Structure of the Potassium Form of CGCGAATTCGCG: DNA Deformation by Electrostatic Collapse around Inorganic Cations[†,‡]

Xiuqi Shui, Chad C. Sines, Lori McFail-Isom, Don VanDerveer, and Loren Dean Williams*

School of Chemistry & Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332-0400

Received August 25, 1998; Revised Manuscript Received October 1, 1998

ABSTRACT: The potassium form of d(CGCGAATTCGCG) solved by X-ray diffraction to 1.75 Å resolution indicates that monovalent cations penetrate the primary and secondary layers of the "spine of hydration". Both the sodium [Shui, X., McFail-Isom, L., Hu, G. G., and Williams, L. D. (1998) *Biochemistry 37*, 8341−8355] and the potassium forms of the dodecamer at high resolution indicate that the original description of the spine, only two layers deep and with full occupancy by water molecules, requires substantive revision. The spine is merely the bottom two layers of a four layer solvent structure. The four layers combine to form a repeating motif of fused hexagons. The top two solvent layers were not apparent from previous medium-resolution diffraction data. We propose that the narrow minor groove and axial curvature of A-tract DNA arise from localization of cations within the minor groove. In general, the results described here support a model in which most or all forces that drive DNA away from canonical B-conformation are extrinsic and arise from interaction of DNA with its environment. Intrinsic forces, originating from direct base−base interactions such as stacking, hydrogen bonding, and steric repulsion among exocyclic groups appear to be insignificant. The time-averaged positions of the ubiquitous inorganic cations that surround DNA are influenced by DNA bases. The distribution of cations depends on sequence. Regions of high and low cation density are generated spontaneously in the solvent region by heterogeneous sequence or even within the grooves of homopolymers. The regions of high and low cation density deform DNA by electrostatic collapse. Thus, the effects of small inorganic cations on DNA structure are similar to the effects of proteins.

DNA conformation is heterogeneous, not uniform and repetitive as in Watson and Crick's fiber structure. The accepted view of conformational heterogeneity, articulated in recent reviews (*1*−*3*), is that two independent classes of forces perturb DNA. In this model, one class of force is intrinsic to DNA and the other is extrinsic. Intrinsic forces, originally proposed by Klug and co-workers (*4*), derive from direct base−base interactions such as stacking, hydrogen bonding, and steric repulsion among exocyclic groups. Intrinsic forces cause spontaneous bends of some DNA sequences (*5*−*9*) and other deformations. Extrinsic forces, by contrast, arise from interaction of DNA with its environment, for example, when the self-repulsive phosphate system collapses around cationic proteins. Extrinsic forces applied to DNA by TATA element binding protein (*10*) and catabolite activator protein (*11*) cause axial bends.

This internal/external duality is challenged by new models in which most or all forces that drive DNA from B-conformation are extrinsic and originate from interaction of DNA with its environment. In the model supported here, the time-averaged positions of the ubiquitous inorganic cations and amines that surround DNA are influenced by DNA bases. The local concentration of cations then depends on sequence. Regions of high and low cation density are generated spontaneously in the solvent region by heterogeneous sequence or in the grooves of homopolymers. Peaks and troughs of cation density deform DNA by electrostatic collapse. In this model, the effects of inorganic cations and amines are similar in kind to the effects of proteins.

Present structural support for electrostatic collapse around inorganic cations is provided by a series of experiments that independently suggest that cations penetrate and bind specifically in the minor groove of A-tract DNA. Hud and Feigon (*12*) used NMR results to support a proposal of axial curvature of A-tracts mediated by specific binding of divalent cations within the minor groove. Young and Beveridge (*13*) used molecular dynamics results to support a proposal of monovalent cation mediation of additional conformational features such as minor groove width. Those modeling experiments are directly supported by our X-ray crystallographic results (*14*). Rouzina and Bloomfield (*15*) used semiquantitive estimates of free energies and an adiabatic approximation to suggest that small multivalent cations can bend DNA by a purely electrostatic mechanism.

A necessary first step for understanding the structural origins of DNA deformation is the unambiguous identification of DNA-bound solvent as water molecules, cations, or hybrids (with partial occupancy by both). A hybrid is observed crystallographically when a population of symmetry equivalent sites is occupied fractionally by cations and fractionally by water molecules. The coordination geometry and scattering characteristics of a hybrid are weighted by

* Address correspondence to this author.

[†] This work was supported by the National Science Foundation (Grant MCB-9056300) and the American Cancer Society (Grant RPG-95-116-03-GMC).

[‡] Atomic coordinates and structure factors have been deposited in the NDB (sodium form entry code BDL084; potassium form entry code BD0005) and PDB.

10.1021/bi982063o CCC: $15.00 © 1998 American Chemical Society
Published on Web 11/11/1998

the fractional occupancies. A sodium−water hybrid is difficult to distinguish from a water molecule, especially if the sodium occupancy is less than 50%. Sodium ions and water molecules (i) carry the same number of electrons, giving electron density peaks with similar volumes, (ii) both have pliable and unpredictable coordination geometries, and (iii) as described here, form hybrids with greater water than sodium occupancy. Our results indicate that some solvent sites traditionally thought to be water molecules are in fact water-monovalent cation hybrids.

X-ray structures of DNA oligonucleotides have given the incorrect impression that nucleic acid molecules are surrounded by neutral aqueous solvent. The majority of the cationic charge expected from charge neutrality appears missing around a DNA dodecamer, even with very high-resolution data. Although well over 100 water molecules can be observed bound to a sodium form of d(CGCGAATTCGCG) refined at 1.4 Å resolution, not a single fully occupied monovalent cation is apparent (*14*) and the majority of the cationic charge required to neutralize the DNA phosphate groups is absent.

We previously proposed a hybrid solvent model, in which DNA solvent sites are occupied by water−cation hybrids (*14*). Each primary site in the spine of hydration of A-tracts appears to be occupied partially by water molecules and partially by monovalent cations. In solution, the fractional occupancy of each site would depend on sequence and on concentrations of sodium and other cations, minor groove binders, proteins, etc. To test the hybrid solvent model and the hypothesis of monovalent cation localization within the minor groove of A-tract DNA, we have conducted potassium substitution experiments. Potassium has an advantage over sodium for identification of monovalent cation-binding sites. A potassium ion, by its larger number of electrons, can be readily distinguished from a water molecule by X-ray diffraction. Potassium substitution for sodium provides a definitive experimental test of the proposal that monovalent cations bind within the minor groove.

The potassium substitution results described here confirm that monovalent cations penetrate the primary layer of the spine. The results further indicate that monovalent cations reside within the secondary layer of the spine. In addition, we report that, in both the sodium and potassium forms, the spine of hydration is merely the bottom two layers of a four-layer structure. The top two solvent layers were not apparent from previous medium-resolution diffraction data (*16*−*18*). In high resolution determinations of the structure of d(CGCGAATTCGCG) (sodium form, NDB structure BDL084; potassium form, NDB structure BD0005), four layers of minor groove hydration (primary, secondary, tertiary, and quaternary) are clearly evident. The four layers combine to form a repeating motif of fused hexagons (Figures 1 and 2). In sum, the original description of a minor groove solvent structure, only two layers deep and with full occupancy by water molecules, requires substantive revision.

MATERIALS AND METHODS

Crystallization, data collection and reduction, and refinement of the high-resolution sodium form of d(CGCGAATTCGCG) were described previously (*14*).

