“Analytical resolution of parallel and antiparallel oligonucleotide triple helices formation and melting processes by means of Multivariate Curve Resolution” Jaumot, J., Eritja, R., Aviñó, A., Tauler, R., Gargallo, R. J. *Biomol. Struct. Dyn.*, 21, 267-278 (2003).

Analytical resolution of parallel and antiparallel oligonucleotide triple helices formation and melting processes by means of Multivariate Curve Resolution

J. Jaumot^a, R. Eritja^b, A. Aviñó^b, R. Tauler^a, R. Gargallo^{a,*}

^a *Departament de Química Analítica, Universitat de Barcelona, Martí i Franqués 1 –11, 08028*

Barcelona, Spain

^b *Institut de Biologia Molecular de Barcelona, CSIC, Jordi Girona 18 - 26, 08034 Barcelona, Spain*

* Author to whom correspondence should be addressed

Fax: +34-934021233

E-mail: raimon@apollo.qui.ub.es

Abbreviations: MCR: Multivariate Curve Resolution by Alternating Least Squares. SVD: Singular Value Decomposition.

ABSTRACT

A procedure is described for the complete resolution of concentration profiles of oligonucleotide triplexes as a function of pH and temperature. The pH and temperature ranges at which triplexes are present and the relative concentrations of all the species involved in acid-base and conformational equilibria are successfully estimated from Multivariate Curve Resolution analysis of experimental spectra recorded along acid-base titrations and melting experiments of single stranded, hairpin and their mixtures. The dependence of formation constants upon pH have been successfully estimated. In this work, the hairpin h26 (5'-GAAGGAGGAGA-TTTT-TCTCCTCCTTC-3'), and the single stranded oligonucleotides s11CT (5'-CTTCCTCCTCT-3'), s11AG (5'-AGAGGAGGAAG-3') and s11TG (5'-TGTGGTGGTTG-3') have been synthesized and their protonation and conformational equilibria studied in detail. The procedure has been shown to be specially useful for the study of triplexes which show a low hypochromicity upon formation.

Keywords: triplexes, Multivariate Curve Resolution, protonation studies, melting, Triplex-Forming Oligonucleotides (TFO)

INTRODUCTION

There is an increasing interest on the biochemistry and biophysics processes involving triplex structures due to their implications on antigene therapy [Praseuth,1999; Frank-Kamenetskii,1995; Sun, 1996; Radhakrishnan, 1994]. Understanding of the cellular role of triplexes, as well as the rational design of third-strand oligonucleotides and analogues for site-specific recognition and modulation of cellular activity at the level of DNA duplex or messenger RNA, requires an understanding of the stabilities of triplexes in the context of their components and of solution conditions [Plum 1997; Mills, 2002]. Frequently, only T_m 's are measured to characterize the triplex formation as a function of solution conditions [Sugimoto,2001]. However, the complex dependence of T_m on salt species and concentrations, on the concentration of oligonucleotide, and on pH makes even comparison of T_m 's between different triplexes problematic, unless the data were collected under identical solution conditions. Moreover, microscopic phenomena, which include strand association / dissociation, release / uptake of counterions, protons, and water molecules, are not necessarily independent of one another. The complex salt- and pH-dependent behavior observed for triple-helix formation emphasizes the need of building multidimensional state diagrams. Only when the state behavior of triplexes is well investigated over a wide range of solution conditions, understanding of triplex thermodynamics is possible [Mills, 2002].

Structurally, depending on the orientation of the third strand with respect to the central polypurine Watson-Crick strand, the triplexes are classified into two main categories: i) parallel and ii) antiparallel. The parallel triplexes (also named pyrimidine-triplexes) are defined by two type of Hoogsteen triads (Figure 1a): d(T-A-T), and d(C-G-C⁺), where the last base refers to the Hoogsteen strand. Protonation of N3 of cytosine is required to form proper Hoogsteen bonding with N7 of guanine. For this reason parallel triplexes are mostly stable under acidic conditions. The antiparallel triplexes (also named purine-triplexes) are based on three reverse-Hoogsteen triads: d(C-G-G), d(T-A-A) and d(T-A-T). Unlike parallel triplexes, antiparallel triplexes requires no protonation and exhibit large pH independent binding.

Figure 1 should be placed here

Most structural studies on DNA triplexes have been focused in the parallel helices [Asensio, 1998; Leitner, 2002; Plum and Breslauer, 1995; Shindo, 1993; Sugimoto, 2001; Torigoe, 1999]. These were first discovered, and transitions involving parallel triplexes are easier to follow by UV melting curves. Accurate structural models of parallel triplexes have been derived from IR and NMR experiments [Asensio, 1998; Gotfredsen, 1998; Rajagopal, 1989; Young, 1991], and molecular dynamics (MD) simulations [Soliva,1999]. This large amount of information about the structure, reactive properties and flexibility of parallel triplexes allowed the design and synthesis of new molecules for the stabilization of the triplex in physiological conditions [Robles,1998]. Scarce amount of structural information exists on the antiparallel triplex [Ji,1996; Radhakrishnan,1993; Mills, 2002]. However, antiparallel triplexes seems to be more promising than the parallel triplexes in the biomedical field, since the formation of antiparallel triplex is pH independent, while that of parallel triplex requires in most cases an acidic pH, which not exists inside the cell. The study of antiparallel triplex formation by UV melting curves is frequently hindered by the fact that only a small hypochromism is observed [Jetter, 1993]. Only a few articles deal with a comparison of all three types of oligodeoxynucleotide third strands [Scaria and Shafer,1996; Mills, 1999; Faucon,1996].

The study of conformational transitions of polynucleotides by means of spectroscopic techniques has been traditionally carried out by monitoring the spectral changes at a single wavelength. This method has certain drawbacks, like the difficulty in the estimation of the number of species or conformations present, mainly when no selective wavelengths for every species of conformation is present. These difficulties can be solved by applying multivariate data analysis methods such as Multivariate Curve Resolution (MCR) [Tauler, 1995; Jaumot, 2002] to multiwavelength measurements. This paper proposes the application of MCR to characterize protonation and denaturation processes involving triplex structures, thereby allowing the estimation of the pure spectra and concentration profiles for all the states involved. From the estimated spectra, qualitative structural information can be obtained and from the concentration profiles an estimation of equilibrium constants can be calculated. Since

those processes involve large, unknown dependencies of the equilibrium constant on the evolution of the process, the proposed procedure is of considerable help for the study of processes involving triplexes. In this work, the hairpin h26 (5'-GAAGGAGGAGA-TTTT-TCTCCTCCTTC-3'), and the single stranded oligonucleotides s11CT (5'-CTTCCTCCTCT-3'), s11AG (5'-AGAGGAGGAAG-3') and s11TG (5'-TGTGGTGGTTG-3') (Figure 1b) have been synthesized and their protonation and conformational equilibria studied in detail.

