Accelerated Publications

Crystallographic Analysis of a Novel Complex of Actinomycin D Bound to the DNA Decamer CGATCGATCG†

Howard Robinson,‡ Yi-Gui Gao,‡ Xiang-lei Yang,‡ Ruslan Sanishvili,§ Andrzej Joachimiak,§,|| and Andrew H.-J. Wang*,‡,⊥

Department of Biochemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, Structural Biology Center, Bioscience Division, Argonne National Laboratory, Argonne, Illinois 60439, Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, Illinois 60208, and Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan, Republic of China

Received December 18, 2000; Revised Manuscript Received February 21, 2001

ABSTRACT: The potent anticancer drug actinomycin D (ActD) acts by binding to DNA, thereby interfering with replication and transcription. ActD inhibits RNA polymerase far more specifically than DNA polymerase. Such discrimination is not easily understood by the conventional DNA binding mode of ActD. We have solved and refined at 1.7 Å resolution the crystal structure of ActD complexed to CGATCGATCG, which contains no canonical GpC binding sequence. The crystal data are space group $P4_32_12$, $a = b = 47.01$ Å, and $c = 160.37$ Å. The structure was solved by the multiple wavelength anomalous diffraction method using a 5-bromo-U DNA. The asymmetric unit of the unit cell contains two independent dimers of a novel slipped duplex complex consisting of two decamer DNA strands bound with two ActD drug molecules. (The DNA in one dimer is numbered C1 to G10 in one strand and C11 to G20 in the complementary strand and in the second dimer, C101 to G110 and C111 to G120, respectively.) The structure reveals a highly unusual ActD binding mode in which the DNA adopts a slipped duplex with the A3-T4/A13-T14 dinucleotides looped out. ActD intercalates between G2-C11* (C11* being from a symmetry-related molecule) and C5-G20 base pairs. Two such slipped duplex−ActD complexes bound to each other by mutually intercalating their T4/T14 bases into the helix cavities (located between C5-G20 and G6-C19 base pairs) of neighboring complexes, forming a dimer of drug−DNA complexes. The binding site mimics the drug binding at the elongation point during transcription. Modeling studies show that the ActD−DNA complex fits snugly in the active site cavity in RNA polymerase but not in DNA polymerase. This may explain the strong preference of ActD inhibition toward transcription.

Actinomycin D (Figure 1) is a potent anticancer drug, and it binds strongly to DNA duplexes, thereby interfering with replication and transcription (*1*, *2*). Earlier studies showed that the binding sequence specificity of ActD[1] is the 5′-GpC site (*3*, *4*), although other sites (e.g., GpG) also have an unusual affinity toward ActD (*5*). ActD binds to DNA by intercalating its phenoxazone ring at a GpC step with the drug's two cyclic pentapeptides located in the DNA minor groove (*3*, *4*). The binding specificity derives from the strong hydrogen bonds found between the NH/C═O groups of threonines of ActD and the corresponding N3/N2 amino group of adjacent guanine bases of the GpC step. The binding affinity to the GpC site is also influenced by the flanking sequences.

The crystal and solution structures of the ActD−GAAGCTTC complex have been analyzed which revealed the *N*-methyl group of MeVal wedging between the bases at the ApG step, resulting in helix kinks on both sides of the intercalator site (*4*, *6*). Surprisingly, ActD forms a very stable complex with GATGCTTC in which the same methyl group fits snugly in a cavity at the TpG step created by the T·T mismatched base pair (*6*). Such structural information helps to understand the sequence preference of ActD toward -XGCY- tetranucleotides.

Binding of ActD to other noncanonical sequences (i.e., those different from 5′-GpC) has been studied. For example, ActD binds to two adjacent GpC sites in GCGC simultaneously, and a kink is induced due to the crowding of the neighboring pentapeptide rings (*7*). ActD appears to bind to certain "single-stranded" sequences with high affinity (*8*),

† This work was supported by the NIH and NSC (Taiwan, ROC) to A.H.-J.W. and the DOE to A.J.

