for the poly(adenylic acid) at pH values within the range 1.5–6.8 in the acid–base titrations. Three of them (species 1–3) are probably due to different conformations of the protonated form of polyA, and the other (species 4) to the unprotonated form of polyA. The midpoint in the transition from protonated to unprotonated forms takes place around pH 5.8. These results are in agreement with previous reports about different conformations of polyA in acidic medium using other experimental approaches.[12,16,17] The single new species detected

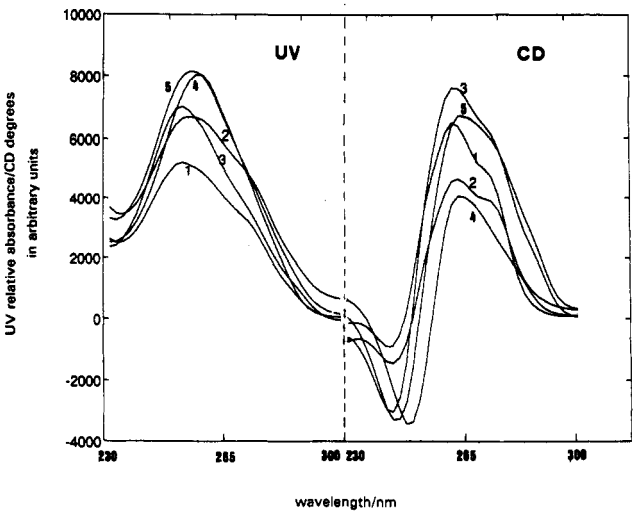

Figure 6. Species spectra resolved using the proposed procedure in the simultaneous analysis of all experimental data. The left part corresponds to the UV spectra, and the right part corresponds to the CD spectra of each species. Curves 1, 2, and 3 are the spectra of different conformations of protonated poly(adenylic acid), curve 4 is the spectrum of deprotonated poly(adenylic acid), and curve 5 is the spectrum of the Cu(II) complex of deprotonated poly(adenylic acid).

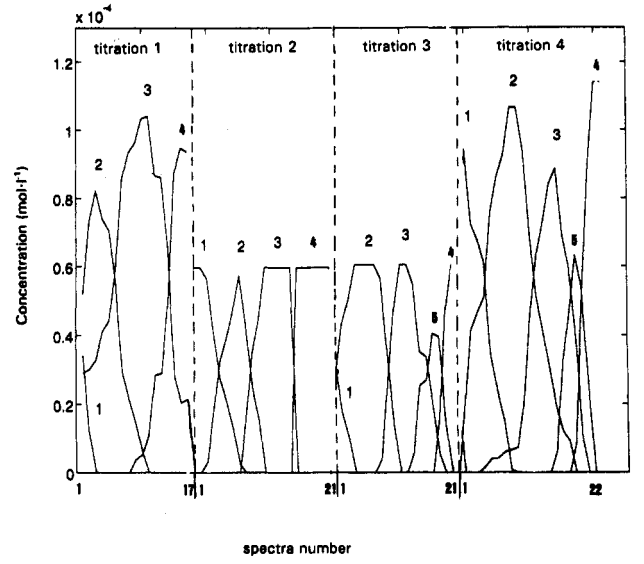

Figure 7. Abstract distribution plot of species (concentration vs pH) obtained from the simultaneous analysis of UV, CD, acid–base, and complexation titrations. The numbers on the curves identify the same species as in Figure 6. The different titrations analyzed are separated by dashed lines. Titration 1: 17 spectra from pH 1.96 to pH 5.92 (without Cu(II) ion). Titration 2: 21 spectra from pH 1.45 to pH 6.88 (without Cu(II) ion). Titration 3: 21 spectra from pH 1.43 to pH 6.05 (with Cu(II) ion). Titration 4: 22 spectra from pH 1.36 to pH 6.18 (with Cu(II) ion).

in the presence of Cu(II) ion (species 5) is obviously a metal complex of the macromolecule.

While the UV spectra of species 1 and 2 are very similar, the CD spectra of these species are more different. This is the reason why the numerical resolution in the treatment of CD titrations is better than that obtained in the UV analysis. In addition, the similarity between both UV unit spectra makes more difficult the determination of the number of different species. Simultaneous treatment of UV and CD spectra improves very much the resolution of the spectral characteristics of the different species. The advantage of CD over UV absorption to solve conformational problems is even more evident in the study of complexation with Cu(II) ion because the spectrum of the complex and that of the unprotonated form of the ligand are much more different in CD than in UV absorption.

The equilibrium reaction considered for the protonation of polyA is

$$H^+ + \text{polyA} \rightleftarrows \text{HpolyA}^+ \qquad (11)$$

where polyA is the unprotonated form of poly(adenylic acid), HpolyA^+ is its protonated form, and H^+ is the hydrogen ion.