MATERIALS AND METHODS

Reagents and Solutions

Sequences h26 (5'-GAAGGAGGAGA-TTTT-TCTCCTCCTTC-3'), s11CT (5'-CTTCCTCCTCT-3'), s11AG (5'-AGAGGAGGAAG-3') and s11TG (5'-TGTGGTGGTTG-3') (Figure 1b) were prepared on 1 μ mol scale using standard 2-cyanoethyl phosphoramidites (Cruachem Ltd.). Syntheses were performed using an automatic DNA synthesizer (Applied Biosystems Mod.392). Sequences were deprotected using standard protocols (concentrated ammonia, 55 °C, overnight). After deprotection, oligonucleotides were desalted using Sephadex G-25 columns. The DNA concentration was determined by UV absorbance measurements (260 nm) at 90 °C, using for the DNA coil state the following extinction coefficients: 7500, 8500, 12500, 15000 M⁻¹ cm⁻¹ for C, T, G and A, respectively. Oligonucleotide samples were kept at 4°C until their use.

Solutions were prepared in Ultrapure water (Millipore) with the appropriate buffer compounds: sodium monohydrogenphosphate (Panreac, a.r.), potassium dihydrogenphosphate (Panreac, a.r.), NaOH (Merck, a.r.), glacial acetic acid (Merck, a.r.), sodium acetate (Probus, a.r.), hydrochloric acid (Merck, a.r.). NaCl (Merck, a.r.) was added to adjust the ionic strength to 150 mM. Salt medium compounds were cacodylic acid sodium salt (Sigma, a.r.) and MgCl₂ (Probus, a.r.).

Apparatus

UV molecular absorption spectra were recorded on a Perkin-Elmer lambda-19 spectrophotometer between 220 and 340 nm. Temperature of the measurements was controlled by a Perkin-Elmer peltier. Instrument control and data acquisition and analysis were carried out using personal computers. pH measurements were performed with a CyberScan 2500 pH meter (with a precision of ± 0.1 mV) and a combined Hamilton pH electrode (Liq-Glass). For all measurements, a Hellma quartz cuvette (path length of 10.0 mm) was used.

Procedure

Experimental on-line set-up for spectroscopically monitored acid-base titrations has already been described elsewhere [Gargallo, 1997]. Titrations were carried out at 25 °C and 10 mM Na⁺ (sodium salt of the cacodylic acid) and 50 mM Mg²⁺ (magnesium chloride). Titrations were carried out by adding increasing volumes of HCl stock solution to a slightly basic solution of oligonucleotide. At each pH value a complete UV molecular absorption spectra was recorded.

Melting experiments were carried out at 3°C increments from 15 to 90 °C with a temperature ramp of 0.6 °C/min. At each temperature, a complete spectrum was recorded. Melting experiments were carried out at same salt medium that titration experiments and appropriate amounts of buffer solution were added to obtain the desired pH. Buffer solutions were prepared as follows: (a) 18 mM HCl, pH 2.0; (b) 3 mM HCl, pH 2.9; (c) 8.0 mM acetic acid, 2.0 mM sodium acetate, pH 3.8; (d) 5.1 mM acetic acid, 5.3 mM sodium acetate, pH 4.9 (e) 3.9 mM KH₂PO₄, 17.4 mM Na₂HPO₄, pH 6.1 (f) 10 mM KH₂PO₄, 10 mM Na₂HPO₄, pH 6.9; (g) 3.1 mM KH₂PO₄, 28.4 mM Na₂HPO₄, pH 8.2. Each sample was allowed to equilibrate at the initial temperature for 30 min before the melting experiment was started. After each experiment, the sample was cooled to initial temperature and the final spectrum was compared with the initial one to confirm the reversibility of the process.

Data Treatment

UV absorbance spectra recorded in one experiment are collected in a table or matrix **D** ($N_r \times N_w$), whose N_r rows are the N_r spectra recorded at successive temperature or pH values, and whose N_w columns are the number of wavelengths measured in every spectrum. The goal of MCR is the recovery of concentration profiles matrix **C** and of pure spectra matrix **S**^T of the spectroscopically active species or conformations present in the system under study. Multiwavelength extension of Lambert-Beer's law (in matrix form) is,

$$\mathbf{D} = \mathbf{C} \mathbf{S}^T + \mathbf{E} \quad (\text{Equation 1})$$

where $\mathbf{E}$ is the matrix of residual absorbance not explained by the model, which should be close to experimental error (Scheme 1). The first step of data analysis procedure is the estimation of the number of spectroscopically distinct species or conformations present in the experiment. This number of conformations (N_s) is initially estimated by rank analysis or singular value decomposition (SVD) of data matrix $\mathbf{D}$. In this estimation, it is assumed that singular values associated with experimental noise are significantly lower than those associated with systematic chemical data variance. Therefore, the number of species or conformations contributing significantly to data variance can be easily distinguished from the plot of the magnitudes or singular values associated to matrix $\mathbf{D}$.

Scheme 1 should be placed here

In cases where is possible to postulate a chemical model describing the interaction between species, the so-called hard-modelling methods like EQUISPEC procedure [Dyson, 1997] can be applied. This method is based on the decomposition of the equation 1 but in this case the concentration matrix $\mathbf{C}$ has to fulfill the postulated chemical model defined by: (1) a set of stoichiometries for the present species, (2) the fulfillment of mass-action law and (3) approximate values of their equilibrium constants. This procedure provides excellent results when applied to the study of well defined chemical equilibria systems of small molecules like nitrogenated bases or nucleosides. In case of more complex systems or when conditions of experiments cannot be totally controlled, the use of soft-modelling methods like the proposed MCR is more appropriate.

MCR is a procedure based on Factor Analysis methods, which can be used to analyse spectroscopic data obtained in biochemical or biophysical process monitoring [Jaumot, 2002; Navea, 2002]. In MCR, Equation 1 is solved iteratively by an Alternating Least-Squares (ALS) algorithm which estimates concentration $\mathbf{C}$ and pure spectra $\mathbf{S}^T$ matrices from experimental data matrix $\mathbf{D}$ using the proposed number of N_s conformations. This iterative optimization requires initial estimations either of $\mathbf{C}$ or $\mathbf{S}^T$, which can be obtained from pure variable detection methods [Windig, 1991] or from Evolving Factor Analysis [Maeder, 1987]. During the ALS optimization, several constraints were applied including non-negativity for concentration profiles $\mathbf{C}$ and for UV absorbance spectra profiles

$\mathbf{S}^T$, unimodality for concentration profiles $\mathbf{C}$, and closure (mass balance) also for concentration profiles $\mathbf{C}$. See previous works for a more detailed explanation of the ALS iterative optimization procedure [de Juan, 1997; Tauler, 1995; Tauler, 2001].