Crystallization of the Potassium Form. The ammonium salt of reversed-phase HPLC-purified d(CGCGAATTCGCG) was purchased from the Midland Certified Reagent Company (Midland, TX). Crystals were grown at 22 °C from sitting drops that initially contained 1.0 mM d(CGCGAATTCGCG), 19 mM potassium cacodylate (pH 6.5), 10 mM $MgCl_2$, 4.8% 2-methyl-2,4-pentanediol (MPD), and 8.9 mM spermine tetrahydrochloride. The droplets were equilibrated by vapor diffusion against a reservoir of 50% MPD. Orthorhombic ($P2_12_12_1$) crystals appeared within several days. The crystal chosen for data collection was 1.3 × 0.3 × 0.15 mm. The crystal was plucked up in a loop and shock cooled by direct transfer into a −110 °C N_2 stream bathing the goniostat.

Data Collection and Reduction. X-ray intensity data were collected at low temperature (−110 °C) with a Siemans CCD smart detector and monochromated Mo Kα radiation (λ = 0.7107 Å).

Refinement. Excluding solvent molecules, coordinates of the high-resolution sodium form of d(CGCGAATTCGCG) were used for the starting model for the potassium structure. The structure was annealed and refined with XPLOR (version 3.851, ref *19*), using parameter file dna-rna-multi-endo.param (*20*, *21*). Xplor was also used to calculate sum ($2F_o - F_c$) and difference ($F_o - F_c$) Fourier maps and to determine difference peak intensities. The potassium model was refined against all data between 10 and 1.5 Å resolution (6479 unique reflections). $|F_o|$ were not scaled to $|F_c|$ by any means other than a single isotropic scale factor. The data collection and refinement statistics for the fully refined model are given in Table 1. Electron density maps of the DNA are sharp, clean, and continuous (not shown). The intensities of difference peaks in Table 2 are from the final model with all solvent molecules removed. The final model contains water molecules, one hydrated magnesium ion, and a partial spermine, but no monovalent cations. Therefore, the model is not biased toward monovalent cation occupancy within the minor groove.

RESULTS

The minor groove of A-tract DNA supports a fused hexagon motif of solvent sites (Figure 1). The primary layer of the motif is denoted by P, the secondary layer by S, the tertiary layer by T and the quaternary layer by Q. The positions of solvent sites of the motif are unambiguous in high-resolution electron density maps of the sodium and potassium forms (Figure 2). The fused hexagon motif is composed of a monolayer of 18 sites on a smooth but curved surface. The surface is not wrinkled as predicted by the tetrahedral coordination geometry preferred by water molecules or by the octahedral coordination geometry preferred by monovalent cations. Each hexagon is very nearly planar (Table 1S, Supporting Information). Each site is a member of one or two hexagons. Much of the curvature is between rather than within hexagons. The fused hexagons splay on the curved floor of the groove, such that the minor groove bound edges of the hexagons are compacted and the exposed edges are elongated. The elongation is most pronounced at the center of the motif.

Monovalent cations bind in the minor groove. A series of self-consistent indications of hybrid cation−water solvent within the minor groove of the A-tract of d(CGCGAATTCGCG) is given by potassium difference peaks, thermal factors, sodium specific valencies, and coordination geometry.

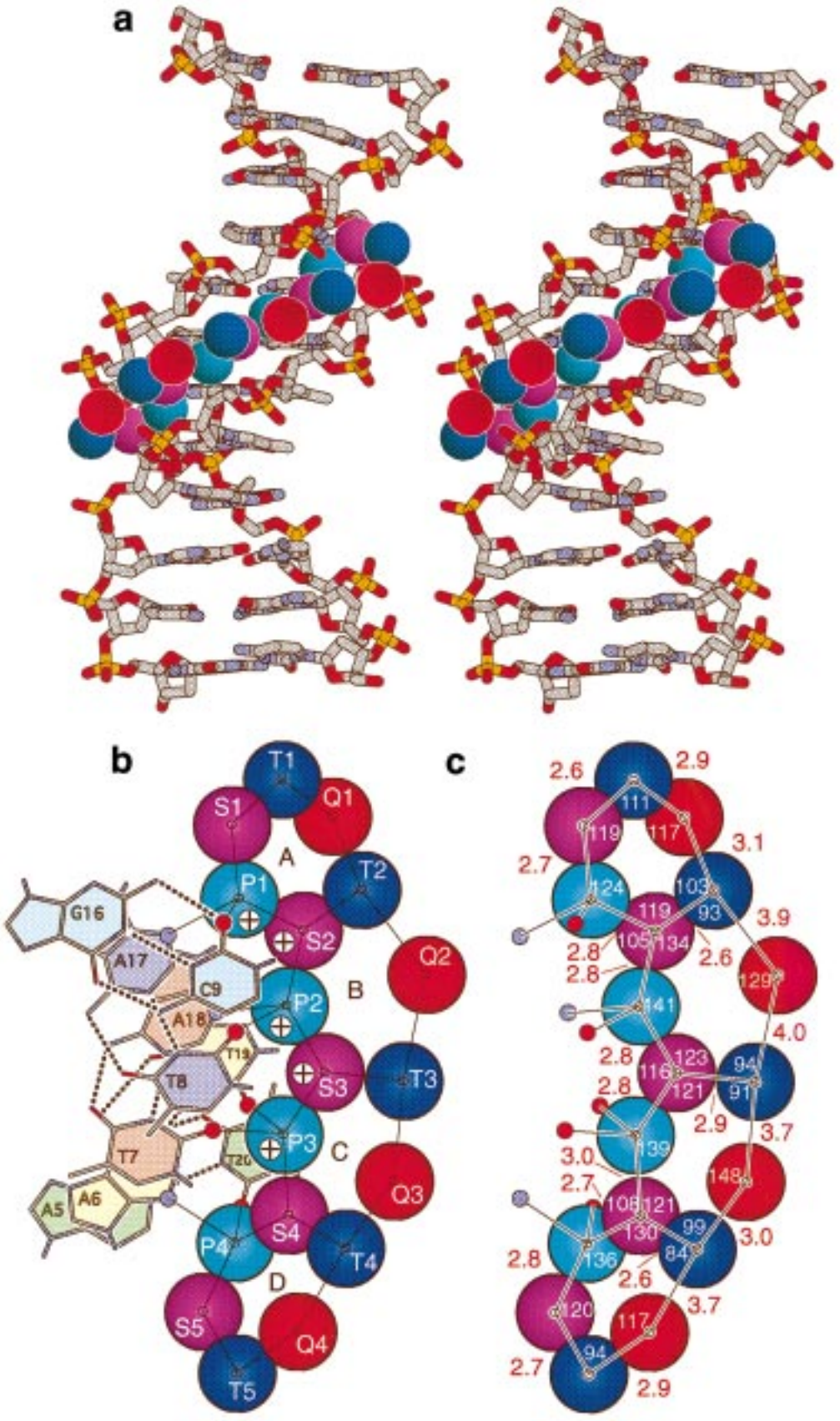

FIGURE 1: The fused hexagon motif of A-tract DNA. The four layers are coded by color with the primary layer light blue, the secondary layer magenta, the tertiary layer blue, and the quaternary layer red. The fused hexagon motif is shown in space filling representation, with van der Waal radii of oxygen atoms. (a) Stereoview into the minor groove of the DNA. The DNA is colored by CPK and shown in stick representation. (b) View across the groove, approximately down the normal of the central hexagons. Sites of potassium occupancy are indicated by plus signs. The DNA bases are shaded. Base functional groups that interact with the fused hexagon motif are indicated by circles. (c) The geometry of the sodium form fused hexagon motif. Distances are in red and angles are in white.

