

# A Bridging Water Anchors the Tethered 5-(3-Aminopropyl)-2'-deoxyuridine Amine in the DNA Major Groove Proximate to the N+2 C•G Base Pair: Implications for Formation of Interstrand 5'-GNC-3' Cross-Links by Nitrogen Mustards<sup>†,‡</sup>

Feng Wang,<sup>§</sup> Feng Li,<sup>||</sup> Manjori Ganguly,<sup>⊥</sup> Luis A. Marky,<sup>⊥</sup> Barry Gold,<sup>⊥,\*,+</sup> Martin Egli,<sup>||</sup> and Michael P. Stone<sup>\*,§</sup>

Department of Chemistry, Center for Structural Biology, Center in Molecular Toxicology, and Vanderbilt-Ingram Cancer Center, Vanderbilt University, Nashville, Tennessee 37235, Department of Biochemistry, Center for Structural Biology, Center in Molecular Toxicology, School of Medicine, Vanderbilt University, Nashville, Tennessee 37232, Eppley Institute for Research in Cancer, University of Nebraska Medical Center, Omaha, Nebraska 68198-6805, Department of Pharmaceutical Sciences, University of Nebraska Medical Center, Omaha, Nebraska 68198-6025, and Department of Pharmaceutical Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15261

Received March 4, 2008; Revised Manuscript Received April 2, 2008

**ABSTRACT:** Site-specific insertion of 5-(3-aminopropyl)-2'-deoxyuridine (Z3dU) and 7-deaza-dG into the Dickerson–Drew dodecamers 5'-d(C<sup>1</sup>G<sup>2</sup>C<sup>3</sup>G<sup>4</sup>A<sup>5</sup>A<sup>6</sup>T<sup>7</sup>T<sup>8</sup>C<sup>9</sup>Z<sup>10</sup>C<sup>11</sup>G<sup>12</sup>)-3'•5'-d(C<sup>13</sup>G<sup>14</sup>C<sup>15</sup>G<sup>16</sup>A<sup>17</sup>A<sup>18</sup>T<sup>19</sup>T<sup>20</sup>-C<sup>21</sup>Z<sup>22</sup>C<sup>23</sup>G<sup>24</sup>)-3' (named DDD<sup>Z10</sup>) and 5'-d(C<sup>1</sup>G<sup>2</sup>C<sup>3</sup>G<sup>4</sup>A<sup>5</sup>A<sup>6</sup>T<sup>7</sup>X<sup>8</sup>C<sup>9</sup>Z<sup>10</sup>C<sup>11</sup>G<sup>12</sup>)-3'•5'-d(C<sup>13</sup>G<sup>14</sup>C<sup>15</sup>G<sup>16</sup>A<sup>17</sup>A<sup>18</sup>-T<sup>19</sup>X<sup>20</sup>C<sup>21</sup>Z<sup>22</sup>C<sup>23</sup>G<sup>24</sup>)-3' (named DDD<sup>2+Z10</sup>) (X = Z3dU; Z = 7-deaza-dG) suggests a mechanism underlying the formation of interstrand N+2 DNA cross-links by nitrogen mustards, e.g., melphalan and mechlorethamine. Analysis of the DDD<sup>2+Z10</sup> duplex reveals that the tethered cations at base pairs A<sup>5</sup>•X<sup>20</sup> and X<sup>8</sup>•A<sup>17</sup> extend within the major groove in the 3'-direction, toward conserved Mg<sup>2+</sup> binding sites located adjacent to N+2 base pairs C<sup>3</sup>•Z<sup>22</sup> and Z<sup>10</sup>•C<sup>15</sup>. Bridging waters located between the tethered amines and either Z<sup>10</sup> or Z<sup>22</sup> O<sup>6</sup> stabilize the tethered cations and allow interactions with the N + 2 base pairs without DNA bending. Incorporation of 7-deaza-dG into the DDD<sup>2+Z10</sup> duplex weakens but does not eliminate electrostatic interactions between tethered amines and Z<sup>10</sup> O<sup>6</sup> and Z<sup>22</sup> O<sup>6</sup>. The results suggest a mechanism by which tethered N7-dG aziridinium ions, the active species involved in formation of interstrand 5'-GNC-3' cross-links by nitrogen mustards, modify the electrostatics of the major groove and position the aziridinium ions proximate to the major groove edge of the N+2 C•G base pair, facilitating interstrand cross-linking.

Nitrogen mustards, such as melphalan and mechlorethamine, were among the first effective clinically used anticancer compounds (1). Their biological activity is derived

from DNA alkylation. Among the numerous adducts formed, N7-dG to N7-dG interstrand cross-links (ICLs)<sup>1</sup> are a main contributor to cytotoxicity since they prevent strand separation required for DNA replication and transcription. The initial step in ICL formation is reaction with DNA (Scheme 1) to afford a reactive monofunctional N7-dG adduct, which can either react with another N7-dG site on the complementary strand, if one is available, to afford an ICL or undergo solvolysis to yield a stable monofunctional lesion. The first and second alkylation steps in the cross-linking involve aziridinium ions (Scheme 1).

The ~7.5 Å length of the -(CH<sub>2</sub>)<sub>2</sub>N(R)(CH<sub>2</sub>)<sub>2</sub>- linkage suggests that the ICL should form between N7-dG at nearest-neighbor base pairs in 5'-GC-3' repeats (Scheme 1) (2). Molecular modeling confirms that this would be the predicted

<sup>†</sup> This work was supported by NIH Grants CA-76049 (B.G., L.A.M., and M.P.S.) and GM-55237 (M.E.), National Cancer Institute Center Grants CA-36727 (University of Pittsburgh) and CA-68485 (Vanderbilt University), and NSF Grant MCB-0315746 (L.A.M.). Crystallographic data were collected on beamline 22-ID by the Southeast Regional Collaborative Access Team (SER-CAT) at the Advanced Photon Source, Argonne National Laboratory (Argonne, IL). Supporting institutions may be found at [www.ser-cat.org/members.html](http://www.ser-cat.org/members.html). Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract W-31-109-Eng-38.

<sup>‡</sup> Crystallographic coordinates and structure factors have been deposited in the Protein Data Bank as entries 2QEG (DDD<sup>Z10</sup>) and 2QEF (DDD<sup>2+Z10</sup>).

\* To whom correspondence should be addressed. E-mail: [michael.p.stone@vanderbilt.edu](mailto:michael.p.stone@vanderbilt.edu). Phone: (615) 322-2589. Fax: (615) 322-7591.

<sup>§</sup> Center in Molecular Toxicology, and Vanderbilt-Ingram Cancer Center, Vanderbilt University.

<sup>||</sup> Center for Structural Biology, Center in Molecular Toxicology, School of Medicine, Vanderbilt University.

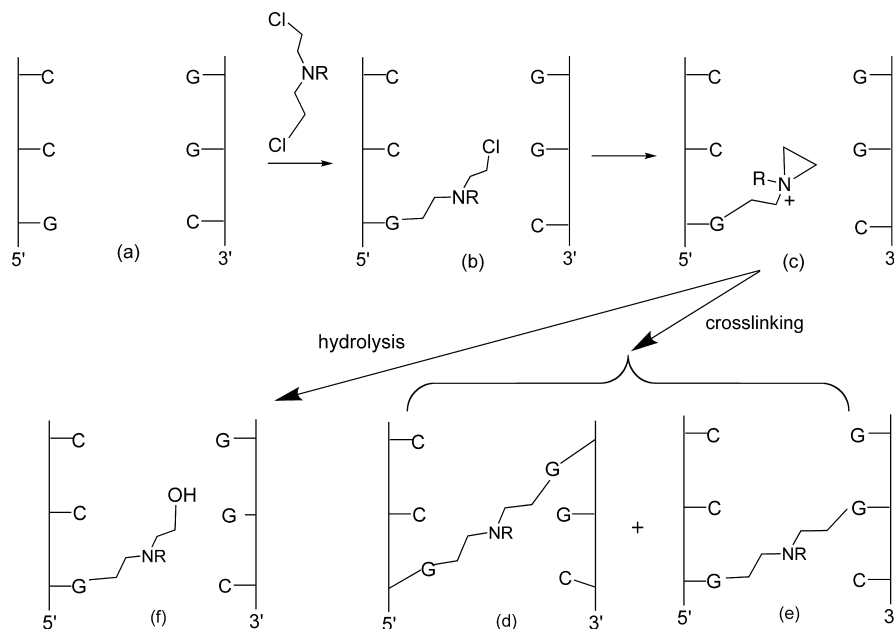
<sup>⊥</sup> Eppley Institute for Research in Cancer, University of Nebraska Medical Center.

<sup>+</sup> Department of Pharmaceutical Sciences, University of Nebraska Medical Center.

<sup>+</sup> Department of Pharmaceutical Sciences, University of Pittsburgh.

<sup>1</sup> Abbreviations: CCD, charge-coupled device; 7-deaza-dG, 7-deaza-2'-deoxyguanosine; DDD, Dickerson–Drew dodecamer; 5'-d(C<sup>1</sup>G<sup>2</sup>C<sup>3</sup>G<sup>4</sup>A<sup>5</sup>A<sup>6</sup>T<sup>7</sup>T<sup>8</sup>C<sup>9</sup>G<sup>10</sup>C<sup>11</sup>G<sup>12</sup>)-3'•5'-d(C<sup>13</sup>G<sup>14</sup>C<sup>15</sup>G<sup>16</sup>A<sup>17</sup>A<sup>18</sup>T<sup>19</sup>T<sup>20</sup>-C<sup>21</sup>G<sup>22</sup>C<sup>23</sup>G<sup>24</sup>)-3'; DDD<sup>Z10</sup>, Dickerson–Drew dodecamer containing 7-deaza-dG at the G<sup>10</sup> nucleotide position; DDD<sup>2+Z10</sup>, Dickerson–Drew dodecamer containing 7-deaza-dG at the G<sup>10</sup> nucleotide position and Z3dU at the T<sup>8</sup> position; DDD<sup>4+</sup>, Dickerson–Drew dodecamer containing Z3dU at the T<sup>7</sup> and T<sup>8</sup> positions; DSC, differential scanning calorimetry; ICL, interstrand cross-link; MPD, 2-methyl-2,4-pentandiol; Z3-dU, 5-(3-aminopropyl)-2'-deoxyuridine.

