DNA–Ditercalinium Interactions: Implications for Recognition of Damaged DNA†

Loren Dean Williams*,‡ and Qi Gao§

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received October 7, 1991; Revised Manuscript Received February 10, 1992

ABSTRACT: Ditercalinium is unique among known DNA-binding chemotherapeutic agents. Ditercalinium treatment of *Escherichia coli* causes cell death by provoking malfunction of the (A)BC exinuclease excision DNA repair system. In this report, we describe the three-dimensional X-ray structure of a ditercalinium–$[d(CGCG)]_2$ complex in detail with an analysis of both the structure and its implications. Ditercalinium bisintercalates in the DNA fragment; the positively-charged linker of the drug interacts with the major groove. The DNA retains an underwound, right-handed, double-helical conformation with a bend of around 15° in the helical axis. One striking feature of the complex is the extensive interaction of ditercalinium with guanines in contrast to the near absence of interaction with cytosines. The terminal cytosines in particular are "unstacked" from the rest of the complex. A systematic comparison of the three-dimensional structure of the DNA–ditercalinium complex with those of triostin A and nogalamycin (chemotherapeutic agents that act by conventional mechanisms) allows us to suggest a general model for recognition by the (A)BC exinuclease excision repair system. It is commonly hypothesized that the same distorted conformation of DNA results from modification by each member of a diverse family of DNA-damaging agents. This specific conformation of DNA would then be recognized by the (A)BC exinuclease excision repair system. Alternatively, we propose that each of these damaging agents causes local instability of DNA (but not necessarily a common conformation) and that the (A)BC exinuclease excision repair system recognizes excessive or unusual deformability of damaged DNA in comparison to normal DNA.

Many chemotherapeutic agents act at the DNA level. DNA-binding antibiotic, antiviral, antitrypanasomal, and anticancer agents can inhibit such processes as DNA replication or transcription. Chemotherapeutic agents bind to DNA within the grooves, ordinarily within the minor rather than the major groove, and/or intercalate between the bases. In general, molecules that intercalate in DNA also interact with the minor groove and thus should be designated intercalating minor groove binders. Examples of clinically important intercalating minor groove binders are daunomycin and the closely related adriamycin, currently two of the most widely used anticancer chemotherapeutic agents (Acramone & Penco, 1989).

In a search for new and better chemotherapeutic agents, DNA intercalators have been oligomerized to form compounds with DNA affinities as great as 10^{14} M^{-1} (Laugaa et al., 1985). One result of this type of search is the synthetic 7*H*-pyridocarbazole dimer, ditercalinium (Figures 1A and 2A,B). Using rigid linkers to couple planar compounds, ditercalinium was designed/discovered by Roques, Le Pecq, and co-workers (Pelaprat et al., 1980a,b) in a systematic quest for improved chemotherapeutic agents.

Ditercalinium is unique among known chemotherapeutic agents, causing cell death by an unprecedented mechanism. Unlike most other DNA-binding chemotherapeutic agents, the mechanism of action of ditercalinium does not involve inhibition of DNA replication or transcription processes. Instead, DNA–ditercalinium complexes provoke malfunction of DNA repair systems (Lambert et al., 1989, 1988, 1990). The three-dimensional structure of a DNA complex with ditercalinium should help us understand the unusual activity of this drug. In general, the increasing number of three-dimensional structures of DNA–drug complexes provides the opportunity to relate differences in activity to differences in molecular structure.

We have recently reported in a preliminary fashion (Gao et al., 1991) the three-dimensional X-ray structure of a complex of ditercalinium bound to the double-stranded DNA fragment $[d(CGCG)]_2$. Space-filling representations of this complex are shown in Figure 3. In this report, we describe the complex in detail with an analysis of both the structure and its implications. A systematic comparison of the three-dimensional structure of the DNA–ditercalinium complex with those of chemotherapeutic agents that act by conventional mechanisms, triostin A (Figures 1B and 2C,D) and nogalamycin (Figure 1C), allows us to suggest certain aspects of a general model for recognition by DNA repair systems. Space-filling representations of a triostin A complex are shown in Figure 4, and space-filling representations of nogalamycin and a DNA–nogalamycin complex can be found in Egli et al. (1991a).

Materials and Methods

Crystals of the ditercalinium–$[d(CGCG)]_2$ complex were grown at room temperature in sitting drops using the vapor diffusion method with crystallization conditions as described previously (Gao et al., 1991). The complex crystallized in space group $P4_12_12$ with unit cell dimensions $a = b = 26.88$ Å and $c = 82.60$ Å. Details on data collection and reduction were described previously.

Patterson maps revealed that the number of stacked units (base pairs plus intercalators) was six and that the helical axis

† The coordinates of the DNA–ditercalinium complex have been deposited in the Brookhaven Data Bank under Identifier Code 1D32.

* Address correspondence to this author.

‡ L.D.W. acknowledges fellowship support by the National Institutes of Health and the Medical Foundation, Boston, MA.

§ Current address: Analytical Research Department, Bristol-Meyers Squibb Corporation, Wallingford, CT 06492.

0006-2960/92/0431-4315$03.00/0

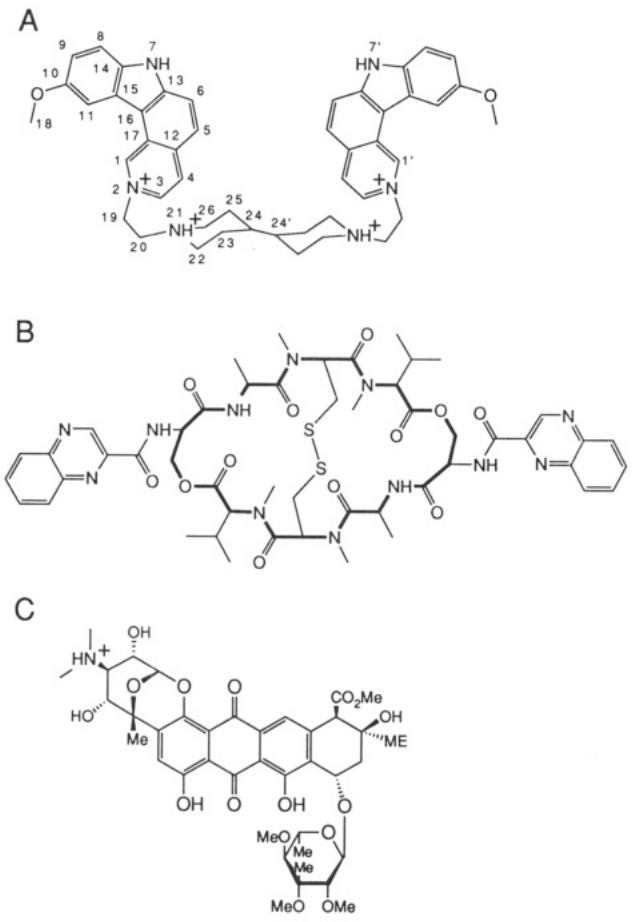

FIGURE 1: Structures of (A) ditercalinium, (B) triostin A, and (C) nogalamycin.

was aligned nearly parallel to the crystallographic *c*-axis. Assuming that the complex consisted of one DNA tetramer duplex plus one ditercalinium molecule and considering the volume per base pair (either base pairs or chromophores), it was deduced that the unit cell contained eight complexes which thus must adopt general positions. A model of ditercalinium was constructed on the basis of ideal geometry as there were no crystal structure data available. This model of ditercalinium was then docked onto a tetramer duplex which was obtained by linking two CpG steps derived from the crystal structure of the d(CGATCG)–adriamycin complex (Frederick et al., 1990). Many features of the search model, such as the location of the linker within the major groove, were derived from published NMR experiments (Delbarre et al., 1987a,b, 1985; Delepierre et al., 1988, 1989). The search model was refined to stereochemical restraints.

