© 1998 Oxford University Press
Nucleic Acids Research, 1998, Vol. 26, No. 3 **831–838**

Solution structure of a highly stable DNA duplex conjugated to a minor groove binder

Surat Kumar[+], Michael W. Reed[1], Howard B. Gamper Jr[1], Vladimir V. Gorn[1], Eugeny A. Lukhtanov[1], Matthew Foti, John West, Rich B. Meyer Jr[1] and Barry I. Schweitzer*

Walt Disney Memorial Cancer Institute at Florida Hospital, 12722 Research Parkway, Orlando, FL 32826, USA and [1]Epoch Pharmaceuticals Inc., 1725 220th Street SE, #104, Bothell, WA 98021, USA

Received August 28, 1997; Revised and Accepted December 1, 1997

PDB 1aul.pdb.

ABSTRACT

The tripeptide 1,2-dihydro-(3*H*)-pyrrolo[3,2-*e*]indole-7-carboxylate ($CDPI_3$) binds to the minor groove of DNA with high affinity. When this minor groove binder is conjugated to the 5′-end of short oligonucleotides the conjugates form unusually stable hybrids with complementary DNA and thus may have useful diagnostic and/or therapeutic applications. In order to gain an understanding of the structural interactions between the $CDPI_3$ minor groove binding moiety and the DNA, we have determined and compared the solution structure of a duplex consisting of oligodeoxyribonucleotide 5′-TGATTATCTG-3′ conjugated at the 5′-end to $CDPI_3$ and its complementary strand to an unmodified control duplex of the same sequence using nuclear magnetic resonance techniques. Thermal denaturation studies indicated that the hybrid of this conjugate with its complementary strand had a melting temperature that was 30°C higher compared with the unmodified control duplex. Following restrained molecular dynamics and relaxation matrix refinement, the solution structure of the $CDPI_3$-conjugated DNA duplex demonstrated that the overall shape of the duplex was that of a straight B-type helix and that the $CDPI_3$ moiety was bound snugly in the minor groove, where it was stabilized by extensive van der Waal's interactions.

INTRODUCTION

Long single-stranded DNA and RNA have considerable secondary and tertiary structure that can interfere with the use of oligodeoxyribonucleotides (ODNs) as diagnostic probes or antisense agents (1–3). One approach to overcoming these barriers involves the design of chemically modified ODNs that form unusually stable hybrids or triplexes that can favorably compete with existing secondary structures. Examples include ODNs conjugated to an intercalator (4), peptide nucleic acids (5) and ODNs containing C-5 propynyl (6) or N3′-P5′ (7) groups.

ODNs conjugated to a minor groove binder represent another strategy to form stabilized hybrids. (+)-CC-1065 (Fig. 1) is a natural product that binds to and alkylates DNA in the minor groove and exhibits potent antitumor and antibiotic activity (8–9). Recently conjugation of *N*-3-carbamoyl-1,2-dihydro-3*H*-pyrrolo[3,2-*e*]indole-7-carboxylate ($CDPI_3$; Fig. 1), a non-reactive analog of CC-1065, to the ends of various ODNs has been reported (10). This conjugate was shown to increase the melting temperature of DNA hybrids by as much as 44°C when covalently linked to one strand of an A/T-rich duplex. These conjugates also formed stabilized hybrids with complementary RNA targets and with G/C-rich DNA (11). Subsequent experiments demonstrated that $CDPI_3$ covalently linked to the 5′-end of ODNs effectively blocked primer extension by a DNA polymerase (12). Polymerase blockage was abolished when a single mismatch was introduced into the oligomer or when the minor groove binder was either removed or replaced by a 5′-acridine group. More recently $CDPI_3$ conjugates as short as 8–10mers have been used as primers in PCR to amplify DNA with good specificity and efficiency (13). These results and the other properties mentioned above suggest that DNA conjugates of this type may have useful diagnostic and/or therapeutic applications.

There is great potential for the versatility of DNA conjugates to be significantly enhanced by refining the tethered minor groove binding moiety. In order for this refinement to be carried out in a rational manner it is important to gain an understanding of the structural interactions between the minor groove binding moiety and the DNA. In the present study we have determined the solution structure of a duplex consisting of an oligodeoxyribonucleotide 5′-TGATTATCTG-3′ conjugated at the 5′-end to $CDPI_3$ and the complementary strand. We compare this structure with the structure of an unmodified control duplex of the same sequence in an effort to explain the extraordinary stability of the hybrid.

MATERIALS AND METHODS

Sample preparation

Synthesis of the DNA oligomer d(TGATTATCTG) containing a minor groove binding drug, $CDPI_3$, conjugated to the 5′-end of the dT1 residue with a hexamethylene linker and its complementary

*To whom correspondence should be addressed. Tel: +1 407 380 9977; Fax: +1 407 380 9978; Email: barry@sneezy.fhis.net

[+]Present address: Department of Chemistry, Long Island University, Brooklyn, NY 11201, USA

