10. K. Gallagher and K. Lambeck, *Basin Res.* **2**, 115 (1989).
11. K. Gallagher, T. A. Dumitru, A. J. W. Gleadow, *ibid.* **6**, 77 (1994).
12. J. R. Cochran and M. Talwani, *Geophys. J. R. Astron. Soc.* **50**, 495 (1977).
13. F. Pribac, thesis, Australian National University, Canberra (1991).
14. S. R. Hart, *Nature* **309**, 753 (1984).
15. D. G. Pyle, D. M. Christie, J. J. Mahoney, R. A. Duncan, *J. Geophys. Res.* **100**, 22261 (1995).
16. D. G. Pyle, D. M. Christie, J. J. Mahoney, *Earth Planet. Sci. Lett.* **112**, 161 (1992); M. Rehkämper and A. W. Hofmann, *ibid.* **147**, 93 (1997).
17. E. M. Klein and C. H. Langmuir, *J. Geophys. Res.* **92**, 8089 (1987).
18. M. Tolstoy, A. J. Harding, J. A. Orcutt, J. Phipps Morgan, *Eos (Fall Suppl.)* **76**, 570 (1995).
19. Using the observed difference in crustal thickness (*18*) and a simple isostatic balance (ρ_c = 2900 kg/m^3 and ρ_m = 3250 kg/m^3), we predicted a bathymetric difference of 470 m, smaller than the observed ~750 m. Using a more sophisticated isostatic balance and observed major element chemistry, P. Lecroart, A. Cazenave, Y. Ricard, C. Thoraval, and D. G. Pyle [*Earth Planet. Sci. Lett.* **149**, 49 (1997)] have predicted a 400-m difference between observed and isostatic topography.
20. Subduction off eastern Australia was but a part of the larger Pacific-Gondwanaland converging margin [B. C. Storey, *Nature* **377**, 301 (1995)]. The early Mesozoic margin, which included New Caledonia, New Zealand, Lord Howe Rise, and Norfolk ridge, was typified by convergence and calc-alkaline volcanism that progressively migrated outboard as the style of the Australian margin shifted to rifting in the Cretaceous, as summarized by N. Williams and R. Korsch [in *Extended Abstracts 43* (Geological Society of Australia, Sydney, 1996), pp. 564–568].
21. The Whitsunday volcanic province represents the tail end of rifting in a back arc setting [S. Bryan, in *Extended Abstracts 43* (Geological Society of Australia, Sydney, 1996), pp. 124–133].
22. The reconstructions were based on the hot spot reference frame of R. D. Müller, J.-Y. Royer, and L. A. Lawver [*Geology* **21**, 275 (1993)].
23. The Gondwanaland converging margin did not undergo a substantial shift in paleolatitude from ~200 to ~120 Ma, as seen in paleogeographic reconstructions [L. A. Lawver and L. M. Gahagan, in *Flow and Creep in the Solar System*, D. B. Stone and S. K. Runcorn, Eds. (Kluwer Academic, Dordrecht, Netherlands, 1993), pp. 225–247].
24. L. Moresi and M. Gurnis, *Earth Planet. Sci. Lett.* **138**, 15 (1996)].
25. The continental outlines, isochrons, and referenced rotation poles of R. D. Müller, U. R. Roest, J.-Y. Royer, L. M. Gahagan, and J. G. Sclater [*J. Geophys. Res.* **102**, 3211 (1997)] were used and transformed to a Cartesian domain with an azimuthal Lambert equal area projection.
26. Our models showed that slabs, while in the lower mantle, only moved laterally by about 500 km, and consequently our focus on a segment of the margin confined to beneath Australia and the SEIR was justified.
27. Thermodynamic calculations and petrologic models estimate that oceanic crustal thickness increases 0.06 km for each 1°C increase in mantle temperature [C. H. Langmuir, E. M. Klein, T. Plank, in *Mantle Flow and Melt Generation*, J. Phipps Morgan, Ed. (American Geophysical Union, Washington, DC, 1992), pp. 183–280; P. Asimow, thesis, California Institute of Technology, Pasadena (1997)]. Crustal thickness is computed by sampling the temperature within a 200-km region beneath the ridge.
28. The nominal model had a Rayleigh number of 7.4 × 10^6 (based on upper mantle and transition zone viscosity of 10^{21} Pa·s), no temperature-dependent viscosity, and a 100-fold jump in viscosity at 660 km. The clapeyron slope and jump in density were 2.0 MPa/K and 5% at 410 km and −4.2 MPa/K and 8% at 660 km. The upper 100 km was assumed to be lithosphere, with a viscosity 100 times that of the upper mantle.
29. Geoid height variations are ignored because its amplitude is smaller than the dynamic topography, such that topography is the primary dynamic control on sea-level change [M. Gurnis, *Nature* **344**, 754 (1990); C. Lithgow-Bertelloni and M. Gurnis, *Geology* **25**, 735 (1997)].
30. A. B. Watts and M. S. Steckler, in *Deep Drilling Results in the Atlantic Ocean: Continental Margins and Paleoenvironments*, M. Talwani, W. Hays, W. B. F. Ryan, Eds. (American Geophysical Union, Washington, DC, 1979), pp. 218–234.
31. L. L. Sloss, in *Sedimentary Cover—North American Craton: U.S.*, L. L. Sloss, Ed. (Geological Society of America, Boulder, CO, 1988), pp. 25–51.
32. For temperature-dependent viscosity, an Arrhenius form with an activation energy of 400 kJ/mol with a 100-fold truncation in lateral variation was used.
33. U. Christensen, *Geophys. J. R. Astron. Soc.* **77**, 343 (1984).
34. R. van der Hilst and T. Seno, *Earth Planet. Sci. Lett.* **120**, 395 (1993).
35. R. van der Hilst, *Nature* **374**, 154 (1995); S. Zhong and M. Gurnis, *Earth Interact.* [online] **1** (no. 6) (1997), available at http://earthinteractions.org.
36. BMR Paleogeographic Group, *Cretaceous Paleogeographic Maps Record 1990/30 Paleogeographic Series 14* (Bureau of Mineral Resources, Canberra, Australia, 1990).
37. This report represents contribution 8489 of the Division of Geological and Planetary Sciences, California Institute of Technology. Some of the work reported here was conducted as part of the Australian Geodynamics Cooperative Research Centre (AGCRC) and is published with the consent of the Director, AGCRC. We thank J. Veevers for taking the time to discuss the geologic constraints on Mesozoic subduction beneath Australia and K. Gallagher and J. Veevers for helpful comments on the manuscript.

3 November 1997; accepted 28 January 1998

Crystal Structures of Human Topoisomerase I in Covalent and Noncovalent Complexes with DNA

Matthew R. Redinbo,* Lance Stewart,*† Peter Kuhn, James J. Champoux, Wim G. J. Hol‡

Topoisomerases I promote the relaxation of DNA superhelical tension by introducing a transient single-stranded break in duplex DNA and are vital for the processes of replication, transcription, and recombination. The crystal structures at 2.1 and 2.5 angstrom resolution of reconstituted human topoisomerase I comprising the core and carboxyl-terminal domains in covalent and noncovalent complexes with 22–base pair DNA duplexes reveal an enzyme that "clamps" around essentially B-form DNA. The core domain and the first eight residues of the carboxyl-terminal domain of the enzyme, including the active-site nucleophile tyrosine-723, share significant structural similarity with the bacteriophage family of DNA integrases. A binding mode for the anticancer drug camptothecin is proposed on the basis of chemical and biochemical information combined with these three-dimensional structures of topoisomerase I–DNA complexes.

Topoisomerases are ubiquitous enzymes that solve topological problems generated by key nuclear processes such as DNA replication, transcription, recombination, repair, chromatin assembly, and chromosome segregation (*1*). There are two types of topoisomerases. Type I enzymes are monomeric and transiently break one strand of duplex DNA, allowing for single-step changes in the linking number of circular DNAs (the number of times one strand of DNA crosses the other). Type II enzymes are dimeric and break both strands of a duplex to generate a gate through which another region of DNA can be passed, resulting in linking number changes in steps of two. Type I and type II enzymes are fundamentally different in both mechanism and cellular function (*2–4*). The medical importance of these enzymes is underscored by the fact that they are the specific targets of many promising anticancer drugs (*5*).

Eukaryotic type I topoisomerases differ significantly from their prokaryotic counterparts. Prokaryotic enzymes require magnesium and a single-stranded segment of DNA; additionally, they form a covalent intermediate with the 5′ end of the broken DNA strand and relax only negatively supercoiled DNA. Eukaryotic topoisomerase I enzymes, in contrast, do not require any metal cofactor, need no single-stranded stretch of DNA, form a covalent intermediate with the 3′ end of the broken strand, and are able to relax both positive and negative supercoils. All available evidence indicates that prokaryotic and eukaryotic type I enzymes share no sequence or structural similarity. Vaccinia virus topoisomerase (*6*), however, shares several mechanistic and sequence features with eukaryotic type I topoisomerases but is much smaller (36 kD) than all known cellular topoisomerases I (*7*), which range from 80 to 110 kD (*8*). So far, no structure of a eukaryotic topo-