Structure determination by molecular replacement was not straightforward using either X-PLOR (Brunger, 1988) or ULTIMA (Rabinovich & Shakked, 1984) in part due to the deviations of the search model from the real complex. Although the initial maps were ambiguous, the two programs yielded nearly the same set of solutions and the correct solution was obtained after careful analysis of the symmetry elements and molecular packing. The solution was confirmed by further refinement.

The next stage consisted of extensive model rebuilding and reorientation of ditercalinium with respect to the DNA. By displaying Fourier sum ($2F_o - F_c$) and difference ($F_o - F_c$) maps on an Evans & Sutherland PS390 graphics station with the program FRODO (Jones, 1982), the chromophores and the linker of the ditercalinium molecule were manually shifted to more closely fit the electron density. During Konnert–Hendrickson least-squares refinement (Hendrickson & Konnert, 1981), extensive efforts were made to definitively establish the conformation of the linker. Parallel refinements on all possible rotamers had the same *R*-factor. Therefore, the most ste-

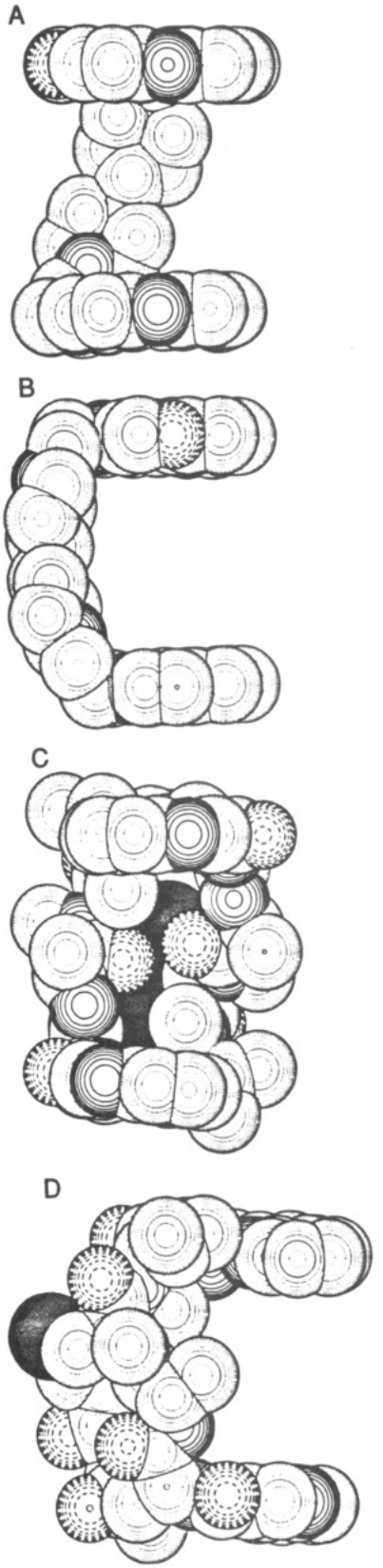

FIGURE 2: Space-filling representations of ditercalinium viewed (A) into the best plane of the linker and (B) across the best plane of the linker, and triostin A viewed (C) into the best plane of the linker and (D) across the best plane of the linker. Carbon atoms are dotted, oxygen atoms are dashed, and nitrogen atoms are concentric circles. The sulfur atoms of triostin A are shaded.

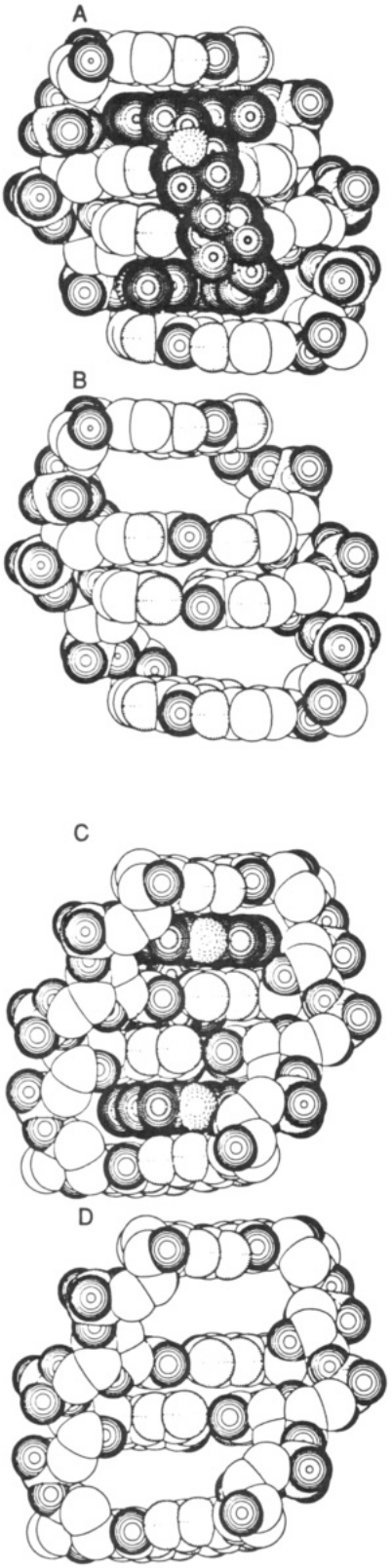

FIGURE 3: Space-filling representations of the ditercalinium–[d-(CGCG)]$_2$ complex: (A) the complete complex viewed into the major groove, (B) the same view with the ditercalinium molecule omitted, (C) the complete complex viewed into the minor groove, and (D) the same view with the ditercalinium molecule omitted. In the DNA, carbon atoms are white, nitrogen atoms are dotted, oxygen atoms are lined, and phosphorus atoms are dashed. In the ditercalinium molecule, concentric circles with radial lines were used for all atoms except for the nitrogen atoms, which are dashed.

reochemically reasonable (most energetically favorable) conformation was chosen for final refinement.

