Structural Consequences of a Carcinogenic Alkylation Lesion on DNA: Effect of O^6-Ethylguanine on the Molecular Structure of the d(CGC[e^{6}G]AATTCGCG)–Netropsin Complex†,‡

M. Sriram,§ Gijs A. van der Marel,‖ Harlof L. P. F. Roelen,‖ Jacques H. van Boom,‖ and Andrew H.-J. Wang*,§

Biophysics Division and Department of Cell and Structural Biology, University of Illinois at Urbana–Champaign, Urbana, Illinois 61801, and Gorlaeus Laboratories, Leiden State University, 2300RA Leiden, The Netherlands

Received May 14, 1992; Revised Manuscript Received August 5, 1992

ABSTRACT: Exposure of cells to alkylating agents produces DNA lesions, most of which are repaired. However some alkyl lesions persist and play a role in inducing point mutations and the subsequent carcinogenic conversion. O^6-Ethylguanine (e^{6}G) is a relatively persistent alkylation lesion caused by the exposure of DNA to *N*-ethyl-*N*-nitrosourea. We study the consequence of the e^{6}G incorporation in DNA by X-ray crystallography. We have obtained crystals of the modified DNA dodecamer d(CGC[e^{6}G]AATTCGCG) and the unmodified d(CGCGAATTCGCG), complexed to the minor groove binding drug netropsin. The space group of both crystals is $P2_12_12_1$, isomorphous to other related dodecamer DNA crystals. The structures have been solved by the molecular replacement method and refined by the constrained least-squares procedure to *R*-factors of ~16% at resolution of ~2.5 Å. The two independent e^{6}G-C base pairs in the DNA duplex adopt different base-pairing schemes. The e^{6}G4-C21 base pair has a configuration similar to a normal Watson–Crick base pair, except with one three-centered hydrogen bond pair and one direct hydrogen bond between e^{6}G4 and C21. In contrast, the e^{6}G16-C9 base pair adopts a wobble configuration. The ethyl group is in the *proximal* orientation (to N^7) in both base pairs. These observations enrich and support those found in the crystal structure of d(CGC[e^{6}G]AATTCGCG), complexed to minor groove binding drugs Hoechst 33258 and Hoechst 33342 [Sriram et al. (1992) *EMBO J. 11*, 225–232]. We suggest that a dynamic equilibrium between these two configurations for the e^{6}G-C base pair is likely and would present an ambiguous signal to the cellular transcription, replication, or repair mechanisms. In contrast, thymine can pair with e^{6}G in only one way, albeit imperfect, mimicking a Watson–Crick base pair. This may be a plausible explanation of why thymine is found preferentially incorporated across the e^{6}G during replication. In addition, we analyze the influence of the alkylation lesion on DNA and the molecular details of netropsin–DNA interaction. In the present two new netropsin complexes, the netropsin spans across five base pairs (starting halfway between C3-G22 and e^{6}G4-C21 base pairs and ending at T8-A17 base pair) in the narrow minor groove. This is in contrast to the earlier crystal structure of netropsin complexed with another DNA dodecamer having the same AATT central core sequence, d(CGCGAATT[br^{5}C]GCG) [Kopka et al. (1985) *J. Mol. Biol. 272*, 390–395]. In the latter structure, the netropsin lies between G4-br^{5}C21 and br^{5}C9-G16 base pairs. Our structural analysis suggests that netropsin can occupy the minor groove in two orientations in the crystal, as in the case of the netropsin–d(CGCGATATCGCG) complex [Coll et al. (1990) *Biochemistry 28*, 1022–1029]. Only a fraction of the amide nitrogens of netropsin form three centered hydrogen bonds with acceptor atoms of DNA in all structures.

DNA is the hereditary material, and modifications of its structure through covalent and noncovalent interactions have significant functional consequences. Interaction of DNA with various ligands such as proteins and drugs has been the focus of study by several techniques. In addition, the structural polymorphism of DNA and the environmental variables which affect these conformational transitions have been intensively explored. An important class of DNA modifications with potential of causing subsequent carcinogenic conversion is the alkylation lesions of DNA. These lesions are caused by alkylating agents, some of which are chemical carcinogens, e.g., *N*-methyl-*N*-nitrosourea (MNU) and *N*-ethyl-*N*-nitrosourea (EtNU). The structural and functional consequences of these alkylating reactions have been studied (Singer et al., 1978; Thomale et al., 1990).

The nucleophilic sites on DNA bases (e.g., N^7 and N^3 of purine) are the targets for attack by the methyl or ethyl cation generated via a nonenzymatic heterolytic reaction by MNU and EtNU (Singer et al., 1978). Alkylation on those sites is potentially deleterious to the function of DNA. Cells have developed efficient repair systems to remove such lesions. Nevertheless, some lesions persist, either because of their abundance or the difficulty in the recognition/removal process. The O^6-ethylguanine (e^{6}G)[1] lesion is relatively persistent and is repaired by the suicide enzyme O^6-alkylguanine alkyltransferase (AGT, EC 2.1.1.63) which removes the alkyl group and regenerates an intact G (Lindahl et al., 1988; Thomale et al., 1990). O^6-alkylguanine is mutagenic in vivo and induces a G → A transition (Loechler et al., 1984; Ellison et al., 1989), and the relative capacity of cells for its repair is a critical determinant for their risk of malignant conversion by *N*-nitroso carcinogens (Jurgen et al., 1990).

† This work was supported by NIH GM-41612 and CA-52506 (A.H.-J.W.). M.S. was supported in part by a grant from DCCA of the state of Illinois to the Beckman Institute at University of Illinois at Urbana–Champaign.

‡ Atomic coordinates have been deposited at Brookhaven Protein Databank.

* Address correspondence to this author.

§ University of Illinois at Urbana–Champaign.

‖ Leiden State University.

0006-2960/92/0431-11823$03.00/0

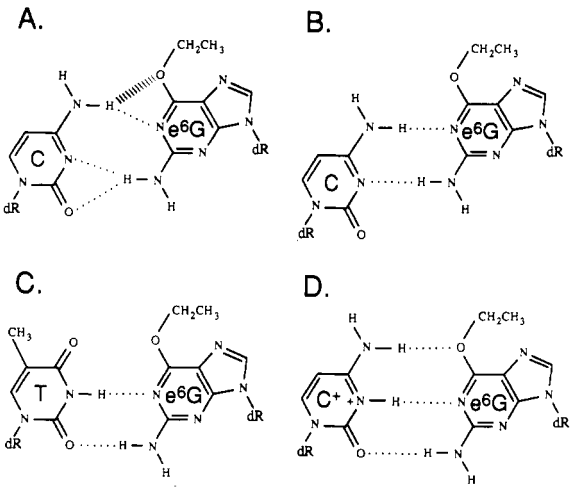

FIGURE 1: Possible hydrogen bonding configurations of e^{6}G-C and e^{6}G-T base pairs. The dotted line indicates hydrogen bonds. (A) Bifurcated e^{6}G-C base pair with one or two three-centered hydrogen bonds. The thick striped line indicates that this hydrogen bond may be missing in some cases. (B) Wobble e^{6}G-C base pair with two hydrogen bonds. (C) e^{6}G-T base pair with a shape similar to a Watson–Crick base pair. (D) Watson–Crick e^{6}G-C$^+$ base pair. The e^{6}G base pairs with a protonated cytosine.

When the O^6 position of guanine is alkylated, the tautomeric form of guanine changes such that N^1 no longer has a proton and cannot be a hydrogen bond donor. Furthermore, the *O*^{6}-alkyl group may adopt two possible orientations, proximal or distal to N^7. The distal configuration is energetically more favorable as predicted by theoretical calculations (Pedersen et al., 1990). In fact, the crystal structure of the free nucleoside *O*^{6}-methylguanosine showed unequivocally that the methyl lesion adopted the distal configuration (Parthasarathy & Fridey, 1986). However, a m^{6}G/e^{6}G in a DNA double helix participates in base-pairing interactions and a distal orientation of the methyl/ethyl group is expected to hinder the hydrogen bond formation with the opposing base, causing uncommon base-pairing schemes which would destabilize the helix. Figure 1A–D shows some of the possibilities for e^{6}G-C and e^{6}G-T base pairs.

This critical issue about the type of base-pairing schemes that m^{6}G/e^{6}G may form with other bases in DNA has been studied recently. In the Z-DNA crystal structure of d(CGC-[m^{6}G]CG), the type in Figure 1D, which requires a protonated C, was found (Ginell et al., 1990). In the B-DNA structure of d(CGC[m^{6}G]AATTTGCG), the type in Figure 1C was found (Leonard et al., 1990). However, until recently there has been no definitive conclusion about which [m^{6}G/e^{6}G]-C base pairing scheme exists in B-DNA under physiological conditions. On the basis of the NMR studies of the DNA dodecamer d(CGC[e^{6}G]AGCTCGCG) in solution, Kalnik et al. (1989) tentatively proposed that the wobble type of Figure 1B is adopted in the helix. To answer this question more definitively, we have determined the structures of the complexes of Hoechst 33258 (H258) and Hoechst 33342 (H342) with both d(CGCGAATTCGCG) and d(CGC[e^{6}G]AATTCGCG) (Sriram et al., 1992).