* Correspondence should be addressed to this author (tel, 8862-2788-1981; fax, 8862-2788-2043; e-mail, ahjwang@gate.sinica.edu.tw).
‡ University of Illinois at Urbana-Champaign.
§ Argonne National Laboratory.
|| Northwestern University.
⊥ Academia Sinica.

[1] Abbreviations: ActD, actinomycin D; MAD, multiple wavelength anomalous diffraction; NMR, nuclear magnetic resonance; rmsd, root-mean-square deviation; PEG400, poly(ethylene glycol) 400; br^{5}U, 5-bromodeoxyuridine.

10.1021/bi002859z CCC: $20.00
Published on Web 04/17/2001

Table 1: Crystal and Refinement Data of the CGATCGA[br^{5}U]CG−ActD Complex

	0.9530 Å	0.9537 Å low-remote	0.9197 Å inflection point	0.9196 Å peak	0.8563 Å high-remote
crystallographic data					
$a = b$ (Å)	47.03	47.02	46.97	46.97	46.97
c (Å)	160.37	160.38	160.29	160.28	160.28
space group	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$
resolution (Å)	20−1.7	20−2.42	20−2.33	20−2.33	20−2.33
no. of reflections [$>0\sigma(I)$]	19747	7404	8187	8169	8238
$\langle I/\sigma(I)\rangle$	23.3 (3.0)[a]	26.7	48.7	48.9	37.7
R_{merge} (%)	5.7 (53.3)[a]	3.1	3.2	3.5	2.4
completeness (%)	95.2 (88.3)[a]	99.2	98.5	98.5	99.1
refinement data					
no. of reflections [$>2.0\sigma(I)$]	17034				
$R_{working}/R_{free}$ (5% data)	0.231/0.270				
rmsd of bonds (Å)/angles (deg)	0.008 /1.8				
no. of DNA/drug/water atoms	770/360/135				

[a] The numbers in parentheses are from the outermost shell (1.76−1.70 Å) data.

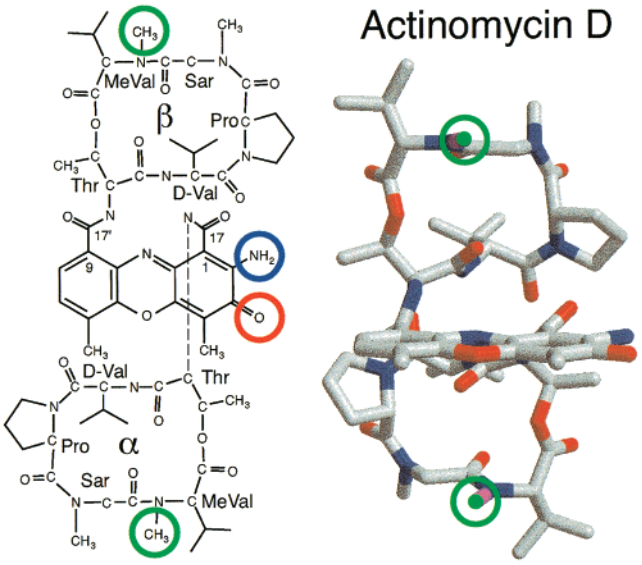

FIGURE 1: Chemical (left panel) and molecular structure (right panel) of actinomycin D (ActD). The five amino acids (threonine, *N*-methylvaline, sarcosine, proline, and D-valine) and the atomic numbering on the phenoxazone ring are shown. The amino acids on the quinonoid ring are labeled α, and those from the benzenoid ring are labeled β. The *N*-methyl groups on two *N*-methylvalines which play the important role of wedging open base pairs are marked with green circles. The 2-amino and 3-keto groups on the phenoxazone ring are marked with blue and red circles. The chemical structure is drawn in a manner resembling the three-dimensional structure of ActD shown in the right panel.

which might suggest that ActD induces a local double helical structure at the binding site. However, no structural information is available on the binding mode of the latter system.