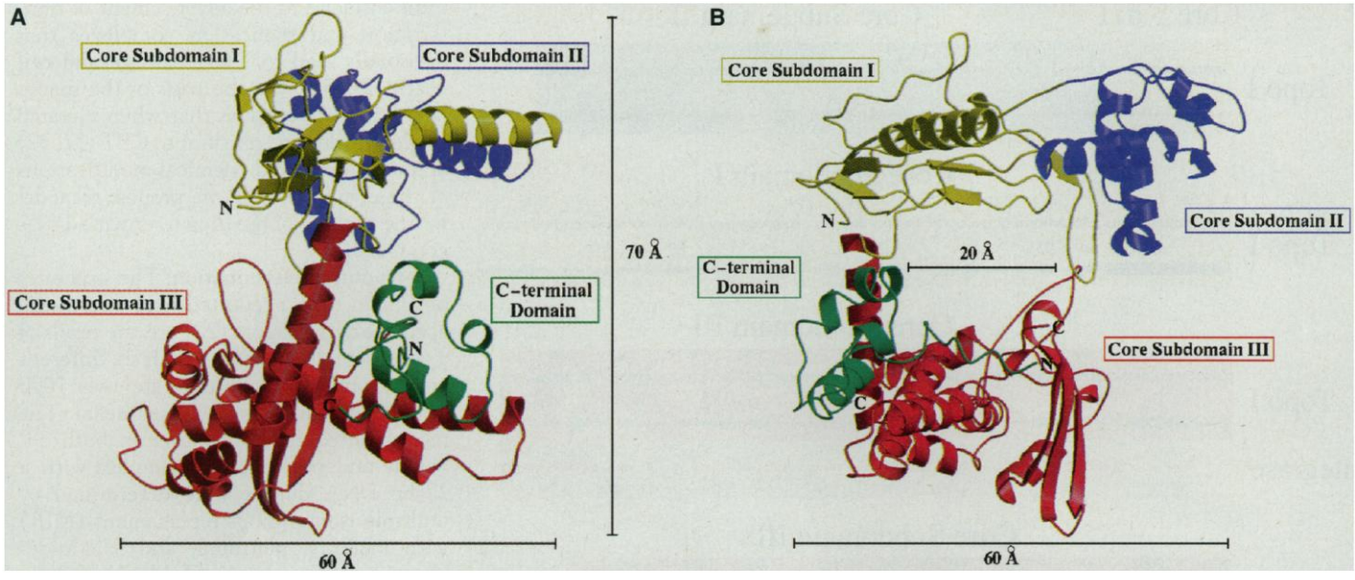


Fig. 1. The structure of reconstituted human topoisomerase I. (**A**) Viewed perpendicular to the pore. Core subdomains I, II, and III are shown in yellow, blue, and red, respectively, and the COOH-terminal domain is shown in green. The enzyme consists of a "cap" consisting of core subdomains I and II (with the extended nose-cone helices facing to the right), and the closely associated complex of core subdomain III and the COOH-terminal domain at the bottom. (**B**) A view of the enzyme down the pore, rotated 90° about the vertical axis relative to (A), and with the domains colored as indicated above. The central pore of the molecule varies from 15 to 20 Å in diameter and provides an extensive, highly positively charged binding region for the substrate DNA. The overall dimensions of the reconstituted enzyme are 70 Å by 60 Å by 60 Å. The cap of the molecule and core subdomain III contact one another at two lips (residues 367 to 369 of subdomain I and residues 497 to 499 of subdomain III), as shown on the right central part of this figure.

isomerase I containing all elements essential for catalysis has been reported (*9*).

Human topoisomerase I (topo I) is a monomeric protein of 765 amino acids (*10*). Limited proteolysis studies of human topo I (*11*), and the crystallographic structure determination of an inactive fragment of the yeast enzyme (*2*), indicate that eukaryotic topo I consists of four major regions: the NH_2-terminal, core, linker, and COOH-terminal domains. The ~210-residue NH_2-terminal region of human topo I is highly charged, contains very few hydrophobic amino acids, and is largely disordered (*12*). This domain contains several nuclear-targeting signals (*13*) and has been implicated in nucleolar localization through interactions with nucleolin (*14*). The COOH-terminal domain, comprising residues ~713 to 765, contains the essential catalytic Tyr^{723}, which forms a phosphoester bond with the 3′ phosphate at the site of cleavage of the scissile strand of the substrate DNA (*15*). Purified COOH-terminal domain can be combined with isolated core domain composed of residues ~200 to ~635 to reconstitute near full enzyme activity (*16*). Thus, residues ~636 to 712, which form the so-called linker domain, contribute to but are not required for activity (*16*). Hence, human topo I is a multidomain enzyme that contains two highly conserved globular domains (the core and the COOH-terminal domain) that are crucial for catalytic activity, and two other regions (NH_2-terminal and linker) that are not strictly required for its catalytic and relaxation functions.

The "reconstituted topo I" used in these studies was obtained by mixing the purified 58-kD core domain (consisting of residues 175 to 659, expressed in the baculovirus-insect cell system) and the 6.3 kD COOH-terminal domain (residues 713 to 765, expressed in *Escherichia coli*) to reconstitute a tightly associated 1:1 complex (*16*). The reconstituted enzyme exhibits the same specific activity as either the NH_2 terminally truncated enzyme (residues 175 to 765) or the full-length protein when assayed under physiological salt conditions (*16*, *17*).

Although human topo I does not exhibit a high degree of sequence specificity for its target cleavage sites (*18*), a sequence from the ribosomal DNA of *Tetrahymena thermophilus* has been described by Westergaard and his co-workers that is a high-affinity binding and cleavage site for all eukaryotic type I enzymes (*19*). Here we describe high-resolution structures of two variants of the reconstituted enzyme, each in association with a slightly different 22–base pair (bp) version of the *Tetrahymena* DNA sequence. The first is the structure of the $Tyr^{723} \rightarrow Phe$ inactive mutant in a noncovalent complex at 2.5 Å resolution, and the second is the covalent complex at 2.1 Å resolution. Together with the structure of the 70-kD NH_2 terminally truncated form of human topo I in complex with DNA described in the accompanying report (*20*), these complexes provide detailed insights into the mode of DNA binding and DNA manipulation by this key human enzyme.

The structure of human topo I is of considerable biomedical importance because this enzyme is the sole target of the camptothecin (CPT) family of anticancer drugs (*5*, *21*–*23*). Two analogs of CPT, topotecan and irinotecan, have been used successfully in the treatment of several human cancers (*5*, *24*, *25*). CPT and its derivatives act by prolonging the lifetime of the nicked intermediate in the topo I reaction (*22*, *26*–*28*). The resulting stabilized covalent intermediates are presumed to

M. R. Redinbo and L. Stewart are at the Biomolecular Structure Center and Department of Biological Structure, Box 357742, School of Medicine, University of Washington, Seattle, WA 98195, USA. P. Kuhn is at Stanford Synchrotron Radiation Laboratory, Stanford University, Stanford, CA 94309, USA. J. J. Champoux is in the Department of Microbiology, Box 357242, School of Medicine, University of Washington, Seattle, WA 98915, USA. W. G. J. Hol is in the Departments of Biological Structure and Biochemistry, Howard Hughes Medical Institute and Biomolecular Structure Center, Box 357742, School of Medicine, University of Washington, Seattle, WA 98195, USA.

*These authors contributed equally to this work.
†Present address: Emerald BioStructures, 7865 Northeast Day Road West, Bainbridge Island, WA 98110, USA. E-mail: emerald_biostructures@rocketmail.com
‡To whom correspondence should be addressed at Biomolecular Structure Center, School of Medicine, University of Washington, Seattle, WA 98195, USA. E-mail: hol@gouda.bmsc.washington.edu

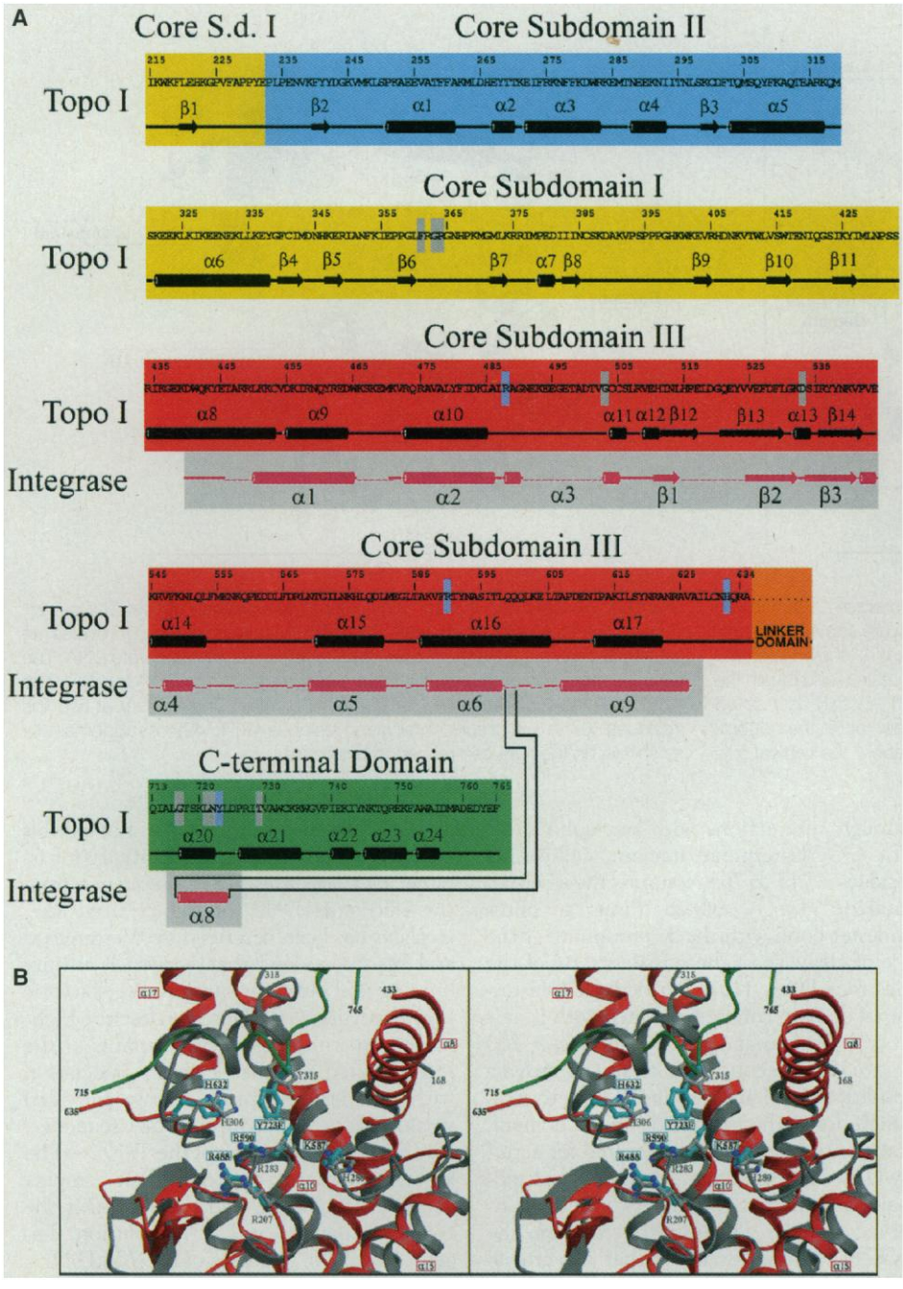


Fig. 2. Structural similarities between human topo I and HP1 integrase. (**A**) The sequence and secondary structural elements of reconstituted human topo I are indicated in the standard coloring scheme of the domain architecture of the enzyme (Fig. 1A), and the structurally similar regions of HP1 integrase are shown in red with gray background. Catalytically relevant residues of human topo I are highlighted in cyan, and the positions of known CPT-resistant mutations in human, hamster, and yeast topoisomerases I are shown in gray. α-Helices 18 and 19 are not depicted because these correspond to the linker domain (*20*), which is not present in the reconstituted enzyme. (**B**) Stereoview of the superposition of core subdomain III (red) and the COOH-terminal domain (green) of human topo I and bacteriophage HP1 integrase (gray) (*38*). The active-site residues of each enzyme are shown, with the human topo I residues in cyan and the integrase residues in gray. Helices 8, 10, 15, and 17 of core subdomain III of topo I are also indicated. There is no structural equivalent in the integrase for the topo I COOH-terminal domain past the first eight residues, which contain the catalytic Tyr^{723}. The Cα positions of the active-site residues Arg^{488} and Arg^{590} of topo I (*20*) superimpose within 0.6 and 1.9 Å, respectively, of the Cα positions of Arg^{207} and Arg^{283} in the integrase. His^{306} of the integrase superimposes within 3.3 Å on His^{632} of human topo I, but the putative catalytic His^{280} of the integrase superimposes on a noncatalytic residue of human topo I, Lys^{587}. Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