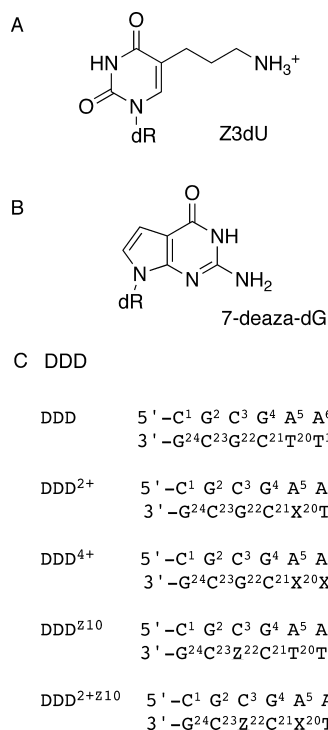
Scheme 1: Reactions of Nitrogen Mustards with DNA Giving Monofunctional and Cross-Linking Adducts



product. However, the Loechler (3, 4) and Hopkins groups (5) independently determined that the actual ICL formation occurs between N7-dG atoms at next nearest-neighbor C•G base pairs in 5'-GNC-3' sequences, i.e., at the  $N + 2$  base pair (Scheme 1d), with only a small amount of cross-linking at the predicted 5'-GC-3' sequences (Scheme 1e). This result was unanticipated because formation of the 5'-GNC-3' ICL requires a significant bending of B-DNA (6, 7); the ICL is  $\sim 1.5$  Å too short to bridge the atoms in a classical B-DNA structure (8). When one considers that the second step in cross-link formation involves an aziridinium ion (Scheme 1), the disparity in the distance is actually closer to 3.0 Å. These observations lead to the conclusion that in the transition state for  $N + 2$  ICL formation, the DNA duplex must deviate from the linear B-form, allowing the covalent linkage between the two atoms that form the second bond.

It is thought that the  $N + 2$  ICL cross-link represents the kinetically favored reaction product (9, 10), but the mechanism of cross-link formation remains obscure. Positioning of a tethered aziridinium ion within the major groove modifies the electrostatics of the DNA duplex, which could transiently bend the duplex, thus positioning the aziridinium ion proximate to the electronegative major groove edge of the  $N + 2$  C•G base pair that is 3' to the point of attachment (11). The tethered 5- $\omega$ -aminopropyl-2'-deoxypyrimidine Z3dU (Scheme 2) serves as a chemically stable model for the monofunctional N7-dG mustard adduct that is the intermediate species formed in the  $N + 2$  cross-linking reaction. The "reach" of Z3dU in DNA as monitored by electrostatic footprinting experiments is longer than what would be predicted by the physical length of the cationic side chain (12–14). A second indication that the electrostatic interactions of the charged amino group with DNA may be important in the formation of the 5'-GNC-3' ICL is the observation that cationic Z3dU substitutions, when appropriately phased on the same side of the helix as an intrinsically bent A-tract, reduce DNA mobility in gels (15, 16). This is suggestive of bending or anisotropic flexing of the DNA, which must be present in the transition state complex

Scheme 2: Structures of (A) Z3dU, 5-(3-Aminopropyl)-2'-deoxyuridine, (B) 7-Deaza-dG, and (C) 7-Deaza-dG and Z3dU-Modified Duplexes (X = Z3dU, and Z = 7-deaza-dG)<sup>a</sup>



<sup>a</sup> The nucleotides of one duplex strand are labeled from 1 to 12, and nucleotides of the complementary strand are labeled from 13 to 24.

for the formation of the  $N + 2$  ICL cross-link (9–11). The electrostatic mechanism responsible for these observations could be similar to that envisioned by Rouzina and Bloomfield for the bending of DNA by divalent cations (17).

To understand how localized cationic charge in the major groove affects structure, we initiated a series of studies to structurally, thermodynamically, and chemically characterize a self-complementary Dickerson-Drew dodocamer 5'-

d(CGCGAATTCGCG)<sub>2</sub>-3' in which either or both of the T residues were replaced with 5-(3-aminopropyl)-2'-deoxyuridine (Z3dU) (Scheme 2). In summary, the structure of DNA and its thermodynamic properties were dependent on where the cationic side chain was attached (14). In DDD<sup>2+</sup>, the proton NMR resonances of the Z3dU moieties were individually resolved, consistent with ordered side chain conformations in the major groove (18). At both X<sup>8</sup> and X<sup>20</sup>, the tethered Z3dU amine exhibited a 3'-orientation with respect to the major groove. This corroborated both the chemical footprinting experiments and the predictions of molecular modeling. The NMR data suggested that the Z3dU cationic amines were transiently located in the vicinity of the electronegative N7 and O<sup>6</sup> atoms at the N + 2 position on the floor of the major groove (18), implicating at least transient axial bending by the tethered cation. A 1.6 Å resolution crystal structure was reported in the presence of Tl<sup>+</sup> (19). The X<sup>8</sup>- and X<sup>20</sup>-tethered amines were also oriented in the 3'-direction. In contrast, the basic amino side chains at positions X<sup>7</sup> and X<sup>19</sup> were directed into the major groove rather than toward the floor of the groove.

Herein, thermodynamic and crystallographic structural characterizations of the modified duplex 5'-d(C<sup>1</sup>G<sup>2</sup>C<sup>3</sup>-G<sup>4</sup>A<sup>5</sup>A<sup>6</sup>T<sup>7</sup>X<sup>8</sup>C<sup>9</sup>Z<sup>10</sup>C<sup>11</sup>G<sup>12</sup>)-3'·5'-d(C<sup>13</sup>G<sup>14</sup>C<sup>15</sup>G<sup>16</sup>A<sup>17</sup>A<sup>18</sup>T<sup>19</sup>-X<sup>20</sup>C<sup>21</sup>Z<sup>22</sup>C<sup>23</sup>G<sup>24</sup>)-3' (X = Z3dU, and Z = 7-deaza-dG) (DDD<sup>2+Z10</sup>) (Scheme 2) are presented. The positioning of 7-deaza-dG at positions Z<sup>10</sup> and Z<sup>22</sup>, designed to alter the electrostatics of the major groove at base pairs C<sup>3</sup>·Z<sup>22</sup> and Z<sup>10</sup>·C<sup>15</sup>, facilitated crystallographic analysis of this dodecamer. The data reveal that the tethered Z3dU cations at base pairs A<sup>5</sup>·X<sup>20</sup> and X<sup>8</sup>·A<sup>17</sup> extend within the major groove in the 3'-direction, toward conserved Mg<sup>2+</sup> binding sites located adjacent to N + 2 base pairs C<sup>3</sup>·Z<sup>22</sup> and Z<sup>10</sup>·C<sup>15</sup>. Bridging water molecules located between the tethered cationic amines and either Z<sup>10</sup> or Z<sup>22</sup> O<sup>6</sup> stabilize the tethered Z3dU cations and allow the Z3dU amines to interact with the N + 2 base pairs without DNA bending. The distances between the cationic amines tethered to X<sup>8</sup> or X<sup>20</sup> and Z<sup>10</sup> O<sup>6</sup> or Z<sup>22</sup> O<sup>6</sup> are ~0.5 Å longer than the distance observed between the cationic amine tethered to X<sup>8</sup> and G<sup>10</sup> O<sup>6</sup> in 5'-d(C<sup>1</sup>G<sup>2</sup>C<sup>3</sup>-G<sup>4</sup>A<sup>5</sup>A<sup>6</sup>X<sup>7</sup>X<sup>8</sup>C<sup>9</sup>G<sup>10</sup>C<sup>11</sup>G<sup>12</sup>)-3'·5'-d(C<sup>13</sup>G<sup>14</sup>C<sup>15</sup>G<sup>16</sup>A<sup>17</sup>A<sup>18</sup>X<sup>19</sup>-X<sup>20</sup>C<sup>21</sup>G<sup>22</sup>C<sup>23</sup>G<sup>24</sup>)-3' (DDD<sup>4+</sup>), which maintains the N7 atom at nucleotides G<sup>10</sup> and G<sup>22</sup> (19). Incorporation of 7-deaza-dG into the DDD<sup>2+Z10</sup> duplex weakens but does not eliminate electrostatic interactions between tethered cationic amines and Z<sup>10</sup> O<sup>6</sup> and Z<sup>22</sup> O<sup>6</sup>. The results support a model in which the presence of tethered N7-dG aziridinium ions, which are the active species involved in formation of interstrand 5'-GNC-3' cross-links by nitrogen mustards, modifies the electrostatics of the major groove and positions the aziridinium ions proximate to the major groove edge of the N + 2 C·G base pair, facilitating interstrand cross-linking. The thermodynamic results indicate that the 3'-orientation of the cationic Z3dU amine in the major groove is accompanied by the release of bound cations from the DDD<sup>2+</sup> duplex, presumably located near the N7 and O<sup>6</sup> atoms of G<sup>10</sup> and G<sup>22</sup>. Overall, the work provides insight into how melphalan- and mechlorethamine-induced ICL formation occurs between N7 G atoms at 5'-GNC-3' sequences, i.e., at the N + 2 base pair.

## MATERIALS AND METHODS

**Materials.** The oligodeoxynucleotides 5'-d(CGCGA-ATTCZCG)-3' (DDD<sup>Z10</sup>) and 5'-d(CGCGAATXCZCG)-3' (DDD<sup>2+Z10</sup>) were synthesized (20) and purified using reversed-phase semipreparative HPLC (Phenomenex, Phenyl-Hexyl, 5 μm, 250 mm × 10.0 mm) equilibrated with 0.1 M ammonium formate (pH 7.0) and desalted using a G25 column. The samples were characterized by capillary gel electrophoresis and mass spectrometry. The oligodeoxynucleotide concentrations were determined using an extinction coefficient of  $1.10 \times 10^5 \text{ M}^{-1} \text{ cm}^{-1}$  at 260 nm (21).

**Temperature-Unfolding Profiles (Melting Curves).** Heat capacities versus temperature profiles were measured with a VP-DSC differential scanning calorimeter (Microcal, Inc., Northampton, MA). The dry oligodeoxynucleotides were dissolved in 10 mM sodium phosphate buffer (pH 7) and adjusted to the desired ionic strength with NaCl for all unfolding experiments. DSC melting curves were obtained in the differential mode, by comparing a duplex solution against a buffer solution, in the temperature range of 0–105 °C. The experimental curve was normalized by the heating rate, and a buffer versus buffer scan was subtracted and normalized for the number of moles. The resulting monophasic, or biphasic, curves were then analyzed by deconvolution with Origin version 5.0 (Microcal); their integration ( $\int \Delta C_p dT$ ) yielded the molar unfolding enthalpy ( $\Delta H_{\text{cal}}$ ), which was independent of the nature of the transition (22). The molar entropy ( $\Delta S_{\text{cal}}$ ) was obtained similarly, using  $\int (\Delta C_p/T) dT$ . The free energy change at any temperature  $T$  was then obtained with the Gibbs equation:  $\Delta G^\circ(T) = \Delta H_{\text{cal}} - T\Delta S_{\text{cal}}$ .

