Nucleic Acids Research, 1997, Vol. 25, No. 22 **4523–4531**

NMR determination of the conformational and drug binding properties of the DNA heptamer d(GpCpGpApApGpC) in aqueous solution

David B. Davies*, Valery I. Pahomov[1] and Alexei N. Veselkov[1]

Department of Chemistry, Birkbeck College, University of London, Gordon House, 29 Gordon Square, London WC1H OPP, UK and [1]Department of Physics and Chemistry, Sevastopol State Technical University, Sevastopol, 335053 Crimea, Ukraine

Received August 6, 1997; Revised and Accepted September 15, 1997

ABSTRACT

1D and 2D NMR spectroscopy (500/600 MHz) has been used to investigate the equilibrium conformational states of the deoxyheptanucleotide 5′-d(GpCpGpApApGpC), as well as its complexation with the phenanthridinium drug ethidium bromide (EB). Quantitative determination (reaction constants and thermodynamic parameters) of the conformational equilibrium of the heptamer in solution and its complexation with EB was based on analysis of the dependence of proton chemical shifts on concentration (at two temperatures, 298 and 308 K) and on temperature (in the range 278–353 K). The experimental results were analysed in terms of a model of the dynamic equilibrium between single-stranded, hairpin and bulged dimer forms of the deoxyheptanucleotide and its complexes with EB. Calculation of the relative amounts of the different complexes reveals important features of the dynamic equilibrium as a function of both temperature and the ratio of the drug and heptamer concentrations. The quantitative analysis also provides the limiting proton chemical shifts of EB in each complex which have been used to determine the most favourable structures of the intercalated complexes of EB with the (GC) sites of both the hairpin and dimer forms of the heptanucleotide.

INTRODUCTION

In addition to the classical regular helical forms, it is known that double-stranded DNA may have different secondary structures, such as hairpins, internal loops and branches. Hairpin structures are common features of RNA molecules (1) and they have been found to exist in DNA in regions with palindromic sequences involved in gene regulation (2,3). For example, hairpins containing 3 nt in the loop occur frequently in native DNAs and RNAs, for example in replication origins of some phages and viruses (4,5), in a promoter region of an *Escherichia coli* heat shock gene (6) and in rRNA genes. The possible involvement of DNA hairpins in biological processes has led to numerous physicochemical investigations of these structures *in vitro* (7–14). A hairpin structure resulting from intramolecular folding of a partially complementary palindromic DNA sequence is composed of a double-helical stem and a single-stranded loop. The double-helical stem contains at least 2 bp and the loop consists of 2–5 nt (7–14). The conformation of the hairpin loop may be determined by different factors, including stacking interactions of the loop residues with base pairs of the stem, the type of terminal base pair of the stem, steric constraints for loop closure and intermolecular interactions (stacking and hydrogen bond formation) of the loop bases (7–14).

It has been shown in recent years that a series of short DNA sequences form extraordinarily stable structures containing only two G-C pairs in the stem (15). Of all these sequences the deoxyheptanucleotide d(GCGAAGC) forms the most stable hairpin structure (melting temperature T_m ~349 K in 0.1 M NaCl) having a trinucleotide (GAA) loop which is also resistant to nucleases (15). Under the same experimental conditions the analogous RNA fragment, r(GCGAAGC), forms a significantly less stable hairpin compared with d(GCGAAGC), which may result from different structures of the stems formed by DNA (B-form) and RNA (A-form) sequences (16).

The three-dimensional structure of the DNA hairpin, d(GCGAAGC), in solution has been determined using 2D NOE data (15), assuming that the heptamer exists in solution in a single conformational state. It has been found previously that for short oligonucleotides there is an equilibrium of different structures of the molecules in solution, including different conformational states of oligonucleotide chains (13,17,18). Hence, it is important to investigate the conformational properties in solution of oligonucleotides forming hairpin structures in order to elucidate any distinctive features of the dynamic equilibrium between the monomeric and associated forms of these molecules. At the same time, it is important to investigate binding with such sequences of biologically active ligands, in particular intercalators, and their influence on the recognition process of non-classical DNA structures by regulatory proteins (19,20).

In this work we report an NMR (500/600 MHz) study of the equilibrium conformational states of the deoxyheptanucleotide d(GpCpGpApApGpC) and its complexation with a typical intercalator, ethidium bromide (EB), in aqueous salt solution. The quantitative determination of the conformational equilibrium of

*To whom correspondence should be addressed. Tel: +44 171 380 7475; Fax: +44 171 380 7464

the heptamer is based on investigation of the dependence of proton chemical shifts of the molecules on concentration and temperature and quantitative determination of the complexation properties of EB with the heptamer is based on investigation of the concentration and temperature dependences of proton chemical shifts of drug–nucleic acid mixtures in solution.

MATERIALS AND METHODS

The deoxyheptanucleotide 5′-d(GpCpGpApApGpC) was synthesized by the Oswel DNA Service (University of Southampton). EB was purchased from Sigma Chemical Co. and used without further purification. The samples were lyophilized from 99.95% D_2O and redissolved in deuterated 0.1 mol/l phosphate buffer, pD 7.15, containing 10^{-4} mol/l EDTA. The concentration of the drug stock solution was measured spectrophotometrically using the molar absorption coefficient ε = 5860 mol/l/cm (λ = 480 nm) (21).

500 MHz ^{1}H NMR spectra were recorded on a JEOL GSX 500 spectrometer with the residual water peak saturated during relaxation. Measurements as a function of concentration of the heptamer solutions as well as of the mixed drug–heptamer solutions were made at two temperatures (298 and 308 K) and measurements as a function of temperature were made at constant concentration in the temperature range 278–353 K. Chemical shifts were measured relative to an internal reference, TMA (tetramethylammonium bromide), and then recalculated with respect to DSS (sodium 2,2-dimethyl-2-silapentane-5-sulfonate), i.e. $\delta_{DSS} = \delta_{TMA} + 3.178$ (p.p.m.).

2D homonuclear TOCSY and NOESY experiments were carried out on a 600 MHz Bruker AMX spectrometer. 2D NOE spectra were recorded at a fixed temperature (T = 298 K) with a standard pulse sequence with different mixing times of τ_{m1} = 90 ms and τ_{m2} = 200 ms over a sweep width of 5400 Hz using 4096 data points in the t_2 dimension and 512 increments in the t_1 dimension. 2D TOCSY spectra were measured at the same temperature using 2048 data points in t_2 and 256 increments in the t_1 dimension with a mixing time of τ_m = 70 ms. The repetition delays were 1.5 and 3.0 s for 2D TOCSY and 2D NOESY experiments respectively. Sixteen scans were collected for each t_1 increment in TOCSY and 32 scans in NOESY experiments. 2D heteronuclear ^{1}H-^{31}P correlation experiments (HMBC) were carried out on a 500 MHz Bruker DRX spectrometer using published pulse sequences (22). The sample temperature was regulated using either Bruker or JEOL variable temperature units, as appropriate.

RESULTS AND DISCUSSION

NMR analysis of different conformational states of the heptamer in aqueous solution

Signal assignment and experimental observations. Complete assignment of all the non-exchangeable proton signals in the PMR spectrum of the deoxyheptanucleotide 5′-d(GpCpGpApApGpC) at a concentration of 5.61 mmol/l has been made using homonuclear 2D TOCSY and 2D NOESY experiments and the chemical shifts are in good agreement with the previous assignments for this heptamer under similar experimental conditions, except for some small differences for the deoxyribose protons, which may be due to the slightly different ionic strength of the solution in the previous work (15).

