The Molecular Structure of a DNA–Triostin A Complex

Andrew H.-J. Wang, Giovanni Ughetto, Gary J. Quigley
Toshio Hakoshima, Gijs A. van der Marel
Jacques H. van Boom, Alexander Rich

Abstract. *The molecular structure of triostin A, a cyclic octadepsipeptide antibiotic, has been solved complexed to a DNA double helical fragment with the sequence CGTACG (C, cytosine; G, guanine; T, thymine; A, adenine). The two planar quinoxaline rings of triostin A bis intercalate on the minor groove of the DNA double helix surrounding the CG base pairs at either end. The alanine residues form hydrogen bonds to the guanines. Base stacking in the DNA is perturbed, and the major binding interaction involves a large number of van der Waals contacts between the peptides and the nucleic acid. The adenine residues in the center are in the* syn *conformation and are paired to thymine through Hoogsteen base pairing.*

A number of biological substances act by binding to DNA. Some of these intercalate planar residues between the base pairs of DNA, and others bind to the outside of the DNA double helix. An interesting class of molecules are those that contain two planar rings with the potential for double intercalation (*1–3*). These molecules are generally antibiotics, and the planar rings are attached to cyclic peptides or depsipeptides if they contain an ester linkage. The quinoxaline antibiotics derived from *Streptomyces* have eight amino acids that form a cyclic depsipeptide (*1–3*). They are highly active against Gram-positive bacteria and also have cytotoxic effects on cultured cells including tumor cells. They bind to DNA and inhibit both DNA replication and DNA-directed RNA synthesis (*4, 5*). There has been considerable interest in the manner in which they carry out these biological activities, but up to now no information has been available about the detailed interaction between these molecules and DNA. We now report on the three-dimensional structure of a complex between the quinoxaline antibiotic triostin A (Fig. 1) and a DNA duplex with the self-complementary sequence d(CpGpTpApCpG) (C, cytosine; G, guanine; T, thymine; A, adenine). The structure has been solved and refined to a resolution of 1.67 Å and reveals a detailed pattern of interactions between the minor groove side of DNA and the antibiotic. These include intercalation by quinoxaline rings, hydrogen bonding between the DNA bases and the peptide backbone, as well as a large number of van der Waals interactions between the amino acids and the nucleic acid. The structure reveals why these particular amino acids, including their modifications, are found in the antibiotic. The DNA is considerably unwound. The antibiotic intercalates on both sides of the dCpG residues and binds in a sequence-specific manner. In addition, a novel rearrangement of the AT base pairs is found outside the intercalative site. These bases are held together by Hoogsteen base pairs involving the N-7 nitrogen atom of adenine rather than Watson-Crick base pairs.

Crystallization and Structure Solution

The oligonucleotides were synthesized by an improved triester method as described (*6*). Co-crystals were made with the use of both triostin A and echinomycin (Fig. 1) (*7*). Neither is water soluble; hence they were initially dissolved in organic solvents [triostin A in a mixture of methanol and chloroform (1:1) and echinomycin in methanol] and then mixed with an aqueous solution of oligonucleotides and ions. In addition to the hexamer a self-complementary octamer, d(GpCpGpTpApCpGpC), was also used, and a variety of crystallization conditions was explored. Crystals were formed by vapor-phase equilibration with a 35 percent 2-methyl-2,4-pentanediol (MPD) solution from a solution containing 2 m*M* DNA hexamer, 20 m*M* sodium cacodylate (*p*H 7), 10 m*M* $MgCl_2$, 1.5 m*M* triostin A, and 5 percent MPD. The chloroform droplets slowly evaporated, and thin rhombic plates began to appear within a week. The plates were allowed to grow until they were almost 1 mm on each edge with a thickness of nearly 0.15 mm. Crystals were mounted in sealed glass capillary tubes with a droplet of mother liquor. The crystals had an orthorhombic lattice with space group F222, a = 31.35(1) Å, b = 62.38(2) Å, and c = 61.26(3) Å. The a and b axes were oriented along the diagonals of the rhombic plate. After the structure was solved, the asymmetric unit was found to contain one strand of DNA plus one molecule of triostin A. Crystal complexes were also obtained with echinomycin in a similar lattice that produced a virtually indistinguishable diffraction pattern, showing that its structure is very similar. Crystals were also obtained with the DNA octamer complexed to triostin A, and initial studies indicate that its structure is generally similar to the triostin A–hexamer complex. Different lattices were obtained with both barium and calcium salts, but the magnesium complex was studied in detail because the crystals were better.

Three-dimensional x-ray diffraction data were collected on a Nicolet P3 with the omega scan mode at −16°C, and the entire data set was taken from one crystal, totaling 2995 reflections with intensity considered to be greater than 1.0 σ(I). A translation-rotation search was carried out in an attempt to solve the structure directly, but all models failed despite considerable effort. Accordingly, several heavy-atom derivatives of the oligonucleotide were synthesized. The d(CpGpTpApCpG) nucleotides were numbered 1 through 6 respectively, and 5-bromouracil was used in place of thymine 3 (T3), 5-bromocytosine was used for both cytosine 1 (C1) and cytosine 5 (C5) residues in place of cytosine, and 5-bromocytosine was used in place of C1 alone and finally in place of C5 alone. These provided four potential heavy-atom derivatives; however, the molecule with 5-bromocytosine substituted for C1 did not crystallize. All the others formed isomorphic crystal complexes, and complete three-dimensional

Drs. Wang, Ughetto, Quigley, Hakoshima, and Rich are in the Department of Biology, Massachusetts Institute of Technology, Cambridge 02139; Drs. van der Marel and van Boom are at Gorlaeus Laboratories, Leiden State University, Leiden, The Netherlands.

data were collected. In the same crystallization dip of the bromo derivatives, pseudo-hexagonal crystals appeared together with the rhombic plates. These crystals were examined and found to be the C1,C5-dibromo- and the C5-bromo-oligonucleotides that had formed crystals of Z DNA. The structure of the 1,5-dibromo derivative has been described (*8*).