Presumably, the potent anticancer activity of ActD is related to its ability to bind to DNA, thereby interfering with replication and transcription. However, ActD inhibits RNA polymerase far more specifically than DNA polymerase. With the structural information on the conventional DNA binding mode of ActD available thus far, it is difficult to understand the molecular basis of such discrimination. Here we show the first high-resolution crystal structure of the complex between ActD and a novel DNA sequence, providing valuable information regarding the inhibitory action of ActD on the transcription process.

MATERIALS AND METHODS

The DNA molecules, including the brominated derivatives, were synthesized and purified by gel filtration column chromatography. ActD was purchased from Sigma Chemical Co. (St. Louis, MO) and dissolved in methanol as stock solutions. The concentration of ActD solutions was determined from its optical density ($\epsilon_{224nm} = 35\,280$). Crystals were obtained from solutions having 1.3 mM DNA single strand, 1.3 mM ActD, 40 mM glycine buffer (pH 4.5), 10 mM $BaCl_2$, 3 mM spermine, and 4% PEG400 solution, equilibrated at 4 °C with 30 mL of 50% PEG400, using the vapor diffusion method (*9*). The crystallographic statistics are listed in Table 1. Diffraction data were measured at 110 K at Structural Biology Center undulator beamline 19ID at the Advanced Photon Source using fundamental undulator harmonics and 3 × 3 mosaic CCD detector as described before (*10*, *11*). Crystallographic data integration and reduction were done with the program package HKL2000 (*12*).

Experimental phases of the ActD−CGATCGA[br^{5}U]CG complex crystal were obtained by the MAD method using an approach similar to that described recently (*10*, *11*). MAD data were collected from four wavelengths using bromine as the anomalous scattering atoms. Bromine sites were found using the Patterson heavy atom search method, and the phases were calculated semiautomatically as implemented in the crystallographic suite CNS (*13*). The figure of merit is 0.69/0.91, before/after density modification, respectively. The resulting electron density map at 2.3 Å resolution was of high quality (Figure 1S, Supporting Information), which allowed the initial models to be built easily.

The structure has been refined, first by the simulated annealing procedure incorporated in X-PLOR (*14*) and then by SHELX97 (*15*). The DNA force field parameters of Parkinson et al. (*16*) with modifications to allow sugar pucker to vary were used. The force field of ActD was generated using the atomic coordinates of its high-resolution crystal structure (*3*). Water molecules were located by the procedure incorporated in SHELX97. We were careful in the criteria of the inclusion of water molecules. Only first shell waters and well-defined higher shell waters were included. All atoms were treated with isotropic temperature factors. The procedure of SWAT in SHELX97 by Moews and Kretsinger (*17*) has been applied to model diffuse solvent. The atomic coordinates of the crystal structure have been deposited at

