Effects of Cationic Charge on Three-Dimensional Structures of Intercalative Complexes: Structure of a *bis*-Intercalated DNA Complex Solved by MAD Phasing†,‖

Xiuqi Shui[1], Mary E. Peek[1,2], Leigh Ann Lipscomb[1,3], Qi Gao[4], Craig Ogata[5], Bernard P. Roques[6], Christiane Garbay-Jaureguiberry[6], Angus P. Wilkinson[1] and Loren Dean Williams[1]*

(1) School of Chemistry & Biochemistry, Georgia Institute of Technology, GA 30332-0400, (2) Current Address; Bioprocess & Bioanalytical Research, Merck & Company, Inc., West Point, PA 19486, (3) Current Address; Department of Biochemistry & Molecular Biology, University of Georgia, B216A Life Sciences Bldg., Athens, GA 30602, (4) Bristol-Myers Squibb, Wallingford, CT 06492-7660, USA (5) NSLS, Brookhaven National Laboratory, Upton, NY 11973, (6) Universite Rene Descartes (Paris V), UFR des Sciences Pharmaceutiques et Biologiques, Laboratory de Chemie Organic, 4, Avenue de L'Observatoire, 75270 Paris Cedex 06, France

Abstract: We characterize intercalative complexes as either "high charge" and "low charge". In low charge complexes, stacking interactions appear to dominate stability and structure. The dominance of stacking is evident in structures of daunomycin, nogalamycin, ethidium, and triostin A/echinomycin. By contrast in a DNA complex with the tetracationic metalloporphyrin CuTMPyP4 [copper (II) meso-tetra(N-methyl-4-pyridyl)porphyrin], electrostatic interactions appear to draw the porphyrin into the duplex interior, extending the DNA along its axis, and unstacking the DNA. Similarly, DNA complexes of tetracationic ditercalinium and tetracationic flexi-di show significant unstacking. Here we report x-ray structures of complexes of the tetracationic bis-intercalator D232 bound to DNA fragments d(CGTACG) and d(^{Br}CGTA^{Br}CG). D232 is analogous to ditercalinium but with three methylene groups inserted between the piperidinium groups. The extension of the D232 linker allows it to sandwich four base pairs rather than two. In comparison to CuTMPyP4, flexi-di and ditercalinium, stacking interactions of D232 are significantly improved. We conclude that it is not sufficient to characterize intercalators simply by net charge. One anticipates strong electrostatic forces when cationic charge is focused to a small volume or region near DNA and so must consider the extent to which cationic charge is focused or distributed. In sum, ditercalinium, with a relatively short linker, focuses cationic charge more narrowly than does D232. So even though the net charges are equivalent, electrostatic charges are expected to be of greater structural significance in the ditercalinium complex than in the D232 complex.

†*The atomic coordinates have been deposited in the Nucleic Acid Databank (IC9Z).*

‖ *This research was funded by the American Cancer Society (RPG-95-116-03-GMC) and National Science Foundation, Molecular Biophysics Program (MCB-9056300).*

Nucleic acid conformational classes are distinguished by spatial relationships among bases. Useful parameters for describing inter-base rotations and translations include helical twist, base pair roll, slide, rise, buckle, and propeller twist [1]. Relatively low roll and slide characterize the B-conformation. Relatively high roll and slide characterize the A-conformation.

Sequence-specific variations in short-range stacking interactions between bases are thought by some to be causative of global phenomena such as DNA bending and groove width variation [2-5]. These “heterocycle-centric models” are of the following general form. (i) Inter-base spatial relationships are dominated by short-range stacking interactions. (ii) Propeller twist increases stacking between adjacent base pairs. (iii) Slide increases cross-stand purine stacking and induces coplanarity of the stacked purines. (iv) Coplanarity of cross-strand stacked purines is equivalent to base pair roll, if propeller twist is

*Address correspondence to this author at the School of Chemistry & Biochemistry, Georgia Institute of Technology, GA 30332-0400, USA

0929-8673/00 $19.00+.00

maintained. (v) Roll causes axial bending. One factor that might induce one to pursue such models is that once a structure is determined, spatial relationships between bases are determined quickly and easily by computer programs [6,7].

Long-range electrostatic forces within DNA have proven to be much less amenable to structural analysis than short-range stacking interactions. The inorganic cations required to neutralize the phosphate groups are generally missing from x-ray and NMR structures. Un-neutralized phosphate groups are glaring thermodynamic flaws, and would denature DNA in solution and explode a crystal.

In polyelectrolyte models, structure details are minimized and electrostatic forces rule. Conformational changes are described by changes in axial charge density [8]. The high degree of thermodynamic accuracy of polyelectrolyte models suggests that the short-range stacking forces emphasized in heterocycle-centric models are not of primary significance in determining global conformation.

X-ray structures and polyelectrolyte models seem to be orthogonal views of reality. Structural models focus on ordered components such as bases and backbone. Inorganic cations are obscured by water and possibly disordered, and so most are commonly omitted from structural models. By contrast, polyelectrolyte models omit structural detail, but include cations within delocalized 'atmospheres'.

We have begun to attempt to experimentally reconcile structural and polyelectrolyte models by locating inorganic cations within DNA X-ray structures [9-11]. We have proposed that sequence-specific, long range, electrostatic interactions modulate nucleic acid conformation. In this model, sequence-specific peaks and troughs of cation density within the counterion atmosphere deforms DNA. The deformations can be global, as in axial bending, and local, as in groove-width variation. The mechanism, originally proposed by Mirzabekov and Rich to explain spontaneous DNA bending around proteins [12-14], is known as "electrostatic collapse".

Some features of DNA conformation and interactions in intercalative complexes similarly appear to arise from electrostatic forces. The effects of electrostatic interactions are evident from comparison of a series of intercalative complexes with various charge distributions.

Several highly charged intercalators have recently been crystallized in complexes with DNA. In the structure of a DNA complex with the tetracationic metalloporphyrin [15], CuTMPyP4 [copper (II) meso-tetra(N-methyl-4-pyridyl)porphyrin; shown in Fig. (**1A**)], electrostatic interactions appear to draw the porphyrin into the duplex interior, in spite of unfavorable steric clashes. The clashes extend the DNA along its axis, unstacking the DNA. Structures of tetracationic ditercalinium [16,17] [Fig. (**1B**) & (**1C**)] and tetracationic flexi-di [18,19] [Fig. (**1B**) & (**1D**) bound to fragments of DNA also show significant unstacking. Each compound in this ditercalinium sub-family is composed of two monocationic 7H-pyridocarbazole units and a charged linker. Structure-activity studies indicate that the cationic charges [20] and linker rigidity [20-24] are critical to activity. Ditercalinium has a rigid diethylbipiperidine linker and is active in both prokaryotes and eukaryotes [23]. Flexi-di has a flexible spermine-like linker and is active only in prokaryotes [24]. Several features of the X-ray complexes of ditercalinium and flex-di were unanticipated, and distinguish these complexes from previous complexes of intercalators with lower cationic charge. Unanticipated structural features are poor base-base and base-chromophore stacking, groove reversal, with placement of the linker in the major groove rather than generally more favorable the minor groove, and axial bending into the minor groove. The structural and energetic origins of these features are not necessarily clear, even with the 3-D structures in hand. For example we hypothesized that specific hydrogen bonding and electrostatic interactions of the linker with the floor of the major groove contribute to groove reversal. However Gago and coworkers [25] discount the importance of these hydrogen bonding interactions in water. They argue for the importance of stacking interactions in groove reversal. They note that preferred orientations of intercalating chromophores relative to adjacent bases could arise from quadrapole interactions. They also suggest that the observed axial bend is artifactual, and may be due to lattice forces.