The concentration (at T = 298 K for example) and temperature dependence of the chemical shifts of some aromatic protons and H1′ of the deoxyribose rings of the heptamer 5′-d(GpCpGpApApGpC) are shown in Figure 1. The experimental data indicate that most of the protons experience chemical shift displacement to low frequency on increasing the concentration in the range 0.1–5.6 mmol/l. It was previously assumed that the heptamer d(GCGAAGC) only formed a hairpin structure in solution (15); in this case no concentration dependence of proton chemical shifts should be observed, which is not consistent with the present experimental results at 298 and 308 K. It is known that the probability of aggregation depends on the concentration, ionic strength, temperature of solution and molecular weight of the oligomer. Investigations carried out by Freier *et al.* (23) provide evidence that formation of higher order aggregates is negligible at the concentrations and relatively high ionic strength used in the present work.

The experimentally observed dependence of proton chemical shifts on concentration (Fig. 1) is a result of intermolecular interactions between heptamer molecules, i.e. most probably the formation of bulged dimer complexes rather than hairpin dimers in solution.

Conformational model and analysis. The equilibrium conformational states of 5′-d(GpCpGpApApGpC) in aqueous solution can be represented by the following reaction scheme (Fig. 2):

$$A \overset{K_l}{\Leftrightarrow} A_1 \qquad A + A \overset{K_a}{\Leftrightarrow} A_2 \qquad \mathbf{1}$$

where A, A_l, A_2 are monomer, hairpin and dimer of the heptamer respectively and K_l, K_a are the equilibrium reactions constants of formation of hairpin and dimer respectively.

Using the mass conservation law for the heptamer and taking into consideration the mass law equations for reactions **1**, the total concentration of heptamer $[A_0]$ is given by equation **2**:

$$[A_0] = [A] + K_l[A] + 2K_a[A]^2 \qquad \mathbf{2}$$

where [A] is the equilibrium molar concentration of the oligonucleotide in the monomer form.

The additive model (24,25) was used to describe the dependence of proton chemical shifts on concentration and temperature:

$$\delta = \delta_m f_m + \delta_l f_l + \delta_a f_a \qquad \mathbf{3}$$

where f_m, f_l, f_a are equilibrium mole fractions and δ_m, δ_l, δ_a are the proton chemical shifts in the monomer (A), hairpin (A_l) and dimer (A_2) forms of the deoxyheptanucleotide respectively. The validity of such a model assumes fast exchange between all conformations and interacting molecules, especially at lower temperatures. Our NMR experiments have shown that there was no significant line broadening for all the systems studied as a function of both concentration and temperature. Therefore, it is concluded that the additive model for description of the observed proton chemical shifts may be used to analyse the experimental results on the conformational properties of the heptamer 5′-d(GCGAAGC), as well as for complexation of the phenantridinium drug EB to the heptamer (see later).

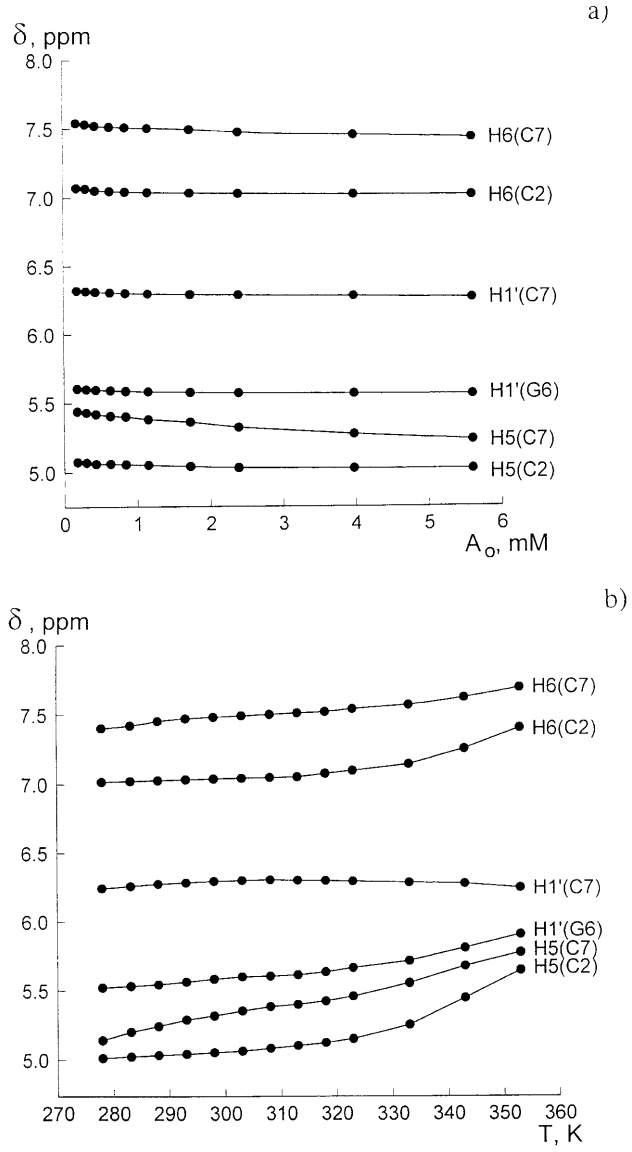


Figure 1. Dependence of proton chemical shifts (**a**) on concentration at constant temperature, (T = 298 K) and (**b**) on temperature at constant concentration ([A_0] = 2.55 mmol/l) of some of the non-exchangeable protons of the deoxyheptanucleotide 5′-d(GpCpGpApApGpC) in 0.1 M phosphate buffer, pD 7.15.

Equations **2** and **3** can be rewritten in the form of equations **4** and **5** respectively:

$$2K_a[A_0]f_m^2 + (1 + K_l)f_m = 1 \qquad \mathbf{4}$$

$$\delta = (\delta_m + \delta_l K_l)f_m + 2\delta_a K_a[A_0]f_m^2 \qquad \mathbf{5}$$

By eliminating the mole fraction f_m from equations **4** and **5**, the following dependence of proton chemical shift on concentration $[A_0]$ of the oligonucleotide in solution is derived:

$$\delta = \delta_a + \frac{2(\delta_0 - \delta_a)}{1 + \sqrt{(1 + 8z[A_0])}}$$

where

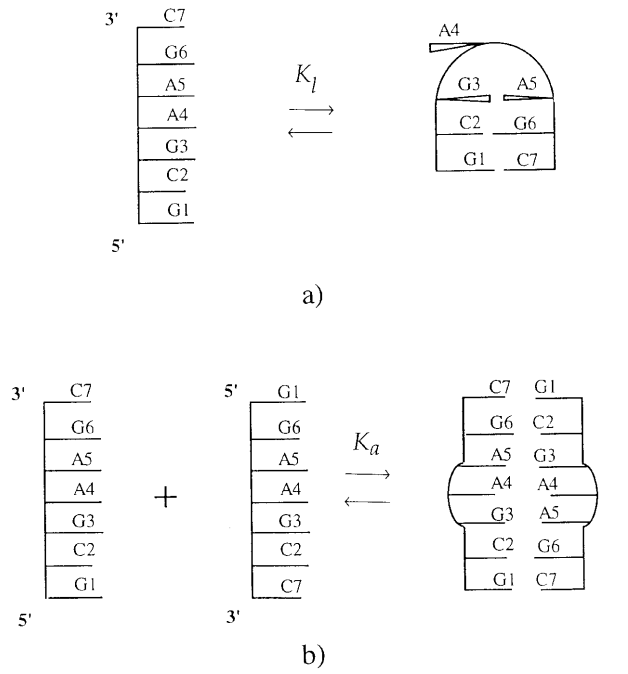


Figure 2. Schematic representation of the formation reactions of (**a**) hairpin and (**b**) bulged duplex from monomer states of the deoxyheptanucleotide 5′-d(GpCpGpApApGpC).

$$\delta_0 = \frac{\delta_m + \delta_l K_l}{1 + K_l} \quad \text{and} \quad z = \frac{K_a}{(1 + K_l)^2} \qquad \mathbf{6}$$

It is seen from equation **6** that the parameter δ_0 corresponds to the limiting value of the concentration dependence of the proton chemical shift when $[A_0] \rightarrow 0$ and the value of z determines the slope and curvature of the $\delta = \delta[A_0]$ dependence.