Three-dimensional difference Patterson maps were used to locate the bromine atoms. The origin was fixed by using cross-Patterson functions (*9*). The multiple isomorphous replacement (MIR) phases were derived from these three derivatives, and the best MIR map was calculated at 2 Å and used to trace the chain. The coordinates of the complex were measured from a Kendrew model built in a Richards optical comparator (*9*). The molecule was then refined to a resolution of 1.67 Å by means of the Konnert-Hendrickson program (*10*), which revealed the presence of 135 solvent atoms in the asymmetric unit. After many cycles of refinement, the final residual factor was 18.9 percent. The crystal was nearly 50 percent solvent and contained large solvent channels that passed through the lattice parallel to the *c* axis.

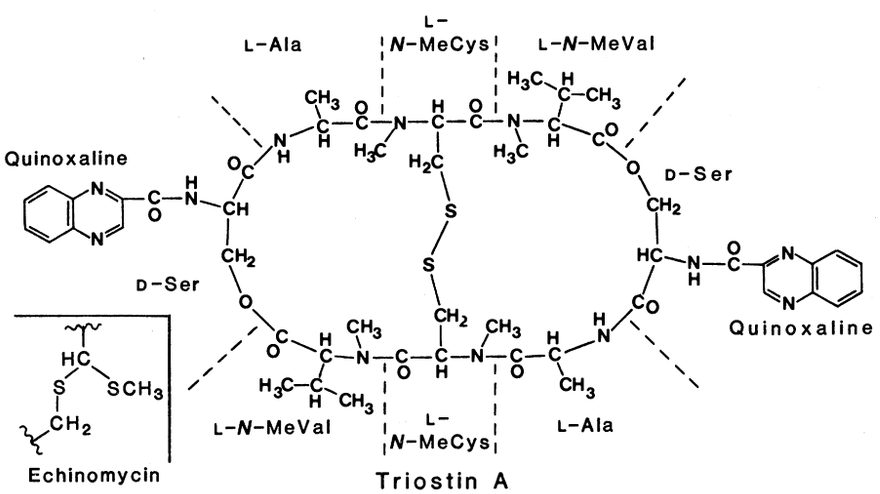

Fig. 1. The molecular formula of triostin A. The inset shows the modification of the disulfide bridge found in the closely related echinomycin.

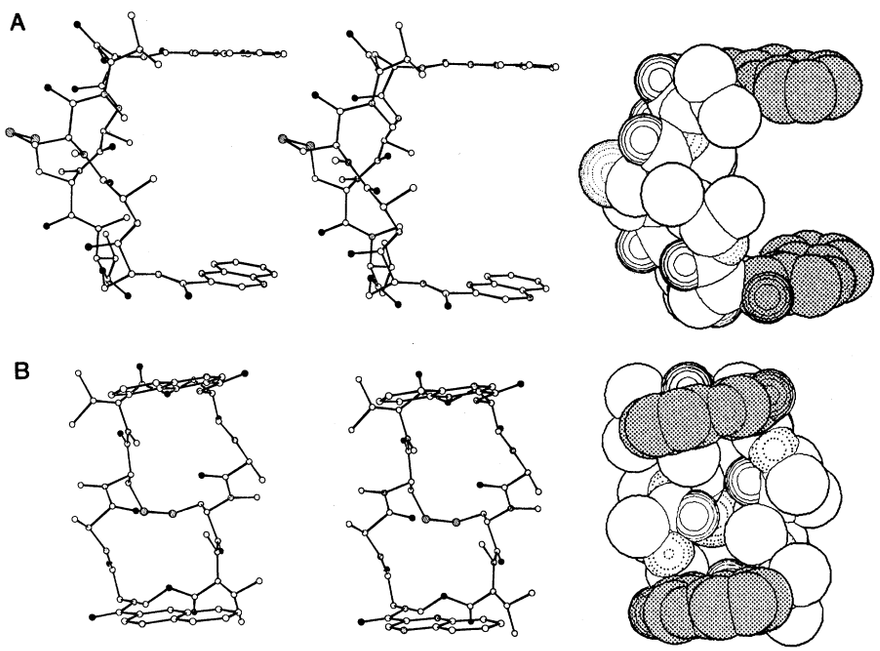

Fig. 2. Van der Waals and stereo views of the triostin A molecule in the complex. (A) Side view of the complex showing the quinoxaline rings extending perpendicular to the cyclic depsipeptide backbone. (B) A view looking down at the triostin A molecule into the surface that binds against the nucleic acid. The quinoxaline rings are both projecting directly toward the reader. In the van der Waals diagram, nitrogen atoms are stipled, oxygen atoms are drawn with concentric circles, and sulfur atoms are drawn with dotted concentric circles. The quinoxaline rings are shaded.

Conformation of Triostin A

In the complex, the DNA forms a six–base pair double helical fragment bound to two triostin A molecules. The antibiotic has the form of a bis intercalator with the quinoxaline rings bracketing the two CG base pairs at either end of the hexamer. The triostin A molecule is C-shaped with a central region containing the cyclic depsipeptide, and the two quinoxaline rings are oriented at right angles at either end (Fig. 2). Viewed from the side (Fig. 2A), the quinoxaline rings are seen to be virtually parallel to each other. Figure 2B is a view directly into the C-shaped molecule toward the residues in the cyclic chain that are in contact with the minor groove of the double helix. The cyclic depsipeptide has a rectangular shape maintained by the cystine disulfide bond near the center of the rectangle with the sulfur atoms projecting away from the nucleic acid binding interface. The serine ester linkages to valine are at the top and bottom of the rectangular molecule. The quinoxaline rings are attached to the serine α-amino groups at diagonally opposite corners of the rectangle. The serine must be D-serine to have its amino group oriented so that the quinoxaline rings are perpendicular to the cyclic depsipeptide and parallel to each other at the ends of the rectangular structure. The quinoxaline carbonyl group is attached to the α-NH_2 of D-serine through a planar amide bond in a *trans* orientation. This arrangement projects the quinoxaline rings a considerable distance away from the cyclic rectangular structure so that they are in a favorable position to intercalate. The valine side chains are found at the other pair of diagonal corners of the rectangular structure, and its *N*-methyl group projects toward the nucleic acid surface.

