Acta Chemica Scandinavica, 1998: 52: 641–650
Printed in UK – all rights reserved

Copyright © Acta Chemica Scandinavica 1998

ACTA
CHEMICA SCANDINAVICA
ISSN 0904–213X

Bis-Intercalation of Homodimeric Thiazole Orange Dyes in Selective Binding Sites of DNA Studied by ^{1}H NMR Spectroscopy

Lisbeth K. Jensen, Charlotte H. Gotfredsen, Kent Bondensgaard and Jens Peter Jacobsen*

Department of Chemistry, Odense University, Odense M, DK-5230 Denmark

Jensen, L. K., Gotfredson, C. H., Bondensgaard, K. and Jacobsen, J. P., 1998. Bis-Intercalation of Homodimeric Thiazole Orange Dyes in Selective Binding Sites of DNA Studied by ^{1}H NMR Spectroscopy. – Acta Chem. Scan. 52: 641–650. © Acta Chemica Scandinavica 1998.

The thiazole orange dye 1,1′-(4,4,8,8-tetramethyl-4,8-diazoniaundecane-1,11-diyl)bis[4-(3-methyl-2,3-dihydro-1,3-benzothiazol-2-ylidenemethyl)quinolinium] tetraiodide (TOTO) binds to double stranded DNA (dsDNA) in a sequence-selective bis-intercalation. TOTO binds preferentially to oligonucleotides containing a $(5'\text{-CTAG-}3')_2$ binding site. The preference of TOTO for the $(5'\text{-CTAG-}3')_2$ is caused by a delicate balance between various different binding contributions. In order to examine these contributions in further detail we have studied the binding to analogs of the oligonucleotide **1** using the oligonucleotides $d(CGCTAICG)_2$ (**3**) (I = inosine, 2-desaminoguanosine) and $d(CGCUAGCG)_2$ (**4**) (U = uridine, 5-desmethylthymidine).

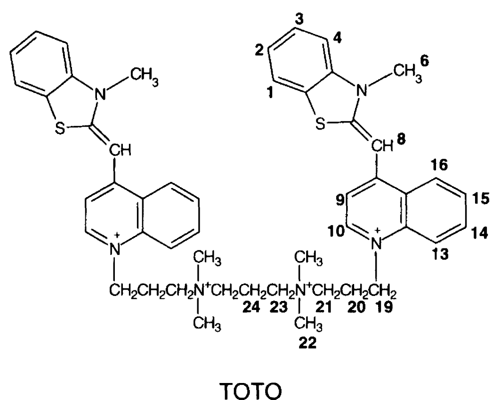
In the modified $(5'\text{-CTAG-}3')_2$ binding sites examined in this work the preferential binding of TOTO was maintained to the $(5'\text{-CTAI-}3')_2$ and the $(5'\text{-CUAG-}3')_2$ sequences. However, the selectivity in the binding to these sites is lower than that of the $(5'\text{-CTAG-}3')_2$ site. These results show that the binding constant is lowered when exchanging a thymidine for a uridine. This implies that the methyl group on T4 in the $(5'\text{-CTAG-}3')_2$ of the **1**–TOTO complex actually contributes to the binding energy by van der Waals interaction. The lowering of the binding constant on going from the $(5'\text{-CTAG-}3')_2$ to the $(5'\text{-CTAI-}3')_2$ site is less and is not easily explained as due to a change in a single contribution to the binding energy. It is probable that local changes between the conformations of the oligonucleotides **1** and **3** cause this effect.

We have used two-dimensional ^{1}H NMR spectroscopy to determine the solution structure of the **4**–TOTO complex in order to compare it with the solution structure of the **1**–TOTO complex. The determination of the structure was based on total relaxation matrix analysis of the NOESY cross peak intensities using the program MARDIGRAS. NOE-derived distance restraints were applied in restrained molecular dynamics calculations. Twenty final structures each were generated for the TOTO complex from both A-form and B-form dsDNA starting structures.

A variety of natural products bind strongly to double stranded DNA (dsDNA) by intercalation. Dimers with two intercalating chromophores linked with a chain of appropriate length (bis-intercalators) show high dsDNA binding affinities, greater by several orders of magnitude than those of the corresponding monomers.[1] Most naturally occurring intercalators consist of rigid planar aromatic chromophores attached to bulky side groups. In some cases it has been shown that such intercalators bind preferentially to specific dsDNA sequences (site selectivity).[2] However, this site selectivity is attributed to either major or minor groove interactions of the bulky side groups with the dsDNA. In general, synthetic compounds consisting of intercalating chromophores without bulky side groups do not exhibit selectivity in binding to various dsDNA oligonucleotides.

Recent work in our laboratory has focused on ^{1}H NMR studies of the binding-mode and sequence selectivity of 1,1′-(4,4,8,8-tetramethyl-4,8-diazoniaundecane-1,11-diyl)bis-[4-(3-methyl-2,3-dihydro-1,3-benzothiazol-2-ylidenemethyl)quinolinium] tetraiodide (TOTO, Scheme 1) with various dsDNA oligonucleotides.[3–7] The synthesis and characterization of TOTO was initiated by the search for non-radioactive DNA stains that are stable under gel electrophoretic conditions.[8–10] TOTO forms highly fluorescent and stable non-covalent complexes

* To whom correspondence should be addressed.



Scheme 1. TOTO numbering scheme.

with double-stranded DNA. The enhancement of the fluorescence quantum yield of TOTO upon complex formation with dsDNA is greater than 1000,[10] and $t_{1/2}$ for the dissociation of the dsDNA–TOTO complex under gel electrophoretic conditions is approximately 11 h.[9] This makes TOTO an excellent choice as a marker with a detection limit of ≈4 pg of DNA.[9]

We have shown that TOTO binds preferentially to oligonucleotides containing a (5′-CTAG-3′)$_2$ binding site in the d(CGCTAGCG)$_2$ (**1**) oligonucleotide. There is at least a 100-fold preference for binding to the (5′-CTAG-3′)$_2$ sequence over any other sites present in the oligonucleotides we have studied containing this sequence.[3,5] TOTO also binds preferentially to a (5′-CCGG-3′)$_2$ sequence in the d(CGCCGGCG)$_2$ (**2**) oligonucleotide but less favorably than to the (5′-CTAG-3′)$_2$ sequence in **1**.[5]

The solution structure of the complex of TOTO bis-intercalated in the (5′-CTAG-3′)$_2$ binding site of the oligonucleotide **1** shows that the sequence selectivity of TOTO is caused by the ability of the TOTO chromophores to adapt to the propeller twist of the nucleobases.[4] This is possible since the benzothiazole ring can twist relative to the quinolinium ring due to the flexible cyanine methine bond in between. The benzothiazole is intercalated between two pyrimidine bases and the quinolinium ring between the two purine bases. In contrast with other intercalators, the sequence selectivity of TOTO arises directly from the interaction of the intercalated chromophores with the nucleobases. Although the linker of TOTO adds significantly to the binding strength it is only a spectator with respect to the sequence selectivity.