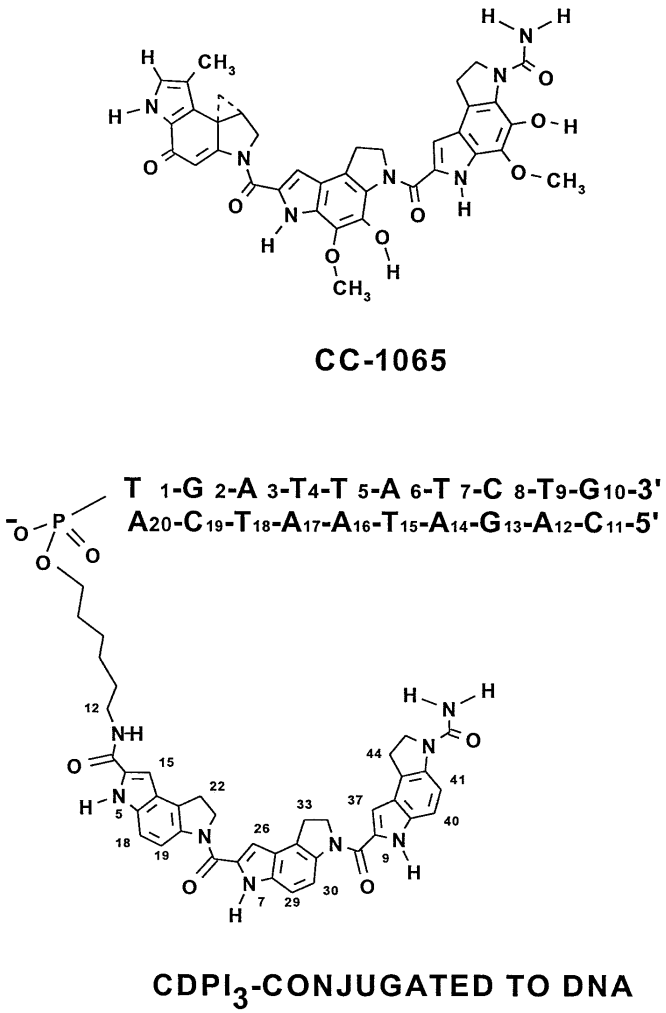

Figure 1. Structure of CC-1065 (top) and the DNA duplex conjugated to *N*-3-carbamoyl-1,2-dihydro-3*H*-pyrrolo[3,2-*e*]indole-7-carboxylate ($CDPI_3$) (bottom).

sequence d(CAGATAATCA) was accomplished as reported earlier, with some modifications (10). Specifically, two methods for preparation of the 5′-$CDPI_3$ conjugates were compared. The first method used a 10 µmol trityl-off synthesis of the 5′-aminohexyl-modified ODN and treatment of the CPG with 5 equiv. (38 mg) $CDPI_3$ activated ester. Ammonia hydrolysis gave a mixture of products that were separated by reverse phase HPLC as described earlier to give 1.54 mg (5% yield) of the desired conjugate. The second method used HPLC-purified 5′-aminohexyl-modified ODN from a 10 µmol scale synthesis (7.4 mg, 2.3 µmol). The triethylammonium form of the ODN was dissolved in 0.5 ml DMSO and treated with 3 equiv. (6.3 mg) $CDPI_3$ activated ester and 50 µl triethylamine. After 22 h the crude conjugate was precipitated with 2% sodium perchlorate/acetone. Purification by reverse phase HPLC gave 3.96 mg product (42% yield from amine-modified ODN). Re-precipitation gave the sodium salt. The purified conjugates from each method were combined and used for further experiments. The unmodified control duplex was purchased from DNAgency Inc. The purity of these sequences was checked by ^{1}H NMR and the control and $CDPI_3$-modified duplexes were prepared by annealing each oligomer with its complementary sequence as described (14).

Thermal denaturation studies

Hybrids formed between minor groove binder conjugates or unmodified oligonucleotides and their complements were melted at a rate of 0.5°C/min in 1× PBS (9.2 mM disodium phosphate, 0.8 mM monosodium phosphate, 131 mM NaCl, pH 7.2) on a Lambda 2S (Perkin Elmer) spectrophotometer with a PTP-6 automatic multicell temperature programer. Each oligonucleotide (4 µM) was mixed with sufficient complement to give a 1:1 ratio. Prior to melting samples were denatured at 100°C and then cooled to the starting temperature over a 10 min period. The melting temperatures of the duplexes were determined from the derivative maxima.

NMR experiments

All NMR experiments were performed on a Varian 600 MHz Unity Plus NMR spectrometer at a regulated temperature of 20°C. NOESY, DQF-COSY and TOCSY data were acquired in the phase-sensitive mode with 2048 (4096 for the DQF-COSY) complex points in t_2 and 512 (1024 for the DQF-COSY) complex points in t_1, a relaxation delay of 2.5 s (during which time the residual HDO peak was irradiated) and a spectral width of 6250 Hz. The TOCSY experiment was carried out with a 120 ms mixing time. Four NOESY spectra for the control and $CDPI_3$-conjugated DNA samples with mixing times of 80, 120, 160 and 200 ms were each collected within separate 3 day periods, without removing the sample from the spectrometer. A 200 ms NOESY spectra for the control and the $CDPI_3$-conjugated DNA in H_2O were acquired with an excitation sculpting pulse sequence for solvent suppression (15). Selective 180° pulses were applied with a SEDUCE profile to suppress the water resonance. Gradient pulses were applied along the *z*-axis for 1 ms at 15 and 3 G/cm with 50 µs delays before and after the gradient pulses. These experiments were collected with States-TPPI phase cycling with 2048 complex points in t_2 and 1024 complex points in t_1, a relaxation delay of 2.5 s and a spectral width of 12 500 Hz.

NMR data were processed using FELIX 95.0 (Molecular Simulations Inc.). NOESY and TOCSY data sets acquired in D_2O or H_2O were processed with a squared sinebell shifted 90° window function in the direct dimension and a squared sinebell shifted 90° window function in the indirect dimension. Data in both dimensions were zero filled to 2048 real points, Fourier transformed, phased to pure absorption and baseline corrected with a fifth order polynomial. NOE cross-peak volumes were obtained by cross-peak integration using FELIX software. The DQF-COSY data sets were apodized with unshifted sinebells in both dimensions, followed by zero filling to 2048 real points in both dimensions prior to Fourier transformation. ^{1}H resonances are referenced to an external sample of sodium 2,2-dimethyl-2-silapentane-5-sulfonate in identical buffer.

Structure calculations

Interproton distances were calculated from initial NOE build-up rates by fitting the cross-peak volumes at 80, 120, 160 and 200 ms mixing times to a linear equation. These rates (R_{ij}) were converted to distances (r_{ij}) using the cytosine H5–H6 distance of 2.45 Å as the reference distance (r_{ref}), the averaged rate calculated from well-

resolved H5–H6 cross-peaks (R_{ref}) and the relationship $r_{ij} = r_{ref}(R_{ref}/R_{ij})^{1/6}$. Cross-peak volumes obtained from the exchangeable protons in the 200 ms NOESY acquired in H_2O were also referenced to the H5–H6 interproton distance of 2.45 Å. These distances were given upper and lower bounds according to the following guidelines: for $r_{ij} < 2.0$ Å the errors given were –0.2/+0.2 Å; for 2.0 Å < r_{ij} < 2.5 Å they were –0.2/+0.4 Å; for 2.5 Å < r_{ij} < 3.3 Å they were –0.3/+0.5 Å; for 3.3 Å < r_{ij} < 6.0 Å they were –0.5/+0.7 Å (16,17). All distances involving exchangeable protons were given upper and lower bounds of 0.7 Å. Base pairs were kept hydorgen bonded in the Watson–Crick form by using distance restaints between the bases as previously described (16).