The alanine residues are positioned on the long sides of the rectangular molecule with its beta carbon methyl group and carbonyl group pointing toward the inside of the C-shaped structure toward the nucleic acid interface. The planar peptide linkage to serine is oriented so that the alanine amino group -NH is also pointing in the same direction.

Triostin A has a repeated sequence of four amino acids, which suggests the possibility of the molecule having a two-fold rotation axis. The conformation is close to that, but it is clear from Fig. 2

that there are significant deviations. For example, the carbonyl group of valine projects away from the nucleic acid at one end of the rectangle and toward the nucleic acid at the other end. Likewise, the disulfide bridge is skewed somewhat away from a central twofold rotation position. All the peptide groups in this molecule are in the *trans* conformation.

Table 1. Hydrogen-bond distances in d(CGTACG)–triostin A complex.

Base pairs	Distance (Å)	Triostin A–guanine	Distance (Å)
C1 N-4–G12 O-6	2.89		
C1 N-3–G12 N-1	2.84	Ala1 N–G12 N-3	2.91
C1 O-2–G12 N-2	2.78	Ala1 O–G12 N-2	3.07
G2 O-6–C11 N-4	2.80	Ala2 N–G2 N-3	3.01
G2 N-1–C11 N-3	2.79		
G2 N-2–C11 O-2	2.83		
T3 O-4–A10 N-6	2.94		
T3 N-3–A10 N-7	2.74		

Structure of the Complex

In the van der Waals diagram of the DNA hexanucleotide alone viewed from the minor (Fig. 3A) and major (Fig. 3C) groove, a space is seen between the second and third and the fourth and fifth base pairs, and the nucleic acid forms a ladder-like structure due to considerable unwinding. Triostin A binds to the minor groove side of the DNA (Fig. 3, B and D) as shown in detail in the stereo diagrams (Fig. 4). These show the manner in which the quinoxaline rings appear to be stacked on the CG base pairs. There are many van der Waals contacts found between the peptide and the oligonucleotide without the intervention of any water molecules (Fig. 5).

Alanine is a key residue with several important interactions (Fig. 5). The beta carbon methyl group side chains point toward the polynucleotide chain, wedged between the sugar residues C1 and guanine 2 (G2) on one side and sugar G12 and guanine base 12 (G12) on the other (Fig. 5). Intrusion of the methyl group is responsible for the tilt of 22° between the plane of guanine base 12 and cytosine C11; the other methyl group produces a tilt of 20° between G2 and C1. The backbone NH of both alanines form hydrogen bonds with N-3 of G2 and G12 as indicated by the dotted lines with enhanced shading (Fig. 5). The alanine residue at the lower right (Fig. 5) has its carbonyl group pointed toward the nucleic acid and forms a hydrogen bond with the amino group at N-2 of G12. However, the carbonyl group of alanine on the left side (Fig. 5) does not form a hydrogen bond because it is more than 4.1 Å away from amino group at N-2 on G2. This is another example of the lack of twofold rotational symmetry in the triostin A molecule.

The valine side chains at the corners of the rectangular cyclic peptide lie more or less tangential to the nucleic acid surface with van der Waals contact of one gamma carbon atom only on the left side (Fig. 5). The *N*-methyl groups of both valines project toward the bases where they are in contact with the cytosine O-2 in C11 and C1.

The base planes are considerably modified by the interaction of the drug with the oligonucleotide. The plane of C1 is approximately parallel to C11 (4°) even though their hydrogen-bonding guanine bases are tilted out of the plane by more than 20°. The tilting of the guanine base leads to some van der Waals contacts between it and the quinoxaline rings. Despite the deviation from planarity in the base pairs, their hydrogen bond lengths are essentially normal (Table 1). The two quinoxaline rings Q13 and Q14 are virtually parallel to each other (6°). The torsion angles of the oligonucleotide chain and the cyclic depsipeptide are listed in Table 2.

A Novel Rearrangement of DNA

Two AT base pairs in the center of the oligonucleotide that are not covered by triostin A molecules have undergone a remarkable transformation. They are no longer held together by Watson-Crick base pairs but instead by base pairs involving N-7 and N-6 of adenine instead of N-6 and N-1. This type of hydrogen bonding was first visualized in a single crystal analysis by Hoogsteen (*11*). To form Hoogsteen base pairing, the adenine residues have rotated about their glycosyl bonds and adopted the *syn* conformation instead of the *anti* conformation found for all other nucleosides in the