The apparent constant of protonation of polyA, $(K_{app})_{prot} = [\text{HpolyA}^+]/([H^+][\text{polyA}])$, is calculated from the concentrations of the protonated and unprotonated forms of polyA (species 3 and 4, respectively) obtained from the constrained alternating least-squares procedure. The hydrogen ion concentration is measured experimentally.

The degree of protonation (α_{prot}) is obtained from the ratio of the concentration of protonated sites of polyA to the total concentration of basic sites: $\alpha_{prot} = [\text{HpolyA}^+]/C_T = [\text{HpolyA}^+]/([\text{polyA}] + [\text{HpolyA}^+])$.

The equilibrium reaction considered for the complexation between Cu(II) ion and polyA is

$$Cu^{2+} + \text{polyA} \rightleftarrows (\text{Cu-polyA})^{2+} \qquad (12)$$

where polyA is the unprotonated form of the ligand, Cu^{2+} is the free Cu(II) ion, and $(\text{Cu-polyA})^{2+}$ is the complex formed. The apparent complexation constant ($(K_{app})_{comp}$) is calculated from the equation

$$(K_{app})_{comp} = \frac{[(\text{Cu-polyA})^{2+}]}{[Cu^{2+}][\text{polyA}]} = \frac{[(\text{Cu-polyA})^{2+}]}{([Cu^{2+}_T] - [(\text{Cu-polyA})^{2+}])[\text{polyA}]} \qquad (13)$$

where the concentrations of the complex form and the unprotonated form of polyA (species 5 and 4) are obtained from the constrained alternating least-squares procedure and the concentration of free Cu(II) ion is calculated from the difference between the total concentration of Cu(II) ion and the concentration of complexed Cu(II) ion (species 5).

The degree of complexation (α_{comp}) is obtained from the equation $\alpha_{comp} = [(\text{Cu-polyA})^{2+}]/[Cu_T]$, where $[(\text{Cu-polyA})^{2+}]$ is the concentration of complex form and $[Cu_T]$ is the total concentration of Cu(II) ion.

The plot of the logarithm of the apparent protonation constant ($(\log K_{app})_{prot}$) versus the degree of protonation (α_{prot}) is shown in Figure 8, and the plot of the logarithm of the apparent complexation constant ($\log K_{app})_{comp}$ versus the degree of complexation (α_{comp}) is shown in Figure 9. The sets of data of both figures have been subjected to least-squares fitting and to some statistical tests to confirm that linear dependences between these values exist (Table 1).

The value of the logarithm of the protonation constant extrapolated to $\alpha_{prot} = 0$, i.e., when the molecule of polyA is completely unprotonated, is 4.87. The value of the logarithm of the complexation constant extrapolated to $\alpha_{comp} = 0$ is 2.94. These values are both higher than the stability constants for the protonation and Cu(II) complexation of adenine (4.02 and 2.25, respectively).[41]

The linear dependence observed for the protonation constant, however, is contrary to what would be expected if the presence of polyelectrolytic effects were considered. During protonation of poly(adenylic acid) an increase of positive charge is produced over the macromolecule surface, which would, in principle, oppose the continuation of the protonation process, whereas what is observed is

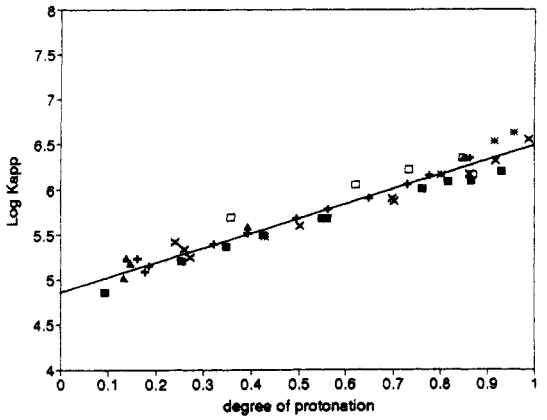

Figure 8. Plot of the variation of the logarithm of the protonation constant of poly(adenylic acid) with respect to the degree of protonation for six different titrations (UV and CD). The results were obtained from the equilibrium concentrations of the third and fourth species in the species distribution plot in Figure 7. The (×, ▲) symbols represent the values obtained in the CD experiments, and the (■, +, ∗, □) symbols represent the values obtained in the UV experiments.