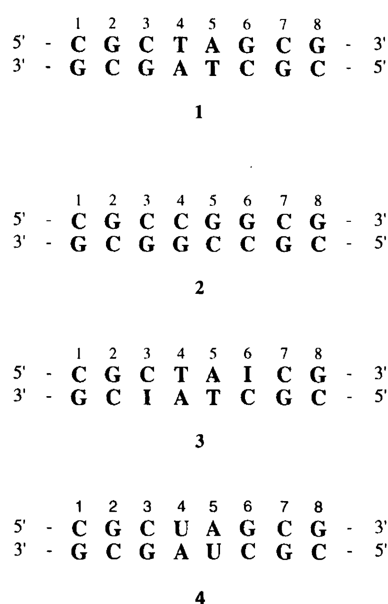
The preference of TOTO for (5′-CTAG-3′) is caused by a delicate balance between various binding contributions. It seems evident that the preference for a pyrimidine–pyrimidine–purine–purine sequence is related to the optimal stacking of the chromophores of TOTO in the intercalation site, but other contributions select between different binding sites of this kind. One of these may be the van der Waals interaction of the T4 CH_3 group and the benzothiazole ring. Others may, in a more obscure way, be related to the variation of the local structure of the different binding sites. In order to examine this problem further we have studied the binding to analogs of the oligonucleotide **1** in which the T–A base pair has been replaced with a U–A base pair (U = uridine, 5-desmethylthymidine) or a G–C base pair with a I–C base pair (I = inosine, 2-desaminoguanosine). Thus, the oligonucleotides d(CGCTAICG)$_2$ (**3**) and d(CGCUAGCG)$_2$ (**4**) were used and compared with the previously reported results on oligonucleotide **1** and **2**.[5]

Experimental

Materials. Purified DNA oligonucleotides were purchased from DNA Technology, Aarhus, Denmark, and used without further purification. The dsDNA oligonucleotides used in this work are shown in Scheme 2.

TOTO is almost insoluble in water. The complexes with dsDNA therefore cannot be made by simple titration of the dsDNA with a TOTO solution. Consequently, the experimental procedure developed was to dissolve TOTO in [$^2H_6$]DMSO and add an appropriate amount of this solution to approximately 5 ml of a diluted solution of the oligonucleotide in water. The addition was done dropwise with rapid stirring, and the reaction mixture was lyophilized on a Speedvac immediately following the addition.

The NMR samples were prepared by dissolving the complexes in 0.5 ml of 10 mM sodium phosphate buffer (pH 7.0) and 0.05 mM sodium EDTA. For experiments carried out in D_2O the solid complex, lyophilized twice from D_2O, was redissolved in 99.96% D_2O (Cambridge Isotope Laboratories). A mixture of 90% H_2O and 10% D_2O (0.5 ml) was used for experiments examining exchangeable protons. The final concentration of the complexes was 2 mM.



Scheme 2. Numbering scheme for the dsDNA duplexes used.

Methods. All NMR experiments were performed at 500 MHz on a Varian Unity 500 spectrometer. NOESY spectra were acquired with various mixing times from 500 to 200 ms in D_2O using 1024 complex points in t_2 and a spectral width of 5000 Hz. Five hundred and twelve t_1 experiments were recorded using the States phase cycling scheme. Normally 64 scans were acquired for each t_1 value. NOE build-up experiments of the **4**–TOTO complex were carried out by collecting NOESY spectra with mixing times of 25, 50, 100, 150, and 200 ms and with 64 scans acquired for each t_1 value. The NOESY spectra were obtained sequentially without removing the sample from the magnet. The TOCSY experiments with mixing times of 30 and 90 ms were carried out in the TPPI mode using 1024 complex points in t_2, 512 t_1 experiments and by acquiring 64 scans for each t_1 value. The NOESY spectra in H_2O were acquired with a spectral width of 12 000 Hz in case of **3** and 10 000 Hz in case of **4** using 2048 complex points and a pulse sequence where the last 90° pulse was replaced by a pulse containing a notch to suppress the solvent signal.[11] The acquired data were processed using FELIX (version 2.35, Biosym Technologies, San Diego). The TOCSY and NOESY spectra were assigned by conventional methods[12–17] as described earlier.[3,4] The integration of the signals in the imino region was performed by the line fitting procedure in FELIX.

Structure calculations. The upper and lower diagonal parts of each of the five assigned NOESY spectra of the NOE build-up experiment were integrated separately with FELIX yielding a total of ten peak intensity sets. The RANDMARDI procedure[18] of the complete relaxation matrix analysis method, MARDIGRAS,[19,20] was used to calculate interproton distance bounds from the resulting integrated peak intensities for use in restrained molecular dynamic (RMD) calculations as described previously.[4]

The distance restraints obtained from the MARDIGRAS calculations were incorporated into an RMD procedure. The RMD and energy minimization calculations were performed using DISCOVER (version 2.9.5) with modified AMBER force-field potentials as described previously.[4] The models were displayed using INSIGHTII (version 97.0) (Biosym Technologies, San Diego). Helix parameters were calculated with the program CURVES 5.1.[21,22]

Results

Oligonucleotide (CGCTAICG)$_2$ (**3**). The fluorescent orange complex between TOTO and the oligonucleotide **3** containing the (5′-CTAI-3′)$_2$ binding site yields a ^{1}H NMR spectrum with sharp lines. Observation of lines for just one DNA strand and one thiazole orange (TO) chromophore shows that the TOTO complex has dyad symmetry. NOESY and TOCSY spectra were used to assign dsDNA and TOTO resonances. Parts of the NOESY spectrum are given in Figs. 1 and 2.

