Molecular Dynamics Simulations of Disulfide Cross-Linked DNA Decamers

Ann E. Ferentz, Joanna Wiorkiewicz-Kuczera, Martin Karplus,* and Gregory L. Verdine*

Contribution from the Department of Chemistry, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138

Received March 9, 1993

Abstract: Molecular dynamics simulations with the CHARMM program are reported for the decamer d(GCGAATTCGC)$_2$ and two derivatives of this duplex that contain disulfide cross-links attached to the exocyclic amines of the adenines in the central 5′-AT-3′ base pairs. The average structures from 20 ps of dynamics are compared to one another and to the crystal and NMR solution structures of the dodecamer d(CGCGAATTCGCG)$_2$. The duplexes all remain in the B-family of DNA, but cross-linking does affect some structural details. Most notable are changes in roll, tilt, and propeller twist near the center of the sequence that allow the cross-link to be accommodated in the major groove. The dynamics calculations show that there are some differences in the fluctuations near the site of cross-linking: in particular, there is a significant constraint on twisting induced by the shorter of the two cross-links.

Introduction

Site-specific cross-linking of DNA is a potentially valuable tool for the study of protein–DNA interactions. Many DNA-binding proteins, such as those involved in transcription, replication, recombination, and restriction and modification systems, require strand separation or local helical distortion to perform their functions.[1-4] Mechanistic studies of these proteins would be facilitated by the introduction of interstrand cross-links in DNA that are site-specific and can be incorporated into native sequences without severe structural distortions. Such cross-links could allow intermediates in replication or recombination to be trapped and characterized. Psoralen is a popular cross-linking reagent that has been used to trap the elongation complex of T7 RNA polymerase bound to DNA.[5] Although it reacts preferentially with thymines in 5′-TA-3′ sequences,[6] psoralen cannot be targeted to one specific site within a long piece of random DNA. Thus psoralen cross-links suffer from a lack of absolute sequence specificity. An additional drawback is that psoralen is known to distort the DNA surrounding the cross-link.[7] This means that the effect of covalently linking the two strands cannot be distinguished from the effect of the structural distortion that results from cross-linking.

To avoid these problems, a method has been developed for synthesizing site-specific interstrand disulfide cross-links between the exocyclic amines of the adenines within 5′-AT-3′ sequences in synthetic oligonucleotides (Figure 1d).[8] By relying on phosphoramidite chemistry to introduce the cross-link, absolute site-specificity is insured. Cross-links of two different lengths have been designed to be accommodated in the major groove with minimal distortion of the DNA. To examine the effects of the cross-links in more detail, we have undertaken molecular dynamics simulations[9,10] of the decamer with the two different cross-links and the corresponding native DNA, all of which have been studied experimentally. The calculation of the average structures of these oligonucleotides enables us to predict whether or not significant structural perturbations are introduced into the cross-linked DNA. Moreover, the simulations also permit us to investigate whether there are changes in the dynamic behavior of the DNA. If so, it may be possible to use cross-linking as a tool for constraining the mobility of DNA. The dynamic aspects of DNA structure may be a critical factor in protein–DNA interactions, just as is the average structure. One example is the

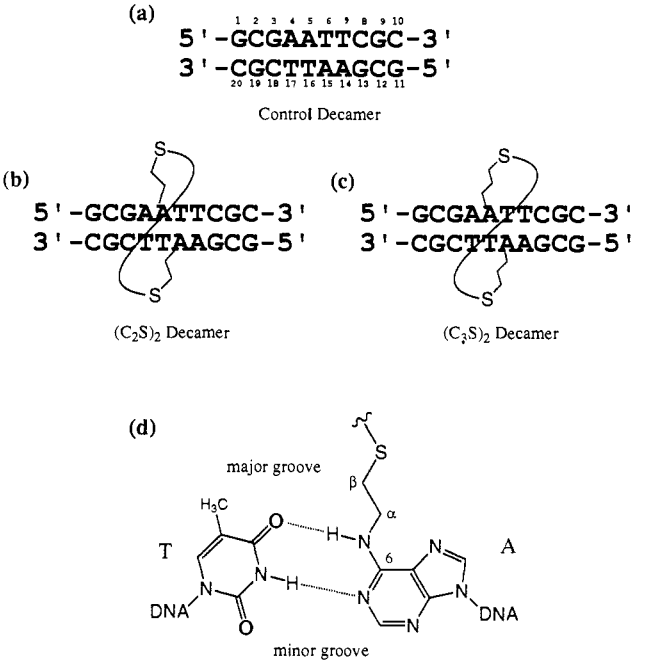

Figure 1. Schematic representation of DNA duplexes studied using molecular dynamics: (a) control decamer, (b) the $(C_2S)_2$ decamer, containing the C_2 cross-link, (c) the $(C_3S)_2$ decamer containing the C_3 cross-link, and (d) attachment of the C_2 cross-link to adenine via the exocyclic amine.

(1) von Hippel, P. H.; Bear, D. G.; Morgan, W. D.; McSwiggen, J. A. *Annu. Rev. Biochem.* **1984**, *53*, 389–446.
(2) McClure, W. R. *Annu. Rev. Biochem.* **1985**, *54*, 171–204.
(3) Nossal, N. G. *Annu. Rev. Biochem.* **1983**, *52*, 581–615.
(4) Cox, M. M.; Lehman, I. R. *Annu. Rev. Biochem.* **1987**, *56*, 229–262.
(5) Shi, Y.; Gamper, H.; Hearst, J. E. *J. Biol. Chem.* **1988**, *263*, 527–534.
(6) Gampar, H.; Piette, J.; Hearst, J. E. *Photochem. Photobiol.* **1984**, *40*, 29–34.
(7) Tomic, M. T.; Wemmer, D. E.; Kim, S.-H. *Science* **1987**, *238*, 1722–1725.
(8) Ferentz, A. E.; Verdine, G. L. *J. Am. Chem. Soc.* **1991**, *113*, 4000–4002. Ferentz, A. E.; Keating, T. A.; Verdine, G. L. *J. Am. Chem. Soc.*, in press.
(9) Brooks, C. L.; Karplus, M.; Pettit, B. M. *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*; Wiley: New York, 1988.
(10) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins & Nucleic Acids*; Cambridge University Press: New York, 1987.

0002-7863/93/1515-7569$04.00/0 © 1993 American Chemical Society

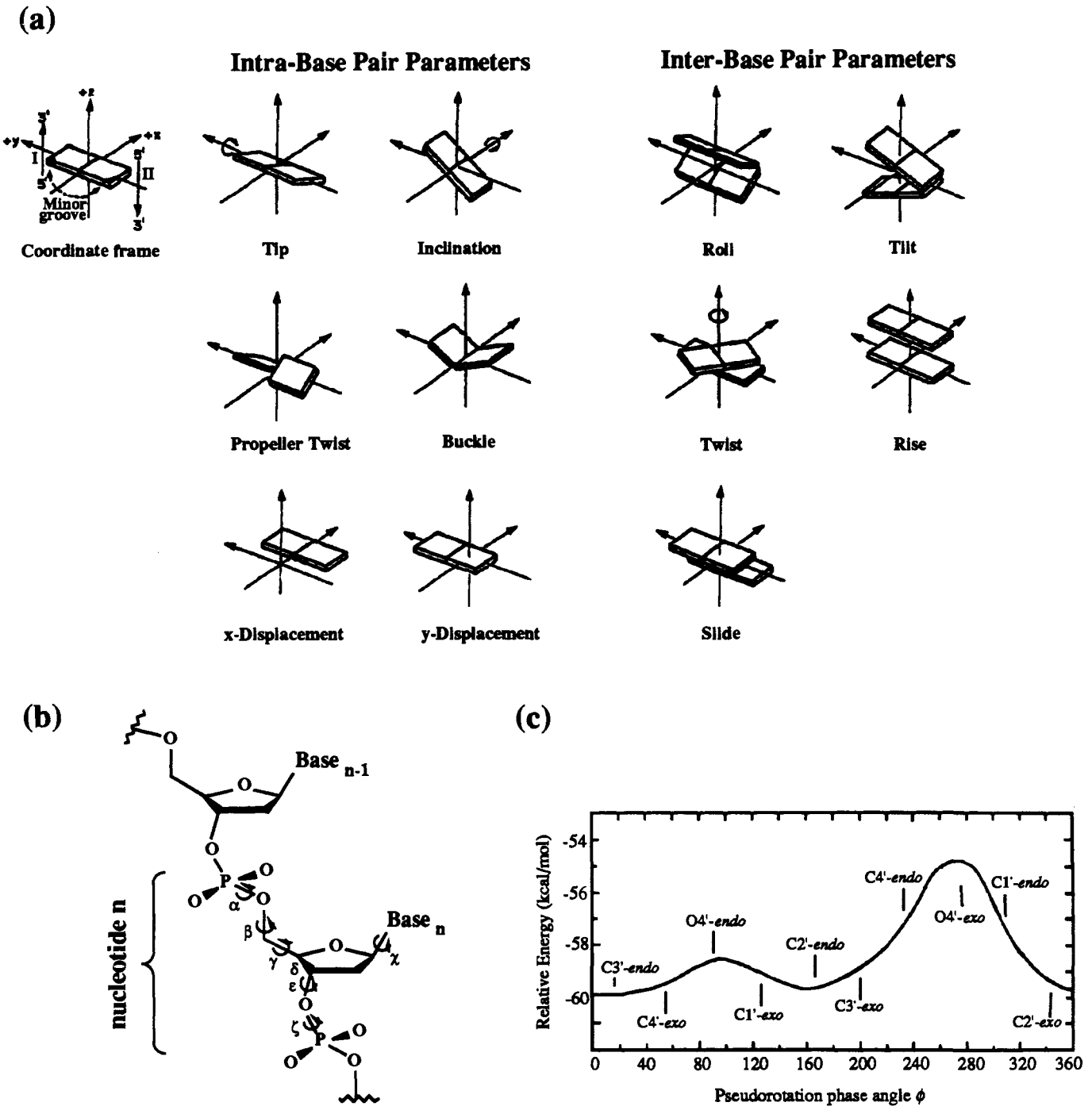

Figure 2. (a) Base pair parameters calculated by NewHelix (adapted from ref 38), (b) definition of backbone torsional angles, and (c) correlation of pseudorotation angle ϕ with sugar pucker (adapted from ref 39).

binding of the 434 repressor to its operator. The affinity of this protein is known to be affected by the sequence of the central bases in the operator even though these are not contact residues.[11] There is some controversy as to whether the preference for AT-rich sequences in the operator arises from their inherently greater flexibility in twisting or their overall lower rigidity as compared to GC-rich regions of DNA.[12,13] Knowledge of the dynamic features of DNA is necessary to address these issues, and the ability to constrain certain motions in DNA without affecting others would be very valuable in this regard.

The sequence used corresponds to the central ten base pairs of the "Dickerson dodecamer", d(CGCGAATTCGCG)$_2$, which has been extensively characterized. It is also of interest because the central six base pairs (GAATTC) constitute the recognition site for EcoR I endonuclease.[14] The initial crystal structure and analysis of the dodecamer were performed by Dickerson and co-workers.[15,16] This has been followed by many structural studies of derivatives of this sequence, either containing modified bases[17-19] or bound to drugs.[20-23] In addition, the duplex has been

(11) Koudelka, G. B.; Harbury, P.; Harrison, S. C.; Ptashne, M. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4633–4637.
(12) Koudelka, G. B.; Carlson, P. *Nature* **1992**, *355*, 89–91.
(13) Hogan, M. E.; Austin, R. H. *Nature* **1987**, *329*, 263–266.
(14) Kim, Y.; Grable, J. C.; Love, R.; Green, P. J.; Rosenberg, J. M. *Science* **1990**, *249*, 1307–1309.
(15) Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Dickerson, R. E. *Nature* **1980**, *287*, 755–758.
(16) Drew, H. R.; Wing, R. M.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 2179–2183.
(17) Drew, H. R.; Dickerson, R. E. *J. Mol. Biol.* **1981**, *151*, 535–556.
(18) Drew, H. R.; Samson, S.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 4040–4044.
(19) Fratini, A. V.; Kopka, M. L.; Drew, H. R.; Dickerson, R. E. *J. Biol. Chem.* **1982**, *257*, 14686–14707.