Absorption versus temperature profiles (UV melts) for each duplex were measured at 260 nm using a thermoelectrically controlled UV-vis Aviv 14DS or Lambda 40-Perkin-Elmer spectrophotometers. The temperature was scanned at heating rates of 0.75–1.00 °C/min. UV melts were measured in the salt range of 10–200 mM NaCl at pH 7, and at a constant total strand concentration of 7 μM, to determine the differential binding of counterions,  $\Delta n_{\text{Na}^+}$ , which accompanied their helix-coil transitions. The  $T_M$  values of biphasic transitions were determined from the differential UV melts and only using the  $T_M$  of the duplex → random coil transition. This linking number was measured experimentally with the assumption that counterion binding to the helical and coil states of each oligonucleotide took place with a similar type of binding using the relationship (23)  $\Delta n_{\text{Na}^+} = 1.1(\Delta H_{\text{cal}}/RT_M^2)(\partial T_M/\partial \ln[\text{Na}^+])$ . The numerical factor corresponded to the conversion of ionic activities into concentrations. The first term in parentheses,  $(\Delta H_{\text{cal}}/RT_M^2)$ , was a constant determined directly from DSC experiments, where  $R$  was the gas constant. The second term in parentheses was also determined experimentally from the dependencies of  $T_M$  on salt concentration.

**Crystallization.** Crystals of the DDD<sup>Z10</sup> and DDD<sup>2+Z10</sup> DNAs were grown at 18 °C by the hanging drop vapor diffusion method, using the Nucleic Acid Miniscreen (Hampton Research) (24). Droplets containing 0.6 mM oligodeoxynucleotide, 5% 2-methyl-2,4-pentanediol (MPD), 20 mM sodium cacodylate (pH 6.0), 6 mM spermine tetrahydrochloride, 40 mM KCl, and 10 mM MgCl<sub>2</sub> were equilibrated against a reservoir of 35% MPD. Orthorhombic crystals in space group  $P2_12_12_1$  appeared within 7 days and grew to a



Table 1: Thermodynamic Profiles for the Folding of the Dodecamers<sup>a</sup>

[NaCl] (mM)	$T_M$ (°C)	$\Delta H_{cal}$ (kcal/mol)	$T\Delta S_{cal}$ (kcal/mol)	$\Delta G^{\circ}_{20}$ (kcal/mol)	$\Delta n_{Na^+}$ (mol/mol)
DDD					
10	33.3	-116	-109	-6.9	$-2.3 \pm 0.2$
100	57.7	-110	-94.1	-15	$-1.8 \pm 0.1$
DDD <sup>Z10</sup>					
10	35.7	-106	-99.9	-6.1	$-1.7 \pm 0.1$
100	50.0	-87.7	-77.0	-11	$-1.3 \pm 0.1$
DDD <sup>2+</sup>					
10	29.8	-68.0	-64.7	-3.3	$-1.5 \pm 0.1$
100	45.3	-60.0	-52.1	-7.9	$-1.2 \pm 0.1$
DDD <sup>2+Z10</sup>					
10	31.9	-74.3	-70.6	-3.7	$-1.2 \pm 0.1$
100	42.2	-70.2	-61.2	-9.0	$-1.1 \pm 0.1$

<sup>a</sup> All parameters were measured from UV ( $T_M$ ) and DSC melting curves in 10 mM sodium phosphate buffer (pH 7) and at the indicated NaCl concentrations. All  $T_M$ 's correspond to a strand concentration of 10 mM. The experimental uncertainties were as follows:  $\pm 0.5$  °C for  $T_M$ ,  $\pm 5\%$  for  $\Delta H_{cal}$ ,  $\pm 7\%$  for  $\Delta G^{\circ}_{20}$ , and  $\pm 5\%$  for  $T\Delta S_{cal}$ .

size of 0.1 mm  $\times$  0.1 mm  $\times$  0.3 mm over 2–3 days. A wide range of Mg<sup>2+</sup> and spermine concentrations and pH values were assayed. It was concluded that both modified duplexes showed a propensity to crystallize at low cationic strengths.

**Data Collection and Processing.** Diffraction data were collected on beamline 22-ID at SER-CAT, APS (Argonne National Laboratory, Argonne, IL), using a MAR-165 CCD detector at -160 °C. Data were scaled, integrated, and reduced with HKL2000 (25).

**Refinement.** Both structures were determined by the molecular replacement method using EPMR (26) with the DDD structure with PDB entry BDL084 (27) as a search model. The structures of DDD<sup>Z10</sup> and DDD<sup>2+Z10</sup> were initially refined with simulated annealing by using CNS (28), and subsequently by geometric constraint/maximum-likelihood and isotropic temperature factor refinement for individual atoms by using REFMAC 5.0/CCP4 (29, 30). In later refinement stages of both structures, solvent water molecules were added on the basis of Fourier  $2F_o - F_c$  sum and  $F_o - F_c$  difference electron density maps, and accepted on the basis of standard distance and B-factor criteria. X<sup>8</sup> in the DDD<sup>2+Z10</sup> structure and Z<sup>10</sup> in both structures were introduced into DNA duplexes on the basis of Fourier electron density maps; corresponding geometry/topology files were adapted, and anisotropic temperature factor refinement was carried out afterward. TURBO-FRODO (31) was used to display electron density maps, manually rebuild the duplex models, and add or delete water molecules. The geometries of the DNA duplexes were analyzed with CURVES (version 5.3) (32).

## RESULTS

**Unfolding Thermodynamics of Dodecamer Duplexes.** Heat capacity thermograms for the unfolding of each duplex are shown in panels A and B of Figure 1 at overall Na<sup>+</sup> concentrations of 16 and 116 mM, respectively. At the lower salt concentration, all curves exhibited biphasic transitions that shifted to monophasic transitions at higher temperatures with the increase in salt concentration. This was due to the stronger effect of salt on the first (duplex) transition, relative

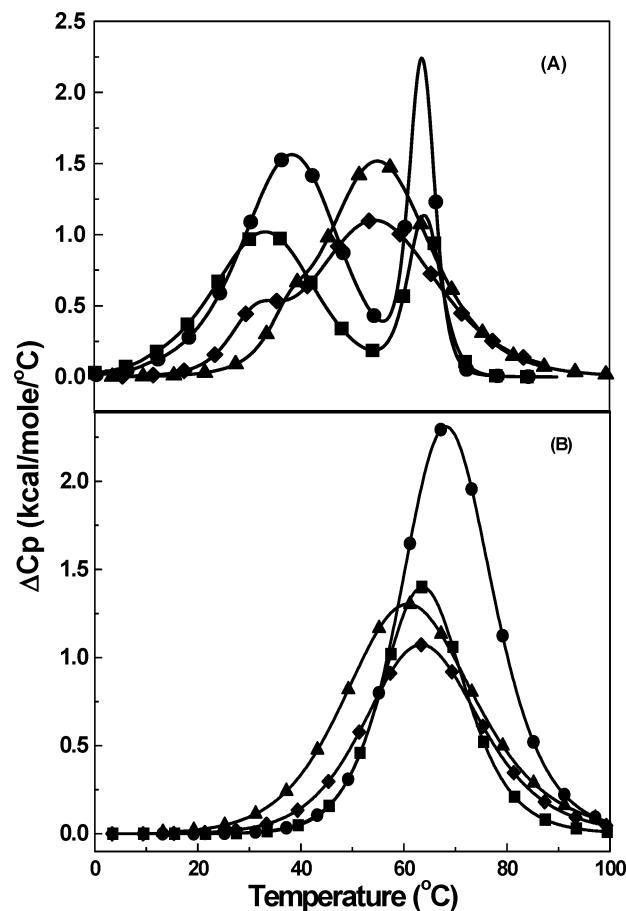


FIGURE 1: DSC curves of duplexes in 10 mM sodium phosphate buffer (pH 7) and total Na<sup>+</sup> concentrations of (A) 16 and (B) 116 mM: DDD (●), DDD<sup>2+</sup> (■), DDD<sup>Z10</sup> (▲), and DDD<sup>2+Z10</sup> (◆).

to the second (hairpin) transition, confirming the helix  $\rightarrow$  hairpin  $\rightarrow$  random coil transitions of each dodecamer duplex (33). Deconvolution analysis of differential scanning calorimetry (DSC) curves yielded the associated enthalpies; however, only the model-independent enthalpies of the duplex  $\rightarrow$  random coil transitions are reported in Table 1. At low salt concentrations, enthalpies in the following order were obtained: DDD (116 kcal/mol)  $>$  DDD<sup>Z10</sup>  $>$  DDD<sup>2+Z10</sup>  $>$  DDD<sup>2+</sup> (68 kcal/mol). The increase in salt concentration yielded lower enthalpy values but a similar trend: DDD (110 kcal/mol)  $>$  DDD<sup>Z10</sup>  $>$  DDD<sup>2+Z10</sup>  $>$  DDD<sup>2+</sup> (60 kcal/mol). Relative to DDD, the dT  $\rightarrow$  Z3dU substitution, the dG  $\rightarrow$  7-deaza-dG substitution, or both were destabilizing at low and high salt concentrations (34–36). The lowering of the enthalpies at higher salt concentrations suggested the presence of heat capacity effects. The heat capacities were in the following order: DDD ( $-0.8$  kcal K<sup>-1</sup> mol<sup>-1</sup>)  $\sim$  DDD<sup>Z10</sup>  $<$  DDD<sup>2+</sup>  $\sim$  DDD<sup>2+Z10</sup> ( $-0.3$  kcal K<sup>-1</sup> mol<sup>-1</sup>). Comparison of DDD with DDD<sup>2+</sup> and the second pair, DDD<sup>Z10</sup> with DDD<sup>2+Z10</sup>, suggested that the decrease in heat capacity was due to weakened hydrophobic interactions from the incorporation of the charged Z3dU chain in the helical states of these oligodeoxynucleotides.

**Thermodynamic Profiles for the Formation of Each Duplex.** Table 1 lists thermodynamic profiles for the formation of each duplex at 20 °C. The magnitude of the  $\Delta G^{\circ}_{20}$  terms indicated that each oligodeoxynucleotide formed a stable duplex. The formation of each duplex resulted from the partial compensation of favorable enthalpy and unfavor-

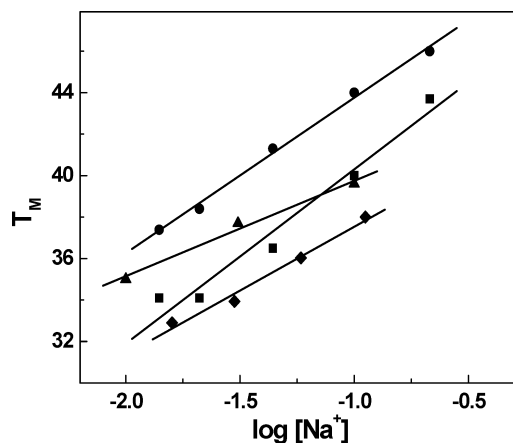


FIGURE 2:  $T_M$  dependencies on salt concentration of duplexes in 10 mM sodium phosphate buffer (pH 7) and a total strand concentration of 7  $\mu$ M. All  $T_M$  values correspond to the duplex  $\rightarrow$  random coil transition: DDD ( $\bullet$ ), DDD<sup>2+</sup> ( $\blacksquare$ ), DDD<sup>Z10</sup> ( $\blacktriangle$ ), and DDD<sup>2+Z10</sup> ( $\blacklozenge$ ).

able entropy contributions. Favorable enthalpy contributions included formation of base pairs and base-pair stacks, the immobilization of water around charges (electrostricted hydration), and the release of structural water around polar and nonpolar groups, while the unfavorable entropy term arose from contributions of the ordering of two strands, counterion condensation, and immobilization of water molecules. The free energy changes for duplex formation were in the following order: DDD (−6.9 kcal/mol) < DDD<sup>Z10</sup> < DDD<sup>2+Z10</sup> < DDD<sup>2+</sup> (−3.3 kcal/mol) at low salt concentrations and DDD (−15.0 kcal/mol) < DDD<sup>Z10</sup> < DDD<sup>2+Z10</sup> < DDD<sup>2+</sup> (−7.9 kcal/mol) at high salt concentrations. Relative to DDD, both dT  $\rightarrow$  Z3dU and dG  $\rightarrow$  7-deaza-dG substitutions were destabilizing at low and high salt concentrations.