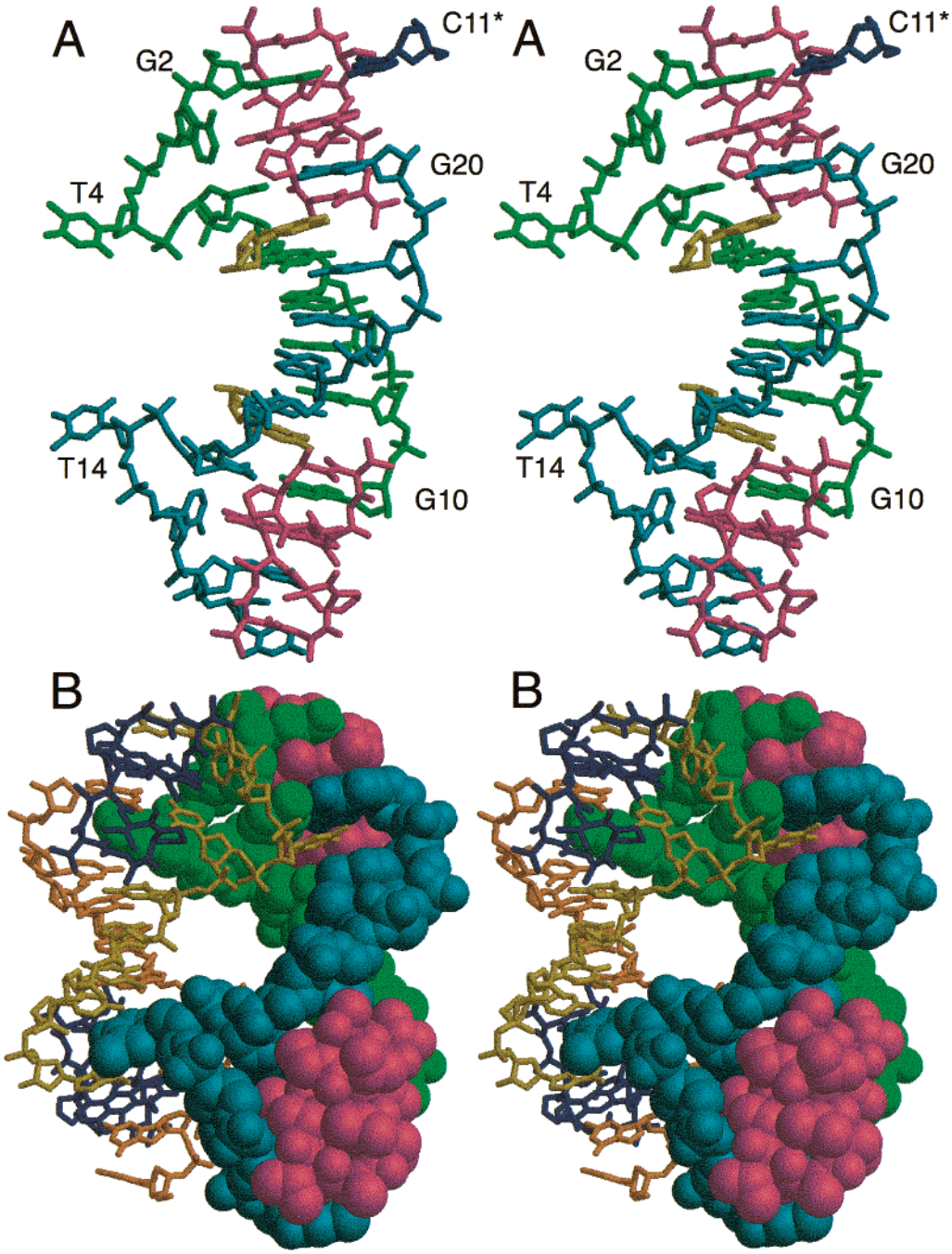

FIGURE 2: (A) The refined structure of the 2:2 ActD−CGATCGA[br^{5}U]CG complex. (B) The dimer formed by two complexes (one in skeletal model and one in van der Waals model) related by a local 2-fold symmetry. The mutually intercalated T4/104 and T14/T114 nucleotides locked the two complexes together. The numbering of nucleotides is defined in the Abstract.

Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (accession number 1I3W).

RESULTS AND DISCUSSION

The asymmetric unit of the unit cell contains two independent dimers of a novel slipped duplex complex consisting of two decamer DNA strands bound with two ActD drug molecules (Figure 2S). The structure of the ActD−CGATCGATCG complex (Figure 2A) contains many unexpected features. First, the DNA is a slipped duplex, with two overhanging 5′-end C1-G2-A3-T4 and C11-G12-A13-T14 tetranucleotides. Of the entire decamer DNA, only the (G6-A7-T8-C9)·(G16-A17-T18-C19) fragment adopts the B-DNA conformation. The C1 nucleotide is disordered, and the C11 is base paired with G2* from a (noncrystallographic) 2-fold-related complex. In the 5′-C1-G2-A3-T4 part, the A3-T4 (and A13-T14 of the opposite strand) nucleotides are extruded, with the A3 (A13) base plane parallel to the helix axis and contacting other bases and ActD on the major groove side. The T4 (T14) base is flipped out, pointing away from the complex. The (C5-G6-A7-T8-C9-G10)·(C15-G16-A17-T18-C19-G20) duplex part is also distorted. The T104 base from the (noncrystallographic) 2-fold-related complex intercalates into the (C5pG6)·(C19pG20) step.