A starting structure for the $CDPI_3$-conjugated duplex was generated by constructing the DNA duplex in canonical B-type form and then docking the $CDPI_3$ moiety into the minor groove of the duplex using InsightII software (Molecular Simulations Inc.). The structures for both the control and $CDPI_3$ duplexes were refined by restrained molecular dynamics using the program X-PLOR 3.1 (18) as previously described (16). A total of 210 or 175 experimentally determined interproton distance restraints were used in the control and $CDPI_3$ refinements, respectively. At this stage of the calculations an effort was made to eliminate those distances from interproton interactions where contamination from spin diffusion was most likely [e.g. H3′(*i*)–H8/H6(*i+1*)]. As described in more detail below, sugar torsion angles were approximated by a qualitative estimation of $J_{H2''–H3'}$ and $J_{H3'–H4'}$ from high resolution DQF-COSY data and from NOESY-determined distances involving sugar protons (19). In addition to the experimentally determined distance and sugar dihedral restraints, the α, β, γ, ε and ζ backbone torsion angles were restrained to a range covering all right-handed DNAs (16,17). The restrained molecular dynamics procedure was repeated six times from each starting structure, using different random number seeds for each subsequent molecular dynamics run. The structure calculations were carried out in an iterative manner by adding more $CDPI_3$–DNA NOEs to each cycle to resolve potentially ambiguous assignments. Each restrained molecular dynamics structure was further refined using complete relaxation matrix refinement using the RELAX option of X-PLOR as reported earlier (16); a total of 1080 or 840 NOE intensities were used in refinement of the control and $CDPI_3$ structures, respectively. No backbone torsion angles were used in this phase of the calculations. Analyses of helical parameters were performed with the program Dials-and-Windows (kindly provided by Dr G.Ravishanker, Wesleyan University) (20).

RESULTS AND DISCUSSION

In order to understand the mode of binding, stability and sequence specificity of DNA conjugated to $CDPI_3$ we have undertaken a structural study of a 5′-$CDPI_3$ conjugate of the DNA decamer 5′-TGATTATCTG-3′ hybridized to its complementary sequence 3′-ACTAATAGAC-5′. Preparation of these minor groove binder–oligodeoxyribonucleotide conjugates has been previously reported by Lukhtanov *et al.* (10) and was accomplished by reaction of the 2,3,5,6-tetrafluorophenyl ester of $CDPI_3$ with a 5′-aminohexyl phosphate ester. Two methods were compared for conjugating $CDPI_3$ to the amine-modified ODN. Treatment of the CPG-immobilized ODN was less labor intensive, but HPLC separation of the desired product was more difficult. Modification of pre-purified amine-modified ODN was found to give easier isolation, better yields and consumed less of the valuable $CDPI_3$ activated ester. UV melting studies on the conjugated DNA indicated a melting temperature of 61°C for this duplex versus a melting temperature of 30°C for an unmodified control duplex of the same sequence. This large increase in stability of the conjugated duplex was comparable with increases previously observed in other $CDPI_3$-conjugated duplexes (10–13).

NMR resonance assignments and chemical shift analysis

The non-exchangeable base protons (purine H8 and pyrimidine H6) and the sugar protons (H1′, H2′, H2″, H3′ and H4′) for the control and $CDPI_3$-conjugated duplexes were assigned on the basis of an analysis of through-space distance connectivities in NOESY data sets as a function of mixing time and through-bond connectivities in COSY data sets recorded in D_2O buffer. The expanded NOESY (400 ms mixing time) contour plot establishing sequential connectivities between the base protons and the sugar H1′ protons of the $CDPI_3$-conjugated duplex is plotted in Figure 2. A similar sequential tracing could be made from the base protons to the H2′/H2″ protons (not shown).

Exchangeable proton (amino and imino) resonance assignments were made in 90% H_2O/10% D_2O at 10°C using a gradient-enhanced NOESY experiment that employs a double spin-echo sequence for water suppression with minimal saturation exchange (15). Imino–imino and imino–H2 proton NOEs observed in these spectra indicated normal base stacking in both the control and $CDPI_3$-conjugated duplexes. Furthermore, interstrand imino proton and H2 proton to amino proton NOEs indicated normal Watson–Crick base pairing in both DNAs.

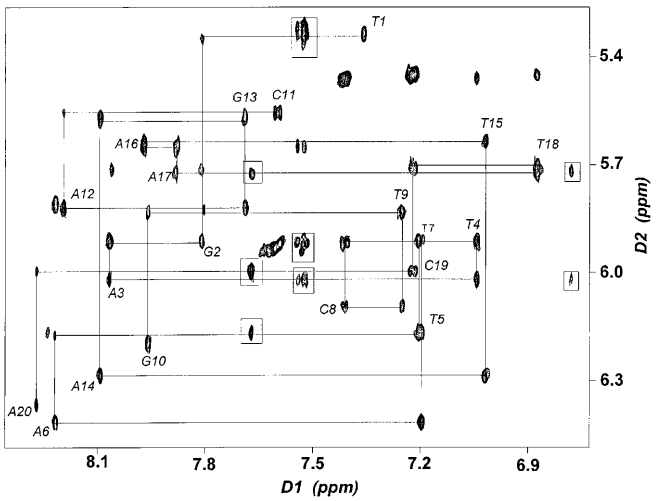

Figure 2. Expanded NOESY (400 ms mixing time) contour plot of the $CDPI_3$-conjugated duplex in D_2O at 25°C. The sequential connectivity of the base H8/H6 and deoxyribose H1′ protons is diagramed. Boxed cross-peaks at 6.79, 7.53 and 7.68 p.p.m. (D1) represent NOEs between DNA protons and protons H26, H15 and H37 of the $CDPI_3$ moiety, respectively.