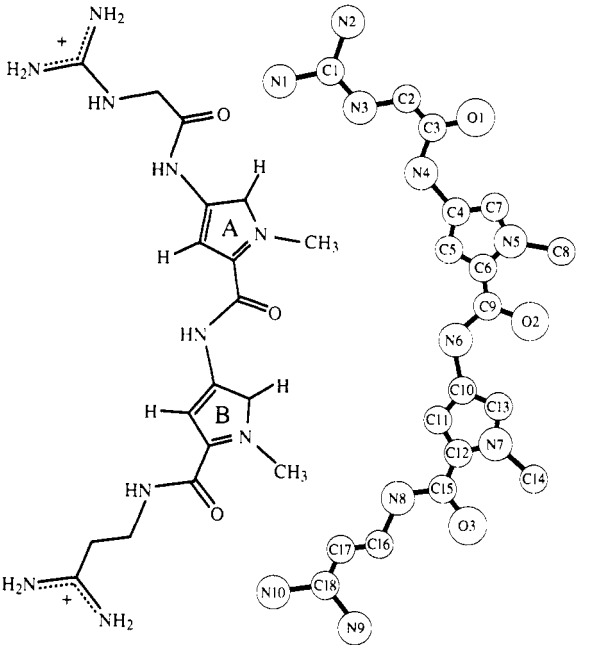

FIGURE 2: Molecular formula and atom numbering scheme of netropsin, an antitumor/antiviral antibiotic isolated from *Streptomyces netropsis*. The drug has a binding preference for AT-rich sequences of DNA (Zimmer & Wahnert, 1986). The DNA-binding mechanisms of netropsin and related drugs have been extensively studied by chemical footprinting experiments (Taylor et al., 1984; Dervan, 1987), NMR (Patel et al., 1981, 1982, 1986a,b; Klevit et al., 1986; Leupin et al., 1986), and single crystal X-ray diffraction (Kopka et al., 1985; Coll et al., 1987, 1989). Netropsin has a pseudo-2-fold rotation symmetry axis which lies near the middle amide bond.

In this paper, we expand our study on the influence of ethylation at the O^6 position of guanine on the B-DNA conformation as well as the netropsin–DNA interaction by solving two new crystal structures of complexes of d(CGC-[e^{6}G]AATTCGCG) and d(CGCGAATTCGCG) with netropsin (Figure 2). The comparison of netropsin binding in the present structures with the already-published d(CGC-GAATT[br^{5}C]GCG)–netropsin (br^{5}C–DNET) and d(CGC-GATATCGCG)–netropsin (AT–DNET) complexes indicates that they have different netropsin binding modes to DNA. The available structural information taken together allows us to propose possible functional consequences due to the alkylation of DNA. In all three e^{6}G-modified DNAs, the two independent e^{6}G-C base pairs adopt different base-pairing schemes in the B-DNA double helix. The *O*^{6}-ethyl group plays an important role of influencing the conformation of the e^{6}G-C base pair and its neighbors. The corresponding unmodified DNA and netropsin complex was also solved to serve as an experimental control.

MATERIALS AND METHODS

The procedure of Roelen et al. (1992) was followed for the synthesis of *O*^{6}-ethyldeoxyguanosine. The nucleoside was then converted into the phosphoramidite precursor and incorporated into the oligonucleotides on a Pharmacia DNA synthesizer. The synthesized DNA fragment was purified by Sephadex G-50 column chromatography. HPLC was used to check the

[1] Abbreviations: A, T, G, C, adenine, thymine, guanine, cytosine or their respective nucleotide in DNA; AT–DNET, d(CGCGATATCGCG)–netropsin complex; br^{5}C–DNET, d(CGCGAATT[br^{5}C]GCG)–netropsin complex; DH258, d(CGCGAATTCGCG)–Hoechst 33258 complex; DH342, d(CGCGAATTCGCG)–Hoechst 33342 complex; DNET, d(CGCGAATTCGCG)–netropsin complex; e^{6}G, *O*^{6}-ethylguanine or *O*^{6}-ethyldeoxyguanosine; e^{6}G–DH258, d(CGC[e^{6}G]AATTCGCG)–Hoechst 33258 complexes; e^{6}G–DH342, d(CGC[e^{6}G]AATTCGCG)–Hoechst 33342 complex; e^{6}G–DNET, d(CGC[e^{6}G]AATTCGCG)–netropsin complex; ED, electron density; H258 Hoechst 33258; H342, Hoechst 33342; m^{6}G, *O*^{6}-methylguanine or *O*^{6}-methyldeoxyguanosine; NMR, nuclear magnetic resonance; pr tw, propeller twist.

purity of the final product which was judged to be greater than 95%. Netropsin was kindly provided by the National Cancer Institute.

The DNA duplexes containing O^6-alkylguanine are notably difficult to crystallize. This is in part due to the destabilization of the DNA duplex caused by the O^6-alkylguanine base pairing. The sequence d(CGC[e^{6}G]AATTCGCG) was selected as we believed that the molecule could be crystallized in the lattice of the native d(CGCGAATTCGCG) crystal (Drew & Dickerson, 1981). Our unsuccessful attempts to crystallize d(CGC[e^{6}G]AATTCGCG) alone prompted us to try the crystallization of DNA–ligand complexes, inspired by our earlier studies (Coll et al., 1989; Aymami et al., 1990). Crystallization experiments using the procedure of Wang and Gao (1990) were carried out. We were able to obtain crystals of complexes of d(CGC[e^{6}G]AATTCGCG) with Hoechst 33258 (H258), Hoechst 33342 (H342), netropsin, and DAPI. Crystals of the first three complexes were useful, whereas the crystal of the DAPI complex was not. The crystallization solution in general contained 0.8 mM dodecamer (single-strand concentration), 31 mM cacodylate buffer at pH 6.0, 4 mM $MgCl_2$, 1 mM spermine, 0.8 mM drug, and 2% 2-methyl-2,4-pentanediol (2-MPD), and it was equilibrated against 50% 2-MPD by vapor diffusion at room temperature. Large crystals with somewhat irregular shape appeared after 4–8 weeks.

For comparison, we also crystallized the complexes of netropsin and Hoechst 33342 with unmodified d(CGC-GAATTCGCG) and determined their structures. The crystals of netropsin complexed with d(CGCGAATTCGCG) could be obtained in the presence of Ba^{2+}, Ca^{2+}, or Mg^{2+} ions. Crystals of d(CGCGAATTCGCG)–netropsin (DNET) and d(CGC[e^{6}G]AATTCGCG)–netropsin (e^{6}G–DNET) complexes grown in the presence of magnesium icn gave better crystals. Crystal of each complex was mounted in a thin-walled glass capillary and sealed with a droplet of the crystallization mother liquor for data collection. All of them are in the isomorphous orthorhombic space group $P2_12_12_1$ and have unit cell dimensions of $a \sim 26$ Å, $b \sim 42$ Å, and $c \sim 64$ Å. The diffraction data set was collected at room temperature on a Rigaku AFC-5R rotating-anode diffractometer, equipped with a copper anode and a graphite monochromator, at a power of 50 kV and 180 mA. An ω-scan mode was used for data collection with CuK_α radiation (1.5406 Å), and the data set was collected to a resolution of 2.0 Å. Lorentz-polarization, absorption, and decay corrections were applied to obtain the structure factor amplitudes.

It was evident from the unit cell dimensions and the diffraction pattern that the crystals were closely related to those of other dodecamer–drug complexes (Wang & Teng, 1990). Molecular replacement was used to solve the structure. A B-DNA dodecamer without drug or solvent molecules, from the d(CGCGAATTCGCG)–Hoechst 33258 (DH258) crystal structure, was used as the starting model for refinement. The entire dodecamer duplex is in the asymmetric unit. Therefore, the two strands of the duplex are not identical in their conformation. The model was refined using the Konnert–Hendrickson constrained refinement procedure (Hendrickson & Konnert, 1979; Westhof et al., 1985).

The overall position of the molecule was refined by several cycles of least-squares refinement using low-resolution data (3.5 Å) and strong stereochemical constraints (target distance RMSD < 0.005 Å). The position of the molecule and its component groups were checked on a Fourier $(2|F_o| - |F_c|)$ sum map, and then higher resolution reflections were gradually included in the refinement. After many cycles of refinement with all available data, the R-factor reached ~30% at a resolution better than ~2.5 Å. A difference Fourier $(|F_o| - |F_c|)$ map was calculated, and a clear continuous residual density could be seen in the minor groove of the DNA near the central AATT site. As the position and size of the density were unambiguous, netropsin molecule was located from this map using the program FRODO/TOM (Jones, 1978) and included in the refinement. Well-ordered solvent molecules were located from the Fourier $(2|F_o| - |F_c|)$ maps, excluding those in the minor groove or near the e^{6}G4 or e^{6}G16 residues, and gradually added in the subsequent refinement cycles. Only solvent molecules well located in density with reasonable geometry and thermal factors were retained. When the R-factor was about ~18%, a difference Fourier $(|F_o| - |F_c|)$ map was used to locate the position of the ethyl groups, for the e^{6}G-containing complexes. From this stage, until the completion of the refinement, difference Fourier $(|F_o| - |F_c|)$ maps were calculated at the end of each refinement cycle to monitor and fine-tune the orientation of the ethyl groups and the components groups of the drug. No hydrogen bond distance constraints were imposed on the two e^{6}G-C base pairs during the refinement in an attempt to get an unbiased base-pairing scheme.