form obstacles to the advancement of transcription and replication complexes that eventually lead to DNA damage and cell death (*29*, *30*). On the basis of the spatial distribution of residues that when mutated render the enzyme resistant to CPT (*31*, *32*) and on the allowed chemical modifications of the compound (*5*), we propose a model for the binding of the drug to enzyme-DNA complexes.

Structure determination. The structures of human topo I reported here and in the accompanying report (*20*) are the result of a considerable effort in which six different crystal forms were produced and over 1000 crystals examined in the x-ray beam. The reconstituted (*33*) structures in both covalent and noncovalent complex with a 22-bp DNA duplex were determined by multiple isomorphous replacement (MIR) with mercury, platinum, and 5-iodo-deoxyuridine derivatives (Table 1). In the structure of the noncovalent complex with DNA (crystal form 2), the protein was inactivated by replacement of Tyr^{723} with Phe. The covalent complex (crystal form 6) was generated by introducing a 5′-bridging phosphorothioate group at the DNA cleavage site, which effectively traps the enzyme covalently attached to the DNA via a phosphoester bond to the active-site tyrosine (*34*). Nonisomorphism, large mosaic spread, and poor crystal quality made the structure determination extremely difficult, and only the use of precise crystal growth and cryo-cooling protocols, synchrotron radiation, and phasing with the program SHARP (*35*) allowed the structures to be elucidated (Table 1). One crystal of the covalent complex (form 6) diffracted to 2.1 Å resolution, and the structure of this crystal form was solved by MIR. Repositioning of the form 6 model into the nearly isomorphous form 2 cell allowed the structure of the noncovalent complex to be determined to 2.5 Å resolution. In both structures, the protein was well defined for almost the entire length of the polypeptide chain, and the DNA oligonucleotide was clearly defined over its entire length.

The human topo I protein. Reconstituted human topo I is composed of four domains that temporarily clamp around the DNA. In its closed form it contains a central DNA-binding pore of 15 to 20 Å in diameter (Fig. 1, A and B). Subdomains I, II, and III constitute the 56-kD core domain of the molecule, which in turn associates closely with the COOH-terminal domain to form the complete reconstituted enzyme. Subdomain I consists of residues 215 to 232 and 320 to 433 and is made up of two α helices and nine β strands (Fig. 2A). Subdomain II runs from

residue 233 to 319 and is composed of five α helices and two β strands. Subdomains I and II fold tightly together, burying ~1000 Å^2 of solvent-accessible surface area on both subdomains, and form the top half or "cap" of the enzyme (Fig. 1, A and B). These domains are further characterized by two long "nose-cone" helices (α5 and α6) that come together in a "V" at a point 25 Å away from the body of the molecule. Subdomains I and II are similar in structure to the 26-kD yeast topo I fragment (*2*), with a root-mean-square (rms) deviation of 1.7 Å for 212 Cα positions and 53% sequence identity.

Subdomain III runs from residue 434 to 635 and is a complex arrangement of 10 α helices and 3 β strands. This subdomain contains all the active-site residues, with the exception of the catalytic Tyr^{723} (Fig. 2A). Subdomain III extends from the top half of the molecule downward by way of two long α helices (8 and 9) that most likely function as the hinge that opens and closes the enzyme around the DNA (Fig. 1B). Subdomains I and III interact via two short "lips" opposite from these long-hinge helices. This is the only point of contact between the top and bottom halves of the molecule on that side of the clamp, and involves three residues from each subdomain and one salt bridge. The core of the enzyme interacts closely with the COOH-terminal domain primarily by way of α helices 8, 16, and 17 of subdomain III (Fig. 1A). This largely hydrophobic interface buries ~1300 Å^2 of solvent-accessible surface on each domain and is made up of 11 hydrophobic residues, as well as five specific hydrogen bonds and three salt bridges. The compact COOH-terminal domain runs from residues 713 to 765, is composed of five short α helices, and contains the active-site Tyr^{723} (Fig. 2A).

The three subdomains of the core and the COOH-terminal domain form a 64.3-kD particle with overall dimensions of 70 Å by 60 Å by 60 Å. The close integration between the four domains that is observed in the structure even in the absence of the 77-residue linker domain is in agreement with the observed stability of the reconstituted enzyme (*16*). The central pore of the molecule is composed largely of positively charged residues, 15 lysines and 8 arginines, which give rise to a high positive electrostatic potential surrounding the pore. The catalytic residues of the enzyme, including Tyr^{723}, are contained within this highly positively charged channel region.

The structures of the reconstituted human topo I proteins in our two complexes are nearly identical, with an rms deviation of 0.45 Å on Cα positions. The only significant differences between the structures occur in the active-site region (as discussed below); in addition, residues 628 to 635 of subdomain III and 715 to 719 of the COOH-terminal domain are ordered in the noncovalent complex but are disordered in the covalent complex.

A search for structural similarity with other proteins with the program DALI (*36*) indicated that subdomain II of topo I was similar in structure to the homeodomain family of DNA-binding proteins. For example, residues 244 to 314 of human topo I superimpose on the POU homeodomain of the human Oct-1 transcription factor (*37*) with an rms deviation of 3.0 Å over 55 Cα positions, but show only 11% sequence identity. However, subdomain II of human topo I does not interact with the substrate DNA in either of the protein-DNA complexes we have solved, so any possible DNA binding function within this region of the enzyme still awaits further investigation.

A second structural similarity was identified by DALI between subdomain III of human topo I and the bacteriophage HP1 integrase (*38*). Residues 440 to 614 of topo I superimpose on residues 168 to 330 of integrase with an rms deviation of 3.8 Å for 129 Cα atoms with a sequence identity of 12% (Fig. 2, A and B). Although it was known that eukaryotic type I topoisomerases and several bacteriophage integrases share similar biochemical properties (*39*), a close structural similarity between the enzymes has not been described. Not only do 90% of the residues in subdomain III correspond with HP1 integrase, but also residues Gly^{717} to Leu^{724} of the COOH-terminal domain of topo I appear to correspond with residues 309 to 316 of integrase after subdomain III of topo I and integrase are superimposed (Fig. 2B). Most importantly, despite short insertions in each sequence, the active sites of the two enzymes line up remarkably well. The Cα position of the catalytic Tyr^{723} of human topo I falls within 3.9 Å of the corresponding atom in Tyr^{315} of the integrase after the two structures are superimposed. We observed that the absolutely conserved and catalytically important His^{632} of topo I corresponds to His^{306} of the integrase (*20*). The putative catalytic His^{280} of HP1 integrase superimposes on Lys^{587} of human topo I, which interacts with a phosphate away from the active site and is likely not a catalytic residue.

The DNA oligonucleotide structure. The 22-bp DNA oligonucleotides were fully visible in the electron density maps of both the covalent and the noncovalent complexes. In both crystal forms, the blunt-ended duplexes stack head-to-tail to form a pseudo-continuous helix along the crystallographic *c* axis (Fig. 3). This stacking gives rise to the observed 72 Å length for the *c* axis in both crystal forms, which is shorter than the ~75 Å expected for a 22-bp stretch of perfectly canonical B-DNA. Superpositions of the DNA duplexes with B-DNA reveal that the inner 10 bp (−4 to +6) deviate little from canonical B-form (with an rms deviation of 1.2 Å for 460 atoms), whereas the 6 bp at either end of the central region show more deviations from ideal B-DNA (rms deviation of 2.0 Å for 486 atoms). These re-

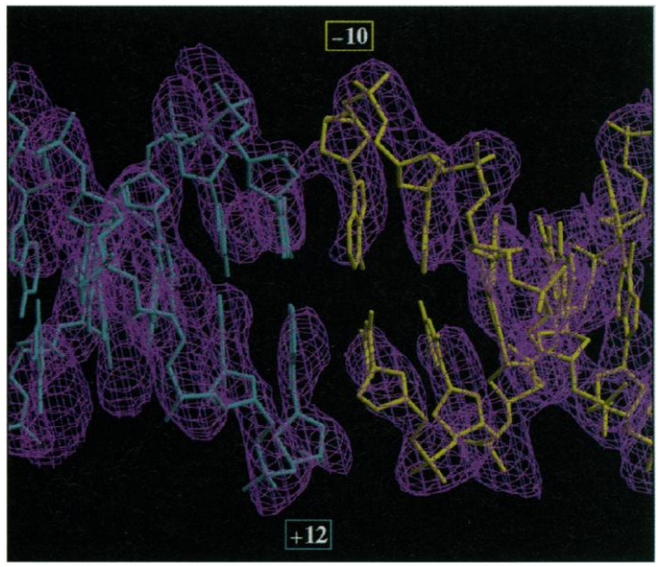


Fig. 3. Experimental electron density map of symmetry-related DNA duplexes. Representative electron density from the experimental MIR with anomalous scattering map at 3.2 Å resolution is shown in purple (contoured at 1.2σ). The tail-to-head stacking arrangement of symmetry-related DNA oligonucleotides along the monoclinic *c* axis is evident (+8 to +12 region of DNA oligonucleotide downstream of the cleavage site in green; −10 to −6 region of symmetry-related DNA oligonucleotide upstream of the cleavage site in yellow). The entire DNA oligonucleotide, as well as the majority of the protein atoms, could be built into this experimental map.