**Differential Thermodynamic Binding of Counterions.** Increasing the salt concentration shifted the melting curves to higher temperatures, consistent with the expectation that the duplex states had higher charge density parameters. The linear  $T_M$  dependencies on salt concentration are shown in Figure 2. The slopes of these lines in conjunction with the  $\Delta H/RT_M^2$  terms allowed measurement of differential counterion binding. The  $\Delta n_{Na^+}$  values for the formation of each duplex (Table 1) were found in the following order: −2.3 (DDD) < −1.7 (DDD<sup>Z10</sup>) < −1.5 (DDD<sup>2+</sup>) < −1.2 (DDD<sup>2+Z10</sup>) in 10 mM NaCl. A similar trend was observed at 100 mM NaCl; however, the  $\Delta n_{Na^+}$  values (Table 1) were lower due to higher screening of the phosphates. The main effect was that the incorporation of Z3dU, 7-deaza-dG, or both into the dodecamer duplexes, assuming similar random coils at high temperatures, caused a weaker binding of counterions (35, 36).

**DNA Conformation in the Solid or Crystalline State. Crystal Packing.** The crystals of the modified DDD<sup>Z10</sup> and DDD<sup>2+Z10</sup> dodecamers were isomorphous with those of the native DDD dodecamer. The two strands were not symmetry-related, and the lattice interactions between the chemically modified DDD<sup>Z10</sup> and DDD<sup>2+Z10</sup> duplexes were preserved, as compared to those in the crystals of the native DDD dodecamer. Within the crystal, the DDD<sup>2+Z10</sup> duplexes were related by crystallographic 2<sub>1</sub> symmetries. They were arranged in a head-to-tail fashion and formed extended columns

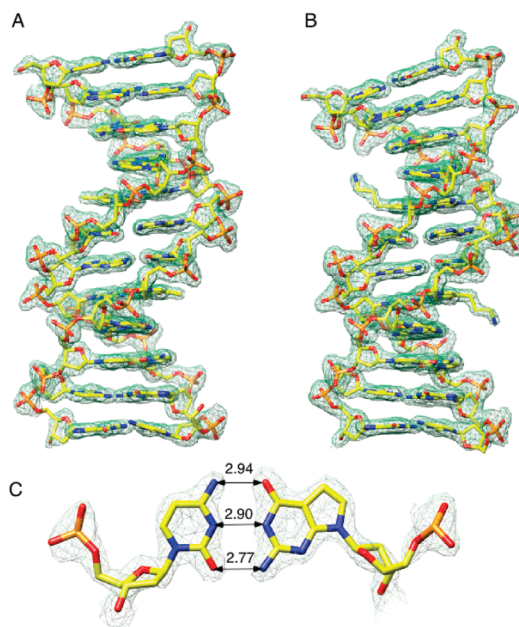


FIGURE 3: Sum electron density contoured at the 1.5 $\sigma$  level (green meshwork) around modified duplexes (A) DDD<sup>Z10</sup> and (B) DDD<sup>2+Z10</sup> and (C) the 7-deaza-dG•dC base pair with Watson-Crick geometry.

along the crystallographic  $c$ -axis. Their ends formed 2 bp overlaps, involving direct hydrogen bonds between adjacent guanines and resulting in base triples (Figure S1 of the Supporting Information). Each G•G pair was stabilized by two N<sup>2</sup>-H...N<sup>3</sup> hydrogen bonds. This packing contact is a common feature of orthorhombic  $P2_12_12_1$  crystals of Dickerson–Drew-type dodecamers (37, 38). Similar features were observed for the modified DDD<sup>Z10</sup> duplex (Figure S2 of the Supporting Information).

**Overall Structures.** The structures of both DDD<sup>Z10</sup> and DDD<sup>2+Z10</sup> were refined to a resolution of 1.6 Å. The quality of the final electron density maps in both instances was excellent (Figure 3). The rmsd values for bond distances were 0.006 Å in both instances and for angles 1.6° in both instances (Table 2). The high resolutions achieved for the structures allowed a detailed analysis of geometric and conformational changes in the duplexes as a consequence of the introduced modifications. In addition to a single duplex per crystallographic asymmetric unit, a total of 110 and 95 water molecules were placed in the case of the DDD<sup>Z10</sup> and DDD<sup>2+Z10</sup> structures, respectively (Table 2). Neither Mg<sup>2+</sup> nor spermine were located in the DDD<sup>Z10</sup> or DDD<sup>2+Z10</sup> dodecamers.

**Comparison of the DDD<sup>Z10</sup> and DDD Duplexes.** A superimposition of the DDD<sup>Z10</sup> and DDD (27) (PDB entry BDL084) duplexes is shown in Figure 4. Bending was evident at one end of the duplex in the case of the DDD duplex (27) compared to the DDD<sup>Z10</sup> duplex. In the native DDD duplex structure, there was an ordered Mg<sup>2+</sup> ion bound at the G<sup>2</sup>•C<sup>23</sup>  $\rightarrow$  C<sup>3</sup>•G<sup>22</sup> base-pair step (Figure 5). The cation contacted the N7 and O<sup>6</sup> atoms of G<sup>2</sup> and G<sup>22</sup> in the opposite strands via coordinated waters (39, 40). It also bridged the O2P oxygens of P6c and P7c of an adjacent molecule and stabilized the close interduplex contact between P2 and P7c (39–41). The bend at this end of the DDD duplex was associated with the coordination of a Mg<sup>2+</sup> ion. In the modified DDD<sup>Z10</sup> duplex, this ordered Mg<sup>2+</sup> was not

Table 2: Selected Crystal Data and Refinement Statistics

	DDD <sup>Z10</sup>	DDD <sup>2+Z10</sup>
(A) Crystal Data		
unit cell ( $\alpha$ , $\beta$ , and $\gamma$ angles are 90°)		
<i>a</i> (Å)	24.19	24.86
<i>b</i> (Å)	40.33	41.06
<i>c</i> (Å)	65.69	66.62
space group	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
(B) Data Collection		
temperature of data collection (°C)	−160.0	−160.0
total no. of reflections	164005	359861
no. of independent reflections	9994	9574
completeness (%)	96.6	93.4
maximum resolution (Å)	1.42	1.42
(C) Structure Refinement		
resolution range (Å)	34.36–1.60	34.94–1.60
no. of reflections used in refinement	7824	8368
no. of reflections used in the test set	471	497
rmsd of bonds from ideal (Å)	0.006	0.006
rmsd of angles from ideal (deg)	1.6	1.6
no. of DNA atoms	486	492
no. of water molecules	110	95
<i>R</i> -factor(test set) <sup>a,b</sup>	0.262	0.233
<i>R</i> -factor(work set) <sup>b</sup>	0.179	0.178
<i>R</i> -factor(work+test) <sup>b</sup>	0.186	0.180

<sup>a</sup> Calculated using 6% of the reflection data that were not used in the refinement. <sup>b</sup>  $R\text{-factor} = \frac{\sum |F_o(hkl)| - F(hkl)_o}{\sum |F_o(hkl)|}$ .

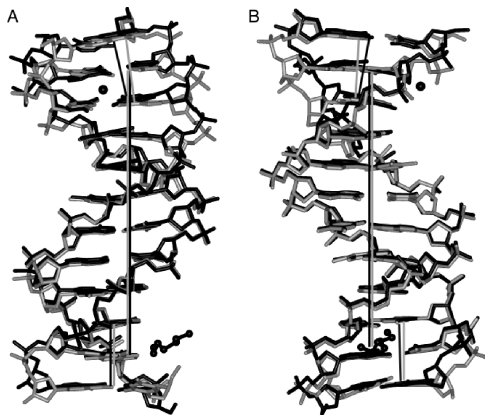


FIGURE 4: Superimposition of the DDD (PDB entry BDL084) (black) and DDD<sup>Z10</sup> (gray) duplexes. The view is into the (A) minor groove and (B) major groove, with respect to the central base pairs. Mg<sup>2+</sup> and a partial spermine molecule in the structure of the native DDD are shown in a ball-and-stick representation. The helix axes are based on the central hexamer duplexes and are coincident (straight gray and black lines for DDD<sup>Z10</sup> and DDD, respectively). In both duplexes, the helix axes defined by base pairs G<sup>10</sup>•C<sup>15</sup>, C<sup>11</sup>•G<sup>14</sup>, and G<sup>12</sup>•C<sup>13</sup> at one end are parallel with those of the central hexamers. Conversely, the helix axis defined by base pairs C<sup>1</sup>•G<sup>24</sup>, G<sup>2</sup>•C<sup>23</sup>, and C<sup>3</sup>•G<sup>22</sup> at the other end is inclined more in the native DDD than in DDD<sup>Z10</sup> relative to the respective helix axis.

observed. The degree of the bend where the Mg<sup>2+</sup> was located in the DDD dodecamer was reduced (Figure 4 and Figure S3 of the Supporting Information), and disorder was observed for the phosphate group between C<sup>1</sup> and G<sup>2</sup>.

The substitution of 7-deaza-dG was accompanied by changes in hydration in the major groove (Figure 5). Two waters contacting G<sup>2</sup> and G<sup>4</sup> O<sup>6</sup> and N<sup>7</sup> in the DDD<sup>Z10</sup> duplex maintained their positions as in the DDD duplex, while two waters close to Z<sup>22</sup> O<sup>6</sup> C<sup>7</sup> reorganized versus the natural dodecamer. The phosphodiester linkage between C<sup>1</sup> and G<sup>2</sup> flipped to the major groove with the disappearance of the Mg<sup>2+</sup> ion at the G<sup>2</sup>•C<sup>3</sup> step (Figure 5). Changes were evident

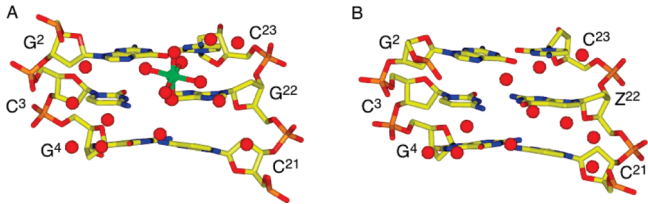


FIGURE 5: (A) Unmodified DDD duplex (PDB entry BDL084), showing the contact of Mg<sup>2+</sup> with O<sup>6</sup> and N<sup>7</sup> of G<sup>2</sup> and G<sup>22</sup> through coordinated waters at the G<sup>2</sup>•C<sup>23</sup> → C<sup>3</sup>•G<sup>22</sup> base-pair step. (B) DDD<sup>Z10</sup> duplex, showing the reorganization of waters and the absence of the Mg<sup>2+</sup> ion. The Mg<sup>2+</sup> ion is colored green in panel A, and the waters are colored red.