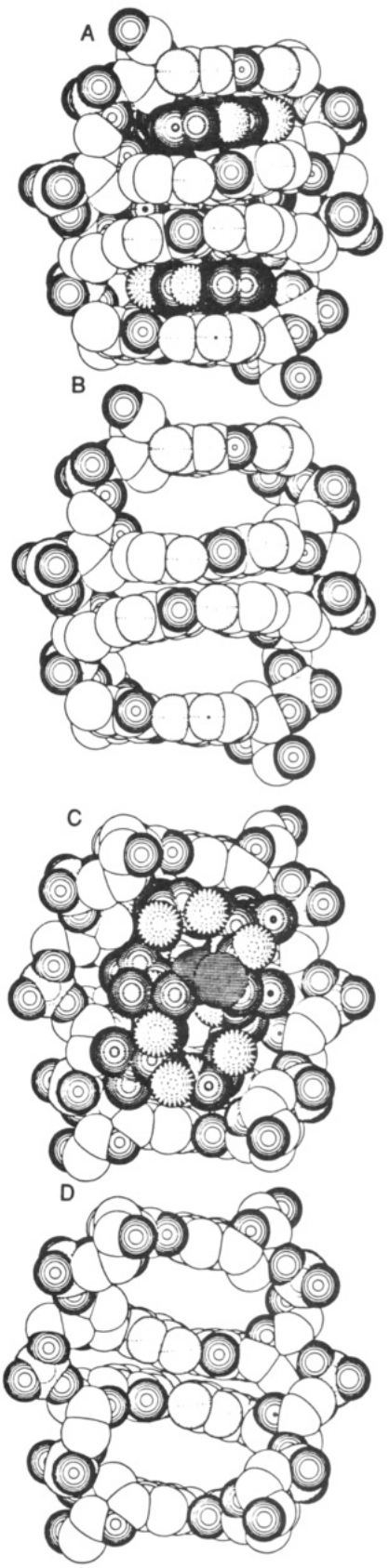

FIGURE 4: Space-filling representations of the triostin A–d-(GCGT)·d(ACGC) complex: (A) The complete complex (i.e., the asymmetric unit of the complex) viewed into the major groove, (B) the same view with the triostin A molecule omitted, (C) the complete complex viewed into the minor groove, and (D) the same view with the triostin A molecule omitted. In the DNA, carbon atoms are white, nitrogen atoms are dotted, oxygen atoms are lined, and phosphorus atoms are dashed. In the triostin A molecule, concentric circles with radial lines were used for all atoms except for the nitrogen atoms, which are dashed, and the phosphorous atoms, which are shaded.

Table I: Local Inter-Base-Pair Parameters[a]

step	shift (Å)	slide (Å)	rise (Å)	tilt (deg)	roll (deg)	twist (deg)
C(1)·G(8)–G(2)·C(7)[b]	0.1	−0.3	6.9	0.7	−0.7	18.8
G(2)·C(7)–C(3)·G(6)	0.2	−0.1	3.6	−3.4	−8.4	28.6
C(3)·G(6)–G(4)·C(5)[b]	−0.8	−0.6	6.8	0.6	−1.1	24.3

[a] Calculated with the program Curves (Lavery & Sklenar, 1989). [b] Intercalation step.

Table II: Global Base–Base Parameters[a]

base pair	shear (Å)	stretch (Å)	stagger (Å)	buckle[b] (deg)	propeller[b] twist (deg)	opening (deg)
C(1)·G(8)	0.3	−0.1	0.0	4.3	−8.4	7.8
G(2)·C(7)	−0.5	−0.2	−0.4	7.0	−1.9	6.6
C(3)·G(6)	−0.2	−0.2	0.1	5.9	−8.7	2.4
G(4)·C(5)	0.2	−0.1	−0.1	5.1	−5.0	3.6

[a] Calculated with the program Curves (Lavery & Sklenar, 1989). [b] Estimated values.

After structure refinement was completed, a comparison of the search model with the refined structure indicated that the difficulty in solving the structure by molecular replacement was due primarily to an incorrect orientation of the ditercalinium relative to the DNA. Specifically, in the search model the linker ran roughly parallel to the helical axis, while in the final structure the linker runs diagonally across the major groove.

At the later stages of the refinement, an analysis of the intensity data revealed anisotropy in the diffraction power of the crystal. In the direction of c^*, which is approximately parallel to the helical axis of the DNA, the crystal diffracts to 1.6 Å. Perpendicular to c^*, the crystal diffracts to only 2.5 Å. A method (Shakked, 1983; Sheriff & Hendrickson, 1987) was applied to evaluate overall anisotropy in the crystal to bring the $|F_c|$ values calculated from an isotropic model into optimal agreement with anisotropic $|F_o|$ observations. Symmetry restrictions on tensor elements were taken into account, e.g., $b_{11} = b_{22}$ and $b_{13} = b_{23} = 0$. A scale factor and the elements of the anisotropic tensor were derived by a least-squares fit of the structure model and the observed structure factor moduli $|F_o|$. It appears that the thermal motion in the a^* (and b^*) direction is 25% higher then that along c^* direction. The application of the overall anisotropic scaling on the diffraction data resulted in a 5% drop of the R-factor.

The structure refinement was terminated with 84 water molecules/asymmetric unit. The calculated crystal density is 1.00g/cm^3 with water content of 34% w/w. The R-factor converged to 22.5% at 1.7-Å resolution. The final electron density maps showed clean and continuous density around the complex. The rms deviation from ideal bond lengths is 0.019 Å. The coordinates of the complex have been desposited in the Brookhaven Data Bank under Identifier Code 1D32.

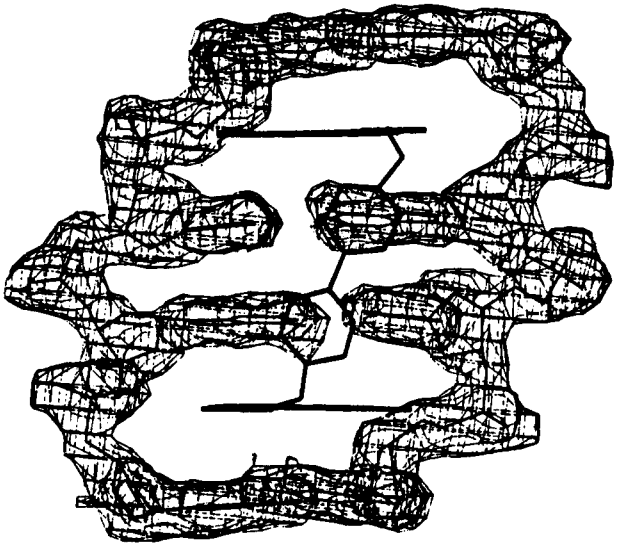

FIGURE 5: The ditercalinium–$[d(CGCG)]_2$ complex plus the electron density sum map surrounding the DNA.

Results

In the X-ray structure of ditercalinium bound to the duplex DNA fragment, $[d(CGCG)]_2$, the ditercalinium molecule bisintercalates at the two CpG steps of the DNA fragment. The linker is located in the major groove. The Fourier electron density ($2F_o - F_c$) maps are clean and continuous around the ditercalinium molecule (Gao et al., 1991) and around the DNA (Figure 5). Many of the gross characteristics of the X-ray structure (such as bisintercalation and the major groove location of the linker) were correctly predicted from a series of NMR experiments (Delbarre et al., 1987a,b, 1985; Delepierre et al., 1988, 1989). In the X-ray structure, the DNA is unwound (i.e., twist < 36°) at both intercalation steps and to a lesser extent at the central step (Table I). The base pairs are not distorted (Table II) beyond what is generally observed in B-DNA and less than what has been observed in certain monointercalated complexes (see, for example, Egli et al. (1991a)). The high degree of planarity of the base pairs in the ditercalinium complex is consistent with our recent observation (Williams et al., 1992) that when an intercalator binds with the long axis of the chromophore oriented nearly parallel to the long axes of the flanking base pairs (as in the ditercalinium complex), the base pairs flanking the intercalator remain relatively planar.

Certain torsion angles of the X-ray structure (Table III) were predicted with some success by a molecular mechanics modeling study (Maroun et al., 1989). In particular, the values of the torsion angles β, ϵ, and ζ in the X-ray structure correspond somewhat to those in the model. However the values

Table III: Deoxyribo-Phosphate Backbone and Glycosidic Torsion Angles[a] (in Degrees) and Deoxyribo-Sugar Puckers[b]

residue	α	β	γ	δ	ϵ	ζ	χ	P^c	pucker
C(1)	−71	−115		91	−133	−35	−166	331	C2′-exo
G(2)	−53	131	49	157	−166	−126	−71	194	C3′-exo
C(3)	−83	−156	70	95	−176	−78	−157	81	O1′-endo
G(4)			83	138			−90	156	C2′-endo
C(5)	−61	−148		81	−138	−61	−166	41	C4′-exo
G(6)	−47	144	53	142	−103	169	−79	157	C2′-endo
C(7)	−77	−167	5	101	172	−71	−126	63	C4′-exo
G(8)			86	140			−97	144	C1′-exo
B-DNA[d]	−63	171	54	123	−169	−108	−117		

[a] Backbone torsion angles for the bonds in the backbone P–O5′–C5′–C4′–C3′–O3′–P are α, β, γ, δ, ϵ, and ζ, respectively, and the glycosidic angle is χ. [b] Calculated with the program Curves (Lavery & Sklenar, 1989). [c] Pseudorotation phase angle. [d] From Saenger (1984).

of other torsion angles correspond poorly.