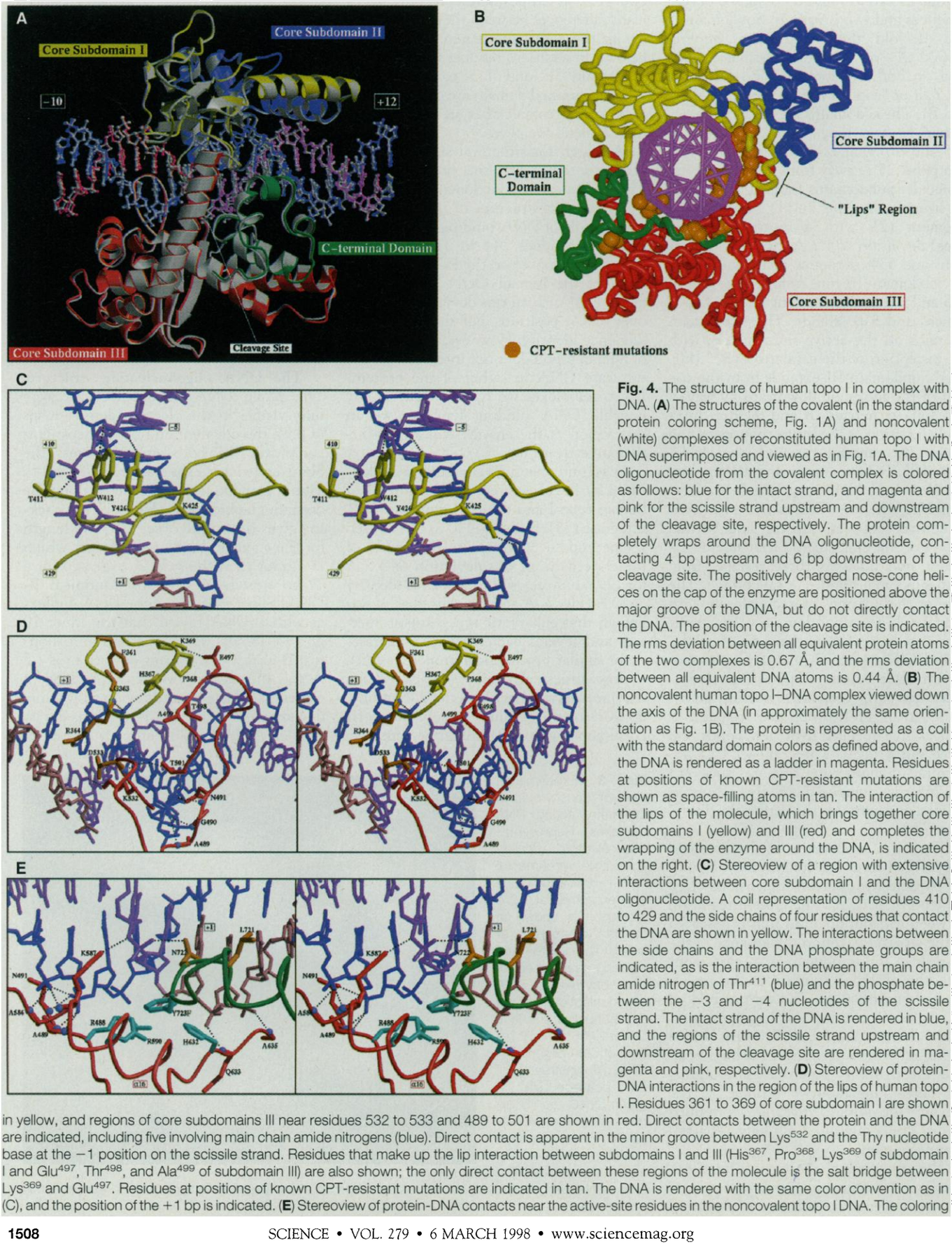
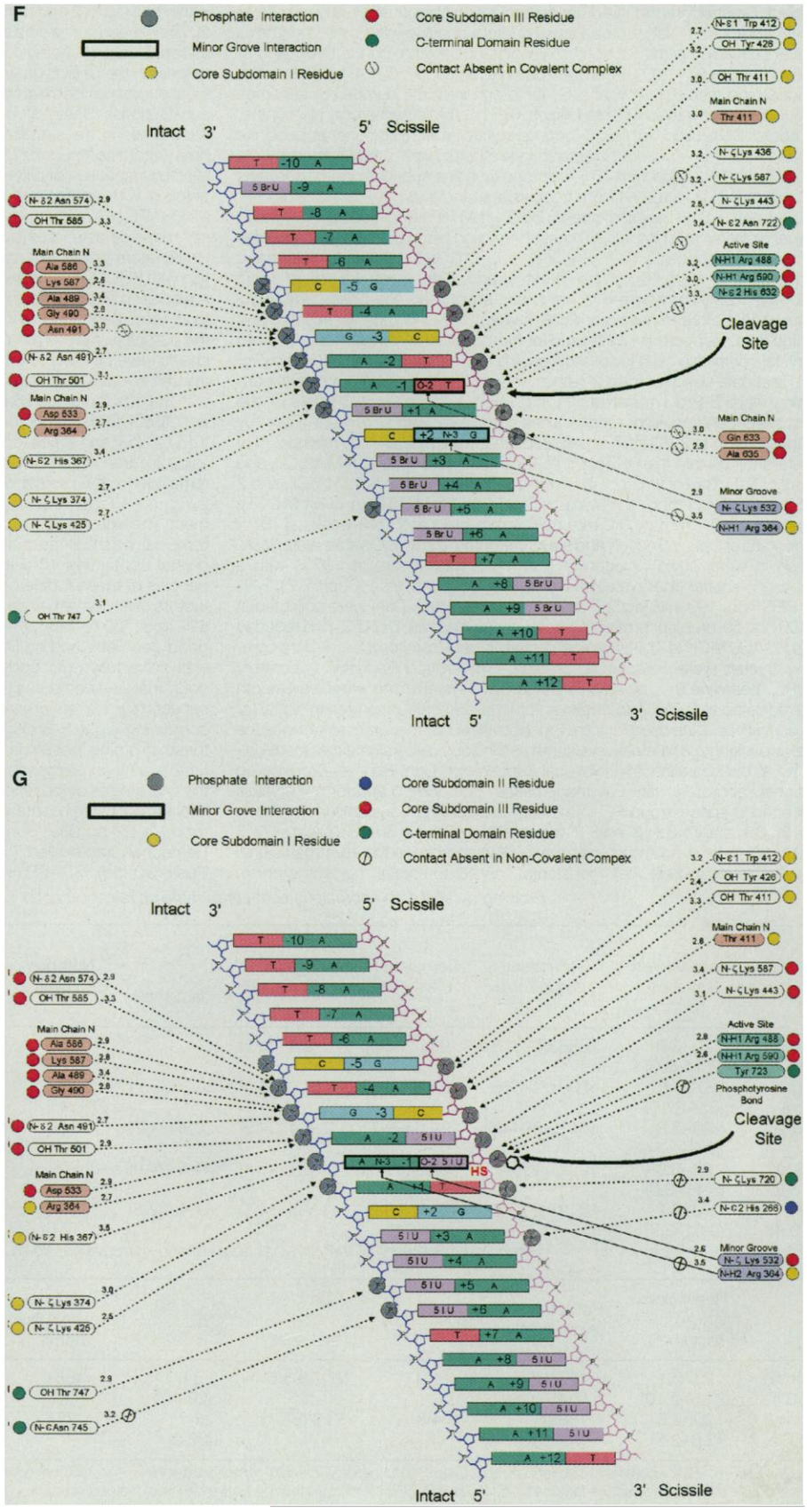


Fig. 4. The structure of human topo I in complex with DNA. (**A**) The structures of the covalent (in the standard protein coloring scheme, Fig. 1A) and noncovalent (white) complexes of reconstituted human topo I with DNA superimposed and viewed as in Fig. 1A. The DNA oligonucleotide from the covalent complex is colored as follows: blue for the intact strand, and magenta and pink for the scissile strand upstream and downstream of the cleavage site, respectively. The protein completely wraps around the DNA oligonucleotide, contacting 4 bp upstream and 6 bp downstream of the cleavage site. The positively charged nose-cone helices on the cap of the enzyme are positioned above the major groove of the DNA, but do not directly contact the DNA. The position of the cleavage site is indicated. The rms deviation between all equivalent protein atoms of the two complexes is 0.67 Å, and the rms deviation between all equivalent DNA atoms is 0.44 Å. (**B**) The noncovalent human topo I–DNA complex viewed down the axis of the DNA (in approximately the same orientation as Fig. 1B). The protein is represented as a coil with the standard domain colors as defined above, and the DNA is rendered as a ladder in magenta. Residues at positions of known CPT-resistant mutations are shown as space-filling atoms in tan. The interaction of the lips of the molecule, which brings together core subdomains I (yellow) and III (red) and completes the wrapping of the enzyme around the DNA, is indicated on the right. (**C**) Stereoview of a region with extensive interactions between core subdomain I and the DNA oligonucleotide. A coil representation of residues 410 to 429 and the side chains of four residues that contact the DNA are shown in yellow. The interactions between the side chains and the DNA phosphate groups are indicated, as is the interaction between the main chain amide nitrogen of Thr^{411} (blue) and the phosphate between the −3 and −4 nucleotides of the scissile strand. The intact strand of the DNA is rendered in blue, and the regions of the scissile strand upstream and downstream of the cleavage site are rendered in magenta and pink, respectively. (**D**) Stereoview of protein-DNA interactions in the region of the lips of human topo I. Residues 361 to 369 of core subdomain I are shown in yellow, and regions of core subdomains III near residues 532 to 533 and 489 to 501 are shown in red. Direct contacts between the protein and the DNA are indicated, including five involving main chain amide nitrogens (blue). Direct contact is apparent in the minor groove between Lys^{532} and the Thy nucleotide base at the −1 position on the scissile strand. Residues that make up the lip interaction between subdomains I and III (His^{367}, Pro^{368}, Lys^{369} of subdomain I and Glu^{497}, Thr^{498}, and Ala^{499} of subdomain III) are also shown; the only direct contact between these regions of the molecule is the salt bridge between Lys^{369} and Glu^{497}. Residues at positions of known CPT-resistant mutations are indicated in tan. The DNA is rendered with the same color convention as in (C), and the position of the +1 bp is indicated. (**E**) Stereoview of protein-DNA contacts near the active-site residues in the noncovalent topo I DNA. The coloring

gions of the oligonucleotide are more compacted and slightly underwound, probably giving rise to the shorter than canonical length along the crystallographic *c* axis. However, because these regions of the DNA oligonucleotide are not in contact with the protein (as described below), it seems likely that these observed deviations from ideal B-form DNA are crystal-packing effects rather than specific changes induced by the protein.