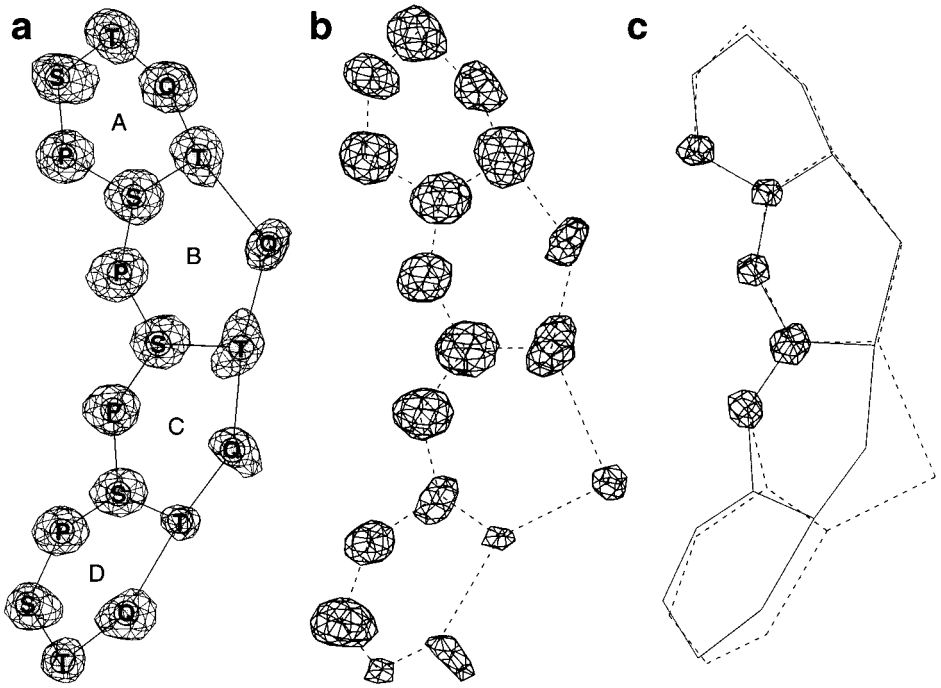

FIGURE 2: (a) Sodium form (1.4 Å) Fourier sum map ($2F_o - F_c$) contoured at 1σ surrounding sites that compose the fused hexagon motif. Solid lines connecting the sites indicate molecular interactions within the motif, either hydrogen bonds or cation−ligand interactions. (b) Potassium form (1.75 Å) Fourier sum map contoured at 1σ surrounding the members of the fused hexagon motif. Dashed lines connect the sites. (c) The potassium form Fourier difference maps ($F_o - F_c$), contoured at 4.5σ indicating the sites of highest potassium occupancy. Both the sodium (solid) and potassium (dashed) form connectivities are indicated.

Partial potassium occupancy is indicated by difference peak intensity. We are able to directly calculate potassium difference maps because the potassium and sodium forms of the DNA duplex d(CGCGAATTCGCG) crystallize in the same crystal form and unit cell and with similar DNA conformation. We have collected high-resolution X-ray intensity data from both the sodium form (1.4 Å resolution) and potassium form (1.75 Å resolution). A potassium form DNA model was obtained by simulated annealing and refinement of the sodium structure of d(CGCGAATTCGCG), lacking all solvent, against the potassium form data. Sum Fourier ($2F_o - F_c$, where F_o indicates observed potassium form structure factor amplitudes and F_c indicates calculated structure factor amplitudes from the potassium model) peaks within the minor groove indicate that the geometry of fused hexagon motif is conserved in the potassium and sodium forms. To date, we have obtained better diffraction data from the sodium form than from the potassium form. The general features of the sodium form and potassium form fused hexagons are similar (Figure 2). The electron density maps of the higher resolution sodium form are sharper, more geometric, and no doubt more accurate. Thus, the specific geometric descriptions here refer to the sodium form.

Intense peaks within the fused hexagon motif are observed in the potassium form Fourier difference maps (Figure 2c). The intensities of some of these peaks are greater than expected for water molecules. The intensities of each difference peak within the fused hexagon motif are given in Table 2. The most reasonable explanation for the intensity of difference peaks is partial occupancy by potassium ions.

A useful reference for intensities of difference peaks is provided by the intensity of the magnesium ion peak. In the major groove of the potassium form, a difference peak (intensity 7.7σ) is located at the position of a magnesium ion observed previously in the sodium form. This difference peak and a surrounding octahedra of less intense difference peaks confirm that magnesium binds at the same site in the sodium and potassium forms. A water molecule, with a relatively diffuse electron cloud, is expected to give a significantly less intense difference peak than a magnesium ion. The intensities of the difference peaks that we have identified as partially occupied potassium ions range from 6.9−6.2σ, and are nearly as intense as the magnesium ion peak. With the exception of a single site in the major groove, adjacent to O2P of G2, other difference peaks within the map are significantly less intense than the magnesium peak so that additional sites of occupancy by potassium cannot be definitively established.

Within the primary layer, three intense potassium difference peaks are observed. One of these peaks (intensity 6.9σ) is located at site P1, bridging from O2 of C9 to N3 of A17. The second peak (6.5σ) is located at site P2, bridging from O2 of T8 to N3 of A18. The third peak (6.7σ) is located at site P3, bridging from O2 of T7 to O2 of T19. Within the secondary layer, two intense potassium difference peaks are observed. The first (6.2σ) is at site S2, spanning the P1 and P2 sites. The second peak (6.5σ) is at site S3, spanning the P2 and P3 sites.

A second indication of monovalent occupancy is given by the results of thermal factor analysis. These results are in accord with those given by difference peaks. When refined as a water molecule, a hybrid solvent site will give an anomalously low thermal factor because the electrons of a cation occupy a smaller volume (sodium) and are in greater number (potassium) than the electrons of a water molecule. The anomaly should increase with the number of electrons of the cation. Therefore, monovalent cation-binding sites

Table 1: Crystallographic and Refinement Statistics of Potassium Form

unit cell	
α, β, γ (deg)	90
a (Å)	25.296
b (Å)	40.244
c (Å)	65.939
DNA (asymmetric unit)	[d(CGCGAATTCGCG)]2
space group	$P2_12_12_1$
temperature of data collection (°C)	−110
no. of reflections	61 696
no. of unique reflections	11 376
no. of reflections used in refinement ($F > 3\sigma F$)[a]	6479
R (int) (%)[b]	10.1
rms deviation of bonds from ideality (Å)	0.004
rms deviation of angles from ideality (Å)	1.02
max resolution of observed reflections (Å)	1.50
max resolution of refinement (Å)	1.50
no. of DNA atoms	486
no. of water molecules, excluding magnesium first shell	126
no. magnesium ions plus coordinating water molecules	7
no. of spermine atoms	7
R-factor (%)	21.8%
R-free (%)	28.2%

resolution range (Å)	no. of reflections	*R*-factor	completeness[c] (%)
2.98−10.0	1211	17.2	80.73
2.37−2.98	1092	21.7	76.36
2.08−2.37	934	23.2	66.19
1.89−2.08	814	24.5	57.94
1.75−1.89	682	24.9	49.42
1.65−1.75	514	22.6	37.03
1.57−1.65	375	23.9	27.27
1.50−1.57	190	27.6	13.79