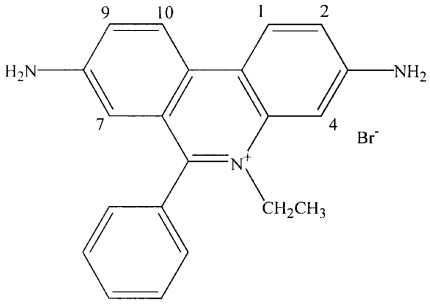
The parameters δ_m, δ_l, δ_a, K_l and K_a in equation **6** and the thermodynamical parameters $\Delta H°$, $\Delta S°$ of formation of hairpin and dimer (reactions **1**) in solution were calculated by using, at the same time, the concentration (at temperatures 298 and 308 K) and temperature dependence of proton chemical shifts of the heptamer in solution. In the calculations the equilibrium constants K_l and K_a were expressed in terms of the corresponding parameters $\Delta H_l°$, $\Delta S_l°$ and $\Delta H_a°$, $\Delta S_a°$:

$$K_l(T) = \exp(\Delta S_l°/R - \Delta H_l°/RT) \qquad \mathbf{7}$$

$$K_a(T) = \exp(\Delta S_a°/R - \Delta H_a°/RT) \qquad \mathbf{8}$$

assuming that the values of $\Delta H°$ and $\Delta S°$ do not depend substantially on temperature in the range studied.

The calculations were carried out by the variational method (26) using the minimization of the joint quadratic discrepancy functions between the experimental and calculated values of δ for nine non-exchangeable protons of the deoxyheptanucleotide d(GCGAAGC) as a function of concentration and temperature: H8(G1), H6(C2), H8(G6), H5(C7), H6(C7), H1′(C2), H1′(G6) and H1′(G7). These protons experience relatively large concentration and temperature changes and the positions of their resonances in the PMR spectra have been determined with sufficient accuracy for the calculations as discussed in our previous work (17,24). In the analysis of the temperature dependence of chemical shifts it was assumed that δ_m and δ_l are



Scheme 1. Structural formula of the phenantridium drug ethidium bromide (EB).

monotonic functions of temperature due to changes in intramolecular base stacking in the different conformational states of the heptamer in solution. Linear and quadratic approximations for $\delta_m(T)$ and $\delta_l(T)$ respectively were used in the calculations (23,25–27). It should be noted that the dependence of δ_m on temperature is also observed for self-complementary and non-self-complementary deoxytetranucleotides of different base sequence (28–31). The results of calculations for different protons of the heptamer d(GCGAAGC) are presented in Table 1.

Table 1. Calculated values of enthalpy $\Delta H°$ (kJ/mol) and entropy $\Delta S°$ [J/(mol K)] of formation of hairpin (A_l) and bulged dimer form of the heptanucleotide 5′-d(GCGAAGC)[a]

Nucleotide proton	$-\Delta H_l°$	$-\Delta S_l°$	$-\Delta H_a°$	$-\Delta S_a°$
H8(A4)	70.0	198.4	195.8	554
H2(A4)	66.1	187.2	185.4	524
H8(G6)	57.7	163.5	122.2	320
H6(C2)	65.3	185.2	143.9	385
H6(C7)	65.7	186.2	169.3	470
H1′(C2)	65.2	184.8	169.7	472
H1′(G6)	58.6	166.2	182.7	522
H5(C2)	64.1	181.7	174.8	490
H5(C7)	66.2	187.4	179.1	502
Mean values	64.3 ± 2.8	182.3 ± 8.3	169.2 ± 16.1	471 ± 53

[a]In 0.1 M phosphate buffer, pD 7.15, T = 298 K.

According to the analysis of the molecular equilibrium in solution summarized in Figure 2 it is assumed that the equilibrium constant for either hairpin or dimer formation (K_l or K_a) is equal for all protons studied. The mean values of these parameters at T = 298 K are respectively K_l = 57.2 ± 4.7 and K_a = 112.6 ± 17.7 × 10^3 l/mol. The calculated value of parameter K_a is larger than the corresponding values for self-complementary tetramers under the same solution conditions (17,18) and similar to the values of the self-association constants of complementary oligonucleotide sequences containing 5 or 6 nt in the chain (23,27); this result may be explained by the possibility of non-Watson–Crick base pairing in the centre of the bulged duplex of the heptamer (Fig. 2) in solution.

Properties of the conformational equilibrium. The relative content of each of the conformational states of the deoxyheptanucleotide d(GCGAAGC) has been calculated from the equilibrium constants as a function of concentration of the heptamer in solution at 298 and 308 K; an example of such calculations at

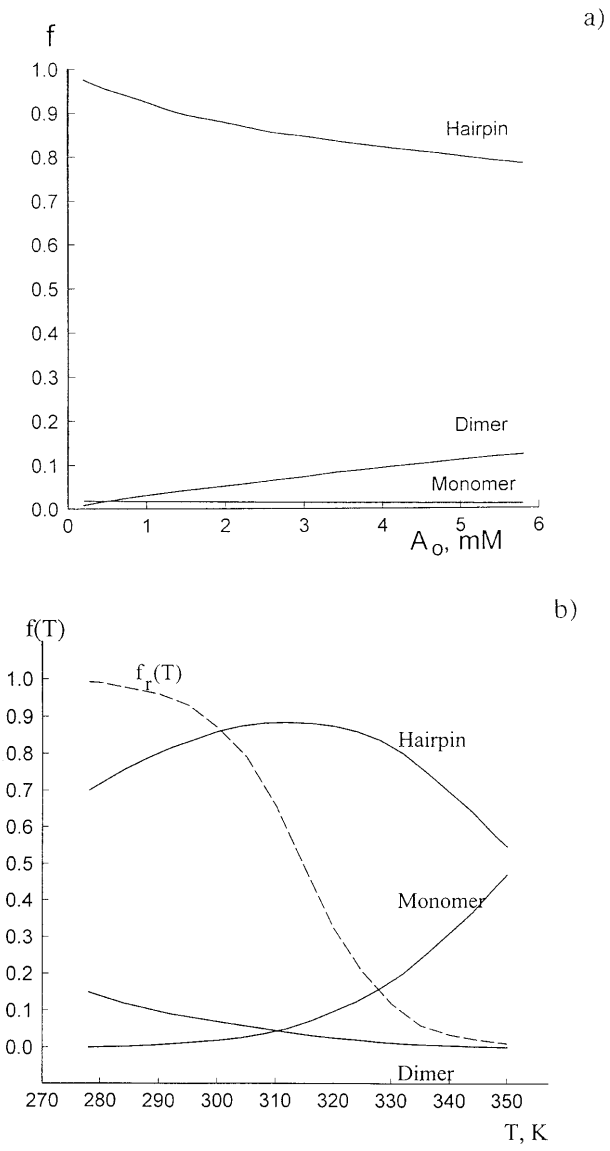


Figure 3. Calculated relative content (f) of different conformers (monomer, hairpin, dimer) as a function of (**a**) concentration ($[A_0]$) at T = 298 K and (**b**) temperature at constant concentration ($[A_0]$ = 2.55 mmol/l) of the deoxyheptanucleotide 5′-d(GpCpGpApApGpC); $f_r(T)$ is the relative content of the dimer compared with the sum of monomer and dimer forms of the heptamer (dashed line).