Concentration profiles $\mathbf{C}$ and pure spectra $\mathbf{S}^T$ resolved for each conformation in the analysis of individual data matrices may differ from the true ones because of possible unresolved underlying factor analysis ambiguities [Tauler, 1995]. This means that concentration profiles and pure spectra may be only a solution within a band of feasible solutions that are bounded by the constraints applied in the calculation [Tauler, 2001]. Some of these ambiguities are more easily solved by means of the simultaneous MCR analysis of multiple experiments under different conditions [Tauler, 1995].

The model used for MCR simultaneous analysis of two different melting experiments at different pH, giving respectively experimental data matrices $\mathbf{D}^1$ and $\mathbf{D}^2$, is described by equation (Scheme 1):

$$\begin{bmatrix} \mathbf{D}^1 \\ \mathbf{D}^2 \end{bmatrix} = \begin{bmatrix} \mathbf{C}^1 \\ \mathbf{C}^2 \end{bmatrix} [\mathbf{S}]^T + \begin{bmatrix} \mathbf{E}^1 \\ \mathbf{E}^2 \end{bmatrix} \quad (\text{Equation 2})$$

This simultaneous analysis of matrices $\mathbf{D}^1$ and $\mathbf{D}^2$ is more powerful than the separate individual analysis of matrices $\mathbf{D}^1$ or $\mathbf{D}^2$, and allows improvement of the resolution of complex experimental data structures. MCR analysis of column-wise augmented data matrices has been shown to give more reliable solutions, eventually removing totally rotational ambiguities and rank deficiency problems [Amrhein, 1996; Tauler, 1995].

All MCR calculations were performed using in-house MATLAB (version 6, The Mathworks Inc, Natick, MA) routines (Codes and tutorials are freely available at the electronic address: www.ub.es/gesq/mcr/mcr.htm).

From the MCR resolved concentration profiles it is possible to evaluate an approximate value of the apparent formation constant for the triplex structures in acidic and neutral pH values. The global reaction considered at acidic pH values is:

The apparent constant for this global reaction is defined as:

$$K_{app1} = [\text{triplex}]/[s11^+][h26^+]$$

The same equation is valid for the three triplex structures studied in this work. The concentrations are expressed in molarity.

On the other hand, the global reactions considered at neutral pH values is:

and the apparent constants is defined as:

$$K_{app2} = [\text{triplex}]/[s11][h26]$$

for the three triplexes studied in this work.

The strand molar concentration of s11 and of protonated s11 ($[s11]$ and $[s11^+]$, respectively) are estimated from their respective MCR resolved concentration profiles multiplied by the total concentration of s11 (C_{s11} , expressed in strand molarity). A similar procedure is applied to calculate the strand molar concentration of h26 and of protonated h26 ($[h26]$ and $[h26^+]$, respectively). In this case, the total concentration of h26 (C_{h26} , expressed in strand molarity) is used instead of C_{s11} . Finally, the molar concentration of triplex ($[\text{triplex}]$) is calculated by applying the mass balance equation:

$$[\text{triplex}] = C_{s11} - [s11] - [s11^+]$$

K_{app1} and K_{app2} values are calculated at pH value where the considered species are at significant concentrations. From the $\log K_{app1}$ and $\log K_{app2}$ vs. pH representation, it is possible to estimate the dependence of the apparent constants estimated with pH.

Gel-retardation assay

Nondenaturing polyacrylamide gel electrophoresis was carried out at 4 °C. The 15% polyacrylamide gels [29:1 acrylamide:bis(acrylamide)] contained 90 mM Tris-borate-EDTA (TBE) pH 8.0. All DNA samples were preheated at 90 °C for 5 min, slowly cooled, and loaded in 90 mM TBE pH 8.0, 5% glycerol, containing bromophenol blue (BPB) dye. Gels were stained for 20 min in a 0.1 mg/ ml solution of stains-all in 15% formamide in water, briefly washed with distilled water, destained with a IR lamp and photographed.

RESULTS AND DISCUSSION

s11 oligonucleotides

First, an acid-base study for each single stranded oligonucleotide was carried out to determine their pKa values at the experimental conditions and to define the UV absorption spectra of the protonated and deprotonated species of each oligonucleotide. The spectroscopically monitored acid-base titration of s11CT (pH range 2.1 – 8.9) showed a slightly increasing of the absorption intensity upon protonation, and a red shift of the maximum from 267 nm at basic pH values to 274 nm at acid pH values (Figure 2a). Experimental spectra were arranged in a data matrix **D** with dimensions $N_r = 30$ pH values, and $N_w = 121$ wavelengths measured from 220 to 340 nm. This data matrix **D** was analyzed by EQUISPEC program [Dyson,1997] according to equation 1, yielding a pKa value of 4.32 ± 0.04 for the protonation of cytosine bases on the N3 position. The result agreed with those obtained in previous works for an identical oligonucleotide [Xodo, 1991; Manzini, 1990]. Figure 2b shows the resolved concentration profiles **C**, and Figure 2c the resolved species spectra **S^T** explaining the trends previously observed in experimental spectra.

Figure 2 should be placed here

Analogously, the experimental spectra recorded along the acid-base titration of the s11TG oligonucleotide (pH range 2.6 – 7.7) were analyzed by the same procedure. A pKa value of 3.66 ± 0.08 , corresponding to the protonation of the N7 position of the guanine base was obtained. Finally, a pKa value of 3.7 ± 0.2 was obtained for the protonation of s11AG in the pH range 2.0 – 9.0. This pKa corresponds in fact to the average pKa values for the simultaneous protonation of guanine (pKa around 3.6) and adenine (pKa 3.9). It was not possible to obtain separately the pKa values for both positions because their protonation are almost simultaneous and the pure spectra for each acid-base species are very similar in shape and intensity, hindering the mathematical resolution by EQUISPEC.

h26 oligonucleotide

Next, the acid-base equilibria of the h26 oligonucleotide were studied. Upon protonation, a blue shift (from 265 nm to 260 nm) and a decrease of the absorbance maximum were observed (Figure 3a). Matrix **D** was analyzed by EQUISPEC considering the presence of two or three acid-base species. The reliability on those models was analyzed according to the final results obtained with each one of them, i.e., according to the chemical sense of the calculated pKa values, concentration profiles, the pure spectra and to the residual magnitude of the variance not explained by each model. Finally, the system including only two acid-base species, i.e., one pKa value, was finally considered to be the most chemically meaningful. For this system, the calculated pKa value was 3.9 ± 0.2 . Figure 3b shows the corresponding calculated concentration profiles **C**. The species distribution in function of pH is as follows: at neutral pH values, guanine at N1 and thymine at N3 are protonated, allowing the formation of Watson-Crick hydrogen bonds. At more acid pH values, the oligonucleotide shows simultaneous protonation of adenine at N1, cytosine at N3 and guanine at N7. The resolved species spectra (Figure 3c) explained the trends previously observed in experimental spectra.