The difference Fourier maps of the e^{6}G–DNET and DNET structures were calculated by removing netropsin from the phase contribution and the drug molecule can be seen to fit nicely in the residual caterpillar-shaped electron density envelope in both structures (Figure 3). The density was sufficiently well-resolved to define the position of the drug accurately, but the polarity of the drug molecule in the duplex remains ambiguous. The netropsin (Figure 2) structure is nearly 2-fold symmetrical. At the nominal resolution of ~2.5 Å it is not possible to distinguish between a nitrogen and a carbon, and we believed that the question about the orientation of the drug cannot be unequivocally addressed. The netropsin molecule could be refined equally well in both orientations in the d(CGCGATATCGCG)–netropsin (AT–DNET) structure (Coll et al., 1989). To confirm this, we refined netropsin in the opposite orientation also for the DNET structure. It could be refined to the same R-factor of 16.4% and had a similar temperature factor for the drug as well as other atoms. However, in keeping with the presentation of the first netropsin structure br^{5}C–DNET, we show the same orientation for netropsin in all structures. Occupancy refinement of the netropsin coexisting in two orientations simultaneously would lead to an unjustifiably large number of parameters for the observed raw data. For most purposes, the netropsin molecule is sufficiently symmetric to leave most of the DNA–drug interactions similar in either orientation if the atom numbering scheme is disregarded. Therefore, we choose not to address the question of orientation of netropsin when bound to the DNA dodecamer in the crystal lattice. No magnesium or spermine ions could be explicitly identified. The relevant crystal data along with the final refined parameters are listed in Table I. The summarized breakdown of the number of reflections observed by resolution shell and R-factors for each are listed in Table 1S (supplementary material). The final atomic coordinates of these two structures have been deposited at Brookhaven Protein Databank.

RESULTS AND DISCUSSION

Overall Structure of the Netropsin Complexes. The stereoscopic van der Waals and skeletal drawings of the e^{6}G–

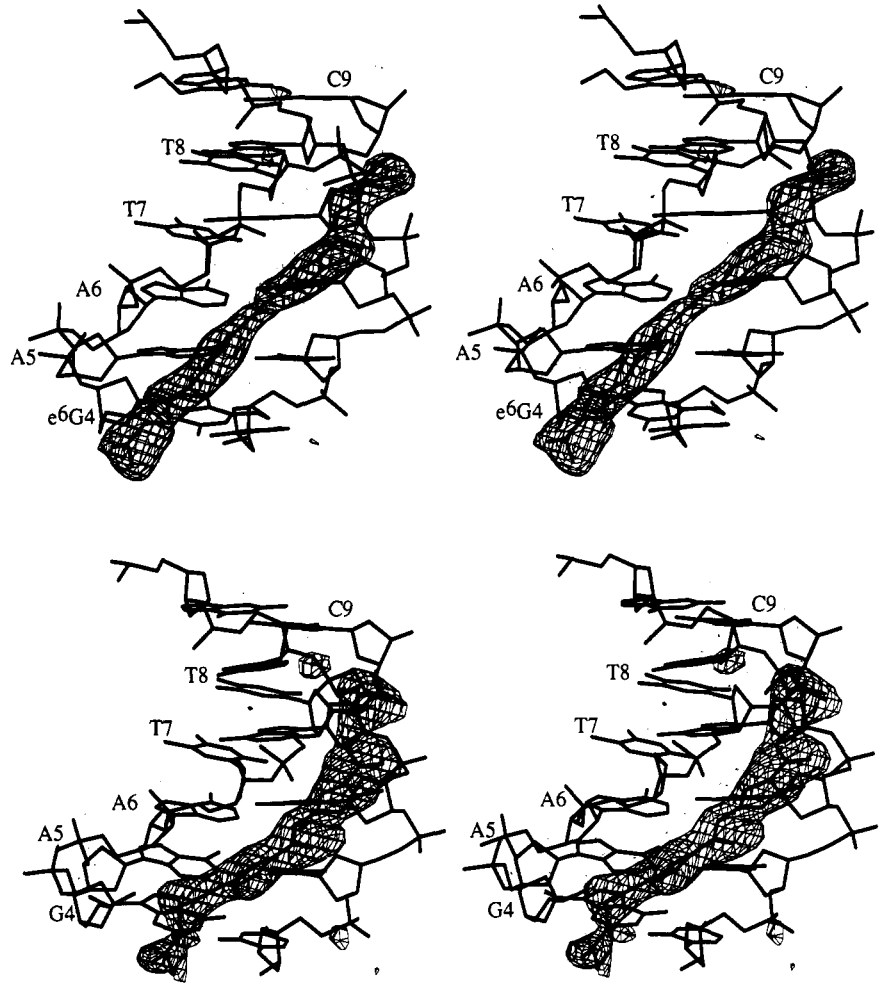

FIGURE 3: Stereoscopic ($|F_o| - |F_c|$) difference Fourier maps displaying in detail the central six base pairs GAATTC of the two new complexes where the drug is removed from the phase contribution. (Top) The d(CGC[e^{6}G]AATTCGCG)–netropsin complex. (Bottom) The d(CGCGAATTCGCG)–netropsin complex. The density is well-resolved for the position of the drug, but the polarity of the drug is not evident. In both structures, the netropsin-binding site lies shifted from the central AATT location.

Table I: Relevant Crystal Data and Final Refinement Parameters of Four Dodecamer–Netropsin Complexes

complex[a]	space group	no. of reflections $F_o > 2\sigma(F_o)$	resolution (Å)	*R*-factor	no. of waters located	RMS bond deviation (Å)
e^{6}G–DNET	$P2_12_12_1$	1514	~2.5	15.6%	89	0.019
DNET	$P2_12_12_1$	2056	~2.2	16.4%	74	0.020
AT–DNET	$P2_12_12_1$	1848	~2.4	20.0%	60	0.024
br^{5}C–DNET	$P2_12_12_1$	2528	~2.2	21.1%	75	na

[a] Complexes: e^{6}G–DNET, CGC[e^{6}G]AATTCGCG + netropsin complex (this work); AT–DNET, CGCGATATCGCG + netropsin complex (Coll et al., 1989); br^{5}C–DNET, CGCGAATT[br^{5}C]GCG + netropsin complex (Kopka et al., 1985); DNET, CGCGAATTCGCG + netropsin complex (this work).

DNET and DNET complexes (Figures 4 and 5) illustrate the overall structure of the two complexes. The structures are similar to those of the other related dodecamer–drug complexes (Kopka et al., 1985; Coll et al., 1987, 1989; Carrondo et al., 1989; Larsen et al., 1989; Teng et al., 1988; Sriram et al., 1992). They all have a characteristically narrow minor groove at the AATT region which acts as the ligand-binding site (Figure 6A). Several other gross features, such as the ~19° bend of the helix axis of the duplex, are conserved. However, the changes in the DNA conformation in the e^{6}G–DNET complex relative to that of the other dodecamer–netropsin complexes are numerous and distributed throughout the helix. An estimate of the overall variation among the four DNA–drug complexes (e^{6}G–DNET, AT–DNET, br^{5}C–DNET, and DNET) is shown by the least-squares fit of the common atoms of the four DNA duplexes and the drug molecules (Table II). The root mean square deviations of the DNA duplexes range from 0.92 Å to 1.38 Å, while those of netropsin range from 0.75 Å to 1.11 Å. Tables 2Sa and 2Sb (supplementary material) list the backbone torsion angles of the e^{6}G–DNET and DNET complex.

The variations in the dodecamer structures are reflected in the global features (e.g., P–P distance across the minor groove) (Figure 6A) and local features (e.g., base-pair buckle and propellar twist angles) (Figure 6B). For comparison, six e^{6}G-related complexes {e^{6}G–DNET, DNET, d(CGC[e^{6}G]AATTCGCG)–Hoechst 33258 [e^{6}G–DH258], d(CGCGAATTCGCG)–Hoechst 33258 [DH258], d(CGC[e^{6}G]AATTCGCG)–Hoechst 33342 [e^{6}G–DH342], d(CGCGAATTCGCG)–Hoechst 33342 [DH342]} are used. From Figure 6B it can be seen that nearly every base pair in the six e^{6}G-related duplexes has either the buckle or the propeller twist