The ditercalinium molecule forms 73 van der Waals contacts (criterion is interatomic distance < 3.5 Å) with the DNA fragment (Table IV). The overwhelming majority of these contacts, 68 of them, are with guanines. In particular, the internal guanines [residues G(2) and G(6)] are well-stacked upon the chromophores of ditercalinium. Every atom of the base of G(2) and every atom, except the N2 and N3 atoms, of the base of G(6) are in van der Waals contact with the ditercalinium molecule. In contrast, the bases of the internal cytosines [residues C(3) and C(7)] are poorly stacked upon the chromophores of ditercalinium. Combined, the internal cytosines form only five van der Waals contacts with the ditercalinium molecule. One of the most surprising features of the complex is the complete unstacking of the bases of the terminal cytidines [residues C(1) and C(5)] from the chromophores of ditercalinium. The terminal cytosines do not form a single van der Waals contact with the ditercalinium molecule. This unstacking of the terminal cytosines is observable as a gap between a portion of each terminal base pair and the adjacent chromophore (Figure 3A,C).

Stacking interactions in DNA are commonly evaluated by viewing along the helical axis [axial slices of the ditercalinium complex are shown in (Gao et al., 1991)]. However, in the ditercalinium complex, stacking interactions are lost in part because the complex is spread out roughly *along the helical axis*. In these circumstances, axial views will prove to be less than illuminating.

One end of the linker, but not the other, forms extensive van der Waals interactions with the major groove of the DNA. This end (marked with primes in Figure 1 and in Table IV) forms six van der Waals contacts (two are hydrogen bonds, see below) with the major groove while the other end "floats" in the groove forming only a single van der Waals contact with the DNA. The interactions of the linker do not appear to contain a large hydrophobic contribution. Each of the contacts is to a hydrophilic position of the DNA, either O6 or N7. The linker does not form van der Waals contacts with the cytosines.

Only two of the van der Waals contacts between ditercalinium and the DNA are hydrogen bonds (criterion is interatomic distance < 3.5 Å with reasonable geometry and hydrogen bond donor/acceptor arrangements). The proton of one of the piperidinic nitrogens (N21′) of ditercalinium is directed toward the floor of the major groove. This nitrogen atom appears to form a bifurcated hydrogen bond, donating a proton simultaneously to both N7 and O6 of residue G(6). The distance is significantly less to the N7 than to the O6 (Table IV). The ditercalinium molecule does not form hydrogen bonds to cytosines.

The minor groove of the complex is unobstructed. The shape and positions of the chromophores of ditercalinium present van der Waals surfaces which are flush with the floor of the minor groove. The indole nitrogens (N7 and N7′ in Figure 1) of ditercalinium simulate N2 atoms of guanines. Viewing into the minor groove (Figure 3C), the N2 atoms of guanines and the N7 and N7′ atoms of ditercalinium appear to form a continuous string of hydrogen bond donors lined up roughly parallel to the helical axis.

Although both the ditercalinium molecule and the DNA fragment contain the potential to adopt 2-fold symmetry, in the X-ray structure the complex is asymmetric. The conformation and DNA interactions of one half of the ditercalinium molecule are different from those of the other half. The lack of 2-fold symmetry is discernible in Table IV, where it is shown, for example, that there is a single van der Waals contact between C1 of ditercalinium and residue G(2) of the DNA. In contrast, there are four van der Waals contacts between C1′ of ditercalinium (the pseudosymmetry mate of

Table IV: van der Waals Contacts[a] between Ditercalinium and DNA

ditercalinium atom	DNA residue	DNA atom	distance (Å)
C1	G(2)	O6	3.37
C1′	G(6)	C6	3.23
		O6	3.28
		C5	3.31
N2′	G(6)	C5	3.44
		N7	3.15
C3	G(2)	N7	3.30
C3′	G(6)	N7	3.21
		C8	3.26
C4	G(2)	N7	3.24
		C8	3.20
		O5′	3.00
C4′	G(6)	O1′	3.44
		N9	3.41
		C8	3.43
C5	G(2)	O1′	3.10
		C5′	3.34
		N9	3.33
		C4	3.32
		O5′	3.32
C5′	G(6)	O1′	3.49
C6	G(2)	O1′	3.46
		N3	3.35
		C4	3.48
	G(8)	N2	3.44
C6′	G(4)	N2	3.31
		C2	3.49
N7	G(2)	N2	3.43
	G(8)	N3	3.44
		C2	3.23
		N2	3.25
N7′	G(4)	C4	3.39
		N3	3.34
		C2	3.45
C8	G(8)	C4	3.32
C8′	C(3)	C2	3.36
		O2	3.44
	G(4)	C5	3.42
		N7	3.37
C9	G(8)	C5	3.48
C10′	C(3)	N3	3.49
		C4	3.34
O10	C(7)	C5	3.42
C12	G(2)	C4	3.43
		C5	3.29
		N7	3.38
		C8	3.48
C12′	G(6)	C4	3.37
C13	G(2)	N3	3.44
		C2	3.35
	G(8)	N2	3.28
C13′	G(4)	C2	3.37
	G(6)	C2	3.44
C14	G(8)	N3	3.47
		C2	3.35
C14′	G(4)	C5	3.38
C15	G(2)	N1	3.38
	G(8)	N1	3.41
C15′	G(6)	N1	3.38
C16	G(2)	N1	3.34
		C6	3.34
C16′	G(6)	N1	3.24
		C6	3.44
C17	G(2)	C6	3.32
		C5	3.26
C17′	G(6)	C6	3.26
		C5	3.36
C20′	G(6)	N7	3.17
N21′	G(6)	O6	3.44[b]
		N7	2.95[b]
C22	G(2)	O6	3.32
C22′	G(6)	O6	3.31
C23′	G(6)	O6	2.93

[a] The criterion for a van der Waals contact is a nonbonded interatomic distance less than 3.5 Å. [b] Hydrogen bonds.

Table V: van der Waals Contacts between Water Molecules and Oxygens/Nitrogens of the Complex

water molecule	B (Å^2)	residue	atom	distance (Å)
13W10	17.1	G(2)	N7	2.64
13W17	57.7	C(3)	O1′	3.41
			O3′	3.27
13W18	43.8	G(4)	O1′	2.84
			O3′	3.24
13W29	26.6	C(3)	O1P	2.50
			O2P	2.59
14W08	31.1	G(2)	O3′	3.00
			O2P	2.48
14W12	19.5	G(2)	O2P	2.42
14W13	28.4	G(8)	N2	2.86
			N3	2.47
		D	N7′	3.50
15W07	45.9	C(7)	O1P	3.42
15W13	52.4	G(2)	O1P	3.04
16W03	36.0	G(2)	O1′	2.91
			O3′	3.40
16W04	45.9	G(2)	O1′	2.52
			N3	2.69
16W14	41.5	G(8)	O3′	2.79
16W15	30.2	D	N7	3.44
16W16	49.3	C(1)	O2	2.50

C1) and residue G(6) of the DNA [the pseudosymmetry mate of residue G(2)]. In a second example, one of the piperidinic nitrogens forms hydrogen bonds to the DNA while the other does not. The molecular basis of the asymmetry appears to arise at least in part from stereochemical constraints of the bipiperidine ring system. When one piperidinic nitrogen (N21′) forms a hydrogen bond to the floor of the major groove, a combination of constraints (chair plus rotamer) direct the proton of the second piperidinic nitrogen (N21) away from the floor of the major groove.