Figure 3 should be placed here

Several melting experiments of the h26 oligonucleotide were carried out to confirm the reliability of the proposed concentration profiles **C**, and to assess the existence of a hairpin structure before the study of the triplex formation is attempted. Hence, the melting behaviour of h26 was studied at several pH values. T_m values were 71 ± 1 °C in the pH range 4.9 – 6.9, which agree with the values previously reported for the melting of duplexes of similar length and base composition [Sugimoto, 2001].

s11CT-h26 triplex

In order to study the possible formation of triple helix structures and after the study of each oligonucleotide alone, the mixtures of oligonucleotide h26 with each one of the single stranded oligonucleotides, s11CT, s11TG and s11AG were studied. From previous works [Xodo,1991; Leitner, 2000; Pilch, 1991; Wu, 2002; Sugimoto, 2001; Xodo, 1994], the interaction between h26 and

s11CT was known to produce a Hoogsteen base pairing between the purine strand of the hairpin and the single stranded oligonucleotide, which lied parallel to the purine strand. The Hoogsteen base pairing is only possible when cytosines are protonated and therefore triplex was only possible in a narrow region at acid pH values.

UV absorption data of the titration for the mixture of oligonucleotides showed a small increase of absorbance at pH values over 6.5 (Figure 4a). From pH 6.5 to 4.5, the absorbance maximum shifted from 262 nm to 264 nm and its intensity decreased. A strong shift of the maximum to 275 nm and a dramatic increase of the absorbance were observed at pH values below 4.5. Unlike the previous cases for s11 and h26, multivariate data analysis could not be carried out in this case by using hard-modelling methods like EQUISPEC because of the difficulty in postulating a chemical model based on mass-action law for the formation of triplex structures showing cooperative effects [Gargallo, 1997]. For this reason, the application of soft-modelling methods was necessary to obtain the model-free concentration profiles and spectra of the species present in the system.

Rank analysis of matrix **D** by SVD and EFA indicated $N_s = 3$. Individual MCR analysis according to Scheme 1a allowed a preliminary estimation of concentration profiles for the three species postulated (Figure 4b). The distribution of these species is as follows. At pH 8, a mixture of deprotonated s11 and h26 is present. Upon protonation, the concentration of this mixture decreased and disappeared at pH near 4. At pH values 4-5, the major species is assumed to be a mixture of triplex and of small concentrations of neutral hairpin h26 and s11. Upon protonation, the triplex structure is lost and a mixture of protonated h26 and s11 is present at pH values below 3.

Figure 4 should be placed here

These concentration profiles described the formation of a triplex structure and their dependence of the pH but the resolution can be even improved. Hence, the species at acidic pH values in Figure 4b is in fact a mixture of both protonated h26 and s11CT. Similarly, the species at neutral pH values is a mixture of both deprotonated h26 and s11CT. The complete resolution of the system should allow the recovery of the concentration profiles for all the species involved in the system. This is a typical example of rank deficiency problem where the number of mathematically detected species is lower

than the number of chemical species present in the system. This problem is detected when concentration profiles and / or spectra of the species in a system showed very similar properties and do not allow their mathematical differentiation. The better way to overcome this problem and to achieve a complete resolution of the species is the simultaneous analysis of experimental data from several experiments, each set providing additional selective information. In this case, UV absorbance data from three experiments were analyzed simultaneously to build a column-wise augmented data matrix, according to Scheme 1b. Those experiments were the titrations of h26, s11CT and the mixture of h26 and s11CT.

Preliminary analysis of the augmented data matrix showed $N_s = 5$ species. MCR resolved concentration profiles and pure spectra are shown in Figure 4c-d. As expected, the resolved concentration profiles for the titrations of h26 and s11CT showed practically identical results to those obtained previously in the individual analysis of these experiments. The resolution power of the simultaneous analysis is shown in the concentration profiles for the titration of the mixture. Now, the results allowed a more accurate description of the acid-base equilibria present in the mixture than the results previously obtained in the individual analysis of only the mixture data. At basic pH, the mixture of h26 and s11CT with deprotonated bases is obtained. The concentration of both oligonucleotides remains constant for pH values higher than 7. Below this pH value, protonation of cytosines bases have begun and the triplex structure appeared. The concentration of the triplex increased until pH 5, where it is clearly the major species. At more acidic pH values, protonation of h26 started and the triplex structure disappeared, yielding a mixture of the protonated h26 and s11CT. This species distribution agrees with a pH dependent formation of the triplex structure due to the protonation of cytosine and to the formation of Hoogsteen base pairings C-G-C⁺ and T-A-T. The proposed concentration profile agrees with that proposed by Plum and Breslauer for the intramolecular triple helix of a 31-mer oligonucleotide [Plum and Breslauer, 1995].

From MCR resolved concentration profiles for the s11CT-h26 triplex (Figure 4c), the apparent constants at the experimental conditions (25 °C, 10 mM sodium cacodilate, 50 mM Mg²⁺) were calculated according to the procedure described in the Data Treatment section (see Table 1). The

value of $\log K_{app1}$ is positively correlated to pH, i.e., the formation of the s11CT-h26 triplex is more favoured in mildly acid media, when s11 is already protonated. On the contrary, the formation of this triplex is hindered at lower acid pH values, when both s11 and h26 are fully protonated. A different behavior is observed for $\log K_{app2}$, which can be defined at slightly acidic and neutral pH values. In this pH region, $\log K_{app2}$ is negatively correlated with pH, i.e., the formation of s11CT-h26 triplex is favoured at slightly acidic pH values, and hindered at neutral and basic pH values [Shindo et al., 1993]. This fact is clearly related to the protonation of cytosine bases in s11CT, which is a necessary step for the formation of the triplex structure. Torigoe et al., (Torigoe, 1999) estimated a $\log K_{app2}$ 7.44 for a system including a 15-mer oligonucleotide containing cytosines and a 23-mer duplex (10 mM sodium cacodylate, 20 mM magnesium chloride, 200 mM sodium chloride, pH 4.9, 25 °C). This value is in good agreement with the estimated value calculated in this work according to the equation for $\log K_{app2}$ given in Table 1 (7.4). Mills et al., (Mills, 2002) estimated a $\log K_{app2}$ 6.02 for the formation constant of an intramolecular triplex of 9 base triplets at 40 °C, 10 mM Mg^{2+} and pH 6.9. Again, this value is in good agreement with the estimated value calculated using the procedure proposed in this work and given in Table 1 (5.8).

Resolved spectra showed a red shift of the absorbance maximum of h26 (from 260 nm to 265 nm) and of s11CT (from 266 nm to 273 nm) and the resolved spectra S^T of the triplex showed a maximum of the absorption band at 263 nm (Figure 4d).