Table 1. $CDPI_3$ ^1H chemical shifts in the $CDPI_3$-conjugated duplex

Residue	H6/H8	H5/CH_3	H2	H1′	H2′	H2″	H3′	H4′	Imino
T1	7.36	1.31		5.33	1.82	2.16	4.58	3.95	
G2	7.81			5.93	2.08	2.45	5.24	4.18	12.97
A3	8.08		6.00	6.03	2.38	2.65	5.34	4.18	
T4	7.05	1.16		5.93	1.95	2.23	5.49	n.a.	14.12
T5	7.21	1.60		6.18	2.45	2.86	4.98	3.70	13.70
A6	8.23		5.60	6.42	2.67	2.96	5.03	4.03	
T7	7.21	1.38		5.92	1.94	2.23	4.62	n.a.	13.42
C8	7.42	5.47		6.10	1.94	2.42	5.00	3.90	
T9	7.25	1.65		5.85	2.01	2.37	4.89	4.06	14.04
G10	7.96			6.20	2.39	2.68	4.73	4.20	
C11	7.60	5.94		5.54	1.65	2.19	4.64	4.20	
A12	8.20		6.46	5.84	2.73	2.83	5.00	4.07	
G13	7.70			5.57	2.55	2.73	4.99	4.38	12.65
A14	8.10		5.75	6.30	2.54	2.96	5.03	4.23	
T15	7.03	1.42		5.65	2.03	2.37	5.07	3.96	13.28
A16	7.98		6.17	5.66	2.29	2.38	4.65		
A17	7.89		6.08	5.74	2.18	2.42	4.70	4.23	
T18	6.88	1.19		5.71	1.76	1.99	5.48		13.62
C19	7.23	5.47		6.01	1.80	1.88	4.52		
A20	8.28			6.37	2.48	2.64	4.92	4.03	

Table 2. $CDPI_3$ ^1H chemical shifts in the $CDPI_3$-conjugated duplex

$CDPI_3$ proton	Chemical shift (p.p.m.)
C15-H	7.53
C18-H	7.62
C19-H	8.31
C22-H1	3.61
C22-H2	3.75
C23-H1	3.13
C23-H2	2.45
C26-H	6.79
C29-H	7.49
C30-H	8.27
C37-H	7.68
C40-H	7.49
C41-H	7.93

Assignment of the $CDPI_3$ ^1H resonances was accomplished using DQF-COSY, TOCSY and NOESY experiments. The J-coupled aromatic protons (H18–H19, H29–H30 and H40–H41) could be identified by the presence of COSY, TOCSY and NOESY cross-peaks near the diagonal in the aromatic region of the spectra. Assignment of these protons was based upon assignment of the ^1H NMR spectrum of the methyl ester form of free $CDPI_3$ in DMSO (10). These assignments should be considered as tentative, since the chemical shifts of these resonances could be affected by the solvent change from DMSO to H_2O. No restraints involving these protons were used in the structural calculations described below. Assignments of other $CDPI_3$ protons were assisted by correlating observed NOESY cross-peaks between the DNA and $CDPI_3$ moiety with those expected from our preliminary model of the $CDPI_3$-conjugated DNA duplex. Intraresidue NOEs for the $CDPI_3$ moiety (e.g. between H15 and H22, between H26 and H33 and between H37 and H44) also facilitated the assignment process. Assignments were confirmed by gradually including restraints involving these protons into the structural calculations (see Materials and Methods). Some aliphatic protons could not be unambiguously assigned due to overlap in the crowded aliphatic region of the DNA spectrum.

The DNA and $CDPI_3$ ^1H resonance assignments for the $CDPI_3$-conjugated duplex are shown in Tables 1 and 2, respectively. Some of the differences between the DNA ^1H chemical shifts in the $CDPI_3$-conjugated duplex versus those of the control duplex are plotted in Figure 3. While the changes in chemical shift are fairly complex, most of the differences are localized to protons in the minor groove of the duplex. These chemical shift changes and the appearance of NOE cross-peaks between $CDPI_3$ protons and DNA protons in the minor groove of the DNA support the presence of $CDPI_3$ in the minor groove of the DNA. In addition, most of the larger chemical shifts occur at residues 2–5 in the conjugated strand (Fig. 3A) and residues 16–19 in the complementary strand (Fig. 3B). Most of the chemical shifts in this region of the $CDPI_3$-conjugated duplex were upfield relative to the control duplex, which is consistent with a shielding effect of the aromatic rings of the $CDPI_3$ moiety on the DNA protons in this region. This effect was especially apparent in the complementary strand (Fig. 3B).

Solution structure of the $CDPI_3$-conjugated DNA duplex

Qualitative analysis of the NOE data obtained for both the control and $CDPI_3$-conjugated duplexes revealed the general pattern of $NOE_{H2'(i)–H8/H6(i)} >> NOE_{H2''(i-1)–H8/H6(i)} > NOE_{H2'(i-1)–H8/H6(i)}$. This pattern is characteristic of right-handed B-type structures. Analysis of the DQF-COSY data indicated that the control duplex contained a variety of sugar pucker conformations (C1′-*exo*, C2′-*endo* and O4′-*endo*) found in the B-type DNA family.