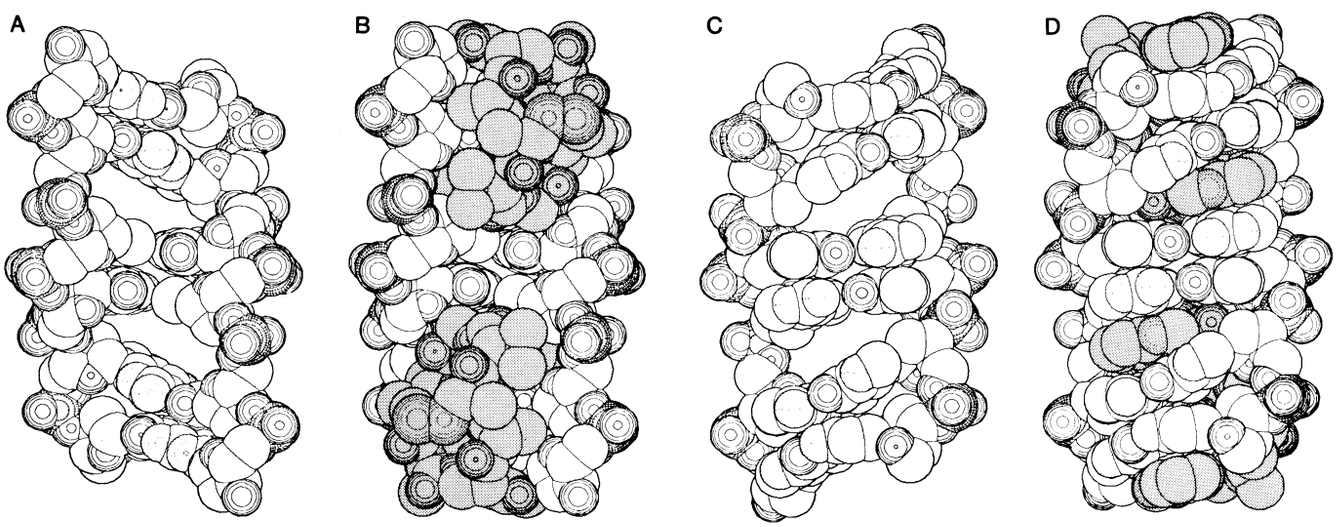

Fig. 3. Van der Waals diagram of the polynucleotide chain alone (A and C) and the oligonucleotide chain complexed with triostin A (shaded) (B and D). (A) and (B) are viewed with the minor groove facing the reader. The cyclic depsipeptide of triostin A largely fills the minor groove. In (C) and (D), the molecule is viewed from the major groove. In (D) the shaded ends of the quinoxaline rings can be seen projecting between the base pairs. All these views are oriented down the twofold axis of the molecule.

molecule. Because of this, the distance between the C-1′ carbon atoms of adenine 10 (A10) and thymine 3 (T3) are 2 Å closer together than they would be with Watson-Crick base pairing. This altered form of hydrogen bonding generates stabilization due to the close packing of the oligonucleotide around the end of the triostin A molecule. For example, the valine carbonyl group is in van der Waals contact with the sugar and base of A10, while the sugar of T3 is in contact with serine alpha and beta carbon atoms. These contacts would be lost if the two sugar rings were farther apart with Watson-Crick hydrogen bonding.