Table I. Atomic rms Differences (Å) between Initial, Minimized, and Average Structures from Molecular Dynamics[a]

	minimized	average	A-DNA	B-DNA
		Control Decamer		
initial	0.38(0.31)	2.1(1.3)	5.3(1.8)	1.1(0.8)
minimized		2.1(1.4)	5.3(1.8)	1.2(0.9)
average			5.7(2.0)	2.3(1.2)
		$(C_2S)_2$ Decamer		
initial	0.38(0.34)	1.7(1.0)	5.3(1.8)	1.1(0.8)
minimized		1.6(1.0)	5.3(1.8)	1.2(0.9)
average			4.8(1.9)	2.0(1.1)
		$(C_3S)_2$ Decamer		
initial	0.40(0.37)	1.7(1.3)	5.3(1.8)	1.1(0.8)
minimized		1.7(1.3)	5.3(1.8)	1.2(0.9)
average			5.4(2.1)	1.9(1.2)

[a] Values in parentheses are rms differences between the atoms in the central AATT sequence only.

characterized in solution by NMR spectroscopy,[24-26] and a solution structure has been determined.[27] There appear to be significant differences between the solution and crystal structures. The dodecamer has also been the subject of molecular dynamics studies using a range of force fields: two sets of results using AMBER 3.0[28,29] and simulations using GROMOS with and without solvent have been reported.[30,31] All of the simulations show the dodecameric duplex remaining in the B-form, but none of them reproduce the helical unwinding relative to A- or B-DNA that has been suggested by the three-dimensional NMR structure.[27] Accounting for solvent in some way, either by explicit inclusion of counterions and water molecules or by a dielectric term, appears to be important for maintaining proper groove structure.[29] In the present simulations, we reduce the charges on the phosphate backbone in accord with the Manning condensation model[32] and use a distance-dependent dielectric.[33] This approach has been shown to give satisfactory results in simulations of B and Z hexamers.[33]

Methods

Structures Studied. The initial coordinates for the calculations were generated from the crystallographic coordinates of the Dickerson dodecamer, d(CGCGAATTCGCG)$_2$,[16] by removing the terminal nucleotides. For the cross-linked molecules, an interstrand tether was then appended to the amino groups of the adenines in the central 5′-AT-3′ sequence (A5–A15; see Figure 1). We refer to the native sequence as the control decamer (Figure 1a), to the duplex containing a two-carbon linker (the C_2 cross-link) between the adenine amine and the disulfide as the $(C_2S)_2$ decamer (Figure 1b), and to the duplex containing a three-carbon linker (the C_3 cross-link) as the $(C_3S)_2$ decamer (Figure 1c).

Computational Details. All molecular dynamics calculations were performed with the program CHARMM.[34] The potential energy function used is the sum of quadratic terms describing bond and bond angle deformations, cosine terms for torsions, and quadratic terms for improper torsions, Lennard-Jones 6–12 terms for van der Waals interactions, Coulomb terms for electrostatic interactions, and hydrogen-bond terms;[35] the function used thus has the form

$$V_{tot} = \sum_{bonds}\frac{1}{2}K_b(b-b_0)^2 + \sum_{\substack{bond\\angles}}\frac{1}{2}K_\theta(\theta-\theta_0)^2 + \sum_{\substack{dihedral\\angles}}K_\phi[1-\cos(n\phi-\delta)] + \sum_{\substack{improper\\torsions}}\frac{1}{2}K_\omega(\omega-\omega_0)^2 + \sum_{\substack{nonbonded\\pairs}}\left\{4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12}-2\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]+\frac{332q_iq_j}{Dr_{ij}}\right\}+\sum_{\substack{H\text{-}bonded\\pairs}}\left(\frac{A_{ij}}{r_{ij}^{6}}-\frac{B_{ij}}{r_{ij}^{4}}\right)\cos^4\phi_{DHA}\cos^2\phi_{AAAH}$$

where K_b, K_θ, K_ϕ, and K_ω are the bond stretch, angle, dihedral, and improper dihedral force constants, respectively; b_0, θ_0, and ω_0 are the equilibrium values for bond lengths, valence angles, and improper dihedral angles, respectively; and n and δ are the multiplicity and phase in the torsional potential. In the expressions for the nonbonded terms r_{ij} is the interatomic distance, ϵ_{ij} and σ_{ij} define the van der Waals energy well depth, D is an effective distance dependent dielectric function (r_{ij}), and q is the partial atomic charge. In the hydrogen bond term A_{ij} and B_{ij} define the depth of the hydrogen bond well, ϕ_{DHA} and ϕ_{AAAH} are the donor–hydrogen–acceptor and acceptor–antecedent–acceptor–hydrogen angles, respectively.

The DNA parameters were taken from the work of Nilsson and Karplus.[35] Polar hydrogen atoms (in –NH, NH_2, and –OH groups) were included explicitly in the calculations, while –CH, $-CH_2$, and $-CH_3$ groups were treated as extended atoms. The parameter and topology files were extended to include descriptions of sulfur and disulfides from the protein potential function; the required parameters were taken from CHARMM 19.

Nonbonded interactions were calculated for atoms separated by three or more covalent bonds. The interaction energy between atoms separated by three bonds (1,4-interactions) was not scaled. Nonbonded interactions were computed with a switch truncation for the van der Waals interactions between 9.5 and 10.5 Å with a cutoff (list) distance of 11.5 Å, a shift truncation for electrostatic interactions at 10.5 Å, and a switching function applied between 4.0 and 7.5 Å for hydrogen bonds, with a cutoff distance of 7.0 Å.

The systems studied consisted of 450 atoms for the control decamer, 454 atoms for the $(C_2S)_2$ decamer, and 456 atoms for the $(C_3S)_2$ decamer. To reduce electrostatic interactions in the absence of explicit solvent and counterions a Manning-type condensation model[32] was adapted: the charge on each phosphate group was reduced to –0.32e, and a $1/r$ shielding distance dependent function was used as the effective dielectric parameter D.[33,35,36]

The energies of the crystal structure of the control decamer and the starting structures of the cross-linked decamers were minimized by the adopted basis set Newton–Raphson (ABNR) method for 100 steps to eliminate poor bond lengths and angles or bad contacts.[34] Dynamics simulations of the energy-minimized structures were performed by numerical integration of Newton's

(20) Wing, R. M.; Pjura, P.; Drew, H. R.; Dickerson, R. E. *EMBO J.* **1984**, *3*, 1201–1206.
(21) Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. *J. Mol. Biol.* **1985**, *184*, 553–563.
(22) Pjura, P. E.; Grzeskowak, K.; Dickerson, R. E. *J. Mol. Biol.* **1987**, *197*, 257–271.
(23) Teng, M.; Usman, N.; Frederick, C. A.; Wang, A. H.-J. *Nucl. Acids Res.* **1988**, *16*, 2671–2690.
(24) Hare, D. R.; Wemmer, D. E.; Chou, S.-H.; Drobny, G. *J. Mol. Biol.* **1983**, *171*, 319–336.
(25) Patel, D. J.; Kozlowski, S. A.; Marky, L. A.; Broka, C.; Rice, J. A.; Itakura, K.; Breslauer, K. J. *Biochemistry* **1982**, *21*, 428–436.
(26) Ott, J.; Eckstein, F. *Biochemistry* **1985**, *24*, 2530–2535.
(27) Nerdal, W.; Hare, D. R.; Reid, B. R. *Biochemistry* **1989**, *28*, 10008–10021.
(28) Rao, S.; Kollman, P. *Biopolymers* **1990**, *29*, 517–532.
(29) Srinivasan, J.; Withka, J. M.; Beveridge, D. L. *Biophys. J.* **1990**, *58*, 533–547.
(30) Ravishanker, G.; Swaminathan, S.; Beveridge, D. L.; Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1989**, *6*, 669–699.
(31) Swaminathan, S.; Ravishanker, G.; Beveridge, D. L. *J. Am. Chem. Soc.* **1991**, *113*, 5027–5040.
(32) Manning, G. S. *Quarterly Rev. Biophys.* **1978**, *11*, 179–246.
(33) Tidor, B.; Irikura, K. K.; Brooks, B. R.; Karplus, M. *J. Biomol. Struct. Dyn.* **1983**, *1*, 231–252.
(34) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
(35) Nilsson, L.; Karplus, M. *J. Comput. Chem.* **1986**, *7*, 591–616.
(36) Nilsson, L.; Karplus, M. In *Proceedings of the NATO Advanced Research Workshop on 3D Structure and Dynamics of RNA*; Plenum Press: New York, 1986; p 151ff.

Table II. Base Pair Parameters in Starting Structure of Decamer[a]

Intrabase Pair Parameters												
base pair	tip	(rms)	incl	(rms)	x-dsp	(rms)	y-dsp	(rms)	PrTw	(rms)	buckle	(rms)
G1–C20	2.02	(1.41)	9.40	(0.62)	–0.09	(0.08)	–0.55	(0.12)	–10.83	(1.56)	4.41	(3.29)
C2–G19	–7.24	(1.01)	8.32	(0.76)	0.00	(0.04)	–2.69	(0.14)	–18.09	(1.21)	3.95	(1.20)
GC–C18	–5.33	(2.39)	5.43	(0.63)	–1.05	(0.05)	–0.04	(0.11)	–10.37	(1.24)	–8.14	(0.92)
A4–T17	–1.66	(0.78)	2.06	(0.31)	–1.18	(0.13)	0.46	(0.05)	–16.22	(0.89)	–5.51	(1.10)
A5–T16	–1.37	(0.82)	0.00	(0.32)	–0.99	(0.07)	0.69	(0.05)	–17.52	(2.91)	–3.16	(1.71)
T6–A15	–7.40	(0.85)	–0.88	(1.43)	–0.63	(0.05)	0.55	(0.03)	–17.02	(3.14)	–2.03	(1.96)
T7–A14	–5.97	(1.10)	–2.46	(0.94)	–0.57	(0.09)	0.70	(0.06)	–17.07	(2.98)	0.20	(1.65)
C8–G13	–0.47	(1.32)	–4.71	(0.78)	–0.11	(0.16)	1.00	(0.08)	–15.97	(2.09)	9.60	(5.79)
G9–C12	3.62	(1.72)	–4.35	(0.59)	1.02	(0.23)	1.89	(0.10)	–4.68	(1.24)	–1.56	(3.29)
C10–G11	–2.63	(0.92)	–6.73	(1.00)	1.33	(0.25)	2.08	(0.12)	–16.93	(2.20)	3.14	(1.40)

Interbase Pair Parameters										
base pair step	roll	(rms)	tilt	(rms)	twist	(rms)	rise	(rms)	slide	(rms)
(G1–C20)/(C2–G19)	–2.83	(1.20)	–1.56	(0.84)	37.82	(0.62)	2.96	(0.04)	14.56	(6.04)
(C2–G19)/(G3–C18)	3.79	(2.43)	–3.63	(0.87)	27.88	(1.81)	4.14	(0.02)	–7.82	(3.70)
(G3–C18)/(A4–T17)	5.81	(1.73)	–1.21	(0.32)	36.83	(1.20)	3.02	(0.07)	0.04	(0.06)
(A4–T17)/(A5–T16)	0.91	(0.44)	–1.00	(0.87)	37.19	(0.52)	3.25	(0.06)	–0.40	(0.04)
(A5–T16)/(T6–A15)	–6.06	(1.93)	1.55	(1.12)	31.82	(0.39)	3.12	(0.10)	–0.62	(0.08)
(T6–A15)/(T7–A14)	0.30	(1.54)	2.58	(1.18)	35.59	(0.31)	3.26	(0.03)	–0.31	(0.05)
(T7–A14)/(C8–G13)	2.67	(1.06)	0.14	(1.31)	40.67	(0.27)	3.00	(0.01)	–0.10	(0.07)
(C8–G13)/(G9–C12)	1.72	(0.50)	–0.44	(0.63)	30.59	(0.22)	4.10	(0.11)	0.81	(0.03)
(G9–C12)/(C10/G11)	–9.82	(2.57)	–2.61	(0.90)	40.27	(1.03)	3.00	(0.05)	0.65	(0.03)

[a] The data listed here are the results of NewHelix analysis of the crystal structure of the Dickerson dodecamer at 290 K (Drew et al., 1981a) reported in the Brookhaven Protein Data Bank. Note that the numbering of the bases differs from that in the dodecamer (e.g., G2 in the dodecamer is G1 in the decamer). The rms deviations come from analysis of the four sets of crystal coordinates for the native dodecamer in the Protein Data Bank.

equations of motion using the VERLET algorithm with a timestep of 1 fs.[37] Covalent bonds involving hydrogens were constrained to their equilibrium values with the SHAKE algorithm to shorten computation time.[38] In all simulations, the system was heated from 0 to 300 K over a period of 6 ps and then equilibrated at 300 K for 3 ps. Production coordinates were saved every 50 fs for a period of 20 ps. The average structures from the last 5 ps of dynamics were minimized by the ABNR method for 200 steps to eliminate incorrect structural features introduced by the averaging. Duplicate runs were performed for the control decamer and the cross-linked decamers to determine which of the differences found for the three systems are likely to be significant. All runs were analyzed by the same methods. For the most part, results from the two data sets agreed closely, and, for the sake of space, only one set of results is presented here. Discrepancies between duplicate runs are mentioned explicitly in the text and serve as a measure of the possible error in the calculations. In the analysis, we regard as significant differences that appear in both sets of simulations and are greater than the rms fluctuations.