The NOESY spectrum in Fig. 1 exhibits the characteristic features of dsDNA sequential connectivities from aromatic H6/H8 protons to both intra- and inter-residue H1′ and H2′/H2″. However, the sequential connectivities are interrupted at the 5′-C3pT4–3′ and 5′-A5pI6–3′ base pair steps as clear evidence of bis-intercalation. Interruption of the sequential NOE connectivities is also observed in the cross peak pattern of the methyl group of T4 for which the peak to C3 H6 is missing. This is shown in the top part of the NOESY spectrum given in Fig. 1. Aromatic (H6, H5, H8, and H2) and deoxyribose proton (H1′, H2′, H2″, H3′, H4′, and H5′/H5″) resonances of dsDNA in the complex were assigned in the conventional way.[13–17] Some of the assignments are given in Table 1. The NOESY spectrum of the complex in H_2O exhibits the normal Watson–Crick NOE connectivity pattern, again interrupted at the 3–4 and 5–6 base pair steps.[12] Chemical shift values of the labile protons are included in Table 1.

The internal NOE connectivities in the chromophores of TOTO are a distinct feature of the NOESY spectra of the complex. Cross peak patterns connect H16–H15–H14–H13, H8–$CH_3$6–H4–H3–H2–H1 and H9–H10 (Fig. 2), respectively. A very strong cross peak between H8 and H16 and a cross peak pattern connecting H10–H19–H13 establish the assignments of the individual protons and conformation of the TOTO ligand. Resonances in the linker of TOTO were assigned by combined use of TOCSY and NOESY cross peaks. Chemical shift values of TOTO in the complex are given in Table 1. Based on the cross peak pattern in the NOESY spectrum we concluded that the relative conformation of the two ring systems in each chromophore is that indicated in Scheme 1.

A large number of cross peaks between TOTO protons and protons on the oligonucleotide were observed. These intermolecular cross peaks clearly show that TOTO is bis-intercalated in the (5′-CTAI-3′)$_2$ site. Cross peaks in the NOESY spectra unambiguously show that the polypropylene-amine linker is positioned in the minor groove. This is demonstrated by a large number of cross peaks: the H2 of the adenines give cross peaks to the linker protons H21, H22 and H23, the H1′ of the adenines to the linker protons H21, H22, and H23 and the linker H19 protons to the H1′ of the guanines in the intercalation site. The H4′ of the adenines and the guanines in the binding site also show cross peaks to linker protons.

The oligonucleotide (CGCUAGCG)$_2$ (**4**). The 1D NMR spectra of the TOTO complexes with the oligonucleotide **4** also exhibit fairly sharp lines indicating the existence of only one complex. It is possible to follow the characteristic sequential dsDNA connectivity pattern in the NOESY spectra of the **4**–TOTO complex. Similar to the **3**–TOTO complex the sequential NOE connectivities from aromatic H(6)/H(8) protons to H(1′) and

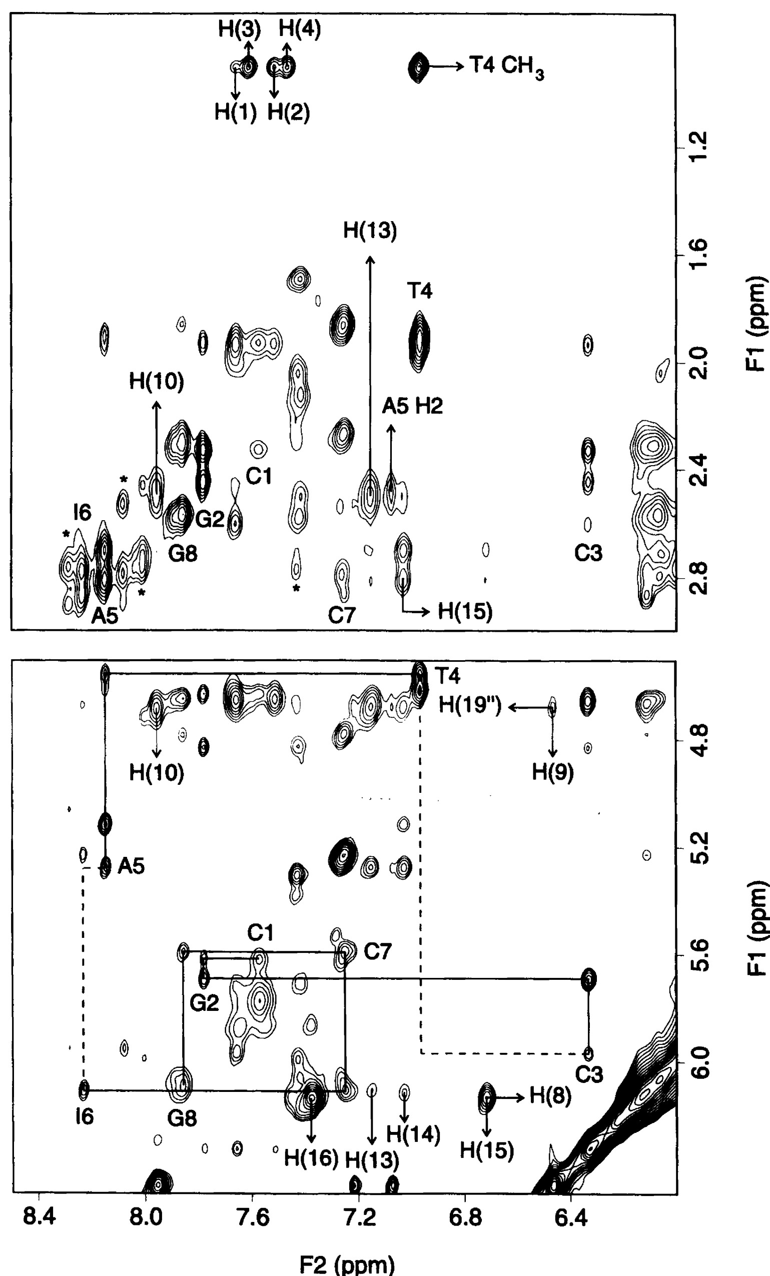


Fig. 1. The H(2′)/H(2″) to aromatic part (top) and the H(1′) to aromatic part (bottom) of the 200 ms NOESY spectrum of the **3**–TOTO complex obtained at 10 °C. The assignments of the H6/H8 resonances for the various nucleotides are indicated together with the resonances of the thymidine methyl protons and some of the TOTO protons (denoted by the letter H). The sequential H1′–H6/H8–H1′ connectivity pathway is indicated by a solid line. The interrupted connectivities at the 5′-C3pT4-3′ and 5′-A5pG6-3′ base pair steps due to intercalation are indicated by a dotted line. A few of the observed cross peaks between TOTO and the oligonucleotide are also marked. The free dsDNA protons are marked with an asterisk.