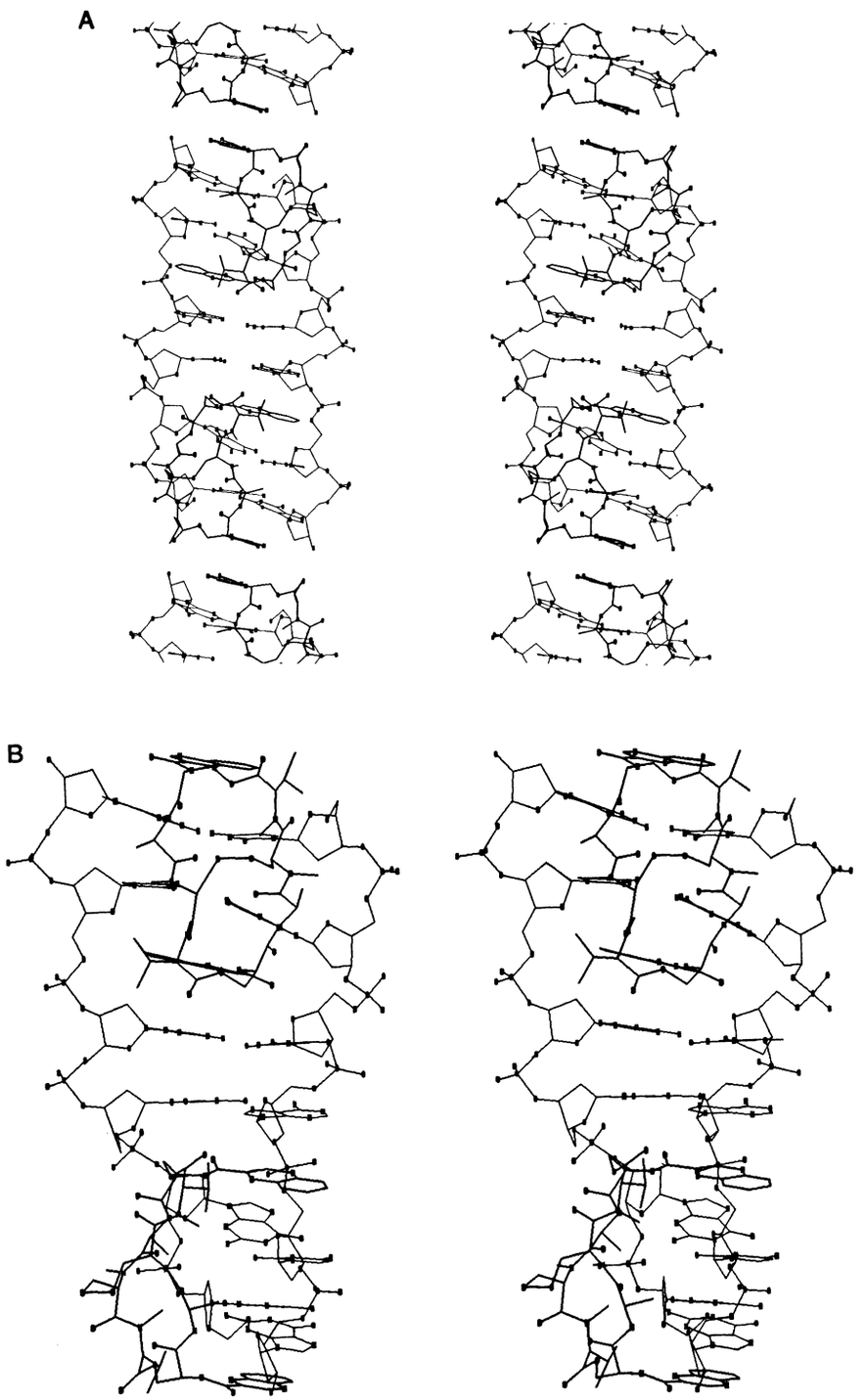

Fig. 4. Stereo views of the triostin A–oligonucleotide complex. (A) A view looking down the twofold axis into the minor groove (corresponds to the view shown in Fig. 3B). The ends of adjacent molecules in the crystal lattice above and below show the stacking of the quinoxaline rings. (B) The triostin A–oligonucleotide complex viewed rotated 30° away from that shown in (A). The complex is seen in two different orientations.

Discussion

Triostin A is made by *Streptomyces* and is used as part of the continuing biological struggle between microorganisms. It binds to DNA and acts in vivo to block replication and transcription (*4*, *5*). In the structure of this complex we can examine the role of biological design in this interaction and ask why this particular group of amino acids is used. Triostin A is a member of a large class of peptide antibiotics that are not made on ribosomes but instead by special enzymatic systems. Because of this they are able to incorporate a variety of modified amino acids, including D-residues.

The structure of this complex provides definite proof that triostin A is a bis intercalator with two base pairs between the quinoxaline rings. This is the first peptide antibiotic to have its structure solved complexed to an oligonucleotide in a presumably functional interaction. The structure of a synthetic triostin A analog has been determined, in which all four *N*-methyl groups were absent (*12*). By comparing it with the present complex, we can see that triostin A complexed to DNA has a somewhat similar conformation but with many different details, especially the orientation of the quinoxaline rings. In addition, absence of the *N*-methyl groups leads to intramolecular hydrogen bonding. The structure of a related decadepsipeptide antibiotic, luzopeptin A, has also been solved (*13*). All three of these antibiotic structures have some similarity in the organization of their cyclic depsipeptide backbones. This is in contrast to the depsipeptide ionophores such as valinomycin (*14*), which have a fundamentally different organization in that all their hydrophilic groups are on the inside of the molecule and the hydrophobic groups are outside, as is suitable for its role as an ionophore. In triostin A, all the carbonyl oxygen atoms other than those of alanine are on the outer surface of the cyclic depsipeptide, away from the nucleic acid (Fig. 2). The structure of the DNA-binding depsipeptides triostin A and luzopeptin (with only one large ring) differ considerably from the two smaller peptide rings found in actinomycin D, whose structure has been solved complexed to deoxyguanosine (*15*) or dGpC (*16*).

It had been suggested that triostin A binds to the minor groove of DNA because it could also bind to glycosylated bacteriophage T2 DNA (*17*). Those observations are supported by the present structure. The alanine residue of triostin A has been suggested to play an impor-

tant role in its binding to DNA (*12, 17*). A similar interpretation was implied by the observation that the alanine NH is hydrogen bonded in solution to guanosine and adenosine nucleosides (*18*). Those suggestions are borne out by the present structure, although the details of the suggested and observed interactions are not identical. A model has been proposed for the structure of echinomycin in solution on the basis of results from nuclear magnetic resonance (NMR) studies as well as energy minimization calculations (*19*). This model has structural features similar to those observed in the present complex. The crystal of the complex of echinomycin with the same DNA hexamer is isomorphous with that of triostin A, and the similar distribution of diffraction intensity indicates that its binding mode is the same. Cocrystals have also been made of triostin A complexed to the octamer d(GCGTACGC), and preliminary analysis reveals that this forms a complex similar to that described with the exception that there is an additional base pair covering the outer quinoxaline ring.

The interactions between triostin A and DNA can be described in several different categories: intercalation, hydrogen bonding between the peptide backbone and the nucleic acid bases, and extensive van der Waals interactions or hydrophobic bonding. All of the bonding interactions between the antibiotic and the nucleic acid are direct without the intervention of water molecules, which have been found in the binding of other compounds to DNA (*20*).

A closer look at the interaction between triostin A and the double helical DNA fragment reveals that a number of compromises have taken place in terms of stabilizing and destabilizing interactions. The stacking of bases is a stabilizing interaction, and interplanar spacing between unsaturated ring systems is generally near 3.4 Å, which is the thickness of the unsaturated aromatic ring. However, the interaction of triostin A with the double helix has modified the stacking interactions considerably. Although the rings lie over each other in such a position that they could have stabilizing interactions with their pi electron clouds, in many cases the distance has increased so that there is a corresponding decrease in the amount of stabilization. For example, the quinoxaline ring Q14 stacks under the pyrimidine ring of adenine 10 (Fig. 5), but the average distance between the planes is 3.53 Å. Similarly, Q14 lies partly over cytosine C11, but here the distance is 3.97 Å, which means that the stabilization due to stacking is likely to be diminished considerably. A