Analysis of Trajectories. Production data were analyzed using CHARMM,[34,35] NewHelix,[39] and Dials and Windows.[30] Parameters examined included backbone torsional angles, torsional angles in the cross-links, hydrogen bond lengths, and the base pair parameters shown in Figure 2a. The base pair parameters, defined according to the Cambridge Convention,[39] were calculated with NewHelix for each 50 fs time frame of the dynamics trajectory and then averaged. The backbone torsional angles were analyzed from the production data using Dials and Windows, while interatomic distances (i.e., hydrogen bond lengths and groove widths) were calculated using CHARMM.

Results

I. Structural Results. The simulations of the control and cross-linked decamers indicate that they remain Watson–Crick hydrogen bonded B-form DNA duplexes. The atomic rms deviations of the average structure during the last 5 ps of dynamics from the starting structure (crystal coordinates), minimized structure, and canonical B- and A-DNA are summarized in Table I. All three duplexes remain within 2.3 Å of canonical B-DNA and near the starting structure (within 2.1 Å), with the average structures of the cross-linked decamers lying slightly closer to the initial coordinates than the control decamer. When the highly fluctuating terminal base pairs are not considered, the rms deviations are significantly smaller, with all the average structures lying within 1.2 Å of canonical B-DNA and 1.3 Å of the crystal structure. To examine fine details of the structures, the helical and base pair parameters for each are compared to one another and to the starting crystal structure. The results are presented first for the control decamer and then for the cross-linked decamers, with a brief comparison of the structures to each other and to the crystal structure. A second section focuses on the dynamic properties of the simulations and compares the results with and without a cross-link. A comparison of the dynamic results to crystal structures, other simulations, and the solution structure is presented in the Discussion.

A. Structure of the Unmodified Decamer, 5′-GCGAATTCGC-3′. Table II lists the values of the base pair parameters in the central ten base pairs of the crystal structure of 5′-CGCGAATTCGCG-3′ at 290 K and shows the rms deviations for these parameters, based upon the four crystal structures of this sequence that have been deposited in the Protein Data Bank.[16,18,41,42] Although these sets of coordinates vary in the temperature at which data were obtained or in the refinement method used, they provide an indication of the flexible parameters in the oligomers. Table III summarizes the average values of these parameters for the decamer during 20 ps of dynamics at 300 K. Average values are displayed graphically in Figure 3, along with those for the Drew crystal structure. In essentially all cases the root mean square fluctuations of the parameters during dynamics are larger than the difference between their average values and those for the crystal structure. Since the averages do not accurately represent the values for parameters that undergo abrupt transitions during the course of dynamics, these cases are examined individually; see Dynamics Results below. Overall, the simulations show more variation from base to base than does the crystal structure. Whether this is due to the short duration of the simulation or whether it is inherent to the model is not clear.

(37) Verlet, L. *Phys. Rev.* **1967**, *159*, 98–103.
(38) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311–1327.
(39) Dickerson, R. E. *J. Biomol. Struct. Dyn.* **1989**, *6*, 627–634.
(40) Levitt, M.; Warshel, A. *J. Am. Chem. Soc.* **1978**, *100*, 363–384.
(41) Holbrook, S. R.; Dickerson, R. E.; Kim, S.-H. *Acta Crystallogr., Sect. B* **1985**, *41*, 255–262.
(42) Westhof, E. *J. Biomol. Struct. Dyn.* **1987**, *5*, 581–600.

Table III. Base Pair Parameters in Control Decamer

Intrabase Pair Parameters

	tip		inclination		*x*-displacement		*y*-displacement		propeller twist		buckle	
base pair	mean	(rms)	mean	(rms)	mean	(rms)	mean	(rms)	mean	(rms)	mean	(rms)
G1–C20	78.75	(33.97)	–1.11	(2.03)	2.43	(0.62)	0.23	(0.53)	–96.54	(11.12)	–14.42	(11.48)
C2–G19	–8.55	(4.94)	–2.12	(4.98)	1.44	(0.52)	–1.44	(0.52)	–16.13	(8.53)	–9.82	(8.82)
G3–C18	–4.46	(4.37)	–8.02	(3.89)	–0.34	(0.34)	–0.59	(0.46)	–11.89	(7.11)	–10.98	(7.49)
A4–T17	1.84	(4.20)	–5.81	(2.80)	0.20	(0.31)	–0.05	(0.33)	–14.12	(5.87)	–5.72	(7.00)
A5–T16	0.29	(4.30)	–5.27	(2.71)	–0.19	(0.37)	0.61	(0.36)	–24.99	(6.70)	8.76	(7.12)
T6–A15	3.45	(4.63)	–0.16	(2.71)	1.32	(0.47)	0.17	(0.38)	–15.91	(6.96)	11.27	(9.01)
T7–A14	–6.27	(5.24)	1.00	(3.51)	1.34	(0.46)	–0.68	(0.45)	–14.69	(6.71)	4.84	(9.89)
C8–G13	–4.90	(5.16)	–5.10	(5.04)	0.29	(0.38)	–1.03	(0.46)	–13.23	(6.70)	3.07	(8.04)
G9–C12	7.94	(6.16)	–4.74	(5.05)	–0.36	(0.51)	–1.15	(0.60)	–9.29	(6.73)	7.28	(13.28)
C10–G11	1.99	(9.40)	4.81	(8.82)	–1.42	(1.01)	0.45	(0.59)	–12.24	(13.05)	–5.10	(21.38)
A-DNA	–0.4		19.7		–4.5		–0.1		–11.2		0.0	
B-DNA	0.0		–4.6		0.2		0.0		–1.3		0.0	

Interbase Pair Parameters

	roll		tilt		twist		rise		slide	
base pair step	mean	(rms)	mean	(rms)	mean	(rms)	mean	(rms)	mean	(rms)
(G1–C20)/(C2–G19)	–20.87	(35.99)	–21.90	(12.23)	52.31	(2.31)	1.86	(0.71)	0.25	(0.34)
(C2–G19)/(G3–C18)	3.53	(5.09)	–3.92	(2.26)	17.42	(3.06)	4.12	(0.20)	0.71	(0.35)
(G3–C18)/(A4–T17)	–0.17	(4.89)	3.12	(2.75)	49.19	(2.58)	2.31	(0.27)	0.07	(0.21)
(A4–T17)/(A5–T16)	–5.75	(4.42)	–0.46	(2.38)	46.07	(4.47)	2.63	(0.22)	0.26	(0.27)
(A5–T16)/(T6–A15)	1.71	(4.42)	4.05	(1.92)	30.02	(3.14)	3.33	(0.22)	–0.37	(0.27)
(T6–A15)/(T7–A14)	–8.83	(4.79)	2.01	(2.59)	36.29	(3.13)	2.02	(0.32)	0.12	(0.28)
(T7–A14)/(C8–G13)	0.07	(4.88)	–2.29	(2.93)	38.10	(3.50)	3.28	(0.22)	0.10	(0.30)
(C8–G13)/(G9–C12)	9.67	(5.72)	–0.51	(2.79)	34.09	(3.79)	3.79	(0.26)	–0.38	(0.25)
(G9–C12)/(C10–G11)	–5.37	(8.16)	4.42	(4.97)	46.33	(4.10)	2.46	(0.27)	0.44	(0.53)
A-DNA[a]	10.5		0.0		32.7		2.3		–2.7	
B-DNA[a]	–2.8		0.0		36.0		3.4		–1.7	

[a] Ideal values are from models of ideal A- and B-form DNA generated in Quanta and analyzed using NewHelix.

Helical Structure. The base pair parameters that contribute to the helicity of the DNA are the interbase pair parameters twist and rise and the intrabase pair parameters *x*-displacement and *y*-displacement. These relate directly to the groove widths and the general form of the DNA (A or B). Examination of the twist and rise shows that the structure from dynamics has an increased twist and reduced rise across base pair steps G3/A4 and A4/A5, while the symmetry-related steps exhibit twists very close in magnitude to those in the crystal. During another dynamics run (data not shown) the values of twist to either side of the central bases are reversed (i.e., the twist at A4/A5 is nearly the value in the crystal structure, and T6/T7 is more highly twisted). The asymmetry in twisting is also present in the crystal structure, indicating that changes in twist are easily achieved, i.e., are "soft" distortions. The trend toward higher twist and/or lower rise in the simulation as compared to the crystal structure results in a "shorter, fatter" helix relative to the starting structure and is seen in the structures of the cross-linked decamers as well as the control decamer (Figure 4).

The change in helical conformation compared to the crystal structure is also reflected in the groove widths and *x*-displacement of the base pairs relative to the helical axis. Groove widths were calculated as the shortest phosphorus–phosphorus distances across the grooves (Figure 5). At the center of the control decamer both the major and minor grooves are consistently narrower than in canonical B-DNA (Table VI), in agreement with the crystal structure. But the major groove is somewhat narrower than in the crystal structure, while the AATT tract in the minor groove is wider than in the crystal. In the solution structure, both the major and minor grooves are somewhat wider at the center of the duplex (25–26 and 12 Å, respectively) than in either the crystal structure or the simulation. This holds for all the molecules simulated, although there are some significant differences between the behavior of the cross-linked decamers and the control decamer (see below). Accompanying this narrowing of the grooves with respect to B-DNA is a significant positive *x*-displacement of the bases toward the major groove; this is not observed in the crystal or solution structures.

Backbone Geometry. The values for the torsional angles along the backbone of the decamer during dynamics are depicted in Figure 6a.

The torsional angle δ (C5′–C4′–C3′–O3′) (Figure 2b) and pseudorotation phase angle ϕ (Figure 2c) are indicative of the sugar puckers along the DNA backbone. Throughout the dynamics trajectories, the sugar conformations are either $C_{2'}$-*endo* or $C_{3'}$-*endo* (characteristic of canonical B-DNA or A-DNA, respectively). Transitions between these two regions do occur, mainly in a concerted fashion, as can be seen at $\delta 6/\phi 6$ of the decamer (Figure 6a), but most of these angles remain stable throughout the run. A third set of torsional angles that are correlated with the sugar conformation are the glycosidic bond angles χ [O4′–C1′–N1–C2 (pyrimidines) or O4′–C1′–N9–C4 (purines)]. These angles orient the bases with respect to their sugar rings and maintain values close to those normally found in A- or B-DNA throughout the runs.

The backbone angles may be considered in two groups, α, β, and γ to the 5′ side of the sugar and ϵ and ζ to the 3′ side of the sugar (Figure 2b). The most stable parameters throughout the simulations are the backbone angles γ, which remain in the g^+ conformation in all of the trajectories, just as they do in the crystal structure. Similarly, α and β exhibit values in the (g^-,t) range, close to those seen in the crystal. The only base that exhibits values drastically different from the crystal structure is A15, which has a gauche conformation of β. This change would tend to bring A14 and A15 closer together than in the crystal structure. The predominant conformation of ϵ and ζ throughout all three trajectories is (t,g^-), the B_I conformation[19] seen in the crystal structure. Transitions to the alternate B_{II} conformation, (ϵ,ζ) = (g^-,t), are observed in a few instances (e.g., the transient spike in $\epsilon 6/\zeta 6$), and there are a few unusual values: (ϵ,ζ)4, (ϵ,ζ)-15, and (ϵ,ζ)18 lie close to a (t,t) configuration. This allows the regions of the backbone containing these angles to be elongated and the bases to either side to lie farther away from one another than in the crystal structure.