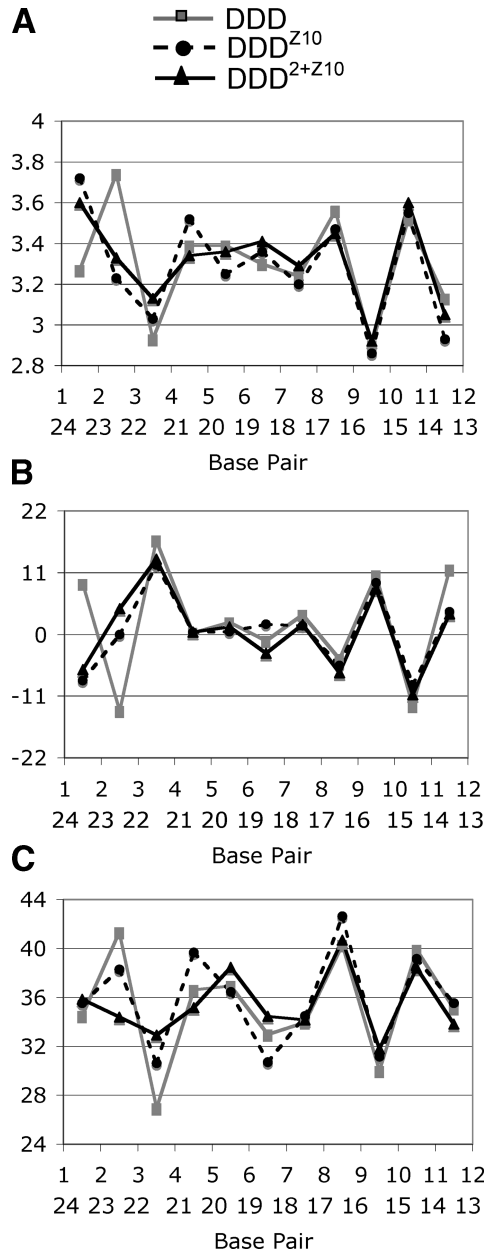


FIGURE 6: Interbase pair parameters: (A) helical rise, (B) roll, and (C) twist for the DDD<sup>Z10</sup>, DDD<sup>2+Z20</sup>, and DDD (PDB entry BDL084) duplexes.

in the rise and roll and, to a smaller degree, in the twist parameters (Figure 6). In contrast, the other end of the duplex was not perturbed (Figure 4 and Figure S3 of the Supporting Information). In the unmodified DDD duplex, one of the internal nitrogen atoms of a spermine molecule was bound



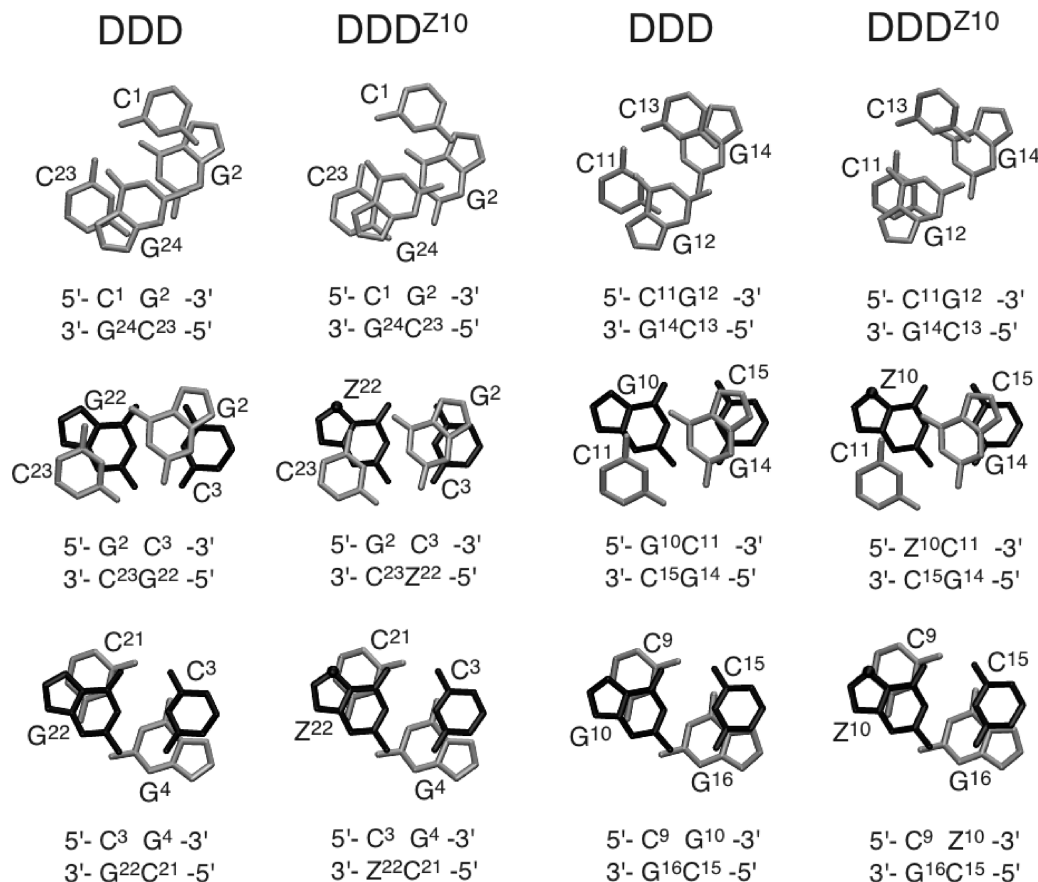


FIGURE 7: Base stacking interactions in the CGCG portions at both ends of the DDD (PDB entry BDL084) and DDD<sup>Z10</sup> duplexes. The conformations of the tetramer duplexes in the structures of the DDD and DDD<sup>Z10</sup> duplexes indicate adjustments in the relative orientations of the modified and flanking base pairs in the latter. The modified base pair in DDD<sup>Z10</sup> and the respective base pair in the unmodified DDD duplex are colored black, and atom C7 of 7-deaza-dG is highlighted with a black ball.

to G<sup>10</sup> N<sup>7</sup> (3.4 Å) (Figure 4) (27). The 7-deaza-dG modification at G<sup>10</sup> in the DDD<sup>Z10</sup> or DDD<sup>2+Z10</sup> duplex made this H-bond impossible, and indeed, no ordered spermine was present.

**Structural Perturbations from the 7-Deaza-dG Modification.** The C<sup>3</sup>•Z<sup>22</sup> and C<sup>15</sup>•Z<sup>10</sup> base pairs were maintained in the DDD<sup>Z10</sup> or DDD<sup>2+Z10</sup> duplexes despite dG being substituted with 7-deaza-dG (Figure 3). In the case of the C<sup>3</sup>•Z<sup>22</sup> base pair, the distances from N<sup>4</sup> to O<sup>6</sup>, from N3 to N1, and from O<sup>2</sup> to N<sup>2</sup> were 2.9, 2.9, and 2.8 Å, respectively. For the Z<sup>10</sup>•C<sup>15</sup> base pair, the distances from N<sup>4</sup> to O<sup>6</sup>, from N3 to N1, and from O<sup>2</sup> to N<sup>2</sup> were 2.9, 2.9, and 2.7 Å, respectively. These distances differed slightly from those observed for unmodified duplex C<sup>3</sup>•G<sup>22</sup> (2.8, 2.9, and 2.7 Å from N<sup>4</sup> to O<sup>6</sup>, from N3 to N1, and from O<sup>2</sup> to N<sup>2</sup>, respectively) and G<sup>10</sup>•C<sup>15</sup> (3.1, 2.9, and 2.7 Å from N<sup>4</sup> to O<sup>6</sup>, from N3 to N1, and from O<sup>2</sup> to N<sup>2</sup>, respectively). In terms of the base stacking interactions between the C<sup>1</sup>•G<sup>24</sup> and G<sup>4</sup>•C<sup>21</sup> pairs in the DDD<sup>Z10</sup> duplex as compared to those in the DDD duplex (27), the base overlaps at the C<sup>3</sup>•Z<sup>22</sup> → G<sup>4</sup>•C<sup>21</sup> steps were similar in the two duplexes (Figure 8B). However, the G<sup>2</sup>•C<sup>23</sup> → C<sup>3</sup>•Z<sup>22</sup> and C<sup>1</sup>•G<sup>24</sup> → G<sup>2</sup>•C<sup>23</sup> steps were slightly perturbed. Similar features were observed for the base pairs at the other end of the duplex (Figure 7).

**Comparison of the DDD<sup>Z10</sup> and DDD<sup>2+Z10</sup> Duplexes.** A superimposition of the modified DDD<sup>Z10</sup> and DDD<sup>2+Z10</sup> duplexes is shown in Figure 8B. While neither the DDD<sup>Z10</sup> duplex nor the DDD<sup>2+Z10</sup> duplex exhibited a bound Mg<sup>2+</sup>, an increased axial bending of 20° was observed at one end

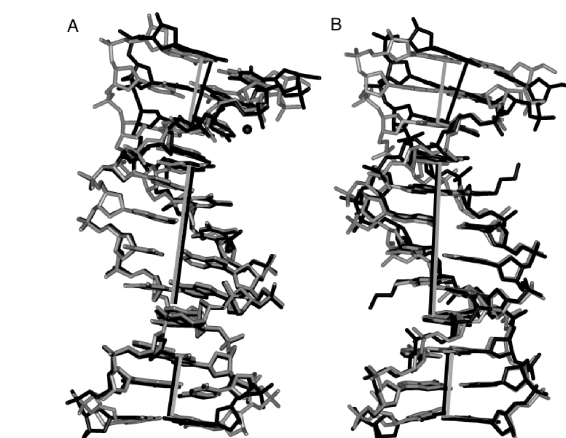


FIGURE 8: (A) Superimposition of the DDD (PDB entry BDL084, in black) and DDD<sup>Z10</sup> (gray) duplexes. (B) Superimposition of the DDD<sup>2+Z10</sup> (black) and DDD<sup>Z10</sup> (gray) duplexes. The view is into the major groove. The Mg<sup>2+</sup> ion in the structure of the native DDD duplex is shown in a ball-and-stick representation. The helix axes based on the central hexamer duplexes are coincident (black line for DDD and DDD<sup>2+Z10</sup> and gray line for DDD<sup>Z10</sup>). For both, the helix axes defined by base pairs G<sup>10</sup>•C<sup>15</sup>, C<sup>11</sup>•G<sup>14</sup>, and G<sup>12</sup>•C<sup>13</sup> are parallel with those of the central hexamers. The helix axes defined by base pairs C<sup>1</sup>•G<sup>24</sup>, G<sup>2</sup>•C<sup>23</sup>, and C<sup>3</sup>•G<sup>22</sup> in either DDD<sup>2+Z10</sup> or DDD are bent as compared to that the DDD<sup>Z10</sup> duplex.

of the duplex for DDD<sup>2+Z10</sup>, as compared to an axial bend of 10° for the DDD<sup>Z10</sup> duplex (Figure 8B and Figure S3 of the Supporting Information). The conformational change was accompanied by alterations in the helical twist (Figure 6C).