The unusual DNA structure is associated with extensive changes of backbone torsion angles from those of B-DNA

(listed in Table 1S). The extrusion of the A3-T4 nucleotides is mostly due to the nonhelical ζ/α combination of the A3pT4(A13pT14) steps (*trans*/*gauche*$^+$) and T4pC5(T14pC15) steps (*gauche*$^+$/*gauche*$^+$). The extension of the backbone at the intercalation cavity at the (C5pG6)·(C19pG20) step is produced by the γ angle adopting a *gauche*$^-$ conformation and a change of both C19/G20 sugar puckers to the N-type.

Two complexes in the asymmetric unit, each with two ActD drugs bound to one severely distorted decamer duplex, are locked together into a dimer by mutually intercalating the flipped-out T4/T14 nucleotides into each other (Figure 2B). The 2-fold-related A3/A103 extruded adenines are stacked against each other (vide infra). The two complexes within a dimer are related by a local noncrystallographic 2-fold axis. The interactions among the crystallographic-related dimers in the crystal lattice are shown in Figure 2S. There are four independent copies of ActD and CGATC-GATCG in the asymmetric unit. The consistency of the conformations is high, with the averaged rmsd derived from those four copies being 0.43 Å for DNA and 0.34 Å for ActD. (Unless otherwise stated, all numbers discussed are the average of four numbers.)

It is intriguing to ask whether the dimer, resulting from four mutually intercalated thymines and stacking of adenines, is stable in solution or not. Such an interlocking motif between two duplexes resulting from the binding of ActD may have general biological consequences. The binding of ActD at a G-rich region in DNA induces destabilization of base pairs due to the wedging of *N*-methyl of MeVal in the flanking regions of the intercalation site (vide infra). In the present structure, the destabilization attracts neighboring duplexes to exchange nucleotides to mutually intercalate into each other. Therefore, the binding of ActD may produce a hot spot, at which two duplexes attach, mimicking a recombination junction (*18*). Structures of exchanging nucleotides (to form base pairs) between neighboring drug−DNA complexes have been determined recently (*19*, *20*), but they differ in details from the present strcuture.

The conformations of the two ActDs, bound at both ends of the complex (Figure 2A), are slightly different from that of the free ActD (*3*). There is a rotation (∼20°) about the C1−C17 (and C9−C17′) bond in ActD such that the proline side of the pentapeptide ring is moved away from the phenoxazone plane by about 1 Å relative to that in the free ActD. No evidence of disordering of ActD is found, unlike that seen in the structure of the ActD−GAAGCTTC complex (*4*). The phenoxazone ring is intercalated between the G2-C11* and C5-G20 base pairs from the minor groove side, with its 3-keto group nudging against the extruded A3 base. The 2-amino group of phenoxazone is hydrogen bonded (3.01 Å) to O4′ of the adenine A3/A13 residue, whose N6 amino group (A3 only) is in turn hydrogen bonded (3.05 Å) to O1P of G12. Remarkably, there is no direct backbone linkage between adjacent base pairs. The phenoxazone ring is mostly sandwiched between two guanines (G2 and G20) with a manner similar to that in the 2:1 deoxyguanosine:ActD crystal structure (*3*).

Figure 3 shows the close-up view of the complexes at the ActD intercalation site. In the complex, ActD maintains its pseudo-2-fold symmetry, with the rmsd between the two penptadepsipeptide rings being 0.23 Å (using all eight ActDs). Two interpeptide NH···O═C hydrogen bonds (2.88 Å) between the D-valines of peptide rings α and β provide the conformational rigidity of ActD. The strong intermolecular ActD−DNA hydrogen bonds between the G2-N2/N3 and the CO/NH of Thr(α) and the G20-N2/N3 and the CO/NH of Thr(β), respectively, define the strong sequence preference for two adjacent 5′-G nucleotides from opposite strands of DNA. The outer edges of the peptide rings (i.e., the Sar-MeVal dipeptide part) from ActD reach beyond the flanking base pairs in the minor groove.