Base Pair Geometry. The average values for the base pair parameters from NewHelix calculations are presented in Table

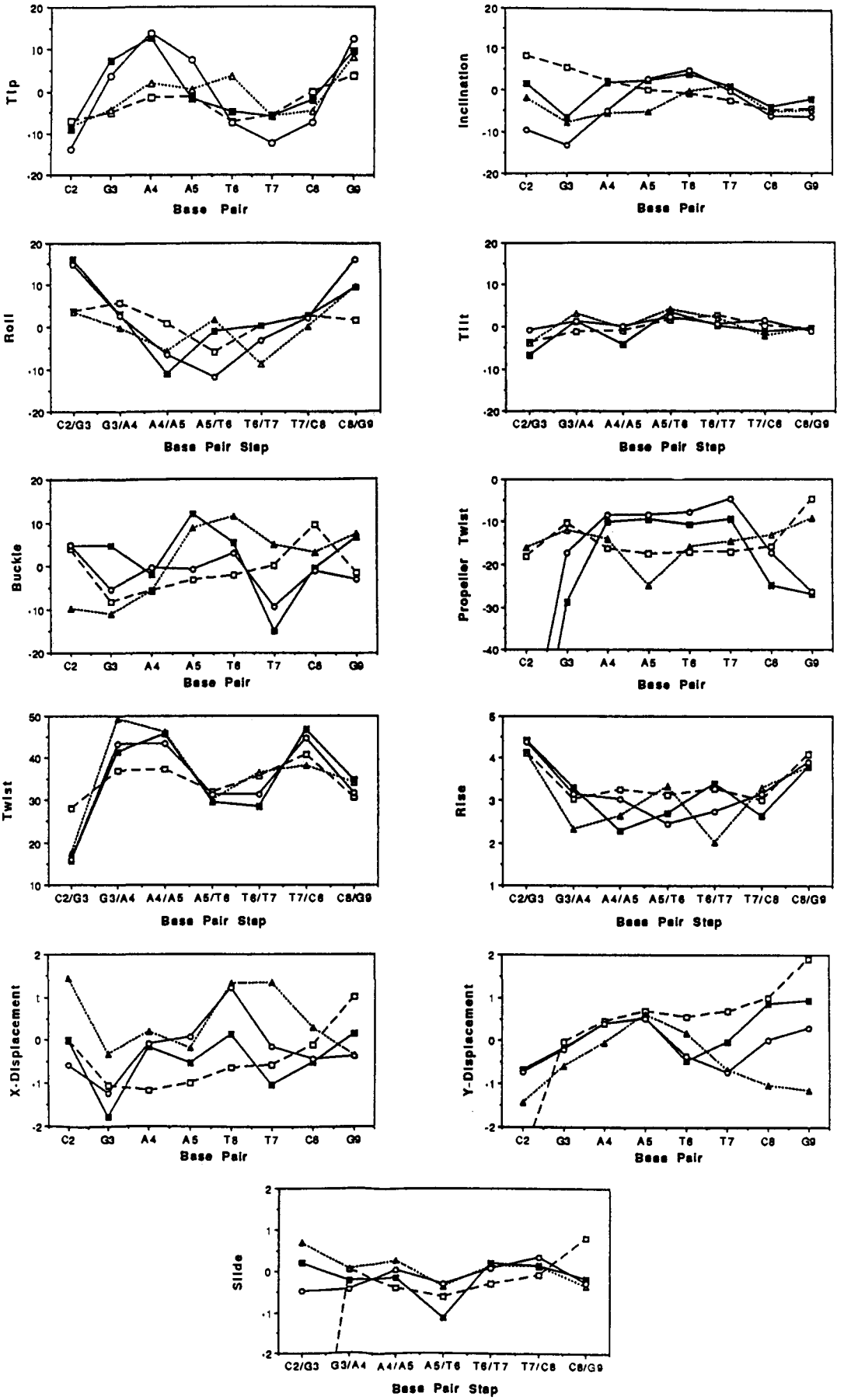

Figure 3. Average values of helical and base pair parameters from molecular dynamics runs of the decamer (△), $(C_2S)_2$ decamer (■), and $(C_3S)_2$ decamer (○) as compared to the crystal structure (□). Base pairs are referred to by the lower-numbered base in the pair.

III and Figure 3. The tip angles for the decamer parallel those for the crystal structure with the exception of T6–A15, which shows a larger tip than in the crystal. This difference is correlated with an increase in the roll between base pairs 5 and 6, the only

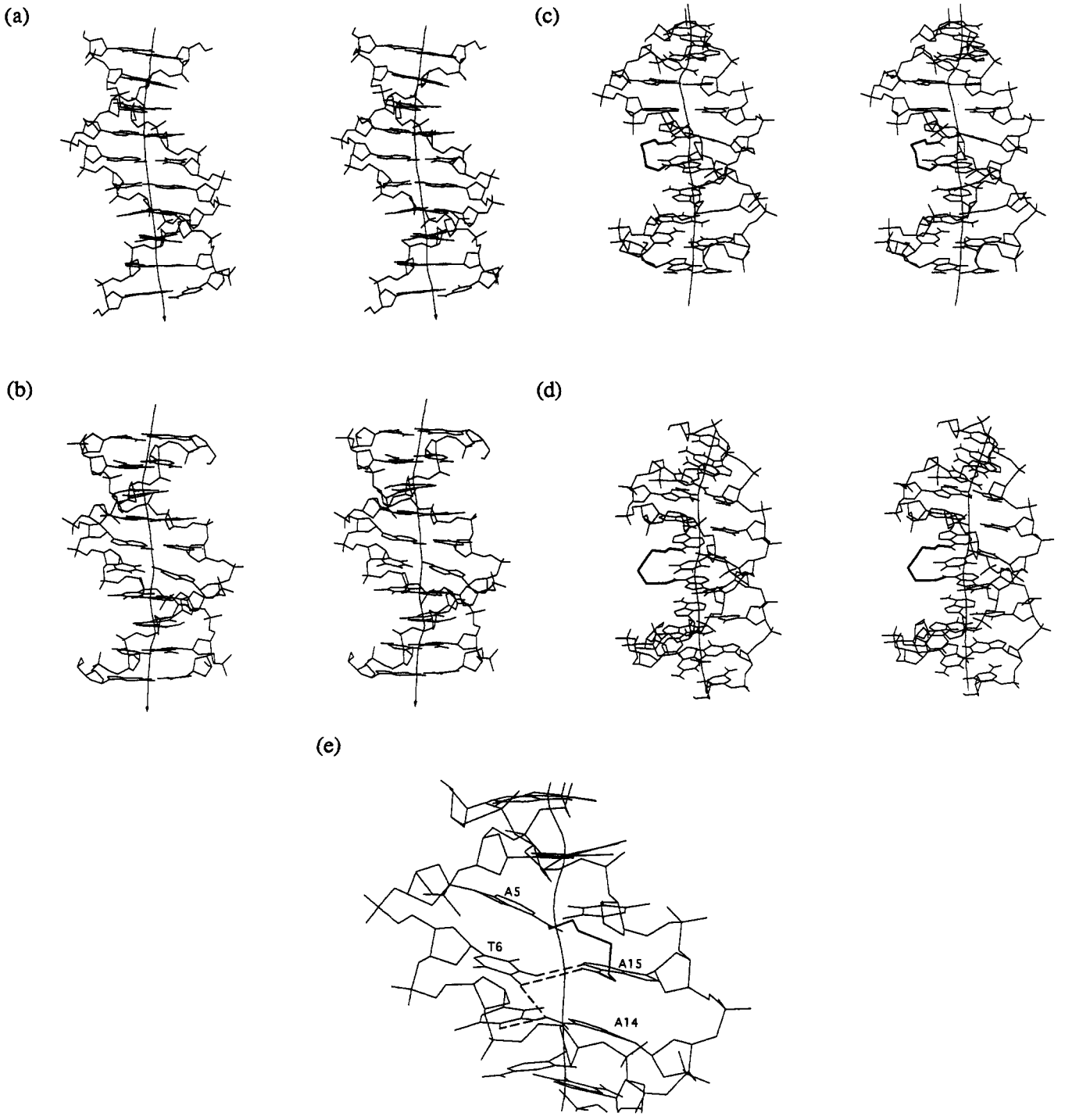

Figure 4. Stereoviews of (a) the starting crystal structure, as compared to the average structures from the last 5 ps of dynamics for (b) the control decamer, (c) the $(C_2S)_2$ decamer, and (d) the $(C_3S)_2$ decamer and (e) a closeup view of the central base pairs with the $(C_2S)_2$ decamer. The helical axis calculated by Dials & Windows using the Curves algorithm[40] is shown as a solid line running vertically through each structure.

roll angle that is larger in the simulation than in the crystal structure. The inclinations of base pairs 3, 4, and 5 are smaller in the structure from dynamics than in the crystal structure, but this difference is not significant, as in another run the inclinations of base pairs 4 and 5 were almost identical to the values in the crystal structure. The values of tilt do not differ significantly from those in the crystal structure, except for the terminal base pairs.

The relative positions of bases within a base pair are described by propeller twist and buckle. Buckling shows a wider range of values in the average structure from dynamics than in the crystal structure, but the dynamic fluctuations are so large (see below) that the differences may not be significant. Propeller twisting remains almost the same in the simulation as in the crystal structure, with the exception of the high propeller twist at A5–T16.

B. Structures of the Cross-Linked Decamers. Helical Structure. During the course of dynamics, the cross-linked decamers exhibit the same overall widening of the helix relative to the crystal structure as does the control decamer (Figure 4). In terms of twisting, the cross-linked structures show values very close to one another and to the native sequence. They have slightly lower degrees of twisting at the G3/A4 and T6/T7 steps and slightly higher twist at T7/C8 than does the control (Figure 3). Decreased twisting is accompanied by an increase in rise at these steps with respect to the native decamer, but in general the values of rise

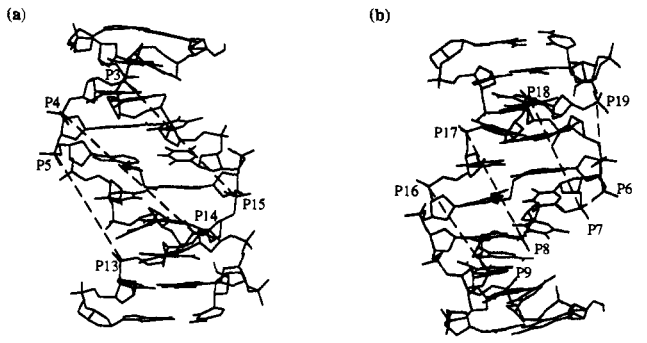

Figure 5. Measurement of groove widths: (a) the major groove and (b) the minor groove.

between the central six base pairs all lie between 2.0 and 3.4 Å and do not show any conclusive trends. The decamers containing cross-links also show parallel trends in terms of x-displacement. Like the control, the cross-linked decamers show positive x-displacement at base pair 6; however, in the other simulation of the control decamer this x-displacement is smaller than that in the $(C_3S)_2$ decamer.

Groove widths are considerably affected by the presence of either cross-link. The minor groove is generally wider in the cross-linked decamers than in the control, with values approaching those found in canonical B-DNA (Table VI). Both cross-linked species show a similar widening of the minor groove across the site of cross-linking (P8–P17 distances; see Figure 5b) relative to the decamer. To either side of the center of the helix, the cross-linked decamers differ from one another: on average, the $(C_2S)_2$ decamer has a longer P7–P18 distance and shorter P9–P16 distance than the $(C_3S)_2$ decamer (though the latter is due to a transition in the trajectory and may not be significant; Figure 9). The major groove width is also altered near the site of cross-linking. At the center of the duplex the C_2 cross-link causes a widening of the groove (P4–P14 distance) relative to the control decamer or the $(C_3S)_2$ decamer, so that it becomes nearly as wide as in the crystal structure. This may be because one of the Cα–Cβ torsional angles (see Figure 1d) must assume a gauche conformation in order for the disulfide torsional angle to remain close to 90° (Figure 4c,e). This would tend to push the major groove edges of the linked adenines away from each other causing a widening of the groove. On one side of the center, the P5–P13 distance is made slightly shorter by cross-linking, while on the other side the P3–P15 distance is elongated (however, another run showed that the latter difference is not significant).