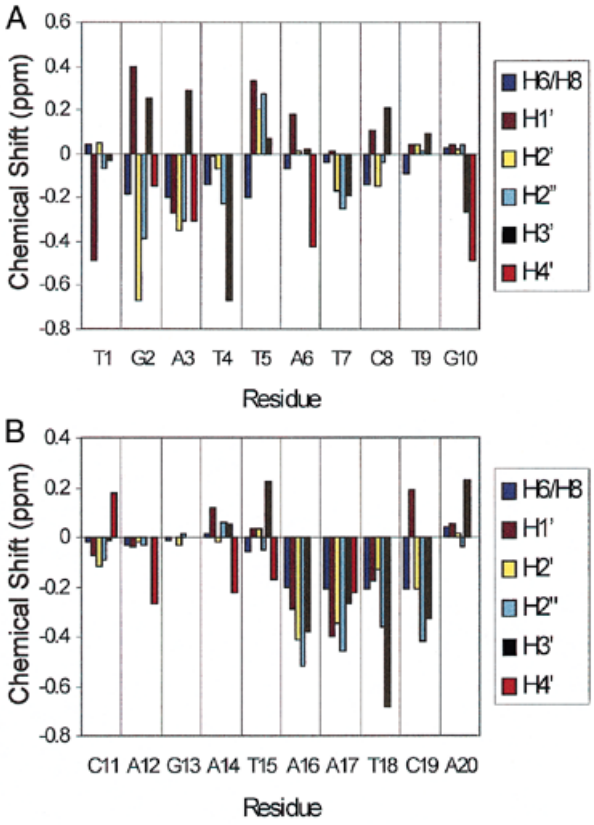

Figure 3. Histogram bar plot of differences in selected ^{1}H chemical shifts between the CDPI$_3$ conjugate and the non-conjugated control decamer duplexes. Positive shift differences represent a downfield shift of the resonance in the CDPI$_3$–decamer conjugate relative to the control duplex. (**A**) Chemical shift differences between the conjugated DNA strand and the homologous strand in the control duplex. (**B**) Chemical shift differences between the DNA strand complementary to the conjugated strand and the homologous DNA strand in the control duplex.

The complete absence of H3′–H4′ and H2″–H3′ cross-peaks in the DQF-COSY spectrum of the CDPI$_3$-conjugated duplex (data not shown) suggests that the sugar puckers of the modified DNA were primarily in a C2′-*endo* conformation (21). Based on these qualitative observations, B-type duplexes were used as starting structures for the molecular dynamics structure calculations. In these calculations sugar puckers were restrained with relatively wide bounds (P = 120–160°).

Using distance data, NOE cross-peak volumes and sugar pucker dihedral angle restraints obtained from DQF-COSY and NOESY experiments, structures of the modified and control duplexes were obtained using restrained molecular dynamics and complete relaxation matrix refinement (Fig. 4). The pairwise root mean square (r.m.s.) deviations for the set of control and CDPI$_3$-conjugated structures were 0.54 ± 0.13 and 0.83 ± 0.10 Å respectively. Structural statistics for the starting structures, ensemble of restrained molecular dynamics structures and ensemble of relaxation matrix-refined structures are shown in Table 3. The small deviations from idealized geometry indicate that the calculations have not resulted in any distortion of the covalent structure. Following relaxation matrix refinement there are small increases in the r.m.s. deviation from the experimental distance and dihedral restraints. This is due to the isolated spin approximation used to calculate the distances used in the restrained molecular dynamics calculations. The lack of backbone dihedral restraints during the relaxation matrix refinement procedure leads to the increase in the r.m.s. deviation from the dihedral angle restraints. The $R^{1/6}$ factors (which reflect the degree of agreement between experimental NOE volumes and the NOE volumes back-calculated from the structures) were reduced from 0.129 and 0.156 in the starting structures to 0.086 ± 0.001 and 0.083 ± 0.001 in the final ensembles of relaxation matrix-refined control and CDPI$_3$-conjugated duplex structures respectively. This reduction in $R^{1/6}$ factors for the refined structures in comparison with the starting structures demonstrates convergence to structures that are in good agreement with the experimental data.

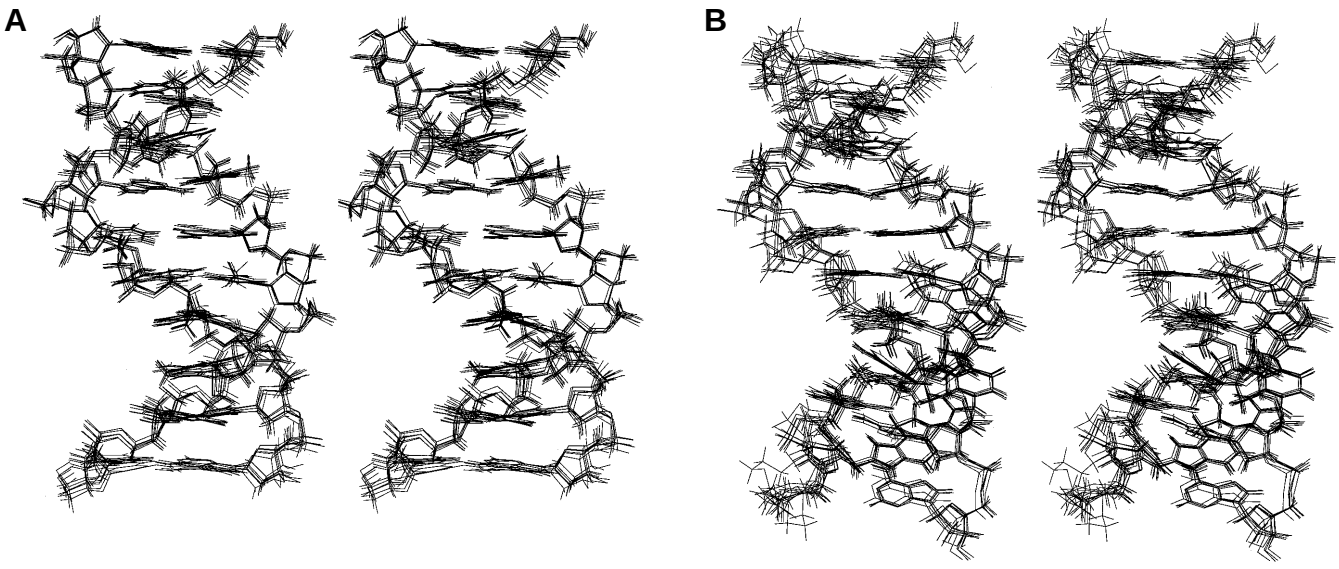

Figure 4. Stereoviews of the best fit superimpositions of the ensembles of relaxation matrix-refined control (**A**) and CDPI$_3$-conjugated (**B**) duplexes.