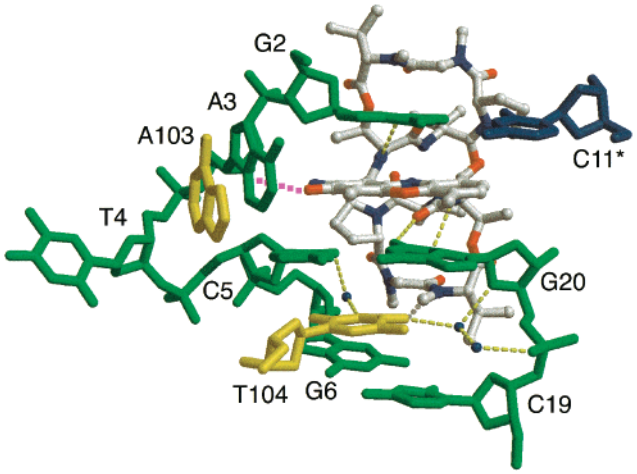

FIGURE 3: Close-up view of the refined structure of the ActD−DNA complex at the ActD intercalation site. The phenoxazone ring is intercalated at the (G2...C5)·(G20...C11*) step, with its 2-amino group hydrogen bonded to the O4′ of the extruded A3 nucleotide and the 3-keto pointing into the A3 base. Two adenine bases (A3/A103) from 2-fold-related complexes are stacked. The T4 (purple) from the partner complex within a dimer is intercalated between the (C5pG6)·(C19-G20) step, with three conserved bridging water molecules.

In fact, there is no base pair immediately beyond the flanking G2-C11* and C5-G20 base pairs in the present structure. The structural basis for this observation may be understood as follows. While the GpC site is the major preferred binding site, other sites such as $(G)_n$ (*5*) or isolated G in single-stranded DNA (*8*) have been noted to have an unusual affinity toward ActD. Moreover, the flanking sequences at the GpC binding site play an important role in the binding affinity of ActD. In the ActD−GAAGCTTC complex, the *N*-methyl of MeVal (marked with green circles in Figure 1) is in close contact with the base A3, forcing the A3pG4 step to be wedged open by the methyl group (*4*, *6*). Such a wedging interaction by a methyl group is structurally destablizing, which may be overcome in different ways, e.g., by protein binding (*21*, *22*). Alternatively, ActD binds GATGCTTC with high affinity due to the unique T·T base pair conformation which generates a cavity in the minor groove where the *N*-methyl of MeVal is located, avoiding the wedging of the T3pG4 step (*6*, *23*).

In the present structure, no base pair is found beyond the G2-C11* base pair. The ActD bound at the G12* end of the 4_3-axis-related complex stacks over the π system of the G2-C11* base pair using the (ring α) D-Val C_β methyl and the peptide *N*-methyl atoms. An additional intercalation cavity is found beyond the C5-G20 base pair where the T104 from the partner complex within a dimer is inserted. There is no direct hydrogen bond between the intercalated T104 and the nearby atoms. Three bridging water molecules are found (Figure 3). The *N*-methyl of MeVal is in contact with