H(2′),H(2″) are interrupted at the intercalation site showing that TOTO bis-intercalates in the (5′-CUAG-3′)$_2$ site. The cross peak pattern in the NOESY spectra of the **4**–TOTO complex is equivalent to the cross peak pattern of the **3**–TOTO complex. A spectral assignment was performed yielding the chemical shift values given in Table 2.

Mixtures of the oligonucleotides. In order to establish the relative preference of TOTO for the (5′-CTAG-3′)$_2$, the (5′-CCGG-3′)$_2$ and the two other sequences presented in this work, we undertook competition studies between the oligonucleotides. In the competition experiments isolated imino proton resonances from both the two free oligonucleotides and the two complexes must be observed. Therefore experiments were performed where **3** was compared with **1** while **4** was compared with **2**. In an earlier study, **1** was compared with **2**.[5] One-dimensional NMR spectra of the appropriate mixtures were recorded. It was observed that an equilibrium of TOTO

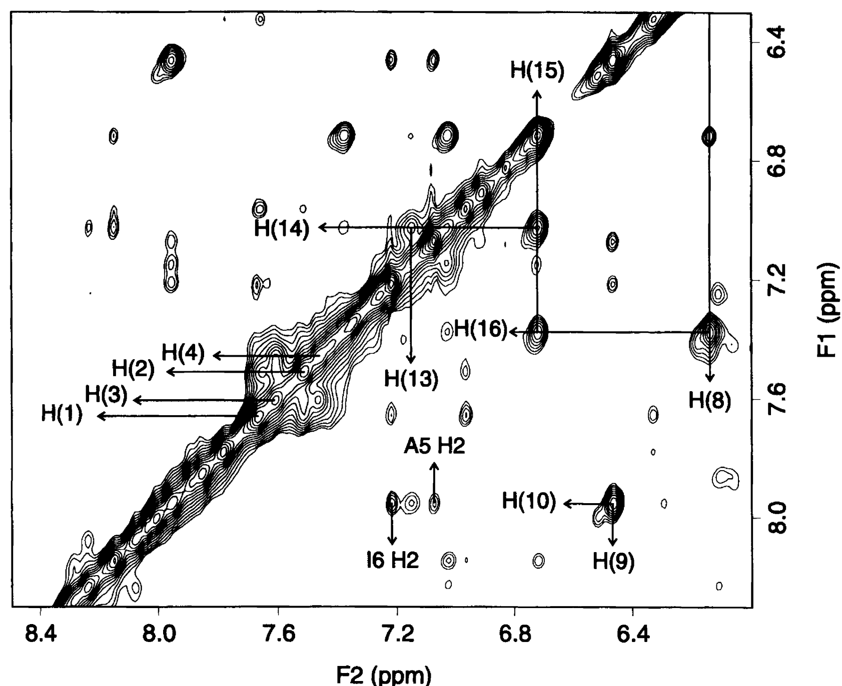


Fig. 2. Aromatic to aromatic part of the 200 ms NOESY spectrum of the **3**–TOTO complex obtained at 10 °C. The sequential paths of H(8)–H(16)–H(15)–H(14)–H(13) and H(1)–H(2)–H(3)–H(4) are indicated together with the connectivities between the two H(9) and H(10) protons. A few of the observed cross peaks between TOTO and the oligonucleotide are also marked.

Table 1. ¹H Chemical shift values (in ppm) of the **3**–TOTO complex at 10 °C compared with the free oligonucleotide given in parentheses. The values are given relative to 2,2-dimethyl-2-silapentane-5-sulfonate (DSS).

	C1	G2	C3	T4	A5	I6	C7	G8
H6/H8	7.56 (7.68)	7.77 (8.01)	6.33 (7.44)	6.96 (7.40)	8.14 (8.28)	8.22 (8.08)	7.25 (7.28)	7.85 (7.91)
H5/H2/CH_3	5.77 (5.93)		4.38 (5.32)	0.90 (1.70)	7.08 (6.94)	7.22 (7.67)	5.23 (5.27)	
H1′	5.62 (5.77)	5.70 (5.99)	5.96 (6.07)	4.56 (5.73)	5.28 (5.95)	6.10 (6.07)	5.60 (5.65)	6.09 (6.15)
H2′	1.92 (2.04)	2.34 (2.73)	1.94 (2.04)	1.91 (2.13)	2.80 (2.77)	2.76 (2.53)	1.86 (1.89)	2.56 (2.59)
H2′	2.33 (2.47)	2.44 (2.77)	2.60 (2.58)	1.91 (2.50)	2.69 (2.90)	2.85 (2.80)	2.27 (2.31)	2.31 (2.34)
H1/H3		12.62 (12.88)		13.67 (13.67)		14.88 (15.42)		
H4/H6			6.71 (7.77)		7.33 (7.70)		8.17 (8.35)	
H4/H6			6.45 (6.66)		6.82 (6.51)		6.61 (6.62)	
	TOTO		TOTO		TOTO		TOTO	
	H1	7.66	H8	6.13	H15	6.72	H21′/21″	3.62/3.74
	H2	7.51	H9	6.47	H16	7.37	CH_3(22)	3.30
	H3	7.61	H10	7.95	H19′	4.19	H23′	3.45
	H4	7.46	H13	7.15	H19″	4.68	H23″	3.52
	CH_3(6)	3.62	H14	7.02	H20′/20″	2.48	H24′/24″	2.48

binding to the two oligonucleotides present was established almost instantaneously.

An example of the imino part of the spectrum in H_2O of such a mixture is given in Fig. 3. Resonance lines from the **4**–TOTO complex, the **2**–TOTO complex and the uncomplexed oligonucleotides were identified in the spectra. The equilibrium (1) between the two complexes has an equilibrium constant defined by eqn. (2).

$$\mathbf{4}\text{–TOTO} + \mathbf{2}(\text{free}) \rightleftarrows \mathbf{2}\text{–TOTO} + \mathbf{4}(\text{free}) \quad (1)$$

$$K = \frac{[\mathbf{2}\text{–TOTO}][\mathbf{4}(\text{free})]}{[\mathbf{4}\text{–TOTO}][\mathbf{2}(\text{free})]} \quad (2)$$

Similar equations are relevant for the other mixture. The values of K were measured from the integral of the imino-proton NMR signals. The values obtained at 10 °C were $K = 1.1(2)$ in the mixture of TOTO with the oligonucleotides **2** and **4**, while $K = 1.8(4)$ in favor of the **1**–TOTO complex in the mixture of TOTO with the oligonucleotides **1** and **3**. This implies that in a mixture of equal amounts of the two oligonucleotides and TOTO the ratio between the **1**–TOTO and the **3**–TOTO complexes is 1.3(2). In a similar mixture of **2**, **4** and TOTO equal amounts of the **2**–TOTO and the **4**–TOTO complexes are formed.