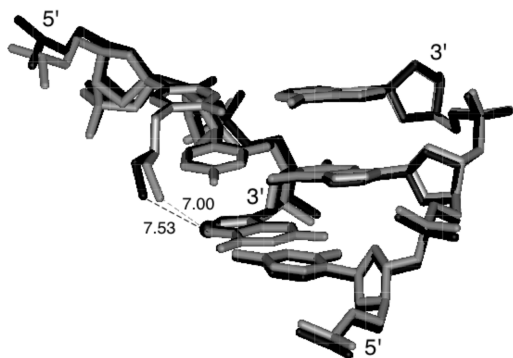


FIGURE 9: View of the superimposed  $\text{DDD}^{2+Z10}$  (black) and  $\text{DDD}^{4+}$  (gray) (PDB entry 1Z5T) duplexes. The distances between the tethered Z3dU cations at  $\text{X}^8$  and  $\text{G/Z}^{10} \text{O}^6$  in the two duplexes are in angstroms. Atom C7 in 7-deaza-dG is highlighted as a black sphere.

The other end of the duplex was not perturbed. In both instances, the asymmetric bending was attributed to a combination of cation binding and lattice interactions (41, 42).

**Orientation and Conformation of the Tethered Cationic Side Chains.** Both Z3dU modifications could be identified in sum and difference electron density maps (Figure 3). Both tethers extended toward the 3'-direction in the major groove. The two side chains at  $\text{X}^8$  and  $\text{X}^{20}$  adopted different conformations. Compared to that of the  $\text{DDD}^{Z10}$  duplex, a  $20^\circ$  bend was induced by the tethered cation at  $\text{X}^{20}$  in  $\text{DDD}^{2+Z10}$  at one end of the duplex. The effect of this tethered cation was similar to that of the  $\text{Mg}^{2+}$  ion bound in the major groove close to  $\text{G}^{22} \text{O}^6$  in the unmodified DDD duplex (27) (Figure 8). The conformation of the  $\text{C}_\alpha\text{--C}_\beta\text{--C}_\gamma\text{--N}^+$  torsion angle for the tethered cationic side chain at  $\text{X}^{20}$  was  $177^\circ$ , which was close to the ideal *trans* conformation. At the opposite end of the duplex, the tethered cation of  $\text{X}^8$  did not induce a conformational change. The conformation of the  $\text{C}_\alpha\text{--C}_\beta\text{--C}_\gamma\text{--N}^+$  torsion angle for the tethered cationic side chain at  $\text{X}^8$  was  $200^\circ$ , somewhat perturbed from the ideal *trans* conformation.

A superimposition of the  $\text{DDD}^{2+Z10}$  and  $\text{DDD}^{4+}$  (19) (PDB entry 1Z5T) duplexes is shown in Figure 9. For the 7-deaza-dG-containing duplex, there was a  $0.5 \text{ \AA}$  increase in distance between the tethered Z3dU amine at  $\text{X}^8$  and the  $\text{Z}^{10} \text{O}^6$  atom. There were two waters observed at  $<5 \text{ \AA}$  with respect to the Z3dU amine at the bent end of the  $\text{DDD}^{2+Z10}$  and  $\text{DDD}^{4+}$  duplexes (Figure 10B,D). In the  $\text{DDD}^{4+}$  duplex, one water molecule mediated the interaction between the Z3dU amine and  $\text{G}^{10} \text{O}^6$  (Figure 10D), with distances of  $4.0 \text{ \AA}$  between  $\text{H}_2\text{O O}$  and  $\text{G}^{10} \text{O}^6$  and  $3.9 \text{ \AA}$  between  $\text{H}_2\text{O O}$  and Z3dU  $\text{N}^+$ . The 7-deaza-dG modification in  $\text{DDD}^{2+Z10}$  resulted in the relocation of the bridging water molecule to a position further removed from the  $\text{Z}^{10} \text{O}^6$  atom (Figure 10B,D). The distance between  $\text{H}_2\text{O O}$  and  $\text{Z}^{10} \text{O}^6$  was lengthened to  $5.3 \text{ \AA}$ . This reflected a weakened interaction between the Z3dU amine and  $\text{G}^{10} \text{O}^6$ . The other water molecule was farther from the Z3dU amine, with distances of  $4.7$  and  $4.1 \text{ \AA}$  in the  $\text{DDD}^{2+Z10}$  and  $\text{DDD}^{4+}$  duplexes, respectively (Figure 10B,D). This water molecule mediated the interaction between the Z3dU amine and  $\text{G}^{16} \text{O}^6$  in the  $\text{DDD}^{2+Z10}$  and  $\text{DDD}^{4+}$  modified duplexes.

The positions and conformations of the tethered cationic side chains were affected by the water organization in the major groove. The  $\text{C}_\alpha\text{--C}_\beta\text{--C}_\gamma\text{--N}^+$  torsion angles of

the Z3dU moieties were  $200^\circ$  and  $207^\circ$  for  $\text{DDD}^{2+Z10}$  and  $\text{DDD}^{4+}$ , respectively. At the other end of the duplexes, the bridging water molecules were also relocated due to the 7-deaza-dG modifications at  $\text{G}^{22}$  (Figure 10A,C). Thus, the  $3.6 \text{ \AA}$  distance between the Z3dU amine and  $\text{Z}^{22} \text{O}^6$  was greater than the distance of  $3.3 \text{ \AA}$  between the Z3dU amine and  $\text{G}^{22} \text{O}^6$ . The closest water molecule to the  $\text{X}^{20}$ -tethered amine made a  $6.8 \text{ \AA}$  bridge to  $\text{Z}^{22} \text{O}^6$ , as compared to a  $7.5 \text{ \AA}$  bridge from the  $\text{X}^8$ -tethered amine to  $\text{Z}^{10} \text{O}^6$ .

## DISCUSSION

The formation of the DNA double helix creates an electrostatic environment that results in the sequence-dependent binding of inorganic cations and reactive electrophiles in the major and minor grooves (17, 40, 43–48), in addition to the array of cations that surround the polyanionic phosphate backbone (49–53). The effect of the environment on DNA conformation is dramatic. While quantitative and qualitative changes in mobile cations can globally alter DNA structure (54–57), DNA affinity binding molecules, including proteins and organic ligands, regioselectively modulate the electrostatic environment of DNA and concomitantly cause localized changes in conformation (58–62).

**A Bridging Water between the Tethered Z3dU Amine and  $\text{Z}^{10} \text{O}^6$ .** The observation of bridging water molecules between the tethered Z3dU amines and the dG  $\text{O}^6$  atoms of base pairs  $\text{C}^3\cdot\text{G}^{22}$  and  $\text{C}^3\cdot\text{Z}^{22}$  and base pairs  $\text{G}^{10}\cdot\text{C}^{15}$  and  $\text{Z}^{10}\cdot\text{C}^{15}$ , in the crystallographic structures of the  $\text{DDD}^{4+}$  (19) and  $\text{DDD}^{2+Z10}$  duplexes, respectively, is significant (Figure 10). These bridging water molecules play a structural role in orienting the Z3dU moiety toward the 3'-direction in the major groove. The 7-deaza-dG modification in  $\text{DDD}^{2+Z10}$  results in the relocation of the bridging water molecule and a lengthening of the distance between  $\text{H}_2\text{O O}$  and  $\text{Z}^{10} \text{O}^6$  to  $5.3 \text{ \AA}$ . This is attributed to a weakened interaction between the Z3dU amine and  $\text{G}^{10} \text{O}^6$ . Likewise, the bridging water molecule that mediates the interaction between the Z3dU amine and  $\text{G}^{16} \text{O}^6$  is farther from the Z3dU amine, with distances of  $4.7$  and  $4.1 \text{ \AA}$  in the  $\text{DDD}^{2+Z10}$  and  $\text{DDD}^{4+}$  duplexes, respectively (Figure 10).

**Relevance to  $\text{N} + 2$  ICL Cross-Link Formation by Nitrogen Mustards.** The ability of the bridging waters to effectively lengthen the  $4 \text{ \AA}$  Z3dU tether so that the combination of the Z3dU tether and the bridging water can span the distance to the next-nearest-neighbor base pairs in the 3'-direction relative to the Z3dU-modified nucleotides, without requiring significant bending of the DNA duplex, suggests that formation of  $\text{N} + 2$  ICLs by nitrogen mustards (3–5) might involve a similar mechanism. A similar orientation of the monofunctional N7-dG-tethered aziridinium ion in the major groove at 5'-GNC-3' sequences, involving a water bridge, might also be accommodated with minimal helical bending of the DNA duplex, thus kinetically favoring formation of the bent  $\text{N} + 2$  transition state geometry. These data probably also explain why previous NMR solution data for tethered Z3dU amines were equivocal with respect to DNA bending (18); on average, the solution structure remains linear due to the presence of the bridging water molecules between the cationic amines and dG  $\text{O}^6$  of the next-nearest-neighbor base pair. Exchange of the bridging water molecule



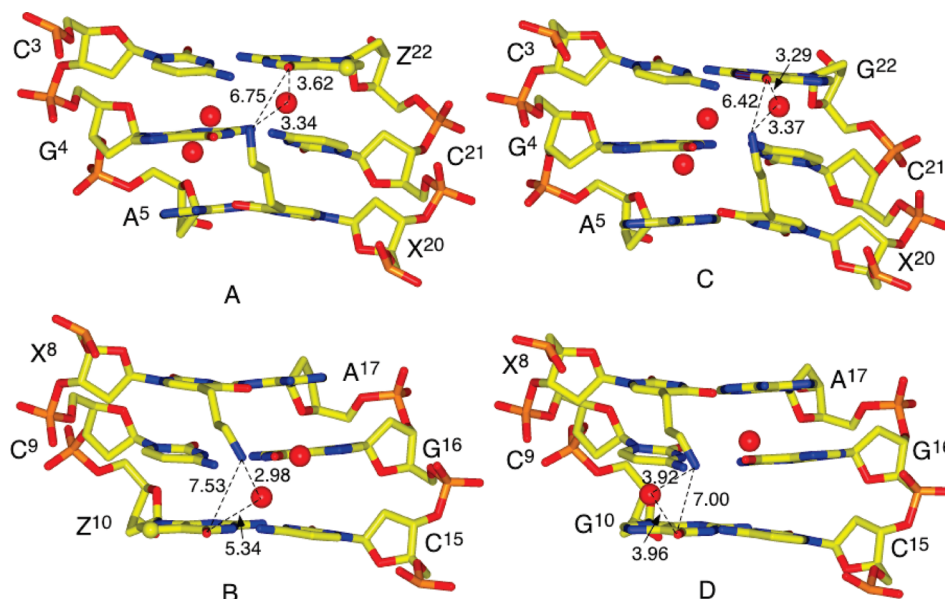


FIGURE 10: Interactions among the Z3dU amine, a bridging water molecule, and G/Z O<sup>6</sup> in the major groove in the (A and B) DDD<sup>2+Z10</sup> and (C and D) DDD<sup>4+</sup> (PDB entry 1Z5T) duplexes.