The solution of 3-D structures of additional complexes, with varying linkers and lattice interactions, can help to resolve some of these issues. Here we report structures of complexes of the tetracationic bis-intercalator D232 [Fig. (**1B**) & (**1E**)] bound to DNA fragments d(CGTACG) and d(^{Br}CGTA^{Br}CG). D232 is analogous to ditercalinium. It has a relatively long, semi-rigid linker, with three methylene groups separating the piperidiniums. The covalent extension of the D232 linker allows it to sandwich four base pairs rather than two via bis-intercalation and with geometry roughly similar to that determined by NMR. Biologically D232 is essentially inactive, even though it binds to DNA with high affinity via bis-intercalation [26].

Methods

The ammonium salts of reverse-phase HPLC-purified native d(CGTACG) and derivative

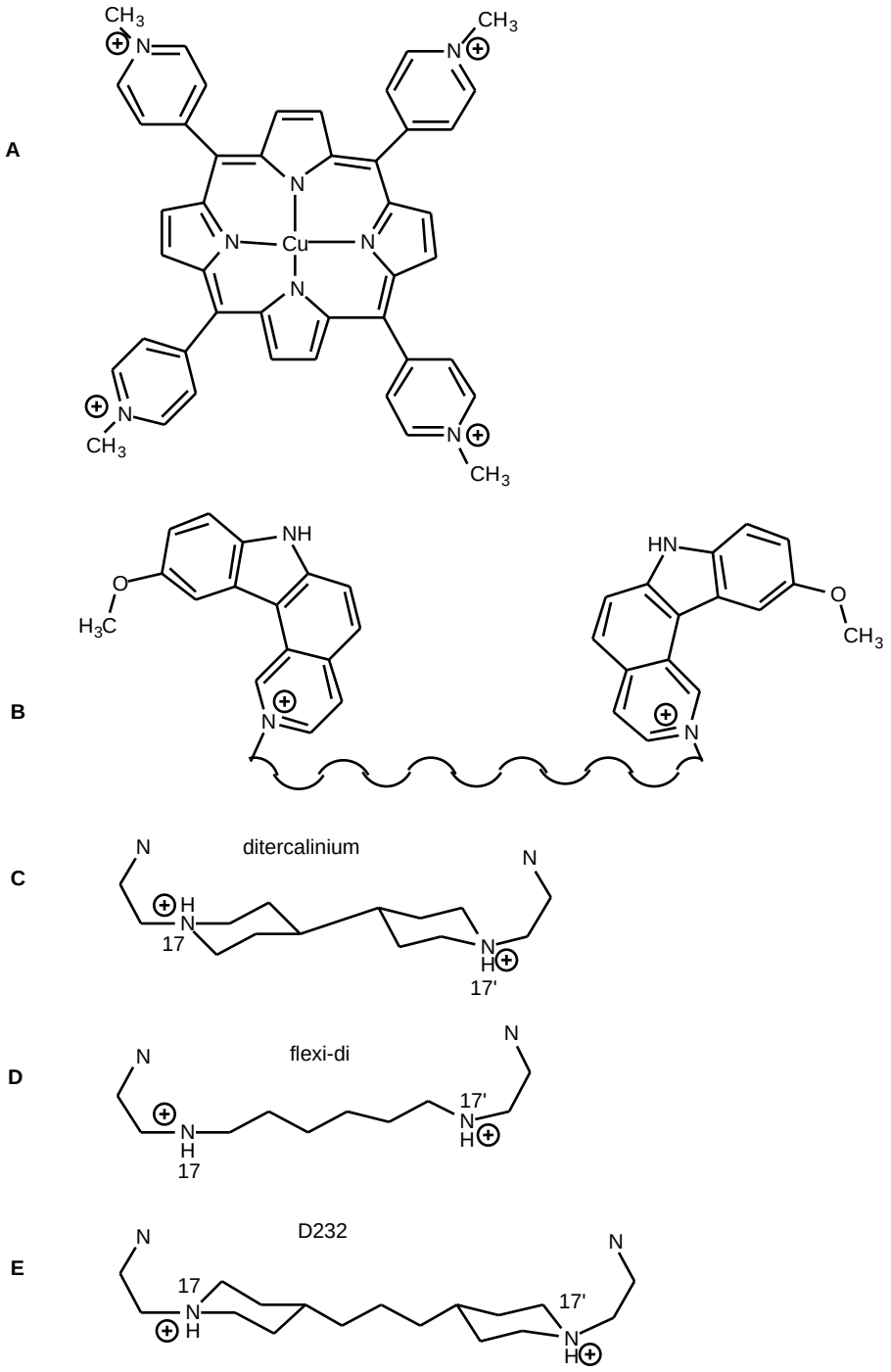

Fig. (1). Chemical structures of (A) CuTMPyP4 [copper (II) meso-tetra(N-methyl-4-pyridyl)porphyrin]; (B) the intercalative chromophores of ditercalinium, flexi-di and D232; the linkers of (C) ditercalinium; (D) flexi-di; and (D) D232.

d(^{Br}CGTA^{Br}CG) were purchased from the Midland Certified Reagent Co., (Midland, TX). D232 was synthesized as described [23]. Native crystals suitable for diffraction were crystallized from sitting drops that initially contained 0.8 mM single strand DNA, 15.8 mM sodium cacodylate (pH 6.0), 2.6% 2-methyl-2,4-pentanediol (2-MPD), 0.51 mM spermine, 0.66 mM $MgCl_2$, and 0.37 mM D232. These droplets were equilibrated against a reservoir of 20% 2-MPD. The complex crystallized in space group $P6_{5(1)}22$. A crystal with the approximate dimensions of 0.66 mm x 0.15 mm x 0.10 mm was sealed in a glass capillary and mounted on a Huber goniostat. Data were collected in the ω scan mode at 22 °C using a San Diego Multiwire Systems (SDMS) area detector. Copper Kα radiation (1.5418 Å) was generated with a fine-focus Rigaku RU200 rotating anode. Data were merged and reduced by SDMS software.