Melting experiments of the mixture of h26 and s11CT were carried out at several pH values to confirm the reliability of the proposed concentration profiles. The experimental absorption spectra were analyzed with MCR according to Scheme 1a [Gargallo, 2001; Jaumot, 2002; Navea, 2002]. T_m values were determined from the resolved concentration profiles. At pH values 3.8 and 4.9, T_m values were 44 and 71 (± 1) °C. The first T_m value was assigned to the dissociation of Hoogsteen strand ~~breaking of the triplex structure~~, and agreed well with previously determined values in the literature [Manzini, 1990]. The second T_m value was assigned to the denaturation of hairpin ~~breaking of the hairpin structure~~. At pH values 6.9 and 8.2, only the melting of the hairpin was observed (T_m 69 °C).

s11AG-h26 and s11TG-h26 triplexes

Recently it has been described that substitution of purines by 8-aminopurines in parallel triplexes produce a high stabilization of the triplex [Güimil García, 1999 and Soliva, 2000]. (----). During the process of analyzing the triplex-stabilizing properties of such oligonucleotides in antiparallel triplexes several attempts to measure the melting temperatures of s11AG-h26 and s11TG-h26 triplexes were carried out. UV melting curves at 260 nm or 275 nm gave only duplex-to-random-coil transition in spite of observing triplex formation by gel-shift assays (Figure 5). This fact could be related to the low hypochromicity associated with the formation of antiparallel triplexes. In this work, the multivariate approach was applied to study the formation of antiparallel triplex structures. Hence, acid-base titrations of mixtures of hairpin oligonucleotide h26 with s11AG or s11TG were carried out and the results analyzed by MCR.

Figure 5 should be placed here

The analysis of absorbance data from the acid-base titration of the mixture of h26 and s11AG was performed in the same way as described above for s11CT. MCR analysis of the augmented data matrix according to Scheme 1b allowed the complete resolution of the concentration profiles and pure spectra for the triplex, hairpin and single-stranded oligonucleotides (Figure 6).

Figure 6 should be placed here

The resolved concentration profiles **C** allowed the description of the h26 and s11AG acid-base equilibria (Figure 6a). At neutral pH values, a complex mixture of the triplex and the free deprotonated oligonucleotides h26 and s11AG is observed. Their relative concentrations remained approximately constant until pH 4. Differently to the triplex formed by h26 and s11CT, the existence of the triplex at neutral pH values is possible because the formation of non-protonated? reverse Hoogsteen base pairs, allowing the formation of the C-G-G and the T-A-A triads. At pH 4, protonation of the bases induced the ~~breaking of the hydrogen bonds~~ disappearance of hydrogen bonds and the destabilization of the triplex. At more acid pH values, only both, protonated h26 and s11AG species are present.

From MCR resolved concentration profiles for the s11AG-h26 triplex (Figure 6a), the apparent constants at the experimental conditions (25 °C, 10 mM sodium cacodylate, 50 mM Mg^{2+}) were calculated according to the procedure described in the Data Treatment section (see Table 1). The value of $\log K_{\text{app1}}$ shows a weak negative correlation with pH, in contrast to the observed behavior for the s11CT-h26 triplex. The calculated mean value of $\log K_{\text{app1}}$ is 3.6 ± 0.3 , at the pH range 2.02 – 4.05. On the contrary, the value for the $\log K_{\text{app2}}$ did not show any dependence with pH. The calculated mean value is 7.0 ± 0.1 at the pH range 5.75 – 8.01, which is in the same order of magnitude than previous values reported by other authors for triplex involving homopurine strands containing A,G bases. Hence, Reither et al. (Reither, 2002) estimated a value for the formation constant of 6.4 for an intermolecular triplex of 13 base triplets at 37 °C, 10 mM Mg^{2+} and pH 7.5. Aich et al. (Aich, 1998) estimated a value of 7.8 for the formation constant of an triplex of 11 base triplets at 22 °C and pH 7.0. Finally, Mills et al. (Mills, 2002) estimated a value of 5.8 for the formation constant of an intramolecular triplex of 9 base triplets at 40 °C, 10 mM Mg^{2+} and pH 6.9.

The resolved pure spectra (Figure 6b) reflect the low hypochromism associated with the formation of triplex structure between h26 and s11AG. This fact was already observed in a previous work where only the univariate measurements at 260 nm were used. Hence, Jetter and Hobbs already observed triplex formation with an oligonucleotide containing only thymine and 7,8-dihydro-8-oxoadenine without significative hypochromicity, which made it difficult to be detected in conventional melting experiments [Jetter, 1993]. No melting transitions for GT- and GA-triplex forming oligonucleotides was also described by Faucon et al. in a similar context [Faucon, 1996] in spite than triplex formation was evident by gel retardation assays and footprinting experiments.

The formation of triplex structure was confirmed by melting experiments. Two T_m values (44 ± 1 and 72 ± 2 °C) were determined from the MCR resolved concentration profiles for each melting experiment. As in the case of parallel triplex, the first value was identified as the dissociation of Hoogsteen strand and the second one as the denaturation of hairpin.

The results for the mixture of h26 and s11TG were very similar to those for the mixture of h26 and s11AG because the formation of triplex was also a consequence of reverse Hoogsteen base pairing

bonds (Figure 6c). At basic and neutral pH, triplex is the major species and, upon protonation, triplex structure is lost yielding a mixture of protonated h26 and s11TG.

Finally, as in the previous cases, the dependence of $\log K_{app1}$ and $\log K_{app2}$ were calculated for the s11TG-h26 triplex formation (see Table 1). The calculated dependence and mean values were similar to those estimated for the s11AG-h26 triplex. Again, they were in the same order of magnitude than the few formation constant values reported in literature for triplex structures involving homopurine strands containing T, G bases. Mills et al., (Mills, 2002) estimated a $\log K_{app2}$ of 6.10 for the formation constant of an intramolecular triplex of 9 base triplets at 40 °C, 10 mM Mg^{2+} and pH 6.9.

The resolved pure spectra (Figure 6d) reflect again the lack of hypochromism associated with the formation of triplex structure between h26 and s11TG.

T_m values for the thermal denaturation of the triplex formed by h26 and s11TG were 28 ± 1 and 72 ± 2 °C. The first value was 16 °C lower than the T_m determined for the other two triplexes. This difference of melting temperature is in agreement with previous studies [Scaria, 1996, He, 1997, Faucon, 1996] and could be explained because the antiparallel G-G-C triplet is more closely isomorphous to A-A-T than reverse Hoogsteen T-A-T (figure 1).