Human topo I in complex with DNA. Human topo I engages in an intimate interaction with its DNA substrate, wrapping completely around the DNA and burying a total of ~4700 Å^2 of solvent-accessible surface area (~2400 Å^2 of the protein and ~2300 Å^2 of the DNA) upon complex formation (Fig. 4, A and B). All three subdomains (I, II, and III) of the core plus the COOH-terminal domain interact with DNA, burying approximately 950, 80, 1050, and 330 Å^2 of solvent-accessible surface area of each of these protein elements, respectively. Clearly, DNA interactions involving subdomain II



scheme is identical to that of Fig. 4, C and D, with the addition of the active-site residues in cyan and the coil in green to represent the COOH-terminal domain. The tight network of protein-DNA phosphate interactions, mostly involving main chain amide nitrogens, is apparent just across the minor groove from the active site. These interactions are further stabilized by a helix dipole from the 20-residue α helix 16 to the phosphate group between the −4 and −5 nucleotides of the intact strand. (**F**) Schematic representation of the protein-DNA interactions in the noncovalent complex of human topo I with DNA. Interactions between protein side chain and main chain atoms with DNA phosphate groups and minor-groove atoms closer than 3.5 Å are indicated. The protein-DNA contacts are limited almost exclusively to protein-phosphate interactions, and the protein only contacts the central 10 bp around the cleavage site. The side chains of Lys^{436}, His^{632}, and Asn^{722} and the main chain amide nitrogens of Asn^{491}, Gln^{633}, and Ala^{635} make contacts closer than 3.5 Å in this complex but not in the covalent complex (Fig. 4G; note that residues 628 to 635 are disordered and are not present in the structure of the covalent complex). In addition, the side chain of Arg^{364} makes a different contact in this complex relative to the covalent complex (Fig. 4G). The sugar-phosphate backbone of the oligonucleotide is colored as described in (A). (**G**) Schematic representation of the protein-DNA interactions in the covalent complex of human topo I with DNA. The coloring and numbering scheme are the same as in (F). The side chains of His^{266}, Lys^{720}, and Asn^{745} make contacts closer than 3.5 Å in this complex but not in the noncovalent complex (Fig. 4F). The two schematics in (F) and (G) provide exact information about the positions of the modified and unmodified bases used in the two reconstituted complexes. None of the substituted iodine or bromine atoms is in contact with the protein.

Table 1. Crystallographic structure determination of reconstituted human topo I–DNA complexes. Native data for heavy-atom phasing were collected from a crystal containing eight 5-iodo-deoxyuridine nucleotides on the DNA oligo at positions +8, +9, +10, +11 on the scissile strand and +3, +4, +5, +6 on the intact strand (see Fig. 4, F and G, for oligonucleotide numbering scheme). Heavy-atom derivatives were as follows: Hg-1: 1 mM ethylmercury phosphate, 14 hours; DNA oligo with six 5-iodo-deoxyuridine substitutions, at positions −2, −1, +8, +9, +10, +11 on the scissile strand; Hg-2: 1 mM thimerosal, 24 hours; Iodo: DNA oligo with 15 5-iodo-deoxyuridine substitutions, at positions −2, −1, +8, +9, +10, +11 on the scissile strand and −9, −8, −7, −6, −4, +3, +4, +5, +6 on the intact strand; Pt: 0.1 mM PIP [di-μ-iodobis (ethylenediamine) diplatinum (II) nitrate], 28 hours; DNA oligo with 10 5-iodo-deoxyuridine substitutions, at positions −2, −1, +8, +9, +10, +11 on the scissile strand and +3, +4, +5, +6 on the intact strand. The mean figures of merit [$\langle|\Sigma P(\alpha)e^{i\alpha}/\Sigma P(\alpha)|\rangle$, where α is the phase and $P(\alpha)$ is the phase probability distribution] before density modification were 0.57 and 0.60 for acentric and centric data, respectively. Data used for form 6 refinement were collected from a crystal containing 10 5-iodo-deoxyuridine nucleotides on the DNA oligo (Fig. 4G), and data used for form 2 refinement were collected from a crystal containing eight 5-BrdU substitutions (Fig. 4F). Crystallization: Crystals of the noncovalent and covalent reconstituted complexes with DNA were grown by vapor diffusion at 22°C from sitting drops that were prepared by mixing 3 μl of crystallant [100 mM tris-HCl (pH 7.7), 24% v/v PEG 400, 100 mM $MgCl_2$, and 10 mM DTT], 3 μl of water, 1 μl of duplex oligo (0.1 mM in 6 mM NaCl), and 2 μl of protein [5 mg/ml in 10 mM tris-HCl (pH 7.5), 1 mM EDTA, and 5 mM DTT]. Crystals of ~0.1 mm by 0.2 mm by 0.4 mm grew in 2 to 3 weeks at room temperature. Data collection and structure determination: Crystals were flash-frozen in either liquid nitrogen or a nitrogen gas stream at 100 K with a stabilization buffer {the crystallant without DTT [100 mM tris-HCl (pH 7.7), 24% v/v PEG 400, 100 mM $MgCl_2$]} as the cryo-protectant. Data were collected at ~100 K at several synchrotron sources, processed with DENZO, and reduced with SCALEPACK (*47*). Native and derivative data from form 6 (covalent complex) crystals were collected at Stanford Synchrotron Radiation Laboratory (SSRL), beamline 9-1, by a MAR image plate system. The x-ray wavelength was adjusted to 0.980 Å to obtain an optimal anomalous scattering signal for Hg derivatives. Data used for form 2 (noncovalent complex) refinement was the peak wavelength of a multiwavelength anomalous dispersion experiment collected at the beamline X4A (National Synchrotron Light Source, Brookhaven National Laboratory) with Fuji image plates. Both form 6 and form 2 crystals belonged to space group $P2_1$ with nearly identical cell constants (a = 72.0 Å, b = 66.6 Å, c = 71.8 Å, β = 98.3° for form 6; a = 71.8 Å, b = 66.3 Å, c = 71.8 Å, β = 98.4° for form 2) and one protein-DNA complex in the asymmetric unit. Despite isomorphous cell constants, crystal-to-crystal nonisomorphism within both form 2 and form 6 crystals was revealed by high R_{iso} values (15 to 20%) between native data sets. Eventually, out of nearly 80 data sets collected, a form 6 data set from a crystal containing eight 5-iodo-deoxyuridine substitutions on the DNA oligo was chosen as the native, and four form 6 data sets were chosen as isomorphous derivatives (two Hg, one Pt, and one with only I substitutions). Heavy-atom derivatives were prepared by washing crystals three times in the stabilization buffer to remove DTT, followed by soaks in heavy-atom compounds dissolved in this buffer. Heavy-atom sites were identified by inspection of difference Pattersons. Initial heavy-atom refinement and phasing with anomalous signal were carried out at 3.2 Å resolution with MLPHARE (*48*), with phase improvement by DM (*49*). Resultant electron density maps revealed some structural features but were not clearly traceable. Heavy-atom refinement and phasing with SHARP (*35*) and density modification with SOLOMON (*50*) yielded a readily traceable electron density map at 3.2 Å resolution (Fig. 3). The improvement gained with SHARP may have been due to its refinement of nonisomorphism parameters. All 22 bp of the DNA, nearly the complete main chain of the protein, and more than half the protein side chains were built with this MIRAS map and O (*51*). The covalent protein-DNA attachment was also evident in this initial MIRAS map. The map was consistent with the position of the 26-kD yeast topo I fragment (*2*) in the form 6 cell found with AMORE (*52*). Structural refinement: The form 6 model was refined with X-PLOR (*53*) and simulated annealing (*54*). Seven percent of the data were set aside for free-R cross-validation (*55*) before any structural refinement. The remainder of the protein was built and improved with SigmaA-weighted (*56*) $2|F_{obs}| - |F_{calc}|$ and $|F_{obs}| - |F_{calc}|$ maps iteratively with X-PLOR refinement. The high-resolution (2.1 Å) data set used for the refinement of the form 6 model was collected at SSRL beamline 9-1 and was obtained from a crystal containing 10 5-iodo-deoxyuridine substitutions (Fig. 4G). Occupancies of these iodine positions were refined alternately with individual B factors until convergence. The final form 6 model contains residues 215 to 627 and 720 to 765 of human topo I, 22 DNA bp, and 360 waters, with good geometry and no Ramachandran outliers (*57*). The form 2 structure was solved by rigid body adjustment of the form 6 model, followed by coordinate refinement with X-PLOR and model rebuilding with O. The data set used for the refinement of the form 2 model was obtained from a crystal containing eight 5-BrdU substitutions (Fig. 4F), and the occupancies of these bromine positions were also refined alternately with individual B factors until convergence. A bulk solvent correction (X-PLOR) was used in the final stages of the refinement of both the form 2 and form 6 models. The final form 2 model contains residues 215 to 635 and 715 to 765 of human topo I, a 22-bp DNA duplex, and 63 waters, with good geometry and no Ramachandran outliers (*57*). Figures were created with MOLSCRIPT (*58*), Raster3D (*59*), GRASP (*60*), and ALSCRIPT (*61*).