[a] Systematic absences and reflections with a resolution of >10 Å were deleted. A 3σ cut-off (F) was applied. [b] $R(\text{int}) = \sum|F_o^2 - F_o^2(\text{mean})|/\sum|F_o^2|$. [c] The structure was refined using the parameter file dna-rna-multi-endo.param, which includes distances and angles recently described by Berman and coworkers (*20*, *21*). The deviations are from those parameters. The completeness excludes ten per cent of reflections used for cross validation.

will be characterized by lower thermal factors in the potassium form than in the sodium form. All else being equal, the difference thermal factor [$\Delta B = B(\text{K}) - B(\text{Na})$] at each common solvent site should be negative for hybrid solvent sites and zero for homogeneous water sites. However, the poorer resolution and data quality of the potassium data set causes the average solvent thermal factor to be greater in the potassium form (average 28.6 Å^2 for 126 water molecules) than in the sodium form (average 24.6 Å^2; for 160 water molecules). Notable exceptions to the $\Delta B > 0$ rule are observed for each of the five sites identified by difference intensity. Those sites show negative ΔBs. All other sites within the fused hexagon motif show positive ΔB with the exception of site T3. The significance of the T3 result is discounted by its relatively large thermal factors and small ΔB (in magnitude).

The positions of monovalent cations within the minor groove were initially inferred from their sodium-specific valence (*14*). In essence, the trend in sodium valence indicates that the coordination geometry tightens and/or the coordination number increases as one approaches the floor of the minor groove. In general, regions with intense potassium difference peaks and negative ΔB conform to regions with high sodium valence. For example, average potassium difference peak intensity and ΔB and the average sodium valence change smoothly from one layer to the next within the fused hexagon motif (Table 2). Both are at maxima within the primary layer, fall off with each succeeding layer, and reach minima at the quaternary layer. The correlation also holds within the primary layer. The P4 site shows the lowest valence, the lowest difference peak intensity, and the only positive ΔB within the primary layer. The P4 site has the lowest monovalent cation occupancy of any of the primary sites. Within the secondary layer, the correlation is weaker. The S3 site has a lower valence than expected from difference peak intensity and ΔB. This discrepancy arises from (i) differences in the binding of potassium and sodium, (ii) experimental error in difference peak intensities, (iii) experimental error in atomic positions, (iv) experimental error in thermal factors, and/or (v) failure of the valence model to predict relative sodium occupancy. In fact the failure of the sodium valence model is not surprising for the secondary layer. The valence is strongly dependent on the number of coordinating ligands. The S2 and S3 sites are 3 coordinate and by definition would show a low sodium valence. However, the coordination geometry, especially of the S2 site, is excessively tight for a water molecule (Figure 1c). Therefore, the geometry of the secondary layer is consistent with partial occupancy by monovalent cations.

The coordination geometries of the primary sites are hybrids of those expected for sodium ions and for water molecules. Each primary site is either 5- or 6-coordinate, interacting with two hydrogen bond acceptors on the floor of the groove, with one or two O4′s of deoxyriboses, and with two secondary sites. The geometry of the 6-coordinate P3 site is shown in Figure 3. The distances between the P3 site and its ligands are longer than expected for an ideal sodium ion, but more numerous and in some cases slightly shorter than expected for an ideal water molecule.

The fused hexagon motif fills the minor groove. Each layer within the fused hexagon motif engages in characteristic interactions with layers above and below. Differences between the fused hexagon motifs of the sodium and potassium forms are subtle, and the discussion here refers to both forms. If one temporarily assumes full occupancy of the fused hexagon motif by water molecules, many hydrogen positions can be assigned even though hydrogen atoms are not observed in our experiments. The compilation of water stereochemistries in peptide and nucleotide hydrates by Jeffrey and co-workers (*22*, *23*) provides a dictionary of possible interactions of nearest neighbors. That dictionary and the observed positions of DNA hydrogen bond acceptors allow us to determine allowable locations of many hydrogen atoms within the fused hexagon motif.

Primary Layer (Light Blue in Figure 1). The primary layer (sites P1−P4) interacts extensively with the DNA, forming a template that directs assembly of the three succeeding layers. Essentially, all contacts between DNA and the fused hexagon motifs are mediated by the primary layer. The upper layers, with the very few exceptions noted below, do not contact the DNA. As noted above, each primary site is

Table 2: Thermal Factors, Difference Peak Intensities, and Sodium-Specific Valencies of the Fused Hexagon Motif

site	atom label	Na^+ form thermal factor	K^+ form thermal factor	Δ thermal factor (Å^2)	K^+ difference peak intensity (σ)	Na^+ specific valence (e−)
P1[a]	**W35**	**5.90**	**4.29**	**−1.61**	**6.88**	**0.46**
P2	**W32**	**9.00**	**8.77**	**−0.23**	**6.49**	**0.49**
P3	**W60**	**10.00**	**5.69**	**−4.31**	**6.74**	**0.49**
P4	W38	10.20	20.97	10.77	3.79	0.45
avg		8.78	9.93	1.16	5.98	0.47
SD		1.99	7.59	6.63	1.47	0.02
S1	W36	13.00	27.11	14.11	4.57	0.42
S2	**W55**	**9.10**	**8.57**	**−0.53**	**6.20**	**0.32**
S3	**W59**	**12.70**	**11.06**	**−1.64**	**6.53**	**0.28**
S4	W48	16.80	25.95	9.15	2.87	0.33
S5	W67	15.20	15.44	0.24	4.62	0.32
avg		13.36	17.63	4.27	4.96	0.33
SD		2.91	8.50	6.98	1.47	0.05
T1	W65	18.40	20.33	1.93	3.99	0.39
T2	W62	16.80	18.62	1.82	4.37	0.29
T3	W137	24.40	24.30	−0.10	2.95	0.28
T4	W134	33.80	74.76	40.96	2.75	0.21
T5	W115	24.30	45.71	21.41	3.11	0.38
avg		23.54	36.74	13.20	3.43	0.31
SD		6.68	23.87	17.83	0.71	0.08
Q1	W95	22.30	28.48	6.18	<2.5	0.27
Q2	W113	36.00	63.64	27.64	3.06	0.07
Q3	W140	30.50	48.24	17.74	<2.5	0.07
Q4	W86	25.50	53.17	27.67	<2.5	0.29
avg		28.58	48.38	19.81	<2.5	0.18
SD		5.99	14.74	10.22		0.12

[a] Sites of monovalent cation occupancy are in bold.