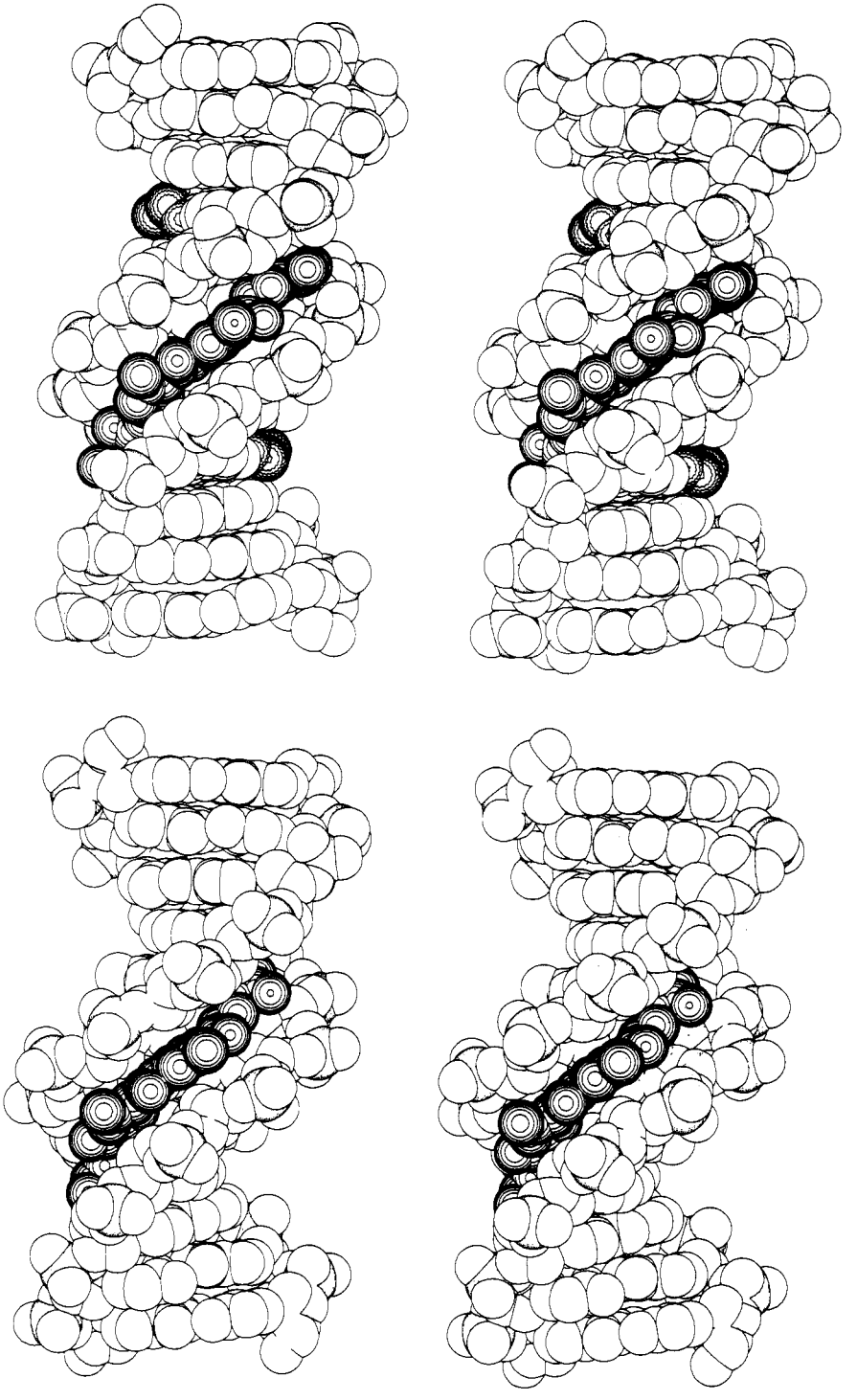

FIGURE 4: Stereoscopic van der Waal drawing of the structure of the d(CGC[e^{6}G]AATTCGCG)–netropsin complex (top) and d(CGCGAATTCGCG)–netropsin complex (bottom). The netropsin molecule drawn in a darker shade binds in the minor groove of the B-DNA duplex. The two ethyl groups and the O^6 atoms of e^{6}G's are drawn as spiked balls for emphasis.

angle greater than 10°, with the exception of C3-G22 and G12-C13 base pairs. The two G-C base pairs at both ends of the helix are involved in the interlocking lattice interactions using the G14-G24# and G12-G2# (the symbol # indicates a symmetry-related duplex) hydrogen bonding pairing in the minor groove. As noted before (Coll et al., 1990), this type of G-G pairing is associated with a high dehedral angle between the two guanines. This may impose conformational distortion in the participating (terminal and penultimate) base pairs. The terminal C1-G24 base pair has a high buckle (av −12°), whereas the penultimate base pairs have a high pr tw angle ω with G2-C23 averaging −12° and C11-G14 averaging −11°.

The base pairs in the central AT region have high propeller twist angles in most of the dodecamer structures. Three-centered hydrogen bonds from the N^6 amino group of an adenine simultaneously to the O^4 atoms of two thymine in the opposite strand are formed (one from the Watson–Crick mate and the other from its adjacent 5′T) (Coll et al., 1987; Nelson et al., 1987). This could be a possible reason for the unusual property (e.g., DNA bend) associated with the A_nT_n sequence (Crothers et al., 1990). In the six e^{6}G-related structures, only the A6-T19 maintains a very high pr tw (av −22°). The average distance in these two structures between N^6 of A5 and O^4 of T20 or O^4 of T19 is 2.94 Å and 3.13 Å, respectively,

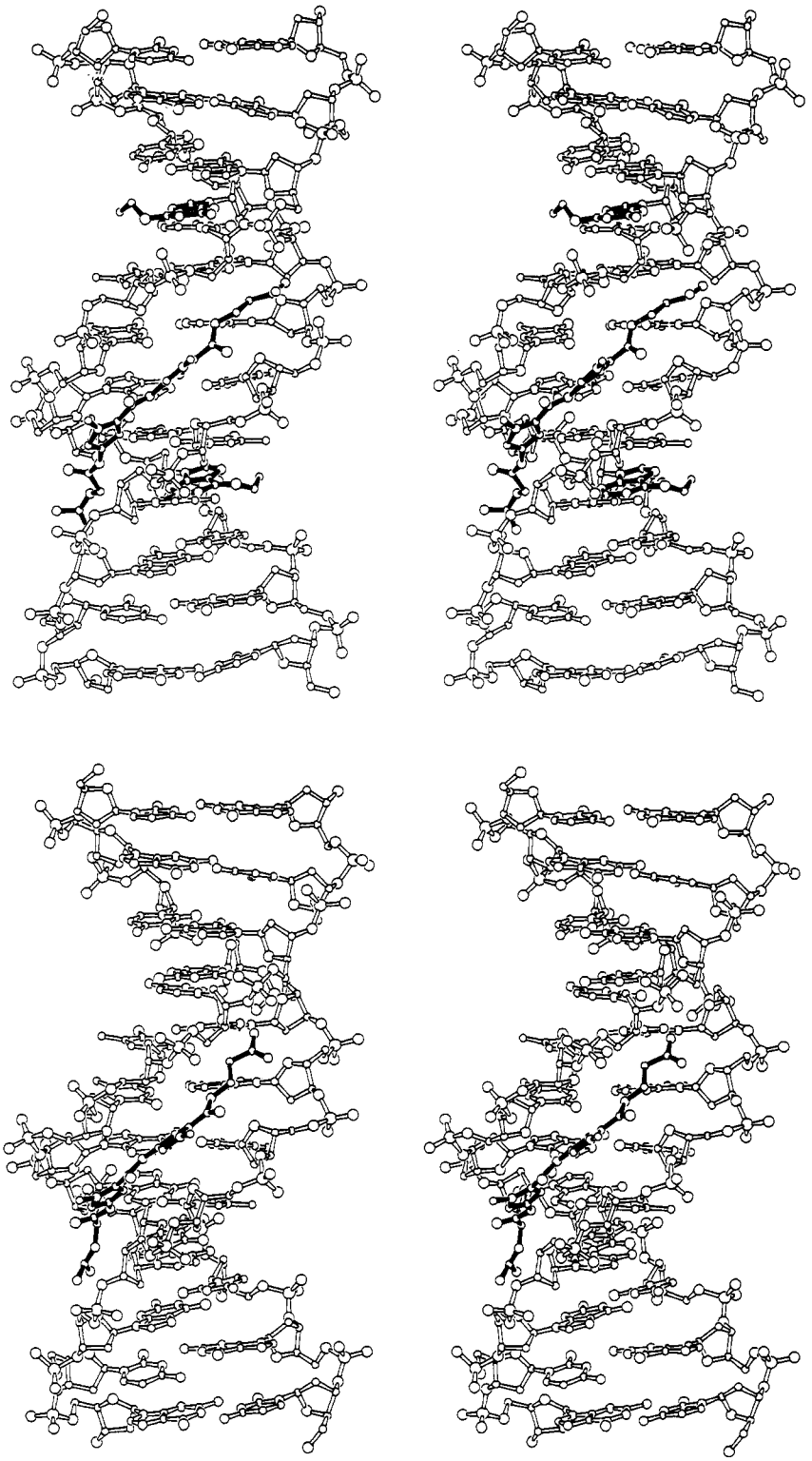

FIGURE 5: Stereoscopic skeletal drawing of the structure of the d(CGC[e^{6}G]AATTCGCG)–netropsin complex (top) and d(CGCGAATTCGCG)–netropsin complex (bottom). The netropsin molecule drawn with filled bonds is in the minor groove of the B-DNA duplex. The two e^{6}G's are drawn with filled bonds, and their ethyl groups are located in the major groove.

satisfying the condition of the interbase three-centered hydrogen bond. Interestingly, the pr tw of T7-A18 decreases in the e^{6}G complexes relative to the regular complexes, from av −15° (in DH258 and DH342) to −3° in e^{6}G–DH258 and 3° in e^{6}G–DH342 and from −19° in e^{6}G–DNET to −15° in DNET. The increase in the buckle of adjacent A6-T19 from av −2° in the normal to av −14° in the e^{6}G complexes may be related and compensatory.

e^{6}G-C Base Pairs. The ($|F_o| - |F_c|$) difference Fourier ED of the two independent e^{6}G-C base pairs in the dodecamer duplex of the e^{6}G–DNET complex is displayed in Figure 7. The entire base (e^{6}G) in one view and the ethyl groups in the other view were not included in the phase contribution. Figure 7 clearly shows the structure of the two e^{6}G-C base pairs. The e^{6}G–DH258 and e^{6}G–DH342 complexes have similar results (Sriram et al., 1991, 1992). The ED maps were examined very carefully, especially in and around the base pairs, to obtain a reliable interpretation of the configuration of these modified base pairs and the ethyl groups. There are very clear ED for these ethyl groups and the bases. The orientation

