A molecular model for proflavine-DNA intercalation

S.Neidle*, L.H.Pearl, P.Herzyk** and Helen M.Berman¹

Cancer Research Campaign Biomolecular Structure Unit, Institute of Cancer Research, Sutton, Surrey SM2 5PX, UK and ¹Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, PA 19111, USA

Received July 4, 1988; Revised and Accepted August 22, 1988

ABSTRACT

A molecular model has been derived for the intercalation of proflavine into the CpG site of the decamer duplex of d(GATACGATAC). The starting geometry of the intercalation site was taken from previous crystallographic studies on the d(CpG)-proflavine complex, and molecular mechanics used to obtain a stereochemically acceptable structure. This has widened grooves compared to standard A- or B- double helices, as well as distinct conformational, roll, twist and tilt features.

INTRODUCTION

The intercalative interactions of drugs with double-helical nucleic acids have been intensively studied and characterised by a large number of biophysical methods (1-3). The cytotoxic effects of compounds such as amsacrine and Adriamycin have been ascribed not only to their direct DNA binding, but to their effect on a ternary complex involving the enzyme DNA topoisomerase II (4,5).

The inference from these studies is that the drugs induce perturbations in DNA structure, which are the features responsible for recognition by the enzyme and formation of a "cleavable complex." Footprinting studies of drugs binding to defined sequences with both natural and synthetic cleavage agents, have frequently shown the presence of drug-induced conformational effects at sequences remote from the actual preferred binding sites (6,7). These are manifest in abnormally high backbone cleavage, and are presumed to arise as a consequence of changes in groove dimensions, itself a sequence-dependent property (8,9).

Crystallographic studies on a number of intercalation complexes with short lengths of double-helical DNA sequence have been reported. More than a dozen crystal structures of ribo- and deoxydinucleoside monophosphates, complexed with drugs such as proflavine, ethidium bromide and acridine orange, have been determined (10). All have very similar backbone conformations;

the values of the conformational angles of the C5'-O5' and the glycosidic bonds of the 3' ends of the duplex are each increased by about 50° over the standard A-RNA values assumed by the rest of the bonds in the backbone. There is only one crystal structure of a complex between a mono-intercalator and a longer sequence - that of daunomycin and the self-complementary hexanucleoside pentaphosphate duplex of d(CGTACG). In this structure, two drug molecules are intercalated per double helix, each between terminal CG base pairs (11). In all of these structures, the sequence at the intercalation site is pyrimidine-3', 5' purine. Thus none of these structures can show the long range structural effects of intercalation at both 5' and 3' ends.

There have been numerous molecular modelling studies of intercalation. The original conceptual Lerman model was first translated into detail by a (hand-built) structure for the single helix axis (12). Computer-generated stereochemically-plausible structures for proflavine intercalation into A- and B-DNA polynucleotide helices also preserve this overall helix (13,14). Both models have backbone torsion angles α and γ (about the P-O5' and C5'-C4' bonds) at the intercalation site in the trans range. Extrapolations have been made from the crystal structures of actinomycin-deoxyguanosine and ethidium-dinucleoside complexes, using CPK models (15,16); in both of these, the final nucleotide backbone conformations at the intercalation site were substantially different from those found experimentally. A modelling study that maintained the dinucleoside intercalation geometry in its experimental conformation produced two categories of tetranucleotide duplexes, each with a central opened-up site (17) and disrupted helicity. This also retained the experimentally-derived feature of 5' and 3' end asymmetry at the intercalation site (18); although these models are of ribotetramers, this asymmetry was subsequently observed in the crystal structures of deoxydinucleoside complexes (19), and (though not identical in detail) in the structure of the daunomycin-hexamer complex (11). A theoretical study (20) has indicated that there are probably a considerable number of approximately equi-energetic ways of altering an A- or B- form nucleotide backbone so as to increase base-pair separation to 6.8Å for intercalation.