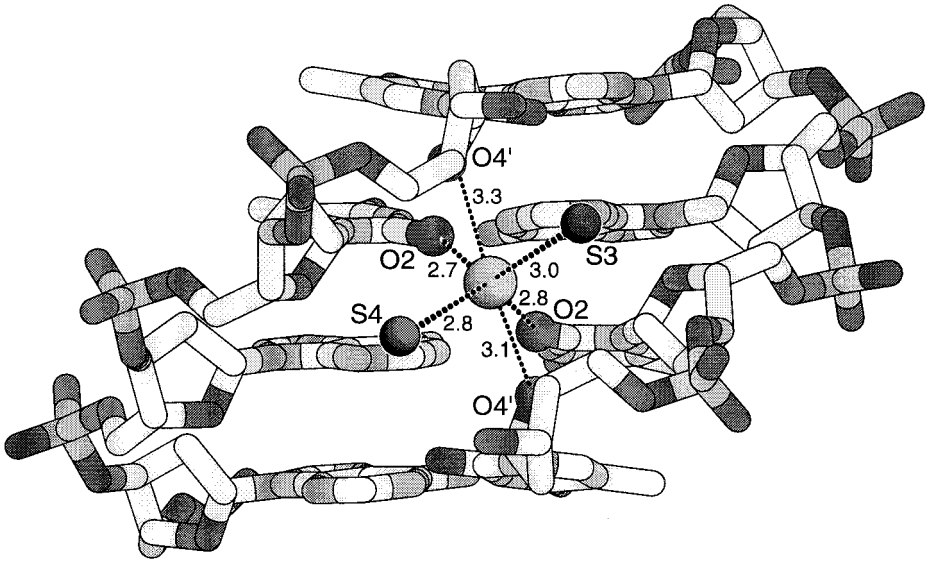

FIGURE 3: A view into the minor groove of the sodium form of dCGCGAATTCGCG showing the coordination geometry of the P3 site, which is located at the 5′ ApT 3′ step. The atoms are shaded according to type with O (dark) > P > N > C (light). The ligands of the P3 site, represented as spheres, are two O4′ atoms, two carbonyl oxygen atoms, and the occupants of two S3 sites. The sphere representing the P3 site is larger than the other six spheres. The distances indicated are in Angstroms.

either 5- or 6-coordinate, interacting with two hydrogen bond acceptors on the floor of the groove, with one or two O4′s of deoxyriboses, and with two secondary sites. Primary sites form bridges between N3 (adenine) and O2 (thymine or cytosine) positions on the floor of the minor groove (Figures 1 and 4). The interactions extend beyond the A and T base pairs of the A-tract. As shown in Figure 1B, the O2 of a cytosine (C9) supports the P1 site. Changing the C9−G16 base pair to G9−C16 in the present dodecamer would disrupt the hexagon motif and repel cations from the vicinity of the P1 site.

For a primary site occupied by a water molecule (as assumed for all sites throughout the following descriptions of the layers), the water molecule is a hydrogen bond donor in all interactions with the DNA. No other type of hydrogen bonding interaction is possible because the minor groove of A-tract DNA is absolutely devoid of hydrogen bond donors. Where the primary sites are 6-coordinate, each hydrogen bond with the DNA is three centered. These hydrogen atoms are bonded by a covalent bond to one atom and by hydrogen bonds to two other atoms (*23*). As noted by Jeffrey, three centered hydrogen bonds result from a deficiency of proton

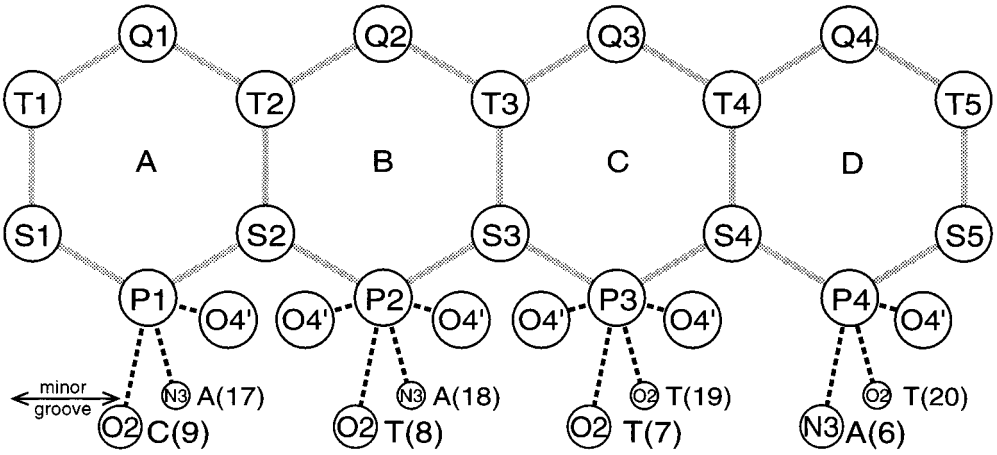

FIGURE 4: Schematic diagram showing solvent sites and the connectivities within idealized fused hexagons. The four layers are indicated by P (primary), S (secondary), T (tertiary), and Q (quaternary). O4′ atoms are primary site ligands.

donors relative to acceptors. The singularity in hydrogen bond polarity within the primary layer restricts water rotations and allows unambiguous assignment of some hydrogen positions.

Secondary Layer (Magenta). The secondary layer (sites S1−S5) does not directly contact the DNA and forms bridges between adjacent primary sites. The hydrogen atoms of primary water molecules are sequestered by the DNA and are unavailable for interaction with the secondary layer. Therefore, secondary water molecules interact exclusively as hydrogen bond donors to the primary layer. Water molecules in the primary and secondary layers are linked by simple two-centered bonds; each secondary hydrogen atom interacts with one primary acceptor atom. The coordination environment around secondary sites is very nearly planar. In the most planar case, T3 (W137) is only 0.04 Å from the plane of P2 (W32), S3 (W59), and P3 (W60).

Tertiary Layer (Blue). The tertiary layer (sites T1−T5) rests on a linear array of hydrogen bond acceptors provided by the secondary water molecules. The tertiary layer was not apparent at low resolution (*16*−*18*) where the positions and behavior of upper level solvent molecules appeared to arise primarily from non-sequence-specific interactions with phosphate groups. Each tertiary site perches on a single supporting secondary site. The perching geometry of the tertiary layer on the secondary layer differs fundamentally from the bridging geometry of the secondary layer on the primary layer. Each tertiary site is juxtaposed over a supporting secondary site whereas each secondary site forms a bridge between two supporting primary sites.

The perching geometry of the tertiary layer arises from steric repulsion between the primary and tertiary layers. One might imagine that the tertiary layer could shift over and drop down slightly toward the floor of the minor groove so that tertiary sites would form bridges between secondary sites. In fact, reasonable secondary/tertiary distances and angles can be achieved by this shift. However, the shift also causes short repulsive contacts (<2.3 Å) between the primary and tertiary sites.

Two limiting hydrogen-bonding schemes are consistent with the observed geometry of the tertiary layer. These schemes differ by rotations of the tertiary water molecules. In a simple two-center bonding scheme, one hydrogen atom of a tertiary water molecule would interact with each secondary water acceptor. In this scenario, each secondary water molecule would accept a single tertiary hydrogen and donate two hydrogen atoms to two primary water acceptors. This 3-coordinate, two-center hydrogen-bonding scheme is consistent with planar geometry (*22*) as observed here. Alternatively, the tertiary layer could interact with the secondary layer via a chelating scheme, in which both hydrogen atoms of a tertiary water molecule interact with a secondary acceptor. In this scenario, each secondary water molecule would accept two tertiary hydrogen atoms and donate two hydrogen atoms to two primary acceptors. These two hydrogen-bonding schemes are not mutually exclusive.

