Miller, J. A., Serio, G. F., Howard, R. A., Bear, J. L., Evans, J. E., & Kimball, A. P. (1979) *Biochim. Biophys. Acta 579*, 291-297.
O'Sullivan, W. J., & Perrin, D. D. (1964) *Biochemistry 3*, 18-26.
Prosen, D. E., & Cech, C. L. (1986) *Biochemistry 25*, 5378-5387.
Rhodes, G., & Chamberlin, M. J. (1974) *J. Biol. Chem. 249*, 6675-6683.
Roeder, R. G. (1976) in *RNA Polymerase* (Losick, R., & Chamberlin, M., Eds.) pp 285-329, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Sabat, M., Cini, R., Harmony, T., & Sundaralingam, M. (1985) *Biochemistry 24*, 7827-7833.
Scatchard, C. T. (1949) *Ann. N.Y. Acad. Sci. 51*, 660-672.
Scrutton, M. C., Wu, C.-W., & Goldthwait, D. A. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 2497-2501.
Steck, T. L., Caicuts, M. J., & Wilson, R. G. (1968) *J. Biol. Chem. 243*, 2769-1778.
Storer, A. C., & Cornish-Bowden, A. (1977) *Biochem. J. 165*, 61-69.
Stryer, L. (1978) *Annu. Rev. Biochem. 47*, 819-846.
Sweetser, D., Nonet, M., & Young, R. A (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 1192-1196.
Tabor, S., & Richardson, C. C. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 4076-4080.
Tabor, S., & Richardson, C. C. (1990) *J. Biol. Chem. 265*, 8322-8328.
Tyagi, S. C. (1991) *J. Biol. Chem. 266*, 17936-17940.
Tyagi, S. C., & Wu, F. Y.-H. (1987) *J. Biol. Chem. 262*, 10684-10688.
von Hippel, P. H., Bear, D. G., Morgan, W. D., & McSwiggen, J. A. (1984) *Annu. Rev. Biochem. 54*, 389-446.
White, C. E., Ho, M., & Weimer, E. Q. (1960) *Anal. Chem. 32*, 438-440.
Wu, C.-W., & Goldthwait, D. A. (1969a) *Biochemistry 8*, 4450-4458.
Wu, C.-W., & Goldthwait, D. A. (1969b) *Biochemistry 8*, 4458-4464.
Wu, C. W., Wu, F. Y.-H., & Speckhard, D. C. (1971) *Biochemistry 16*, 5449-5454.
Wu, F. Y.-H., & Tyagi, S. C. (1987) *J. Biol. Chem. 262*, 13147-13154.
Yager, T. D., & von Hippel, P. H. (1991) *Biochemistry 30*, 1097-1118.
Yarbrough, L. R., Schlageck, J. G., & Maughman, M. (1979) *J. Biol. Chem. 254*, 12069-12073.

A ^{1}H-NMR Study of the DNA Binding Characteristics of Thioformyldistamycin, an Amide Isosteric Lexitropsin†

Malvinder P. Singh,‡ Surat Kumar,‡ Tomi Joseph,‡ Richard T. Pon,§ and J. William Lown*,‡
Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada, and Regional DNA Synthesis Laboratory, University of Calgary, Calgary, Alberta T2N 4N1, Canada
Received December 20, 1991; Revised Manuscript Received April 15, 1992

ABSTRACT: The interaction of thioformyldistamycin, an amide isostere of the naturally occurring antibiotic distamycin A, with a self-complementary decadeoxynucleotide duplex, d(CGCAATTGCG)$_2$, has been examined using a variety of high-field ^{1}H-NMR techniques. The ligand exhibits two forms in solution arising from geometric isomerism due to restricted rotation around the thioformamide bond. Only the thermodynamically more stable *Z*-form is shown to bind to the oligonucleotide along its minor groove at the central 5′-AATT segment with the end groups of the ligand extending into the flanking GC regions but without any close contact at the amidinium terminus. Cross-peaks involving characteristic intra- and interresidue proton connectivities in the 2D experiments (COSY and NOESY) were employed to assign individual resonances of both strands in the asymmetric DNA–drug complex. The solution structure of the complex was constructed by molecular mechanics calculations based upon initial estimates of drug–DNA NOE contacts and further refined through energy minimization. These results complement previous structural studies on distamycin and other lexitropsins with oligonucleotides. The exchange of the ligand between two equivalent binding sites on the DNA sequence was estimated to occur at 40 s^{-1} with a free energy of activation of 16.5 kcal·mol^{-1} at 321–326 K. There was no evidence of formation of a 2:1 drug–oligomer complex, in contrast to the case of the natural product, which is attributed to steric demands of the larger sulfur atom.

Several compounds including distamycin A (Arcamone et al., 1967; Hahn, 1975), netropsin (Julia & Preau-Joseph, 1963), anthelvencin A (Probst et al., 1965), noformycin (Diana, 1973), and the kikumycins A and B (Takahishi et al., 1972) comprise the pyrrole amidine class of antitumor antibiotics. These oligopeptides appear to share a common molecular mechanism of binding to double-stranded B-DNA (Wartell et al., 1974) and are known to inhibit the DNA and RNA polymerase activity in vitro through interactions with the DNA template, thereby blocking the synthesis of DNA (Hahn, 1980; Zimmer et al., 1971, 1983). Both distamycin A and netropsin have been of widespread interest in biochemical and biophysical investigations, and the reader is directed to extensive reviews on such studies (Hahn, 1980;

† This investigation was supported by a grant (to J.W.L.) from the National Cancer Institute of Canada.
* Address correspondence to this author.
‡ University of Alberta.
§ University of Calgary.

0006-2960/92/0431-6453$03.00/0

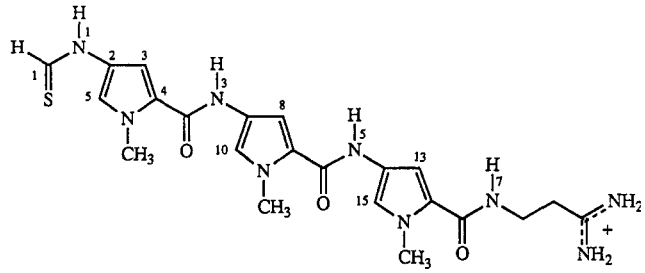
Baguley, 1982; Zimmer & Wahnert, 1986; Lown, 1988). These two antibiotics are also the most studied from the point of view of structural studies on their complexes with synthetic oligodeoxynucleotides, and a number of NMR (Patel, 1982; Klevit et al., 1986; Patel & Shapiro, 1986a,b; Pelton & Wemmer, 1988, 1989) and X-ray crystallographic studies (Kopka et al., 1985; Coll et al., 1985, 1987) indicate their high preference for the AT-rich sequences in the minor groove of duplex DNA.

On the basis of the factors contributing to such sequence specificity (Lown & Krowicki, 1985; Kopka et al., 1985), a number of lexitropsins, or synthetic analogs of the parent antibiotics, have been developed with the aim of altering the sequence selection in a predictable manner (Lown, 1989). As part of this ongoing program in our laboratories, structural modifications in both the side-chain residues and the skeletal *N*-methylpyrrole units have been examined systematically (Lown, 1990, and references therein). An important extension of these efforts with particular relevance to their biological potency is to overcome the inherent susceptibility of the terminal amide groups in these amido ligands toward cellular degradation. To this end, we recently reported on an isosteric thioamide surrogate of distamycin A, which shows DNA binding footprints similar to that of the parent molecule and is stable toward acid and enzymatic hydrolysis (Zimmermann et al., 1991). We now report the details of the high-field ^{1}H-NMR study of interactions between thioformyldistamycin and a synthetic decadeoxynucleotide sequence.