Backbone Geometry. The cross-linked duplexes, like the control, maintain $C_{2'}$-*endo* and $C_{3'}$-*endo* sugar puckers throughout dynamics, with occasional transitions between these conformations (Figure 6b,c). In accord with the behavior observed in the control decamer, γ is very stable throughout the runs. The angles α and β reside mostly in the (g^-,t) conformation, with the exception of T16. T16 also shows an unusual (t,t) conformation of ϵ and ζ throughout most of the simulations of the cross-linked molecules. This extended (t,t) conformation also occurs on the other side of the C_2 cross-link at A4. The C_3 cross-link sometimes causes T6 to reside in the (t,t) conformation and A4 to explore the alternate B_{II} (g^-,t) conformation.

Base Pair Geometry. (See Figure 3 and Tables IV and V.) Tip and roll across the central base pairs are affected by cross-linking. The tip at A4 is markedly higher for the cross-linked molecules than for either the crystalline conformation or the control decamer, as is the tip of A5 in the $(C_2S)_2$ decamer. Meanwhile, T6 exhibits a slightly smaller value of tip with either cross-link than in the native duplex, and T7 maintains this trend in the $(C_3S)_2$ decamer. The positive tip of A5 and negative tip of T6 in the $(C_3S)_2$ decamer combine to yield a more negative roll angle at this step than is seen in the control. The $(C_2S)_2$ decamer also shows a slightly smaller roll angle at this step than does the control, but the difference is not significant in this run; it was much more pronounced in the other simulation, where the mean A5/T6 roll angle was 8° lower in the $(C_2S)_2$ decamer than in the control decamer. This more negative roll at the center of the helix is accompanied by a slightly more positive roll at the adjacent T6/T7 step in the cross-linked duplexes. The central base pairs tend to have higher inclination angles when cross-linked than in the control decamer, but because these values parallel those in the decamer and the tilt angle is measured as the difference between the inclination angles in adjacent base pairs, the tilt angles have values very close to those observed in the control and the crystal structure.

Buckling shows a high degree of fluctuation (see Dynamics Results below), so the only significant difference between the cross-linked decamers and the control is at T7–A14, where the cross-linked species show negative buckling relative to the control. The average buckle at A5–T16 is smaller in the $(C_3S)_2$ decamer than in the other two duplexes, but the difference from the $(C_2S)_2$ decamer is not significant. Propeller twisting, however, is significantly lower throughout the central AATT region in the presence of either cross-link.

Cross-Link Parameters. The cross-links themselves are asymmetric on the picosecond timescale, just as is the whole duplex. In both cross-links, the angles near the disulfide approach their optimal values: the disulfide dihedral angles remain close to ±90°, and the dihedral angle adjacent to the disulfide (Cβ–S for the C_2 cross-link and Cγ–S for the C_3 cross-link) lies in the range of 90–135° (Figure 10). The rest of the angles in the C_3 cross-link can assume extended conformations, but the C_2 cross-link takes on a more asymmetric configuration having one of the Cα–Cβ torsional angles nearly trans and the other gauche$^-$.

II. Dynamics Results. Overall the dynamical properties (i.e., the fluctuations of variables of interest) are similar in the native decamer and in the two cross-linked species. However, there are some differences, particularly in the region of the cross-links. Values are listed in Tables III–V and are illustrated in Figures 6–10.

Helical Structure. Cross-linking changes the fluctuations in twist and rise (Tables III, IV, and V). Constrained motion in twisting is observed throughout the central ATTC sequence in the $(C_2S)_2$ decamer, but the C_3 cross-linked tether does not induce such a consistent trend. There are no abrupt transitions from one value to another during the course of the simulation.

The magnitude of helical rise remains stable throughout the simulation. Fluctuations in the rise between base pairs A5–T16 and T6–A15 are increased somewhat by cross-linking, while fluctuations at the T6/T7 step are decreased by the C_3 cross-linked tether. The other rise parameters all have approximately the same fluctuations in the three decamers simulated.

The minor groove widths of the decamer are stable during the simulation, although some interstrand distances show much larger fluctuations than others (Figure 9). Groove widths in the cross-linked molecules are not so stable: there are very large fluctuations in distances P7–P18 and P8–P17 in the $(C_2S)_2$ decamer, and in the P8–P17 distance in the $(C_3S)_2$ decamer. In addition, there is an abrupt increase in the minor groove width P9–P16 in the presence of the C_3 cross-link at 10 ps, although the fluctuations around the averages before and after the transition are not particularly large. P6–P19 is the one distance which seems to fluctuate less in the cross-linked molecules than in the control.

The fluctuations in major groove width are comparable for all three systems. Differences in the root mean square deviation evident in Table VI result from a gradual drift in the parameter with time rather than a dramatic difference in fluctuation around the average value at any given time (e.g., the P3–P15 distance in the $(C_2S)_2$ decamer in Figure 9). There are also no significant differences in fluctuations of x-displacement in the cross-linked and control decamers. The y-displacements of the A5 and T6 base pairs are slightly more constrained in both cross-linked

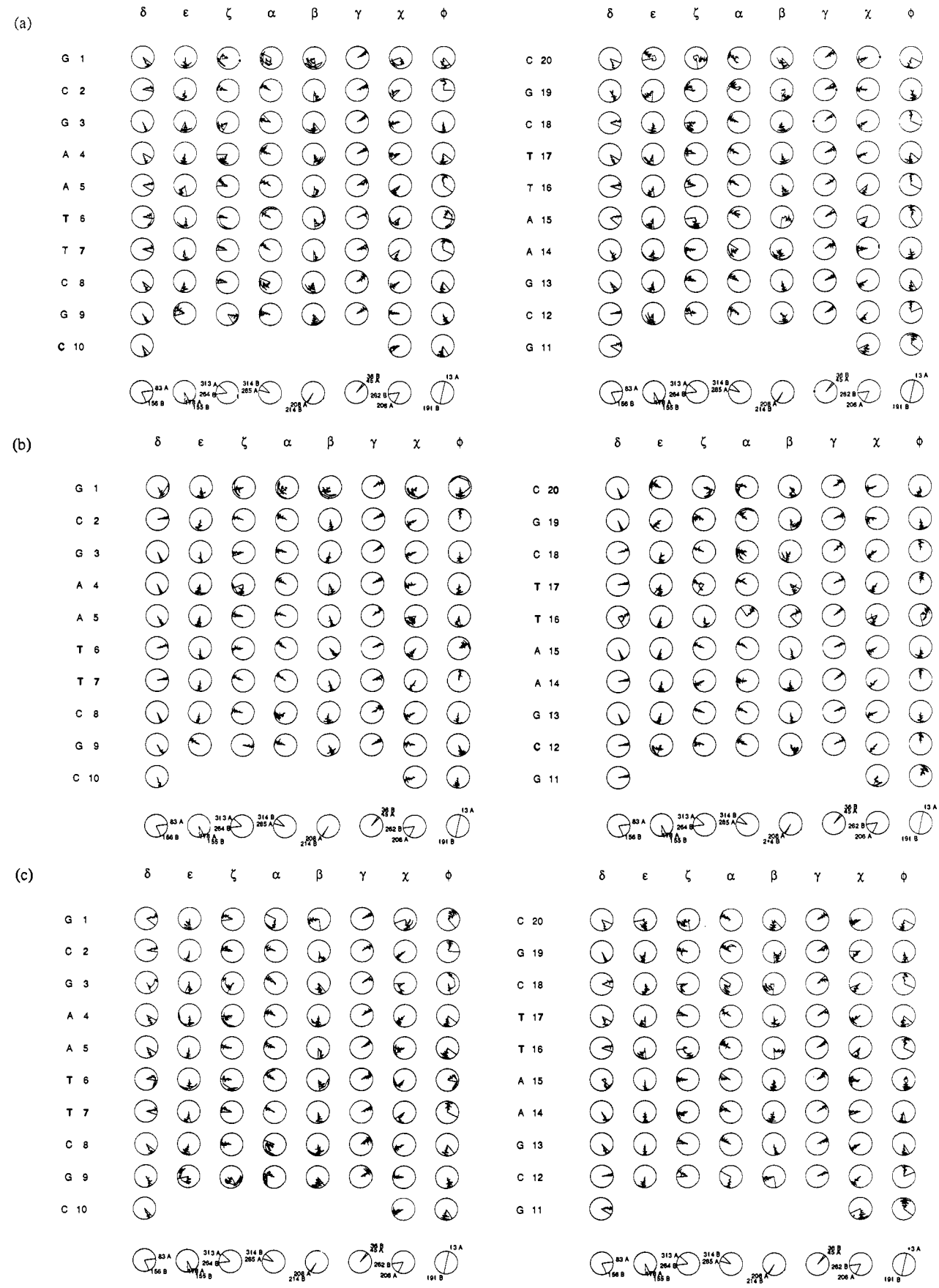

Figure 6. Dials representations of the backbone angles during molecular dynamics of the control decamer (a), $(C_2S)_2$ decamer (b), and $(C_3S)_2$ decamer (c). Plots were generated using the program Dials and Windows[29] and depict each torsional angle as a function of time, time being the radial coordinate with t = 0 ps at the center of the dial. The value in the crystal structure[15] is indicated in each dial by a solid line, and the values for canonical A- and B-DNA are shown in the dials at the bottom of each set of plots.

molecules than in the control, as is the y-displacement of base pair T7 in the $(C_2S)_2$ decamer.

Backbone Geometry. There are some differences in the behavior of the backbone parameters in the cross-linked decamers as

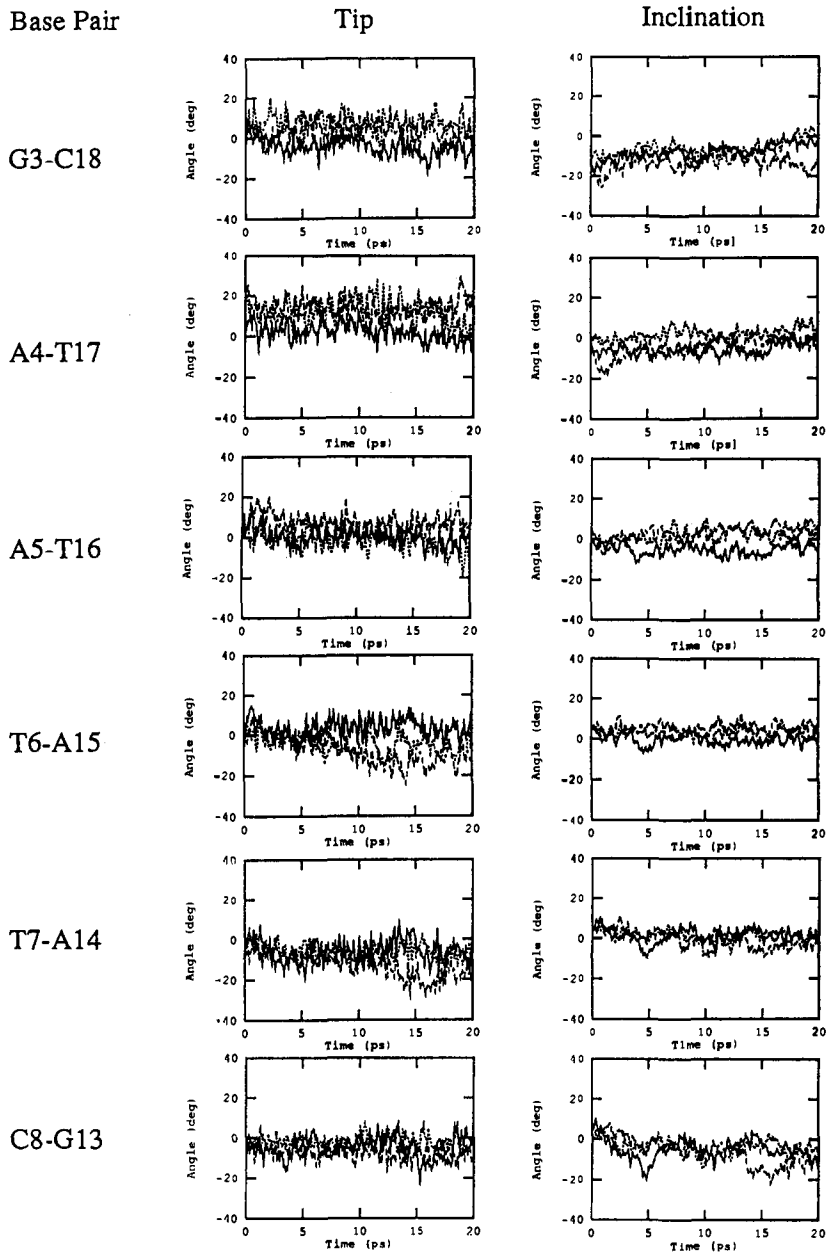

Figure 7. Sample time courses for the intrabase pair parameters tip and inclination (in deg). Data are shown for the decamer (solid line), $(C_2S)_2$ decamer (dotted line), and $(C_3S)_2$ decamer (dashed line).

compared to the control. In all three systems, occasional abrupt transitions are observed from one state to another in the backbone during the course of dynamics (e.g., at T6 in the control decamer; Figure 6). Several are transitions from C_3'-*endo* to C_2'-*endo* sugar pucker (reflected in transitions in δ, χ, and ϕ), and one (G3 in the $(C_3S)_2$ decamer) is a transition in the opposite direction, from C_2'-*endo* to C_3'-*endo*. Most of these changes are accompanied by local backbone distortions in ϵ, ζ, α, and β. In general, changes in α and β are correlated, an increase in α being accompanied by a decrease in β (e.g., $\alpha 19/\beta 19$ in the control decamer).