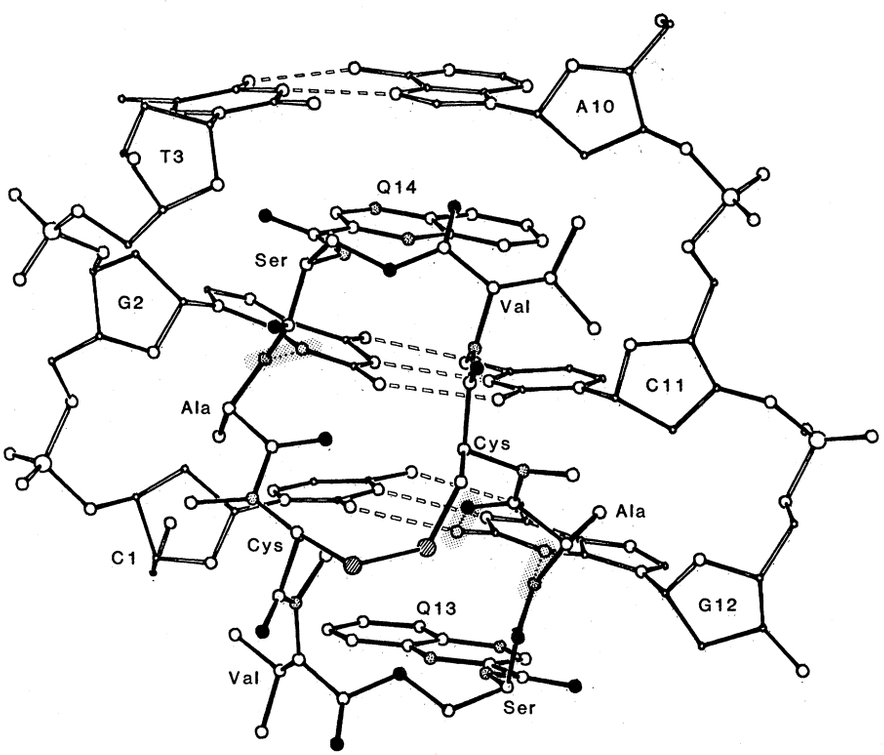

Fig. 5. A perspective view showing the interaction of triostin A with three base pairs of the oligonucleotide complex. Hydrogen bonds are represented by dashed lines. The nucleotide residues are labeled by letters in the deoxyribose rings. The numbering proceeds from 1 through 6 on one molecule and 7 through 12 on its hydrogen-bonded complement, with only three of the six base pairs shown in the diagram. The quinoxaline rings are labeled Q13 and Q14. The amino acids are labeled as indicated. The hydrogen bonding between the alanine residues and the guanine bases is emphasized by shading.

Table 2. Torsion angles of the d(CpGpTpApCpG) oligonucleotide backbone and the triostin A cyclic depsipeptide chain. Torsion angles are defined as $P\overset{\alpha}{-}O\text{-}5'\overset{\beta}{-}C\text{-}5'\overset{\gamma}{-}C\text{-}4'\overset{\delta}{-}C\text{-}3'\overset{\varepsilon}{-}O\text{-}3'\overset{\zeta}{-}P$ and χ for the glycosyl (in DNA), O-1′—C-1′—N-9—C-4 (in purines), and O-1′—C-1′—N-1—C-2 (in pyrimidines). Torsion angles in peptides are defined as described (*25*).

Species	Base	α	β	γ	δ	ε	ζ	χ	ϕ	ψ	ω
Oligonucleotide backbone	C1			37.9	119.7	−166.7	−27.4	−114.9			
	G2	−148.7	123.6	131.1	138.6	−104.7	179.9	−79.7			
	T3	−30.7	159.1	6.2	133.4	176.9	−72.3	−87.4			
	A4	−86.7	177.6	55.3	88.3	−163.4	−53.2	67.1			
	C5	−154.3	142.9	−175.4	83.6	−168.3	−75.1	−98.5			
	G6	−76.6	174.4	50.1	96.7			−105.1			
Triostin A–depsipeptide chain	D-Ser								105.1	−30.8	−156.7
	Ala								−101.6	161.5	−161.3
	Cys								−132.8	76.0	−167.4
	Val								−117.4	31.5	165.6*
	D-Ser								104.0	−52.9	−158.6
	Ala								−62.8	177.0	175.8
	Cys								−139.5	77.6	160.3
	Val								−65.6	−54.8	−168.6*

*Ester linkage.

similar distance is found between Q13 and C1. In contrast, the end of ring Q14 (and Q13) has van der Waals interactions with one part of the guanine 2 (G12) because of its tilt. As mentioned above, the large dihedral angles between G2 and C1 and G12 and C11 on the other side of the helix mean that these residues have largely lost their stacking interactions.

A major stabilizing interaction between the antibiotic and the DNA occurs through a series of 19 van der Waals contacts between atoms of the cyclic depsipeptide and components of the nucleic acids. Some of these, as mentioned above, include the beta carbon methyl group of alanine in close contact with the deoxyribose of G12 on one side and G2 on the other side. Likewise, the *N*-methyl groups of both valines have close contacts with cytosines 11 and 1, which is probably responsible for these cytosines no longer having close stacking interaction with the quinoxaline rings. The DNA molecule has changed its conformation in several ways because of the antibiotic binding. This can also be seen in the modifications of the sugar pucker, which is normally C-2′-*endo* in B DNA. Here the sugars show a variety of puckers: C1, O-1′-*exo*; G2, C-2′-*endo*; T3, C-3′-*exo*; A4, O-1′-*endo*; C5, C-3′-*endo*; and G6, C-3′-*endo*.

In looking at the complex, it is interesting to ask why these particular eight amino acids are used in triostin A. Part of the answer can be seen in the detailed interactions found between components of the amino acids and the nucleic acids. Alanine plays a central role because its beta carbon methyl group has van der Waals binding to the sugar-phosphate backbone. At the same time its NH-group forms hydrogen bonds to the nucleic acid on both sides of the ring, and the carbonyl group of one residue also forms a hydrogen bond. It is not clear whether the lack of twofold rotational symmetry in triostin A is associated with its asymmetric binding to the oligonucleotide fragment or whether it is an intrinsic part of the conformation. The disulfide linkage between the L-cystine residues deviates significantly from the position it would have if there were a dyad axis, suggesting that it may be an intrinsic feature of the molecule. The methylation of the cystine NH group effectively prevents its being involved in hydrogen bonding. Because the cystine NH group is connected to the alanine carbonyl group through a planar peptide linkage, the methylation may allow the alanine carbonyl group to be oriented in such a position that it can hydrogen bond to guanine at N-2. If that *N*-methyl group were absent, hydrogen bonding by the cystine NH might give rise to an altered conformation as seen in the structure of the synthetic des-*N*-tetramethyltriostin A (*12*).