The ditercalinium–[d(CGCG)]$_2$ complex, with 84 ordered water molecules observable by X-ray diffraction, is well hydrated. Table V lists the hydrogen bonding contacts between the nitrogen/oxygen atoms of the complex and the first shell of water molecules. Intramolecular water-mediated hydrogen bonds are numerous (Table V). In three cases a water molecule bridges the O1′ and O3′ atoms of a single deoxyribose. The complex contains the potential to adopt a "spine of hydration" within the minor groove (employing the indole nitrogens of ditercalinium). However, no compelling spine of hydration is observed. There is a single water-mediated hydrogen bond between the DNA and the ditercalinium molecule.

The view down the *c*-axis of the unit cell, which is nearly parallel to the helical axis (Figure 6A), shows that there are no direct lateral contacts between complexes in the crystal. However, two views perpendicular to the *c*-axis (panels B and C of Figure 6) show that the complexes stack to form infinite helices. In previous crystals of DNA fragments, both possible end-to-end stacking orientations have been observed. Helices can stack with 3′,5′ and 5′,3′ overlap, (which approximates the stacking arrangement within DNA) or with 3′,3′ and 5′,5′ overlap (Williams et al., 1990b). The ditercalinium complexes stack with 3′,5′ and 5′,3′ overlap as observed internally in DNA.

Discussion

Ditercalinium is a combination bisintercalator and major groove binder. As shown schematically in Figure 7, the positively-charged linker of the drug is located within, and binds to, the major groove of the DNA. As we described previously (Gao et al., 1991), the DNA in the complex retains a right-handed, although underwound, double-helical con-

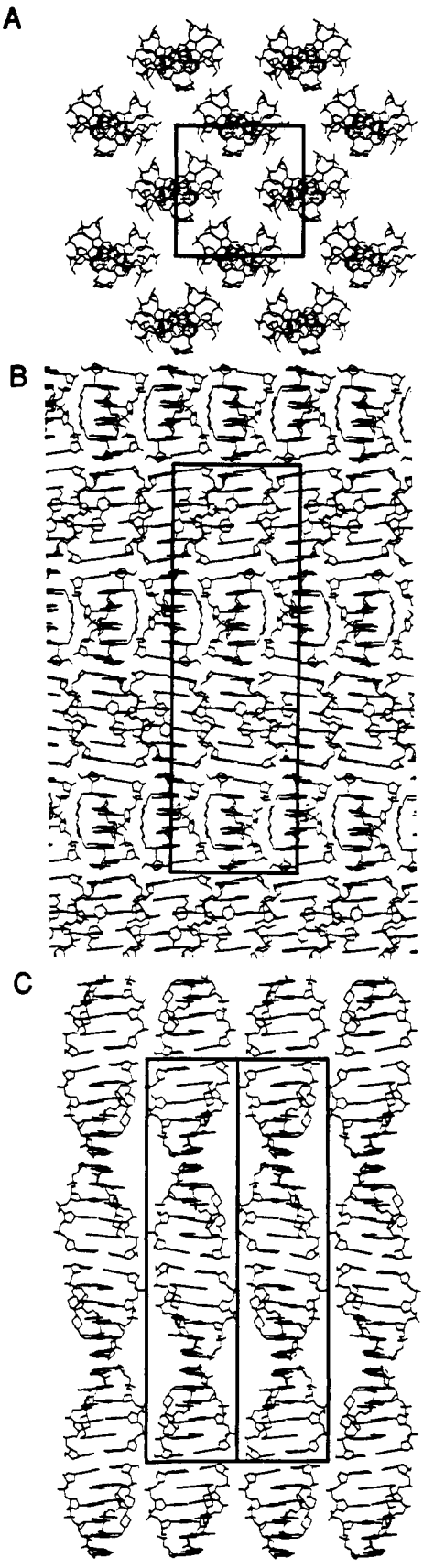

FIGURE 6: Packing arrangement in the crystal lattice of the ditercalinium–[d(CGCG)]$_2$ complex. The view (A) down the *z*-axis, (B) down the *x*-axis, and (C) down the line bisecting the *x*- and *y*-axes. Each drawing shows a layer of approximately 20 Å thickness. The unit cell is outlined by heavy lines.

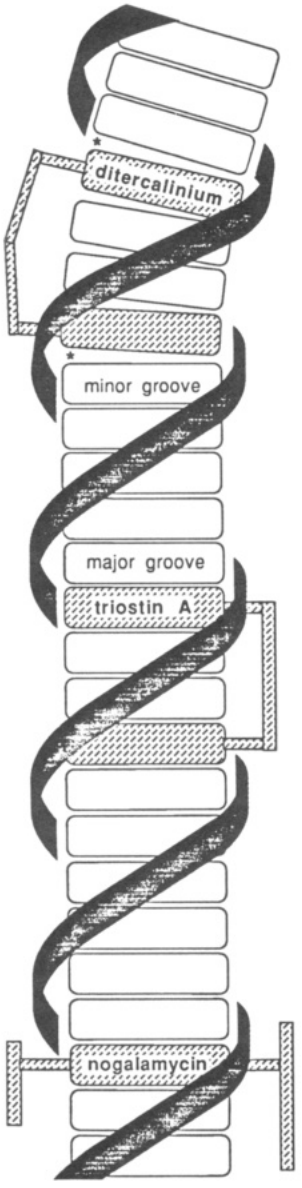

FIGURE 7: Schematic diagram of intercalated DNA complexes with ditercalinium (top), triostin A (center), and nogalamycin (bottom). The drug molecules are shaded. The asterisks indicate sites of unstacking.

formation. The DNA is kinked, with an abrupt bend of around 15° in the helical axis. The direction of this static bend is toward the minor groove.

From the minor groove, the ditercalinium complex mimics normal DNA. Within the minor groove, both the shape of the van der Waals surface and the arrangement of hydrogen bonding functionalities of the ditercalinium complex simulate those of DNA base pairs. Although the complex is well-hydrated, there is no compelling "spine of hydration" within the minor groove.

One striking feature of the complex is the selective interaction of ditercalinium with guanines but not cytosines. Both the intercalative and groove-binding segments of ditercalinium interact extensively with the guanines of the DNA fragment and do not interact to a significant extent with the cytosines. The terminal cytosines [C(1) and C(5)] are completely excluded from van der Waals contacts with the chromophores of ditercalinium. Thus, the terminal cytosines are "unstacked" from the rest of the complex (Figure 7).

A comparison of the three-dimensional structure of the ditercalinium complex with complexes of triostin A and nogalamycin allows us to propose a general model for biological recognition of one class of damaged DNA.

Triostin A. Triostin A (Figures 1B and 2C,D) and echinomycin, which differs from triostin A by a minor modification of the sulfur cross-bridge, are naturally occurring bisintercalators. Triostin A is formed by a cyclic depsipeptide linking two quinoxaline rings. As can be seen in space-filling representations of triostin A (Figure 2C,D), this drug has certain structural features in common with ditercalinium. Both triostin A and ditercalinium are composed of two intercalative moieties joined by a rigid linker. The three-dimensional shapes, each resembling a claw, are similar. However, unlike ditercalinium, the mechanism of action of triostin A appears to involve interaction of DNA–triostin A complexes with DNA replication and transcription systems (Gale et al., 1981). Triostin A is active against Gram-positive bacteria and is cytotoxic to cultured tumor cells.