Conclusions

In this work, we have studied the effect of the pH and the relative stability of three different triplex based on the same target sequence. The pH and temperature ranges at which triplexes are present and the relative concentrations of all the species involved in acid-base and conformational equilibria are successfully estimated from Multivariate Curve Resolution analysis of experimental spectra. This study provides a detailed comparison of the different types of DNA triplexes under the same conditions. It has been described that the formation of parallel triplexes is accompanied by a significant hypochromism that provides a convenient signal for studying the relative stability of triplexes by UV melting. In contrast, the antiparallel triplexes shows a much smaller change in hypochromism presumably because the purine third strand is largely stacked in the free state [He et al. 1996]. In our case, the melting temperatures of the different triplexes were able to be measured only by using Multivariate Curve Resolution analysis. As judged by the melting temperatures the

most stable triplexes at the optimal pH were the parallel triplex (44 °C, pH 3.8-4.9) and the antiparallel triplex formed by the s11AG (44°C, pH 4-8) and the least stable was the antiparallel triplex formed by the s11TG (28 °C, pH 4-8). This result is important because it shows that the stability of antiparallel triplexes formed by (G, A)-oligonucleotides have been underestimated probably due to the difficulties to measure the melting temperatures using monochromatic UV measurements and the fact that G-rich oligonucleotides adopt unusual tetrameric structures which compete for triplex formation .

ACKNOWLEDGEMENTS

A. Aviñó (CSIC, Barcelona), K. Lorand (Institut Universitaire Technologique, Lannion, France), R. Pilni and J.Topinkova (Mazaryk University, Brno, Czech Republic) are thanked for their help carrying out preliminary experimental work. This study was supported by the Ministerio de Ciencia y Tecnología (grant BQU2000-0788 and BQU2000-0649) and the Generalitat de Catalunya (2001SGR-00056 and 2001SGR-0049). J.J. also thanks MCYT for a Ph.D. grant.

REFERENCES

- Aich, P., Ritchie, S., Bonham, K. and Lee, J.S. 1998. Thermodynamic and kinetic studies of the formation of triple helices between purine-rich deoxyribo-oligonucleotides and the promoter region of the human c-src proto-oncogene. *Biochem.* 26: 4173 - 4177
- Amhrein, M., Srinivasan, B., Bonvin, D. and Shumacher, M.M. 1996. On the rank deficiency and rank augmentation of the spectral measurement matrix. *Chemom. Intel. Lab. Syst.* 33: 17 - 33.
- Asensio, J.L., Lane, A.N.; Dhesi, J., Bergqvist, S. and Brown, T. 1998. The contribution of cytosine protonation to the stability of parallel DNA triple helices. *J. Mol. Biol.* 275: 811 – 822
- de Juan, A., van der Heyden, Y., Tauler, R. and Massart, D.L. 1997. Assessment of new constraints applied to alternating least squares method. *Anal. Chim. Acta* 346: 307- 318
- Dyson, R., Kaderli, S., Lawrence, G.A., Maeder, M. and Zuberbühler, A.D. 1997. Second order global analysis: the evaluation of series of spectrophotometric titrations for improved determination of equilibrium constants. *Anal. Chim. Acta* 353: 381 - 393
- Faucon, B.; Mergney, J.L. and Hélène, C. 1996. Effect of third strand composition on triple helix formation: purine versus pyrimidine oligodeoxynucleotides. *Nucleic Acids Res.* 24, 3181 - 3188
- Frank-Kamenetskii, M.D. and Mirkin, S.M. 1995. Triplex DNA structures. *Annu. Rev. Biochem.* 64: 65 - 95
- Gargallo, R., Tauler, R. and Izquierdo-Ridorsa, A. 1997. Acid-base and copper II complexation equilibria of polyinosinic-polycytidylic acid. *Biopol.* 42: 271 - 283
- Gargallo, R., Vives, M., Tauler, R. and Eritja, R. 2001. Protonation Studies and Multivariate Curve Resolution on Oligodeoxynucleotides Carrying the Mutagenic Base 2-Aminopurine. *Biophys. J.* 81: 2886 – 2896
- Güimil García, R., Ferrer, E., Macías, M.J., Eritja, R., and Orozco, R. 1999. Theoretical calculations, síntesis and base pairing properties of oligonucleotides containing 8-amino-2'-deoxyadenosine. *Nucleic Acids Res.* 27: 1991-1998.

- Gotfredsen, C.H.; Schultze, P. and Feigon, J. 1998. Solution Structure of an Intramolecular Pyrimidine-Purine-Pyrimidine Triplex Containing an RNA Third Strand. *J.Am.Chem.Soc.* 120: 4281-4289
- He, Y., Scaria, P.V. and Shafer, R.H. 1997. Studies on Formation and Stability of the $d[G(AG)_5] \cdot d[G(AG)_5] \cdot d[C(TC)_5]$ and $d[G(TG)_5] \cdot d[G(AG)_5] \cdot d[C(TC)_5]$ Triple Helices. *Biopolymers* 41: 431-441
- Jaumot, J., Escaja, N., Gargallo, R., González, C., Pedroso, E. and Tauler, R. 2002. Multivariate Curve Resolution: A powerful tool for the analysis of conformational transitions in nucleic acids. *Nucleic Acids Res.* 30(17): e92
- Jetter, M.C. and Hobbs, F.W. 1993. 7,8-Dihydro-8-oxoadenine as a replacement for cytosine in the third strand of triplexes. Triplex formation without hypochromicity. *Biochem.* 32: 3249 - 3254
- Ji, J., Hogan, M.E. and Gao, X. 1996. Solution structure of an antiparallel purine motif triplex containing a T.CG pyrimidine base triple. *Structure* 4: 425 - 435
- Leitner, D., Schroeder, W. and Weisz, K. 2000. Influence of Sequence-Dependent Cytosine Protonation and Methylation on DNA Triplex Stability. *Biochem.* 39: 5886 - 5892
- Maeder, M. 1987. Evolving factor analysis for the resolution of overlapping chromatographic peaks. *Anal. Chem.* 59: 527 - 531
- Manzini, G., Xodo, L.E., Gasparotto, D., Quadrifoglio, F., Van der Marel, G.A. and Van Boom, J.H. 1990. Triplex formation by oligopurine-oligopyrimidine DNA fragments. Electrophoretic and thermodynamic behavior. *J. Mol. Biol.* 213: 833 - 843
- Mills, M., Arimondo, P., Lacroix, L., Garestier, T., Hélène, C., Klump, H.H. and Mergny, J.L. 1999. Energetics of strand-displacement reactions in triple helices: a spectroscopic study. *J. Mol. Biol.* 291: 1035 - 1054
- Mills, M.; Arimondo, P.B.; Lacroix, L.; Garestier, T.; Lump, H. and Mergny, J.-L. 2002. Chemical modification of the third strand: differential effects on purine and pyrimidine triple helix formation. *Biochem.* 41: 357 - 366
- Navea, S., de Juan, A. and Tauler, R. 2002. Detection and resolution of intermediate species in protein folding processes using fluorescence and circular dichroism spectroscopies and multivariate curve resolution. *Anal. Chem.* 74: 6031 - 6039