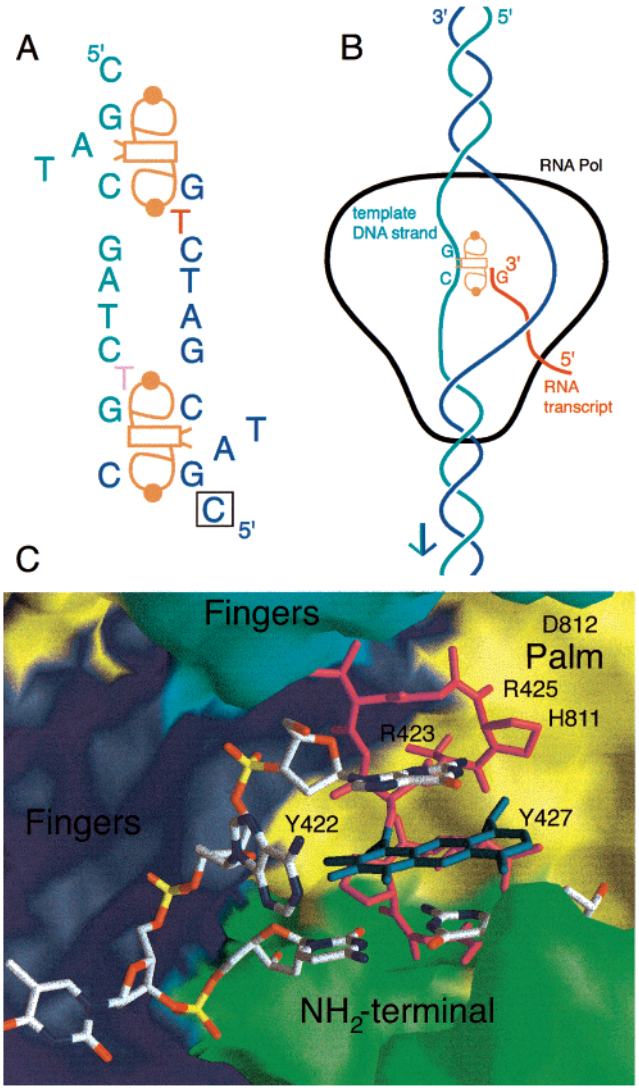

FIGURE 4: (A) Schematic diagram showing the binding of ActD to CGATCGATCG. (B) Schematic diagram showing the binding of ActD to the elongation locus during transcription. (C) A model showing the binding of ActD in the T7 RNA polymerase elongation site during transcription. The active site cavity at the elongation point is surrounded by the thumb (red), palm (green), and "fingers" (yellow) parts of the enzyme. The model was constructed by overlaying the phenoxazone ring in the present structure with the (C+2)-(G+2) base pair in the complex of Cheetham and Steitz (*24*). The two bulky peptapetide rings of ActD, located in the minor groove side of the heteroduplex, require that the base pairs move away from the floor of the T7 RNA polymerase by about 4 Å. The back side of the bound ActD is contacting the amino acids of Y639, M635, and W422 of the T7 RNA polymerase.

T104O4 (3.2 Å). Parenthetically, the two adjacent intercalation cavities observed in this structure here suggest that nearest-neighbor bis-intercalation is possible, but some backbone distortions (e.g., looped-out nucleotides) may be required.

The novel structure of the complex may be depicted schematically (Figure 4A), whose implication in the mechanism of transcription is immediately evident (Figure 4B). The ActD binding site at the end of the complex resembles strikingly the drug binding to the elongation locus during transcription. Model building studies based on the crystal structure of the complex between T7 RNA polymerase, the T7 promoter DNA, and a RNA trinucleotide transcript (*24*) provide clues to the inhibitory function of ActD (Figure 4C). In the model of Cheetham and Steitz (*24*) only three base pairs (+1 to +3) of the template strand DNA and the RNA transcript heteroduplex may be accommodated in the active site of the "close form" complex. The incoming 5′-NTP forms a base pair with the +4 nucleotide of the template DNA. It is easy to envision that ActD diffuses into the active site pocket when the enzyme returns to the "open form" and binds to a C (*N*th nucleotide of the template strand)·G (3′-end of the RNA transcript) base pair, similar to that seen in Figure 3A. The two pentapeptides point toward the bottom surface of the active site. If the template ($N + 1$)th nucleotide is a G, specific hydrogen bonds will form between ActD and ($N + 1$)G, resulting in a tight complex in the active site which prevents the enzyme to proceed for the elongation step. Figure 4C shows the model of this trapped complex induced by ActD binding. If the template ($N + 1$)th nucleotide is not a G, it may be extruded until a G is found, again not unlike the present crystal structure. There is sufficient room in the region next to the active site in the polymerase to accommodate the extruded template DNA nucleotides.

The minor groove binding of antitumor drugs in the active site of the enzyme−DNA−RNA complex prevents the drug from exiting easily, thus providing a more effective way for inhibition. This may be relevant in the observation that most antitumor drugs are minor groove binders.

In the structure of the DNA polymerase−DNA complex, the open cleft of the active site is filled by a continuous duplex DNA at least five base pairs long (reviewed in ref *25*). The intact duplex presents a less favorable binding site for ActD due to the wedging of methyl group from MetVal discussed above.