Table 2. ^{1}H Chemical shift values (in ppm) of the **4**–TOTO complex at 25 °C compared with the free oligonucleotide given in parentheses. The values are given relative to DSS.

	C1	G2	C3	U4	A5	G6	C7	G8
H6/H8	7.50 (7.68)	7.74 (8.03)	6.47 (7.47)	7.11 (7.62)	8.15 (8.26)	7.82 (7.70)	7.26(7.31)	7.88 (7.96)
H5/H2/CH_3	5.76 (5.96)		4.52 (5.46)	4.53 (5.53)	7.01 (7.40)		5.31 (5.37)	
H1′	5.65 (5.84)	5.64 (5.99)	5.99 (6.02)	4.68 (5.71)	5.32 (6.05)	5.82 (5.75)	5.69 (5.78)	6.08 (6.20)
H2′	1.83 (2.03)	2.43 (2.76)	1.95 (2.22)	1.95 (2.15)	2.80 (2.77)	2.64 (2.51)	1.84 (1.92)	2.58 (2.40)
H2″	2.27 (2.48)	2.32 (2.81)	2.56 (2.49)	1.95 (2.54)	2.71 (2.92)	2.70 (2.64)	2.27(2.36)	2.33 (2.64)
H1/H3		12.77 (13.05)		13.67 (13.79)		12.04 (12.78)		
H4/H6			7.91 (8.30)		7.31 (—)		8.22 (8.36)	
H4/H6			6.58 (6.59)		6.70 (—)		6.44 (6.49)	

TOTO		TOTO		TOTO		TOTO	
H1	7.60	H8	6.13	H15	6.68	H21′/21″	3.65
H2	7.44	H9	6.49	H16	7.35	CH_3(22)	3.25
H3	7.60	H10	7.96	H19′	4.14	H23′	3.36
H4	7.44	H13	7.14	H19″	4.71	H23″	3.50
CH_3(6)	3.60	H14	7.03	H20′/20″	2.45	H24′/24″	2.45

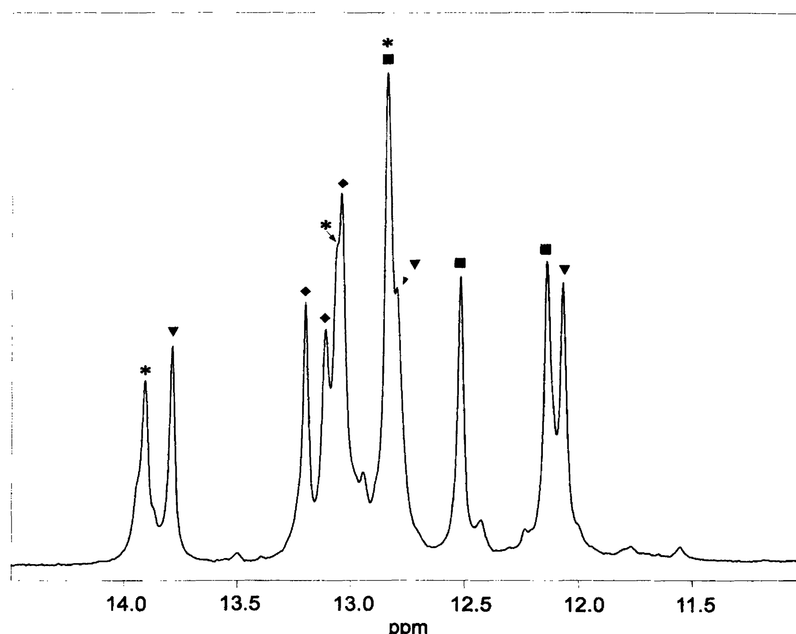


Fig. 3. The imino region of the ^{1}H NMR spectrum obtained in H_2O at 10 °C of a sample consisting of a mixture of the **4**–TOTO (1 μmol) complex and the uncomplexed oligonucleotide **2** (1 μmol). The lines are marked as follows: free oligonucleotide **2** (◆), **2**–TOTO complex (■), free oligonucleotide **4** (✻), and **4**–TOTO complex (▼).

Structure calculations. More than 300 NOE cross peaks were observed in the NOESY spectrum of the **4**–TOTO complex obtained with a mixing time of 200 ms at 25 °C. Some of them resulted predominantly from spin diffusion and were consequently not observed in NOESY spectra with shorter mixing times. The accuracy of the integration of some cross peaks was hampered by spectral overlap. Such cross peaks were therefore not included in the MARDIGRAS calculations. The integrated intensities from 240 NOESY cross peaks were used in the total relaxation matrix analysis. Because of the symmetry in the complex, the number of cross peaks used in the MARDIGRAS calculations was actually doubled to 480. Integrations of the cross peaks in the five NOESY spectra used were done separately for each side of the diagonal. This yielded a total of 4800 measured NOE intensities in 10 sets that were converted into distance restraints using the RANDMARDI procedure. The MARDIGRAS calculations were performed as described earlier[4] and returned 218 interproton distances.

The bounds for use in RMD simulations were determined by combining the results from all the individual MARDIGRAS calculations performed during the RANDMARDI procedure into one set. The distance bounds were calculated individually for each proton pair corresponding to a NOESY cross peak included in the RANDMARDI calculations. Each individual restraint was generated from the average distance calculated for

that proton pair from the MARDIGRAS calculations, ±one standard deviation calculated for that proton pair. No symmetry was enforced during the RANDMARDI procedure. The bounds for symmetry related restraints were calculated by averaging the symmetry related upper and lower bounds for use in the RMD calculations. No additional symmetry was enforced during the RMD calculations.

NOESY spectra obtained for samples in H_2O showed that normal Watson–Crick hydrogen bonding was present for all of the base pairs in the complex, justifying inclusion of 22 hydrogen bond distance restraints. An additional 54 restraints with loose boundaries were derived by inspection of the spectra for a total of 294 restraints. Twenty final structures were generated for the **4**–TOTO complex starting from dsDNA in either B-form or A-form. All the structures converged to the same conformation. The root-mean-square (rms) deviation of the coordinates of the twenty structures obtained from B-DNA was 1.01 Å. The resulting total violations summed to 7.54 Å. Six violations greater than 0.2 Å, but non greater than 0.26 Å, were found in any of the structures. Different views of the calculated structures are given in Figs. 4 and 5.