Interestingly, the values assumed by C5'-O5' and the glycosidic angles in the x-ray studies were not found in this analysis because of unfavorable steric contacts. A number of other theoretical intercalation geometries that are quite distinct from and unrelated to the experimental ones, have

been derived. One type (21) has tetranucleotide duplexes, with only the ends constrained to have B-DNA geometry. Intercalation geometries have been constructed (22,23) in a B-DNA hexanucleotide, by changing the conformation at the central dinucleoside so that in particular, the torsion angles α and γ are in the trans domain. Analogous structures have also resulted from a molecular dynamics simulation of intercalation (24).

The present study reports on a new model for intercalation into a relatively long (decamer) length of B-DNA with a drug embedded in the centre of the sequence. The geometry for the intercalation site itself has been deliberately taken from X-ray crystallographic results on dinucleoside complexes (19); the four base pairs on either side have been allowed to relax from their initial B-DNA conformation so as to compensate for the helical disruption at the intercalation site. We have chosen to use a molecular mechanics force-field (25) that has been widely employed in the study of DNA and DNA-ligand structure, and is generally accepted as being capable of reliably predicting experimental data when it is available, such as from NMR studies (26). We detail here the structural ramifications produced throughout the helix by the intercalation site, and the sequence-dependent features of the intercalated decamer.

METHODS

The crystal structure of the d(CpG)-proflavine intercalation complex (19) was used as a starting point, together with two pentanucleotides in a classic B-DNA conformation. Segments were positioned relative to the intercalated dimer by transforming their coordinates by the rotation matrix and translation vector that best superimposed the atoms of the base pair at the end of the segment, with the corresponding base pair in the intercalated dimer. A combined coordinate set was then produced by deleting the sugar and base atoms of the overlapped ends of each segment. This resulted in a double helix of sequence d(GpApTpApCpGpApTpApC); the CG base pairs at the ends were used in order to minimize "fraying" during subsequent energy refinements. A proflavine molecule was docked on a computer graphics display into the intercalation site at approximately the position observed in the dimer crystal structure (19). The geometry of this initial structure was generally satisfactory, with poor geometry restricted to the phosphate backbone at the junctions between the end segments and the central intercalated dimer. Regularization was

accomplished by means of a molecular mechanics force-field, which removed all sterically unacceptable close contacts between atoms. The structure was energy-minimized by means of a conjugate-gradients technique. The total energy of the system was considered as the sum of several component energies:

$$E_{\text{TOTAL}} = \sum k_r (r - r_0)^2 + \sum k_\theta (\theta - \theta_0)^2 + \sum V_n / 2 [1 + \cos(n\phi - \gamma)] \\ + \sum [A_{ij}/R_{ij}^{12} - B_{ij}/R_{ij}^6] + \sum q_i q_j / \epsilon R_{ij}$$

where the right-hand side terms represent bond length, bond angle, torsional, non-bonded and electrostatic potentials. The all-atom parameter set used for the DNA was essentially identical to that derived by Kollman and his associates (25). Partial charges for proflavine were obtained by the CNDO/2 method using the crystal-structure coordinated for the drug (27). A distance-dependent dielectric constant, of $\epsilon = 4R_{ij}$ was used; other workers have found that this gives satisfactory results in the absence of explicit solvent and counter-ion shielding (28). Convergence was judged to have been achieved after each component of the gradient vector had an absolute value less than machine precision. In practice, this meant that at convergence the change in total energy was less than 10^{-3} kcal mole $^{-1}$ Å $^{-1}$.

The energy refinements were performed with the program EMPMDS (29), implemented on a VAX-11/750 computer. Molecular graphics manipulations were made on a Silicon Graphics IRIS workstation using the GEMINI molecular modelling package (30). Calculations of helix and base parameters were made with the helix analysis programs of Rosenberg and Dickerson (31). Refined coordinates for the decamer model are available from S. Neidle.

RESULTS

A. General Features

The initial structure has a number of close intramolecular contacts; the energy refinement has removed these and resulted in a stereochemically plausible, low-energy structure. This is shown in Fig.1. There are significant changes to many of the backbone torsion angles. The result is a smoothly-deformed structure with a greatly widened major groove (Table 1). The maximum width is now 17.6 Å, compared to 11.4 Å in classic B-DNA (32). More surprisingly, the minor groove width has also increased to a maximum of 9.1 Å, which compares to that of classic A-DNA (11.0 Å), as much as

Fig. 1. Stereo plot of the decamer structure, viewed into the major groove. The proflavine molecule is shown with its long axis aligned to the horizontal axis.