The difference in the ease with which ActD binds at the active site between RNA and DNA polymerases may explain the strong preference of ActD inhibition toward RNA polymerase. Our results here should be useful in the understanding of the transcription mechanism and in the design of new derivatives of ActD or related peptide antibiotics (e.g., triostin/echinomycin) (*26*).

SUPPORTING INFORMATION AVAILABLE

One table of the torsion angles of the DNA and two figures of MAD and refined electron density maps and the crystal packing diagram. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Hurwitz, J., Furth, J. J., Malamy, M., and Alexander, M. (1962) *Proc. Natl. Acad. Sci. U.S.A. 48*, 1222−1230.
2. Aivasashvilli, V. A., and Beabealashvilli, R. S. (1983) *FEBS Lett. 160*, 124−128.
3. Sobell, H. M., and Jain, S. C. (1972) *J. Mol. Biol. 68*, 1−20.
4. Kamitori, S., and Takusagawa, F. (1992) *J. Mol. Biol. 225*, 445−456.
5. Bailey, S. A., Graves, D. E., Rill, R. L., and Marsch, G. A. (1993) *Biochemistry 32*, 5881−5887.
6. Lian, C., Robinson, H., and Wang, A. H.-J. (1996) *J. Am. Chem. Soc. 118*, 8791−8801.
7. Chen, H., Liu, X., and Patel, D. J. (1996) *J. Mol. Biol. 258*, 457−479.

8. Wadkins, R. M., Vladu, B., and Tung, C.-S. (1998) *Biochemistry 37*, 11915−11923.
9. Wang, A. H.-J., and Gao, Y.-G. (1990) *Methods 1*, 91−99.
10. Walsh, M. A., Evans, G., Sanishvili, R., Dementieva, I., and Joachimiak, A. (1999) *Acta Crystallogr. D55*, 1726−1732.
11. Gao, Y.-G., Robinson, H., Sanishvili, R., Joachimiak, A., and Wang, A. H.-J. (1999) *Biochemistry 38*, 16452−16460.
12. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
13. Brünger, A. T., et al. (1998) *Acta Crystallogr. D54*, 905−921.
14. Brünger, A. T. (1992) *X-PLOR 3.1, A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.
15. Sheldrick, G. M. (1997) *SHELX-97, crystallographic refinement program*, University of Gottingen, Germany.
16. Parkinson, G., Vojtechovsky, J., Clowney, L., Brunger, A. T., and Berman, H. M. (1996) *Acta Crystallogr. 52*, 57−64.
17. Moews, P. C., and Kretsinger, R. H. (1975) *J. Mol. Biol. 91*, 201−225.
18. Ortiz-Lombardia, M., Gonzalez, A., Eritja, R., Aymami, J., Azorin, F., and Coll, M. (1999) *Nat. Struct. Biol. 6*, 913−917.
19. Yang, X.-L., Robinson, H., Gao, Y.-G., and Wang, A. H.-J. (2000) *Biochemistry 39*, 10950−10957.
20. Adams, A., Guss, J. M., Collyer, C. A., Denny, W. A., and Wakelin, L. P. (2000) *Nucleic Acids Res. 28*, 4244−4253.
21. Robinson, H., Gao, Y.-G., McCrary, B. S., Edmondson, S. P., Shriver, J. W., and Wang, A. H.-J. (1998) *Nature 392*, 202−205.
22. Murphy, F. V., 4th, Sweet, R. M., and Churchill, M. E. (1999) *EMBO J. 18*, 6610−6618.
23. Chen, F.-M. (1998) *Biochemistry 37*, 3955−3964.
24. Cheetham G. M., and Steitz, T. A. (1999) *Science 286*, 2305−2309.
25. Steitz, T. A. (1999) *J. Biol. Chem. 274*, 17395−17398.
26. Wang, A. H.-J., Ughetto, G., Quigley, G. J., Hakoshima, T., van der Marel, G. A., van Boom, J. H., and Rich, A. (1984) *Science 225*, 1115−1121.

BI002859Z