Helix parameters for the 40 final structures were analyzed with the program CURVES 5.1.[21,22] The CURVES algorithm allows the determination of local structural parameters for irregular nuclei acids structures. Some of the results are shown in Fig. 7.

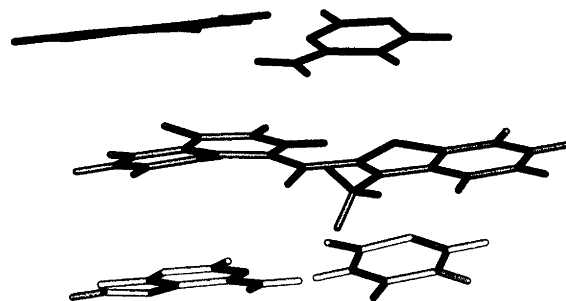


Fig. 5. Stacking of the TOTO chromophore (middle part) in the (5′-CpU-3′):(5′-GpA-3′) binding site viewed perpendicular to the helix axis. Only the nucleobase and the TOTO chromophores are shown.

Discussion

The selectivity of TOTO. TOTO binds preferentially to oligonucleotides containing a $(5'\text{CTAG-}3')_2$ binding site. TOTO also binds preferentially to a $(5'\text{-CCGG-}3')_2$ site but less favorably than to the $(5'\text{-CTAG-}3')_2$ site.[3–5] For the oligonucleotide **1** we estimate that there is at least a 100-fold preference for binding to the $(5'\text{-CTAG-}3')_2$ site than to any other site present in that oligonucleotide.[3–5] The spectra of the complex formed with the oligonucleotide **2** show that binding to the $(5'\text{-CCGG-}3')_2$ predominates over binding to any other site in this oligonucleotide by a factor of at least 50. The distribution of TOTO binding to the $(5'\text{-CTAG-}3')_2$ site compared with the

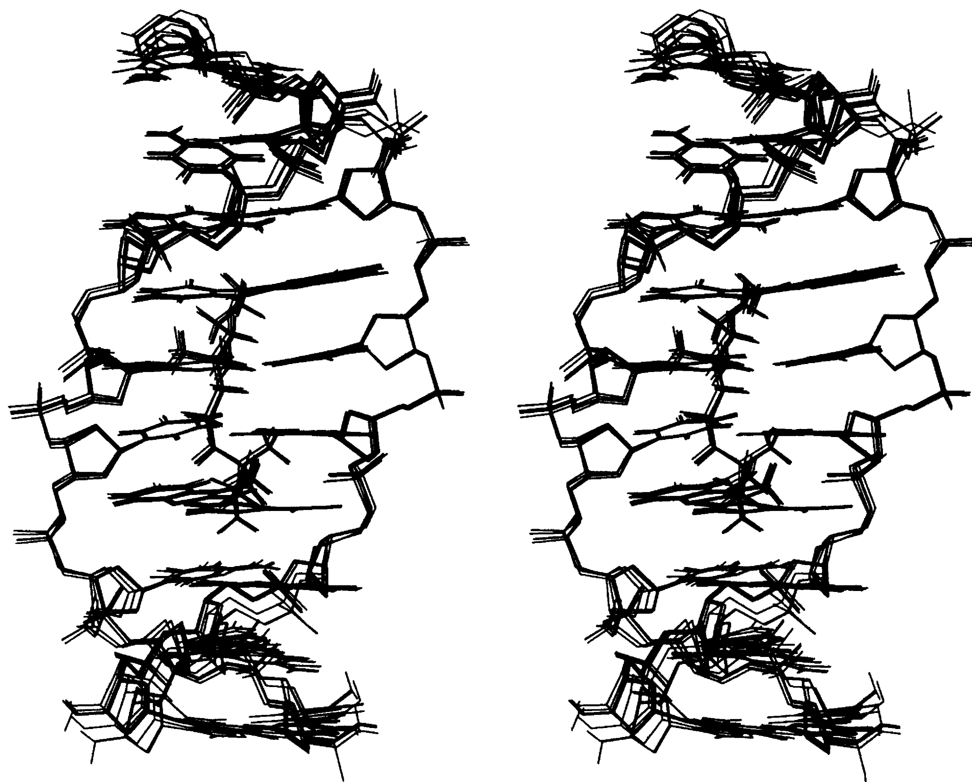


Fig. 4. Stereoview of the stick plot of the structure of the TOTO–DNA complex. A superposition of ten of the forty structures obtained by RMD calculations. Deoxyribose protons have been omitted for clarity.

$(5'\text{-CCGG-}3')_2$ site in the oligonucleotides **1** and **2** has been shown to have an equilibrium constant of $K = 6.2(1.2)$.[5] The ratio between the **1**–TOTO and the **2**–TOTO complexes was found to be 2.7(5) in a mixture with equal amounts of **1**, **2** and TOTO.[5]

In the modified $(5'\text{-CTAG-}3')_2$ binding sites examined in this work, preferential binding is maintained to the $(5'\text{-CTAI-}3')_2$ and the $(5'\text{-CUAG-}3')_2$ sequences. However, the selectivity in the binding to these sites is lower than that to the $(5'\text{-CTAG-}3')_2$ site. TOTO prefers the $(5'\text{-CUAG-}3')_2$ to the same extent as the $(5'\text{-CCGG-}3')_2$ site, whereas binding to the $(5'\text{-CTAI-}3')_2$ is stronger but not as strong as to the $(5'\text{-CTAG-}3')_2$ site.

These results show that the binding constant is lowered when a thymidine is replaced by a uridine. This implies that the methyl group on T4 in the $(5'\text{-CTAG-}3')_2$ of **1**–TOTO complex actually contributes to the binding energy by a van der Waals interaction as indicated by the structure. The lowering of the binding constant on going from the $(5'\text{-CTAG-}3')_2$ to the $(5'\text{-CTAI-}3')_2$ site is less pronounced and is not easily explained as being due to a change in a single contribution to the binding energy. It is probable that local changes between the conformations of the oligonucleotides **1** and **3** cause this effect.

*Description of the structure of the **4**–TOTO complex.* The structure of the **4**–TOTO complex is similar to the structure of the **1**–TOTO complex and reveals that TOTO bis-intercalates in the $(5'\text{-CUAG-}3')_2$ site with the benzothiazole ring system sandwiched between the pyrimidine bases and the quinolinium ring system between the purine bases. The *N*-methyl group on the benzothiazole is centered in the major groove. The linker between the two chromophores is positioned in the minor groove crossing from one side of the groove to the other. This probably introduces van der Waals contacts between the linker chain *N*-methyl groups and the walls of the groove. Owing to the very few restraints involving linker protons used in the calculation, it is more dispersed than the rest of the structure. The length of the linker matches the dsDNA structural requirements to fulfil nearest neighbor bis-intercalation. Twofold symmetry of the parameters about the center of the helix is evident and consistent with the nature of the complex.