One valine side chain has van der Waals contacts with a sugar through one gamma carbon methyl group. The other gamma carbon methyl group is not in van der Waals contact. Triostins are found as a family of antibiotics with additions of various aliphatic groups on the valine side chains (*1*, *2*), and these are the only side chains where this is possible. The valine *N*-methyl groups make van der Waals contacts with cytosine rings. If the *N*-methyl group were not present, water would probably compete with the nucleic acid binding by forming hydrogen bonds with the backbone NH group.

The NH group of serine remains unmodified and is not hydrogen bonded to any atom, not even a water molecule. This represents a destabilization, but there is no room for any additional material in triostin A in this region. Finally, as mentioned above, the D form of serine is necessary in order to have the quinoxaline rings project perpendicular to the rectangular cyclic depsipeptide. In an L residue the quinoxaline rings could not occupy this position, which is needed for an intercalative interaction.

Peptide antibiotics have many *N*-methylated amino acids and other amino acids with free NH groups. The suggestion from this structure is that some of the latter may be used for forming hydrogen bonds to the nucleic acids, thereby generating binding specificity. In the complex, many of the base-stacking interactions are decreased, but they are more than compensated by the large number of tight van der Waals (including hydrophobic) interactions between the cyclic depsipeptide and the nucleic acid. The total of the stabilizing interactions is the sum of a large number of contacts.

Even though triostin A is a bis intercalator, the actual stabilization due to intercalative interactions is less than would otherwise be imagined since the numerous close contacts between the peptide backbone and the nucleic acids have disrupted the stacking interactions. This is in marked contrast to the results of numerous structural studies of simple intercalators, which in general show close contacts between base pairs on both sides of the intercalating base (*21*). This is true even when there are fairly complex side chains attached to the intercalated planar ring, as for example in the stucture of a daunomycin-oligonucleotide complex (*19*).

Binding studies in solution of triostin A and echinomycin show a preference for double-stranded DNA with higher GC content (*2*). The hydrogen bonding of the peptide backbone to the two CG base pairs located between the two quinoxaline rings is in general agreement with this. A recent study (*22*) using a footprinting technique has shown that echinomycin has a binding-site size of four base pairs. Fragments of the plasmid pBR322 were used and a number of binding sites were analyzed. The tightest binding was found when the sequence dCpG was in the central two base pairs and when AT or TA base pairs were found in the outer two. The sequence specificity dCpG is different from that of actinomycin D, for example, which binds to dGpC (*15*, *16*). From the structure of the complex, we can see that the major interactions that determine sequence specificity are the hydrogen bonding interactions of the alanine residues. The hydrogen bonding is not symmetric in that only one of the two alanine carbonyl groups receives hydrogen bonds from the N-2 of guanine. It is possible that this asymmetry is intrinsic and that triostin A or echinomycin never achieve complete twofold symmetry in binding to DNA. However, this may also be a consequence of triostin A binding to the outer two base pairs of the hexamer, and the asymmetry may be an end effect. This question may be resolved in the analysis of the triostin A–octamer complex, or, alternatively, it may be resolved by NMR studies in solution.

Solution studies have revealed a DNA unwinding angle of 45° to 55° per echinomycin molecule (*23*). In the complex, the observed unwinding of the base pairs around quinoxaline 14 is 12°; however, the unwinding between the base pairs G2 and C11 and C1 and G12 (Fig. 5) is 27°. The observed unwinding between the central AT and TA base pairs is about 3°. If we assume that quinoxaline 13 would produce an unwinding similar to that of Q14 in a long DNA molecule, then the total unwinding per molecule in the complex is very close to 55° per antibiotic molecule. The major component of the unwinding does not occur in the base pairs surrounding the intercalator but rather one base pair removed because of the interactions with the rest of the molecule. A similar situation was found to occur in the structure of daunomycin complexed to an oligonucleotide (*20*).

A totally unexpected discovery in this structure is the remodeling of the DNA duplex in its center, in which the two adenine rings have rotated around to a *syn* conformation so that they form Hoogsteen base pairs with the thymine

residues. This type of base pairing is stabilized by the van der Waals binding to the sugar residues at the end of the triostin A molecule (Fig. 5) near the AT base pairs. Solution studies of echinomycin binding to DNA (*22*) show that the tightest tetrameric footprinting sequence is one in which AT or TA base pairs are found surrounding the central dCpG sequence. The presence of flanking AT base pairs in the solution binding suggests that these may form the Hoogsteen pairs on both sides of bis intercalators. GC base pairs can also form a similar type of Hoogsteen hydrogen bonding with guanine in the *syn* conformation, but it requires protonation of the cytosine residue in the N-3 position in order to form two hydrogen bonds. This type of bonding might be stabilized at a somewhat lower *p*H. We examined the structure of the two AT base pairs in the center of the molecule and conclude that it is possible to continue this type of interaction to build a DNA double helix model that is right-handed and incorporates such an alternative hydrogen bonding systematically into its structure. This is the second double-helical structure with nucleotides in the *syn* conformation. In B DNA all of the nucleotides adopt the *anti* conformation. However, in left-handed Z DNA, alternate residues along the polynucleotide chain have the *syn* conformation (*24*). It is possible that *anti-syn* conformational changes are widespread in DNA and may appear elsewhere as well.