Materials and Methods

Synthesis. The synthesis of the self-complementary decadeoxynucleotide d(CGCAATTGCG)$_2$ has been reported previously (Lee et al., 1988a). The detailed procedure for the preparation of thioformyldistamycin has also been communicated recently form our laboratories (Zimmerman et al., 1991). Briefly, the parent compound distamycin A (Sigma Chemical Co.) was hydrolyzed under mild acidic conditons to obtain deformyldistamycin (Arcamone et al., 1967; Lown & Krowicki, 1985), which was thioformylated with potassium dithioformate. The precipitated product was washed with water, dried in vacuo, and characterized by ^{1}H-NMR and FAB-MS.

NMR Spectroscopy. For NMR studies the oligonucleotide was dissolved in 400 μL of 99.8% D_2O containing 10 mM NaCl, 0.1 mM EDTA, and 35 mM phosphate buffer, pH 7.0 (uncorrected). The solution was lyophilized twice from 99.96% D_2O and finally dissolved in 400 μL of 99.996% D_2O. Owing to the sparse solubility of thioformyldistamycin in water, its stock solution was prepared in D_2O containing 20% DMSO-d_6. All NMR spectral studies were carried out at 21 °C on a Bruker AM-400 (interfaced to an Aspect 3000 computer) and a Varian Unity spectrometer operating at 499.84 Hz. All experiments were temperature regulated to ±1 °C and were run on nonspinning samples. Phase-sensitive spectra were acquired using the hypercomplex method on the Varian-500 spectrometer (States et al., 1982) and the time-proportional phase increments (TPPI) method on the Bruker spectrometer. Two-dimensional correlation (COSY; Marion & Wuthrich, 1983) and NOESY (Bodenhausen et al., 1984) spectra were acquired using spectral windows of 3800 Hz (at 400 MHz) or 4200 Hz (at 500 MHz) and 1K data points in t_2 dimension. A total of 64–80 scans were acquired for each of 256 t_1 increments with a 1.4- or 2.4-s delay between scans. The contour maps for the COSY experiment were constructed after apodization with an unshifted square sine-bell function in both dimensions while a 90°-shifted square sine-bell function was used for NOESY. A series of NOESY experiments were run at different mixing time intervals of 80, 100, 150, 200, 250, and 400 ms.



5' - C G C A A T T G C G - 3'

I IIIIIV V VIV III II I

3' - G C G T T A A C G C - 5'

FIGURE 1: Structures and numbering schemes for thioformyldistamycin and the decadeoxyribonucleotide. The roman numerals reflect the 2-fold symmetry of the self-complementary duplex sequence.

For detection and analysis of the exchangeable imino protons, the samples were evaporated and redissolved in a 9:1 H_2O–D_2O mixture and the spectra recorded using the 1–3–3–1 pulse sequence (Hore, 1983a,b). The 1D-NOE difference spectra were similary obtained using the same sequence for the imino protons using an acquisition window of 8000 Hz and collecting 1240 fid's using selective irradiation intervals of 350 ms.

Molecular Mechanics and Modeling. The molecular modeling and force field calculations were carried out using the Insight II software of Biosym. We used the standard Biosym force field (CVFF) in all energy calculations. For nonbonded energy calculations a smooth cutoff extending up to 15 Å was used. We used a distance-dependent dielectric constant of $2r$, where r is the interatomic distance, to calculate the electrostatic energies.

The structure of thioformyldistamycin was modeled using the Builder module of Insight II. The partial charges were calculated using AMPAC. The structure thus obtained was energy minimized using the Discover module, until the derivatives converged to 0.2 kcal (mol·Å)$^{-1}$. The B-DNA duplex d(CGCAATTGCG)$_2$ was generated using the Biopolymer module of Insight II. The structure thus obtained was further energy minimized using the Discover module, until the derivatives converged to 0.5 kcal (mol·Å)$^{-1}$.

The docking of the ligand to the DNA duplex was carried out with the aid of 3D graphics on the IRIS 4D-70/GT workstation. A combined translational/rotational operation was used to insert the ligand, thioformyldistamycin, at the AATT site of the DNA helix. For orientation of the ligand in the minor groove, five experimentally observed NOEs were used as a guide. The naturally curved shape of the ligand facilitated a comfortable fit of the ligand in the DNA groove. Finally the ligand–DNA complex was subjected to another energy minimization until all the derivatives converged to 0.5 kcal (mol·Å)$^{-1}$.

Results

The structures and numbering schemes for the decadeoxynucleotide and thioformyldistamycin are shown in Figure 1. The roman numerals designate the imino protons of the oli-

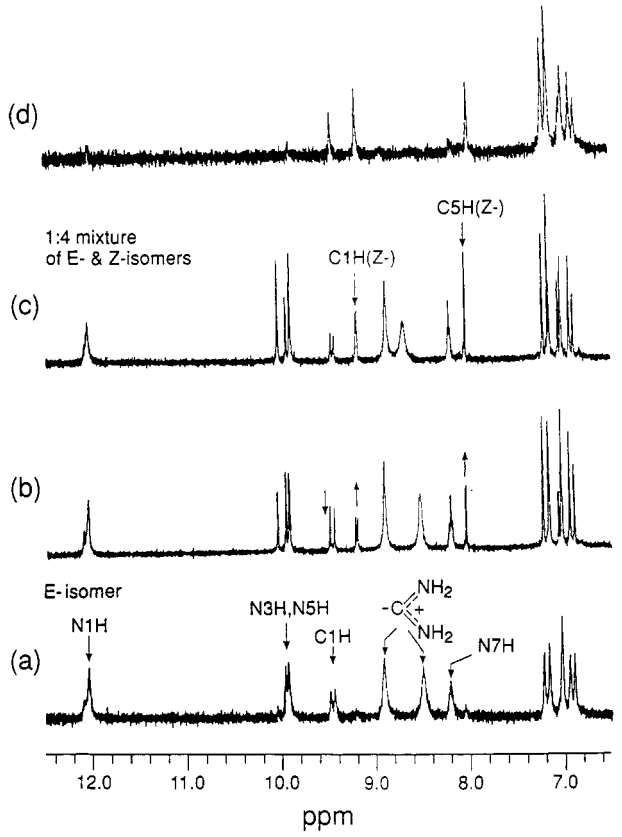


FIGURE 2: Partial 6.5–12.5 ppm region of the ^{1}H-NMR spectra for thioformyldistamycin, recorded in DMSO-d_6: (a) a freshly prepared sample indicating exclusively the *E*-configuration at the thioamide bond; (b) spectrum after the sample was allowed to stand for 14 h at room temperature and (c) after 24 h, further to which no changes were observed; (d) spectrum after exchange with D_2O. The relevant protons are identified, with reference to their numbering as in Figure 1.

gonucleotide and indicate the 2-fold symmetry of the self-complementary sequence. The complete assignments of both exchangeable and nonexchangeable proton resonances of the duplex d(CGCAATTGCG)$_2$ have been described previously (Lee et al., 1988a,b) and were confirmed for the present sample using 2D-NOESY experiments.