The tertiary layer in the sodium form engages in two contacts with the DNA. Excluding interactions of the primary layer, these are the only DNA contacts of the fused hexagon motif. In one hydrogen bond, at the junction of rings A and B, site T2 (W62) interacts with a phosphate oxygen [G10 (O1P), 3.1 Å]. This hydrogen bond confers near-tetrahedral geometry to site T2, inducing the greatest observed deviation of the monolayer from planarity. Q1 (W95) is further from its hexagonal plane than any other site in the motif. In the second hydrogen bond, at the junction of rings C and D, site T4 (W134) forms a hydrogen bond to a different phosphate oxygen [C21 (O1P), 3.6 Å]. This hydrogen bond is longer and appears to have less effect on the planarity of the fused hexagon motif than the interaction of T2.

Quaternary Layer (Red). The foundation for the quaternary layer (sites Q1−Q4) is provided exclusively by the tertiary layer, with no contributing interactions from the DNA. The quaternary layer forms bridges between tertiary sites. There are no repulsive quaternary to secondary steric contacts to force a perched rather than a bridged configuration.

The interactions of the quaternary layer with its foundation are the weakest links of the hexagonal motif. The curvature of the DNA groove causes the hexagons to splay, lengthening hydrogen bonds from the tertiary to the quaternary layer. In fact some distances from the quaternary layer to the supporting layer below are longer than anticipated for stable hydrogen bonds (*23*). Within the quaternary layer, site Q2 (W113) is further from its means of support (3.9 and 4.0 Å)

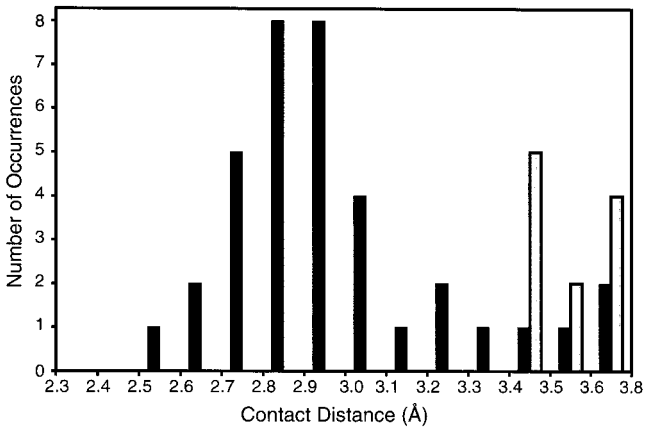

FIGURE 5: The distribution of distances (<3.8 Å) between magnesium first shell water molecules and DNA amino groups (shaded) and between magnesium first shell water molecules and oxygen atoms/non-amino group nitrogen atoms (black).

than any other member of the motif. It is conceivable that some solvent molecules in the tertiary are disordered within a broad potential well, and interact with either one T site or the other. Disorder is consistent with the relatively high thermal factors of the quaternary layer (Table 2). The thermal factors suggest that the quaternary sites are the least localized of the fused hexagon motif.

Hydrated magnesium ions interact unfavorably with DNA amino groups. To determine how divalent cations sequester around DNA and modulate conformation and to correlate the effects of divalent cations with those of monovalent cations, we have conducted a structural survey of interactions of magnesium with each type of DNA functional group. The survey is of high resolution (better than 1.7 Å resolution) B-DNA-magnesium complexes contained in the Nucleic Acid Data Base (*24*). The results show that, when bound to DNA, magnesium ions remain fully hydrated. Water molecules in the first hydration shell of magnesium (W_{Mg} molecules) interact favorably with DNA oxygen and nitrogen atoms. The contact distances can be clearly separated into two classes. Those with non-amino group nitrogen atoms and with oxygen atoms are distinct from those with amino group nitrogen atoms. The favorable non-amino group nitrogen/oxygen interactions with magnesium are characterized by a Gaussian distribution of W_{Mg} to oxygen/nitrogen contact distances centered at around 2.9 Å (Figure 5). W_{Mg} molecules interact with both phosphate oxygen atoms (33%) and with non-phosphate oxygen and nitrogen atoms (67%). DNA amino groups are excluded from favorable interactions with magnesium. The envelope of favorable W_{Mg} to oxygen/nitrogen contacts centered at 2.9 Å does not overlap a single amino group to W_{Mg} distance. Each amino group to W_{Mg} distance is long; none is less than 3.4 Å (Figure 5).

DISCUSSION

Dickerson and co-workers determined the first single-crystal X-ray structure of a complete double-helical turn of B-DNA [d(CGCGAATTCGCG), NDB structure BDL001] (*25*). In the central A-tract, consisting of AT and TA base pairs, they observed substantial deviations from canonical B-DNA conformation and hydration. It is now known that A-tracts narrow the minor groove (*16*−*18*), bend DNA when repeated in phase with the helical repeat (*5*−*9*), propeller twist base pairs (*26*), and support a "spine of hydration" (*16*−*18*). The A-tract spine of hydration, as described by Dickerson and Drew, is composed of two layers of water molecules that cover the floor of the minor groove but do not interact with the deoxyribose backbones. The primary layer of the spine forms bridges between N3 (adenine) and O2 (thymine) positions on the floor of the minor groove (Figure 4). The secondary layer forms bridges between members of the primary layer.

High-resolution diffraction data allowed us to determine the three-dimensional structure of sodium (*14*) and potassium (here) forms of d(CGCGAATTCGCG) and its associated solvent molecules and ions accurately and in high detail. We observe that the minor groove of the AT-tract of the DNA supports a fused hexagon motif of solvent sites, with primary, secondary, tertiary, and quaternary layers (Figures 1 and 4). Essentially all contacts between DNA and the fused hexagon motifs are mediated by the primary layer. The upper layers, with few exceptions, do not contact the DNA. The motif is a smooth, curved monolayer of highly geometric water molecules. The "spine of hydration" is the lower two levels of the fused hexagon motif and is populated by monovalent cations. The observation of monovalent cations within the minor groove confirms the correctness and relevance of original dinucleotide structure of Rich and co-workers (*27*), who identified a sodium ion near the floor of that abbreviated minor groove.

Water molecules previously identified by X-ray diffraction of DNA are cation−water hybrids. Sodium valence (*28*), thermal factors, and potassium difference peaks support the hybrid solvent model, in which solvent sites contain both water molecules and monovalent cations, each in partial occupancy. Thus, cation density is smeared-out over many or even all solvent sites that are observed by crystallography. Partial monovalent cation occupancy of solvent sites is clearly indicated in the potassium form of d(CGCGAATTCGCG) described here, which yielded less than complete 1.75 Å resolution data. Potassium, with an ionic radius of 1.33 Å, is slightly larger than sodium (0.97 Å) and is more easily dehydrated, but should bind DNA in nearly the same fashion as sodium. In fact, we have observed subtle shifts of monovalent cations toward the floor of the minor groove with increasing ionic radii (C.C.S., L.M.-I., and L.D.W., unpublished results).

Our crystallographic observation of smeared-out cation density is consistent with the predictions of the polyelectrolyte model of Manning and Record (*29*), which describes a statistical atmosphere of cations around polyanions. The polyelectrolyte model correctly applies to polymers in solution, not oligonucleotides in crystals, so the specific cation distributions around the two types of anions should differ. Yet both distributions are statistical, with smeared-out cation atmospheres. One goal of our structural analysis then becomes an effort to map out regions of high and low cation density and to determine the effects of uneven cation density on nucleic acid structure. The Manning−Record model predicts displacement of monovalent cations by divalent cations. However, the combined structural results indicate that monovalent and divalent ions bind to different sites. We believe that displacement occurs by a combined mechanism of direct interaction between nearby sites and by

solvent reorganization. Displacement should be observable via a series of crystal structures with difference cation occupancies.