Phasing by MIR with anomalous scattering of the covalent complex (crystal form 6)

	Resolution (Å) (last shell)	Reflections (unique/total)	Complete (%) (last shell)	R_{sym}* (%) (last shell)	R_{iso}† (%) (last shell)	Number of sites	*PhP*‡		R_{cullis}§	
							Acentric (iso/anom)	Centric	Acentric (iso/anom)	Centric
Native (8 I)	20–3.0 (3.11–3.0)	12863/27375	92.0 (98.5)	9.3 (31.2)	–	–	–	–	–	–
Hg-1 (6 I)	20–3.15 (3.2–3.15)	12004/42666	98.8 (99.8)	11.6 (36.1)	33.3 (39.7)	6 Hg 2 I, 4 I⁻**	1.20/1.41	0.76	0.82/0.88	0.81
Hg-2	20–2.8 (2.9–2.8)	16138/38449	94.5 (99.7)	6.4 (34.5)	26.5 (31.2)	3 Hg 8 I⁻**	0.98/1.20	0.97	0.88/0.93	0.91
Iodo (15 I)	20–3.0 (3.11–3.0)	12921/18547	92.9 (95.0)	8.1 (28.4)	24.5 (24.9)	7 I	2.21/–	1.84	0.66/–	0.70
Pt (10 I)	20–3.2 (3.31–3.2)	10797/20475	94.4 (93.2)	7.7 (34.3)	17.1 (22.5)	3 Pt 2 I	2.20/–	2.42	0.67/–	0.70

Structural refinement of the covalent (form 6) and noncovalent (form 2) complexes

	Resolution (Å) (last shell)	Reflections unique/total	Redundancy	Complete (%) (last shell)	R_{sym}* (%) (last shell)	Number of protein, DNA atoms	Number of waters	R‖ (%)	R_{free}¶ (%)	rmsd# bonds (Å)	rmsd# angles (°)
Form 6 (10 I)	20–2.1 (2.15–2.10)	42786/ 79570	1.9	92.7 (95.5)	5.5 (23.4)	4681	360	24.9	34.6	0.012	1.60
Form 2 (8 Br)	20–2.5 (2.6–2.5)	18834/ 78779	4.2	83.9 (87.7)	7.7 (26.9)	4768	63	22.5	33.8	0.012	1.61

*$R_{sym} = \Sigma|I - \langle I\rangle|/\Sigma I$, where I is the observed intensity and $\langle I\rangle$ the average intensity of multiple symmetry-related observations of that reflection. †$R_{iso} = \Sigma||F_{PH}| - |F_P||/\Sigma|F_P|$, where $|F_P|$ is the protein structure factor amplitude and $|F_{PH}|$ is the heavy-atom derivative structure factor amplitude. ‡*PhP* (phasing power) = (rms $|F_h|$)/(rms E), where $|F_h|$ is the heavy-atom structure factor amplitude and E the residual lack of closure error. §$R_{cullis} = \Sigma|E|/\Sigma||F_{PH}| - |F_P||$. ‖$R = \Sigma||F^{obs}{}_P| - |F^{calc}{}_P||/\Sigma|F^{obs}{}_P|$. ¶$R_{free} = \Sigma||F^{obs}{}_P| - |F^{calc}{}_P||/\Sigma|F^{obs}{}_P|$ where $|F^{obs}{}_P|$ is from a test set not used in structural refinement (7% of the observed data). #rmsd values are root-mean-square deviations from ideal bond lengths and bond angles. **$|^-$ denotes iodine positions with negative occupancies.

and the COOH-terminal domain are quite limited but those involving subdomains I and III are very extensive. Subdomains I and III make up in large part the molecular clamp that closes around the DNA before cleavage and covalent attachment. In keeping with the nearly identical protein and DNA conformations in the two crystal forms, the protein-DNA contacts made in the noncovalent complex are essentially the same as those made in the covalent complex (Fig. 4, A, F, and G).

The nose-cone helices of the "cap" of the enzyme offer a highly positively charged surface to the sugar-phosphate backbone, yet the DNA declines to interact directly with these positively charged residues in both the covalent and noncovalent complexes. Elsewhere, extensive protein-DNA contacts do occur, including residues 410 to 429 from subdomain I (Fig. 4C), which position β strand 11 into the major groove roughly opposite the cleavage site. The regions containing the lips between subdomains I and III are also in close contact with the DNA (Fig. 4, B and D). As expected, residues near the active site make several DNA contacts, especially directly across the minor groove from the cleavage site (Fig. 4E). Moreover, the dipole of helix 16 helps to stabilize the protein-DNA interactions by bringing its NH_2-terminus close to the phosphate between the −4 and −5 nucleotides of the intact strand.

The protein contacts only the central 10 bp of the DNA (that is, positions −4 to +6), and the vast majority of these contacts involve the DNA phosphate groups (Fig. 4, F and G). The total number of protein-DNA contacts closer than 3.5 Å is 30 for the noncovalent complex and 27 for the covalent complex, and several DNA contacts are made by main chain amide nitrogen atoms as hydrogen bond donors (Fig. 4, C to G). A total of 26 residues in the noncovalent complex and 24 residues in the covalent complex participate in these direct protein-DNA contacts. There is only one base-specific contact, a 2.6 to 2.9 Å hydrogen bond in the minor groove between Lys^{532} and the O-2 carbonyl oxygen of the −1 thymidine base on the scissile strand (Fig. 4D). This interaction is probably related to the preference observed for thymidines at the scissile −1 position in strong breaks sites (*18*), because thymidine is the only nucleotide base in which the two lone electron pairs of the O-2 atom are freely available to accept a hydrogen bond from the ε-amino group of Lys^{532}. In general, topo I makes extensive protein-DNA contacts, similar to the way certain restriction enzymes wrap around their target DNA, but with only one base-specific contact and little evidence for sequence-specific DNA recognition.

The essential Tyr^{723}. The electron density at the active site of the covalent reconstituted complex reveals that the catalytic Tyr^{723} is indeed covalently attached by a phosphodiester bond with the 3′ end of the scissile strand (Fig. 5A). This covalent attachment was evident even in the original maps with experimental phases at 3.2 Å resolution and was confirmed by capillary electrophoresis of washed and dissolved crystals (not shown). The free 5′-sulfhydryl of the cleaved strand had not carried out the religation reaction, in accordance with biochemical studies (*34*). The phosphate of the tyrosine-DNA phosphodiester bond makes close interactions with the guanidinium groups of Arg^{488} (at a distance of 2.8 Å) and Arg^{590} (at 2.6 Å). Both arginines are implicated in the reaction mechanism of the enzyme, as described in the accompanying report (*20*).

The electron density at the active site in the noncovalent complex shows unequivocally that residue 723 is a phenylalanine in the mutant enzyme and that the DNA strand is intact (Fig. 5B). When the covalent and noncovalent complexes are superimposed, it is observed that (i) the phenyl rings of Tyr^{723} and Phe^{723} are rotated ~20° relative to one another and the Cα positions of these residues are shifted by ~0.5 Å; (ii) the conserved residues Arg^{488} and Arg^{590} are displaced by rms deviations, respectively, of 0.45 and 0.27 Å when averaged over all atoms; (iii) the 5′ bridging oxygen of the +1 nucleotide of the scissile strand is shifted ~0.8 Å with respect to the free 5′-sulfhydryl group; and (iv) the phosphorous atoms of the scissile phosphates have moved 1.6 Å with respect to one another. Thus, the differences between the covalent and noncovalent complexes of human topo I with DNA are confined mainly to the catalytic tyrosine and the scissile phosphate group. The residues around the active site, including the catalytic arginines 488 and 590, are shifted minimally, as are the deoxyribose rings upstream and downstream of the cleavage site. Hence, the formation of the covalent complex does not seem to require marked structural changes in the protein or the DNA.

Camptothecin binding. Camptothecin (Fig. 6A) is a cytotoxic plant alkaloid that rapidly blocks both DNA and RNA synthesis in treated cells and has emerged as a potent anticancer compound (*5*, *21*, *22*, *30*). Topo I is the sole intracellular target for CPT and CPT derivatives (*23*), which act by stabilizing the covalent topo I–DNA complex. In in vitro studies, CPT and its analogs inhibit plasmid relaxation by human topo I and increase the yield of covalent intermediates when reactions are stopped with SDS (*22*, *26*, *40*). The observations reviewed below provide the rationale for a provisional placement of CPT into the structure of the covalent topo I–DNA complex shown in Fig. 6B.

Numerous structure-activity studies have identified several design principles for CPT analogs and have also established a direct correlation between the ability of

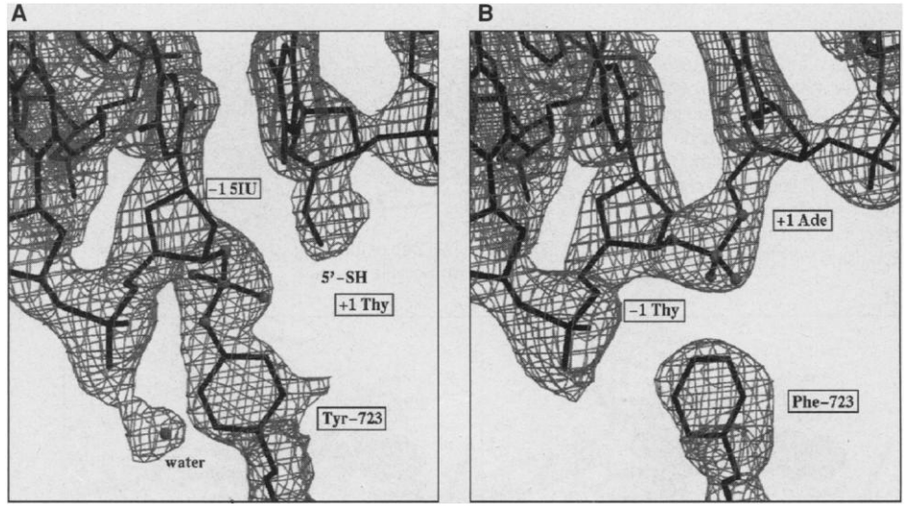


Fig. 5. Covalent and noncovalent complexes. **(A)** Close-up of the 2.1 Å $2|F_{obs}| - |F_{calc}|$ electron density map of the covalent complex near the active site, contoured at 1.5σ, revealing the covalent 3′-phosphate-tyrosine bond between Tyr^{723} and the broken strand of the DNA. **(B)** Close-up of the 2.5 Å $2|F_{obs}| - |F_{calc}|$ electron density map of the noncovalent complex near the active site, contoured at 1.5σ, revealing the inactive Tyr^{723}→Phe mutation and the intact DNA oligonucleotide.

CPT derivatives to stabilize the covalent topo I–DNA intermediate and their ability to kill cancer cells (*40*, *41*). The structural features of CPT that are essential for activity include the 20(*S*)-hydroxyl (*42*), the pyridone moiety of the D-ring, the lactone moiety of the E-ring (*24*), and the planarity of the five-membered ring system (Fig. 6A) (*43*). Hence, the C-D-E rings of CPT cannot be altered without severely affecting its ability to block religation by topo I. In contrast, modifications to the 9, 10, and 11 positions of the A-ring and the 7 position of the B-ring are generally well tolerated and in many cases enhance the potency of the CPT analog in both in vitro and in vivo studies (*5*). Furthermore, substantial evidence indicates that CPT binds reversibly only after cleavage and covalent attachment of the enzyme to the DNA (*27*).