Table 3: Differential Thermodynamic Profiles for Pairs of Dodecamer Duplexes

[NaCl] (mM)	$\Delta\Delta H_{\text{cal}}$ (kcal/mol)	$\Delta(T\Delta S_{\text{cal}})$ (kcal/mol)	$\Delta\Delta G^{\circ}_{20}$ (kcal/mol)	$\Delta\Delta n_{\text{Na}^+}$ (mol of Na <sup>+</sup> /mol)
Substitution of dG with 7-Deaza-dG in DDD (DDD <sup>Z10</sup> minus DDD)				
10	10.0	9.2	0.8	0.6
100	21.8	17	4.8	0.5
Incorporation of Z3dU into DDD (DDD <sup>2+</sup> minus DDD)				
10	48.0	44.4	3.6	0.8
100	49.5	41.9	7.6	0.6
Incorporation of Z3dU into DDD <sup>Z10</sup> (DDD <sup>2+Z10</sup> minus DDD <sup>Z10</sup> )				
10	31.7	29.3	2.4	0.5
100	17.5	15.8	1.7	0.2

then releases the tethered Z3dU moiety, facilitating formation of a high-energy bent structure, in which the Z3dU amine is positioned adjacent to dG N7 and O<sup>6</sup> of the next-nearest-neighbor base pair. This structure models the intermediate leading to N + 2 ICL formation by nitrogen mustards. Alternatively, the dislocation of the bridging water releases the Z3dU from the floor of the major groove.

**Thermodynamic Consequences of Positioning Z3dU in the Major Groove.** The differential thermodynamic profiles for the incorporation of the tethered Z3dU cation in the major groove of DDD (last four rows of Table 3) indicate that inclusion of two Z3dU chains in the DDD duplex is destabilizing ( $\Delta\Delta G = 5.6$  kcal/mol), resulting from the compensation of an unfavorable enthalpy contribution ( $\Delta\Delta H = 48.8$  kcal/mol) with a favorable entropy contribution [ $\Delta(T\Delta S) = 43.2$  kcal/mol (63)] and average counterion release of 0.7 mol of Na<sup>+</sup>/mol of duplex. In comparison, the incorporation of two Z3dU chains in DDD<sup>Z10</sup> yielded lower magnitudes for each parameter:  $\Delta\Delta G = 2.1$  kcal/mol,  $\Delta\Delta H = 24.6$  kcal/mol,  $\Delta(T\Delta S) = 22.5$  kcal/mol, and  $\Delta\Delta n_{\text{Na}^+} = 0.35$  mol of Na<sup>+</sup>/mol of duplex. The unfavorable enthalpy terms explain differences in both base stacking and hydration contributions between each pair of duplexes, while the

favorable entropy terms are consistent with the release of counterions. The lower magnitudes of the latter profile are consistent with the perturbations caused by the inclusion of two 7-deaza-dG bases.

The placement of the Z3dU side chains in the major groove of DDD or DDD<sup>Z10</sup> reduces the level of base-pair stacking interactions and reorganizes both counterions and water molecules. These effects can be explained by the fact that the physical presence of the tethered propyl chain displaces counterions and waters in the major groove proximal to the site of attachment, while the positively charged amine indirectly neutralizes negative phosphate charges (17, 35). For instance, if it was assumed that the incorporation of two amino propyl chains effectively neutralized two negative phosphate charges of the 22 formal charges of DDD, then the average residual charge per phosphate predicted from Manning condensation theory (49) would correspond to a reduction in the level of counterion binding of 0.5 counterions for DDD<sup>2+</sup>, which is in good agreement with the experimental  $\Delta\Delta n_{\text{Na}^+}$  value of 0.8. A smaller release of counterions,  $\sim 0.3$  mol of Na<sup>+</sup>/mol of duplex, was obtained for the DDD<sup>Z10</sup> duplex, confirming the weaker binding of counterions by the duplex state of DDD<sup>Z10</sup>. This is consistent with the dG  $\rightarrow$  7-deaza-dG substitution into DDD at G<sup>10</sup> eliminating this ion binding site (45). The overall thermodynamics and structural data seem to suggest that it is the electrostatic attraction of the electron dense region at the edge of G<sup>10</sup> that propels the tethered aminopropyl chain toward the 3'-direction rather than out into solvent (18, 19, 36).

**Structural and Thermodynamic Consequences of the 7-Deaza-dG Modification.** The rationale for incorporation of 7-deaza-dG was to alter the electrostatics of the major groove at base pairs C<sup>3</sup>•Z<sup>22</sup> and Z<sup>10</sup>•C<sup>15</sup>. This incorporation of 7-deaza-dG at two sites in DDD facilitated these crystallographic studies. Thus, the results presented here also provide the first crystallographic characterization of the 7-deaza-dG modification in duplex DNA. It results in minimal structural perturbation of the Dickerson dodecamer, with Watson–Crick pairing being maintained at base pairs

$C^3 \cdot Z^{22}$  and  $Z^{10} \cdot C^{15}$ . The 7-deaza-dG modification does alter the electrostatics at the site typically occupied by a  $Mg^{2+}$  hexahydrate ion in unmodified DDD crystal lattices (19) and removes a proton acceptor from the major groove. Neither  $Mg^{2+}$  nor spermine is located in the crystal structure of the  $DDD^{Z^{10}}$  dodecamer, as compared to the native DDD dodecamer (Figure 5). The contact of the bound  $Mg^{2+}$  ion with O<sup>6</sup> and N7 of G<sup>2</sup> and G<sup>22</sup> through coordinated waters at the  $G^2 \cdot C^{23} \rightarrow C^3 \cdot G^{22}$  and  $C^{11} \cdot G^{14} \rightarrow G^{10} \cdot C^{15}$  steps may stabilize these base pairs with respect to breathing. NMR studies on the  $DDD^{Z^{10}}$  duplex revealed the rapid exchange of the Watson–Crick hydrogen-bonded imino protons at the 5′-neighboring base pairs relative to the 7-deaza-dG-modified bases, the  $G^2 \cdot C^{23}$  and  $C^{11} \cdot G^{14}$  base pairs (36). This suggests that in solution, the 7-deaza-dG modification increases the level of breathing at the 5′-neighbor base pair, related to the role of major groove cations in stabilizing DNA structure. The reduction in stability ( $\Delta\Delta G \sim 2$  kcal/mol) due to the 7-deaza-dG substitution is dominated by an unfavorable  $\Delta\Delta H_{cal}$  term, which suggests changes in stacking and/or hydration in the 7-deaza-dG-containing duplexes. The unfavorable  $\Delta\Delta H$  is compensated, in part, by a favorable  $\Delta(T\Delta S_{cal})$ . This is consistent with a reduced level of release of salts and H<sub>2</sub>O from the 7-deaza-dG-substituted duplexes and chemical footprinting showing enhanced accessibility of the  $G^2 \cdot C^{23}$  and  $C^{11} \cdot G^{14}$  base pairs (36).

**Summary.** These structural and thermodynamic studies suggest an electrostatic mechanism underlying the formation of interstrand N+2 cross-links by the nitrogen mustards, e.g., melphalan and mechlorethamine. The tethered Z3dU cations extend within the major groove in the 3′-direction, toward conserved  $Mg^{2+}$  binding sites located adjacent to N+2 base pairs  $C^3 \cdot Z^{22}$  and  $Z^{10} \cdot C^{15}$ . Bridging waters located between the tethered cationic amines and either  $Z^{10}$  or  $Z^{22}$  O<sup>6</sup> stabilize the tethered Z3dU cations and allow the Z3dU amines to interact with the N+2 base pairs without DNA bending. On the basis of the results, we envision a model in which the presence of tethered N7-dG aziridinium ions, which are the active species involved in formation of interstrand 5′-GNC-3′ cross-links by nitrogen mustards, modifies the electrostatics of the major groove and positions the aziridinium ions proximate to the major groove edge of the N+2 C•G base pair, facilitating interstrand cross-linking.

## SUPPORTING INFORMATION AVAILABLE

Hydrogen bonding interactions mediating direct contacts between neighboring DDD duplexes in the crystal lattice for duplex  $DDD^{Z^{10}}$  (Figures S1 and S2), comparison of the global axis curvature (Figure S3), and comparison of backbone and glycosyl torsion angles in the crystal structures of DDD,  $DDD^{Z^{10}}$ , and  $DDD^{Z^{2+Z^{10}}}$  duplexes (Figure S4). This material is available free of charge via the Internet at <http://pubs.acs.org>.