Table 3. Structural statistics for initial and calculated structures

	Initial[a]	rMD[b]	RMR[c]
Control			
r.m.s.deviation from experimental			
Distance restraints (Å)			
NOE (210) + H bond (25)	0.32	0.12 ± 0.004	0.23 ± 0.004
Dihedral restraints (°)	6.00	1.48 ± 0.20	4.17 ± 2.03
r.m.s. deviation from idealized covalent geometry			
Bonds (Å)	0.033	0.004 ± 0.0001	0.010 ± 0.0005
Angles (°)	2.80	0.62 ± 0.004	0.96 ± 0.01
Impropers (°)	2.18	0.23 ± 0.01	0.50 ± 0.04
$CDPI_3$-modified			
r.m.s.deviation from experimental			
Distance restraints (Å)			
NOE (175) + H bond (25)	0.77	0.26 ± 0.01	0.33 ± 0.007
Dihedral restraints (°)	6.90	2.73 ± 0.08	10.80 ± 1.42
r.m.s. deviation from idealized covalent geometry			
Bonds (Å)	0.009	0.008 ± 0.0001	0.010 ± 0.0002
Angles (°)	0.817	0.92 ± 0.003	1.26 ± 0.01
Impropers (°)	0.20	0.28 ± 0.02	0.72 ± 0.04

[a]Initial starting structure used in subsequent calculations.
[b]rMD, ensemble of six restrained molecular dynamics structures.
[c]RMR, ensemble of six relaxation matrix-refined structures.

The structure of the final $CDPI_3$-conjugated duplex (Fig. 5) has the overall appearance of right-handed B-DNA. However, we observed some of the following salient features. The bulk of the $CDPI_3$ moiety spans the region of DNA from residues 2–5 on the conjugated strand and 16–19 on the complementary strand, in agreement with the chemical shift data shown in Figure 3. Examination of Figures 4 and 5 shows that there is little overall distortion in the structure of the $CDPI_3$-conjugated DNA double helix compared with the control duplex. However, the T1·A20 base pair of the $CDPI_3$-conjugated duplex exhibited a much greater degree of incline, tip and buckle compared with the same base pair in the control structure. Mean values of the backbone torsion angles α, β, γ, δ and ζ, as well as the glycosidic torsion angle χ and the pseudorotation angle *P* for the ensembles of control and $CDPI_3$-conjugated duplexes are given in Table 4. Inspection of this table reveals that the greatest deviations from the control structure are located at the 5′-end of the $CDPI_3$-conjugated strand and the 3′-end of the unconjugated strand.

Comparison with other minor groove binders

The minor groove binding drugs netropsin (22), distamycin (23) and their structural analogs (24–25) have been reported to bind isohelically into the AT-rich minor groove of the DNA double helix, owing to their crescent-shaped structure. These molecules have planar *N*-pyrrole moieties and fit edge-on into the minor groove, replacing the spine of hydration and follow the natural curvature of the DNA. This type of molecular recognition is at least partially driven by formation of hydrogen bonds between the drug and the DNA.

In the case of CC-1065 binding of the drug in the minor groove of the DNA is accompanied by opening of the cyclopropyl ring of the drug to form a covalent bond with the N3 position of an adenine base. Using gel electrophoresis, hydroxyl radical footprinting and NMR studies, Hurley and co-workers have extensively characterized the structural consequences of CC-1065 bonding to DNA (26–28). They concluded from these studies that covalent modification of the DNA by CC-1065 results in kinking of the DNA by 17–22° with an overall direction and magnitude of bending similar to those of an A_6 tract sequence (27). It was also reported that a narrowing of the minor groove at the binding site occurred as a result of (+)-CC-1065 binding, possibly as a consequence of hydrophobic contacts between the drug and the minor groove of the DNA. This region of narrowing was sandwiched between widened regions (27). Finally, it was proposed that the inherent propensity of a DNA sequence to form a bent structure was related to its preference as a target site for binding of (+)-CC-1065 (28).

In an NMR study on another analog of CC-1065, CPI-$CDPI_2$, Powers and Gorenstein (29) reported the presence of two conformers of the drug–DNA adduct in solution. The major isomer had a kink of 60° at the site of alkylation, similar to the sequence-specific bends reported in several DNA structures (30–31). The general binding mode of the CPI-$CDPI_2$ moiety to the DNA was similar to that reported by Hurley and co-workers for the CC-1065 adduct, which was not surprising since both of these molecules formed an N3 adduct with adenosine. In the CPI-$CDPI_2$ structure, however, bending of the DNA was much greater than that observed in CC-1065-modified DNA structures; furthermore, the minor groove was widened at the site of CPI-$CDPI_2$ binding rather than narrowed, as in the case of CC-1065.

The molecule under study in this paper has two major features that distinguish it from those described above. First, no adduct is formed between the $CDPI_3$ moiety and any group in the minor groove of the DNA. Second, the $CDPI_3$ moiety is attached to one strand of the DNA duplex by a flexible hexamethylene linker.

Figure 5. Connolly surface representation of the averaged relaxation matrix-refined CDPI$_3$–decamer conjugate structure. The orientation of the structure shown in this figure is rotated 90° relative the view shown in Figure 4. The CDPI$_3$ moiety is shown in red. The ensemble of CDPI$_3$-conjugated DNA structures was averaged and the averaged structure was refined again using complete relaxation matrix refinement to generate the structure shown here. The coordinates for this structure have been deposited with the Brookhaven Protein Data Bank.

Table 4. Backbone torsion angles[a], glycosidic torsion angle χ and pseudorotation angle χ of relaxation matrix-refined structures[b]