CPT preferentially enhances DNA breakage at sites with a guanine base at the +1 position on the scissile strand, immediately downstream of the cleavage site (*44*). Moreover, studies by Pommier and coworkers have shown that chemically reactive analogs of CPT that contain a chloromethyl addition to the C-7 position can alkylate the N-3 nitrogen of the +1 guanine base after cleavage (*45*). It is inferred from these and other studies (*32*) that upon cleavage and covalent attachment, the +1 guanine can assume a conformation in which it is stacked against the five-membered CPT ring system such that the guanine N-3 is in close proximity to the C-7 of CPT (*45*).

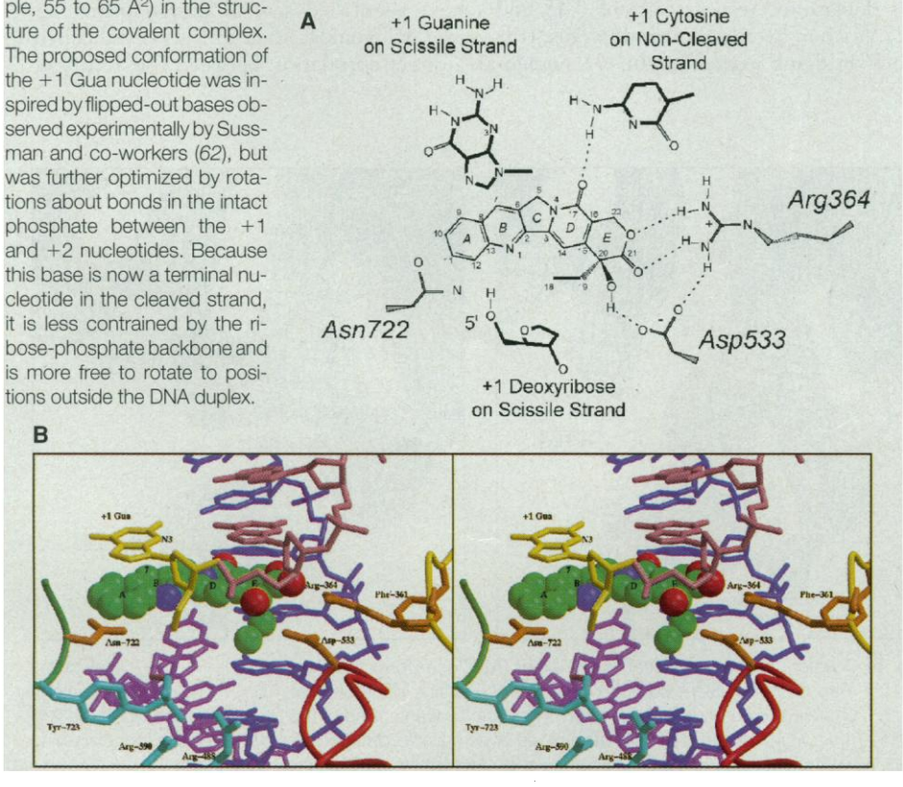


Fig. 6. Proposed CPT binding mode. (**A**) A schematic representation of the key hydrogen bond and ring-stacking interactions made between the human topo I–DNA covalent complex and CPT in the proposed CPT binding mode. The atomic nomenclature for CPT is also indicated. (**B**) Stereoview of the proposed binding mode of CPT to the covalent human topo I DNA complex. The active lactone form of CPT (20-*S*-camptothecin, in green) is shown stacked between the terminal +1 guanine nucleotide from the cleaved strand (+1 Gua, in yellow, which is reoriented from the observed position as described below), and the side chain of Asn^{722}, which provides interactions with the A-ring of CPT (the cleaved strand is rendered in light and dark magenta upstream and downstream of the cleavage site, respectively). The carbonyl oxygen at the 17 position in CPT makes a hydrogen bond with the NH_2 group on the pyrimidine ring of the +1 cytosine. The side chains of active-site residues Tyr^{723}, Arg^{488}, and Arg^{590} are shown in cyan. The side chain residues that, if singly mutated, result in a CPT-resistant phenotype [Phe^{361}, Gly^{363}, and Arg^{364} of region 1 (see text); Asp^{533} and Asn^{722} of region 2] are shown in tan. The side chain conformations of Arg^{364} and Asp^{533} have been altered slightly from the final structure of the covalent complex to allow for optimal hydrogen bonding to the double-bonded lactone oxygen and the hydroxyl at the 20-*S* chiral center of CPT, respectively. Modifications to the 10 and 11 positions of CPT may require some minor shifts in the positions of residues Lys^{720} and Leu^{721} of topo I, which exhibit relatively high temperature factors (for example, 55 to 65 Å^2) in the structure of the covalent complex. The proposed conformation of the +1 Gua nucleotide was inspired by flipped-out bases observed experimentally by Sussman and co-workers (*62*), but was further optimized by rotations about bonds in the intact phosphate between the +1 and +2 nucleotides. Because this base is now a terminal nucleotide in the cleaved strand, it is less contrained by the ribose-phosphate backbone and is more free to rotate to positions outside the DNA duplex.

A small number of CPT resistance mutants (CPT^r) have been reported for the topo I enzymes from human, yeast, and hamster (*31*, *32*) (Figs. 2A and 4, B, D, and E). Projecting all known CPT escape mutants from the various species onto the three-dimensional structure of human topo I suggests that there are four regions of the protein that can be mutated to produce a CPT^r enzyme: region 1: residues Gly^{717} to Asn^{722}, which lie in close proximity to the active-site Tyr^{723} (Fig. 4E); region 2: the interface between α helix 13 (residues Lys^{532} to Ser^{534}) in core subdomain III, which contains the human CPT^r mutant $Asp^{533}\rightarrow Gly$, and a loop segment (residues Phe^{361} to Met^{370}) between β strands 6 and 7 of core subdomain I, which contains the human CPT^r mutations at positions 361, 363, and 364 (Fig. 4D); region 3: Thr^{729}, which lies deep within the hydrophobic core of the COOH-terminal domain (Fig. 4B); and region 4: a conserved hydrophobic segment of core subdomain III, near α helix 13, involving residue Gly^{503} (Fig. 2A). The amino acid residues at positions 503 and 729, which lie in regions 3 and 4, are far removed from the active site and from the DNA oligonucleotide and hence are not directly taken into account in our model for CPT binding. These residues may play a structural role in the proper packing of the COOH-terminal and core domains and may affect CPT efficacy by interfering with the positioning of catalytic or CPT-binding residues.

Combining our crystallographic information with these chemical and biochemical observations, we propose a hypothetical CPT binding mode (Fig. 6A). We chose to model CPT into the structure of the covalent complex because it is the relevant target of CPT binding. The locations of CPT^r side chains suggest that the CPT binding site spans the DNA duplex from region 1 of the protein to region 2. An attractive feature of this binding mode is that carbon positions 7 and 9 of CPT, which are amenable to chemical modification, are facing out into open space. Thus, modifications to the C-7 and C-9 positions of CPT are allowed and could form favorable interactions with distant protein or DNA atoms that would explain the increased potency such derivatives have relative to unmodified CPT. Insertion of CPT between regions 1 and 2 requires the movement or displacement of the +1 bp immediately downstream of the cleavage site. This movement could be achieved by repositioning the guanine nucleotide at the +1 position of the scissile strand outside of the DNA duplex and allowing CPT to occupy the space vacated by the base, such that a reasonable hydro-

gen bonding network can be established between the Asp^{533} and Arg^{364} side chains and the immutable 20(S)-hydroxyl and lactone moieties of CPT (Fig. 6, A and B). Mutations of the Phe^{361} and Gly^{363} residues are likely to mediate CPT resistance by disrupting the conformation of the loop that holds Arg^{364}.

In this binding mode, the terminal carbon (C-18) of the CPT ethyl group has been rotated slightly relative to its position in the CPT crystal structure (*46*) to remove steric clashes with the side chain of Asp^{533} and the −1 nucleotide base. A CPT molecule docked in this way would explain the necessity for an S configuration at the C-20 position. The inactive *R*-diastereoisomer of CPT would fail to bind to right-handed DNA in this manner because of major steric clashes with the +2 bp.

Support for the displacement of the +1 guanine base from its normal base-paired position comes from the alkylation studies, which indicate that the N-3 position of this guanine is in close proximity to the C-7 position of CPT in the covalent complex (*45*). This proximity has been achieved by the aforementioned repositioning of the terminal guanine of the broken strand, which allows the base to stack on top of the planar CPT (Fig. 6B). In this configuration, the A-ring of CPT forms an additional stacking interaction with the planar conjugated side chain of Asn^{722}. The intact cytosine nucleotide at the +1 position in the noncleaved strand further stabilizes this binding mode by forming a hydrogen bond between its N-3 amino group and the pyridone oxygen of the D-ring of CPT. In the proposed topo I-DNA:CPT complex, the free 5′ hydroxyl group is over 4.5 Å removed from the phosphate group to be attacked for religation. Hence, the enzyme is trapped in its covalent complex with DNA.