The fluctuations along the backbone do not vary dramatically. At the site of cross-linking, the glycosidic bond angle χ shows high fluctuations at A5 in both cross-linked duplexes and at A15 in the $(C_3S)_2$ decamer. A15 in the presence of either cross-link also shows a decrease in the fluctuations of α as compared to the control. A14, when adjacent to the C_2 cross-link, shows smaller fluctuations in χ, δ, and ϕ than in the control.

Base Pair Geometry. Throughout the molecular dynamics runs of all three molecules, there are no abrupt transitions in tip, roll, inclination, or tilt, but there are some cases in which the value of a parameter changes gradually over the course of the simulation (e.g., tip 6 and roll 5 in the $(C_3S)_2$ decamer; Figures 7 and 8). In terms of fluctuations, no significant differences in tip in the cross-linked duplexes versus the native sequence are apparent, and in all cases the fluctuations are larger than variations between the crystal structures of the Dickerson dodecamer (Table II). The roll angle exhibits no change in fluctuation across the site of cross-linking (roll A5/T6), but there are differences to either side of the cross-link: the C_2 cross-link induces a slight increase in fluctuation relative to the control at roll A4/A5 and a decreased fluctuation at roll T6/T7. The fluctuation of roll T7/C8 is smaller in both cross-linked decamers than in the control decamer.

Inclination and tilt also remain stable throughout the course of dynamics and show few differences in fluctuation between the cross-linked and control decamers. Again, there is no impact of cross-linking upon the A5/T6 step, but there are slight changes to either side of the cross-link: the fluctuation of tilt in A4–T17 is larger in the $(C_2S)_2$ decamer than in the control, and the tilt of T7–A14 shows slightly restrained fluctuations in both cross-linked species as compared to the control.

Fluctuations in propeller twisting and buckling within the central base pairs are affected by cross-linking. Propeller twisting

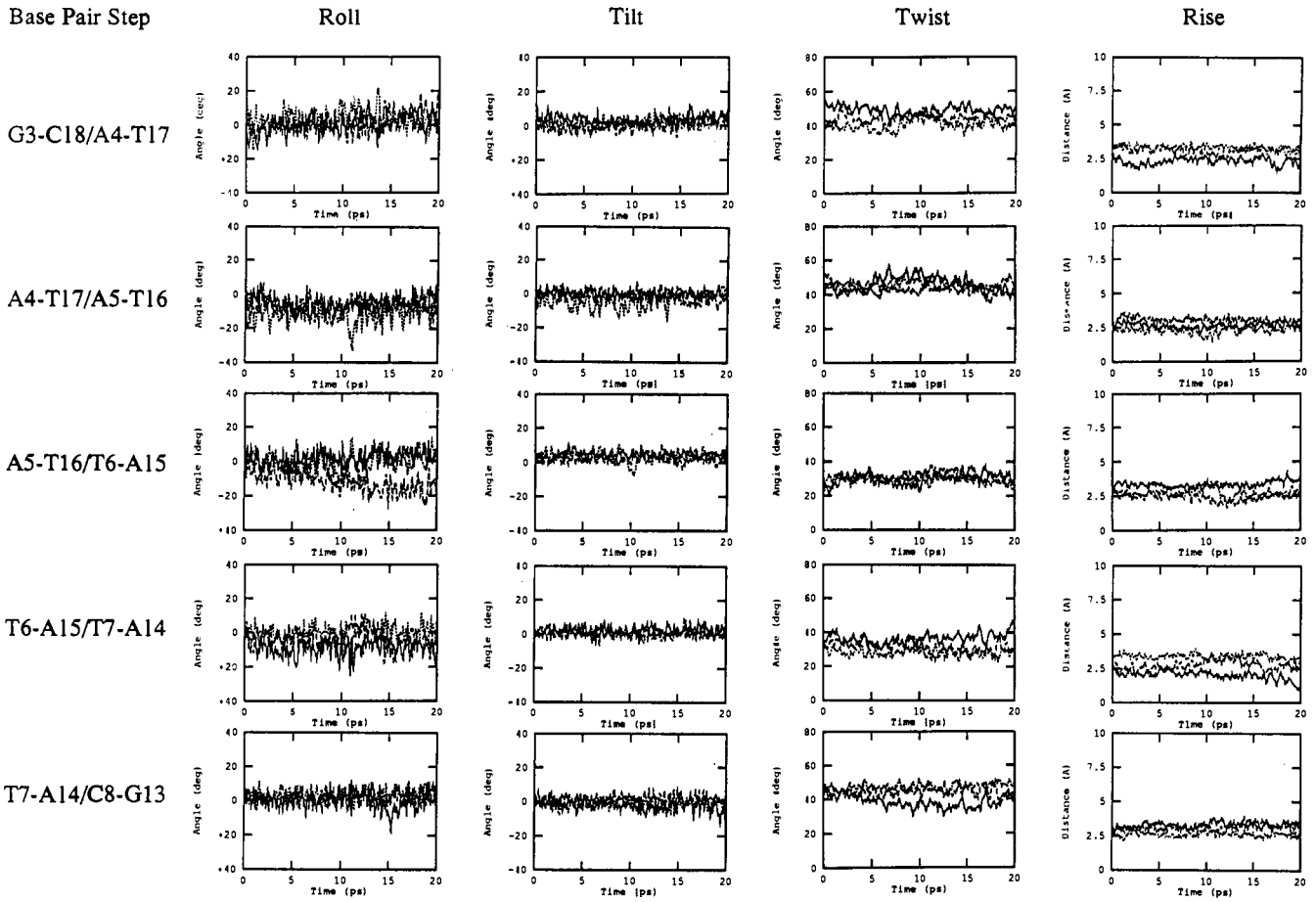

Figure 8. Sample time courses for the interbase pair parameters roll, tilt, twist, and rise (Å) in the decamer (solid line), $(C_2S)_2$ decamer (dotted line), and $(C_3S)_2$ decamer (dashed line).

exhibits high fluctuations in all the simulations, with some increased fluctuations in the AATT sequence in the presence of a cross-link. In terms of buckling, A5–T16 also shows higher fluctuations in both cross-linked molecules than in the control, but increased fluctuations in buckle at T6–A15 occur only in the $(C_2S)_2$ decamer. Meanwhile both cross-links constrain the buckle at T7–A14, but buckling of A4–T17 shows decreased fluctuation only with the C_2 cross-link.

Slide remains stable throughout all the simulations, and there are no significant differences in fluctuations between the cross-linked and control decamers.

Cross-Link Parameters. In general the C_3 cross-link exhibits higher fluctuations than the C_2 cross-link (Figure 10). The Cβ–Cγ torsional angles and the disulfide itself show particularly high fluctuations in the C_3 cross-link. The rms values (supplementary material) are on the order of magnitude of those found in protein side chains.

Discussion

Molecular dynamics simulations of three oligonucleotide decamers differing only in the presence of disulfide cross-links of two different lengths have been presented. During 20 ps of molecular dynamics at 300 K using the CHARMM potential energy function, all three structures remained stably Watson–Crick base paired. Two identical runs and analyses were performed for each system to obtain some measure of the significance of the differences among them (average structures from the duplicate runs are within 0.5–1.5 Å of each other). The atomic rms deviations of the average structure during the last 5 ps of dynamics from the starting structure (crystal coordinates), minimized structure, and canonical B- and A-DNA are summarized in Table I. All three duplexes remain in the B-form (within 2.3 Å) and near the starting structure (within 2.1 Å), with the average structures of the cross-linked decamers lying slightly closer to the initial coordinates than the control decamer. All three structures lie much farther away from A-DNA (at least 4.8 Å). In all cases, the deviations from the X-ray structure are smaller when one considers only the central AATT sequence rather than the whole molecule; the values are in the range 1.0–1.3 Å, as given in Table I. This is expected, since it has been established that the terminal two base pairs undergo anomalously large fluctuations during molecular dynamics.[33] Therefore, the following structural and dynamic analyses exclude the terminal GC residues and focus on the central AATT sequence within the decamers.

The present results from simulations for the decamer compare favorably with previously reported simulations using other force fields: simulation of the Dickerson dodecamer using GROMOS *in vacuo*[30] led to a structure 5 Å from the crystal structure or canonical B-DNA, while the addition of solvent[31] brought the structure within 2.3 Å of canonical B-DNA; simulation *in vacuo* using AMBER resulted in a structure 3.6 Å from the crystal coordinates.[29] It is striking that the present simulations with reduced charges on the phosphates and a distance-dependent dielectric constant (but no explicit solvent or counterions) yield average structures that are so close to the crystallographic results.

The rms differences between the average structures of the control decamer and the cross-linked decamers are 1.6 and 1.7 Å for the central AATT sequence in the $(C_2S)_2$ and $(C_3S)_2$ decamers, respectively. These regions of the average structures of the two cross-linked decamers are within 1.0 Å of each other. This may imply that cross-linking causes some structural distortions that are common to the two cross-links of different lengths. In particular, the two cross-linked decamers have very similar twist, propeller twist, and *y*-displacement in the central AATT region. The differences between the cross-linked and

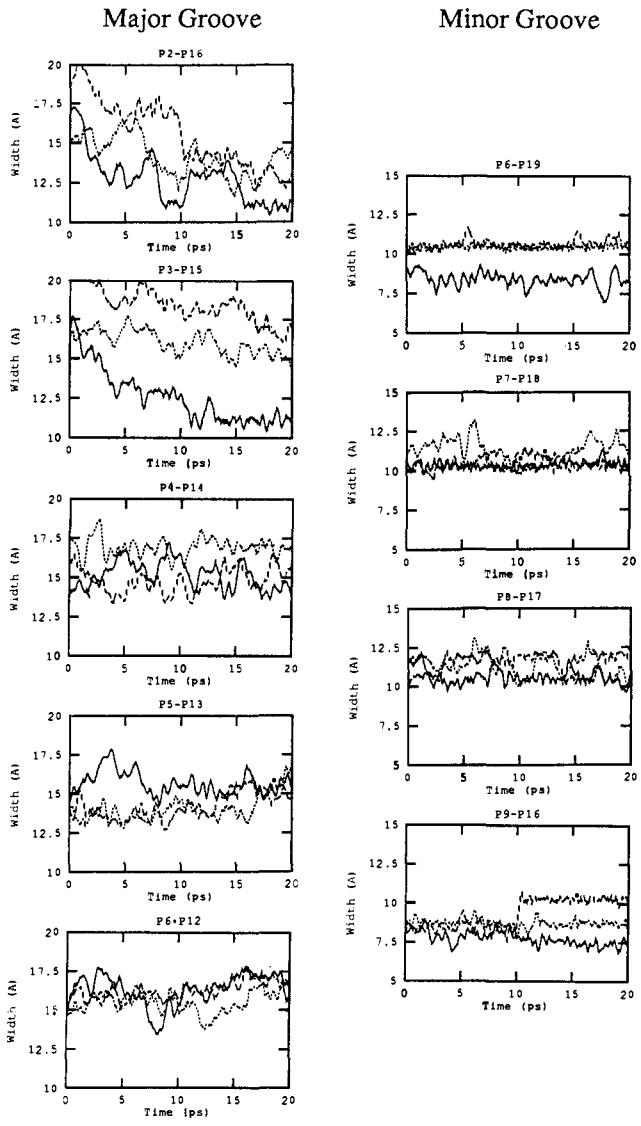

Figure 9. Sample time courses for the major and minor groove widths (Å) in the decamer (solid line), $(C_2S)_2$ decamer (dotted line), and $(C_3S)_2$ decamer (dashed line).

native duplexes suggests that there is a global constraint of motion in the presence of a cross-link, though many of the local parameters undergo similar fluctuations in the presence and absence of the cross-links.