## REFERENCES

- Colvin, M. (1980) The alkylating agents, in *Pharmacological Principles of Cancer Treatment* pp 276–308, W. B. Saunders, Philadelphia.
- Brookes, P., and Lawley, P. D. (1961) The reaction of mono and di-functional alkylating agents with nucleic acids. *Biochem. J.* 80, 496–503.
- Ojwang, J. O., Grueneberg, D. A., and Loechler, E. L. (1989) Synthesis of a duplex oligonucleotide containing a nitrogen mustard interstrand DNA–DNA cross-link. *Cancer Res.* 49, 6529–6537.
- Grueneberg, D. A., Ojwang, J. O., Benasutti, M., Hartman, S., and Loechler, E. L. (1991) Construction of a human shuttle vector containing a single nitrogen mustard interstrand, DNA–DNA cross-link at a unique plasmid location. *Cancer Res.* 51, 2268–2272.
- Millard, J. T., Raucher, S., and Hopkins, P. B. (1990) Mechlorethamine cross-links deoxyguanosine residues at 5′-GNC sequences in duplex DNA fragments. *J. Am. Chem. Soc.* 112, 2459–2460.
- Rink, S. M., Solomon, M. S., Taylor, M. J., Rajur, S. B., McLaughlin, L. W., and Hopkins, P. B. (1993) Covalent structure of a nitrogen mustard-induced DNA interstrand cross-link. *J. Am. Chem. Soc.* 115, 2551–2557.
- Rink, S. M., and Hopkins, P. B. (1995) A mechlorethamine-induced DNA interstrand cross-link bends duplex DNA. *Biochemistry* 34, 1439–1445.
- Arnott, S., and Hukins, D. W. L. (1972) Optimised parameters for A-DNA and B-DNA. *Biochem. Biophys. Res. Commun.* 47, 1504–1509.
- Dong, Q., Barsky, D., Colvin, M. E., Melius, C. F., Ludeman, S. M., Moravik, J. F., Colvin, O. M., Bigner, D. D., Modrich, P., and Friedman, H. S. (1995) A structural basis for a phosphoramidate mustard-induced DNA interstrand cross-link at 5′-d(GAC). *Proc. Natl. Acad. Sci. U.S.A.* 92, 12170–12174.
- Bauer, G. B., and Povirk, L. F. (1997) Specificity and kinetics of interstrand and intrastrand bifunctional alkylation by nitrogen mustards at a G-G-C sequence. *Nucleic Acids Res.* 25, 1211–1218.
- Gold, B. (2002) Effect of cationic charge localization on DNA structure. *Biopolymers* 65, 173–179.
- Liang, G., Encell, L., Nelson, M. G., Switzer, C., and Gold, B. (1995) The role of electrostatics in the sequence selective reaction of charged alkylating agents with DNA. *J. Am. Chem. Soc.* 117, 10135–10136.
- Dande, P., Liang, G., Chen, F. X., Roberts, C., Nelson, M. G., Hashimoto, H., Switzer, C., and Gold, B. (1997) Regioselective effect of zwitterionic DNA substitutions on DNA alkylation: Evidence for a strong side chain orientational preference. *Biochemistry* 36, 6024–6032.
- Gold, B., Marky, L. M., Stone, M. P., and Williams, L. D. (2006) A review of the role of the sequence-dependent electrostatic landscape in DNA alkylation patterns. *Chem. Res. Toxicol.* 19, 1402–1414.
- Strauss, J. K., Prakash, T. P., Roberts, C., Switzer, C., and Maher, L. J. (1996) DNA bending by a phantom protein. *Chem. Biol.* 3, 671–678.
- Strauss, J. K., Roberts, C., Nelson, M. G., Switzer, C., and Maher, L. J. (1996) DNA bending by hexamethylene-tethered ammonium ions. *Proc. Natl. Acad. Sci. U.S.A.* 93, 9515–9520.
- Rouzina, I., and Bloomfield, V. A. (1998) DNA bending by small, mobile multivalent cations. *Biophys. J.* 74, 3152–3164.
- Li, Z., Huang, L., Dande, P., Gold, B., and Stone, M. P. (2002) Structure of a tethered cationic 3-aminopropyl chain incorporated into an oligodeoxynucleotide: Evidence for 3′-orientation in the major groove accompanied by DNA bending. *J. Am. Chem. Soc.* 124, 8553–8560.
- Moulaei, T., Maehigashi, T., Lountos, G. T., Komeda, S., Watkins, D., Stone, M. P., Marky, L. A., Li, J. S., Gold, B., and Williams, L. D. (2005) Structure of B-DNA with cations tethered in the major groove. *Biochemistry* 44, 7458–7468.
- Heystek, L. E., Zhou, H. Q., Dande, P., and Gold, B. (1998) Control over the localization of positive charge in DNA: The effect on duplex DNA and RNA stability. *J. Am. Chem. Soc.* 120, 12165–12166.
- Cavaluzzi, M. J., and Borer, P. N. (2004) Revised UV extinction coefficients for nucleoside-5′-monophosphates and unpaired DNA and RNA. *Nucleic Acids Res.* 32, e13.
- Marky, L. A., and Breslauer, K. J. (1987) Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers* 26, 1601–1620.
- Kaushik, M., Suehl, N., and Marky, L. A. (2007) Calorimetric unfolding of the bimolecular and i-motif complexes of the human telomere complementary strand, d(C<sub>3</sub>TA<sub>2</sub>)<sub>4</sub>. *Biophys. Chem.* 126, 154–164.
- Berger, I., Kang, C. H., Sinha, N., Wolters, M., and Rich, A. (1996) A highly efficient 24-condition matrix for the crystallization of nucleic acid fragments. *Acta Crystallogr. D52*, 465–468.
- Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Acta Crystallogr. A276*, 307–326.

26. Kissinger, C. R., Gehlhaar, D. K., and Fogel, D. B. (1999) Rapid automated molecular replacement by evolutionary search. *Acta Crystallogr. D* **55**, 484–491.
27. Shui, X., McFail-Isom, L., Hu, G. G., and Williams, L. D. (1998) The B-DNA dodecamer at high resolution reveals a spine of water on sodium. *Biochemistry* **37**, 8341–8355.
28. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921.
29. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255.
30. Collaborative Computational Project Number 4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763.
31. Cambillau, C., and Roussel, A. (1997) TURBO FRODO, version OpenGL.1, Université Aix-Marseille II, Marseille, France.
32. Ravishanker, G., Swaminathan, S., Beveridge, D. L., Lavery, R., and Sklenar, H. (1989) Conformational and helicoidal analysis of 30 ps of molecular dynamics on the d(CGCGAATTCGCG) double helix: 'Curves', dials, and windows. *J. Biomol. Struct. Dyn.* **6**, 669–699.
33. Marky, L. A., Blumenfeld, K. S., Kozlowski, S., and Breslauer, K. J. (1983) Salt-dependent conformational transitions in the self-complementary deoxydodecanucleotide d(CGCAATTCGCG): Evidence for hairpin formation. *Biopolymers* **22**, 1247–1257.
34. Soto, A. M., Kankia, B. I., Dande, P., Gold, B., and Marky, L. A. (2002) Thermodynamic and hydration effects for the incorporation of a cationic 3-aminopropyl chain into DNA. *Nucleic Acids Res.* **30**, 3171–3180.
35. Shikiya, R., Li, J. S., Gold, B., and Marky, L. A. (2005) Incorporation of cationic chains in the Dickerson-Drew dodecamer: Correlation of energetics, structure, and ion and water binding. *Biochemistry* **44**, 12582–12588.
36. Ganguly, M., Wang, F., Kaushik, M., Stone, M. P., Marky, L. A., and Gold, B. (2007) A study of 7-deaza-2'-deoxyguanosine 2'-deoxycytidine base pairing in DNA. *Nucleic Acids Res.* **35**, 6181–6195.
37. Chiu, T. K., Kaczor-Grzeskowiak, M., and Dickerson, R. E. (1999) Absence of minor groove monovalent cations in the crosslinked dodecamer C-G-C-G-A-A-T-T-C-G-C-G. *J. Mol. Biol.* **292**, 589–608.
38. Tsunoda, M., Kondo, J., Karino, N., Ueno, Y., Matsuda, A., and Takenaka, A. (2002) Water mediated Dickerson-Drew-type crystal of DNA dodecamer containing 2'-deoxy-5-formyluridine. *Biophys. Chem.* **95**, 227–233.
39. Tereshko, V., Minasov, G., and Egli, M. (1999) The Dickerson-Drew B-DNA dodecamer revisited at atomic resolution. *J. Am. Chem. Soc.* **121**, 470–471.
40. Minasov, G., Tereshko, V., and Egli, M. (1999) Atomic-resolution crystal structures of B-DNA reveal specific influences of divalent metal ions on conformation and packing. *J. Mol. Biol.* **291**, 83–99.
41. Egli, M., and Tereshko, V. (2004) Lattice- and sequence-dependent binding of  $Mg^{2+}$  in the crystal structure of a B-DNA dodecamer. *ACS Symp. Ser.* **884**, 87–109.
42. Pallan, P. S., Ittig, D., Heroux, A., Wawrzak, Z., Leumann, C. J., and Egli, M. (2008) Crystal structure of tricyclo-DNA: An unusual compensatory change of two adjacent backbone torsion angles. *Chem. Commun.*, 883–885.
43. Wurdeman, R. L., Douskey, M. C., and Gold, B. (1993) DNA methylation by N-methyl-N-nitrosourea: Methylation pattern changes in single- and double-stranded DNA, and in DNA with mismatched or bulged guanines. *Nucleic Acids Res.* **21**, 4975–4980.
44. Chiu, T. K., and Dickerson, R. E. (2000) 1 Å crystal structures of B-DNA reveal sequence-specific binding and groove-specific bending of DNA by magnesium and calcium. *J. Mol. Biol.* **301**, 915–945.
45. Howerton, S. B., Sines, C. C., VanDerveer, D., and Williams, L. D. (2001) Locating monovalent cations in the grooves of B-DNA. *Biochemistry* **40**, 10023–10031.
46. Hud, N. V., and Polak, M. (2001) DNA-cation interactions: The major and minor grooves are flexible ionophores. *Curr. Opin. Struct. Biol.* **11**, 293–301.
47. Egli, M. (2002) DNA-cation interactions. *Quo vadis? Chem. Biol.* **9**, 277–286.
48. Howerton, S. B., Nagpal, A., and Williams, L. D. (2003) Surprising roles of electrostatic interactions in DNA-ligand complexes. *Biopolymers* **69**, 87–99.
49. Manning, G. S. (1978) Molecular theory of polyelectrolyte solutions with applications to electrostatic properties of polynucleotides. *Q. Rev. Biophys.* **11**, 179–246.
50. Fenley, M. O., Manning, G. S., and Olson, W. K. (1990) Approach to the limit of counterion condensation. *Biopolymers* **30**, 1191–1203.
51. Anderson, C. F., and Record, M. T. (1995) Salt nucleic-acid interactions. *Annu. Rev. Phys. Chem.* **46**, 657–700.
52. Bloomfield, V. A. (1997) DNA condensation by multivalent cations. *Biopolymers* **44**, 269–282.
53. Manning, G. S. (2002) Electrostatic free energy of the DNA double helix in counterion condensation theory. *Biophys. Chem.* **101**, 461–473.
54. Saenger, W. (1984) in *Principles of Nucleic Acid Structure*, Springer, New York.
55. McConnell, K. J., and Beveridge, D. L. (2000) DNA structure: What's in charge? *J. Mol. Biol.* **304**, 803–820.
56. Manning, G. S. (2003) Comments on selected aspects of nucleic acid electrostatics. *Biopolymers* **69**, 137–143.
57. Manning, G. S. (2006) The contribution of transient counterion imbalances to DNA bending fluctuations. *Biophys. J.* **90**, 3208–3215.
58. Paoletta, D. N., Liu, Y., Fabian, M. A., and Schepartz, A. (1997) Electrostatic mechanism for DNA bending by bZIP proteins. *Biochemistry* **36**, 10033–10038.
59. Dickerson, R. E. (1998) DNA bending: The prevalence of kinkiness and the virtues of normality. *Nucleic Acids Res.* **26**, 1906–1926.
60. Williams, L. D., and Maher, L. J. (2000) Electrostatic mechanisms of DNA deformation. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 497–521.
61. Horton, N. C., and Perona, J. J. (2000) Crystallographic snapshots along a protein-induced DNA-bending pathway. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5729–5734.
62. Manning, G. S. (2003) Is a small number of charge neutralizations sufficient to bend nucleosome core DNA onto its superhelical ramp? *J. Am. Chem. Soc.* **125**, 15087–15092.
63. Marky, L. A., and Kupke, D. W. (2000) Enthalpy-entropy compensations in nucleic acids: Contribution of electrostriction and structural hydration. *Methods Enzymol.* **323**, 419–441.

BI800375M