TABLE 1

Groove dimensions in the proflavine-decamer model, in Å

a) Major Groove		
Strand 1	Strand 2	Distance
P2	P17	14.2
P3	P16	14.9
P4	P15	16.6
P5	P14	17.6
P6	P13	14.6
P7	P12	15.0
b) Minor Groove		
P5	P20	4.2
P6	P19	7.6
P7	P18	9.0
P8	P17	7.9
P9	P16	7.4
P10	P15	4.8

Groove width is defined as the shortest inter-strand phosphorus-phosphorus distance minus 5.8Å.

TABLE 2

Conformational features of the decamer intercalation model (a) before and (b) after energy refinement. P is the pseudorotation phase angle for the deoxyribose sugars. The conformational angles are in the standard convention (30)

		P	CHI	ALPHA	BETA	GAMMA	DELTA	EPSILON	ZETA
(a) strand 1	GUA	151	256	0	137	33	142	219	202
	ADE	151	256	326	137	33	142	219	202
	THY	151	256	326	137	33	142	219	202
	ADE	151	256	326	137	33	142	219	273
	CYT	22	198	244	121	54	84	209	289
	GUA	38	251	289	218	46	94	56	50
	ADE	151	256	326	137	33	142	219	202
	THY	151	256	326	137	33	142	219	202
	ADE	151	256	326	137	33	142	219	202
	CYT	151	256	326	137	33	142	168	255
strand 2	GUA	151	256	136	35	33	142	219	202
	THY	151	256	326	137	33	142	219	202
	ADE	151	256	326	137	33	142	219	202
	THY	151	256	326	137	33	142	219	261
	CYT	17	194	252	113	53	79	203	299
	GUA	174	284	287	217	72	149	261	160
	THY	151	256	326	137	33	142	219	202
	ADE	151	256	326	137	33	142	219	202
	ADE	151	256	326	137	33	142	219	202
	CYT	151	256	326	137	33	142	349	59
(b) strand 1	GUA	82	189	0	184	61	86	191	284
	ADE	139	236	288	179	58	129	180	265
	THY	95	219	287	178	59	93	176	272
	ADE	131	237	297	171	61	122	187	271
	CYT	67	230	286	167	56	81	190	279
	GUA	82	242	297	182	66	81	46	52
	ADE	130	240	224	183	59	120	176	267
	THY	133	239	290	177	58	126	190	225
	ADE	160	242	294	162	64	143	182	272
	CYT	113	222	290	170	60	110	178	267
strand 2	GUA	127	221	127	12	59	119	189	271
	THY	71	211	278	173	55	81	176	277
	ADE	161	248	296	174	67	142	182	253
	THY	127	234	289	176	58	120	181	259
	CYT	26	214	277	167	60	75	196	289
	GUA	149	273	294	202	70	145	272	166
	THY	68	212	283	119	54	76	174	273
	ADE	160	249	298	177	62	142	177	267
	THY	123	229	292	173	61	114	182	272
	GYT	118	225	292	167	61	114	294	69
	A-DNA		255	270	210	47	83	185	215
	A-RNA		194	300	185	49	83	213	281

that of B-DNA (6.0Å). On the other hand, the depths of the grooves are closer to those seen in B-DNA in that they are both relatively shallow. For both grooves, the distortions are maximal at the intercalation site and appear to increase in a continuous manner from the ends of the helix. Table 2 details the conformational features of the model. The changes consequent to energy refinement have affected all torsion angles, with a number of alterations from the overall B-type conformational class. The asymmetry of the 3' and 5'

Fig. 2. Correlation plot of glycosidic angles versus the C4'-C3' conformational angles (δ) for the decamer model. Values with open squares have been excluded from the least-squares straight line.