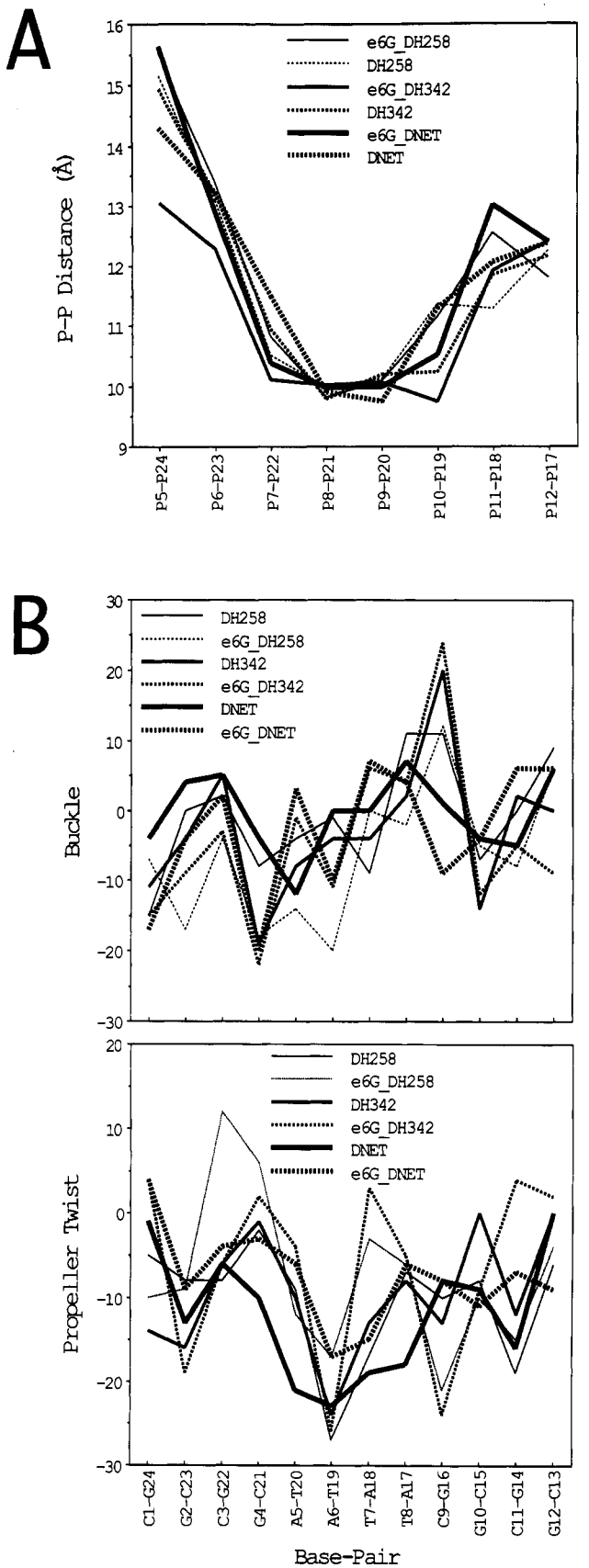

FIGURE 6: Plot of phosphate to phosphate distance across the minor groove in the B-DNA helices of six e^{6}G-related dodecamer–drug complexes. It demonstrates the narrowing of the minor groove in the P^7/P^{22} to the P^{10}/P^{22} region. (B) A plot of buckle (Dickerson et al., 1989) and propeller twist (Dickerson et al., 1989) of base pairs in six e^{6}G-related dodecamer–drug complexes.

of the ethyl groups is different in each of the e^{6}G-containing complexes for a given e^{6}G-C base pair, indicating that the crystal packing forces are not dictating the conformation of the ethyl group. A detailed structure of the two e^{6}G-C with adjacent base pairs for e^{6}G–DNET is shown in Figure 8, and those of e^{6}G-DH258 and e^{6}G-DH342 are included in Figures 1S and 2S (supplementary material). It is interesting to note that the two independent e^{6}G-C base pairs adopt different base-pairing schemes. The overlay of the two independent e^{6}G-C base pairs from e^{6}G-DH258, e^{6}G–DH342, and e^{6}G-DNET is shown in Figure 9A. The base-pairing scheme observed in all the e^{6}G structures falls within the schemes shown in Figure 1A,B for a given e^{6}G-C base pair.

Table II: Least-Squares Fit of DNA (Above Diagonal) and Drug Molecules (Below Diagonal) in Four DNA–Drug Complexes with Root Mean Square Deviations (RMSD) in Angstroms[a]

	e^{6}G–DNET	AT–DNET	br^{5}C–DNET	DNET
e^{6}G–DNET	x	1.15	1.31	1.01
AT–DNET	1.108	x	1.38	0.92
br^{5}C–DNET	0.966	0.894	x	1.06
DNET	0.909	0.788	0.750	x

[a] e^{6}G–DNET, CGC[e^{6}G]AATTCGCG + netropsin complex (this work); AT–DNET, CGCGATATCGCG + netropsin complex (Coll et al., 1989); br^{5}C–DNET, CGCGAATT[br^{5}C]GCG + netropsin complex (Kopka et al., 1985); DNET, CGCGAATTCGCG + netropsin complex (this work). All the common atoms were used in all cases for calculating the root mean square deviation (RMSD).

The e^{6}G4-C21 base pair has a configuration similar to a normal Watson–Crick base pair, but a close inspection of it suggests that it is similar to the bifurcated base-pairing configuration of Figure 1A. The N^4 of C21 is 3.72 Å from the O^6 and 3.27 Å from the N^1 of e^{6}G4, and the N^2 of e^{6}G4 is 2.82 Å from the O^2 and 2.75 Å from the N^3 of C21. The base pair is slightly opened up toward the major groove with the increase in distance between N^4 of C21 and O^6 of e^{6}G4 reducing the hydrogen bond observed in e^{6}G–DH342 and e^{6}G–DH258 to a charge interaction in e^{6}G–DNET. In this case, only the amino group N^2 of e^{6}G4 participates in three-centered hydrogen bonding interactions. The ED map indicates that the ethyl group for the e^{6}G4 base is proximal to N^7 and is out of the best plane of guanine by 0.60 Å. The distance between the N^7 atom and the C_β atom (C_α is the methylene carbon and C_β is the methyl carbon of the ethyl group) is 2.53 Å. The proximal orientation of the ethyl group effectively blocks any access of solvent, metal, or protein to the N^7 position of e^{6}G. The torsion angle about the O^6–C_α bond is in the gauche$^+$ conformation (56°). Notice that the base pair is quite distorted with a buckle of −22°, −18°, −20° and propeller twist of 2°, 6°, −3° for the e^{6}G–DH342, e^{6}G–DH258, and e^{6}G–DNET complexes, respectively.

The e^{6}G16-C9 base pair adopts a wobble configuration of Figure 1B. The N^4 of C9 is 2.86 Å from the N^1 of G16, and the N^2 of G16 is 2.61 Å from the N^3 of the C9. In the e^{6}G–DH342 this is due to the fact that the ethyl group is proximal to N^7, as clearly shown in the difference Fourier ED for the ethyl group (Sriram et al., 1992). In the e^{6}G–DH258, it is again due to the orientation of the ethyl group. But in the e^{6}G–DNET structure the reason is less clear. We noted that the ED envelope is a little more diffuse for the e^{6}G16-C9 base pair than that of the e^{6}G4-C21 base pair in all the e^{6}G complexes. This is markedly so for the e^{6}G–DH342 where the ethyl group may be accommodated in two different orientations. For the e^{6}G–DNET structure, only one ethyl group orientation is possible for the e^{6}G16-C9 base pair (Figure 7), and the ethyl group is out of the best plane of guanine (by 0.72 Å). Here the C^6–O^6–C_α–C_β torsion angle is in the anticlinical$^+$ conformation (118°) and the N1–C^6–O^6–C_α

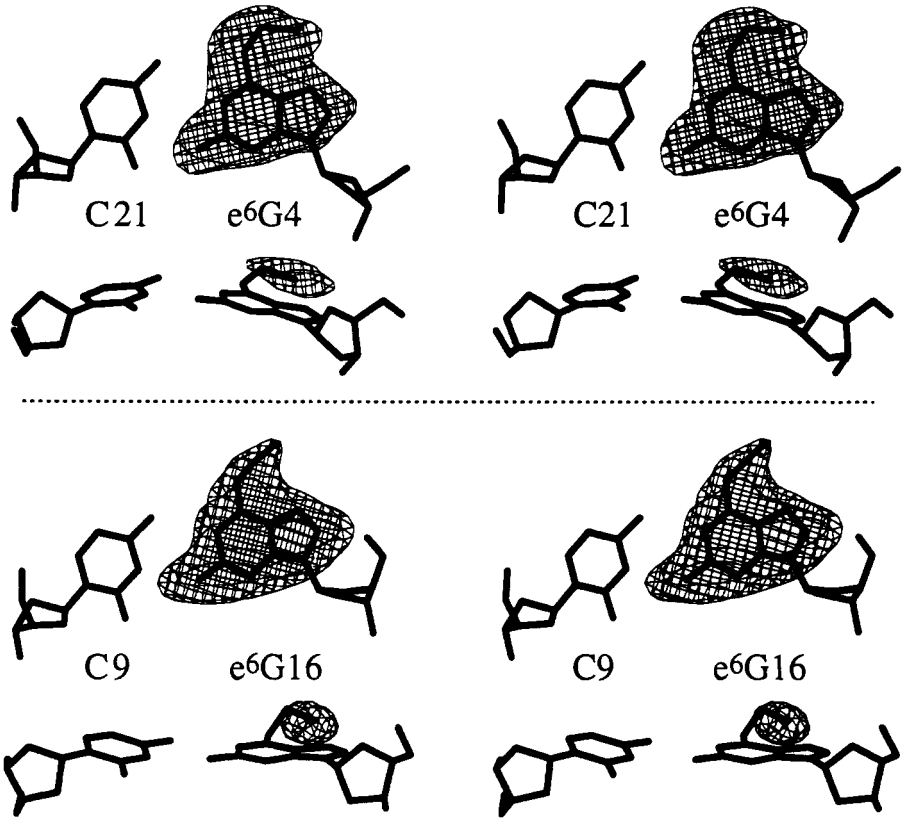

FIGURE 7: Stereoscopic ($|F_o| - |F_c|$) difference Fourier maps displaying in detail the two e^{6}G-C base pairs of the d(CGC[e^{6}G]AATTCGCG)–netropsin complex where the e^{6}G base or the ethyl group is removed from the phase contribution. (Top) The e^{6}G4-C21 base pair adopts an opened bifurcated configuration. (Bottom) The e^{6}G16-C9 base pair adopts the wobble configuration with the ethyl group in the proximal orientation. In both base pairs, the ethyl group is out of the plane of guanine base.