- Pilch, D.S., Levenson, C. and Shafer, R.H. 1991. Structure, Stability, and Thermodynamics of a Short Intermolecular Purine-Purine-Pyrimidine Triplex. *Biochem.* 30: 6081 - 6087
- Plum, G.E. and Breslauer, K. J. 1995. Thermodynamics of an intramolecular DNA triple helix: a calorimetric and spectroscopic study of the pH and salt dependence of thermally induced structural transitions. *J. Mol. Biol.* 248: 679 - 695
- Plum, G.E. 1997. Thermodynamics of Oligonucleotide Triplexes. *Biopolymers* 44: 241 - 256
- Praseuth, D., Guieysse, A. and Hélène, C. 1999. Triple helix formation and the antigene strategy for sequence-specific control of gene expression. *Biochim. Biophys. Acta* 1489: 181 - 206
- Radhakrishnan, I. and Patel, D.J. 1993. Solution structure of a purine-purine-pyrimidine DNA triplex containing G-GC and T-AT triples. *Structure* 1: 135 - 152
- Radhakrishnan, I. and Patel, D.J. 1994. DNA triplexes: solution structures hydration sites, energetics, interactions, and function. *Biochem.* 33: 11405 - 11416
- Rajagopal, P. and Feigon, J. 1989. Triple-strand formation in the homopurine:homopyrimidine DNA oligonucleotides d(G-A)₄ and d(T-C)₄. *Nature* 339: 637 - 640
- Reither, S. and Jeltsch, A. 2002. Specificity of DNA triple helix formation analyzed by a FRET assay. *BMC Biochem.* 3: 27
- Robles, J., Grandas, A., Pedroso, E., Luque, F.J., Eritja, R. and Orozco, M. 2002. Nucleic acid triple helices: Stability effects of nucleobase modifications. *Curr. Org. Chem.* 6: 1333 - 1368
- Scaria, P.V. and Shafer, R.H. 1996. Calorimetric analysis of triplexes targeted to the d(G₃A₄G₃)·d(C₃T₄C₃) duplex. *Biochem.* 35:10985 - 10994
- Shindo, H., Torigoe, H. and Sarai, A. 1993. Thermodynamic and kinetic studies of DNA triplex formation of an oligohomopyrimidine and a matched duplex by filter binding assay. *Biochem.* 32: 8963 - 8969
- Soliva, R., Luque, F.J. and Orozco, M. 1999. Can G-C Hoogsteen-wobble pairs contribute to the stability of d(G-C-C) triplexes? *Nucleic Acids Res.* 27: 2248- 2255

- Soliva, R., Güimil García, R., Blas, J.R., Eritja, R., Asensio, J.L., González, C., Luque, F.J., and Orozco, M. 2000. DNA-triplex stabilizing properties of 8-aminoguanine. *Nucleic Acids Res.* 28: 4531-4539.
- Sugimoto, N., Wu, P., Hara, H. and Kawamoto, Y. 2001. pH and Cation Effects on the Properties of Parallel Pyrimidine Motif DNA Triplexes. *Biochem.* 40:9396 - 9405
- Sun, J.S., Garestier, T. and Hélène, C. 1996. Oligonucleotide directed triple helix formation. *Curr. Opin. Struct. Biol.* 6: 327 - 333
- Tauler, R. 1995. Multivariate curve resolution applied to second-order data. *Chemom. Intell. Lab. Syst.* 30: 133 - 146
- Tauler, R. 2001. Calculation of maximum and minimum band boundaries of feasible solutions for species profiles obtained by multivariate curve resolution. *J. Chemom.* 15: 627 - 646
- Torigoe, H., Shimizume, R., Sarai, A. and Shindo, H. 1999. Triplex formation of chemically modified homopyrimidine oligonucleotides: thermodynamic and kinetic studies. *Biochem.* 38: 14653 - 14659
- Windig, W. and Guilment, J. 1991. Interactive self-modeling mixture analysis. *Anal. Chem.* 63:1425 - 1430
- Wu, P., Kawamoto, Y., Hara, H. and Sugimoto, N. 2002. Effect of divalent cations and cytosine protonation on thermodynamic properties of intermolecular DNA double and triplexes. *J. Inorg. Biochem.* 91: 277 - 285
- Xodo, L.E., Manzini, G., Quadrifoglio, F., van der Marel, G.A. and van Boom J.H. 1991. Effect of 5-methylcytosine on the stability of triple-stranded DNA – a thermodynamic study. *Nucleic Acids Res.* 19: 5625 - 5631
- Xodo, L., Alunni-Fabbroni, M., Manzini, G. and Quadrifoglio, F. 1994. Pyrimidine phosphorothioate oligonucleotides form triple-stranded helices and promote transcription inhibition. *Nucleic Acids Res.* 22: 3322 - 3330
- Young, S.L., Krawczyk, S.H., Matteuci, M.D. and Toole, J.J. 1991. Triple helix formation inhibits transcription elongation in vitro. *Proc. Natl. Acad. Sci. USA.*,88: 10023 - 10026

FIGURE CAPTIONS

Scheme 1. Data matrix arrangement for (a) MCR individual analysis of a matrix **D** recorded along one single experiment, and (b) MCR simultaneous analysis of several data matrices recorded along different experiments.

Figure 1. (a) Triads base pairing: (1) Hoogsteen C-G-C⁺, (2) reverse Hoogsteen C-G-G, (3) reverse Hoogsteen T-A-T, (4) reverse Hoogsteen T-A-A. (b) Schematic representation of the h26 and s11CT, s11AG and s11TG interactions.

Figure 2. Acid-base titration of s11CT (a) Experimental absorbance data, matrix **D**. (b) Concentration profiles calculated by EQUISPEC, **C**. (c) Pure spectra, **S^T**. (—) : protonated s11 species. (....) : deprotonated s11 species.

Figure 3. Acid-base titration of h26. (a) Experimental absorbance data, **D**. (b) Concentration profiles calculated by EQUISPEC, **C**. (c) Pure spectra, **S^T**. (.-.-) : protonated h26 species. (---) : deprotonated h26 species.

Figure 4. Acid-base titration of the mixture of h26 and s11CT. (a) Experimental absorbance data, **D**. (b) Concentration profiles calculated by MCR according to Scheme 1a, **C**. 1: mixture of protonated h26 and protonated s11CT. 2: mixture of species containing the triplex as majority species. 3: mixture of deprotonated h26 and deprotonated s11CT. (c) Concentration profiles for the triplex acid-base titration calculated by MCR according to the simultaneous analysis depicted in Scheme 1b, **C_{triplex}**. (d) Pure spectra, **S^T**. Thick continuous line: triplex. Other lines as in figures 2 and 3.

Figure 5. Photograph of a 15% polyacrylamide gel in 90 mM Tris-borate-EDTA (TBE) pH 8.0 stained with stains-all. Lane 1: bromophenol blue (BPB) dye; lane 2: s11AG; lane 3: h26; lane 4: a stoichiometric mixture of h26 and s11AG; lane 5: a stoichiometric mixture of h26 and s11TG; lane 6: s11TG.