The bromine derivative was crystallized from a sitting drop that initially contained 0.8 mM single-strand DNA, 20 mM sodium cacodylate (pH 6.0), 3.3% 2-MPD, 1.7

Table 1. Data Collection Statistics

	λ (Å)	Unit Cell[a] (a axis, b axis, Å)	No. of ref.	No. of unique ref.	% unique ref. observed	res. cutoff (Å)	R_{sym}
Native	1.5418	28.240 72.744	4698	744	91%	2.4	5.14[b]
Br	1.5418	27.790 74.365	54715	1874	100	1.8	6.59[b]
Br							
λ1:	0.9270	27.723 74.265	6788	1928	88	2.0	4.3[c]
λ2:	0.9204	27.723 74.265	6927	1938	88	2.0	5.7[c]
λ3:	0.9201	27.723 74.265	6198	1986	90	2.0	4.9[c]
λ4:	0.9100	27.723 74.265	6273	1846	84	2.0	4.1[c]

a) The space group is $P6_{1(5)}22$; b) Rsym = ∑ | Fav - Fobs |/∑ Fav; c) Rsym = ∑ (Iav - Iobs)^2/∑ Iobs^2.

mM $MgCl_2$ and 0.3 mM D232. These droplets were equilibrated against a reservoir of 30% 2-MPD. A single wavelength (Kα) data set was collected on a bromine derivative crystal (size 0.29 x 0.29 x 0.18 mm) as described for the native crystal above. Multiple wavelength diffraction data were collected of a crystal of approximate size 0.20 x 0.15 x 0.12 mm at beamline X4A of the National Synchrotron Light Source at the Brookhaven National Laboratory (Upton NY) using Fuji imaging plates and a BAS2000 scanner. The choice of wavelengths near the K-shell edge of bromine was made by observation of the X-ray fluorescence spectrum of the crystal (peak: 0.9201Å; inflection point: 0.9204Å; two remote wavelengths below and above the absorption edge: 0.9270Å and 0.9100Å). The X-ray fluorescence spectrum was used to obtain F' and F" as a function of wavelength with the programs FPRIME [27] and KRAMER (Lieselotte Templeton).

The c* axis of the crystal was aligned along the horizontal spindle direction. "Mirror" geometry was used to measure Bijvoet pairs. Data were collected by a total phi sweep of 48°. The missing cone region was collected by aligning the bisection of a* and b* parallel to the spindle and sweeping phi 15°. Data were processed by DENZO and reduced by SCALEPACK. Data collection statistics are presented in Table 1.

MAD phasing was performed with the MADSYS program suite provided by Wayne Hendrickson's laboratory [28]. Phasing statistics are given in Table 2. Data were placed on a common scale with F2ANO. Local scaling at each wavelength was done by ANLSCL and among four wavelengths by WVLSCL. MADLSQ was used to calculate $|°F_T|$, $|°F_A|$ and $\Delta(\varphi_T - \varphi_A)$. The phasing results were reduced to a unique set by MERGIT ($R_{sym}|°F_T|$ = 0.043, $R_{sym}|°F_A|$ = 0.282 and $\Delta\Delta(\varphi_T-\varphi_A)$ = 22.12°). Eighty per cent of unique data

Table 2. Phasing Statistics

	Observed ratios[a]				Scattering factors[b]		R_{sym} (%)[c]
	λ1	λ2	λ3	λ4	Δf'	Δf"	
λ1	0.048 (0.049)	0.080	0.090	0.063	-4.36	0.51	2.1
λ2		0.058 (0.051)	0.055	0.094	-7.98	1.25	2.8
λ3			0.088 (0.053)	0.099	-8.97	3.18	2.3
λ4				0.105 (0.061)	-4.03	4.05	2.5

a) Bijvoet difference rations are given in the diagonal elements with the centric values in parethesis. The dispersive difference ratios are given in the off-diagonal elements of the table; b) Scattering factors Δf' and Δf" of λ1 were fixed. Scattering factors at the other wavelengths were refined by MADLSQ; c) After local scaling. $R_{sym} = \sum\sum | Ii(h) - <|(h)|> | / \sum\sum Ii(h)$, where for each reflection h, Ii(h) indicates the ith observation and <|h|> indicates the weighted mean of all observations.

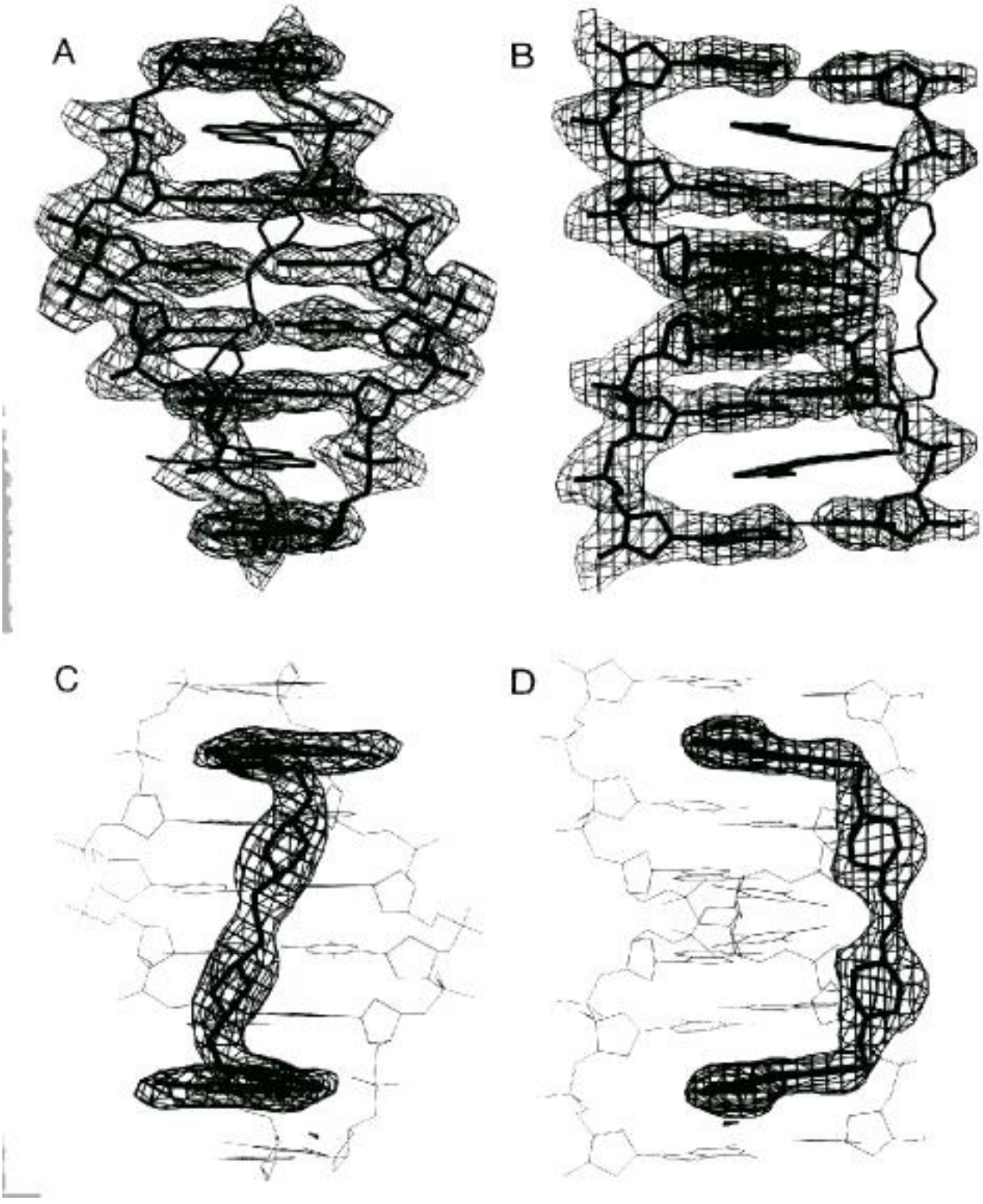

Fig. (2). Sum electron density (2Fo-Fc contoured at 1σ) surrounding the D232-d(CGTACG) complex. (A) Electron density surrounding the DNA only, view into the major groove (B) Electron density surrounding the DNA only, view in (A) rotated by 90°. (C) Electron density surrounding the D232 only, view into the major groove (D) Electron density surrounding the D232 only, view in (C) rotated by 90°.