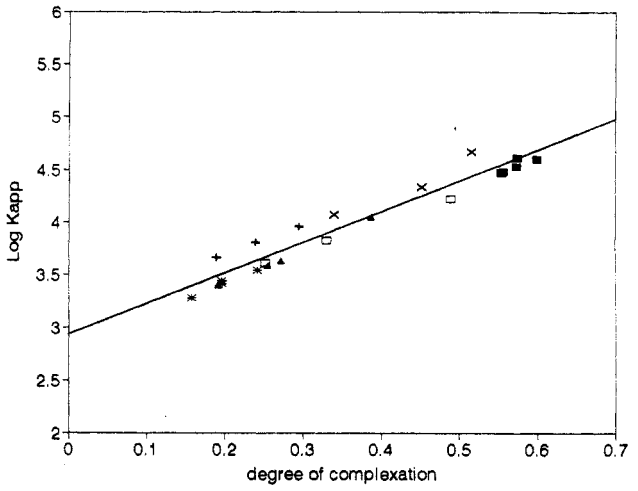

Figure 9. Plot of the variation of the logarithm of the stability constant for the complex of poly(adenylic acid) with Cu(II) ion with respect to the degree of complexation for six different titrations (UV and CD). The results were obtained from the equilibrium concentrations of the fifth and fourth species in the species distribution plot in Figure 7. The (□, ×, ▲) symbols represent the values obtained in the CD experiments, and the (■, +, ∗) symbols represent the values obtained in the UV experiments.

Table 1. Linear Dependence of Calculated log K_{app} Values with Respect to α Values

metal	N^a	NT[b]	intercept (sd)[c]	slope (sd)[c]	sd residuals[d]	% of variance explained by regression	F^e
H^+	44	6	4.87 (0.04)	1.61 (0.06)	0.11	94.43	713
Cu^{2+}	24	6	2.94 (0.06)	2.89 (0.16)	0.12	93.51	317

[a] N is the total number of points (log K_{app}/α) subjected to the analysis. [b] NT is the number of titrations considered in the calculations. [c] The values of the standard deviations obtained by least-squares data fitting are given in parentheses. [d] Value of the standard deviation of the residuals. [e] F statistical test; F_{crit} at the significance level of 5% is 4.0 for the acid–base experiments and 4.3 for the complexation experiments.

just the opposite: the protonation of poly(adenylic acid) is enhanced when the positive charge increases over the macromolecule, and the protonation constant consequently becomes higher. A possible explanation of this phenomenon, deduced from experimental data, is that the stabilization of protonated nitrogen in the adenine base sites occurs because of base pairing through hydrogen bonding. This is probably related with a transition from single-stranded to double-stranded polyadenylic chains or some other conformational changes occurring at acidic pH values described in the literature.[10-12] While unprotonated poly(adenylic acid) has a random coil, single-stranded structure, protonated poly(adenylic acid) is present at least in three different polystranded conformations, all of them highly stabilized by hydrogen bonding between base pairs. The hydrogen bonding involves those nitrogen base sites which are also involved in the acid–base equilibria observed around pH 5. Consequently, the changes in speciation observed from the spectrometric titrations performed are a consequence of conformational changes of the protonated poly(adenylic acid), together with the acid-base equilibrium of the macromolecule and its Cu(II) ion binding equilibrium.

Similar linear behavior is observed too for the apparent complexation constant. The increase of the value of this constant with increasing degrees of complexation should be assigned to bridging of adenine residues by complexed Cu(II) ions.

Conclusions

The proposed method is proved to be a very useful tool for the study of the acid–base and metal ion binding properties of macromolecules in solution. It provides a way to corroborate speciation and conformational changes derived from other completely different and independent approaches in the field. The simultaneous treatment of several titrations using different spectrometric techniques (UV, CD, etc.) has been proved to increase considerably the resolving power and reliability of the data treatment.

In particular, the proposed method allows the deduction of the number of independent macromolecular species present along a spectrometric acid–base titration of solutions containing poly(adenylic acid) in the absence or the presence of Cu(II) ion. Four spectrometrically different species are detected for poly(adenylic acid) between pH 1.5 and pH 6.8. Three of them are different conformational forms of protonated polyA and the other is unprotonated polyA. One new species is detected in the presence of Cu(II) ion, a copper complex of the macromolecule. In addition, the proposed method allows the estimation of the concentration distributions and the unit spectra of the species formed along the titration at each pH value. Stabilization by hydrogen bonding between base pairs is postulated to explain the stabilization observed during the protonation of poly(adenylic acid); similarly, copper(II) ion bridging stabilizes the complex formed.

Acknowledgment. This research has been supported by the CYCIT, Spain (Grant No. PB90-0821). The circular dichroism spectra were obtained with the cooperation of the Technical–Scientific Services of the University of Barcelona.

References and Notes

(1) *Circulation of Metals in the Environment. Metal Ions in Biological Systems*; Sigel, Helmut, Ed.; Marcel Dekker: New York, 1984; Vol. 18.
(2) *Methods Involving Metal Ions and Complexes in Clinical Chemistry. Metal Ions in Biological Systems*; Sigel, Helmut, Ed.; Marcel Dekker: New York, 1983; Vol. 16.
(3) Buffle, J. *Complexation Reactions in Aquatic Systems. An Analytical Approach*; Ellis Horwood Sr., John Wiley and Sons, New York, 1988.
(4) Marinsky, J. A. *Coord. Chem. Rev.* **1976**, *19*, 125.
(5) Barbucci, R.; Casolaro, M.; Nocentini, M.; Corezzi, S.; Ferruti, P.; Barone, V. *Macromolecules* **1986**, *19*, 37.