In the complex the TOTO chromophores extend across the dsDNA helix. U4 and U12 unstack from A5 and A13 and move into the major groove to maximize van der Waals contacts with the benzothiazole ring of TOTO (Figs. 5 and 6). The observed imino proton cross peaks including the cross peaks to the adenine amino protons prove the existence of normal Watson–Crick base pairing in the intercalation site. So even though the uridines are unstacked from the neighboring adenines, the base pairings of the thymidines with the complementary adenines do not appear to be appreciably disturbed.

Compared with the **1**–TOTO complex the structure of the **4**–TOTO complex shows some slight differences. These differences originate mainly from the exact position of the stacked TOTO chromophore in between the C–G and the A–U base pairs (Figs. 5 and 6). In contrast with the oligonucleotide **1**, the oligonucleotide **4** has no methyl group in the binding site to lock the position of the benzothiazole ring system. This implies that the chromophore is moved slightly into the major groove away from the position in the **1**–TOTO complex.

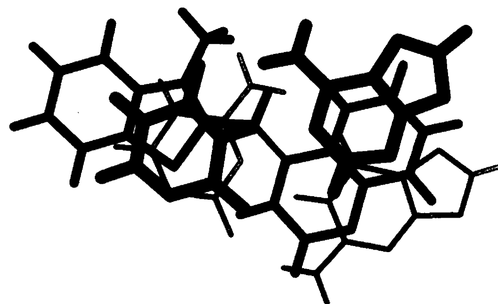


Fig. 6. Stacking of the TOTO chromophore (middle part) in the (5′-CpU-3′):(5′-GpA-3′) binding site viewed down the helix axis. Only the nucleobase and the TOTO chromophores are shown.

Helix parameters. There are three major categories of helix parameter: axis–base pair, intra-base pair and inter-base pair.[23] Some of these parameters for the **4**–TOTO complex are shown in Fig. 7. The reported values for the terminal base pairs are probably less reliable due to dynamic motions and are included only for completeness. The deviation of the calculated helix parameters from symmetry reflects the level of the uncertainty. The helix parameters show in general a similarity to those reported earlier for the **1**–TOTO complex.[4] Many of the helix parameters, such as tip, shear, buckle, shift, slide and tilt reflect the symmetry of the complex with an overall change for the helix of close to zero. This is similar to the standard parameters for both A- and B-DNA. The rise between base pairs are large at the two intercalation sites as expected. The values for twist are smaller at the intercalation site compared with the value of 36° for standard B-DNA. This leads to an overall value of 64° for the unwinding of the helix.

Comparison with other bis-intercalators. There are a number of dsDNA complexes with other bis-intercalators that have been structurally characterized.[2] The quinomycin antitumor antibiotics are a family of cyclic depsipeptides that bis-intercalate in dsDNA. Echinomycin is probably the best known among the quinomycin antitumor antibiotics. The bis-intercalation of echinomycin in various dsDNA oligonucleotides has been studied by Feigon and co-workers.[24–26] They found that echinomycin bis-intercalates selectively on each side of a 5′-CpG-3′ site. Similar results have been found by Gao and Patel.[27]

TANDEM, a fully synthetic compound, belongs to the triostin family of naturally occurring antibiotics that are closely related to quinomycins. While triostin and echino-