Three-dimensional structures of a series of complexes formed by triostin A bound to double-stranded DNA fragments have been solved by X-ray crystallography (Quigley et al., 1986; Ughetto et al., 1985; Wang et al., 1984). The focus here will be the complex of two triostin A molecules bound to $[d(GCGTACGC)]_2$, solved to 2.0-Å resolution (Quigley et al., 1986). Space-filling representations of this complex are shown in Figure 4. As in the case of the ditercalinium complex, the asymmetric unit of the triostin A complex is composed of one bisintercalated drug molecule and four base pairs of DNA [i.e., one triostin A molecule and the DNA fragment d(GCGT)·d(ACGC)]. A schematic diagram of triostin A bound to DNA is shown in Figure 7. The linker of triostin A is located in the minor groove of the DNA rather than in the major groove as in the ditercalinium complex. The unstacking of bases in the ditercalinium complex (Figure 3A,C) is not observed in the triostin A complex (Figure 4A,C). Each base pair is well-stacked either upon another base pair or upon one of the intercalated chromophores of triostin A.

Nogalamycin. Nogalamycin (Figure 1C) is a naturally occurring monointercalator with an unusual three-dimensional shape. This drug, shaped somewhat like a dumbbell, is composed of a planar chromophore substituted on both ends. On one end, nogalamycin contains both a methyl ester and a nogalose sugar and on the other end a fused bicyclo aminosugar with a positive charge. A space-filling representation of nogalamycin and a DNA complex of nogalamycin can be found in Egli et al. (1991a). Ditercalinium and nogalamycin have certain structural features in common. Both are intercalators and both bind to DNA such that a bulky, positively-charged segment of the drug forms hydrogen bonds to the floor of the major groove (Williams et al., 1990a). Unlike ditercalinium, but like triostin A, the mechanism of action of nogalamycin appears to involve interaction of DNA–nogalamycin complexes with DNA replication and transcription systems. Nogalamycin selectively inhibits RNA synthesis in vivo (Fok & Waring, 1972; Li et al., 1979) and is active against Gram-positive bacteria and experimental tumors (Bhuyan, 1967; Bhuyan & Reusser, 1970; Li et al., 1979).

Three-dimensional structures of a series of complexes with two nogalamycin molecules bound to a double-stranded DNA hexamer have been determined (Egli et al., 1991a; Liaw et al., 1989; Williams et al., 1990a). A schematic diagram of nogalamycin bound to DNA is shown in Figure 7. The unstacking of bases in the ditercalinium complex is not observed in the nogalamycin complexes. Each base pair is well-stacked either upon another base pair or upon a chromophore of nogalamycin. The unstacking of bases observed in the diter-

calinium complex is not a feature of any other DNA–drug complex solved by X-ray crystallography thus far.

DNA Repair. In *Escherichia coli*, ditercalinium provokes malfunction of the (A)BC exinuclease excision repair system. This DNA repair system recognizes a wide variety of damage to DNA including pyrimidine dimers, 6–4 photoadducts of pyrimidines, and chemical adducts of psoralen, *cis*- and *trans*-platinum, mitomycin C, benzo[*a*]pyrene, and the products of many alkylating agents possessing alkyl groups larger than methyls (Grossman et al., 1988; Sancar & Sancar, 1988). The capacity of one enzymatic system to recognize such a diverse array of covalent damage to DNA clearly requires detection of an attribute that (a) distinguishes damaged from undamaged DNA and (b) is independent of the intrinsic structure of the lesion. It has been hypothesized (Grossman et al., 1988; Sancar & Sancar, 1988) that a common, distorted conformation of DNA is induced by this diverse family of DNA-damaging agents. As described below, the structure of the ditercalinium complex leads us to suggest an alternative mechanism.

The initial recognition of a lesion by the (A)BC exinuclease excision repair system is provided by the high affinity of the protein uvrA for damaged DNA. However, it appears that essentially all compounds with a reasonable DNA affinity, including both intercalators and groove binders, enhance binding of uvrA to undamaged DNA (Selby & Sancar, 1991). The second, and more subtle, stage in the recognition of a DNA lesion is conferred by the protein uvrB, which appears to have no affinity for DNA but is installed (i.e., kinetically trapped) at a lesion site by uvrA (Orren & Sancar, 1990). uvrA then dissociates, after which DNA incisions are incurred by a third protein, uvrC (Orren & Sancar, 1989).

Ditercalinium appears to be unique in its ability to promote the enzyme-mediated incision of noncovalently damaged DNA (Lambert et al., 1989, 1988, 1990). It is likely that the unusual activity of ditercalinium is manifested at the second stage in recognition; i.e., ditercalinium appears to facilitate kinetic trapping of uvrB. It is important to understand which features of DNA–ditercalinium complexes facilitate this process. Molecular features of the DNA–ditercalinium complex that merit consideration as possible factors in kinetic trapping of uvrB are the following: (a) unstacking of the terminal cytosines, (b) accessibility of the minor groove, (c) mimicry of the minor groove of unperturbed DNA, (d) the bend in the helical axis, (e) bisintercalation, and/or (f) placement and interactions of the linker. Each of these possible contributing factors is discussed below and is shown schematically in Figure 7.

(*a*) *Unstacking of the Terminal Cytosines.* Ditercalinium stabilizes DNA (Delbarre et al., 1987a), and the complete loss of stacking between each chromophore of ditercalinium and the terminal cytosine was an unexpected feature of the three-dimensional structure of the complex. It appears that even though the net effect of ditercalinium is to stabilize duplex DNA, ditercalinium is likely to cause localized instability. Local disruption of stacking distinguishes the ditercalinium complex from those with triostin A and nogalamycin described above and from complexes of other intercalators and groove binders solved thus far by X-ray diffraction. In previous DNA–drug complexes, regardless of severe buckling or Hoogsteen pairing of the base pairs, each base is well-stacked within the helix.

(*b*) *Accessibility of the Minor Groove.* A second exceptional feature of ditercalinium is the lack of interaction of this drug with the minor groove of DNA. Ditercalinium binds with the linker in the major groove; the drug does not protrude into the minor groove. The minor groove of the DNA–ditercalinium complex is completely accessible to other cellular components. In contrast, bulky substituents bind in the minor groove of DNA complexes with triostin A and nogalamycin and in complexes with ethidium, daunomycin/adriamycin, distamycin, and Hoechst 33258, and essentially all other organic compounds with DNA affinity. Ditercalinium appears to be a rare exception to this pattern.

(*c*) *Mimicry of the Minor Groove of Unperturbed DNA.* A related feature of the ditercalinium complex is that, from the minor groove, the complex mimics normal DNA. In both shape and arrangement of hydrogen bonding functionalities, the minor groove of the complex appears strikingly similar to that of unperturbed DNA. This mimicry of the minor groove of normal DNA distinguishes the ditercalinium complex from other DNA–drug complexes solved by X-ray crystallography thus far.

(*d*) *The Bend in the Helical Axis.* The DNA in the complex with ditercalinium is bent by 15°. However, DNA in vivo is severely compacted and the normal biological state of DNA is anything but linear. The static bend induced by ditercalinium is rather subtle in comparison with bends observed in DNA–protein complexes solved by X-ray crystallography (Brennan et al., 1990; Schultz et al., 1991). Proper function requires that repair systems rigorously avoid mistaking a DNA–protein complex for damaged DNA.