T = 298 K is presented in Figure 3a. The contribution to the general equilibrium in solution of different conformational states (monomer, hairpin and dimer forms) of the heptamer d(GCGAAGC) is determined not only by the values of equilibrium reaction constants, K_l and K_a, but also by the concentration of oligonucleotide in solution. With increasing concentration of the heptamer there is a small decrease in content of the monomer in solution, whereas the relative amount of dimer increases and the corresponding fraction of the hairpin decreases.

The calculated temperature dependences of the mole fractions of monomer, hairpin and dimer forms of the heptamer are shown in Figure 3b. It is seen that at low temperatures the amount of dimer is relatively high, but decreases with increasing temperature to give a concomitant increase in monomer concentration in

solution. The melting temperature of the dimer, $T_m \approx 314$ K, is easily observed in Figure 3b, where the content of the dimer (f_r, dashed line) at different temperatures is presented relative to the sum of the fractions of monomer and dimer forms in solution, so that the cooperativity of the melting process of the dimer is emphasized more clearly. The fraction of the hairpin in solution has a maximum at ~314 K; at lower temperatures ($T \leq T_m$ of the dimer) the content of the hairpin increases with temperature as the dimer melts, but at $T > T_m$ a monotonic decrease of the mole fraction of the hairpin structure is observed due to its melting at higher temperatures. The calculated melting temperature of the d(GCGAAGC) hairpin is ~348 K, in agreement with previous results (16).

Thus the present NMR analysis reveals the existence of a complex equilibrium of different conformational states which includes monomeric, hairpin and dimeric forms of the deoxyheptanucleotide d(GCGAAGC) in aqueous solution. Qualitatively this kind of behaviour was used to explain the melting experiments by absorption and fluorescence spectroscopy for the related sequence d(GCGAAACGC) (13). As quantitative analysis of NOE intensities is not valid when multiple conformations are present in fast exchange, there are limitations on structure determination of hairpins such as d(GCGAAGC) based solely on NOE measurements (15).

Thermodynamic parameters. The thermodynamic parameters, summarized in Table 1, confirm that formation of both the hairpin and dimer of the deoxyheptanucleotide in aqueous solution is exothermic. The magnitude of the enthalpy of formation of the bulged dimer for the deoxyheptanucleotide d(GCGAAGC) in aqueous solution, $\Delta H_a = -169.2 \pm 16.1$ kJ/mol, is somewhat higher than that calculated theoretically for four G:C base pairs (32). It is likely that the difference reflects some contribution to ΔH_a° of two mismatched base pairs (A·G) in d(GCGAAGC), as suggested previously (15).

The calculated value of $\Delta H_l = -64.3 \pm 1.3$ kJ/mol for formation of the hairpin having a stem with two G:C base pairs is consistent with $\Delta H_l = -79.5$ kJ/mol (7) for a hairpin with a stem containing three G:C base pairs. At the same time these experimental enthalpies are somewhat smaller (~20%) in absolute values than the theoretical values of ΔH_l determined by nearest neighbour analysis (32). However, in the calorimetric study of oligonucleotide sequences containing 13 deoxynucleotides the magnitude of enthalpy, $\Delta H_l = -155$ kJ/mol, for formation of a hairpin with a stem consisting of four G:C base pairs (33,34) is higher in absolute value (calculated relative to 1 bp) than the enthalpy of hairpin formation with a smaller stem size. In the light of the present study this difference may be partially due to some disorder of the G:C base pairing next to the loop when the size of the stem is relatively small and may be partially due to a higher probability of formation of a bulged duplex in longer oligonucleotides; in the latter case there will be a relatively large contribution of the enthalpy of dimerization to the total thermal effect of hairpin and dimer formation of the oligonucleotide in solution.

The calculated entropy of formation of the hairpin, $\Delta S_l = -182.3 \pm 8.3$ J/mol K, obtained in this work is in good agreement with the results for a hairpin with a short stem (13). It is worth noting that for hairpin structures (33,34) where the length of the stem is double that in d(GCGAAGC) the entropy change of hairpin formation is about double the ΔS_l value obtained in the present work. The large negative value of the entropy of formation of the deoxyheptamer dimer in aqueous solution, $\Delta S_a = -471 \pm 53$ J/mol K (Table 1), indicates substantial ordering of solvent in the vicinity of the dimer, which may be due to a rather bulky structure of the bulged duplex having non-complementary purine base pairs in the centre of the deoxyheptamer d(GCGAAGC) (Fig. 2).

Complexation of ethidium bromide to d(GpCpGpApApGpC) in aqueous solution

Experimental observations. The structural formula of the EB molecule (Scheme 1) shows the atom position of all aromatic protons having different NMR resonance peaks used in the drug binding studies. In the 2D NOESY spectra of mixed solutions obtained at different mixing times ($\tau_{m1} = 90$ ms, $\tau_{m2} = 200$ ms) only a few intermolecular cross-peaks between EB and the heptamer protons were observed, as found previously for EB binding to self-complementary deoxytetranucleotides (17,18). In the present case only cross-peaks of small intensities between the meta/para protons of the EB phenyl ring and cytosine deoxyribose protons H1′(C2) of the heptamer could be detected. Down-field shifts of phosphate resonances have been observed at the sites of intercalation of drug molecules in an oligonucleotide duplex (35), providing evidence of significant conformational perturbations of the oligomer secondary structure on intercalative binding of the drug. An indication of the preferred site of binding of EB with the heptamer is given by comparison of the 2D ^{1}H-^{31}P heteronuclear shift correlation spectra of the deoxyheptanucleotide d(GCGAAGC) in solution and when mixed with EB; there are larger down-field shifts of the two phosphate resonances G(1)pC(2) and G(6)pC(7) on drug binding compared with other phosphate signals in the oligonucleotide chain. Taking into consideration the results of homonuclear (2D NOESY) and heteronuclear ^{1}H-^{31}P 2D NMR spectroscopy, it may be concluded that the prefered sites of drug binding with the heptamer are the GC sites of the 5′-d(GCGAAGC) sequence.

In order to determine quantitatively the structures and thermodynamics of complexation of EB with the deoxyheptanucleotide 5′-d(GpCpGpApApGpC), the chemical shift dependence of the EB chromophore protons was measured as a function of both the concentration of the heptamer (e.g. Fig. 4a at 298 K) and temperature (Fig. 4b).