(6) Eisenberg, H.; Felsenfeld, G. J. *J. Mol. Biol.* **1967**, *30*, 17.
(7) Sorokin, V. A.; Blagoi, Y. P.; Valeev, V. A.; Antonova, O. A. *J. Inorg. Biochem.* **1985**, *24*, 97.
(8) Berger, N. A.; Eichhorn, G. L. *Biochemistry* **1971**, *10*, 1857.
(9) Eichhorn, G. L.; Rifkind, F. M.; Shin, Y. A. *Adv. Chem. Ser.* **1977**, *162*, 305.
(10) Janik, B.; Sommer, R. G.; Bobst, A. M. *Biochim. Biophys. Acta* **1972**, *281*, 152.
(11) Adler, A. J.; Grossman, L.; Fasman, G. D. *Biochemistry* **1969**, *8*, 3846.
(12) Scovell, W. M. *Biopolymers* **1978**, *17*, 969.
(13) Malinowski, E. R. *Factor Analysis in Chemistry*, 2nd ed.; Wiley: New York, 1991.
(14) Tauler, R.; Casassas, E. *J. Chemomet.* **1988**, *3*, 151.
(15) Tauler, R.; Casassas, E. *Anal. Chim. Acta* **1989**, *223*, 257.
(16) Tauler, R.; Casassas, E.; Izquierdo-Ridorsa, A. *Anal. Chim. Acta* **1991**, *248*, 447.
(17) Tauler, R.; Casassas, E. *Analusis* **1992**, *20*, 255.
(18) Sheardy, R. D. *Spectrosc. Int.* **1991**, *3*, 11.
(19) Tauler, R.; Casassas, E. *Chemomet. Intell. Lab. Syst.* **1992**, *14*, 305.
(20) Beltrán, J. L.; Centeno, G.; Izquierdo, A.; Prat, M. D. *Talanta* **1992**, *39*, 981.
(21) Gans, P. *Coord. Chem. Rev.* **1976**, *19*, 99.
(22) Leggett, D. J., Ed.; *Computational Methods for the Determination of Formation Constants*; Plenum: New York, 1985.
(23) Gazier, F. *Coord. Chem. Rev.* **1979**, *27*, 195.
(24) Tauler, R.; Flemminng, S.; Kowalski, B. R. *Anal. Chem.* **1993**, *65*, 2040.
(25) Tauler, R.; Barceló, D. *Trends Anal. Chem.* **1993**, *12*, 319.
(26) Malinowski, E. R. *Anal. Chem.* **1977**, *49*, 612.
(27) Wold, S. *Technometrics* **1978**, *20*, 397.
(28) Forina, M.; Lanteri, S.; Armanino, C.; Leardi, R. *Miscellània Enric Casassas, Pub. Univ. Autònoma de Barcelona, Bellaterra* **1991**, 151.
(29) Tauler, R.; Izquierdo-Ridorsa, A.; Casassas, E. *Chemomet. Intell. Lab. Syst.* **1993**, *18*, 293.
(30) Lawton, W. H.; Sylvestre, E. A. *Technometrics* **1971**, *13*, 617.
(31) Spjotvoll, E.; Martens, H.; Volden, R. *Technometrics* **1982**, *3*, 173.
(32) Kvalheim, O. M.; Liang, Y. Z. *Anal. Chem.* **1992**, *64*, 2735.
(33) Gampp, H.; Maeder, M.; Meyer, Ch. J.; Zuberbühler, A. D. *Talanta* **1985**, *32*, 1133.
(34) Gampp, H.; Maeder, M.; Meyer, Ch. J.; Zuberbühler, A. D. *Talanta* **1986**, *33*, 943.
(35) Cartwright, H. *J. Chemomet.* **1986**, *1*, 111.
(36) Tauler, R.; Casassas, E. *Anal. Chim. Acta* **1988**, *206*, 189.
(37) Scarminio, I.; Kubista, M. *Anal. Chem.* **1993**, *65*, 409.
(38) Sánchez, E.; Kowalski, B. R. *J. Chemomet.* **1988**, *2*, 265.
(39) Sánchez, E.; Kowalski, B. R. *Anal. Chem.* **1986**, *58*, 496.
(40) *Matlab*; The Math Works, Inc.: Natick, MA.
(41) Tauler, R.; Rainer, M. J. A.; Rode, B. M. *Inorg. Chim. Acta* **1986**, *123*, 75.