glycosidic angles at the intercalation site that was observed in dinucleoside phosphate-drug complexes (18) is retained only on strand 2, with a difference of 59° between the two values. The increased value for beta seen in the dinucleoside phosphate complexes is not retained in the refined model. The sugar puckers at the intercalation site were C3'endo-3',5'-C3'endo and C3'endo-3',5'-C2'endo for strands 1 and 2 respectively before regularization of the model. In the final refined model, the cytidine and guanosine sugars at the intercalation site in strand 1, although generally still in the C3'endo domain, can more accurately be described as C4'exo,O4'endo. This pucker is also adopted by the thymidine sugars on the 5' side of the intercalation site in strands 1 and 2, as well as by the thymidine immediately on the 3' side in strand 2. The gauche⁺ value of the C3'-O3' angle on the 3' side of the central guanosine in strand 1 is matched by a gauche⁻ value for the corresponding strand 2 guanosine. The torsion angle around the P-O5' bond in the adenosine immediately 3' to the central CG in strand 1 has an almost trans value of 224° .

The O5'-C5' angle is mostly trans in the decamer, except for the 3' side thymidine in strand 2, which is complementary to this adenosine. Overall then, the imposition of the intercalated dinucleoside backbone geometry on a B-DNA sequence has particularly effects on the torsion angles of the residues on the 3' side.

The values of the backbone torsion angles in the sites other than the intercalated one have refined from their initial B-DNA values to ones closer to A-RNA. The values of the glycosidic torsion angles are intermediate between A- and B-DNA. The sugar puckers are highly variable and range from C3'endo to C2'endo with no clearly discernible pattern. It is important to point out that although many of the backbone angles are in the A-class the overall appearance of the helix remains B-like.

Correlations between conformational parameters are not in general found in this model, in striking contrast to the situation found in the crystal structures of A- and B-type oligonucleotides (31-33). An exception is the plot of δ vs. glycosidic angle (Fig. 2) that shows a high degree of correlation (with a correlation coefficient of 0.9 for the least-squares straight line), once the values for the residues immediately at the intercalation site are excluded. Backbone torsion angles ϵ and ζ are not correlated, nor are α and γ , presumably in large part on account of the changes brought about by the intercalation geometry.

B. Sequence-dependent Features

We have used the definitions of Dickerson et al. (31) for describing local variations in helical and base-pair parameters (Table 3). An overall helix axis was generated using base pairs 2,3 and 8,9 in order to minimize both end-effects and changes around the central CG part of the decamer. The global helical twist angles t_g , defined as the angle subtended between C1'-C1' vectors in successive base pairs, are shown in Fig. 3. There is large variation in t_g along the sequence, with the value at the CpG intercalation site itself being 21°. This is the smallest t_g value in the sequence, which rise to maximum values at the TpA steps. Base-pair propeller twist (Fig. 4), defined as the twist about the long axis, also changes markedly along the decamer sequence, with a minimum value of 2° for the 5' end CpG base pair at the intercalation site.

Surprisingly, the adjacent CG base pair is twisted by 17°. Overall, the clear trend in propeller twist is for increasing values towards the termini of the duplex. Base-pair roll θ_R defined as the change in orientation of the mean plane

TABLE 3
Helical and Base-pair parameters.

a) For base pairs, using the best plane through each.				
Base pair	Local tilt θ_R (°)	Local roll θ_R (°)	Global roll Φ_R (°)	Propeller twist (°)
1	1.5	4.9	-49.8	28.0
2	1.2	-3.8	-37.4	24.5
3	0.0	5.3	-24.1	18.9
4	0.7	-4.9	16.9	10.2
5	-0.6	-5.0	26.4	2.4
6	0.0	0.8	32.5	16.7
7	0.7	-1.3	37.9	15.1
8	-2.3	4.9	30.7	8.9
9	-1.3	1.4	14.0	21.6
10	-	-	-6.7	27.0

Base pairs are defined with respect to the 5' end of strand 1.

<u>Base pairs</u>	<u>Helical twist</u> <u>tg (°)</u>	<u>Rise per base</u> <u>pair (Å)</u>
1/2	28.5	2.1
2/3	35.6	3.4
3/4	57.1	3.3
4/5	33.0	3.1
5/6	20.9	5.3
6/7	28.9	2.4
7/8	25.2	2.1
8/9	44.3	2.2
9/10	34.8	0.7

b) for individual bases.