were successfully phased. The locations of two bromine atoms solved from Patterson maps calculated between 10 - 2.4 Å from the $|°F_A|$ set using HASSP [29]. The positions as well as B factors of the heavy atoms were refined against $|°F_A|$ by ASLSQ. Phased φ_T were computed by MADFAZ to resolution 2.4 Å in both possible space groups ($P6_122$ and $P6_522$). Hendrickson-Lattman phase probability coefficients of all unique reflections were calculated by MADABCD [30]. The initial maps in both space groups were uninterpretable even though the average figure of merit was 0.94. The electron density map of space group $P6_122$ appeared the most promising. This map was improved significantly by phase combination using the base of one brominated cytosine (C1) and the second bromine. The program PHASES [31] was used for phase recombination. A single strand DNA hexamer and drug without the linker were located by iterative phase recombination. The model was refined with reasonable statistics by simulated annealing and several cycles of positional refinement with XPLOR [32]. However the DNA formed a left-handed helix. The model was inverted to the right-handed helix and the hand of the space group switched from $P6_122$ to $P6_522$. This conversion dropped the R-factor from 0.340 to 0.283 and the R-free from 0.380 to 0.329 (Table 3). The refinements were switched to the native and derivative data collected at 1.5418 Å and the structures were re-annealed and refined to that data. The linker was located in $2|F_o|-|F_c|$ and $|F_o|-|F_c|$ Fourier maps. An internal crystallographic 2-fold axis passes through the central methylene carbon atom of the linker. Since macromolecular refinement suits are generally incapable of restraints involving atoms on special positions, the structures were refined at lower crystallographic symmetry (space group $P6_5$) but with forced symmetry about the omitted two-fold. Anisotropic scaling of |Fo| to |Fc| was applied. Three water molecules were located in the final native structure. Ten water molecules were located in the derivative structure. The final models are characterized by R values of 19.6% for the native structure [731 reflections ($3\sigma(F)$) between 10 - 2.4 Å] and of 19.7% for the derivative brominated structure (1761 reflections ($5\sigma(F)$) between 10 - 2.4 Å). The rms deviations from ideality are 0.021 Å (native) and 0.013 Å (derivative) in bond lengths and 3.29° (native) and 2.89° (derivative) in angles. The final electron density maps are clean and continuous around D232 and the DNA [Fig. (**2**)] Both sets of coordinates have been deposited in the Nucleic Acid Data Bank.

Results

D232 *bis*-intercalates into DNA, intercalating at each CpG step of $[d(CGTACG)]_2$ [Figs. (**3**), (**4**), & (**5**)] and at the ^{Br}CpG steps of $[d(^{Br}CGTA^{Br}CG)]_2$. The DNA forms a right-handed but underwound double helix. The linker spans four internal base pairs, so that D232 forms a sandwich around the central $[d(GTAC)]_2$ segment. The complex contains crystallographic two-fold symmetry. The asymmetric unit contains one DNA hexamer and half of a D232 molecule. The central methylene carbon of D232 falls directly on the two-fold rotation axis. The DNA complex is labeled 5' C(1) G(2) T(3) A(4) C(5) G(6) 3' on one strand and 5' C(7) G(8) T(9) A(10) C(11) G(12) 3' on the other. C(1) is base paired to G(12), G(2) is base paired to C(11) and so on. Base pair C(1)-G(12) is crystallographically identical to base pair G(6)-C(7), G(2)-C(11) is identical to C(5)-G(8) and so on.

The linker is located in the major groove, running roughly parallel to the helical axis [Fig. (**3**) & (**4**)]. The twisting of the floor of the groove causes the linker, which is essentially linear, to effectively traverse the major groove. The cationic N17 and N17' amino groups of D232 form hydrogen bonds to the N7 atoms of the internal guanines [G(2) and G(8); 2.9 Å; distances in the native complex are given here, although the structure of the brominated complex is nearly identical]. The N17 and N17' atoms do not appear to form hydrogen bonds with the O6 atoms of the internal guanines (3.7 Å). We previously proposed that the N17-N7 hydrogen bonds of ditercalinium and flexi-di cause a shift of the internal guanine into the major groove relative to the flanking bases. The observed shift in the D232 complex (Dx for the C1pG2 step is 0.7 Å) is consistent with this proposal. In the ditercalinium-$[d(CGCG)]_2$ complex,

Table 3. Resolution of Phase Ambiguity

	R-free	R-factor	Final R-factor	helical twist
$P6_522$	34.0	0.283	0.196 (native, 2.4 Å)	right handed
$P6_122$	38.0	0.329		left handed

The phase ambiguity was broken by cross validation and the assumed right-handed helical twist of intercalated B-DNA. $R_{sym}|^oF_T| = 0.043$, $R_{sym}|^oF_A| = 0.282$, $\Delta\Delta(\varphi_T - \varphi_A) = 22.12^o$. The percentage of unique data phased to 2.4 Å is 75.7%. Br positions and thermal factors were refined upon $|^oF_A|$: $R = \sum | F_o - F_c | / \sum |F_o| = 0.278$. The figure of merit of $|^oF_T|$ phasing for $P6_{1(5)}22$ is 0.941

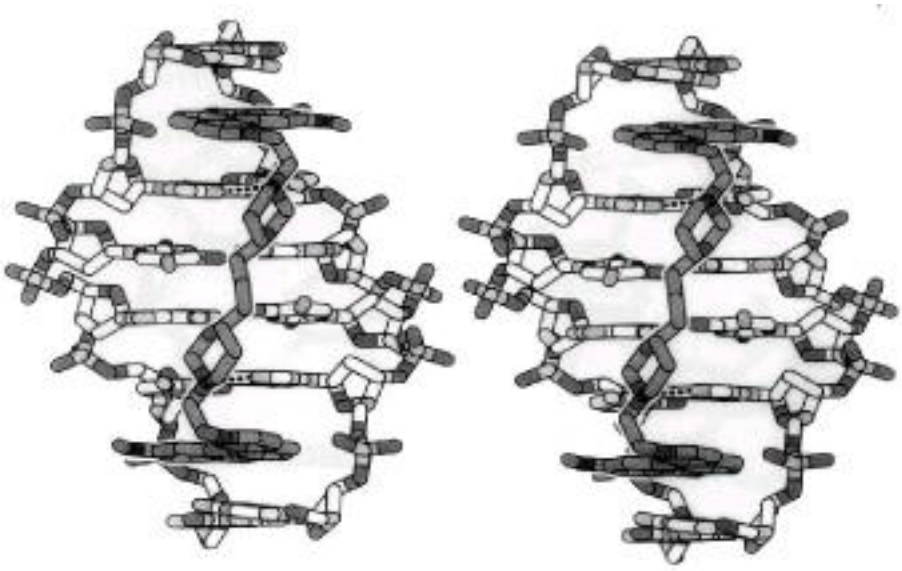

Fig. (3). Stereoview of the D232-DNA complex, in stick representation. The atoms are shaded by type with O (dark) > P > N > C (light). D232 is shaded relative to the DNA. The hydrogen bonds between the N17/N17' of D232 and the N7 of G(2)/G(11) are indicated by dashed lines.

only one end of the linker forms hydrogen bonds to the floor of the major groove. That end of the complex, but not the other end, displays a significant shift of the internal guanine into the major groove. The observation of hydrogen bonds between the linker and the floor of the major groove in all three complexes, in two different crystal froms, suggests that they are not artifacts of the crystal lattice.