residue	α	β	γ	δ	ε	ζ	χ	P
Section A: Control Duplex Structure								
T1	308[5]	196[2]	45[6]	135[1]	171[3]	267[3]	243[4]	141[1]
G2	308[9]	184[5]	35[7]	128[1]	177[4]	263[4]	255[1]	134[1]
A3	335[4]	134[9]	46[2]	134[1]	200[5]	220[8]	245[2]	145[2]
T4	302[11]	167[13]	47[7]	123[1]	195[13]	252[9]	243[2]	127[1]
T5	289[11]	200[8]	45[6]	131[2]	171[7]	265[6]	249[2]	133[2]
A6	353[15]	121[20]	40[4]	133[1]	185[3]	211[15]	254[2]	140[2]
T7	287[5]	180[2]	62[5]	125[1]	177[5]	254[5]	249[2]	128[1]
C8	265[5]	189[4]	73[5]	130[2]	179[3]	271[2]	250[1]	136[2]
T9	295[10]	202[5]	42[6]	126[1]	168[5]	269[3]	256[1]	128[1]
G10				133[1]			257[1]	143[1]
C11	278[8]	198[5]	73[10]	115[1]	160[3]	277[5]	254[3]	115[2]
A12	343[30]	153[25]	7[14]	121[7]	198[8]	245[16]	258[3]	123[9]
G13	310[4]	151[2]	64[2]	125[1]	188[1]	248[2]	239[1]	127[2]
A14	302[4]	189[3]	40[6]	129[1]	172[8]	260[3]	257[1]	133[1]
T15	297[8]	200[9]	49[6]	145[1]	175[6]	260[4]	269[2]	159[1]
A16	301[4]	154[8]	65[2]	126[1]	194[6]	248[5]	244[1]	129[1]
A17	314[10]	190[7]	32[5]	143[1]	175[3]	259[7]	260[1]	159[1]
T18	321[4]	171[3]	45[4]	139[1]	175[2]	238[4]	258[1]	153[1]
C19	318[6]	203[2]	10[9]	145[2]	200[5]	274[3]	275[3]	161[3]
A20				137[2]			228[3]	142[2]
Section B: CDPI$_3$-Modified Duplex Structure								
T1	289[8]	103[1]	95[5]	123[3]	204[6]	208[7]	331[2]	126[5]
G2	287[10]	158[5]	77[10]	142[2]	196[6]	243[5]	244[1]	159[3]
A3	291[7]	189[3]	42[4]	125[2]	184[2]	281[2]	241[5]	128[3]
T4	313[5]	166[4]	43[4]	121[2]	190[7]	253[3]	244[1]	122[2]
T5	258[26]	204[4]	74[25]	117[1]	160[3]	273[6]	243[3]	119[1]
A6	305[12]	167[6]	62[7]	129[4]	173[3]	244[5]	258[4]	131[5]
T7	238[20]	195[5]	91[15]	131[5]	179[3]	268[5]	245[2]	138[8]
C8	306[5]	178[5]	48[4]	125[2]	187[6]	262[6]	243[1]	127[3]
T9	274[15]	234[17]	42[7]	129[3]	147[14]	277[8]	246[3]	137[5]
G10				146[1]			253[4]	162[1]
C11	248[8]	120[7]	99[7]	132[2]	254[4]	277[10]	248[2]	139[4]
A12	275[13]	179[5]	68[9]	126[2]	190[5]	261[5]	241[1]	130[3]
G13	284[6]	191[5]	66[3]	134[2]	174[3]	265[2]	232[2]	141[3]
A14	290[6]	178[5]	53[6]	137[3]	193[5]	255[4]	263[1]	149[5]
T15	313[4]	191[5]	29[2]	137[3]	175[2]	264[3]	244[2]	149[6]
A16	251[8]	192[4]	85[5]	135[1]	178[4]	260[2]	252[2]	143[2]
A17	288[12]	186[6]	60[8]	130[3]	170[7]	269[3]	240[2]	135[4]
T18	290[13]	179[8]	44[9]	127[3]	202[6]	264[4]	250[3]	131[4]
C19	168[12]	175[7]	182[9]	121[1]	167[2]	283[4]	216[2]	121[1]
A20				125[1]			230[1]	127[2]
A-DNA[c]	285	208	45	83	178	313	206	13
B-DNA[d]	314	214	36	156	155	264	262	191

[a]Backbone torsion angles are defined as P(*i*)-α-O5′-β-C5′-γ-C4′-δ-C3′-ε-O3′-ζ-P(*i* + 1). [b]Standard deviations for the ensemble of structures are given in brackets. [c]Canonical A-DNA (36). [d]Canonical B-DNA (37).

Despite these differences, the structure of the CDPI$_3$-conjugated DNA duplex that we report here has some similarities to the structures of other DNA complexes with minor groove binding agents. Like other crescent-shaped minor groove binders, the CDPI$_3$ moiety fits snugly into the minor groove of the DNA duplex (Figs 4 and 5) even though there are no hydrogen bonds nor adduct formation occurring in the minor groove to stabilize this binding mode. In the minor groove van der Waal's forces appear to provide the major factor in stabilizing the conjugated duplex. The compact binding mode that is observed produces a narrowing of the minor groove at the drug binding site and a widening of the groove distal to the binding site; as mentioned above, a similar observation was reported by Lin *et al*. (27).

Several notable variances were observed, however, between our results and those obtained in related studies. First, only one conformer of the CDPI$_3$-conjugated DNA duplex was evident from the NMR data (in Fig. 2 cross-peaks involving proton H15 of the CDPI$_3$ moiety may show evidence of slow exchange); this was in contrast to the two conformers that were observed for the CPI-CDPI$_2$–DNA adduct (29). Indeed, the sugar puckers in the CDPI$_3$-conjugated duplex were all C2′-*endo*, whereas the control duplex exhibited a variety of pseudorotation angles, suggesting that binding of the CDPI$_3$ moiety places additional constraints on mobility of the sugar–phosphate backbone. Second, neither the structure of the control non-conjugated duplex nor the structure of the conjugated DNA duplex is noticeably bent, indicating that binding of the CDPI$_3$ moiety does not induce bending in this DNA sequence. Most adducts of DNA and minor groove binders have been reported to result in bending of the duplex, although a lack of bending was also observed in the structure of a DNA–CPI-lexitropsin adduct determined by Lown and co-workers (32). Finally, neither the Hurley nor the Gorenstein groups used any drug–DNA intermolecular NOE cross-peaks in their structure refinement analyses, whereas 12 intermolecular cross-peaks were employed in the calculations used to generate the structure of the conjugated duplex shown in Figure 4.