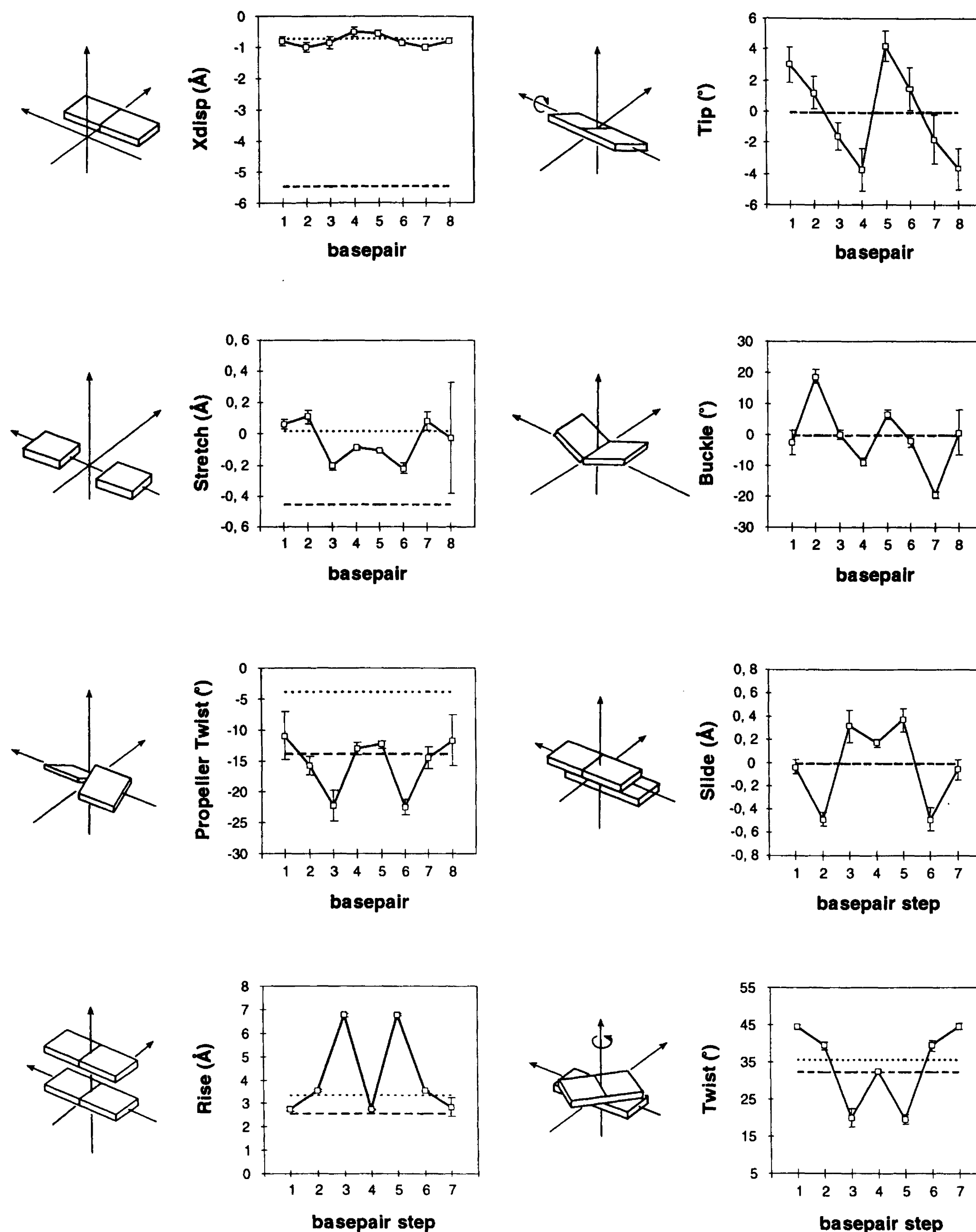


Fig. 7. Helical parameters for the TOTOEt complex calculated using CURVES and compared with canonical A-DNA (- - - - -) and B-DNA (- - - - - - - -).

mycin show high preference for 5′-CpG-3′-rich dsDNA a methylated TANDEM analog preferentially binds to alternating 5′-TpA-3′ and 5′-CpI-3′ sequences.[28–30] Luzopeptin is also a cyclic depsipeptide antibiotic that bis-intercalates in dsDNA. It has been shown that it binds by bracketing two AT base pairs.[31,32]

Common to the bis-intercalator of the quinomycin, triostin and luzopeptin type is the presence of two intercalating chromophores linked with cyclic peptide groups. The site selectivity of these compounds is predominantly ascribed to interaction of the peptide with the dsDNA in the grooves.

TOTO belongs to a new structural class of 'adaptable intercalators' in contrast with the common intercalators. The characteristic feature of the TOTO chromophores is the ability to adapt the base pair propeller twist of the dsDNA.[4] Most other bis-intercalators have rigid aromatic chromophores that cannot adapt to the base pair propeller twist. This makes TOTO unique among bis-intercalators.

References

1. Wakelin, L. P. G. *Med. Res. Rev. 6* (1986) 275.
2. Searle, M. S. *Progr. NMR Spectrosc. 25* (1993) 403.
3. Jacobsen, J. P., Pedersen, J. B., Hansen, L. F. and Wemmer, D. E. *Nucleic Acids Res. 23* (1995) 753.
4. Spielmann, H. P., Wemmer, D. E. and Jacobsen, J. P. *Biochemistry 34* (1995) 8542.
5. Hansen, L. F., Jensen, L. K. and Jacobsen, J. P. *Nucleic Acids Res. 24* (1996) 859.
6. Faridi, J., Nielsen, K. E., Stein, P. C. and Jacobsen, J. P. *J. Biomol. Struct. Dyn. 15* (1997) 321.
7. Stærk, D., Hamed, A. A., Pedersen, E. B. and Jacobsen, J. P. *Bioconjugate Chem. 8* (1997) 869.
8. Benson, S. C., Singh, P. and Glazer, A. N. *Nucleic Acids Res. 21* (1993) 5727.
9. Benson, S. C., Mathies, R. A. and Glazer, A. N. *Nucleic Acids Res. 21* (1993) 5720.
10. Rye, H. S., Yue, S., Wemmer, D. E., Quesada, M. A., Haugland, R. P., Mathies, R. A. and Glazer, A. N. *Nucleic Acids Res. 20* (1992) 2803.
11. Stein, P. C., Jacobsen, J. P. and Spielmann, H. P. *J. Magn. Reson., Ser. B 109* (1995) 93.
12. Wüthrich, K. *NMR of Proteins and Nucleic Acids*, Wiley, New York 1986, pp. 1–290.
13. Feigon, J., Leupin, W., Denny, W. A. and Kearns, D. R. *Biochemistry 22* (1983) 5943.
14. Hare, D. R., Wemmer, D. E., Chou, S.-H., Drobny, G. and Reid, B. R. *J. Mol. Biol. 171* (1983) 319.
15. Scheek, R. M., Boelens, R., Russo, N., Van Boom, J. H. and Kaptein, R. *Biochemistry 23* (1984) 1371.
16. Scheek, R. M., Russo, N., Boelens, R. and Kaptein, R. *J. Am. Chem. Soc. 105* (1983) 2914.
17. Boelens, R., Scheek, R. M., Dijkstra, K. and Kaptein, R. *J. Magn. Reson. 62* (1985) 378.
18. Liu, H., Spielmann, H. P., Ulyanov, N. A., Wemmer, D. E. and James, T. L. *J. Biol. NMR 6* (1995) 390.
19. Borgias, B. A. and James, T. L. *J. Magn. Reson. 87* (1990) 475.
20. Borgias, B. A., Gochin, M., Kerwood, D. and James, T. L. *Progr. NMR Spectrosc. 22* (1990) 83.
21. Lavery, R. and Sklenar, H. *J. Biomol. Struct. Dyn. 6* (1988) 63.
22. Lavery, R. and Sklenar, H. *J. Biomol. Struct. Dyn. 7* (1989) 655.
23. Dickerson, R. E., Bansal, M., Calladine, C. R., Diekmann, S., Hunter, W. N., Kennard, O., von Kitzing, E., Nelson, H. C. M., Lavery, R., Olson, W. K., Saenger, W., Shakked, Z., Soumpasis, D. M., Tung, C.-S., Sklenar, H., Wang, A. H.-J. and Zhurkin, V. B. *EMBO J. 8* (1989) 1.
24. Gilbert, D. E., van der Marel, G. A., van Boom, J. H. and Feignon, J. *Proc. Natl. Acad. Sci. USA 86* (1989) 3006.
25. Gilbert, D. E. and Feigon, J. *Biochemistry 30* (1991) 2483.
26. Gilbert, D. E. and Feigon, J. *Nucleic Acids Res. 20* (1992) 2411.
27. Gao, X. and Patel, D. J. *Q. Rev. Biophys. 22* (1989) 93.
28. Addess, K. J. and Feigon, J. *Biochemistry 33* (1994) 12386.
29. Addess, K. J., Gilbert, D. E., Olsen, R. K. and Feigon, J. *Biochemistry 31* (1992) 339.
30. Addess, K. J. and Feigon, J. *Biochemistry 33* (1994) 12397.
31. Searle, M. S., Hall, J. G., Denny, W. A. and Wakelin, L. P. G. *Biochem. J. 259* (1989) 433.
32. Zhang, X and Patel, D. J. *Biochemistry 30* (1991) 4026.

Received September 29, 1997.