With the exception of the two hydrogen bonds, the linker does not engage in especially tight interactions with the floor of the major groove. Other than these hydrogen bonds, none of linker-major groove contacts are less than 3.6 Å. However, even in the absence of van der Waals contacts the linker appears to engage in hydrophobic interactions with the 5-methyl groups of the central thymines, by excluding water from a portion of the methyl group surface. The linker lies neatly in a cleft formed by the methyl groups.

The inter-chromophore spacing is greater than expected from the increase in the length of the linker, but less than expected from the increase in the length of the DNA. The three methylene insertion adds around 3.4 Å to the length of the linker of D232 in comparison to the linker of ditercalinium. The standard axial rise of B-DNA suggests that the addition of two base pairs to the intercalated sandwich should increase the intercalative spacing by around 6.4 Å. In fact the increase in separation of the chromophores of D232 relative to ditercalinium is 4.2 Å (15.6 Å in D232 and 11.5 Å in ditercalinium). The linker appears to be stretched taut by the DNA.

D232 unwinds the DNA in this complex by 40° (assuming a 36° helical turn for the un-complexed DNA). This value is similar to that observed in the ditercalinium complex, in which $[d(CGCG)]_2$ is unwound by a total of 36° [16,17]. The DNA is unwound severely at the CpG intercalation step by D232, and moderately at the GpT step. Surprisingly, in the D232 complex is slightly overwound at the central TpA step. We believe that the D232 complex represents the first observation of overwinding, even in a local sense, induced by a mono or bis-intercalator. The local overwinding at the central step may be related to the hydrogen bonds formed by the linker and to the tautness of the linker. In the axial projection, the N-H bond on one end of the linker is rotated by 100° relative to that on the other end. This rotation is fixed by the tautness of the linker. Therefore the DNA must wind by a total of 100° in three steps to achieve the proper orientation of the two guanine bases that engage in hydrogen bonds [Fig. (**6**)] The DNA effectively wraps around the linker, to achieve favorable hydrogen bonds, whose relative orientations are fixed by the linker. The helical axis of the D323 complex does not show any notable bend, unlike the ditercalinium and flexi-di complexes.

In comparison to ditercalinium, the stacking interactions of D232 are significantly improved. One measure of extent of stacking is given by the number and length of the interatomic base-chromophore contacts. As expected from the ditercalinium complex, the D232 chromophore appears to be well-stacked on the flanking guanines, and stacks preferentially on the guanine over cytosine bases. The tightest stacking

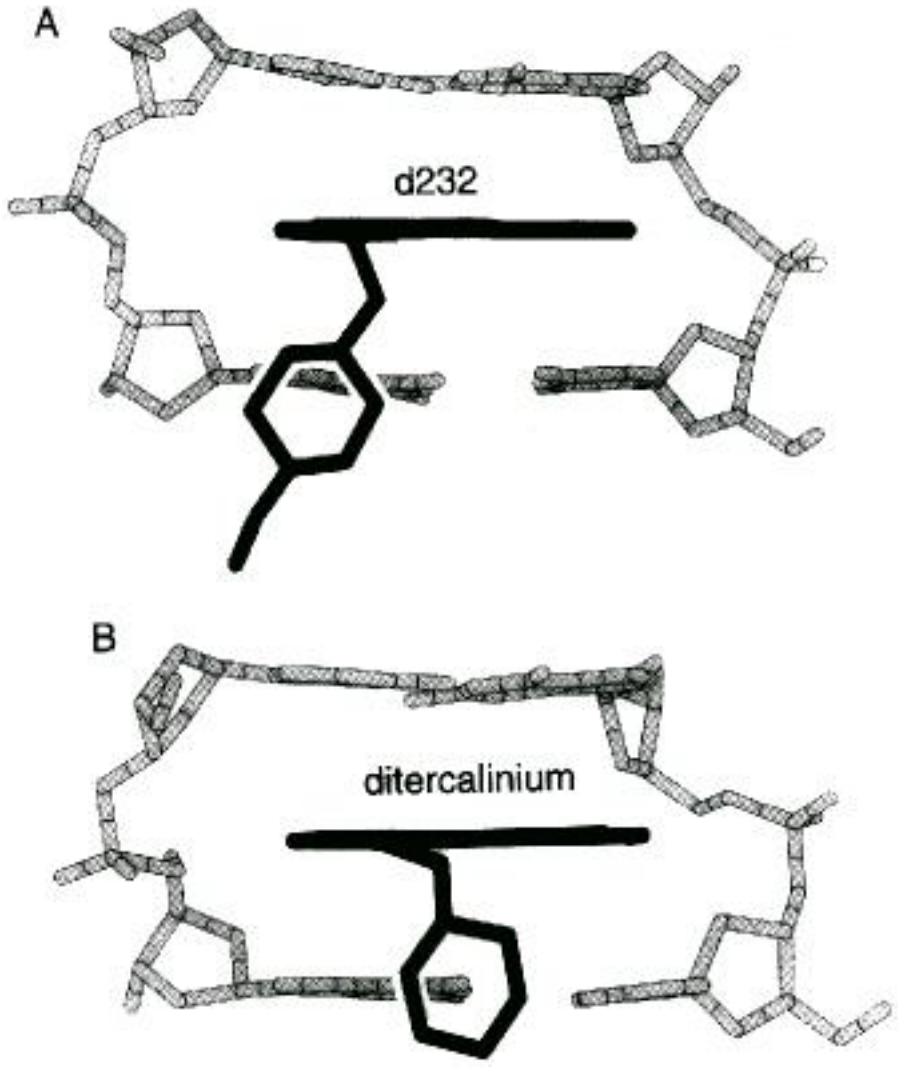

Fig. (4). Views into the major groove, showing the (A) D232, and (B) ditercalinium intercalation sites. The bis-intercalators are shaded.

contacts are with guanines. Three of guanine-chromophore (D232) contacts are less than 3.3 Å in the D232 complex. There are no cytosine-chromophore contacts less than 3.3 Å. However the D232 chromophore is also well-stacked on the flanking cytosine bases. Thirteen (13.5; numbers in parenthesis are from the ditercalinium complex. That complex is asymmetric; the numbers of contacts are averaged for the two different intercalation sites) of the sixteen cytosine base atoms are within 3.7 Å of the chromophore. Ten (7) are with 3.6 Å. Seven (3) are within 3.5 Å. Five (one) are within 3.4 Å. Cytosines are not well-stacked in the ditercalinium complex.

Stacking interactions can be visualized by the extent of apparent overlap when viewing along the normals of the base planes. One must keep in mind that shifts along the helical axis are not observable in the axial view. Axial views of base pair-chromophore in the D232 and ditercalinium complexes are shown in [Fig. (**5**)]. In both complexes the terminal cytosine overlaps the A-ring, and appears to engage in reasonable stacking interactions (Fig. (**5A**) & (**5B**)]. However as described above this cytosine engages in much more favorable interactions with D232 than with ditercalinium. These differences in base-chromophore stacking do not arise from shifts of entire base-pairs

along the helical axis. The helical rise at the intercalation step of D232 is 6.8 Å (6.9, 6.8 Å). The 3.2 Å rise at the next step is considerably less than that of the ditercalinium complex (3.6 Å). The rise at the central step of the D232 complex is 3.5 Å.