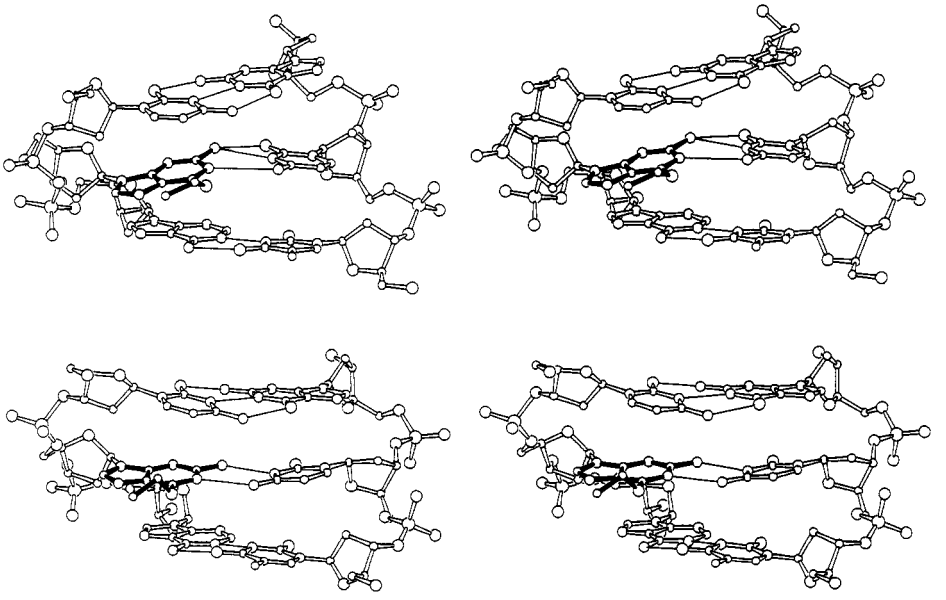

FIGURE 8: Stereoscopic skeletal drawings displaying in detail the two e^{6}G-C base pairs along with the adjacent base pairs of the d(CGC-[e^{6}G]AATTCGCG)–netropsin complex. (Top) The e^{6}G4-C21 base pair adopts the bifurcated configuration with an opening toward the major groove. The hydrogen bond from N^4 of C to O^6 of e^{6}G observed in e^{6}G–DH342 and e^{6}G–DH258 complexes is reduced to a weak charge interaction in the present case. (Bottom) The e^{6}G16-C9 base pair adopts the wobble configuration with the ethyl group in the proximal orientation. The influence of e^{6}G-C base-pairing scheme on adjacent base pairs is demonstrated. Hydrogen bonds are shown as thin lines to emphasize the base-pairing scheme (3.3-Å distance cutoff).

torsion angle is also in the anticlinical$^+$ conformation (121°). The distance between the C_α atom and the N^4 of C9 is 4.8 Å.

The ethyl orientation for the e^{6}G16-C9 base pair of the e^{6}G–DNET structure lies well out of plane of the guanine base and does not seem to be the factor influencing the base pair, but instead the orientation of the ethyl group seems to be the consequence of the wobble base-pairing scheme adopted. The orientation of the ethyl groups is different in each of the d(CGC[e^{6}G]AATTCGCG)–drug complexes. The ethyl groups of the two e^{6}Gs lie in the major groove of the helix, and they are both out of the plane of the respective guanine

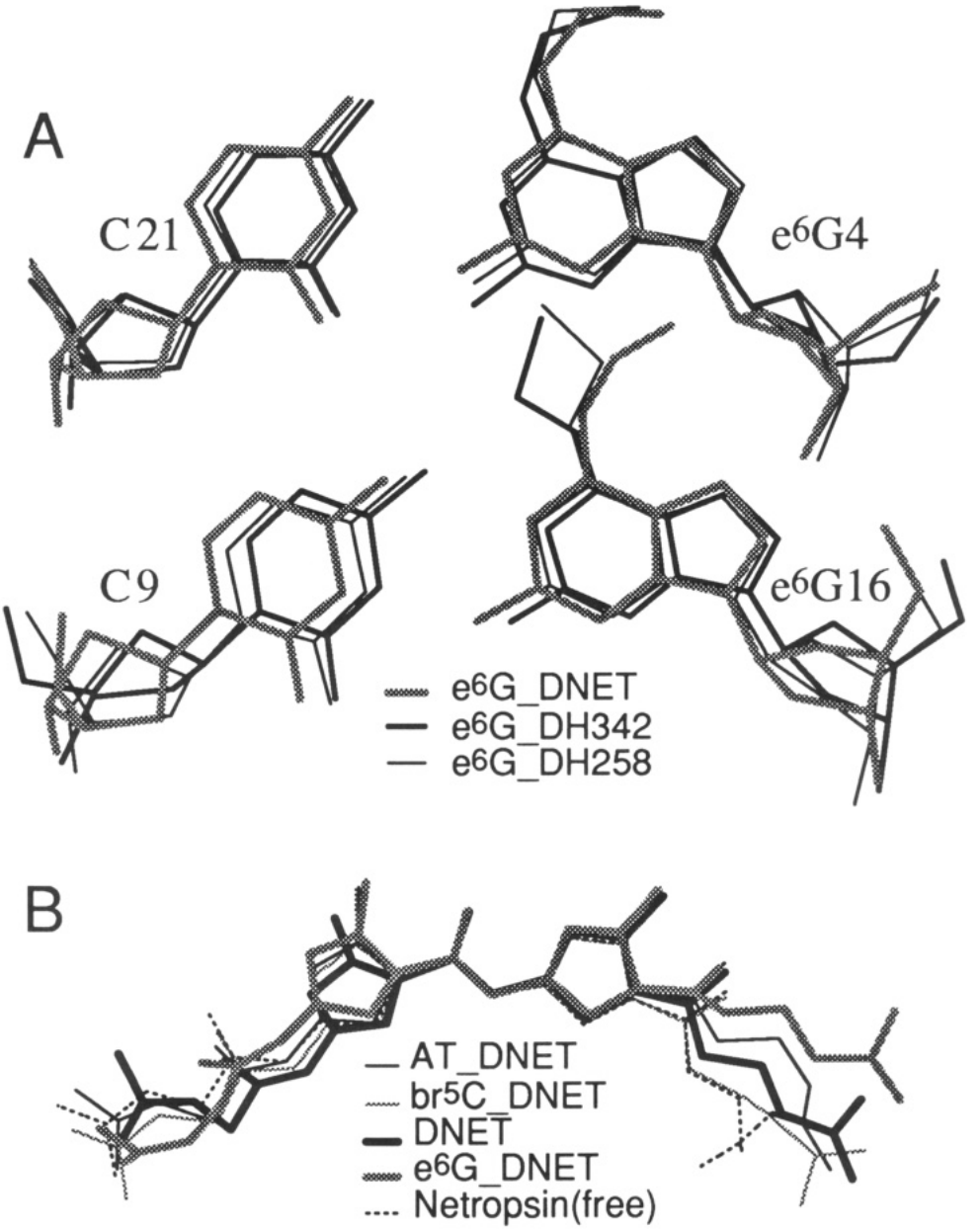

FIGURE 9: (A) Least-squares fit of e^{6}G4-C21 and e^{6}G16-C9 base pairs from e^{6}G–DNET, e^{6}G–DH342, and e^{6}G–DH258 complexes. Each structure shows a slightly different conformation, but the base-pairing falls within a single scheme (bifurcated or wobble type) for a given base pair in all three complexes. The differences observed are a reflection of variation in the local environment around the e^{6}G-C base pairs. (B) A superposition of netropsin drug molecules from six different crystal structures by least-squares fitting their central peptide bond together. All netropsin molecules have slightly different conformation. This reflects the drug's ability to adapt to different local environment found in the minor groove of dodecamer helices for the DNA–netropsin complexes and in the crystal lattice for the free drug.

bases (Figure 7). In the e^{6}G–DH258 and e^{6}G–DH342 structures the ethyl group of e^{6}G16 approaches the N^4 of C9 forcing the base pair to adopt the wobble configuration. The ethyl groups e^{6}G4 and e^{6}G16 are close to the N^4/C^5 of C21 and C9 cytosines, respectively. These close contacts seem to induce a substantial conformational distortion in the e^{6}G-C base pairs of the Hoechst structures.

A significant conformational distortion in the base pair is observed in all cases with a strikingly large buckle or propeller twist relative to the native dodecamer–drug complexes. Interestingly, the buckle undergoes a larger change for the netropsin structure, whereas the propeller twist of the same base pair undergoes a large change for the Hoechst (H342 and H258) structures, when compared to respective unmodified dodecamer–drug complex. The central four AATT base pairs have a significantly larger buckle in the e^{6}G dodecamer–drug complexes, when compared with their respective native dodecamer–drug complex. This indicates that the conformational distortion of the e^{6}G-C base pairs is passed onto the neighboring base pairs as well.

Our results here show unambiguously that the e^{6}G-C base pair may adopt the bifurcated pairing configuration (Figure 1A) or the wobble configuration (Figure 1B) near physiological pH condition. Both types have been suggested as possible base-pairing schemes by theoretical calculations (Pedersen et al., 1990). A common feature associated with the e^{6}G-C base pair at neutral pH is the distorted conformation (high buckle and pr tw angles) in and around the lesion site. This feature may be recognized more easily by the appropriate repair enzymes. However, it is clear that the bifurcated pairing configuration (Figure 1A) is not very different from the normal Watson–Crick G-C configuration, and this would make it more difficult for the repair enzymes to recognize the lesion site. This could also explain why C can still be incorporated

in the daughter strand opposite to the e^6G lesion during replication.