Model of complexation and analysis. It was considered that the basic scheme for molecular complexation had to take into account at least the complexation of one EB molecule with both the single- and double-stranded forms of the heptamer and binding of one and two drug molecules to the hairpin, as the relative content of the hairpin is predominant in solution at all temperatures and concentrations, as shown in Figure 3. The following equilibria were considered for quantitative analysis of complexation of EB with the heptamer:

$$\mathrm{D} + \mathrm{D} \overset{K_d}{\Leftrightarrow} \mathrm{D}_2 \quad \text{(a)}$$

$$\mathrm{A} \overset{K_l}{\Leftrightarrow} \mathrm{A}_l \quad \text{(b)} \qquad \mathrm{A} + \mathrm{A} \overset{K_a}{\Leftrightarrow} \mathrm{A}_2 \quad \text{(c)}$$

$$\mathrm{A} + \mathrm{D} \overset{K_l}{\Leftrightarrow} \mathrm{AD} \quad \text{(d)} \qquad \mathrm{A}_l + \mathrm{D} \overset{K_2}{\Leftrightarrow} \mathrm{A}_l\mathrm{D} \quad \text{(e)} \qquad \textbf{9}$$

$$\mathrm{A}_2 + \mathrm{D} \overset{K_3}{\Leftrightarrow} \mathrm{A}_2\mathrm{D} \quad \text{(f)} \qquad \mathrm{A}_l\mathrm{D} + \mathrm{D} \overset{K_4}{\Leftrightarrow} \mathrm{A}_l\mathrm{D}_2 \quad \text{(g)}$$

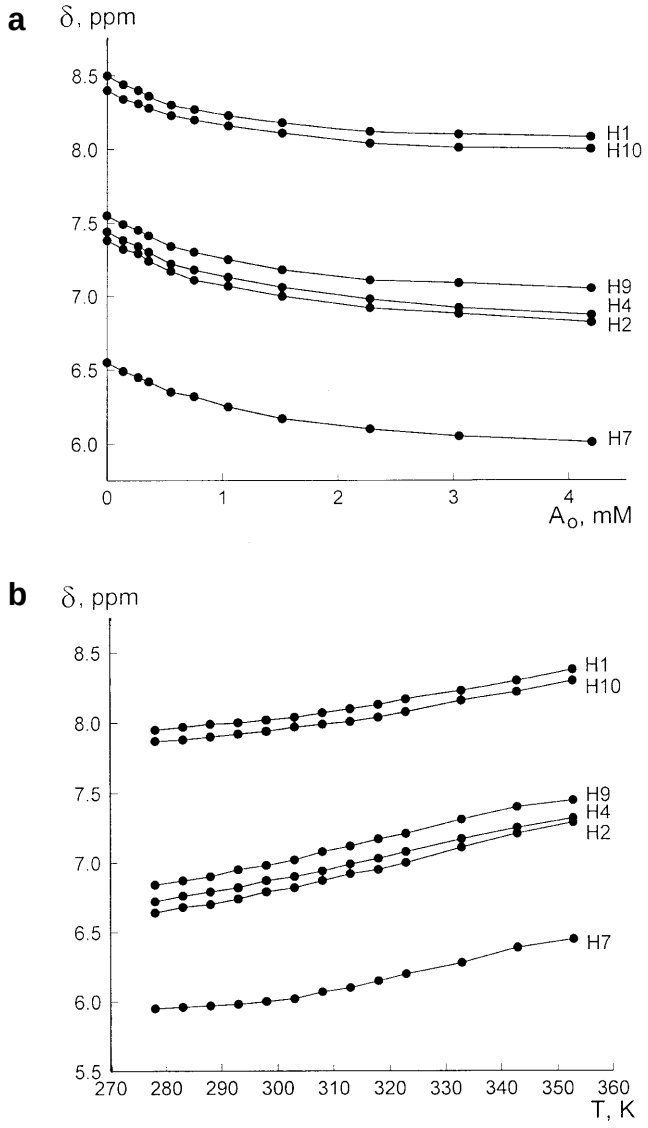


Figure 4. Proton chemical shifts of EB on complexation with the heptanucleotide 5′-d(GpCpGpApApGpC). (**a**) Dependence on concentration of the heptamer at constant drug concentration [D_0] = 1.15 mmol/l in solution, T = 308 K. (**b**) Temperature dependence at [D_0] = 1.15 mmol/l (drug) and A_0 = 2.28 mmol/l (heptamer).

Chemical shifts were calculated using an additive model (the validity of such a model was discussed above):

$$\delta = \delta_m F_m + \delta_d F_d + \sum_{i=1}^{4} \delta_i F_i \qquad \mathbf{10}$$

where δ_m, δ_d, δ_1–δ_4 are the proton chemical shifts of EB in the monomer, dimer, in 1:1 complexes with monomer (AD) and hairpin (A_1D), in the 1:2 complex with heptamer dimer A_2D and in the 2:1 complex with the hairpin (A_1D_2); F_m, F_d, F_1–F_4 are mole fractions of EB in the monomer, dimer and in the above-mentioned complexes respectively. The values of δ_m, δ_d and K_d = 155 ± 7 l/mol at T = 298 K were determined previously (25) from investigations of EB self-association using the dependence on concentration of drug proton chemical shifts in the same solvent system (0.1 mol/l phosphate buffer, pD 7.1).

The dependence of the equilibrium association constants $K(T)$ on temperature was expressed in terms of the thermodynamic parameters $\Delta H°$, $\Delta S°$ using relations equivalent to **7** and **8**. Thus, taking into account the mass law equations and the mass conservation law for reactions **9**, the observed chemical shift **10** is a function of the following parameters: δ_1–δ_4, $\Delta H_1°$–$\Delta H_4°$, $\Delta S_1°$–$\Delta S_4°$. It is reasonable to solve such a multiparametric problem by the method of successive complication of the model (17,18). The computational procedure to determine the parameters of the model by minimizing the discrepancy function between the experimental and calculated values of chemical shifts for the six non-exchangeable protons of the EB chromophore is described in detail in previous work (17,24,26). In the calculations the data on the temperature and concentration (at two different temperatures) dependence of EB proton chemical shifts in mixed solution have been processed jointly. In the present calculations other schemes of complex formation have also been analysed, taking into consideration, for example, additional reactions of a second drug molecule binding with the monomer and dimer forms of the deoxyheptanucleotide. However, the values of the equilibrium constants for such reactions could not be determined with great accuracy due to the relatively small content of these complexes in solution, even though the experimental data are sufficient to calculate these contributions.

Properties of EB complexation with the heptamer. The calculated values of the limiting chemical shifts of EB protons (δ_1–δ_4) in the complexes, equilibrium constants (K_1–K_4) and thermodynamic parameters of formation of different complexes ($\Delta H_1°$–$\Delta H_4°$, $\Delta S_1°$–$\Delta S_4°$) are summarized in Tables 2 and 3. It is seen from Table 2 that the equilibrium constants K_2 and K_4 of complex formation between EB and the hairpin are significantly smaller than the constants K_1 and K_3 for drug complexation reactions with monomer and dimer forms of the heptanucleotide in solution at T = 298 K. The results indicate the relative difficulty of binding of the EB aromatic ligand to the very compact hairpin structure of d(GCGAAGC) compared with binding to the conformationally more flexible structures of the monomer and dimer forms of the heptamer. It is worth noting that the equilibrium constants K_1 and K_3 of EB complexation with the monomer and dimer forms of the heptamer are in good agreement with the corresponding constants for binding of EB with the GC site of single-stranded and double-stranded forms of the deoxytetranucleotide d(AGCT) (18). The value of the constant $K_4 = 0.51 \pm 0.16 \times 10^3$ l/mol shows that the probability of 2:1 EB–hairpin complex formation in solution is substantially smaller than that of the 1:1 complex ($K_2 = 3.9 \pm 1.4 \times 10^3$ l/mol), i.e. binding of a second drug molecule is less favourable compared with the first one. If it is assumed that the first molecule intercalates into the GC site of the stem, then the second drug molecule, in all probability, interacts with the loop of the hairpin and is stabilized by stacking interactions with the guanine (G3) base and unpaired adenine (A4) base. These results are in agreement with conclusions drawn in previous work (33,34) that binding constants of EB with the stem in the hairpins d(GCGCT-$_n$GCGC) are approximately an order of magnitude greater than with loops (T_n) containing different numbers (n = 3, 5 or 7) of thymines in the chain and that drug–loop complexes are stabilized by stacking interactions between thymine base rings and the drug chromophore.