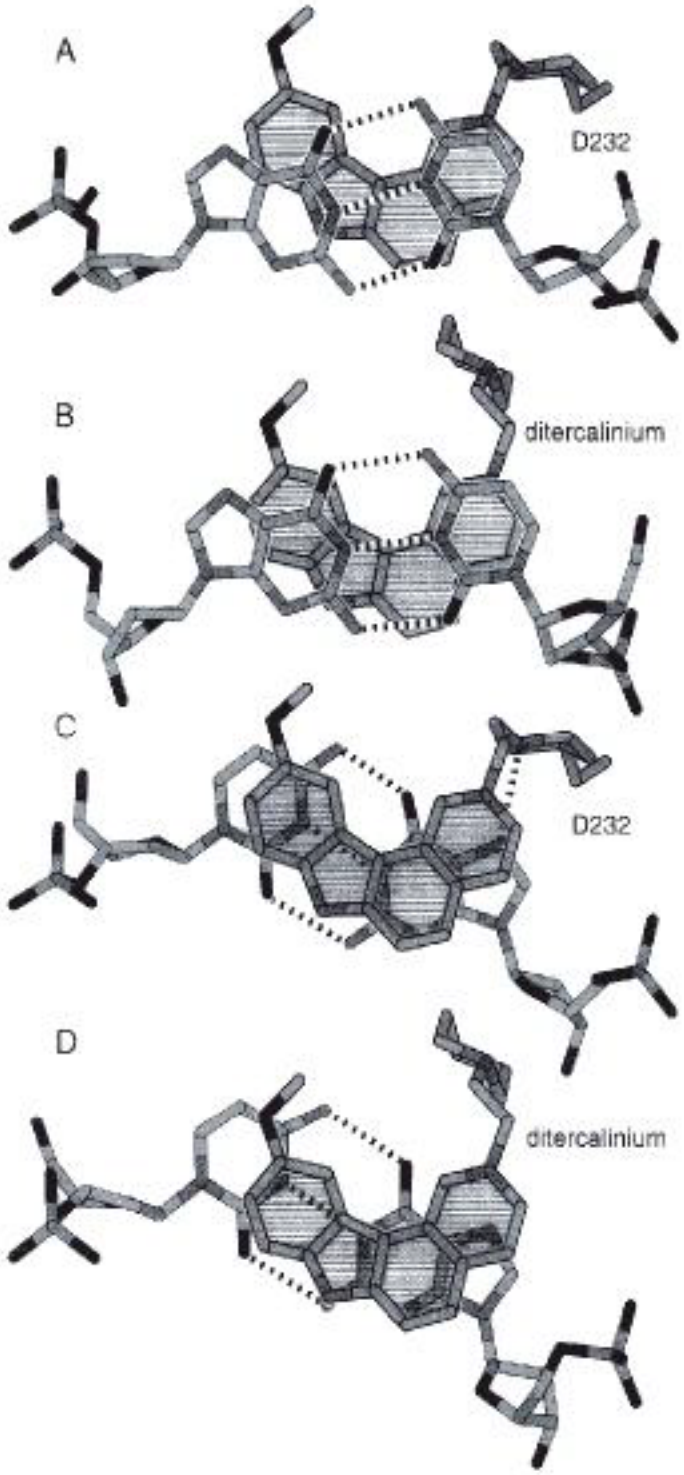

Fig. (5). Views along the helical axis, showing (A) the C(1)-G(12) base pair juxtaposed over the D232 chromophore; (B) the C(1)-G(12) base pair over the ditercalinium chromophore; (C) the D232 chromophore over the G(2)-C(11) base pair; and (D) the ditercalinium chromophore over the G(2)-C(11) base pair.

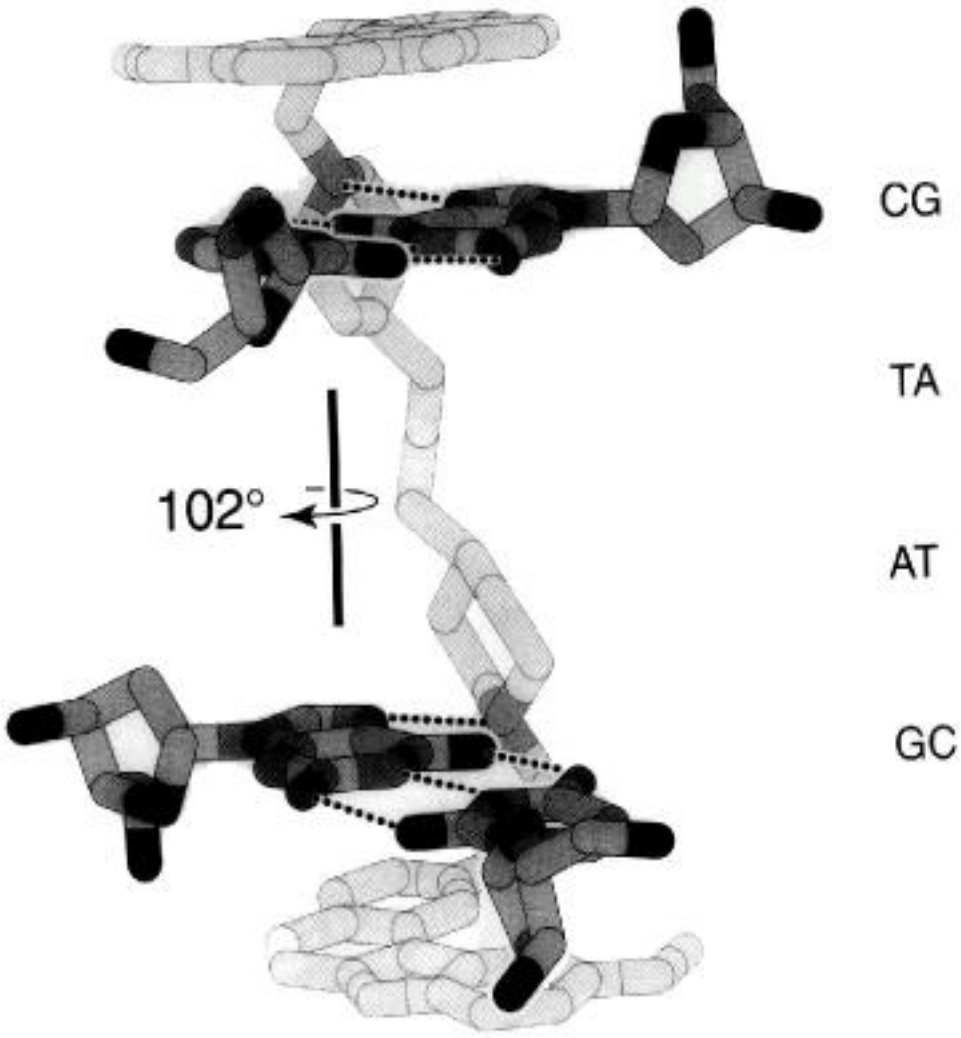

Fig. (6). The relationship between hydrogen bonding and helical twist. Views of D232 and the G(2)-C(11) and C(5)-G(8) base pairs. The hydrogen bonds between the d232 linker and G(2) and G(8) are illustrated by dashed ines. The total DNA helical twist from G(2)-C(11) to C(5)-G(8) is 102°.