The wobble type (Figure 1B) and bifurcated type (Figure 1A) of base-pairing schemes seen in the crystal structures of d(CGC[e^6G]AATTCGCG)–drug complexes seem to explain available experimental observations and biological consequences of the e^6G lesion. The results from the NMR study at neutral pH of d(CGC[e^6G]AGCTCGCG) (Kalnik et al., 1989) suggested a wobble base pair. Williams and Shaw (1987) proposed that m^6G does not pair with a neutral C; instead it only pairs with a protonated C on the basis of their NMR study of nucleosides in deuterated chloroform. A pH-dependent melting study has been carried out on d(CGC-[e^6G]AATTCGCG), the same sequence used in this work, which showed a slightly biphasic melting curve with the highest T_m of 298 K at pH = 5.0 (Leonard et al., 1990). This was interpreted as the result of the formation of the e^6G-C^+ configuration (Figure 1D) in the duplex. Curiously, we have not been able to obtain a suitable crystal of d(CGC[e^6G]AATT CGCG) under low pH condition (<6.0), with or without drug. As noted above, the bifurcated pairing configuration (Figure 1A) is similar to the normal Watson–Crick G-C pairing. Consequently, we believe that the e^6G-C base pair may adopt a conformation ranging from a wobble to a bifurcated type under physiological condition.

Our findings suggest that both proximal and distal orientations (to N^7) of the ethyl group are probable in B-DNA helix. The results from other modified DNAs are consistent with these observations. As mentioned above, in the Z-DNA crystal structure of d(CGC[m^6]CG) (Ginell et al., 1990) and in the B-DNA structure of d(CGC[m^6G]AATTTGCG) (Leonard et al., 1990), the methyl group of the m^6G adopts a proximal orientation, as is the methyl group of the N^6-methyl-A in d(CGCGA[m^6A]TTCGCG) (Frederick et al., 1988). But the methoxy group of the N^4-methoxy-C in d(CG-[N^4-methoxy-C]GCG) adopts a distal orientation (Van Meervelt et al., 1990). This suggests that the conformation of those exocyclic modifications depends on the local environment.

Drug–DNA Interactions. In the complexes, the netropsin molecule is observed to adjust its conformation to adapt to the contour of the narrow minor groove in B-DNA (Wang et al., 1990). Figure 9B shows the superposition of five netropsin drugs (four dodecamer–netropsin complexes and the free drug) by the least-squares fit of the central peptide bond atoms. The netropsin from each structure differs in the overall curvature and the dihedral angles between successive planar elements (pyrroles and amides). The propyl amidinium group of the free drug is bent significantly because of the intermolecular hydrogen bond in the crystal lattice. Conversely, the propyl amidinium group of the e^6G–DNET netropsin is different due to the strong hydrogen bond with a phosphate oxygen of $T19O^2P$.

The propyl amidinium end of netropsin shows a larger variation in curvature than the guanidium end. This feature is conserved even when netropsin refined in the inverted orientation is included in the least-squares fit (data not shown). It is not surprising as this indicates a nonsymmetrical DNA binding pocket available to the drug in the dodecamer crystal lattice. Analysis of the DNA–netropsin interactions indicates that the guanidinium end of netropsin is better anchored by favorable charge interactions in all cases, and this may be responsible for the lesser change in curvature observed at that end. The variation in the drug conformation may have some relevance in the ability of netropsin to bind to nucleic acids in other binding modes, as the one observed in the binding of netropsin to RNA where it was seen to bind across the major groove bridging the phosphates (Rubin & Sundaralingam, 1984). Two binding modes of netropsin to DNA poly(dA-dT) duplexes have been established (Burckhardt et al., 1985).

The detailed interactions between the crescent-shaped netropsin drug with DNA in the four different crystal structures are schematically summarized in Figure 10. The position of netropsin binding to DNA is slightly different in all cases. In the e^6G–DNET and DNET structures the netropsin position is different from that of Kopka et al. (1985) and moves upward (relative to strand 1 of the DNA duplex) by one base pair, making GAATT the binding site sequence. Netropsin actually covers about six base pairs while binding to DNA. In the e^6G–DNET structure the guanidium group of the netropsin is directly hydrogen bonded to the $O^{3'}$ atom (of G12#) at the end of the symmetry-related helix which packs into the minor groove in the crystal lattice. Netropsin interaction with the symmetry-related helix in the DNET structure is bridged by a water. In three of the four dodecamer–netropsin complexes (e^6G–DNET, DNET, and br^5C–DNET) a water molecule bridging interaction with the sugar–phosphate backbone of DNA is found at either end of netropsin. Is the water molecule associated with netropsin a factor in the sequence specific binding of netropsin? Is the new position for netropsin a crystallization condition artifact, or is it a display of the tolerance of netropsin for slightly different sequences so long as at least a core of four AT bases is present? Further studies are necessary to answer these questions.

The binding of the netropsin molecule to DNA is stabilized by several types of forces. All four structures reinforce the observations made previously regarding the molecular basis of the binding mode of minor groove drugs in general. As already discussed in an earlier review (Wang & Teng, 1990), this may be summarized here as (1) electrostatic attraction between the positively charged drug and the negatively charged DNA, (2) van der Waals interaction between the drug atoms with the DNA atoms along the two walls of the minor groove, (3) hydrogen bonds (both direct and bridged) between the NH of the peptide/guanidium/amidinium groups of netropsin and the $T\text{-}O^2$, $A\text{-}N^3$, $O^{4'}$ of sugar or the phosphate oxygens of DNA, (4) stacking interaction between the aromatic pyrrole rings and sugar $O^{4'}$ of DNA, and (5) favorable hydrophobic interaction associated with the formation of the DNA–drug complex. The specificity toward AT sequence is aided by the natural tendency of the AT segment to have a narrow minor groove which provides a favorable surrounding for the interactions described above. Finally, the additional N^2 amino group in guanine presents a steric hindrance toward the drug and pushes the drug away from the floor of the minor groove and this would diminish the binding interactions substantially. Therefore, netropsin binds preferentially to AT sequence over GC sequence. As the drug actually covers nearly six base pairs, it requires at least 4 AT core sequence for tight binding.

CONCLUSION

The structural analyses of the complexes of the DNA-containing e^6G lesions with the minor groove binding drugs (Hoechst 33258, Hoechst 33342, and netropsin) provide important information of how e^6G pairs with C. Our data suggest that the base-pairing scheme adopted can be the wobble or the bifurcated pairing, depending on the local environment. The bifurcated pairing configuration is similar to a normal G-C base pair. This structural similarity may allow the e^6G