NMR Characteristics of Thioformyldistamycin. The partial 6.5–13.0 ppm region of the ^{1}H-NMR spectrum for a freshly prepared solution of the ligand thioformyldistamycin in DMSO-d_6 is shown in Figure 2a, indicating only one form of the compound which is characterized as the *E*-isomer by a characteristic doublet (J = 14 Hz) at 9.47 ppm assigned to the thioformyl proton (coupling due to the adjacent amide NH proton). Three sets of resonances appear between 6.8 and 7.3 ppm, characteristic of the heterocyclic pyrrole ring protons. The two broad signals at 8.5 and 8.9 ppm are assigned to the two terminal amidine NH_2 protons, while the remaining amide NH groups are identified as follows: thioformamide N1H being the most downfield resonance at 12.0 ppm, and aliphatic side chain containing N7H at 8.2 ppm (triplet multiplicity due to coupling with the adjacent methylene group), and the two central N3- and N5Hs at 9.91 and 9.94 ppm, respectively.

Upon standing these signals start to diminish and some new resonances appear (Figures 2b,c). Most marked of these are (a) a doublet at 9.2 ppm (J = 5 Hz, assigned to the thioformyl proton in the *Z*-form) and (b) a pyrrole-like doublet (J = 1 Hz) at 8.05 ppm. Although the appearance of the former is reminiscent of *E*–*Z* isomerism (Figure 3), it is not clear why the other signal corresponding to a pyrrolic proton stands out separately (almost 1 ppm downfield) from the rest of the

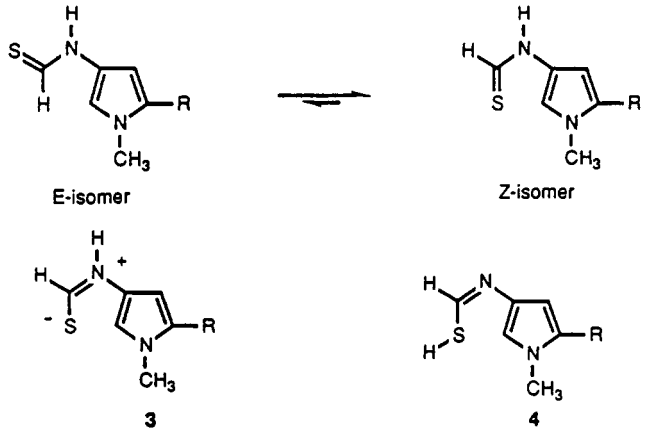


FIGURE 3: Scheme depicting the equilibrium mixture of *E*- and *Z*-isomers of thioformyldistamycin. Structure **3** corresponds to a resonance-delocalized form on the basis of a high degree of polarizability of the thioamide bond. Also shown is a structure based on a tautomeric change of thioamide to the thiolimide form **4**.

pyrrole protons in this configuration. Perhaps, as illustrated in the optimum geometric arrangement of the *Z*-form (Figure 3), the electron cloud around the bulky sulfur atom, which is in close proximity to the pyrrole H5 atom, is able to exert some shielding effects. After still longer times, the spectrum shows predominantly the presence of this *Z*-isomer of thioformyldistamycin (Figure 2c).

NMR Titration of the Ligand onto the Oligonucleotide. The changes observed for the nonexchangeable proton resonances in the ^{1}H-NMR spectrum (in D_2O), upon titration of the free duplex d(CGCAATTGCG)$_2$ to 1 mol equiv of the ligand, are shown in Figure 4. The symmetry in the ligand-free double-stranded decanucleotide is evident from its spectrum, which shows the presence of two thymidine methyl resonances (Figure 4a) and four imino proton signals (when recorded in H_2O solutions, Figure 5a). The imino signal I for the peripheral base pairs was not observed, presumably due to fraying of the duplex at the sequence ends. Also shown is the spectrum for the sample of free ligand consisting of an equilibriated mixture of *E*- and *Z*-isomers (in ~1:4 ratio). It is interesting to note that, at the point of titration to a 1:1 molar ratio of ligand to DNA, the spectrum showed the presence of both these isomers. However, a comparison of the spectra in Figure 4, b and c, reveals that, upon standing for 8 h, the characteristic signal for the *E*-isomer at 9.5 ppm completely disappears, suggesting that in the presence of duplex oligonucleotide the $E \rightleftharpoons Z$ equilibrium lies completely toward the latter, i.e., the *Z*-form in which thioformyldistamycin binds to the DNA sequence.

At a 1:1 molar ratio of ligand to DNA, a loss of degeneracy and doubling of resonances was revealed for selected signals which is evidently due to the asymmetry induced by tight binding of the ligand to the oligomer. Most marked differentiation of the individual strands was observed around the central AATT region, e.g., the thymidine methyl signals at 1.3–1.6 ppm. Some other "doubled" resonances were also identified (from a combination of 2D experiments described later) as C3–H5, A4–H8, A5–H8, A5–H2, T6–H6, T7–H6 and G8–H8. These observations complement several previous reports on removal of the dyad symmetry of self-complementary DNA sequences upon binding of asymmetric ligands (Klevit et al., 1986; Lee et al., 1988a,b; Kumar et al., 1990).

The binding of the ligand in the central region of the nucleotide is also illustrated by the changes in the exchangeable imino protons when the spectrum was recorded in 90% H_2O–D_2O. The A–T imino resonances (13.6–14.2 ppm) were both

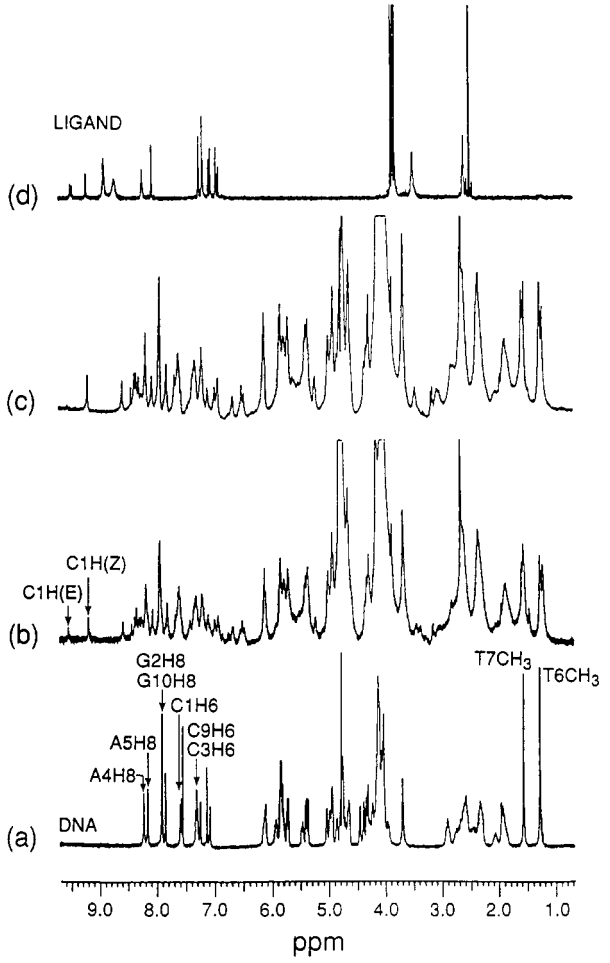


FIGURE 4: Nonexchangeable proton containing regions of the ^{1}H-NMR spectra for the native oligonucleotide (a), the free ligand (d), and a 1:1 ligand–DNA complex (b, c). Spectrum b was recorded immediately after equivalent amounts of ligand and DNA were mixed, while spectrum c corresponds to the same sample after being allowed to stand for 8 h.