Strand 1				Strand 2		
<u>Base Pair</u>	<u>θ_R</u>	<u>θ_T</u>	<u>Φ_R</u>	<u>θ_R</u>	<u>θ_T</u>	<u>Φ_R</u>
1	-2.2	7.1	-62.7	2.5	-1.2	-35.5
2	-4.2	2.4	-52.4	1.0	-3.3	-28.8
3	-1.8	7.3	-31.7	-0.6	-8.6	-14.0
4	1.1	7.2	9.3	-7.7	0.3	19.5
5	-10.2	1.5	25.5	0.4	-2.3	27.9
6	-2.5	0.8	25.4	-5.6	-9.0	42.0
7	1.7	4.9	28.3	-3.7	-1.2	43.1
8	-5.8	6.1	27.4	3.7	-11.9	36.2
9	-0.6	1.6	1.9	0.4	-2.0	23.1
10	-	-	-20.7	-	-	6.2

between adjacent base pairs in the direction of the major or minor groove, is shown in Fig. 5. A positive value for θ_R denotes opening towards the minor groove direction. The plot

Fig. 3. Plot of helical twist angles for steps in the decamer structure, calculated by the method of ref. 30.

Fig. 4. Plot of propeller twist angles for steps in the decamer structure.

Fig. 5. Plot of local base-pair roll angles for steps in the decamer structure.

shows that there is considerable variation in roll along the sequence, with the steps at the intercalation site being opened towards the major groove. Both of the TpA steps, on the 3' and 5' sides, are rolled open in the opposite sense, to the minor groove, and the adjacent ApT ones are rather more gently opened to the major groove. The θ_R values alternate between positive and negative values along the duplex, with the sole exception being at the central CpG steps.

Examination of the local roll values for each strand separately rather than using the best plane through each base pair, as in Fig. 5, shows (Table 2) a rather different picture, with individual bases mostly having negative θ_R values. The opening towards the major groove is concentrated around the intercalation site, although interestingly the two strands have reacted in a very different manner. Strand 1 has rolled open mostly at Step 5, whereas strand 2 has complemented this by opening up at steps 4 and 6. The lack of alternation of θ_R values for each strand compared to the base-pair values is a consequence of the varying propeller twists along the duplex. Local roll θ_R contrasts with global roll, which measures individual base pair inclination about its short axis in relation to the overall calculated helix axis.

Fig. 6 shows that the global roll Φ_R changes from a

Fig. 6. Plot of global roll angles for base pairs in the decamer structure.

minimum value of -50° at the 5' end to a maximum of $+38^\circ$ at the 7th base pair. The overall effect is that the duplex is roughly symmetrically bent away from the major groove (Fig. 7), in a right-handed supercoiling manner. The global roll values for each strand follow this trend.

Base tilt Θ_T , defined as the change in the orientation of an average base-pair plane in the direction of the backbone, is consistently very small along the length of the decamer (Table 2). The increases towards the termini are most likely due to end effects. The local tilt values for individual strands are considerably larger, with the consistent trend of strand 1 having positive Θ_T values and strand 2 having negative ones. Thus, the base pairs in the sequence are bent to varying degrees along their short axis, with in general increasing tilt towards the 5' and 3' ends.

Fig. 7. Stereo plot of the decamer complex structure, rotated 90° from the view in Fig. 1. The major groove is on the right-hand side of this view, and the bend away from it is clearly visible.

C. The Geometry at the Intercalation Site

It is notable that the two CG base pairs at the intercalation site are non-equivalent either in conformation or in helical geometry terms. The 5' side base pair is almost planar, whereas the 3' one has a propeller twist of almost 17° . The latter has especially asymmetric overlap with the intercalated proflavine molecule, one ring of the drug being in almost perfect overlap with the cytosine base whereas the guanine has little effective overlap with it.

This cytosine has essentially zero overlap with its 3' adjacent thymine base, whereas the complementary guanine partially overlaps its adjacent 3' adenine.