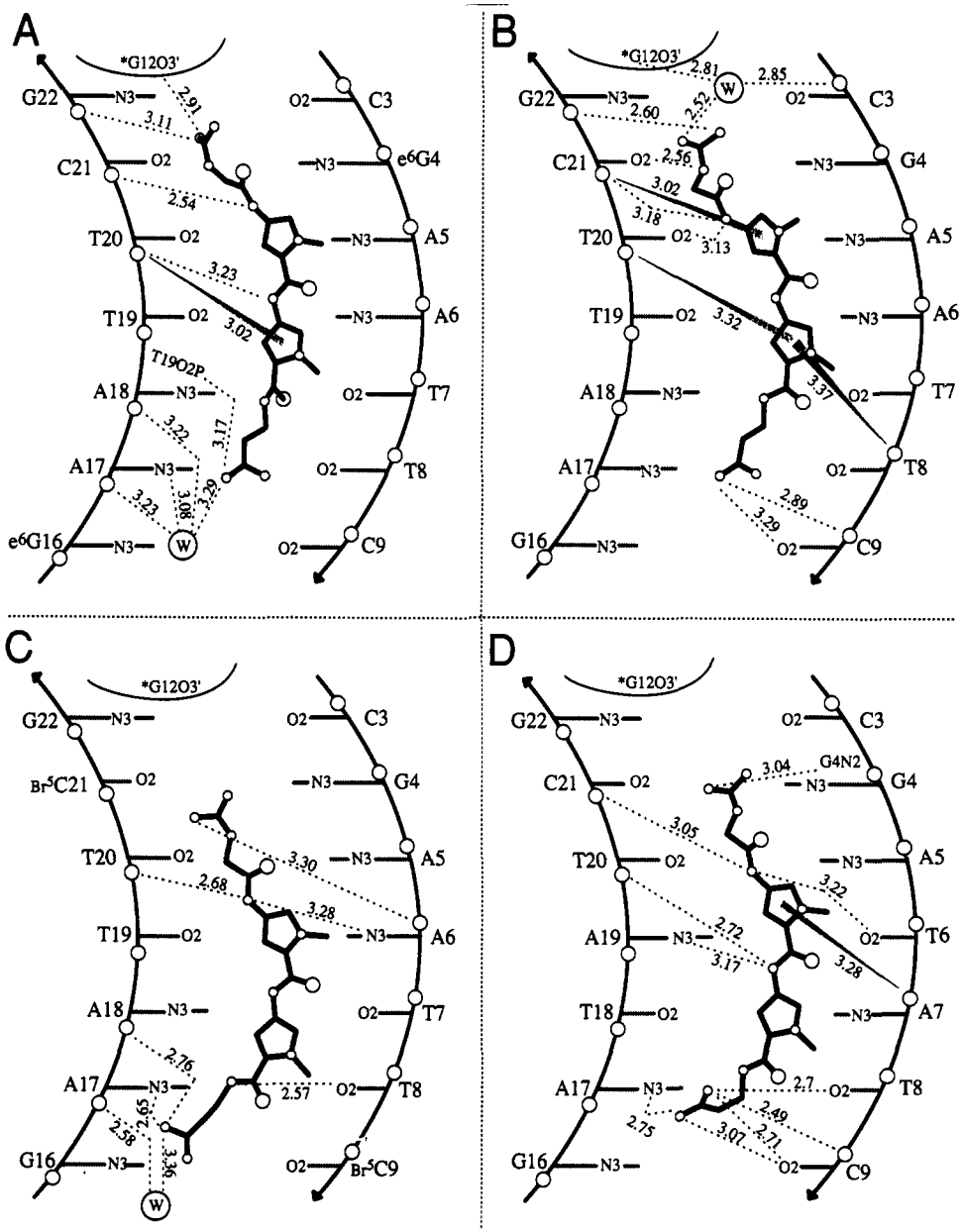

FIGURE 10: Schematic diagram showing the interactions between the netropsin and DNA in the four different structures: (A) e^6G–DNET, (B) DNET, (C) br^5C–DNET, and (D) AT–DNET. In all four structures, netropsin is sandwiched in the minor groove at the AATT site between the two antiparallel backbones of the DNA helix. However, they differ somewhat in the details of position and interaction. Many van der Waals interactions such as the dipole–π interaction (shown as elongated shaded triangles) between the $O^{4'}$ (e.g., from sugars of T8, T20, and C21 in the DNET complex) and the methylpyrroles of the netropsin are used to stabilize the binding. The sugar $O^{4'}$ are shown as open circles on the DNA strand. Hydrogen bonds (shown as dotted lines) contribute to the binding of the drug. A 3.3-Å distance cutoff and a less stringent angle cutoff were used for the hydrogen bonds shown. Water molecules bridging hydrogen bond between the drug and DNA are shown as an encircled letter W.

in DNA to escape the repair system. During replication, either C or T may be inserted in the daughter DNA strand across the e^6G site. However, there may be a dynamic equilibrium between the two configurations of the e^6G-C base pair, which presents an ambiguous signal to the polymerase and is subsequently edited out. In contrast, thymine can pair with e^6G in only one way (with a configuration similar to a regular Watson–Crick G-C base pair), albeit imperfect. This may be a plausible explanation of why thymine is found preferentially incorporated across the e^6G lesion site during replication (Loechler et al., 1984). Finally, the helix with e^6G lesions next to AATT sequence is stabilized by minor groove binding drugs. These results suggest that other lesions including mismatched base pairs may be similarly stabilized. It would be of interest to insert e^6G in different nucleotide sequence to understand more fully the sequence-dependent structural perturbations caused by O^6-alkylated lesions in DNA. The structural consequences of various types of DNA lesions (e.g., 8-oxoguanine) may also be studied in a similar way to aid the understanding of chemical carcinogenesis.

SUPPLEMENTARY MATERIAL AVAILABLE

Three tables listing the breakdown of reflections observed by resolution shell (Table 1S) and the DNA torsion angles of e^6G–DNET and DNET structures (Tables 2Sa and 2Sb, respectively) and two stereoscopic skeletal drawings displaying in detail the two e^6G-C base pairs along with the adjacent base pairs for the d(CGC[e^6G]AATTCGCG)–Hoechst 33258 complex (Figure 1S) and the d(CGC[e^6G]AATTCGCG)–Hoechst 33342 complex (Figure 2S) (5 pages). Ordering information is given on any current masthead page.

REFERENCES

Aymami, J., Coll, M., van der Marel, G. A., van Boom, J. H., Rich, A., & Wang, A. H.-J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 2526–2530.

Burckhardt, G., Votavova, H., Sponar, J., Luck, G., & Zimmer, C. (1985) *J. Biomol. Struct. Dyn. 2*, 721–736.

Carrondo, M., Coll, M., Aymami, J., Wang, A. H.-J., van der Marel, G. A., van Boom, J. H., & Rich, A. (1989) *Biochemistry 28*, 7849–7859.

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.

Coll, M., Sherman, S. E., Gibson, D., Lippard, S. J., & Wang, A. H.-J. (1990) *J. Biomol. Struct. Dyn. 8*, 315–330.

Crothers, D. M., Haran, T. E., & Nadeau, J. G. (1990) *J. Biol. Chem. 265*, 7093–7096.

Dervan, P. B. (1986) *Science 232*, 464–471.

Dickerson, R. E., Bansal, M., Calladine, C. R., Diekman, S., Hunter, W. N., Kennard, O., von Kitzing, E., Lavery, R., Nelson, H. C. M., Olson, W. K., Saenger, W., Shakked, Z., Sklenar, H., Soumpasis, D. M., Tung, C.-S., Wang, A. H.-J., & Zhurkin, V. B. (1989) *Nucleic Acids Res. 17*, 1797–1803.

Drew, H. R., & Dickerson, R. E. (1981) *J. Mol. Biol. 151*, 535–556.

Ellison, K. S., Dogliotti, E., Connors, T. D., Basu, A. K., & Essigmann, J. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8620–8624.

Frederick, C. A., Quigley, G. J., van der Marel, G. A., van Boom, J. H., Wang, A. H.-J., & Rich, A. (1988) *J. Biol. Chem. 263*, 17872–17879.

Ginell, S. L., Kuzmich, S., Jones, R. A., & Berman, H. M. (1990) *Biochemistry 29*, 10461–10465.

Hendrickson, W. A., & Konnert, J. H. (1979) in *Biomolecular Structure, Conformation, Function and Evolution* (Srinvasan, R., Ed.) pp 43–57, Pergamon, Oxford.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–272.

Jurgen, T., Huh, N.-H., Nehls, P., Eberle, G., & Rajewsky, M. F. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 9883–9887.

Kalnik, M. W., Li, B. F. L., Swann, P. F., & Patel, D. J. (1989) *Biochemistry 28*, 6182–6192.

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.

Larsen, T. A., Goodsell, D. S., Cascio, D., Grzeskowiak, K., & Dickerson, R. E. (1989) *J. Biomol. Struct. Dyn. 7*, 477–491.

Leonard, G. A., Thomson, J., Watson, W. P., & Brown, T. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 9573–9576.

Leupin, W., Chazin, W. J., Hyberts, S., Denny, W. A., & Wuthrich, K. (1986) *Biochemistry 25*, 5902–5908.

Lindahl, T., Sedgwick, B., Sekiguchi, M., & Nakabeppu, Y. (1988) *Annu. Rev. Biochem. 57*, 133–158.

Loechler, E. L., Green, C. L., & Essigmann, J. M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 6271–6275.

Nelson, H. C. M., Finch, J. T., Luisi, B. F., & Klug, A. (1987) *Nature 330*, 221–226.

Parthasarathy, R., & Fridey, S. M. (1986) *Carcinogenesis 7*, 221–227.

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.

Patel, D. J., Kozlowski, S. A., Rice, J. A., Broka, C., & Itakura, K. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 7281–7284.

Pedersen, L. G., Darden, T. A., Deerfield, D. W., II, Anderson, M. W., & Hoel, D. G. (1990) *Carcinogenesis 9*, 1553–1562.

Roelen, H. C. P. F., Brugghe, H. F., van der Elst, H., Klien, J. C., van der Marel, G. A., & van Boom, J. H. (1992) *Recl. Trav. Chim. Pays-Bas* (in press).

Rubin, J., & Sundaralingam, M. (1984) *J. Biomol. Struct. Dyn. 2*, 165–174.

Singer, B., Bodell, W. L., Cleaver, J. E., Thomas, G. H., & Rajewsky, M. F. (1978) *Nature 276*, 85–88.

Sriram, M., van der Marel, G. A., van Boom, J. H., & Wang, A. H.-J. (1991) *Seventh Conversion in Biomolecular Stereodynamics*, p a211, Adenine Press, New York.

Sriram, M., van der Marel, G. A., Roelen, H. C. P. F., van Boom, J. H., & Wang, A. H.-J. (1992) *EMBO J. 11*, 225–232.

Taylor, J. S., Schultz, P. G., & Dervan, P. B. (1984) *Tetrahedron 40*, 457–465.

Teng, M.-K., Usman, N., Frederick, C. A., & Wang, A. H.-J. (1988) *Nucleic Acids Res. 16*, 2671–2690.

Thomale, J., Huh, N.-H., Nehls, P., Eberle, G., & Rajewsky, M. F. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 9883–9887.

Van Meervelt, L., Moore, H. M., Lin, P. K. T., Brown, D. M., & Kennard, O. (1990) *J. Mol. Biol. 216*, 773–781.

Wang, A. H.-J., & Gao, Y.-G. (1990) *Methods 1*, 91–99.

Wang, A. H.-J., & Teng, M.-K. (1990) in *Crystallographic and Modeling Methods in Molecular Design* (Bugg, C. E., & Ealick, S. E., Eds.) pp 123–150, Springer-Verlag, New York.

Wang, A. H.-J., Liaw, Y.-C., Robinson, H., & Gao, Y.-G. (1990) in *23rd Jerusalem Symposium in Quantum Chemistry and Biochemistry* (Pullman, B., & Jortner, J., Eds.) pp 1–21, Kluwer Academic Press, Dordrecht, Netherlands.

Westhof, E., Dumas, P., & Moras, D. (1985) *J. Mol. Biol. 184*, 119–145.

Williams, L. D., & Shaw, B. R. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 1779–1783.

Zimmer, C., & Wahnert, U. (1986) *Prog. Biophys. Mol. Biol. 47*, 31–112.