Figure 6. MCR results for the acid-base titrations of the mixtures of h26 and s11AG and of h26 and s11TG. Concentration profiles **C** (a) and pure spectra (b) for the mixture of h26 and s11AG. Concentration profiles **C** (c) and pure spectra (d) for the mixture of h26 and s11TG. Lines as in figures 2, 3 and 4.

Table 1. Calculated values for logKapp1 and logKapp2.

Table 1. Calculated values for $\log K_{app1}$ and $\log K_{app2}$.

s11 oligo in triplex	Observed dependence for $\log K_{app}$ vs. pH	Studied pH range	Number of experimental points	Correlation value (r^2)	Mean value for $\log K_{app}$
-------------------------	--	---------------------	----------------------------------	--------------------------------	----------------------------------

$\log K_{app1}$

s11CT	$1,14 \cdot \text{pH} + 3,74$	2.05 – 3.75	8	0,9712	
s11AG	$-0,34 \cdot \text{pH} + 4,69$	2.02 – 4.05	9	0,9335	3.6 ± 0.3
s11TG	$-0,38 \cdot \text{pH} + 4,62$	3.32 – 4.15	4	0,9152	3.2 ± 0.2

$\log K_{app2}$

s11CT	$-0,80 \cdot \text{pH} + 11,32$	4.53 – 6.49	9	0,9647	
s11AG	$-0,04 \cdot \text{pH} + 7,27$	5.75 – 8.01	9	0,3985	7.0 ± 0.1
s11TG	$-0,14 \cdot \text{pH} + 8,22$	5.67 – 6.83	6	0,9808	7.3 ± 0.1

Scheme 1. Data matrix arrangement for (a) MCR individual analysis of a matrix **D** recorded along one single experiment, and (b) MCR simultaneous analysis of several data matrices recorded along different experiments.

$$\begin{array}{l}
 \text{a)} \quad \begin{array}{c} \lambda_1 \dots \dots \lambda_n \\ pH_1 \\ \vdots \\ pH_m \end{array} \begin{array}{|c|} \hline \mathbf{D} \\ \hline \end{array} = \begin{array}{c} 1 \dots \dots N_s \\ pH_1 \\ \vdots \\ pH_m \end{array} \begin{array}{|c|} \hline \mathbf{C} \\ \hline \end{array} \begin{array}{c} \lambda_1 \dots \dots \lambda_n \\ \mathbf{S}^T \\ \vdots \\ N_s \end{array} + \begin{array}{c} \lambda_1 \dots \dots \lambda_n \\ pH_1 \\ \vdots \\ pH_m \end{array} \begin{array}{|c|} \hline \mathbf{E} \\ \hline \end{array}
 \end{array}$$

$$\begin{array}{l}
 \text{b)} \quad \begin{array}{c} \lambda_1 \dots \dots \lambda_n \\ pH_1^{h26} \\ \vdots \\ pH_m^{h26} \\ pH_1^{s11} \\ \vdots \\ pH_m^{s11} \\ pH_1^{mix} \\ \vdots \\ pH_m^{mix} \end{array} \begin{array}{|c|} \hline \mathbf{D}^{h26} \\ \hline \mathbf{D}^{s11} \\ \hline \mathbf{D}^{mix} \\ \hline \end{array} = \begin{array}{c} 1 \dots \dots N_s \\ pH_1^{h26} \\ \vdots \\ pH_m^{h26} \\ pH_1^{s11} \\ \vdots \\ pH_m^{s11} \\ pH_1^{mix} \\ \vdots \\ pH_m^{mix} \end{array} \begin{array}{|c|} \hline \mathbf{C}^{h26} \\ \hline \mathbf{C}^{s11} \\ \hline \mathbf{C}^{mix} \\ \hline \end{array} \begin{array}{c} \lambda_1 \dots \dots \lambda_n \\ \mathbf{S}^T \\ \vdots \\ N_s \end{array} + \begin{array}{c} \lambda_1 \dots \dots \lambda_n \\ pH_1^{h26} \\ \vdots \\ pH_m^{h26} \\ pH_1^{s11} \\ \vdots \\ pH_m^{s11} \\ pH_1^{mix} \\ \vdots \\ pH_m^{mix} \end{array} \begin{array}{|c|} \hline \mathbf{E}^{h26} \\ \hline \mathbf{E}^{s11} \\ \hline \mathbf{E}^{mix} \\ \hline \end{array}
 \end{array}$$

Figure 1. (a) Triads base pairing: (1) Hoogsteen C-G-C⁺, (2) reverse Hoogsteen C-G-G, (3) reverse Hoogsteen T-A-T, (4) reverse Hoogsteen T-A-A. (b) Schematic representation of the h26 and s11CT, s11AG and s11TG interactions.

Figure 2. Acid-base titration of s11CT (a) Experimental absorbance data, matrix $\mathbf{D}$. (b) Concentration profiles calculated by EQUISPEC, $\mathbf{C}$. (c) Pure spectra, $\mathbf{S}^T$. (—) : protonated s11 species. (....) : deprotonated s11 species.

Figure 3. Acid-base titration of h26. (a) Experimental absorbance data, **D**. (b) Concentration profiles calculated by EQUISPEC, **C**. (c) Pure spectra, **S^T**. (.-.) : protonated h26 species. (---) : deprotonated h26 species.

Figure 4. Acid-base titration of the mixture of h26 and s11CT (a) Experimental absorbance data, **D**. (b) Concentration profiles calculated by MCR according to Scheme 1a, **C**. 1: mixture of protonated h26 and protonated s11CT. 2: mixture of species containing the triplex as majority species. 3: mixture of deprotonated h26 and deprotonated s11CT. (c) Concentration profiles for the triplex acid-base titration calculated by MCR according to the simultaneous analysis depicted in Scheme 1b, C_{triplex} . (d) Pure spectra, S^T . Thick continuous line: triplex. Other lines as in figures 2 and 3.

Figure 5. Photograph of a 15% polyacrylamide gel in 90 mM Tris-borate-EDTA (TBE) pH 8.0 stained with stains-all. Lane 1: bromophenol blue (BPB) dye; lane 2: s11AG; lane 3: h26; lane 4: a stoichiometric mixture of h26 and s11AG; lane 5: a stoichiometric mixture of h26 and s11TG; lane 6: s11TG.

Figure 6. MCR results for the acid-base titrations of the mixtures of h26 and s11AG and of h26 and s11TG. Concentration profiles $\mathbf{C}$ (a) and pure spectra $\mathbf{S}^T$ (b) for the mixture of h26 and s11AG. Concentration profiles $\mathbf{C}$ (c) and pure spectra $\mathbf{S}^T$ (d) for the mixture of h26 and s11TG. Lanes as in figures 2, 3 and 4.