It has been known for some time that a bulky amino group at position 2 in guanine in the minor groove, sterically interferes with binding of minor groove binding agents to DNA sequences containing G·C base pairs (33–34). The sequence employed in

our study is definitely A/T-rich, yet it does have a G·C base pair in the $CDPI_3$ binding site. While we observe some close van der Waal's contacts between the amino group of the G2 residue and atoms of the $CDPI_3$ moiety, this lone G·C base pair certainly does not preclude $CDPI_3$ binding. It is plausible that the attachment of the drug to the DNA by a flexible hexamethylene linker can explain this and some of the other features of the $CDPI_3$-conjugated duplex. The flexibility of this linker permits the $CDPI_3$ moiety to adjust its position to fit snugly in the minor groove despite the presence of the 2-amino group on G2. Formation of DNA adducts by CC-1065 and CPI-$CDPI_2$ upon binding in the minor groove may be accompanied by conformational stress on the DNA backbone that may promote bending of the helix. The lack of bending in $CDPI_3$-conjugated DNA and its very high melting temperature may be at least partially due to the absence of conformational stress that is permitted by attaching the $CDPI_3$ moiety to the DNA strand by a flexible linker (35).

REFERENCES

1 Gamper,H.B., Cimino,G.D. and Hearst,J.E. (1987) *J. Mol. Biol.*, **197**, 349–362.
2 Wyatt,J.R., Puglisi,J.D. and Tinoco,I.,Jr (1989) *BioEssays*, **11**, 100–106.
3 Latham,J.A. and Cech,T.R. (1989) *Science*, **245**, 276–282.
4 Asseline,U., Delarue,M., Lancelot,G., Toulme,F., Thuong,N.T., Montenay-Garestier,T. and Helene,C. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 3297–3301.
5 Nielson,P.E., Egholm,M. and Buchardt,O. (1994) *Bioconjugate Chem.*, **5**, 3–7.
6 Wagner,R.W., Matteucci,M.D., Lewis,J.G., Gutierrez,A.J., Moulds,C. and Froehler,B.C. (1993) *Science*, 260, 1510–1513.
7 Gryaznov,S. and Chen,J.-K. (1994) *J. Am. Chem. Soc.*, **116**, 3143–3144.
8 Reynolds,V.L., McGovern,J.P. and Hurley,L.H.J. (1986) *Antibiotics*, **39**, 319–334.
9 Boger,D.L. and Johnson,D.S. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 3642–3649.
10 Lukhtanov,E.A., Kutyavin,I.V., Gamper,H.B. and Meyer,R.B. (1995) *Bioconjugate Chem.*, **6**, 418–426.
11 Kutyavin,I.V., Lukhtanov,E.A., Gamper,H.B. and Meyer,R.B. (1997) *Nucleic Acids Res.*, **25**, 3718–3723.
12 Afonina,I., Kutyavin,I., Lukhtanov,E., Meyer,R.B. and Gamper,H. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 3199–3204.
13 Afonina,I., Zivarts,M., Kutyavin,I., Lukhtanov,E., Gamper,H. and Meyer,R.B. (1997) *Nucleic Acids Res.*, **25**, 2657–2660.
14 Goljer,I., Kumar,S. and Bolton,P.H. (1995) *J. Biol. Chem.*, **270**, 22980–22987.
15 Callihan,D., West,J., Kumar,S., Schweitzer,B.I. and Logan,T.M. (1996) *J. Magn. Resonance B*, **112**, 82–85.
16 Foti,M., Marshalko,S., Schurter,E., Kumar,S., Beardsley,G.P. and Schweitzer,B.I. (1997) *Biochemistry*, **36**, 5336–5345.
17 Gronenborn,A.M. and Clore,G.M. (1989) *Biochemistry*, **28**, 5978–5984.
18 Brünger,A.T. (1993) *X-PLOR 3.1: A System for Crystallography and NMR*. Yale University, New Haven, CT.
19 Schweitzer,B.I., Mikita,T., Kellogg,G.W., Gardner,K.H. and Beardsley,G.P. (1994) *Biochemistry*, **33**, 11460–11475.
20 Ravishanker,R., Swaminathan,S., Beveridge,R.L. and Sklenar,H. (1989) *J. Biomol. Struct. Dyn.*, **6**, 669–699.
21 Kim,S.-G., Lin,L.-J. and Reid,B.R. (1992) *Biochemistry*, **31**, 3564–3574.
22 Lown,J.W. (1995) *Drug Dev. Res.*, **34**, 145–160.
23 Zimmer,C. and Wahnert,U. (1986) *Prog. Biophys. Mol. Biol.*, **47**, 31–112.
24 Singh,M.P., Kumar,S., Joseph,T., Pon,R.T. and Lown,J.W. (1992) *Biochemistry*, **31**, 6453–6461.
25 Kumar,S., Bathini,Y., Joseph,T., Pon,R.T. and Lown,J.W. (1991) *J. Biomol. Struct. Dyn.*, **9**, 1–21.
26 Lee,C.-S., Sun,D., Kizu,R. and Hurley,L.H. (1991) *Chem. Res. Toxicol.*, **4**, 203–213.
27 Lin,C.H., Sun,D. and Hurley,L.H. (1991) *Chem. Res. Toxicol.*, **4**, 21–26.
28 Lin,C.H., Hill,C.G. and Hurley,L.H. (1992) *Chem. Res. Toxicol.*, **5**, 167–182.
29 Powers,R. and Gorenstein,D.G. (1990) *Biochemistry*, **29**, 9994–10008.
30 DiGabriele,A.D., Sanderson,M.R. and Steitz,T.A. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 1816–1820.
31 Young,M.A., Goljer,I., Kumar,S., Srinivasan,J., Beveridge,D.L. and Bolton,P.H. (1995) *Methods Enzymol.*, **261**, 121–144.
32 Fregeau,N.L., Wang,Y., Pon,R.T., Wylie,W.A. and Lown,J.W. (1995) *J. Am. Chem. Soc.*, **117**, 8917–8925.
33 Wahnert,U., Zimmer,C., Luck,G. and Pitra,C. (1975) *Nucleic Acids Res.*, **2**, 391–404.
34 Marck,C., Kakiuchi,N. and Guschlbauer,W. (1982) *Nucleic Acids Res.*, **10**, 6147–6161.
35 Singh,M.P. and Lown,J.W. (1996) *Prog. Med. Chem.*, **1**, 49–172.
36 Arnott,C. and Hukins,D.W. (1972) *Biophys. Biochem. Res. Commun.*, **47**, 1504–1509.
37 Arnott,C. and Hukins,D.W. (1973) *J. Mol. Biol.*, **81**, 93–105,