Table 2. Calculated values of limiting proton chemical shifts and equilibrium constants for complexes of EB and the heptamer 5′-d(GpCpGpApApGpC) at T = 298 K

EB proton	δ_1 (p.p.m.)	δ_2 (p.p.m.)	δ_3 (p.p.m.)	δ_4 (p.p.m.)	δ_m (p.p.m.)	K_1 (10^3 l/mol)	K_2 (10^3 l/mol)	K_3 (10^3 l/mol)	K_4 (10^3 l/mol)
H1	7.89	8.10	7.93	7.48	8.68				
H10	7.69	7.99	7.53	7.55	8.62				
H9	7.31	6.97	6.51	6.61	7.66	11.7 (± 1.7)	3.9 (± 1.4)	17.7 (± 1.9)	0.51 (± 0.16)
H4	7.18	6.70	6.63	6.40	7.55				
H2	7.16	6.65	6.49	6.66	7.48				
H7	6.37	5.92	5.49	5.26	6.67				

Table 3. Thermodynamic parameters of EB complexation with the deoxyheptanucleotide 5′-d(GpCpGpApApGpC)

Reaction	$-\Delta G°$ (298 K) (kJ/mol)	$-\Delta H°$ (kJ/mol)	$-\Delta S°$ (298 K) (J/mol/K)
K_1			
A + D ↔ AD	23.2 ± 0.3	60.7 ± 7.5	125 ± 24
K_2			
A_l + D ↔ A_lD	20.5 ± 0.7	48.3 ± 7.1	93 ± 22
K_3			
A_2 + D ↔ A_2D	24.2 ± 0.3	69.5 ± 3.8	152 ± 13
K_4			
A_lD + D ↔ $A_l D_2$	15.4 ± 0.7	37.9 ± 9.4	75 ± 34

Using the values of the equilibrium constants (Table 2), the relative content of each of the molecular complexes in solution has been calculated as a function of r (= $[A_0]/[D_0]$, the ratio of initial concentrations of heptanucleotide and drug) at 298 and 308 K. An example is shown at 298 K in Figure 5a, where it can be seen that the contribution of different types of complexes to the general equilibrium in solution is determined not only by the values of equilibrium reaction constants, but also by the value of r, as shown previously for drug binding with self-complementary and non-self-complementary deoxytetranucleotides (17,18,31). The relative amount of the 1:1 complex of EB with the hairpin (A_lD) grows rather quickly with increasing r in the range $0 \leq r \leq 1.5$, because of a relatively large content of hairpin structures compared with other conformational states of the heptamer in solution at lower concentrations (Fig. 3a) and temperatures (Fig. 3b). There is a maximum of the 1:1 A_lD complex at $r \approx 2$ and then a gradual decrease at higher r values, although this complex is always predominant. At the same time, growth of the fraction of the 1:2 complex of the drug with the dimer (DA_2) becomes significant at $r > 1.5$, as is clearly seen in Figure 5a (at T = 298 K). However, due to the relatively low melting temperature of the complex of the drug with the bulged duplex, an increase in temperature of 10°C (T = 308 K) leads to a significantly smaller amount of the 1:2 complex of the drug with the dimer in solution. The relative amount of the 1:1 complex of the drug with the monomer structure of the heptamer (AD) remains small over the whole range of r studied, whereas characteristic maxima of the concentration curves for 2:1 EB–hairpin complexes are observed at r values corresponding to the stoichiometric relations of heptamer and drug concentration for complex formation, i.e. r = ~0.66 for $A_l D_2$ complexes.

The calculated temperature dependencies of the mole fractions of different types of EB–heptamer complexes are shown in Figure 5b. It is seen that at lower temperatures practically all of the drug is in the complexed state, with either the hairpin (A_lD) or dimer (DA_2) forms of the heptamer. Melting of the 1:2 complex of EB with the bulged duplex form of the heptamer takes place at relatively low temperatures and so the relative content of the DA_2 complex decreases rapidly and is not significant at $T \geq$ ~314 K. At the same time there is an increase with temperature in the relative content of the 1:1 complex of EB with the single-strand form of the monomer due to the increase in the fraction of the monomer form of the drug and single-strand form of the heptamer in solution. There is an initial growth with temperature ($T <$ ~314 K) of the relative content of the 1:1 complex of the drug with the hairpin (A_lD), as it is thermally a more stable structure than DA_2, but after a maximum at ~300 K there is a monotonic decrease in the mole fraction of such complexes due to their melting at high temperatures. It should be noted that the melting temperature T_m of the 1:1 drug–hairpin complex is slightly lower (~3–4°C) than the corresponding T_m value for the hairpin in solution. This behaviour is the opposite of that observed for binding of aromatic ligands with double-stranded DNA and self-complementary oligonucleotide sequences, which leads to an increase in the melting temperature of these structures (36). This result suggests that intercalation of a ligand into a very compact and 'tense' hairpin structure to some extent 'relaxes' it, leading to a decrease in the melting temperature of the hairpin on drug binding.

Structures of the complexes

Homonuclear (^{1}H-^{1}H NOESY) and heteronuclear (^{1}H-^{31}P HMBC) NMR experiments indicated qualitatively that EB binds preferentially to the GC base pairs found in the stem of the hairpin and in the duplex form of the heptamer. A more detailed analysis of the structures of the complexes of EB with d(GCGAAGC) has been made using the calculated limiting values of EB proton chemical shifts (Table 2), which enable, in principle, the structures to be determined for each of the four complexes AD (from δ_1), A_lD (δ_2), DA_2 (δ_3) and $A_l D_2$ (δ_4); in practice, it is expected that limiting chemical shifts lead to determination of reliable structures for 1:1 EB binding to GC base pairs in the stem of the hairpin (A_lD) and bulged duplex (DA_2).

The induced chemical shifts of drug protons ($\Delta\delta_i = \delta_m - \delta_i$) are in the range 0.3–1.4 p.p.m. up-field, indicating substantial ring current shielding effects of nucleotide bases on all the EB protons; this is consistent with intercalation of the drug chromophore between the base planes of the heptamer. The three-dimensional structures of the complexes of EB with the deoxyheptanucleotide were calculated using a modified model of equivalent magnetic dipoles (17), which approximates the quantum mechanical isoshielding curves (37), together with the algorithm and software for transformation of coordinates, kindly provided by V.I.Poltev (38,39). The sequence of transformations of coordinates in the double helix consisted of variation of the parameters ω (propeller), κ (buckle), τ (tilt), ρ (roll), D_x (shift), D_y (slide), Ω (twist) and D_z (rise) (40). In this work shielding of the EB protons

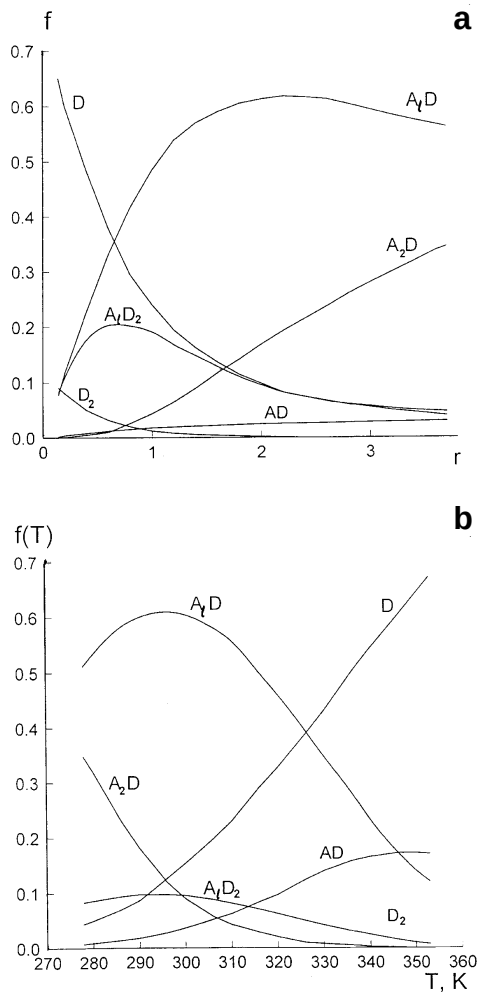


Figure 5. Calculated relative content (*f*) of drug (D) and the different complexes with the heptanucleotide 5′-d(GpCpGpApApGpC) in solution as a function of (**a**) r (= $[A_0]/[D_0]$), the ratio of initial concentrations of heptanucleotide and drug at T = 298 K, and (**b**) temperature in solution at $[D_0]$ = 1.15 mmol/l and $[A_0]$ = 2.28 mmol/l, i.e. r = 1.