The structure of the triostin A–oligonucleotide complex illustrates the extent to which a small molecule containing eight amino acids can rework and modify the structure of DNA in a variety of ways. It also suggests that peptide structures appear to be less flexible than DNA; the DNA changes shape to accommodate the interaction. Many of these interactions are unexpected, especially those dealing with the distortion of the internal geometry of DNA through the binding, apparent loss of base stacking stabilization, and modifications of the base pairing outside the intercalative site. The stabilization of the complex is due to a large number of interactions involving many different forces. We might ask how useful this structure is in leading to information about the nature of interactions that may be found between the nucleic acids and proteins containing many more amino acids. It seems likely that some aspects of these interactions will be found in proteins, namely, hydrogen bonding involving the peptide backbone and quite possibly intercalative interactions involving planar side chains of amino acids. It is also likely that there will be a detailed fitting of some elements of polypeptide chains with the nucleic acids. However, the total number of contacts between a protein and a nucleic acid need not necessarily be much greater than those seen between the triostin A molecule and the oligonucleotide. It is possible that the density of interactions might be less, as they could be, if they were spread over a larger protein surface. It remains to be seen whether there are similar modifications of DNA structure incorporating its apparent structural flexibility associated with protein binding. However, it should be borne in mind that elucidation of the fine details of such interactions will require high-resolution diffraction analysis, which is less likely to occur in studies of larger molecules.

References and Notes

1. K. Katagiri, T. Yoshida, K. Sato, in *Antibiotics*, vol. 3, *Mechanism of Action of Antimicrobial and Antitumor Agents*, J. W. Corcoran and F. E. Hahn, Eds. (Springer-Verlag, Heidelberg, 1975), pp. 234–251.
2. M. J. Waring, in *Antibiotics*, vol. 5, *Mechanism of Action of Antieukaryotic and Antiviral Compounds*, F. E. Hahn, Ed. (Springer-Verlag, Berlin, 1979), part 2, pp. 173–194.
3. M. J. Waring and L. P. G. Wakelin, *Nature (London)* **252**, 653 (1974).
4. D. C. Ward, E. Reich, I. H. Goldberg, *Science* **149**, 1259 (1965).
5. K. Sato, O. Shiratori, K. Katagiri, *J. Antibiot. Ser. A* **20**, 270 (1967).
6. G. van der Marel, C. A. A. van Boeckel, G. Wille, J. H. van Boom, *Tetrahedron Lett.* **22**, 3887 (1981).
7. Triostin A was a gift from T. Yoshida (Shionogi, Osaka, Japan) and echinomycin (Ciba-Geigy) was a gift from M. J. Waring (University of Cambridge).
8. A. H.-J. Wang, T. Hakoshima, G. van der Marel, J. H. van Boom, A. Rich, *Cell* **37**, 321 (1984).
9. T. L. Blundell and L. N. Johnson, *Protein Crystallography* (Academic Press, New York, 1976).
10. W. A. Hendrickson and J. Konnert, *Biomolecular Structure: Conformation, Function and Evolution*, R. Srinivasan, Ed. (Pergamon, Oxford, 1979), vol. 1, pp. 43–57.
11. K. Hoogsteen, *Acta Crystallogr.* **12**, 822 (1959).
12. M. A. Viswamitra *et al.*, *Nature (London)* **289**, 817 (1981); M. B. Hossain *et al.*, *J. Am. Chem. Soc.* **104**, 3401 (1982).
13. E. Arnold and J. Clardy, *J. Am. Chem. Soc.* **103**, 1244 (1981).
14. M. Pinkerton, L. K. Steinrauf, P. Dawkins, *Biochem. Biophys. Res. Commun.* **35**, 512 (1969).
15. H. M. Sobell, S. C. Jain, T. D. Sakore, C. E. Nordman, *Nature (London) New Biol.* **231**, 200 (1971).
16. F. M. Takusagawa, S. Dabrow, S. Neidle, H. M. Berman, *Nature (London)* **296**, 466 (1982).
17. L. P. G. Wakelin and M. J. Waring, *Biochem. J.* **157**, 721 (1976).
18. Y. Kyogoku, N. Higuchi, M. Watanabe, K. Kawano, *Biopolymers* **20**, 1959 (1981).
19. G. Ughetto and M. J. Waring, *Mol. Pharmacol.* **13**, 579 (1977).
20. G. J. Quigley, A. H.-J. Wang, G. Ughetto, G. van der Marel, J. H. van Boom, A. Rich, *Proc. Natl. Acad. Sci U.S.A.* **77**, 7204 (1980).
21. C.-C. Tsai, S. C. Jain, H. M. Sobell, *J. Mol. Biol.* **114**, 301 (1977); S. C. Jain, C.-C. Tsai, H. M. Sobell, *ibid.*, p. 317; S. Neidle, A. Achari, G. L. Taylor, H. M. Berman, H. L. Carrell, J. P. Glusker, W. C. Stallings, *Nature (London)* **269**, 304 (1977); A. H.-J. Wang, J. Nathans, G. van der Marel, J. H. van Boom, A. Rich, *ibid.* **276**, 471 (1978); A. H.-J. Wang, G. J. Quigley, A. Rich, *Nucl. Acids Res.* **6**, 3879 (1979).
22. M. M. van Dyke and P. B. Dervan, *Science* **225**, 1122 (1984).
23. J. S. Lee and M. J. Waring, *Biochem. J.* **173**, 115 (1978).
24. A. H.-J. Wang, G. J. Quigley, F. J. Kolpak, J. H. van Boom, G. van der Marel, A. Rich, *Nature (London)* **282**, 680 (1979); A. Rich, A. Nordheinor, A. H.-J. Wang, *Annu. Rev. Biochem.* 53, 791 (1984).
25. IUPAC-IUB Commission on Biochemical Nomenclature, *Biochemistry* **9**, 3471 (1970).

26. This research was supported by grants from the National Institutes of Health, National Science Foundation, American Cancer Society, National Aeronautics and Space Administration, Netherlands Organization for the Advancement of Pure Research, North Atlantic Treaty Organization (to G.U.), and Istituto Strutturistica Chimica, Consiglio Nazionale Delle Richerche (Italy).

9 May 1984; accepted 28 June 1984