DNA phosphate groups interact with peaks and troughs of cation density. Mirzabekov and Rich proposed that DNA bending around proteins could arise from phosphate neutralization by localized cations (*30*). Covalently localized cations do indeed bend DNA, even in the absence of protein (*31*−*33*). The self-repulsive anionic system of DNA collapses around localized cationic charge. The electrostatic forces that exert such control over DNA conformation are nearly invisible to medium and low-resolution macromolecular X-ray crystallography, which gives the incorrect impression that DNA molecules are surrounded predominantly by seas of neutral aqueous solvent. In fact, conformational deformations appear to be dominated by translation of phosphate groups toward cation density peaks and possibly away from troughs. NMR (*12*), molecular dynamics (*13*), and X-ray diffraction experiments and database analysis (here and ref *14*) indicate that peaks of cation density reside in the minor groove of AT-tracts and the major groove of GC-tracts. The distribution of cations appears to be determined at least in part by unfavorable interactions with electropositive amino groups.

DNA bending is caused by phosphate interaction with sequence-directed peaks and troughs of cation density. The cation dependencies of both A-tract (*5*, *34*) and G-tract bending (*35*) demand a roll for cations in bending mechanism. We propose that DNA bends by collapse around peaks of cation density. Peaks form in the minor groove of A-tracts and in the major groove of G-tracts. The observed directions of bending, A-tracts into the minor groove (*36*) and G-tracts into the major groove (*37*, *38*), are consistent with the effects of electrostatic collapse.

The axial deflections of A-tracts have been described by several related models, termed wedge (*39*) and junction (*7*−*9*, *40*). Electrostatic collapse is essentially consistent with features of both models. The degree and location of axial deflection should depend on location and occupancy of cation-binding sites. Monovalent cations concentrate near the centers of A-tract (this report) and might cause localized axial curvature. Divalent cations concentrate near their termini (*12*) and might cause more dispersed axial curvature. Cations with greater charge would exercise a more powerful effect on conformation than cations with lesser charge, as noted by Hud and Feigon (*12*). Indeed, divalent cations do bend DNA to a greater extent then monovalent cations (*5*, *34*, *35*). The electrostatic collapse model differs from wedge and junction models in ascribing cause to axial bending and in accounting for the observed dependence of DNA bending on cations.

Our results provide a rationale for Hagerman's observation that 5′ [d(GAAAATTTTC)]n 3′ is bent while 5′ [d(GTTTTAAAAC)]n 3′ is linear (*1*). We believe that the minor groove of 5′ dAT 3′ steps provide a singular high-affinity binding site for hard monovalent cations. The site is composed of two carbonyl oxygen ligands, two O4′ oxygen ligands, and two water molecule oxygen ligands (Figure 3). Hard cations such as sodium and potassium bind preferentially to oxygen. By contrast, a 5′ dTA 3′ step lacks a favorable cation-binding site. In this dinucleotide step, two imino nitrogen atoms would bind hard cations poorly.

Solution structures of DNA are explained by phosphate interaction with sequence-directed peaks and troughs of cation density. Electrostatic collapse around small cations is consistent with a series of previous observations. (i) X-ray structures suggest that the minor groove of A-tract DNA is narrowed by phosphate interaction with peaks of cation density within the groove. A-tract DNA has a narrow minor groove in solution (*41*). (ii) X-ray structures suggest that chemical modifications that locate amino groups in the minor groove of A-tracts should displace cations and abolish axial bends. Diekmann and McLaughlin (*42*) used a series of modified bases to determine that in solution amino groups within the minor groove abolish the axial curvature of A-tracts. (iii) In the X-ray structures, the interdependence of solvent sites within the fused hexagon motif suggests that the motif would form cooperatively. Each hexagon of the minor groove motif stabilizes those around it. Each layer stabilizes the others. Thus, axial bending should increase cooperatively with A-tract length. In solution, axial bending does increase cooperatively with A-tract length (*43*, *44*). (iv) X-ray structures suggest that the fused hexagon motif should melt out with a melting temperature at or below that of the DNA. In solution, A-tract DNA shows a distinctive premelting transition (*45*−*47*). (v) X-ray structures suggest that the water and cation interactions of each base pair in the A-tract are dependent on the sequence and solvent structure of surrounding base pairs. The interactions of the fused hexagon motif spill over to the residues that flank the A-tract (Figure 1B). In solution, A-tract bending is dependent on flanking sequence (*48*, *49*). (vi) X-ray structures suggest that minor groove binders should attenuate axial bends of A-tracts. Minor groove binders would displace cations from the minor groove. In solution, minor groove binders straighten A-tract bends (*50*, *51*). (vii) The diol MPD chelates divalent cations (L.M.-I., and L.D.W., data not shown) and so should straighten A-tract bends. In solution, MPD appears to straighten A-tract bends (*52*, *53*).

Specific properties of various divalent cations predict their tendencies to bend DNA. The importance of amino group repulsion in divalent cation interactions suggests certain properties of divalent cations will correlate with tendency to bend DNA. The amount of cationic charge on the outer-sphere of an aquo complex would determine the relative repulsion from electropositive DNA amino groups. The first pK_a of an aquo complex describes with the extent to which cationic charge is distributed onto the outer hydration sphere. We suggest that axial curvature should correlate with the first pK_a of the divalent ion aquo complex. Laundon and Griffith (*34*) showed that A-tract bending of DNA induced by divalent cations varies in the order Ca^{2+}, Mg^{2+} < Mn^{2+} < Co^{2+}, Ni^{2+}, Zn^{2+}. The order of this series does indeed correlate with pK_a [Ca^{2+}, 12.6−13.4; Mg^{2+}, 11.4−12.8; Mn^{2+}, 10.6−10.9; Co^{2+}, 7.6−9.9; Ni^{2+}, 6.5−10.2; Zn^{2+}, 8.2−9.8 (*54*)].

Solution-type bending might not be observable in DNA crystals. Crothers and co-workers (*9*) argue that the dodecamer A-tract described previously by Dickerson at low resolution (*16*−*18*, *25*) and more recently at high resolution by our laboratory (this paper and ref *14*) may not be in the

conformation responsible for DNA bending in solution. We concur. In a crystal, intermolecular electrostatic interactions compete with intramolecular interactions. The distribution of cations around a DNA oligonucleotide might be perturbed by lattice interactions. Oligonucleotide crystallography may be the wrong technique for observation of solution-type axial bending of DNA. However, high-resolution oligonucleotide crystallography is probably the only technique capable of identifying hybrid solvent sites and determining their partial occupancies.

CONCLUSION

The results described here support models of cation−water hybrid solvent and electrostatic collapse of DNA around peaks of inorganic cation density. Observed DNA deformation is consistent with interaction of an uneven cationic cation atmosphere with the self-repulsive anionic system of DNA phosphate groups. We have taken the extreme position here that intrinsic contributions to DNA deformation by sequence-specific base−base interactions (stacking and hydrogen bonding, steric repulsion of the exocyclic groups, etc.) are dwarfed by extrinsic contributions. In the final analysis, both models may have applicability under some conditions.

ACKNOWLEDGMENT

The authors thank Drs. George Jeffrey, Richard Harlow, Nicolas Hud, Juli Feigon, Jonathan Chaires, Donald Crothers, and Paul Hagerman for helpful discussions.