(*e*) *Bisintercalation.* Bisintercalation in DNA is a common mode of interaction, and there is no evidence that other bisintercalators such as triostin A provoke malfunction of the (A)BC exinuclease excision repair system.

(*f*) *Placement and Interactions of the Linker.* Nogalamycin binds to DNA in a manner analogous to that of the binding of ditercalinium in that a bulky, positively-charged segment of nogalamycin is located within the major groove. There is no evidence that DNA–nogalamycin complexes provoke malfunction of the (A)BC exinuclease excision repair system. It may be significant that ditercalinium, *cis*-platinum, and 6–4 photoadducts of pyrimidines each provides the equivalent of a structural tether within the major groove of the DNA.

Model for Recognition by the (A)BC Exinuclease Excision Repair System. As noted above, the capacity of one enzymatic system to recognize a diverse array of covalent damage to DNA clearly requires detection of an attribute that distinguishes damaged from undamaged DNA yet is independent of the intrinsic structure of the lesion.

It is commonly hypothesized (Grossman et al., 1988; Sancar & Sancar, 1988) that the same distorted conformation of DNA results from modification by each of a diverse family of DNA-damaging agents. We propose an alternative hypothesis. We propose that each of these damaging agents causes local instability of DNA (but not necessarily a common conformation) and that the (A)BC exinuclease excision repair system recognizes excessive or unusual deformability of DNA. Specifically, installing uvrB onto DNA may require distortion of DNA to a state that is more easily achieved in the vicinity of certain types of lesions. This proposal is consistent with recent indications that sequence-specific deformability of DNA can be important in DNA recognition by 434 repressor, Cro, and CAP and in other protein–DNA interactions (Drew & Travers, 1985; Gartenberg & Crothers, 1988; Koudelka et al., 1988, 1987; Satchwell et al., 1986; Schultz et al., 1991).

It is reasonable to suggest that localized destabilizatin such as that inferred from the three-dimensional structure of the ditercalinium complex could increase the deformability of the DNA. Although there is currently a lack of X-ray structures

of other substrates of the (A)BC exinuclease excision repair system, from inspection it appears that each would destabilize stacking interactions. For example, thymidine dimers lack the planarity of a normal base, and *cis*-platinum adducts appear to induce a severe kink in the helical axis. Each of these would be likely to destabilize stacking interactions. Thus, we propose that local destabilization of the DNA could be an attribute of damaged DNA that is required for kinetic trapping of uvrB. In addition, the mimicry of the minor groove of normal DNA by the ditercalinium complex is consistent with uvrA installation of uvrB onto the minor groove. The static bend in the helical axis, *bis*intercalation, and the placement of the linker of ditercalinium within the major groove are unlikely to be factors in the kinetic trapping of uvrB.

Influence of Crystal Packing. The influence of crystal packing on molecular conformation appears to be generally more pronounced with DNA than with proteins. For example, crystal lattice-induced bending of DNA has been observed (DiGabriele et al., 1989). Although the DNA conformation of drug complexes may be less influenced than DNA alone due to the tendency of drugs to stabilize and stiffen DNA, lattice contacts may still select for a conformation that is not predominant in solution. In many cases, the effects of lattice on a crystal structure can be determined by crystallizing a given fragment or complex of DNA in more than one crystal form. The lattice contacts and packing forces vary from one crystal form to another. If the DNA conformation in different crystallographic environments is unchanged, then the observed conformation is likely to be near an energy minimum in solution. Further, differences in conformation may reflect conformational heterogeneity in solution. By obtaining multiple crystal forms, we have recently analyzed effects of crystal packing on conformation and hydration of DNA–anthracycline complexes (Williams et al., 1990b) and on Z-DNA (Egli et al., 1991b). It is clear that a complete understanding of ditercalinium requires crystallization in additional crystal forms and with longer fragments of DNA.

ACKNOWLEDGMENTS

We thank Drs. Ezra Abrams and Nassim Usman and Ms. Nidhi Gupta Williams for helpful discussions. Ditercalinium was kindly provided by Professor Bernard P. Roques (Paris). The work described here was performed in the laboratory of Professor Alexander Rich.

REFERENCES

Acramone, F., & Penco, S. (1989) in *Antitumor Natural Products* (Takeuchi, T., Nitta, K. & Tanaka, N., Eds.) pp 1–53, Japan Scientific Press, Tokyo.

Bhuyan, B. K. (1967) in *Antibiotics* (Gottlieb, D., & Shaw, P. D., Eds.) Vol. I, pp 173–180, Springer, New York.

Bhuyan, B. K., & Reusser, F. (1970) *Cancer Res. 30*, 984–989.

Brennan, R. G., Roderick, S. L., Takeda, Y., & Matthews, B. W. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8165–8169.

Brunger, A. (1988) *X-PLOR Software*, The Howard Hughes Medical Institute, Yale University, New Haven, CT.

Delbarre, A., Delepierre, M., Igolen, J., Le Pecq, J. B., & Roques, B. P. (1985) *Biochimie 67*, 823–828.

Delbarre, A., Delepierre, M., Garbay, C., Igolen, J., Le Pecq, J. B., & Roques, B. P. (1987a) *Proc. Natl. Acad. Sci. U.S.A. 84*, 2155–2159.

Delbarre, A., Delepierre, M., Langlois D'Estaintot, B., Igolen, J., & Roques, B. P. (1987b) *Biopolymers 26*, 1001–1033.

Delepierre, M., Igolen, J., & Roques, B. P. (1988) *Biopolymers 27*, 957–968.

Delepierre, M., Maroun, R., Garbay-Jaureguiberry, C., Igolen, J., & Roques, B. P. (1989) *J. Mol. Biol. 210*, 211–228.

DiGabriele, A. D., Sanderson, M. R., & Steitz, T. A. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 1816–1820.

Drew, H. R., & Travers, A. A. (1985) *J. Mol. Biol. 186*, 773–790.

Egli, M., Williams, L. D., Frederick, C. A., & Rich, A. (1991a) *Biochemistry 30*, 1364–1372.

Egli, M., Williams, L. D., Gao, Q., & Rich, A. (1991b) *Biochemistry 30*, 11388–11402.

Fok, J., & Waring, M. J. (1972) *Mol. Pharmacol. 8*, 65–74.

Frederick, C. A., Williams, L. D., Ughetto, G., van der Marel, G. A., van Boom, J. H., Rich, A., & Wang, A. H.-J. (1990) *Biochemistry 29*, 2538–2549.

Gale, E. F., Cundliffe, E., Reynolds, P. E., Richmond, M. H., & Waring, M. J. (1981) in *The Molecular Basis of Antibiotic Action* (Neidle, S., & Waring, M. J., Eds.) pp 285–401, Wiley, London.

Gao, Q., Williams, L. D., Egli, M., Rabinovich, D., Chen, S.-H., Quigley, G. J., & Rich, A. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 2422–2426.

Gartenberg, M. R., & Crothers, D. (1988) *Nature (London) 333*, 824–829.

Grossman, L., Caron, P. R., Mazur, S. J., & Oh, E. Y. (1988) *FASEB J. 88*, 2696–2701.

Hendrickson, W. A., & Konnert, J. H. (1981) in *Biomolecular Structure, Conformation, Function and Evolution* (Srinivasan, R., Ed.) Vol. I, pp 43–57, Pergamon Press, Oxford.