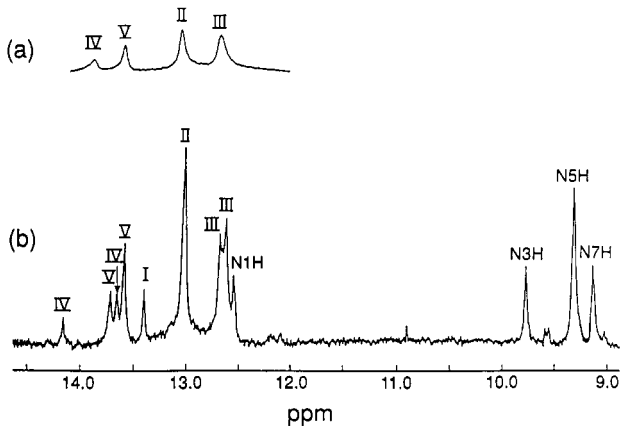


FIGURE 5: Downfield regions in the ^{1}H-NMR spectra of (a) d-(CGCAATTGCG)$_2$ and (b) its 1:1 complex with (*Z*)-thioformyldistamycin at 294 K. The exchangeable imino proton resonances of the oligonucleotide are labeled I–V, and the exchangeable amide protons of the ligand are labeled as in Figure 1. The binomial 1–3–3–1 pulse sequence (Hore, 1983a,b) was employed for suppressing the water signal.

doubled to give four lines together with two of the three G-C imino signals (12.6–13.1 ppm) distinctly separated (Figure 5b). The presence of these imino signals clearly indicates that (i) the oligonucleotide retains the Watson–Crick base-paired duplex conformation and (ii) the larger chemical shift differences for the central base pairs IV and V complement the

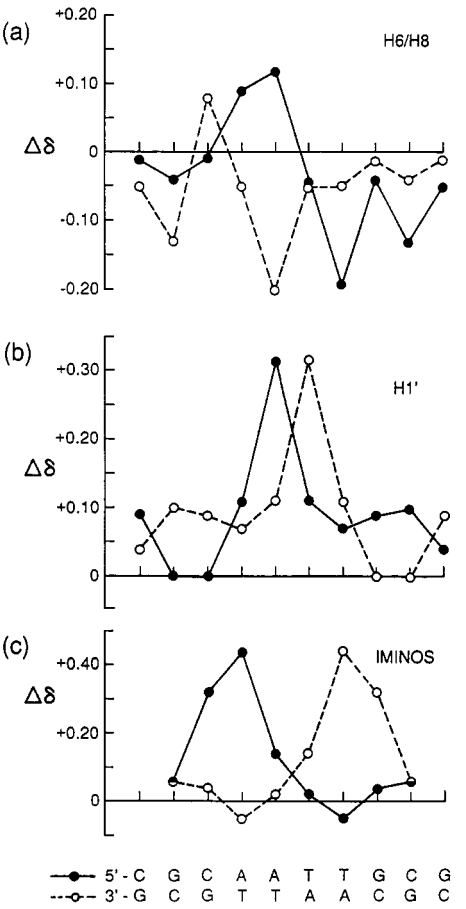


FIGURE 6: Graphical representation of the complexation-induced chemical shift changes for aromatic base H6/H8, anomeric sugar H1′, and imino protons in the DNA sequence. Positive and negative values indicate downfield and upfield shifts, respectively, upon complexation. Both strands of the DNA sequence are drawn at the bottom.

aforementioned change in the local environment around the central AATT site.

Additional support for this greater perturbation around the core segment is derived from the complexation-induced chemical shift data indicating greater chemical shift ($\Delta\delta$) differences for the aromatic base protons and the glycosidic H1′ protons corresponding to the central segments as a result of ligand binding. The difference values were estimated from the assignment of individual resonances in the complex and comparison with the chemical shifts for the ligand-free oligonucleotide (Figure 6). The data presented thus far, i.e., the observation of doubling of resonances in the selected regions of the DNA sequence and the changes in chemical shift values, are consistent with a structure of the complex wherein the ligand is docked along the central section of the oligonucleotide with some structural deformations extending into the adjacent GC regions. The orientation of the asymmetric ligand, with respect to the duplex oligonucleotide, was elucidated from the intermolecular NOE contacts (vide infra).

The nonexchangeable protons of the ligand-bound oligonucleotide complex were assigned using the well-established sequential NOE connectivity approach for B-DNA structures (Gronenborn & Clore, 1985; Kearns, 1987; Reid, 1987; and references cited therein). This sequential mapping approach for B-type oligonucleotide chains essentially depends on the relatively short (3–5 Å) distances between sugar hydrogens (H1′) and aromatic hydrogens (H6/H8) on the same residue and to the ones on their adjacent 3′-neighbor residues. Thymine methyl proton signals provide an additional set of sequential connectivities via their proximity to sugar protons on

Table I: Chemical Shift Assignments (in ppm) of Selected Protons in the 1:1 Thioformyldistamycin–d(CGCAATTGCG)$_2$ Complex at 294 K[a]

	H8/H6[b]	H5/H2/Me[b]	H1′	H2′	H2″	imino
5′ C	7.59 (−0.01)	5.87	5.72 (+0.09)	1.92	2.34	13.41
G	7.90 (−0.04)		5.82 (+0.00)	2.34	2.60	13.02 (+0.06)
C	7.30 (−0.01)	5.39	5.80 (+0.00)	1.89	2.30	12.90 (+0.32)
		5.26*				
A	8.35 (+0.09)	7.14	5.92 (+0.11)	2.61	2.74	14.15 (+0.44)
	8.21* (−0.05)					
A	8.32 (+0.12)	7.70	6.14 (+0.33)	2.74	2.91	13.72 (+0.14)
	8.15* (−0.05)					
T	7.09 (−0.04)	1.27	6.18 (+0.11)	1.96	2.40	13.60 (+0.02)
	6.93* (−0.20)	1.20*				
T	7.09 (−0.19)	1.57	5.83 (+0.07)	2.08	2.44	13.66 (−0.05)
	7.23* (−0.05)	1.64*				
G	7.84 (−0.04)		5.79 (+0.09)	2.42	2.60	12.62 (+0.04)
	7.80* (+0.08)					
C	7.23 (−0.13)	5.37	5.79 (+0.10)	1.86	2.39	13.02 (+0.06)
		5.26*				
G	7.90 (−0.05)		6.14 (+0.04)	2.34	2.60	13.44

[a] The values in parentheses indicate complexation-induced chemical shift changes with respect to the free DNA duplex. [b] Exchange signals are indicated with an asterisk.

the preceding 5′-neighbors. Similarly, the assignment procedure can be extended to identify more intraresidue connectivities from H6/H8 and H1′ to the glycosidic H2′, H2″, and H3′ protons.

As a first step toward the assignment of proton resonances in the complex, a 2D-COSY experiment was performed to identify the aromatic base protons of the three cytosine residues via their strong correlation peaks (data not shown). Also identified in the COSY contour maps were the direct connectivities between sugar protons H1′–H2′ and H1′–H2″, which were helpful in delineating the intranucleotide from the internucleotide connectivities involving H1′ protons in the subsequent 2D-NOESY experiments.

An illustrative NOESY contour map for the complex is shown in Figure 7. In the present study, the previously identified resonances for the cytosine aromatic protons H5 and H6 provided a starting point for analyses of the cross-peaks in the 2D-NOESY spectra. For example, C(1)–H6 gives a cross-peak to C(1)–H1′ with a subsequent connection to G(2)–H8 which in turn shows a cross-peak to G(2)–H1′. The cross-peaks involving A(5)H1′–T(6)Me and T(6)H1′–T(7)Me pairs served as an additional check on the correctness of the sequential connectivities along the length of the oligonucleotide chain. Some of the characteristic intra- and interresidue NOEs involving the nonexchangeable protons are summarized in the form of a NOE grid diagram (Figure 8), and the individual proton resonances based on such assignment strategy are given in Table I.