Furthermore, the same cytosine is relatively coplanar with the drug chromophore, in contrast to the considerable propeller twist of the complementary guanine with respect to the chromophore, imparting to this base pair a high propeller twist. The guanine in turn has only slight roll and tilt with its 3'-adjacent adenine on the same strand, and thus this AT base pair (number 7 in Table 2) similarly has a high propeller twist.

The CG base pair at the 5' side of the intercalation site has generally low overlap with the proflavine chromophore, although the O6 atom of the guanine is stacked almost precisely vertically above the protonated ring nitrogen of the drug, and is 3.3Å distant from it.

DISCUSSION.

Although there is not as yet any direct experimental data to compare with the model resulting from this study, all of the available X-ray crystallographic evidence on short-length oligonucleotide intercalation complexes strongly suggests that the nucleotide residues immediately at the site, are helically asymmetric with respect to each other. The model presented here is distinct from others (21-24) in its direct derivation from the structures obtained from X-ray crystallography. In general the molecular mechanics method does not generally allow structures to move out of local energy minima, as was seen in the earlier intercalation models (23,24) that used the all atom force field (25). In this case the starting decamer model required considerable conformational change in order to remove the initial high-energy non-bonded contacts at the junctions between pure B-DNA and dinucleoside intercalation regions. It should be noted that previous models (13,14,21-24) were constructed with maximum avoidance of steric clashes, largely by means of trans values for the angles α and γ at the intercalation sites - this avoided helical disruption at 3' ends of the sites and thus refined structures were close to starting ones. It had been observed previously that the intercalated dinucleoside phosphate drug complexes fall into the A-RNA conformational class (17). The present study shows that introduction of just two residues with this geometry into an otherwise B-helix has a far reaching effect on all the conformational angles so as to force many of them into the A-conformational class. Of course, the possibility exists that these results are more a consequence of the inherently lower energy of the A-conformation especially considering the absence of solvent and ions in the calculation. A number of modelling studies of DNA oligonucleotides using equivalent force fields and closely analogous methods have shown that both A- and B-type starting points tend to remain in their same local energy domains when they are standard helices (22,40-42) and do not undergo the rather large changes reported here. It thus seems that the helical disruption emanating from the intercalation site is responsible for these changes.

The global effects of intercalation are most marked on groove dimensions, with the major groove in particular increasing its width, by over 6Å at its widest point compared to classic B-DNA. These increased groove dimensions are also much larger than the values observed in the Dickerson dodecamer crystal structure (31,33,34).

They are primarily a consequence of the increase in base-pair separation at the intercalation site combined with a decrease in twist angle t_g , which on purely geometric grounds forces phosphate groups further apart and thus grooves to be widened. The present model does not show the distinctive feature of close phosphate-phosphate distances found in our previous modelling study (17) of an intercalated ribotetranucleotide duplex; the electrostatic component in the energy minimisation used in the present study would tend to maximise phosphate-phosphate separations in the absence of any specific constraints to this. The intercalating drug modelled in this study is bound from the major groove direction - its exocyclic amino groups preclude minor groove binding on account of steric hindrance (35). It would then be unsurprising if a minor groove intercalator such as ethidium or daunomycin produced detailed differences in groove dimensions compared to the model presented here. Calculations with these drugs are currently underway. It is noteworthy that the minor groove in the present model is too wide to accommodate a spine of hydration. It has been pointed out that large propeller twists in B-DNA correlate with narrowing of the minor groove (36,37). To a large extent, we see this feature in the decamer model, with the global minimum propeller twist at the intercalation step of the sequence, being where the grooves are at their widest.

Examination of Figs. 3 and 5 shows a high correlation between helical twist and base-pair roll, with the largest values of each being at the TpA (pyrimidine-3',5'-purine) steps. As in other respects, the central CpG pyrimidine-3',5'-purine step is the exception to this pattern. Comparison with the variations in t_g and θ_R values for the Dickerson dodecamer are instructive; experimentally-observed large roll angles have been observed at CpG steps, as in the present model. On the other hand, the dodecamer has the lowest values for helical twist t_g between base pairs at these steps (37), in total contrast to the pattern found here (Fig. 3). This is probably due to several interrelated factors; firstly, the unwinding at the intercalation site, with its low t_g value appears to produce compensatory over-winding two