Throughout the simulations, Watson–Crick hydrogen bonding is maintained, with most of the distances between donor and acceptor heavy atoms remaining around 3 Å (Table VII). This is in agreement with the observation of hydrogen-bonded imino protons at all the base pairs in the ^{1}H NMR spectra of these duplexes.[8] The hydrogen bonding at the central base pairs of the $(C_2S)_2$ decamer differs from that in the control decamer and the $(C_3S)_2$ decamer: all the average A–N^6 to T–O^4 distances as well as the A15–N1 to T6–N3 distance have been elongated and show larger fluctuations (Table VII). The control also contains an elongated, highly fluctuating A5–N^6 to T16–O^4 hydrogen bond. These changes are consistent with T6 being displaced so that it can form a bifurcated hydrogen bond to A14 in the presence of the C_2 cross-link (T6–O^4 is 3.5 Å from both A15–N^6 and A14–N^6; Figure 4e). To date, NMR spectroscopy has not yielded any experimental confirmation of this observation.

The overall helical parameters of the three duplexes remain near the B-form with respect to twist and rise. Twists range from 28° to 49°, with an average of 40°, somewhat larger than the value of 36° for canonical B-DNA and the crystal structures. This increased winding of the helix contrasts with the much lower values of twist (23–28°) derived from the solution structure[27] and lies a bit higher than the average values seen in other simulations (33–36°). The increased twisting in our simulations is accompanied by decreased average rise (2.9 Å). This value is smaller than that in the crystal structure (3.3 Å) or in canonical B-DNA (3.4 Å) and larger than the typical value of 2.6 Å in A-DNA. Thus, the values of rise from the simulations lie between the A- and B-forms of DNA.

Although the overall structures are B-like, certain parameters maintain values nearer those of canonical A-DNA than B-DNA. Inclination of the central four base pairs tends to be high, as in A-DNA, and the glycosidic bond angles, χ, tend to assume values found in A-form duplexes. The sugar puckers are almost half (17 out of 36) in the C_3'-*endo* conformation characteristic of A-DNA. Of these, ten are thymidine residues (only two thymidines remain in the C_2'-*endo* form), in agreement with the previously observed propensity of thymines for C_3'-*endo* puckering.[28] The high percentage of A-type sugar puckers can be compared to the results of simulations using GROMOS, in which the sugars adopt a pseudorotation angle between the C_2'-*endo* and C_3'-*endo* forms.[30]

The *x*-displacements of the base pairs deviate from the crystal structure in the opposite direction from A-DNA. Positive values of *x*-displacement have also been observed during *in vacuo* simulations using GROMOS, but in that case they were accompanied by collapsing of the minor groove.[30] The minor grooves in the structures described here remain open throughout dynamics, with widths in the range 7.7–11.4 Å, only slightly narrower than the groove in canonical B-DNA and a bit wider than the corresponding distances (P7–P18, P8–P17) in the crystal structure. The average major groove widths are also reasonable for B-DNA, though most are smaller than the canonical value of 17.5 Å. This narrowing of the major groove has been observed in simulations using GROMOS and AMBER as well but has generally been accompanied by widening of the minor groove.[28,29,31] In a visual comparison of the simulations and the crystal structure, these changes lead to an increased overall width of the helices in the former.

The backbone torsional angles fall into some previously observed patterns. Most of the (ϵ,ζ) values in the simulations reported here lie in the B_I (t,g-) or B_{II} (g-,t) conformations of ϵ and ζ.[44,45] The predominance of these forms has been supported experimentally,[19] and they have been observed in previous dynamics calculations.[28] Some transitions between these two conformations are observed in our simulations (see Results above) and at a few positions in the cross-linked decamers (ϵ,ζ) assumes a different, extended (t,t) conformation.

One type of backbone motion that has been noted in other dynamics calculations and in crystal structures, but is absent in the present simulations, is a "crankshaft" movement of α and γ.[28,30,46] Throughout all of the dynamics runs, γ remains remarkably stable as compared to results using other force fields. Transitions in α do occur, but are compensated for by changes in β, ζ, and ϵ rather than in γ. There are no significant trends that distinguish the cross-linked decamers from the control.

The base pair parameters are affected by the presence of the cross-links. The distance between the adenine amino groups (N^6) is increased from 3.0 Å in the control decamer to 3.8 Å in the $(C_2S)_2$ decamer and 3.7 Å in the $(C_3S)_2$ decamer. Tip, roll, and propeller twist all change upon cross-linking in order to accommodate this increased separation. Decreased propeller twisting at the cross-linked base pairs and a decreased roll angle between these base pairs (a result of positive tip in A5-T16 and negative tip in T6–A15) both serve to move the amino groups apart. This effect is seen clearly in the $(C_3S)_2$ decamer and to a lesser extent in the $(C_2S)_2$ decamer. These local distortions lead to a difference

(43) Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1988**, *6*, 63–91.
(44) Olson, W. K. *Nucl. Acids Res.* **1982**, *10*, 777–787.
(45) Prive, G.; Heinemann, U.; Chandrasegaran, S.; Kan, L.; Kopka, M.; Dickerson, R. E. *Science* **1987**, *238*, 498–504.

Table IV. Base Pair Parameters in $(C_2S)_2$ Decamer

Intrabase Pair Parameters												
	tip		inclination		*x*-displacement		*y*-displacement		propeller twist		buckle	
base pair	mean	(rms)	mean	(rms)	mean	(rms)	mean	(rms)	mean	(rms)	mean	(rms)
G1–C20	20.65	(84.65)	0.05	(2.45)	0.89	(0.86)	–0.24	(0.65)	–102.29	(10.56)	–5.67	(8.96)
C2–G19	–8.99	(5.66)	1.43	(4.59)	–0.01	(0.55)	–0.68	(0.65)	–78.09	(9.06)	4.65	(9.08)
G3–C18	7.22	(4.38)	–6.77	(4.24)	–1.77	(0.32)	–0.18	(0.43)	–28.93	(5.34)	4.75	(8.53)
A4–T17	12.38	(6.20)	1.75	(2.88)	–0.17	(0.37)	0.40	(0.31)	–10.07	(7.37)	–1.92	(10.05)
A5–T16	–1.80	(4.61)	2.23	(3.30)	–0.54	(0.49)	0.54	(0.27)	–9.58	(8.69)	11.94	(9.92)
T6–A15	–4.97	(4.55)	3.95	(2.42)	0.13	(0.46)	–0.48	(0.23)	–10.65	(6.62)	5.41	(9.83)
T7–A14	–6.15	(4.12)	1.00	(3.24)	–1.04	(0.43)	–0.03	(0.28)	–9.48	(7.36)	–14.89	(6.00)
C8–G13	–2.19	(2.99)	–3.94	(3.63)	–0.52	(0.34)	0.86	(0.44)	–24.87	(5.11)	–0.46	(7.00)
G9–C12	9.45	(4.13)	–2.15	(3.56)	0.16	(0.36)	0.93	(0.55)	–26.93	(5.10)	6.52	(7.49)
C10–G11	3.10	(7.04)	9.09	(5.60)	0.87	(0.74)	1.75	(0.58)	–19.99	(12.83)	–20.24	(15.96)

Interbase Pair Parameters										
	roll		tilt		twist		rise		slide	
base pair step	mean	(rms)	mean	(rms)	mean	(rms)	mean	(rms)	mean	(rms)
(G1–C20)/(C2–G19)	3.28	(48.35)	–0.74	(27.43)	51.24	(3.54)	1.30	(0.29)	0.23	(0.32)
(C2–G19)/(G3–C18)	16.16	(6.62)	–6.84	(3.53)	15.64	(4.60)	4.43	(0.22)	0.21	(0.27)
(G3–C18)/(A4–T17)	3.15	(5.63)	1.27	(2.70)	41.40	(3.60)	3.25	(0.19)	–0.23	(0.34)
(A4–T17)/(A5–T16)	–11.15	(5.62)	–4.27	(4.36)	45.58	(2.41)	2.16	(0.24)	–0.15	(0.32)
(A5–T16)/(T6–A15)	–0.94	(5.44)	3.48	(2.96)	29.34	(1.97)	2.92	(0.30)	–1.13	(0.28)
(T6–A15)/(T7–A14)	0.17	(4.01)	0.13	(1.71)	28.37	(2.00)	3.37	(0.19)	0.18	(0.23)
(T7–A14)/(C8–G13)	2.56	(3.33)	–1.30	(2.35)	46.79	(2.20)	2.72	(0.17)	0.12	(0.16)
(C8–G13)/(G9–C12)	9.57	(3.70)	–0.41	(1.78)	34.73	(2.43)	3.80	(0.18)	–0.19	(0.21)
(G9–C12)/(C10/G11)	–2.72	(5.78)	4.76	(3.36)	47.33	(4.34)	2.22	(0.27)	1.06	(0.48)

Table V. Base Pair Parameters in $(C_3S)_2$ Decamer

Intrabase Pair Parameters												
	tip		inclination		*x*-displacement		*y*-displacement		propeller twist		buckle	
base pair	mean	(rms)	mean	(rms)	mean	(rms)	mean	(rms)	mean	(rms)	mean	(rms)
G1–C20	–54.36	(63.33)	3.20	(4.79)	1.00	(1.28)	–0.25	(0.61)	–77.38	(8.54)	–3.24	(9.43)
C2–G19	–13.99	(7.87)	–9.75	(5.08)	–0.58	(0.65)	–0.73	(0.99)	–65.15	(4.44)	4.88	(8.56)
G3–C18	3.53	(5.03)	–13.39	(4.16)	–1.24	(0.67)	–0.21	(0.66)	–17.23	(5.65)	–5.41	(7.85)
A4–T17	13.76	(4.62)	–5.12	(4.32)	–0.10	(0.46)	0.39	(0.43)	–8.42	(7.09)	–0.26	(7.91)
A5–T16	7.37	(4.34)	2.73	(3.31)	0.08	(0.37)	0.52	(0.23)	–8.47	(7.80)	–0.78	(9.35)
T6–A15	–7.52	(7.34)	4.85	(2.67)	1.21	(0.57)	–0.36	(0.29)	–7.93	(6.02)	2.85	(8.57)
T7–A14	–12.41	(6.72)	–0.24	(4.19)	–0.16	(0.42)	–0.73	(0.45)	–4.64	(6.54)	–9.34	(7.99)
C8–G13	–7.52	(3.38)	–6.27	(6.60)	–0.43	(0.34)	0.02	(0.79)	–17.40	(5.20)	–1.10	(7.49)
G9–C12	12.21	(4.35)	–6.48	(6.31)	–0.38	(0.51)	0.29	(1.00)	–26.37	(5.50)	–3.15	(8.57)
C10–G11	5.68	(8.79)	15.48	(5.51)	0.77	(1.16)	0.50	(0.63)	–36.27	(17.42)	–34.13	(14.55)

Interbase Pair Parameters										
	roll		tilt		twist		rise		slide	
base pair step	mean	(rms)	mean	(rms)	mean	(rms)	mean	(rms)	mean	(rms)
(G1–C20)/(C2–G19)	30.78	(21.57)	7.27	(22.37)	49.33	(3.88)	2.63	(0.33)	–0.42	(0.34)
(C2–G19)/(G3–C18)	14.76	(7.92)	–0.73	(2.44)	15.94	(6.34)	4.39	(0.28)	–0.47	(0.33)
(G3–C18)/(A4–T17)	2.66	(5.49)	1.24	(3.40)	43.17	(2.82)	3.14	(0.25)	–0.43	(0.32)
(A4–T17)/(A5–T16)	–6.70	(4.21)	–0.09	(2.48)	43.39	(3.00)	3.03	(0.22)	0.04	(0.34)
(A5–T16)/(T6–A15)	–12.00	(5.51)	2.33	(2.40)	31.26	(3.17)	2.44	(0.30)	–0.30	(0.35)
(T6–A15)/(T7–A14)	–3.23	(4.92)	0.62	(1.96)	31.28	(3.63)	2.74	(0.34)	0.07	(0.33)
(T7–A14)/(C8–G13)	2.11	(3.04)	1.51	(2.17)	44.64	(2.49)	3.15	(0.26)	0.32	(0.27)
(C8–G13)/(G9–C12)	15.98	(4.61)	–1.19	(2.34)	31.64	(2.38)	3.90	(0.23)	–0.29	(0.24)
(G9–C12)/(C10/G11)	–1.93	(5.32)	12.47	(3.75)	50.41	(2.44)	2.59	(0.27)	0.28	(0.25)

in the bending of the helical axis near the center of the duplex (Figure 4), but it is not clear whether this is translated into a global bending of the whole oligonucleotide or whether there are compensatory changes near the ends of the DNA.