REFERENCES AND NOTES

1. J. J. Champoux, in *DNA Topology and Its Biological Effects*, J. C. Wang and N. R. Cozzarelli, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1990), pp. 217–242; J. C. Wang, *Annu. Rev. Biochem.* **65**, 635 (1996); J. J. Champoux, *Prog. Nucleic Acids Res. Mol. Biol.*, in press.
2. N. Lue, A. Sharma, A. Mondragon, J. C. Wang, *Structure* **3**, 1315 (1995).
3. A. Sharma, R. Hanai, A. Mondragon, *ibid.* **2**, 767 (1994).
4. C. D. Lima, J. C. Wang, A. Mondragón, *Nature* **367**, 138 (1994); J. M. Berger, S. J. Gamblin, S. C. Harrison, J. C. Wang, *ibid.* **380**, 179 (1996).
5. H.-K. Wang, S. L. Morris-Natschke, K.-H. Lee, *Med. Res. Rev.* **17**, 367 (1997).
6. R. Shaffer and P. Traktman, *J. Biol. Chem.* **262**, 9309 (1987).
7. S. Shuman and B. Moss, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7478 (1987).
8. M. Gupta, A. Fujimori, Y. Pommier, *Biochim. Biophys. Acta* **1262**, 1 (1995); P. R. Caron and J. C. Wang, *Adv. Pharmacol.* **29B**, 271 (1994).
9. The crystal structures of two proteolytic fragments of topo I have been reported, from vaccinia (*3*) and yeast (*2*). However, these structures represent relatively small (9 and 26 kD, respectively), NH_2-terminal fragments of the enzyme that do not contain active-site residues or bound DNA.
10. P. D'Arpa *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2543 (1988).
11. L. Stewart, G. C. Ireton, L. H. Parker, K. R. Madden, J. J. Champoux, *J. Biol. Chem.* **271**, 7593 (1996).
12. L. Stewart, G. C. Ireton, J. J. Champoux, *ibid.*, p. 7602.
13. J. Alsner, J. Q. Svejstrup, E. Kjeldsen, B. S. Sorensen, O. Westergaard, *ibid.* **267**, 12408 (1992).
14. A. K. Bharti, M. O. Olson, D. W. Kufe, E. H. Rubin, *ibid.* **271**, 1993 (1996).
15. K. R. Madden and J. J. Champoux, *Cancer Res.* **52**, 525 (1992).
16. L. Stewart, G. C. Ireton, J. J. Champoux, *J. Mol. Biol.* **269**, 355 (1997).
17. The reconstituted enzyme is more sensitive to ionic strength and more distributive in its mode of action than the full-length enzyme (*16*, *20*).
18. M. D. Been, R. R. Burgess, J. J. Champoux, *Nucleic Acids Res* **12**, 3097 (1984).
19. A. H. Andersen, E. Gocke, B. J. Bonven, O. F. Nielsen, O. Westergaard, *ibid.* **13**, 1543 (1985).
20. L. Stewart, M. R. Redinbo, X. Qiu, W. G. J. Hol, J. J. Champoux, *Science* **279**, 1534 (1998).
21. M. E. Wall *et al.*, *J. Am. Chem.* **88**, 3888 (1966); S. B. Horwitz, C. K. Chang, A. P. Grollman, *Mol. Pharmacol.* **7**, 632 (1971).
22. Y. H. Hsiang, R. Hertzberg, S. Hecht, L. F. Liu, *J. Biol. Chem.* **260**, 14873 (1985).
23. M. A. Bjornsti, P. Benedetti, G. A. Viglianti, J. C. Wang, *Cancer Res.* **49**, 6318 (1989).
24. R. P. Hertzberg *et al.*, *J. Med. Chem.* **32**, 715 (1989).
25. S. Sawada *et al.*, *Chem. Pharmacol. Bull.* **39**, 3183 (1991).
26. S. E. Porter and J. J. Champoux, *Nucleic Acids Res.* **17**, 8521 (1989).
27. R. P. Hertzberg, M. J. Caranfa, S. M. Hecht, *Biochemistry* **28**, 4629 (1989); J. Q. Svejstrup, K. Christiansen, I. I. Gromova, A. H. Andersen, O. Westergaard, *J. Mol. Biol.* **222**, 669 (1991).
28. R. P. Hertzberg *et al.*, *J. Biol. Chem.* **265**, 19287 (1990).
29. C. Bendixen, B. Thomsen, J. Alsner, O. Westergaard, *Biochemistry* **29**, 5613 (1990).
30. Y. H. Hsiang, M. G. Lihou, L. F. Liu, *Cancer Res.* **49**, 5077 (1989).
31. P. Benedetti, P. Fiorani, L. Capuani, J. C. Wang, *ibid.* **53**, 4343 (1993); A. M. Knab, J. Fertala, M. A. Bjornsti, *J. Biol. Chem.* **268**, 22322 (1993); E. Rubin *et al.*, *ibid.* **269**, 2433 (1994); L. F. Wang *et al.*, *Cancer Res.* **57**, 1516 (1997); H. Tamura *et al.*, *Nucleic Acids Res.* **19**, 69 (1991); A. Tanizawa *et al.*, *J. Biol. Chem.* **268**, 25463 (1993); N. Kubota *et al.*, *Biochem. Biophys. Res. Commun.* **188**, 571 (1992); X. G. Li *et al.*, *Bioch. Pharmacol.* **53**, 1019 (1997).
32. F. Leteurtre *et al.*, *J. Biol. Chem.* **269**, 28702 (1994).
33. The reconstituted enzyme was prepared by incubating crude nuclear extracts containing the 58-kD core domain (expressed in insect cells) with the immobilized glutathione-S-transferase (GST)–COOH-terminal domain fusion protein (expressed in *E. coli*) (*16*). The reconstituted core plus COOH-terminal enzyme was liberated by factor Xa cleavage. The reconstituted human topo I was separated from factor Xa and any free GST or COOH-terminal domain by POROS SP20 chromatography as described (*16*). DNA oligonucleotides were synthesized by DNA express (Fort Collins, CO) and purified as described (*11*). Oligonucleotides that contained a 5′-bridging phosphorothioate linkage at the site of topo I cleavage were synthesized according to Burgin *et al.* (*34*), where 5′-(*S*-trityl)-mercapto-5′-deoxythymidine-3′-*O*-(2-cyanoethyl-*N*,*N*-diisopropylamino)phosphite was used for the synthesis of the +1 nucleotide in the cleaved strand of the high-affinity topo I binding site from *T. thermophilus* (*19*). Duplex oligonucleotides were prepared by mixing equimolar concentrations of complementary strands, each at a final concentration of 0.1 mM, in the presence of 6 mM NaCl. Annealing of the two strands was achieved by slow cooling from 80°C to room temperature over a period of 8 hours.
34. A. B. Burgin Jr., B. N. Huizenga, H. A. Nash, *Nucleic Acids Res.* **23**, 2973 (1995).
35. E. de La Fortelle and G. Bricogne, *Methods Enzymol.* **276**, 472 (1997).
36. L. Holm and C. Sander, *Nucleic Acids Res.* **25**, 231 (1997).
37. J. D. Klemm, M. A. Rould, R. Aurora, W. Herr, C. Pabo, *Cell* **77**, 21 (1994).
38. A. B. Hickman, S. Waninger, J. J. Scocca, F. Dyda, *ibid.* **89**, 227 (1997).
39. D. Esposito and J. J. Scocca, *Nucleic Acids Res.* **25**, 3605 (1997); HP1 integrase also uses a transient 3′-phosphate-tyrosine intermediate during its catalytic cycle of integration and excision of the phage DNA and is representative of the family of bacteriophage integrases.
40. C. Jaxel, K. W. Kohn, M. C. Wani, M. E. Wall, Y. Pommier, *Cancer Res.* **49**, 1465 (1989); A. Tanizawa, K. W. Kohn, G. Kohlhagen, F. Leteurtre, Y. Pommier, *Biochemistry* **34**, 7200 (1995).
41. R. T. Crow and D. M. Crothers, *J. Med. Chem.* **35**, 4160 (1992).
42. M. E. Wall *et al.*, *ibid.* **36**, 2689 (1993).
43. K. Lackey *et al.*, *ibid.* **38**, 906 (1995); W. D. Kingsbury *et al.*, *ibid.* **34**, 98 (1991).
44. A. Tanizawa, K. W. Kohn, Y. Pommier, *Nucleic Acids Res.* **21**, 5157 (1993); E. Kjeldsen *et al.*, *J. Mol. Biol.* **202**, 333 (1988).
45. Y. Pommier *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 8861 (1995).
46. A. T. McPhail and G. A. Sim, *J. Chem. Soc.* **B**, 923 (1968).
47. Z. Otwinowski and W. Minor, Eds., *Data Collection and Processing* (Daresbury, UK, 1993).
48. Z. Otwinowski, in *ibid.*, pp. 80–86.
49. K. Cowtan, *Joint CCP4 and ESF-EACBM Newslett. Protein Crystallogr.* **31**, 34 (1994).
50. J. P. Abrahams and A. G. W. Leslie, *Acta Crystallogr.* **D52**, 30 (1996).
51. T. A. Jones, J.-Y. Zou, S. W. Cowan, *ibid.* **A47**, 110 (1991).
52. J. Navaza, *ibid.* **A50**, 157 (1994).
53. A. T. Brunger, *X-PLOR (Version 3.1): A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, CT, 1993).
54. ———, *J. Mol. Biol.* **203**, 803 (1988).
55. ———, *Nature* **355**, 472 (1992).
56. R. J. Read, *Acta Crystallogr.* **A42**, 140 (1986).
57. J. M. Moore, W. J. Chazin, R. Powls, P. E. Wright, *Biochemistry* **27**, 7806 (1988).
58. P. J. Kraulis, *J. Appl. Crystallogr.* **24**, 946 (1991).
59. D. J. Bacon and W. F. Anderson, *J. Mol. Graphics* **6**, 219 (1988); E. A. Merritt and M. E. P. Murphy, *Acta Crystallogr.* **D50**, 869 (1994).
60. A. Nicholls, K. A. Sharp, B. Honig, *Proteins Struct. Funct. Genet.* **11**, 282 (1991).
61. G. J. Barton, *Protein Eng.* **6**, 37 (1993).
62. L. Joshua-Tor *et al.*, *J. Mol. Biol.* **225**, 397 (1992).

63. We thank S. Turley, E. Pohl, S. Suresh, B. Bernstein, A. Ævarsson, D. Chudzik, J. Yeh, F. Athappilly, F. van den Akker, C. Verlinde, and many other members of the Biomolecular Structure Center at the University of Washington for technical assistance and intellectual support. We also thank the staffs of the Stanford Synchrotron Radiation Laboratory, the Brookhaven National Laboratory, the Cornell High Energy Synchrotron Source, and the European Synchrotron Radiation Laboratory for their extensive support with data collection. Thanks are also due to A. Burgin for advice on the synthesis of oligonucleotides containing a 5′-bridging phosphorothioate, and to DNA Express (Fort Collins, CO) for conducting the syntheses. Supported by American Cancer Society grant PF-3905 (L.S.), grant GM49156 from the National Institutes of Health (J.J.C.), by a major equipment grant from the Murdock Charitable Trust to the Biomolecular Structure Center, and grant NCI CA65656 from the National Institutes of Health (W.G.J.H.). The Stanford Synchrotron Radiation Laboratory is founded by the Department of Energy (BES, OBER) and the NIH (NCRR).

1 December 1997; accepted 28 January 1998