base-pairs distant on both 5' and 3' sides, which are maximal at these two TpA steps. Another is probably the transmitted effect of the effectively zero propeller twist and tilt of the drug chromophore at the centre of the duplex - this in effect acts as a pseudo-base and produces increasingly larger propeller twists in the base pairs on going from step 5 in either 3' or 5' direction. The interruption in this progression (Fig. 4) at steps 7 and 8 is probably due to the purine-purine GpA sequence here, which as noted in the dodecamer structure (33,34,38), has minimal steric clash, necessitating only a small amount of relief by propeller twisting. The origin of the differences in propeller twist for the two CG base pairs at the intercalation site is probably at least in part electrostatic.

The 5' side coplanar base pair has the O6 atom of guanine, with a strongly negative electrostatic potential, close to the cationic ring nitrogen atom of the drug, whereas no such close attractive interaction takes place with the 3' side guanine. Thus, the latter is free to adopt its inherently preferred propeller twist, which is probably accentuated by the necessity to relieve a steric clash between the 3' thymine methyl group on strand 2 and a backbone sugar group. How does this (theoretical) structure conform to Calladine's rules? It is clear from the preceding discussion that the nature of the changes in helix structure at various points on a sequence depend on the directionality of the perturbation (i.e. intercalation) produced. In the present case, with major groove insertion at a CpG site, the site itself violates the rules in respect of roll opening. This is unsurprising since the inserted molecule is not a base pair. More remote from the site, the direction of roll (θ_R) openings is in accord with the rules. Features of propeller and helical twist are not, and it can thus be assumed that again, this is due to the profound long-range influence of the drug insertion, producing non-standard A-or B-type conformations in the sequence. It is perhaps then inadvisable to categorise the conformations in the model in A or B terms, since they actually display a continuum of values between these classic extremes.

The long-range effects of the intercalation upon a duplex in terms of groove dimension changes are of direct implication to the observations in footprinting experiments of enhancements to DNase I cutting ability at sequences both 5' and 3' to actual binding (protected) drug sites. This has been recently demonstrated (39) in the case of daunomycin. The crystal structure analysis of a DNase I-oligonucleotide complex (40) has

shown marked minor groove widening to accommodate the enzyme, and it is thus tempting to infer that the minor groove widening found in the present molecular modelling study is the direct cause of the enhancement of enzyme cutting found experimentally. The increase of 3Å in our model for the minor groove width compared to B-DNA agrees remarkably well with that found in the enzyme-DNA complex (40), and suggests that the widening found here is optimal for enzyme attack on a binary drug-DNA complex.

ACKNOWLEDGEMENTS

We are grateful to the Cancer Research Campaign for support to SN and for providing a visiting Fellowship to PH. Richard Dickerson is thanked for provision of his helix analysis program. HMB is grateful for support from grants NIH GM21589, CA06927, RR05539, and from an appropriation of the Commonwealth of Pennsylvania.

*To whom correspondence should be addressed

**On leave from the Institute of Biochemistry and Biophysics, Polish Academy of Sciences, 02-532 Warsaw

REFERENCES

1. Wilson, W.D. and Jones, R.L. (1981) Adv. Pharmacol. Chemotherapy 18, 177-222.
2. Berman, H.M. and Young, P.R. (1981) Ann. Rev. Biophysics 10, 87-114.
3. Neidle, S., Pearl, L.H. and Skelly, J.V. (1987) Biochem. J. 243, 1-13.
4. Pommier, Y., Schwartz, R.E., Zwelling, L.A. and Kohn, K.W. (1985) Biochemistry 24, 6404-6410.
5. Yang, L., Rowe, T.C., Nelson, E.M. and Liu, L.F. (1985) Cell 41, 127-132.
6. Dervan, P.B. (1986) Science 232, 464-471.
7. Ford, K., Ford, K.R., Neidle, S. and Waring, M.J. (1987) Nucleic Acids Res. 15, 2221-2234.
8. Drew, H.R. and Travers, A.A. (1984) Cell 37, 491-502.
9. Drew, H.R. and Travers, A.A. (1985) Nucleic Acids Res. 13, 4445-4467.
10. Conformational details are given in: Neidle, S. (1988) Landolt-Bornstein, (Saenger, W., Ed.), in press.
11. Wang, A.H.-J., Ughetto, G., Quigley, G.J. and Rich, A. (1987) Biochemistry 26, 1152-1163.
12. Fuller, W., and Waring, M.J. (1964) Berichte Bunseng. Physik. Chem. 68, 805-808.
13. Alden, C.J. and Arnott, S. (1975) Nucleic Acids Res. 2, 1701-1717.
14. Alden, C.J. and Arnott, S. (1977) Nucleic Acids Res. 4, 3855-3861.