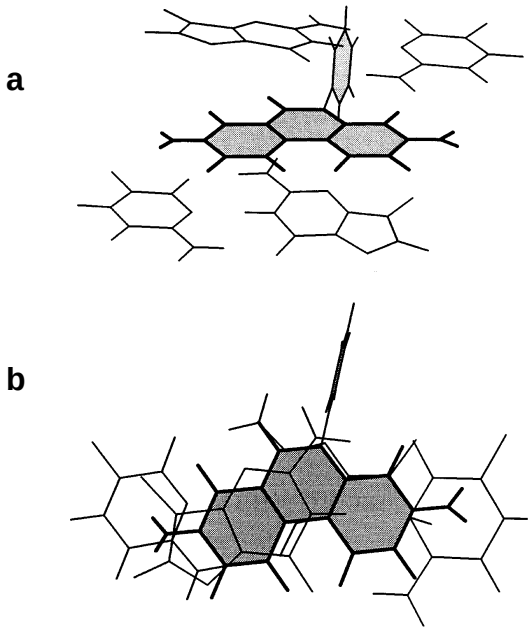


Figure 6. Calculated structure of the fragment of the 1:1 complex of EB and hairpin form of the deoxyheptanucleotide 5′-d(GpCpGpApApGpC) corresponding to intercalation of the drug in the d(GC) site of the stem of the hairpin: (**a**) side view; (**b**) view of the molecular complex looking perpendicular to the planes of the base pairs and the drug chromophore. The drug chromophore and phenyl ring are shaded.

by adjacent base pairs, G:C and C:G, in the intercalated complex in the stem of the hairpin and in the dimer form (GC sites) has been calculated as a function of the conformational parameters of the double helix. The computational procedure for determination of the conformational parameters of the intercalated complex has been described in detail elsewhere (17).

As an example, the most favourable structure calculated for intercalation of EB into the 5′-d(GC) site of the stem of the hairpin is presented in different projections in Figure 6. Spatial representation of the structure was obtained using Mathematica 2.2 software (Wolfram Research Ltd). The intercalated complex is characterized by the following helix parameters: D_z = 6.96 Å, Ω = 6°, ω = 4°, τ = 5°, D_x = –0.04 Å, D_y = 1.25 Å, ρ = 3.5°, κ = 2.0°. The drug chromophore is situated perpendicular to the helix axis at equal distances (3.48 Å) from the upper and lower base pairs in the 5′-d(GC) site of the stem of the hairpin. Unwinding of the adjacent base pairs at the intercalation site is 30° (Ω = 6°), which is somewhat higher than the 25–26° for the 1:2 complexes of EB with duplexes of self-complementary deoxytetranucleotides in solution (17,18), X-ray crystal data of the EB–d$(CpG)_2$ complex (41,42) and the results of theoretical investigations for regular double-helical oligonucleotide sequences (43,44); the difference is probably due to the rather tense hairpin structure being opened up on intercalation of the drug.

The most favourable structure of the fragment of the 1:2 complex of EB intercalated into the 5′-(GC) site of the heptanucleotide dimer is characterized by the following helix parameters: D_z = 6.9 Å, Ω = 10°, ω = 5°, τ = 16°, D_x = 0.17 Å, D_y = 0.05 Å, ρ = 7.5°, κ = 7.0°. Unwinding of the adjacent base pairs in such a complex is 26° (Ω = 10°), which is in good agreement with X-ray crystal data (41,42), solution studies of drug binding to GC sites of self-complementary oligonucleotides (17,18) and the results of theoretical investigations (43,44).

Thermodynamics of complex formation

The thermodynamic parameters presented in Table 3 show that all the reactions of complex formation of EB with the monomer, hairpin and dimer forms of the heptanucleotide (reactions **9**d, e, f and g) are exothermic. The enthalpy values, ΔH_2° = –48.3 ± 7.1 kJ/mol and ΔH_4° = –37.9 ± 9.4 kJ/mol, obtained by NMR for EB binding with the stem and loop of the hairpin d(GCGAAGC) respectively are in good agreement with the results obtained by optical and calorimetric investigations on drug complexation with hairpins d(GCGCT$_n$GCGC), containing four G:C base pairs in the stem and different numbers of thymine residues in the loop (33,34). The enthalpy of complex formation of EB with the bulged duplex of the heptamer 5′-d(GCGAAGC), ΔH_3° = –69.5 ± 3.8 kJ/mol, agrees within error limits with the enthalpy of drug binding to self-complementary deoxytetranucleotides under the same experimental conditions (17,18). The formation of mismatched base pairs (A·G) in the heptamer dimer, considered

previously (15), may influence the energetics of drug binding. The enthalpy of complexation of EB with the monomer form of the deoxyheptanucleotide, $\Delta H_1^\circ = -60.7 \pm 7.5$ kJ/mol, turned out to be substantially lower than that found for binding of EB with the monomer form of self-complementary deoxytetranucleotides (17,18). The differences may be due to the distinctive features of the primary structure of the deoxyheptanucleotide d(GCGAAGC), which contains a fragment of four purine nucleotides (GAAG) in the centre of the sequence, expected to have a relatively small affinity for drug binding (31). As the enthalpy of drug complexation with the monomeric form of the deoxyheptanucleotide is an average over a number of possible binding sites, the existence in the oligonucleotide of a fragment with identical types of bases (i.e. purine-purine or pyrimidine-pyrimidine), which are characterized by small drug binding affinity (31), may decrease the effective values of thermodynamic parameters of complex formation.

The entropies of complexation of one or two EB molecules with the hairpin of the deoxyheptanucleotide d(GCGAAGC), $\Delta S_2^\circ = -93 \pm 22$ J/mol K and $\Delta S_4^\circ = -75 \pm 34$ J/mol K, are substantially smaller in absolute value than $\Delta S_1^\circ = -125 \pm 24$ J/mol K and $\Delta S_3^\circ = -152 \pm 13$ J/mol K for drug binding with the monomer and dimer forms of the heptamer (Table 3). It is likely that hydrophobic interactions play a more significant role in binding of an aromatic ligand to the compact hairpin structure compared with the more flexible monomeric and dimeric forms of deoxyheptanucleotide in aqueous solution; such hydrophobic interactions due to transfer of the drug molecule from solvent to the intercalation site result in positive contributions to entropy and hence lead to entropy change for ΔS_2° and ΔS_4° that are smaller than ΔS_1° and ΔS_3°.