Although our analysis starts with the methods outlined for the B-DNA conformation for the complex, its validity is confirmed by running the 2D-NOESY experiments at progressively shorter mixing time intervals (100, 150, and 200 ms). Inspection of the relative intensities in these experiments showed that the NOESY cross-peaks corresponding to intranucleotide H6/H8–H2′ and internucleotide H2″–H6/H8 proton pairs were present even at short mixing times, while those corresponding to intranucleotide H6/H8–H2″ and internucleotide H2′–H6/H8 start to appear only at mixing times of 100 ms or greater. Such pattern of NOE connectivities is expected of a right-handed B-DNA conformation and is further supported by an overall order of relative cross-peak intensities being H2′ ≫ H1′ > H3′ at relatively short mixing times (100–150 ms) where the intensities are not severely overestimated due to spin diffusion.

It is also evident from the retention of the B-type helical conformation of the DNA that binding of thioformyldistamycin does not induce any significant changes in the sugar ring puckers and the glycosidic bond angles. NOESY cross-peaks corresponding to H6/H8–H1′ protons for the C1, G2, C3, and G10 residues are, however, more intense than for the other nucleotides in the sequence, which leads to the conclusion that the glycosidic bond angles are perturbed to some degree at the binding site. Interestingly, a comparison of the NOESY spectra recorded at short periods of mixing of 80 ms for the complex and free oligonucleotide itself showed a loss of NOE contacts involving the T(6)Me and the adenine (A5) base protons, which is plausibly because of the conformational changes around the ligand binding site. Alternatively, the source of such differentials in the NOE intensities corresponding to selected regions of the DNA duplex could well be a dispersion or line broadening due to the asymmetric ligand. The intrinsic details for such a deformation would

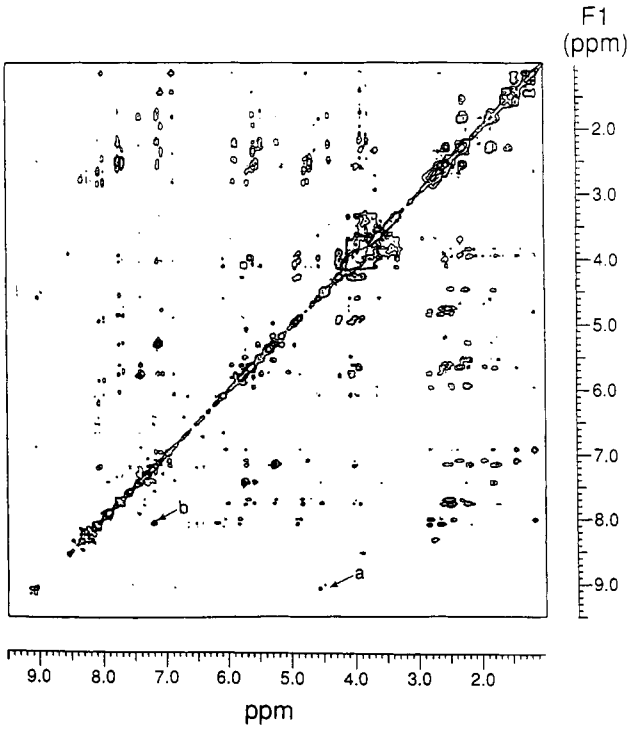


FIGURE 7: Phase-sensitive NOESY (200 ms) contour plot of a 1:1 complex of (Z)-thioformyldistamycin and d(CGCAATTGCG)$_2$ at 294 K. Details for data acquisition and processing are indicated under Materials and Methods. Two intermolecular contacts between ligand and DNA are identified as a and b and listed in Table II.

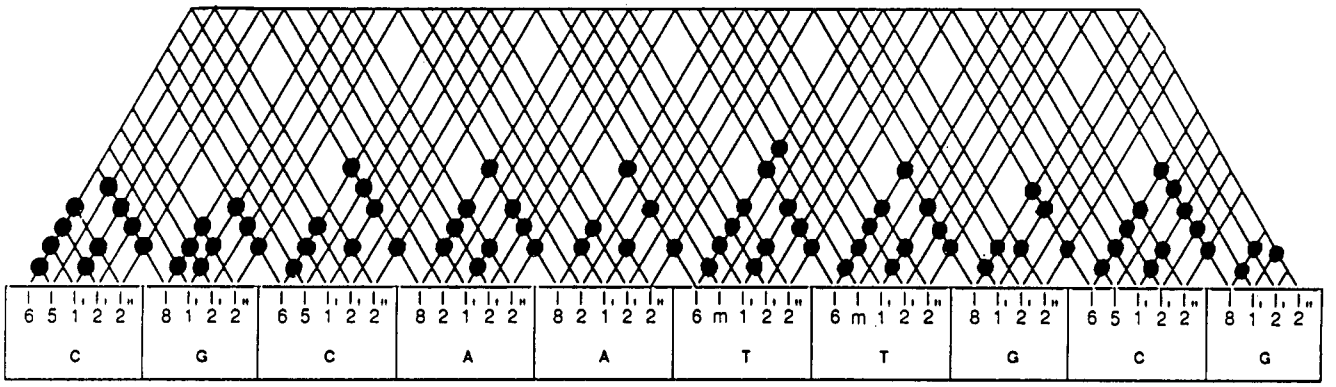


FIGURE 8: Summary of intra- and internucleotide NOE contacts used for assignment of nonexchangeable proton resonances in a 1:1 complex of (*Z*)-thioformyldistamycin and d(CGCAATTGCG)$_2$. Only the contacts involving base and glycosidic H1′, H2′, and H2″ protons are shown.

Table II: Summary of Intermolecular Contacts between (*Z*)-Thioformyldistamycin and d(CGCAATTGCG)$_2$ As Derived from NOE Experiments

	ligand proton	DNA proton
a	C1H	A4H3′
b	C3H	A5H2
c	N1H	imino IV
d	N3H	imino V
e	N5H	imino V
f	N7H	imino IV

require a comprehensive analysis of NOE buildup rates, which was not pursued at present.

Apart from the characteristic intra- and interresidue cross-peaks, a number of intermolecular NOEs were identified (Table II) which provided insight into the relative orientation of the ligand and the actual points of contact between the ligand and the DNA duplex. Thus, two significant cross-peaks corersponded to adenine(4)–H3′ with the ligand formyl proton and to adenine(5)–H2 with the pyrrole H3 proton (Figure 1). Similarly, from the 1D-NOE difference measurements, carried out on the sample in 90% H_2O, the exchangeable ligand amide protons N1H at 12.51 ppm and N7H at 9.12 ppm both showed an NOE enhancement of the imino proton signals corresponding to each of the A–T base pairs marked IV. Also observed was an NOE enhancement in the imino proton signal for each of the AT base pairs marked V upon selective saturation of either ligand N3H (at 9.78 ppm) or N5H (at 9.31 ppm). It is important to note that the NOE contacts between any of the propylamidinium side chain protons with the DNA protons were not observed in any of the NOESY spectra. This is in contrast to the previous NMR studies on distamycin itself and is presumably due in part to a relatively stronger binding of thioformyldistamycin at the thioamide end, arising from a lower pK_a of the thioamide bond which enhances the anchoring of the ligand at one end, thus forcing the other end to come off from the floor of the minor groove (Figure 9). Alternatively, one can envision the same effect due to a relatively faster (on the NMR time scale) free rotation of the aliphatic side chain compared to the exchange between equivalent sites.

Dynamics of Ligand–DNA Binding. A significant observation from this study pertains to the presence of two equivalent and exchanging ligand binding sites present on the self-complementary oligonucleotide sequence, and the next step was to obtain the chemical exchange rates from the shape and coalescence of selected NMR signals. The two T(6)Me signals are well resolved and separated by 18.1 Hz at ambient temperature (294 K), and in the condition of infinitely slow exchange (assumed to be 277 K, the experimental limit for NMR in aqueous solutions) the separation corresponds to 18.9 Hz.