15. Sobell, H.M. and Jain, S.C. (1972) J. Mol. Biol. **68**, 21-34.
16. Sobell, J.M., Tsai, C.-C., Jain, S.C. and Gilbert, S.G. (1977) J. Mol. Biol. **114**, 333-365.
17. Berman, H.M. and Neidle, S. (1979) In Stereodynamics of Molecular Systems, R.H. Sarma, Ed., Pergamon Press, New York, 367-382.
18. Berman, H.M., Neidle, S. and Stodola, R.K. (1978) Proc. Natl. Acad. Sci. USA **72**, 828-832.
19. Shieh, H.-S., Berman, H.M., Dabrow, M. and Neidle, S. (1980) Nucleic Acids Res. **8**, 85-97.
20. Taylor, E.R. and Olson, W.K. (1983) Biopolymers **22**, 2667-2702.
21. Miller, K.J. and Pycior, J.F. (1979) Biopolymers **18**, 2683-2719.
22. Lybrand, T. and Kollman, P. (1985) Biopolymers **24**, 1863-1879.
23. Lybrand, T.P., Brown, S.C., Creighton, S., Shafer, R.H. and Kollman, P.A. (1986) J. Mol. Biol. **191**, 495-507.
24. Prabhakaran, M. and Harvey, S.C. (1985) J. Phys. Chem. **89**, 5767-5769.
25. Weiner, S.J., Kollman, P.A., Nguyen, D.T. and Case, D.A. (1986) J. Comp. Chem. **7**, 230-252.
26. Woodson, S.A. and Crothers, D.M. (1988) Biochemistry **27**, 3130-3141.
27. Jones, T.A. and Neidle, S. (1975) Acta Cryst. **B31**, 1324-1333.
28. Whitlow, M. and Teeter, M.M. (1986) J. Am. Chem. Soc. **108**, 7163-7172.
29. Haneef, I. (1986) Ph.D. Thesis, University of London.
30. Beveridge, A., Pearl, L.H. and Neidle, S., to be published.
31. Fratini, A.V., Kopka, M.L., Drew, H.R. and Dickerson, R.E. (1982) J. Biol. Chem. **257**, 14686-14707.
32. Shakked, Z. and Rabinovich, D. (1986) Prog. Biophys. Molec. Biol. **47**, 159-195.
33. Dickerson, R.E., Kopka, M.L. and Pjura, P. in Biological Macromolecules and Assemblies, Vol. 2 (Jurnak, F.A. and McPherson, A., Eds.), John Wiley, New York, 37-126.
34. Dickerson, R.E. and Drew, H.R. (1981) J. Mol. Biol. **149**, 761-786.
35. Islam, S.A. and Neidle, S. (1984) Acta Cryst. **B40**, 424-428.
36. Dickerson, R.E., Kopka, M.L. and Drew, H.R. (1982) In Structure and Dynamics: Nucleic Acids and Proteins, Clementi, E. and Sarma, R.H., Eds.), Adenine Press, New York.
37. Dickerson, R.E. (1983) J. Mol. Biol. **166**, 419-441.
38. Calladine, C.R. (1982) J. Mol. Biol. **161**, 343-352.
39. Chaires, J.B., Fox, K.R., Herrera, J.E., Britt, M. and Waring, M.J. (1987) Biochemistry **26**, 8227-8236.
40. Suck, D., Lahm, A. and Oefner, C. (1988) Nature **332**, 464-468.