The external guanine partially overlaps the C- and D-rings in both complexes (Fig. (**5A**) & (**5B**)]. A comparison of the positions of the two chromophores reveals that D232 is rotated relative to the chromophore of ditercalinium. This rotation is around a pivot point located near atom N14. The location of the pivot point appears to be related to the hydrogen bond between N14 of the chromophore, and the N7 of the internal guanine. This rotation projects the O-methyl group of D232 into the major groove, and attenuates axial overlap with the external guanine.

In the interior chromophore-base pair step of both complexes (Fig. (**5C**) & (**5D**)], the guanine overlaps the A- and B-rings of the chromophore and cytosine overlaps the D-ring. The internal cytosine is well stacked in the D232 complex [Fig. (**5C**)] but not in the ditercalinium complex [Fig. (**5D**)]. This difference appears to arise in part from the previously described rotation of the chromophore, around a pivot point established by the N14 to N7 hydrogen bond.

Discussion

DNA intercalation, first proposed by Lerman [33,34], is analogous to insertion of a 'false coin' (an intercalator) into a roll of pennies (base pairs). To intercalate, a planar molecule slides between DNA two base pairs, generally without breaking Watson-Crick hydrogen bonds. Intercalation extends DNA along its axis, shifting the base pairs flanking the binding site in opposing directions. Intercalation invariably decreases the axial charge density. Other consequences of intercalation to DNA conformation are variable [35]. For some intercalators DNA unwinding is restricted to the site of intercalation. For other intercalators, DNA is wound normally at the intercalation site, but is unwound

Table 4A. Global Base-Base Parameters; Strand 1 with Strand 2[a]

	Buckle (kappa; deg)	**Propel** (omega; deg)	**Opening** (sigma; deg)
C 1 - G 12	-5.83	-13.7	1.2
G 2 - C 11	-3.40	-0.5	1.8
T 3 - A 10	-0.41	-14.8	-2.0
A 4 - T 9	0.41	-14.8	-2.0
C 5 - G 8	3.37	-0.5	1.7
G 6 - C 7	5.83	-13.7	1.2
Average	0.00	-9.7	0.29

a) Calculated with the program CURVES [6].

Table 4B. Local Inter-Base Pair Parameters; Strand 1 with strand 2

	Shift (Dx; Å)	**Slide (Dy; Å)**	**Rise (Dz; Å)**	**Tilt (tau; deg)**	**Roll (rho; deg)**	**Twist (Omega; deg)**	**Rise (Å)**
C 1/G 2	0.73	-0.05	6.7	4.7	5.8	17.3	6.7
G 2 / T 3	-0.78	-0.28	3.4	-1.1	1.4	33.1	3.4
T 3 / A 4	0.00	-0.01	3.4	0.0	5.5	39.4	3.4
A 4 / C 5	0.78	-0.28	3.4	1.1	1.4	33.1	3.4
C 5 / G 6	-0.73	-0.05	6.7	-4.7	5.7	17.4	6.7
Average	0.00	-0.14	4.7	0.0	4.0	28.0	

in the flanking regions. Intercalation is important in many biological process, and provides a probe for structure and for molecular interactions.

A survey of x-ray structures reveals two rather distinctive classes of intercalative complexes, which we term "high charge" and "low charge" complexes. In low charge complexes, stacking interactions appear to dominate stability and structure. These complexes are characterized by extensive intercalator-base and base-base stacking interactions. This dominance of stacking interactions is evident in structures of daunomycin [36-39], nogalamycin [40], ethidium [41], triostin A/echinomycin [42,43] and D232 (described here).

In high charge intercalative complexes, stacking interactions appear to be of lesser significance. These complexes are characterized by poor intercalator-base and base-base stacking interactions. Disruption of stacking is evident in complexes of ditercalinium [16,17], flexi-di [18,19], and TMPyP4 [15]. We have previously made sterochemical arguments to explain the observation of poor stacking [16,19]. However that analysis now appears to have under-estimated the importance of electrostatic forces between formally charged phosphate groups and cationic centers of the intercalators.

It seems reasonable to propose that electrostatic forces contribute to conformation of some intercalative complexes. In fact, the extent of electrostatic interaction between intercalator and DNA appears to differentiate stacked from un-stacked intercalative complexes. Un-stacking is observed in complexes with high charge intercalators but not in complexes with low charge intercalators. Thus unstacking is associated with strong electrostatic forces.

For classifying intercalators as high and low charge (Table 5), one parameter of obvious importance is the net cationic charge on the intercalator. However it is not sufficient to characterize intercalators simply by net charge. One anticipates strong electrostatic forces when cationic charge is focused to a small volume or region near DNA. One must consider the extent to which cationic charge is focused or distributed. A second parameter, which we call the cationic site size, gives an estimate of the extent to which charge is axially distributed. Here we rather arbitrarily quantize the cationic site size in units of 3.4 Å, the axial rise of B-

Table 5. Charge Characteristics of Intercalators

Intercalator	Charge (-esu)	Cationic Site Size (Å)	Axial Cationic Charge Density (-esu/Å)
Low charge			
Echinomycin	0	10.2	0
D232	4	17.0	0.23
Ethidium	1	3.4	0.29
Daunomycin	1	3.4	0.29
Nogalmycin	1	3.4	0.29
High charge			
Ditercalinium	4	10.2	0.39
Flexi-di	4	10.2	0.39
CuTMPyP4	4	3.4	1.2

DNA. The cationic site size is defined as one (1 x 3.4 Å) for mono-interclators such as ethidium, daunomycin, and TMPyP4. The cationic site size is three (10.2 Å = 3.4 x 3) for bis-intercalators such as triostin A and ditercalinium that span two base pairs. The cationic site size is five (17.0 Å = 3.4 x 5) for bis-intercalators such as D232 that span four base pairs. The cationic site size is not to be confused with exclusion site size, which, for example, differ for ethidium and daunomycin.

These rather simple definitions may not be entirely suitable for comparing radically different intercalators, but clearly illustrate that ditercalinium, with a relatively short linker, focuses cationic charge more narrowly than does D232. So even though the net charges are equivalent, electrostatic charges are expected to be of greater significance in the ditercalinium complex than in the D232 complex. The broad distribution of cationic charge in the D232 complex is consistent with observation that it forms a well-stacked intercalative complex.

These definitions also allow us to identify and compare limiting complexes, such as TMPyP4 (most highly charged) and triostin A (no cationic charge). We have previously proposed that strong electrostatic forces draw CuTMPyP4 into the duplex interior [15]. In the intercalated d(CGATCG)-CuTMPyP4 complex positively charged pyridyl nitrogen atoms are in close proximity to DNA phosphate oxygen atoms. Three cationic pyridyl nitrogen atoms are less than 5.0 Å from phosphate oxygen atoms. In the duanomycin complex, by contrast, all phosphate oxygen atoms are greater than 6.0 Å from the single cationic center (an amino group). In the d(CGATCG)-CuTMPyP4 complex, pyridyl-backbone clashes extend the DNA along its axis and prevent van der Waals stacking contacts in the interior of the complex. It seems clear that in the d(CGATCG)-CuTMPyP4 complex, strong electrostatic forces work in opposition to stacking interactions.