Jones, T. A. (1982) in *Computational Crystallography* (Sayre, D., Ed.) Vol. 11, pp 303–317, Clarendon Press, Oxford.

Koudelka, G. B., Harrison, S. C., & Ptashne, M. (1987) *Nature (London) 326*, 886–888.

Koudelka, G. B., Harbury, P., Harrison, S. C., & Ptashne, M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4633–4637.

Lambert, B., Jones, B. K., Roques, B. P., Le Pecq, J. B., & Yeung, A. T. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 6557–6561.

Lambert, B., Roques, B. P., & Le Pecq, J. B. (1988) *Nucleic Acids Res. 3*, 1063–1078.

Lambert, B., Segal-Bendirdjian, E., Esnault, C., Le Pecq, J. B., Roques, B. P., Jones, B., & Yeung, A. T. (1990) *Anti-Cancer Drug Des. 5*, 43–53.

Laugaa, P., Markovits, J., Delbarre, A., Le Pecq, J. B., & Roques, B. P. (1985) *Biochemistry 24*, 5567–5575.

Lavery, R., & Sklenar, H. (1989) *J. Biomol. Struct. Dyn. 6*, 655–667.

Li, L. H., Kuentzel, S. L., Murch, L. I., Pschigoda, L. M., & Krueger, W. C. (1979) *Cancer Res. 39*, 4816–4822.

Liaw, Y.-C., Gao, Y.-G., Robinson, H., van der Marel, G. A., van Boom, J. H., & Wang, A. H.-J. (1989) *Biochemistry 28*, 9913–9918.

Maroun, R., Delepierre, M., & Roques, B. P. (1989) *J. Biomol. Struct. Dyn. 7*, 607–621.

Orren, D. K., & Sancar, A. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5237–5241.

Orren, D. K., & Sancar, A. (1990) *J. Biol. Chem. 265*, 15796–15803.

Pelaprat, D., Delbarre, A., Le Guen, I., Roques, B. P., & Le Pecq, J. B. (1980a) *J. Med. Chem. 23*, 1336–1343.

Pelaprat, D., Oberlin, R., Le Guen, I., Roques, B. P., & Le Pecq, J. B. (1980b) *J. Med. Chem. 23*, 1330–1335.

Quigley, G. J., Ughetto, G., van der Marel, G. A., van Boom, J. H., Wang, A. H.-J., & Rich, A. (1986) *Science 232*, 1255–1258.

Rabinovich, D., & Shakked, Z. (1984) *Acta Crystallogr. A40*, 195-200.
Saenger, W. (1984) in *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.
Sancar, A., & Sancar, G. B. (1988) *Annu. Rev. Biochem. 57*, 29-67.
Satchwell, S. C., Drew, H. R., & Travers, A. A. (1986) *J. Mol. Biol. 191*, 659-675.
Schultz, S. C., Shields, G. C., & Steitz, T. A. (1991) *Science 253*, 1001-1007.
Selby, C. P., & Sancar, A. (1991) *Biochemistry 30*, 3841-3849.
Shakked, Z. (1983) *Acta Crystallogr. A39*, 278-279.
Sheriff, S., & Hendrickson, W. A. (1987) *Acta. Crystallogr. A43*, 118-121.
Ughetto, G., Wang, A. H.-J., Quigley, G. J., van der Marel, G. A., van Boom, J. H., & Rich, A. (1985) *Nucleic Acids Res. 13*, 2305-2323.
Wang, A. H.-J., Ughetto, G., Quigley, G. J., Hakoshima, T., van der Marel, G. A., van Boom, J. H., & Rich, A. (1984) *Science 225*, 1115-1121.
Williams, L. D., Egli, M., Gao, Q., Bash, P., van der Marel, G. A., van Boom, J. H., Rich, A., & Frederick, C. A. (1990a) *Proc. Natl. Acad. Sci. U.S.A. 87*, 2225-2229.
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. (1990b) *J. Mol. Biol. 215*, 313-320.
Williams, L. D., Egli, M., Gao, Q., & Rich, A. (1992) in *Structure & Function: Proceedings of the Seventh Conversation in Biomolecular Stereodynamics* (Sarma, R. H., & Sarma, M. H., Eds.) Adenine Press, Albany, NY (in press).

Structural and Energetic Consequences of Disruptive Mutations in a Protein Core†

Wendell A. Lim,‡ Dawn C. Farruggio, and Robert T. Sauer*
Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
Received September 13, 1991; Revised Manuscript Received January 14, 1992

ABSTRACT: We have characterized the properties of a set of variants of the N-terminal domain of λ repressor bearing disruptive mutations in the hydrophobic core. These mutations include some that dramatically alter the total core residue volume (by up to six methylene groups) and some that place a single polar residue into the otherwise hydrophobic core. The structural properties of the purified proteins have been studied by CD spectroscopy, biological activity, recognition by conformation-specific monoclonal antibodies, and ^{1}H NMR spectroscopy. The stabilities of the proteins have been measured by thermal and guanidine hydrochloride denaturation. Proteins with disruptive core mutations are found to display a continuum of increasingly nonnative properties. Large internal volume changes cause both significant conformational rearrangements and destabilization by up to 5 kcal/mol. Variants with polar substitutions at core positions no longer behave like well-folded proteins but rather display characteristics of molten globules. However, even proteins bearing some of the most disruptive mutations retain many of the crude secondary and tertiary structural features of the wild-type protein. These results indicate that primitive elements of native structure can form in the absence of normal core packing.

The interiors of proteins are packed very tightly. Such packing has long been thought to play an essential role in stabilizing a protein's native conformation (Richards, 1974, 1977, Chothia, 1975b). Tight packing maximizes van der Waals interactions and, by efficiently excluding solvent, may also contribute to hydrophobic stabilization. In recent studies, site-directed mutagenesis has been used to replace wild-type core residues with other apolar residues that have different steric properties or slightly smaller volumes (Garvey & Matthews, 1989; Karpusas et al., 1989; Kellis et al., 1988, 1989; Lim & Sauer, 1991; Matsumura et al., 1988, 1989; Sandberg & Terwilliger, 1989, 1991a; Shortle et al., 1990). These studies yield two general conclusions. First, packing interactions make significant contributions to the free energy of folding. Mutants with sterically altered residues are often destabilized by more than would be expected solely due to changes in hydrophobic interactions as modeled by transfer free energies. Second, although destabilizing, these types of core mutations alter the overall protein structure relatively little. Mutant proteins that can stably fold usually adopt a structure that is sufficiently native-like to maintain similar spectroscopic properties and at least partial biological activity. Structural analyses show that subtle main-chain rearrangements can allow new side-chain combinations to maintain reasonable packing densities (Eigenbrot et al., 1990; Katz & Kossiakoff, 1990; Daopin et al., 1991). Thus, while modest changes in packing can destabilize a protein, the overall structure seems quite tolerant to such perturbations.

Given the apparent plasticity of protein cores, we were interested in probing the limits to which a protein can tolerate disruptions in its core. How does a protein respond to mutations that severely alter core volume? For example, what happens if much larger side chains are inserted into the tightly packed interior of a protein? Does the structure rearrange and expand to accommodate the greater volume or is the protein simply unable to fold? How does a protein respond to the insertion of polar residues into the core? To investigate these questions, we have constructed and purified a set of potentially disruptive core mutants of the N-terminal domain of λ repressor (Figure 1). These mutant proteins can be

† This work was supported by NIH Grant AI-15706. W.A.L. was a Howard Hughes predoctoral fellow.
* Author to whom correspondence should be addressed.
‡ Current address: Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06511.

0006-2960/92/0431-4324$03.00/0