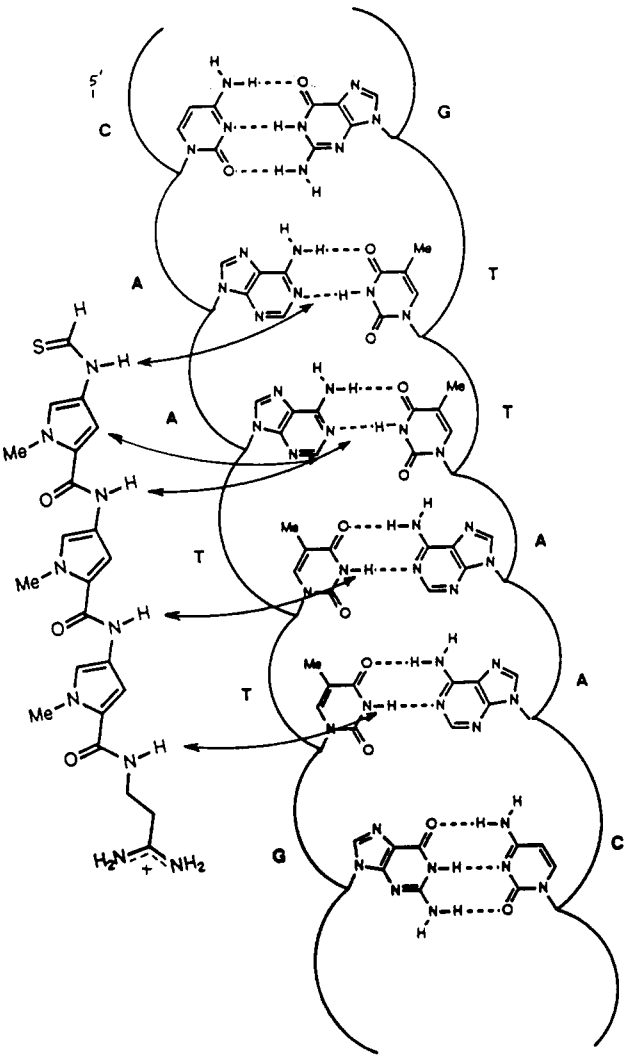


FIGURE 9: Proposed model for the binding of thioformyldistamycin with d(CGCAATTGCG)$_2$, depicting the details at the molecular level. Only the expansion of the minor groove formed by the central 5′-CAATTG segment is shown for clarity. The arrows indicate the intermolecular NOE contacts involving the protons on the concave face of the ligand and selected oligonucleotide sites.

The same two signals were observed to coalesce at 326 K, and the free energy of activation (ΔG^*) for the exchange process is then calculated at coalescence by the equation (Gunther, 1980)

$$\Delta G^* = 4.576T_c[10.32 + \log (T_c/k_c)] \quad (\text{kcal·mol}^{-1}) \qquad (1)$$

where k_c is the rate constant for a two-site exchange at the

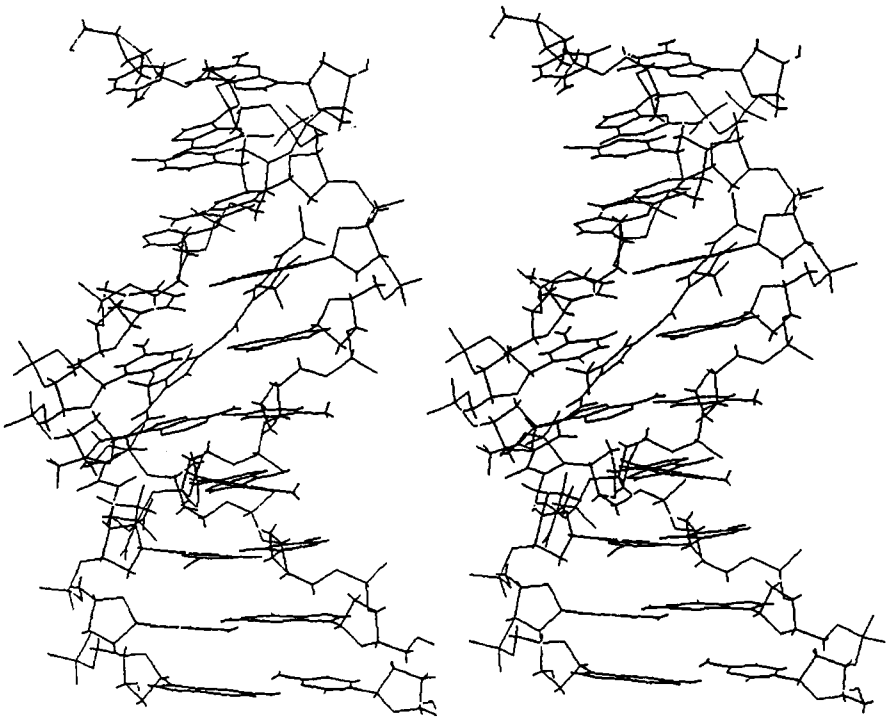


FIGURE 10: Skeletal depiction of an energy-minimized 1:1 complex of ligand–DNA built according to NMR–NOE data. To calculate the nonbonded interactions, a smooth cutoff extending up to 15 Å was used. A distance-dependent dielectric constant of $\epsilon = 2r$ was used to calculate electrostatic energies.

coalescence temperature (T_c) and is estimated from the standard equation (Sutherland, 1971)

$$k_c = \sqrt{2}(\pi/2)|\Delta\nu| \quad (2)$$

where $|\Delta\nu|$ is the chemical shift difference in hertz between two exchanging sites.

From the experimental values of 18.9 Hz and 326 K, the parameters k_c and $\Delta G^{\ddagger}$ are estimated to be 42.1 s^{-1} and 16.7 kcal·mol^{-1}, respectively. A similar analysis on another set of signals corresponding to T(7)Me protons (separated by 17.9 Hz and observed to coalesce at 321 K) gives the values of 39.7 s^{-1} and 16.5 kcal·mol^{-1} for the exchange process. Essentially identical values for the rate constant and free energy of activation imply that the two coalescence phenomena are due to the same exchange process. The mechanism for this exchange might involve a transient dissociation of the complex followed by a reorientation of the ligand before binding again to the second equivalent site.

In order to determine any possibility of a 2:1 ligand–DNA complex formation, a second mole equivalent amount of thioformyldistamycin was added to the sample. This did not produce any changes in the ^{1}H-NMR spectrum, and it is thus concluded that even under forcing conditions thioformyldistamycin is not able to form a 2:1 complex in contrast with the parent antibiotic distamycin A itself (Pelton & Wemmer, 1988, 1989).

Molecular Mechanics and Modeling. The energy-minimized structure of the ligand–DNA complex is shown in Figure 10. The ligand is tightly bound in the minor groove at the AATT site. The thioformyl end lies in the region between the base pairs (III) CG and (IV) AT. In this position the sulfur atom forms a weak hydrogen bond with the hydrogen of the guanine. This type of hydrogen bonding is plausible due to the polarizability of the thioamide group. The bulk of the sulfur atom is located between the base pairs, and it is found to make an angle of ca. 90° with the plane of the adjacent pyrrole ring. Thioformyldistamycin, being a flexible ligand, assumes a slightly twisted conformation to curve along the DNA helix and thereby achieves a snug fit. The conformation of the thioformyl end of the ligand is only slightly different from the energy-minimized conformation of the free ligand itself (Zimmermann et al., 1991). In this arrangement the amide groups of the ligand are in close proximity to several potential hydrogen-bond acceptor sites in the DNA strands. Thus N1H···A4N3, N3H···A5N3, and a three-center bifurcated hydrogen bond N5H···T6O2 and ···T6O2 of the second strand represent three rather strong hydrogen bonds involving the ligand and DNA. It is further observed that all the methyl groups of the ligand are directed away from the floor of the minor groove. The above results are influenced by the starting structure to some extent, and as reported under Materials and Methods, we used a model-built structure for the DNA rather than one based exclusively on NOE data.