SUPPORTING INFORMATION AVAILABLE

A single page table is available that gives inter- and intraring deviations from planarity within the fused hexagon motif (sodium form of CGCGAATTCGCG) (1 page). Ordering information is given on any current masthead page.

REFERENCES

1. Hagerman, P. J. (1990) *Annu. Rev. Biochem. 59*, 755−781.
2. Olson, W. K., and Zhurkin, V. B. (1996) in *Biological Structure and Dynamics* (Sarma, R. H., and Sarma, M. H., Eds.) pp 341−370, Adenine Press, Albany.
3. Allemann, R. K., and Egli, M. (1997) *Chem. Biol. 4*, 643−650.
4. Klug, A., Jack, A., Viswamitra, M. A., Kennard, O., Shakked, Z., and Steitz, T. A. (1979) *J. Mol. Biol. 131*, 669−680.
5. Diekmann, S., and Wang, J. C. (1985) *J. Mol. Biol. 186*, 1−11.
6. Koo, H.-S., Drak, J., Rice, J. A., and Crothers, D. M. (1990) *Biochemistry 29*, 4227−4234.
7. Koo, H.-S., Wu, H.-M., and Crothers, D. M. (1986) *Nature 320*, 501−506.
8. Nadeau, J. G., and Crothers, D. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 2622−2626.
9. Haran, T. E., Kahn, J. D., and Crothers, D. M. (1994) *J. Mol. Biol. 244*, 135−143.
10. Nikolov, D. B., Chen, H., Halay, E. D., Hoffman, A., Roeder, R. G., and Burley, S. K. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 4862−4867.
11. Passner, J. M., and Steitz, T. A. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 2843−2847.
12. Hud, N. V., and Feigon, J. (1997) *J. Am. Chem. Soc. 119*, 5756−5757.
13. Young, M. A., Jayaram, B., and Beveridge, D. L. (1997) *J. Am. Chem. Soc. 119*, 59−69.
14. Shui, X., McFail-Isom, L., Hu, G. G., and Williams, L. D. (1998) *Biochemistry 37*, 8341−8355.
15. Rouzina, I., and Bloomfield, V. A. (1998) *Biophys. J. 74*, 3152−3164.
16. Drew, H. R., Wing, R. M., Takano, T., Broka, C., Itakura, K., and Dickerson, R. E. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 2179−2183.
17. Drew, H. R., and Dickerson, R. E. (1981) *J. Mol. Biol. 151*, 535−556.
18. Kopka, M. L., Fratini, A. V., Drew, H. R., and Dickerson, R. E. (1983) *J. Mol. Biol. 163*, 129−146.
19. Brunger, A. T. (1996) *Methods. Mol. Biol. 56*, 245−266.
20. Clowney, L., Jain, S. C., Srinivasan, A. R., Westbrook, J., Olson, W. K., and Berman, H. M. (1996) *J. Am. Chem. Soc. 118*, 509−518.
21. Gelbin, A., Schneider, B., Clowney, L., Hsieh, S.-H., Olson, W. K., and Berman, H. M. (1996) *J. Am. Chem. Soc. 118*, 519−529.
22. Jeffrey, G. A. and Maluszynska, H. (1990) *Acta Crystallogr., Sect. B* 46, 546−549.
23. Jeffrey, G. A. (1997) *An Introduction to Hydrogen Bonding*, Oxford University Press, New York.
24. Berman, H. M., Olson, W. K., Beveridge, D. L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S.-H., Srinivasan, A. R., and Schneider, B. (1992) *Biophys. J. 63*, 751−759.
25. Wing, R., Drew, H., Takano, T., Broka, C., Takana, S., Itakura, K., and Dickerson, R. E. (1980) *Nature 287*, 755−758.
26. Nelson, H. C. M., Finch, J. T., Luisi, B. F., and Klug, A. (1987) *Nature 330*, 221−226.
27. Rosenberg, J. M., Seeman, N. C., Kim, J. J. P., Suddath, F. L., Nicholas, H. B., and Rich, A. (1973) *Nature 243*, 150−154.
28. Brown, I. D. (1992) *Acta Crystallogr., Sect. B* 48, 553−572.
29. Anderson, C. F., and Record, M. T., Jr. (1995) *Annu. Rev. Phys. Chem. 46*, 657−700.
30. Mirzabekov, A. D., and Rich, A. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 1118−1121.
31. Strauss, J. K., and Maher, L. J. (1994) *Science 266*, 1829−1834.
32. Strauss, J. K., Roberts, C., Nelson, M. G., Switzer, C., and Maher, L. J. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 9515−9520.
33. Strauss-Soukup, J. K., Vaghefi, M. M., Hogrefe, R. I., and Maher, L. J. (1997) *Biochemistry 36*, 8692−8698.
34. Laundon, C. H., and Griffith, J. D. (1987) *Biochemistry 26*, 3759−3762.
35. Brukner, I., Susic, S., Dlakic, M., Savic, A., and Pongor, S. (1994) *J. Mol. Biol. 236*, 26−32.
36. Zinkel, S. S., and Crothers, D. M. (1987) *Nature 328*, 178−181.
37. Dlakic, M., and Harrington, R. E. (1995) *J. Biol. Chem. 270*, 29945−29952.
38. Milton, D. L., Casper, M. L., Wills, N. M., and Gesteland, R. F. (1990) *Nucleic Acids Res. 18*, 817−820.
39. Ulanovsky, L. E., and Trifonov, E. N. (1987) *Nature 326*, 720−722.
40. Selsing, E., Wells, R. D., Alden, C. J., and Arnott, S. (1979) *J. Biol. Chem. 254*, 5417−5422.
41. Price, M. A., and Tullius, T. D. (1992) *Methods Enzymol. 212*, 194−219.
42. Diekmann, S., and McLaughlin, L. W. (1988) *J. Mol. Biol. 202*, 823−834.
43. Diekmann, S. (1986) *FEBS Lett. 195*, 53−56.
44. Leroy, J. L., Charretier, E., Kochoyan, M., and Gueron, M. (1988) *Biochemistry 27*, 8894−8898.
45. Chan, S. S., Breslauer, K. J., Austin, R. H., and Hogan, M. E. (1993) *Biochemistry 32*, 11776−11784.
46. Chan, S. S., Breslauer, K. J., Hogan, M. E., Kessler, D. J., Austin, R. H., Ojemann, J., Passner, J. M., and Wiles, N. C. (1990) *Biochemistry 29*, 6161−6171.

47. Herrera, J. E., and Chaires, J. B. (1989) *Biochemistry 28*, 1993−2000.
48. Hagerman, P. J. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 4632−4636.
49. Zhurkin, V. B. (1985) *J. Biomol. Struct. Dyn. 2*, 785−804.
50. Wu, H. M., and Crothers, D. M. (1984) *Nature 308*, 509−513.
51. Hansma, H. G., Browne, K. A., Bezanilla, M., and Bruice, T. C. (1994) *Biochemistry 33*, 8436−8441.
52. Dlakic, M., Park, K., Griffith, J. D., Harvey, S. C., and Harrington, R. E. (1996) *J. Biol. Chem. 271*, 17911−17919.
53. Ganunis, R. M., Guo, H., and Tullius, T. D. (1996) *Biochemistry 35*, 13729−13732.
54. Kazakov, S. A. (1996) in *Bioorganic Chemistry: Nucleic Acids* (Hecht, S. M., Ed.) pp 244−287, Oxford University Press, New York.

BI982063O