Electrostatic forces must be carefully distinguished from polyelectrolyte contributions. The difference is illustrated by triostin A, which is uncharged and so does not engage in electrostatic interactions with DNA. Triostin A, along with all other intercalators, increases DNA length, and decreases the axial anionic charge density. A decrease in axial charge density releases counter ions to bulk solution, contributing favorably to the free energy of association. Thus polyelectrolyte effects contribute to the association of triostin A with DNA even in the absence of electrostatic forces between triostin A and DNA.

References

[1] Dickerson, R. E. *Nucleic Acids Res.* **1989**, *17*, 1797.

[2] Dickerson, R. E. *Nucleic Acids Research* **1998**, *26*, 1906.

[3] Olson, W. K.; Gorin, A. A.; Lu, X. J.; Hock, L. M.; Zhurkin, V. B. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 11163.

[4] Dickerson, R. E.; Chiu, T. K. *Biopolymers* **1997**, *44*, 361.

[5] Calladine, C. R.; Drew, H. R. *Understanding DNA*; Academic Press: New York, 1992.

[6] Stofer, E.; Lavery, R. *Biopolymers* **1994**, *34*, 337.

[7] Babcock, M. S.; Pednault, E. P. D.; Olson, W. K. *J. Biomol. Struct. Dynam.* **1994**, *11*, 597.

[8] Record, M. T.; Zhang, W. T.; Anderson, C. F. *Advances in Protein Chemistry, Vol 51*; Eds.; 1998; 51, pp. 281.

[9] Shui, X.; Sines, C.; McFail-Isom, L.; VanDerveer, D.; Williams, L. D. *Biochemistry* **1998**, *37*, 16877.

[10] Shui, X.; McFail-Isom, L.; Hu, G. G.; Williams, L. D. *Biochemistry* **1998**, *37*, 8341.

[11] McFail-Isom, L.; Shui, X.; Williams, L. D. *Biochemistry* **1998**, *37*, 17105.

[12] Manning, G. S.; Ebralidse, K. K.; Mirzabekov, A. D.; Rich, A. *J. Biomol. Struct. Dyn.* **1989**, *6*, 877.

[13] Mirzabekov, A. D.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 1118.

[14] Strauss-Soukup, J. K.; Maher, L. J. *Biochemistry* **1998**, *37*, 1060.

[15] Lipscomb, L. A.; Zhou, F. X.; Presnell, S. R.; Woo, R. J.; Peek, M. E.; Plaskon, R. R.; Williams, L. D. *Biochemistryaccelerated publication* **1996**, *35*, 2818.

[16] Gao, Q.; Williams, L. D.; Egli, M.; Rabinovich, D.; Chen, S.-H.; Quigley, G. J.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 2422.

[17] Williams, L. D.; Gao, Q. *Biochemistry* **1992**, *31*, 4315.

[18] Peek, M. E.; Lipscomb, L. A.; Bertrand, J. A.; Gao, Q.; Roques, B. P.; Garbay-Jaureguiberry, C.; Williams, L. D. *Biochemistry* **1994**, *33*, 3794.

[19] Peek, M. E.; Lipscomb, L. A.; Haseltine, J.; Gao, Q.; Roques, B. P.; Garbay-Jaureguiberry, C.; Williams, L. D. *J. Bioorganic and Med. Chem.* **1995**, *3*, 693.

[20] Garbay-Jaureguiberry, C.; Barsi, M. C.; Jacquemin-Sablon, A.; Le Pecq, J. B.; Roques, B. P. *J. Med. Chem.* **1992**, 72.

[21] Leon, P.; Garbay-Jaureguiberry, C.; Barsi, M. C.; Le Pecq, J. B.; Roques, B. P. *J. Med Chem.* **1987**, *30*, 2074.

[22] Leon, P.; Garbay-Jaureguiberry, C.; Le Pecq, J. B.; Roques, B. P. *Anti-Cancer Drug Design* **1988**, *3*, 1.

[23] Pelaprat, D.; Delbarre, A.; Le Guen, I.; Roques, B. P.; Le Pecq, J. B. *J. Med. Chem.* **1980**, *23*, 1336.

[24] Pothier, J.; Delepierre, M.; Barsi, M. C.; Garbay-Jaureguiberry, C.; Igolen, J.; Le Bret, M.; Roques, B. P. *Biopolymers* **1991**, *31*, 1309.

[25] dePascual Teresa, B.; Gallego, J.; Ortiz, A. R.; Gago, F. *J. Med. Chem.* **1996**, *39*, 4810.

[26] Delepierre, M.; Maroun, R.; Garbay-Jaureguiberry, C.; Igolen, J.; Roques, B. P. *J. Mol. Biol.* **1989**, *210*, 211.

[27] Cromer, D. T. *J. Appl. Cryst.* **1983**, *16*, 437.

[28] Hendrickson, W. A. *Science* **1991**, *254*, 51.

[29] Terwilliger, T. C.; Kim, S.-H.; Eisenberg, D. *Acta Cryst.* **1987**, *A43*, 1.

[30] Pahler, A.; Smith, J. L.; Hendrickson, W. A. *Acta Crystallogr. A* **1990**, *46*, 537.

[31] Furey, W.; Swaminathan, S. *Proc. Am. Cryst. Assoc. Natl. Meeting* **1990**, *PA33*, 73.

[32] Brunger, A. *XPLOR Software* **1992**, The Howard Hughes Medical Institute Yale University.

[33] Lerman, L. S. *J. Mol. Biol.* **1961**, *3*, 18.

[34] Lerman, L. S. *J. Cell Comp. Physiol.* **1964**, *64* (suppl. 1), 1.

[35] Williams, L. D.; Egli, M.; Gao, Q.; Rich, A. *Structure & Function, Volume I: Nucleic Acids*; Sarma, R. H. & Sarma, M. H., Eds.; Adenine Press: Albany, **1992**; pp. 107.

[36] Quigley, G. J.; Wang, A. H.-J.; Ughetto, G.; van der Marel, G. A.; van Boom, J. H.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 7204.

[37] Frederick, C. A.; Williams, L. D.; Ughetto, G.; van der Marel, G. A.; van Boom, J. H.; Rich, A.; Wang, A. H.-J. *Biochemistry* **1990**, *29*, 2538.

[38] Williams, L. D.; Frederick, C. A.; Ughetto, G.; Rich, A. *Nucleic Acids Res.* **1990**, *18*, 5533.

[39] Williams, L. D.; Egli, M.; Ughetto, G.; van der Marel, G. A.; van Boom, J. H.; Quigley, G. J.; Wang, A. H.-J.; Rich, A.; Frederick, C. A. *J. Mol. Biol.* **1990**, *215*, 313.

[40] Williams, L. D.; Egli, M.; Gao, Q.; Bash, P.; van der Marel, G. A.; van Boom, J. H.; Rich, A.; Frederick, C. A. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 2225.

[41] Jain, S. C.; Sobell, H. M. *J. Biomol. Struct. Dynam.* **1984**, *1*, 1179.

[42] Ughetto, G.; Wang, A. H.-J.; Quigley, G. J.; van der Marel, G. A.; van Boom, J. H.; Rich, A. *Nucleic Acids Res.* **1985**, *13*, 2305.

[43] Wang, A. H.-J.; Ughetto, G.; Quigley, G. J.; Hakoshima, T.; van der Marel, G. A.; van Boom, J. H.; Rich, A. *Science* **1984**, *225*, 1115.