Discussion

Previous NMR and X-ray studies have indicated that distamycin, containing three *N*-methylpyrrole rings, fits along the minor groove formed by AATT segments. The apparent isohelicity of the ligand with DNA, its hydrogen-bonding ability, the charge-based stabilization of the AT-bound regions all seem to contribute to this inherent specificity (Pullman, 1989; Lown, 1988, 1990). A number of lexitropsins have been reported in an attempt to alter this AT specificity to accommodate GC recognition. These analogs have largely evolved through the structural modifications of the central pyrrole rings, e.g., replacement by imidazole or furan rings [reviewed by Lown (1989)]. One key issue concerning the biological potency, however, pertains to the stability of such peptidic ligands toward hydrolytic and/or enzymatic degradation. In order to circumvent this problem, we recently reported the synthesis and peptidase stability of an analog of distamycin A, obtained by replacing the formyl end group by a thioformyl group (Zimmermann et al. 1991). This report is an extension of that study and provides detailed information on the DNA binding characteristics of thioformyldistamycin, elucidated

from its interaction with a synthetic decadeoxynucleotide d(CGCAATTGCG)$_2$.

Hindered rotation around the nitrogen to carbonyl bonds in amides is a well-studied phenomenon, and analogous results have also been reported for several unsymmetrically substituted secondary and tertiary thioamides. Similarly, in the case of thioformyldistamycin, a significantly slower rotation (on NMR time scale) around the N–C bond permits separate signals to be observed for the *N*-pyrrole and –C(=S)H groups situated syn and anti relative to each other, thereby giving rise to two distinctly rigid geometric isomers, *E* and *Z* (Figure 3). Although the *E*- and *Z*-isomers are shown to be interconverting, the equilibrium lies completely toward the *Z*-configuration in the presence of duplex oligonucleotide d(CGCAATTGCG)$_2$.

The restricted rotation in turn also suggests a strong preference for a delocalized resonance structure like **3** with an extreme case for such polarization (at least in principle) being a tautomeric change into the thiolimide form **4** (Figure 3). It is noteworthy that such high polarizability of the thioamide bond is expected to lower its pK_a value considerably, making it a better hydrogen-bond donor than the parent (oxo)amide counterpart, which presumably facilitates a tighter anchoring of thioformyldistamycin to the AATT segment of DNA at the thioformyl (N1H) end (vide supra).

An NMR titration of the symmetric decanucleotide sequence used in this study with the asymmetric ligand thioformyldistamycin to a 1:1 molar ratio led to an asymmetric ligand-bound complex which is characterized by a doubling of selected DNA proton resonances along much of the sequence. This effect is more pronounced for the central AATT segment of the DNA and is reflected by changes observed in both nonexchangeable and imino proton signals. This contention is further corroborated by a number of intermolecular contacts as identified from the NOE experiments. The retention of a duplex B-DNA-like structure upon ligand binding is evident from a comparison of the relative intensities of characteristic intra- and interresidue NOESY cross-peaks (Kearns, 1987; Reid, 1987). Further, the observation of NOEs between the ligand and the minor groove of the oligonucleotide (Table II) provides direct evidence for a tight binding of (*Z*)-thioformyldistamycin. The tight binding of the ligand is also supported by the appearance and line sharpening of all five imino proton signals even at 294 K, in contrast to their observation in the ligand-free oligonucleotide only at 277 K. This is consistent with a kinetic stabilization of the Watson–Crick base-paired duplex structure where the tightly binding ligand displaces the spine of hydration along the minor groove, thereby decelerating the exchange of the imino protons with the bulk solvent molecules. A similar model has been reported on such dehydration upon complexation of netropsin to a DNA dodecamer (Pardi et al., 1983). It should be noted that the stabilization of the duplex by thioformyldistamycin is consistent with our previous biophysical study (Zimmermann et al., 1991) which showed ΔT_m values of ~20 °C for poly(A·T) and ct-DNA.

Additionally, the ligand shows binding at two equivalent sites in the DNA sequence which are in slow exchange with each other. NMR line-shape analysis provides further information on the dynamics of exchange phenomenon, and average values of 41 s^{-1} and 16.6 kcal·mol^{-1} are estimated for the rate constant and free energy of activation at coalescence temperature.

Recent evidence suggests that under certain conditions distamycin forms a 2:1 complex containing two ligand molecules per duplex, aligned side by side in an antiparallel fashion (Pelton & Wemmer, 1988, 1989). The two molecules in such complexes have such orientations that the amidinium side chains extend into the GC base pairs around the central AT-rich regions where the minor groove is wider and thus accommodates the bulky portions of the molecules. More recent studies (Dwyer et al., 1992) indicate another kind of a 2:1 complex by incorporating a GC base pair in the middle of an AT tract, thereby increasing the width of the minor groove in the central section of the oligomer sequence. Challenging such a sequence with a lexitropsin (i.e., synthetic distamycin analog), obtained by replacing the central *N*-methylpyrrole unit by an allosteric *N*-methylimidazole ring, results in a stable 2:1 complex similar to that reported for distamycin itself. Thioformyldistamycin has a bulky thioformyl group at one end and therefore should also extend into the wider regions formed by GC base pairs. This indeed occurs with the present sequence; however, the 2:1 complex does not form. As pointed out by one of the reviewers, such a direct comparison of binding modes for thioformyldistamycin and distamycin should be interpreted with caution since the ease of 2:1 complexation of distamycin with AT-rich sequences appears to follow the order AAATTT > AAATT > AATT (Pelton & Wemmer, 1989) although the ligand binding site is 4 base pairs long. The lack of a 2:1 binding mode for thioformyldistamycin can be understood on the grounds that an antiparallel arrangement of two ligand molecules aligned side by side shows an unfavorable clash of the amidine side chain with the bulky sulfur atom which is on the outside, and as suggested by molecular modeling, at an angle orthogonal to the rest of the molecule.

References

Arcamone, F., Orezzi, P. G., Barbieri, W., Nicolella, V., & Penco, S. (1967) *Gazz. Chim. Ital. 97*, 1097–1109.

Baguley, B. (1982) *Mol. Cell. Biochem. 43*, 167–181.

Bodenhausen, G., Kogler, H., & Ernst, R. R. (1984) *J. Magn. Reson. 58*, 370–388.

Coll, M., Frederick, C. A., Wang, A. H.-J., & Rich, A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8385–8389.

Coll, M., Aymami, J., van der Marel, G. A., van Boom, J. H., Rich, A., & Wang, A. H.-J. (1989) *Biochemistry 28*, 310–320.

Diana, G. D. (1973) *J. Med. Chem. 16*, 857–859.

Dwyer, T. J., Geierstranger, B. H., Bathini, Y., Lown, J. W., & Wemmer, D. E. (1992) *J. Am. Chem. Soc.* (in press).

Gronenborn, A. M., & Clore, G. M. (1985) *Prog. NMR Spectrosc. 17*, 1–32.

Gunther, H. (1980) in *NMR Spectroscopy*, p 234, Wiley, New York.