In conclusion it should be emphasized that:

(i) NMR analysis of the conformational and drug binding properties of the deoxyheptanucleotide d(GCGAAGC) indicate that there is an equilibrium of molecules in solution, including different conformations of the heptamer (in single-stranded, hairpin and dimer forms) and their complexes with an aromatic ligand such as ethidium bromide (EB);

(ii) the method developed for analysing the concentration and temperature dependencies of NMR experimental parameters of drug–nucleic acid complexation enables the contributions of different reactions of complex formation to be differentiated, as well as determination of the separate enthalpy and entropy effects of formation of each complex in solution;

(iii) the quantitative analysis shows that the relative amount of each molecular complex in solution depends on the ratio of the initial concentrations of heptamer and drug and on the solution temperature;

(iv) the NMR analysis also provides limiting chemical shifts from which the structures of the intercalated complexes of EB with the GC sites of the hairpin and dimer forms of the deoxyheptanucleotide d(GpCpGpApApGpC) can be calculated in terms of different conformational parameters of the helix.

ACKNOWLEDGEMENTS

We thank the University of London Intercollegiate Research Service for access to the 500/600 MHz NMR facilities (at Birkbeck and Queen Mary Colleges) and the Stern Bequest for purchase of the oligonucleotide.

REFERENCES

1 Chastain,M. and Tinoco,I.,Jr (1991) *Prog . Nucleic Acid Res. Mol. Biol*., **41**, 131–177.
2 Rosenberg,M. and Court,D. (1979) *Annu. Rev. Genet*., **13**, 319–351.
3 Wells,R.D., Goodman,T.C., Hillen,W., Hom,G.T., Klein,R.D., Larson,J.E., Muller,U.R., Neuendorf,S.K., Panayotatos,N. and Stirdivan,S.M. (1980) *Prog. Nucleic Acid Res. Mol. Biol*., **25**, 167–267.
4 Arai,K., Low,R., Kobori,J., Shlomai,J. and Korberg,A. (1981) *J. Biol. Chem*., **256**, 5273–5280.
5 Elias,P. and Lehman,I.R. (1988) *Proc. Natl. Acad. Sci. USA*. **85**, 2959–2963.
6 Cowing,D.W., Bardwell,J.C.A., Craig,E.A., Woolford,C., Hendrix,R.W. and Gross,C.A. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 2679–2683.
7 Chattopadhyaya,R., Greskowiak,K. and Dickerson,R.E. (1990) *J. Mol. Biol*., **211**, 189–210.
8 Pramnik,P., Kanhouwa,N. and Kan,L.-S. (1988) *Biochemistry*, **27**, 3024–3031.
9 Xodo,L.E., Manzini,G., Quadrifoglio,F., van der Marel,G. and van Boom,J.H.J. (1991) *Nucleic Acids Res*., **19**, 1505–1511.
10 Uhlenbeck,O.C. (1990) *Nature*, **346**, 613.
11 SantaLucia,J., Kierzek,R. and Turner,D.H. (1992) *Science*, **256**, 217–219.
12 Paner,T.M., Amaratunga,M. and Benight,A.S. (1992) *Biopolymers*, **32**, 1715–1734.
13 Slama-Schwok,A., Teulade-Fichou,M.-P., Vegneron,J.-P., Taillandier,E. and Lehn,J.-M. (1995) *J. Am. Chem. Soc*., **117**, 6822–6830.
14 McMurray,C.T., Wilson,W.D. and Douglass,J.O. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 666–670.
15 Hirao,I., Kawai,G., Yoshizawa,S., Nishimura,Y., Ishido,Y., Watanabe,K. and Miura,K. (1994) *Nucleic Acids Res*., **22**, 576–582.
16 Hirao,I., Nishimura,Y., Tagawa,Y., Watanabe,K. and Miura,K. (1992) *Nucleic Acids Res*., **20**, 3891–3896.
17 Davies,D.B. and Veselkov,A.N. (1996) *J. Chem. Soc. Faraday Trans*., **92**, 3545–3557.
18 Davies,D.B., Karawajew,L. and Veselkov,A.N. (1996) *Biopolymers*, **38**, 745–757.
19 Bianchi,M.E., Beltrame,M. and Paonessa,G. (1989) *Science*, **243**, 1056–1059.
20 Duckett,D.R., Murchie,A.I.H., Bhattacharya,A., Clegg,R.M., Diekmann,S., von Kitzig,E. and Lilley,D.M.J. (1992) *Eur. J. Biochem*., **207**, 285–295.
21 Bresloff,J.L. and Crothers,D.M. (1981) *Biochemistry*, **20**, 3547–3553.
22 Bax,A. and Summers,M.F. (1986) *J. Am. Chem. Soc*., **108**, 2093–2094.
23 Freier,S.M., Albergo,D.D. and Turner,D.H. (1983) *Biopolymers*, **22**, 1107–1131.
24 Davies,D.B., Djimant,L.N. and Veselkov,A.N. (1994) *Nucleosides Nucleotides*, **13**, 637–655.
25 Davies,D.B., Djimant,L.N. and Veselkov,A.N. (1996) *J. Chem. Soc. Faraday Trans*., **96**, 383–390.
26 Veselkov,A.N., Djimant,L.N., Karawajew,L.S. and Kulikov,E.L. (1985) *Studia Biophys*., **106**, 171–180.
27 Petersheim,M. and Turner,D.H. (1983) *Biochemistry*, **22**, 256–263.
28 Veselkov,A.N., Davies,D.B., Djimant,L.N. and Parkes,H.G. (1993) *Biophysics*, **38**, 171–180.
29 Veselkov,A.N., Djimant,L.N., Kodintsev,V.V., Lisyutin,V.A., Parkes,H.G. and Davies,D.B. (1995) *Biophysics*, **40**, 283–291.
30 Veselkov,A.N., Davies,D.B., Djimant,L.N., Parkes,H.G. and Shipp,D. (1991) *Biopolym. Cell*, **7**, 15–23 (in Russian).
31 Davies,D.B., Baranovsky,S.F. and Veselkov,A.N. (1997) *J. Chem. Soc. Faraday Trans*., **93**, 1559–1572.
32 Breslauer,K.J., Frank,R., Blocker,H. and Marky,L.A. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 3746–3750.
33 Rentzeperis,D., Alessi,K. and Marky,L.A. (1993) *Nucleic Acids Res*., **21**, 2683–2689.
34 Rentzeperis,D., Medero,M. and Marky,L.A. (1995) *Bioorg. Med. Chem*., **3**, 751–759.
35 Searle,M.S. and Lane,A.N. (1992) *FEBS Lett*., **297**, 292–296.
36 Gale,E.E., Cundliffe,E., Reynolds,P.E., Richmond,M.N. and Waring,M.J. (1981) *The Molecular Basis of Antibiotic Action*. John Wiley, Chichester, UK.
37 Giessner-Prettre,C. and Pullman,B. (1987) *Q. Rev. Biophys*., **20**, 113–172.
38 Poltev,V.I. and Teplukhin,A.V. (1987) *Mol. Biol*., **21**, 102–115.
39 Poltev,V.I. and Teplukhin,A.V. (1989) *Int. J. Quant. Chem*., **35**, 91–102.
40 Dickerson,R.E. (1989) *J. Biomol. Struct. Dyn*., **6**, 627–634.
41 Sobell,H.M., Tsai,C.-C., Jain,S.C. and Gilbert,S.G. (1977) *J. Mol. Biol*., **144**, 333–365.
42 Jain,S.C. and Sobell,H.M. (1984) *J. Biomol. Struct. Dyn*., **1**, 1161–1177.
43 Lybrand,T. and Kollman,P. (1985) *Biopolymers*, **24**, 1863–1879.
44 Chen,K.-X., Gresh,N. and Pullman,B. (1987) *Biopolymers*, **26**, 831–848.