It is possible that the observed structural changes are a consequence of appending a tether to the adenines and not a result of their being cross-linked. In the crystal structure of the Dickerson dodecamer that has been methylated at the central adenines, some separation of the modified bases was introduced by a decrease in twist and a local lowering in roll across the methylated 5′-AT-3′ sequence.[46] However, solution studies do not indicate a significant difference between the methylated and unmethylated dodecamers.[47]

The structures of the cross-links themselves are largely dictated by the disulfide bond. The torsional angles around the central four atoms are determined by the optimal angles near a disulfide (a ± 90° torsional angle around the disulfide and 90–135° angles around the S–C bonds). Since there is room for the cross-links in the major groove, the rest of the angles in the C_3 cross-link are free to assume extended conformations. In the $(C_2S)_2$ decamer, one Cα–Cβ torsional angle is gauche$^-$ because a trans configuration could not be accommodated into the helix without altering the position of one of the cross-linked adenines. In agreement with this constraint on the orientation of the C_2 cross-link, the

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

(47) Rinkel, L. J.; van der Marel, G. A.; van Boom, J. H.; Altona, C. *Eur. J. Biochem.* **1987**, *163*, 275–286.

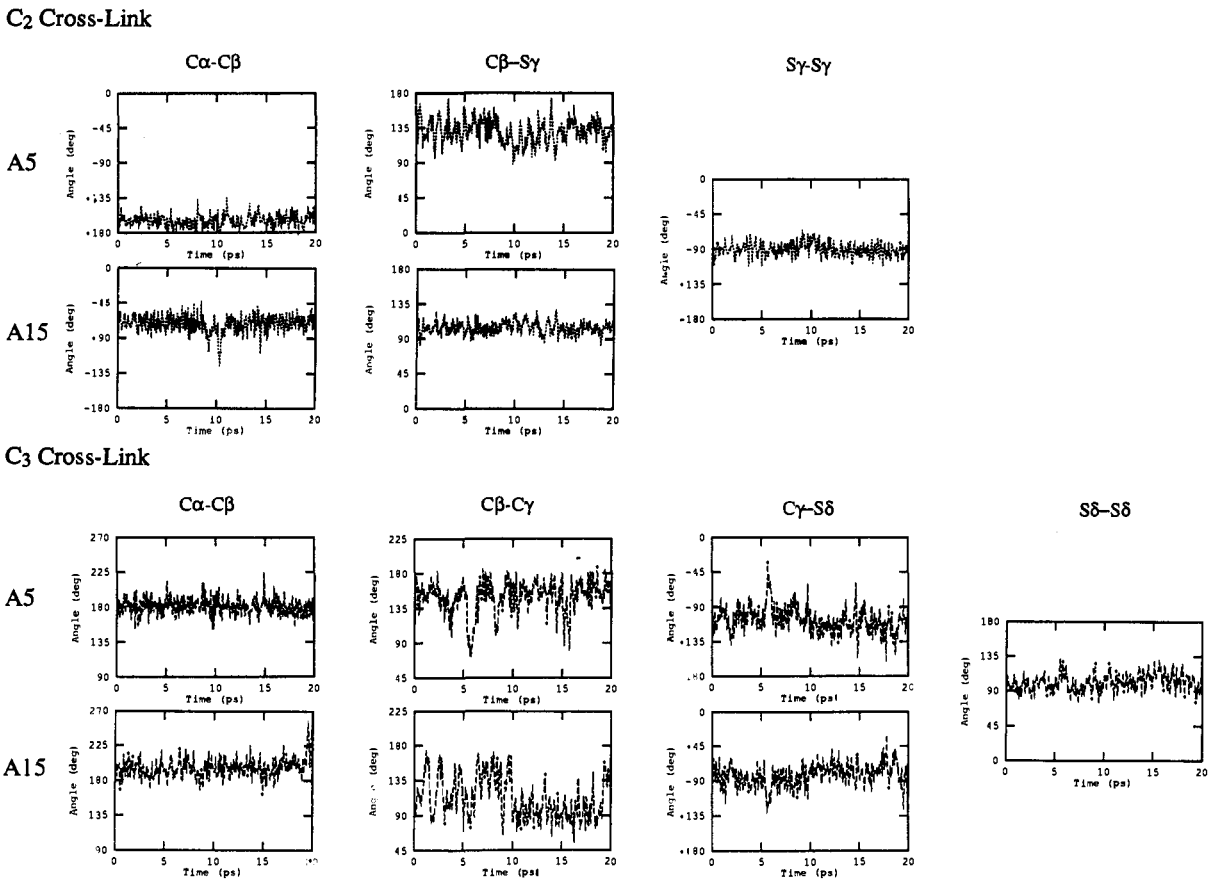

Figure 10. Sample time courses for the torsional angles in the cross-links.

Table VI. Groove Widths in Decamers[a]

		control decamer		$(C_2S)_2$ decamer		$(C_3S)_2$ decamer	
distance	crystal	mean	(rms)	mean	(rms)	mean	(rms)
			Minor Groove				
P5–P20	12.99	10.23	(0.19)	13.41	(1.04)	14.34	(0.75)
P6–P19	11.06	8.38	(0.46)	10.46	(0.19)	10.60	(0.27)
P7–P18	9.99	10.33	(0.23)	11.28	(0.71)	10.42	(0.37)
P8–P17	9.77	10.62	(0.49)	11.42	(0.69)	11.45	(0.65)
P9–P16	8.90	7.75	(0.43)	8.71	(0.24)	9.33	(0.97)
P10–P15	11.05	10.36	(0.27)	10.31	(0.28)	10.57	(0.56)
A-DNA	16.8						
B-DNA	11.5						
			Major Groove				
P6–P12	18.48	16.31	(1.01)	15.48	(0.68)	16.27	(0.59)
P5–P13	17.37	15.67	(0.73)	13.99	(0.76)	14.23	(0.83)
P4–P14	16.78	15.14	(0.86)	16.91	(0.59)	14.77	(0.77)
P3–P15	17.53	12.57	(1.67)	16.05	(0.75)	18.52	(1.22)
P2–P16	17.32	12.73	(1.56)	14.17	(1.30)	15.58	(2.14)
A-DNA	8.5						
B-DNA	17.5						

[a] The intraphosphorus distances tabulated are longer than the frequently cited groove widths by 5.8 Å, twice the van der Waals radius of phosphorus.

Table VII. Hydrogen Bond Lengths in the Three Decamers

		control decamer		$(C_2S)_2$ decamer		$(C_3S)_2$ decamer	
base pair	H-bond	mean	(rms)	mean	(rms)	mean	(rms)
C2–G19	N^2G–O^2C	2.83	(0.09)	2.85	(0.10)	2.83	(0.09)
	N1G–N3C	3.03	(0.14)	2.97	(0.10)	3.04	(0.11)
	O^6G–N^4C	3.24	(0.29)	3.10	(0.22)	3.13	(0.21)
G3–C18	N^2G–O^2C	2.92	(0.14)	2.90	(0.11)	2.89	(0.13)
	N1G–N3C	2.99	(0.10)	2.98	(0.10)	3.03	(0.12)
	O^6G–N^4C	3.02	(0.18)	2.92	(0.14)	3.05	(0.22)
A4–T17	N^6A–O^4T	3.19	(0.23)	3.91	(0.35)	2.98	(0.16)
	N^1A–N^3T	2.91	(0.09)	2.98	(0.12)	2.98	(0.11)
A5–T16	N^6A–O^4T	3.76	(0.47)	3.49	(0.41)	3.15	(0.21)
	N^1A–N^3T	2.95	(0.14)	2.94	(0.11)	2.95	(0.10)
T6–A15	N^6A–O^4T	3.16	(0.25)	4.32	(0.56)	3.18	(0.22)
	N^1A–N^3T	2.96	(0.12)	3.57	(0.33)	2.98	(0.14)
T7–A14	N^6A–O^4T	3.09	(0.19)	3.93	(0.34)	3.18	(0.21)
	N^1A–N^3T	2.92	(0.10)	2.96	(0.10)	2.88	(0.09)
C8–G13	N^2G–O^2C	3.09	(0.25)	2.93	(0.14)	3.03	(0.20)
	N1G–N3C	3.13	(0.18)	3.02	(0.11)	3.09	(0.14)
	O^6G–N^4C	3.07	(0.22)	3.07	(0.20)	3.03	(0.18)
G9–C12	N^2G–O^2C	2.84	(0.10)	2.81	(0.08)	2.84	(0.10)
	N1G–N3C	3.04	(0.19)	2.99	(0.10)	2.98	(0.12)
	O^6G–N^4C	3.13	(0.36)	3.07	(0.20)	3.06	(0.24)

$(C_2S)_2$ decamer exhibits smaller fluctuations than does the $(C_3S)_2$ decamer. Outside of the cross-link itself, the glycosidic bond angles for the cross-linked adenines show large fluctuations that could result from attempting to accommodate the cross-link within the structure. The relatively constrained motion in the $(C_2S)_2$ decamer is reflected in a constraint on the twisting between the central two base pairs. No constraint is observed in other modes of motion, such as fluctuations in roll and tilt.

Conclusions

The structural features and dynamics of two disulfide cross-linked decamers have been studied and compared to the native decamer of the same sequence using molecular dynamics simulations with the CHARMM 19 potential energy function[35] and a Manning-type shielding of the phosphates.[32] The base pair, backbone, and helical parameters have been analyzed. All three decamers remain in the B-form throughout the simulations. The duplexes are Watson–Crick base paired, with the possible addition of a bifurcated hydrogen bond at T6 in the $(C_2S)_2$ decamer. Proper base pairing is in agreement with the observation of base paired imino protons in the ^1H NMR spectrum of the cross-linked decamers.[8] The molecules are all asymmetric on the picosecond time scale of dynamics, though they are observed to be symmetric on the NMR time scale;[8] the crystal structures of the Dickerson/Drew dodecamer are also asymmetric, indicating that such a distortion requires only relatively weak perturbations.

The major and minor grooves are maintained despite the absence of explicit solvent molecules in the simulation. The average structures from the last 5 ps of the 20-ps dynamics runs do not differ radically from one cross-linked species to the other, nor do the cross-linked decamers differ much from the control. There are some local distortions by both cross-links that lead to an increased distance between the amino groups of the modified adenines. These include lower roll, decreased propeller twisting, and more negative x-displacement. The only significant constraint on the motion is a lower fluctuation in twisting in the region near the C_2 cross-link. This is not accompanied by constraints on other degrees of freedom. Thus, the present simulations suggest that a cross-linked oligonucleotide has been designed whose average structure is close to that of the native sequence and whose dynamics show selective constraints on the twisting motion within DNA without affecting its flexibility in terms of bending (roll and tilt). Such a molecule could be useful in solving questions concerning the requirements for DNA flexibility in protein–DNA or drug–DNA complexation.

Acknowledgment. We thank Prof. B. Reid for kindly providing the coordinates for the NMR solution structure of the dodecamer d(CGCGAATTCGCG)$_2$. This work was supported by grants from NIH (to G.L.V.), NSF (Presidential Young Investigator program), and the Chicago Community Trust (Searle Scholars program).

Supplementary Material Available: Tables of the mean values and rms fluctuations for the backbone angles in all three duplexes and the torsional angles in the cross-links (4 pages). Ordering information is given on any current masthead page, or the material may be obtained from the authors by FAX at (617)495-8755.