Hahn, F. E. (1975) in *Antibiotics III. Mechanism of Action of Antimicrobial and Antitumor Agents* (Corcoran, J. W., & Hahn, F. E., Eds.) p 79, Springer-Verlag, New York.

Hahn, F. E. (1980) in *Inhibitors of DNA and RNA Polymerase: International Encyclopedia of Pharmacology and Therapeutics: Section 103* (Sarin, P. S., & Gallo, R. C., Eds.) pp 225–235, Pergamon Press, Great Britian.

Hore, P. J. (1983a) *J. Magn. Reson. 53*, 283–300.

Hore, P. J. (1983b) *J. Magn. Reson. 54*, 539–542.

Julia, M., & Preau-Joseph, N. (1963) *C. R. Acad. Sci. 257*, 1115–1121.

Kearns, D. R. (1987) in *Two-Dimensional NMR Spectroscopy: Applications for Chemists and Biochemists* (Croasman, W. R., & Carlson, R. M. K., Eds.) pp 301–347, VCH Publishers, New York.

Klevit, R. E., Wemmer, D. E., & Reid, B. R. (1986) *Biochemistry 25*, 3296–3303.

Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., & Dickerson, R. E. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1376–1380.

Kumar, S., Bathini, Y., Zimmermann, J., Pon, R. T., & Lown, J. W. (1990) *J. Biomol. Struct. Dyn. 8*, 331–357.
Lee, M., Chang, D. K., Hartley, J. A., Pon, R. T., Krowicki, K., & Lown, J. W. (1988a) *Biochemistry 27*, 445–455.
Lee, M., Pon, R. T., Krowicki, K., & Lown, J. W. (1988b) *J. Biomol. Struct. Dyn. 5*, 939–950.
Lown, J. W. (1988) *Anti-Cancer Drug Des. 3*, 25–40.
Lown, J. W. (1989) *Org. Prep. Proceed. Int. 21*, 1–46.
Lown, J. W. (1990) in *Molecular Basis of Specificity in Nucleic Acid–Drug Interactions* (Pullman, B., & Jortner, J., Eds.) pp 103–122, Kluwer Academic Publishers, The Netherlands.
Lown, J. W., & Krowicki, K. (1985) *J. Org. Chem. 50*, 3774–3779.
Marion, D., & Wuthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Nakamura, S., Karasawa, K., Yonehara, H., Tanake, N., & Umezawa, H. (1961) *J. Antibiot. (Tokyo), Ser. A 14*, 103.
Pardi, A., Morden, K. M., Patel, D. J., & Tinoco, I. (1983) *Biochemistry 22*, 1107–1121.
Patel, D. J. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 6424–6428.
Patel, D. J., & Shapiro, L. (1986a) *J. Biol. Chem. 261*, 1230–1240.
Patel, D. J., & Shapiro, L. (1986b) *Biopolymers 25*, 707–727.
Pelton, J. G., & Wemmer, D. E. (1988) *Biochemistry 27*, 8088–8096.
Pelton, J. G., & Wemmer, D. E. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5723–5727.
Probst, G. W., Hoehn, M. M., & Woods, B. L. (1965) *Antimicrob. Agents Chemother.*, 789.
Pullman, B. (1989) *Adv. Drug Res. 18*, 1–113.
Reid, B. R. (1987) *Q. Rev. Biophys. 20*, 1–34.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.
Sutherland, I. O. (1971) *Annu. Rep. NMR Spectrosc. 4*, 71–235.
Takahashi, T., Sugawara, Y., & Susuki, M. (1973) *Tetrahedron Lett. 19*, 1873.
Wartell, R. M., Larson, J. E., & Wells, R. D. (1974) *J. Biol. Chem. 249*, 6719–31.
Zimmer, Ch., & Wahnert, U. (1986) *Prog. Biophys. Mol. Biol. 47*, 31–112.
Zimmer, Ch., Puschendorf, B., Grunicke, H., Chandra, P., & Venner, H. (1971) *Eur. J. Biochem. 21*, 269–278.
Zimmer, Ch., Luck, G., Birch Hirshfield, E., Weiss, R., Arcamone, F., & Guschlbauer, W. (1983) *Biochim. Biophys. Acta 741*, 15–22.
Zimmermann, J., Rao, K. E., Joseph, T., Sapse, A.-M., & Lown, J. W. (1991) *J. Biomol. Struct. Dyn. 9*, 599–611.

Nucleic Acid Interactive Properties of a Peptide Corresponding to the N-Terminal Zinc Finger Domain of HIV-1 Nucleocapsid Protein†

Martha D. Delahunty, Terri L. South, Michael F. Summers, and Richard L. Karpel*
Department of Chemistry and Biochemistry, University of Maryland Baltimore County, Baltimore, Maryland 21228
Received December 3, 1991; Revised Manuscript Received February 14, 1992

ABSTRACT: An 18-residue peptide (NC-F1) with an amino acid sequence corresponding to the N-terminal zinc finger of human immunodeficiency virus-1 nucleocapsid protein has been shown to bind to nucleic acids by fluorescence and NMR methods. Previously, this peptide has been shown to fold into a defined structure when bound to zinc (Summers et al., 1990). We have used a fluorescent polynucleotide, poly(ethenoadenylic acid), to monitor binding of this peptide to nucleic acids. In the presence of zinc, the peptide had a smaller site size (1.75 nucleotide residues/peptide) than in the absence of the metal ion (2.75). The salt sensitivity of the interaction indicated that two ion pairs are involved in the association of Zn^{2+}(NC-F1) with polynucleotide, whereas one ion pair is found in the metal-free peptide–nucleic acid complex. Competition experiments with single-stranded DNA (ss DNA) in either the presence or absence of Zn^{2+} showed that the peptide bound to ss DNA. Using NMR methods, we monitored the binding of a synthetic oligonucleotide, d(TTTGGTTT), to Zn(NC-F1). The hydrophobic residues F_2 and I_{10}, which are on the surface of the peptide and have been implicated in viral RNA recognition, were shown to interact with the oligomer. In accord with this observation, analysis of the salt dependence of the polynucleotide–peptide interaction indicates a nonelectrostatic component of about −6 kcal/mol, a value consistent with theoretical estimates of stacking energies of phenylalanine with nucleic acid bases.

The nucleocapsid (NC) proteins of retroviruses have been shown to bind to nucleic acids, but the full extent of their participation in the life cycle of the virus remains to be explicated. NC proteins are associated with the RNA genome in the core of the viral particle (Dickson et al., 1984; Karpel et al., 1987). They have also been shown in vitro to promote the annealing of primer tRNA to viral RNA (Barat et al., 1989) and viral RNA dimerization (Prats et al., 1988). Single-stranded (ss) nucleic acids bind nonspecifically to these proteins, with a 15–30-fold greater affinity than double-stranded (ds) nucleic acids (Davis et al., 1976).

Retroviral NC proteins have one or two copies of a conserved sequence that is similar to the zinc binding domain of the eukaryotic transcription factor TFIIIA (Berg, 1986). Classical zinc finger proteins, such as TFIIIA, have a Cys-Cys-His-His motif and form a stable structure when a tetrahedrally coordinated zinc is bound (Miller et al., 1985; Lee et al., 1989). The conserved sequence in retroviral NC proteins is of the Cys-Cys-His-Cys type [C-$(X)_2$-C-$(X)_4$-(H)-$(X)_4$-C]

† This work was supported by grants from the American Cancer Society (NP-671 to R.L.K.) and the National Institutes of Health (GM42561 and AI30917 to M.F.S.).
* Corresponding author.

0006-2960/92/0431-6461$03.00/0