A 26 kDa yeast DNA topoisomerase I fragment: crystallographic structure and mechanistic implications

Neal Lue[1], Amit Sharma[2], Alfonso Mondragón[2]* and James C Wang[1]*

[1]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA and [2]Department of Biochemistry, Molecular Biology, and Cell Biology, Northwestern University, 2153 Sheridan Road, Evanston, IL 60208, USA

Background: Type I DNA topoisomerases, divided mechanistically into two subfamilies, are ubiquitous enzymes that participate in replication and transcription. In addition to its role in these fundamental processes, the biological importance of eukaryotic DNA topoisomerase I is underscored by its identification as the target of the antitumor alkaloid camptothecin. An understanding of the mechanism of catalysis and interactions with camptothecin and other drugs has been hampered by a lack of detailed structural information.

Results: The three-dimensional structure of a 26 kDa fragment (residues 135 to about 363) of *Saccharomyces cerevisiae* DNA topoisomerase I has been determined at 1.9 Å resolution. The fragment has a novel architecture comprising a concave platform and a pair of outlying

V-shaped helices. Photocrosslinking and protein footprinting experiments show that the positively charged concave surface and the junction region of the V-shaped pair of helices contact DNA in the enzyme–DNA complex.

Conclusions: Crystallographic, biochemical and genetic data indicate that this 26 kDa fragment of yeast DNA topoisomerase I is involved in complex formation between the enzyme and DNA, and probably also in camptothecin–enzyme–DNA ternary complex formation. A molecular model for protein–DNA interaction based on these data is proposed. The bipartite DNA-binding regions of the 26 kDa fragment may enable eukaryotic DNA topoisomerase I to adapt to sequence-dependent structural variations in its DNA substrates.

Introduction

DNA topoisomerases catalyze topological transformations of DNA rings via successive trans-esterification reactions [1]. Type I DNA topoisomerases, which effect such transformations through the transient breakage of one DNA strand and the passage of another strand through the broken strand before its resealing, can be classified into two subfamilies. One subfamily, which includes *Escherichia coli* DNA topoisomerases I and III and yeast DNA topoisomerase III, preferentially relaxes negatively supercoiled DNA, forms a 5′ phosphotyrosine linkage in the enzyme–DNA covalent intermediate, and has high affinity for single-stranded DNA [2]. The other subfamily, which includes eukaryotic DNA topoisomerase I, poxviral topoisomerases [3], and probably topoisomerase V of the thermophile *Methanopyrus kandleri* [4], relaxes both positively and negatively supercoiled DNA, forms a 3′ phosphotyrosine linkage in the enzyme–DNA covalent intermediate, and prefers double-stranded DNA substrates [2]. Eukaryotic DNA topoisomerase I is a ubiquitous enzyme that participates in a number of key cellular processes including transcription, replication, and recombination [5]. Interests in this type of enzyme have been further stimulated as a result of its identification as the target of the antitumor alkaloid camptothecin [6].

The three-dimensional structure of a 67 kDa N-terminal fragment of *E. coli* DNA topoisomerase I has been

determined at 2.2 Å resolution [7]. Both structural and biochemical data support an enzyme-bridging mechanism of DNA strand passage by this enzyme (i.e. the enzyme interacts with both sides of the transiently broken DNA strand in its catalysis of DNA strand passage [2,7]). By contrast, structural information for the eukaryotic DNA topoisomerase I subfamily of enzymes is scarce. A 9 kDa N-terminal fragment of the vaccinia virus topoisomerase was recently solved at 1.6 Å [8]. However, the lack of significant sequence homology between this region of the viral enzyme and eukaryotic DNA topoisomerase I makes it difficult to extend the results from the former to the latter [8,9]. We therefore crystallized and solved the structure of a 26 kDa fragment of yeast DNA topoisomerase I comprising residues 135 to about 363. The fragment has a novel architecture which, in combination with photocrosslinking and protein footprinting results, leads to a model for the binding of this fragment to DNA.

Results and discussion

Crystallization and structure determination

The 26 kDa fragment of *Saccharomyces cerevisiae* DNA topoisomerase I comprises residues 135 to about 363. Deletion mutagenesis experiments have previously shown that the N-terminal 135 amino acids of *S. cerevisiae* DNA topoisomerase I are not required for relaxation activity *in vitro*. Complementation experiments in *E. coli* have shown that *S. cerevisiae* DNA topoisomerase I can

---

*Corresponding authors.

substitute for *E. coli* DNA topoisomerase I function *in vivo*. However, in these experiments the N-terminal 135 amino acids of *S. cerevisiae* DNA topoisomerase I are necessary for *in vivo* complementation of the *E. coli* enzyme. Larger deletions of the yeast enzyme extending into the 26 kDa fragment abolish enzymatic activity *in vitro* and *in vivo* [10]. The 26 kDa fragment was obtained by proteolytic treatments of an overexpressed polypeptide (as described in the Materials and methods section), and was found to crystallize in both tetragonal and orthorhombic crystal forms. The crystal structure of the orthorhombic form was solved to 2.8 Å by multiple isomorphous replacement (MIR) and refined to 1.9 Å (Tables 1,2). The structure of the tetragonal form was solved by molecular replacement to 3.0 Å (Tables 1,2); overall the structure in the two crystals forms is similar, although there are significant differences (root mean square deviation [rmsd] of 0.8 Å for all Cα atoms; see below).

Architecture and electrostatics

The 26 kDa fragment forms a platform with a concave surface on one side. The platform consists of two domains (I and II) linked by an extended polypeptide stretch (PPI), and by a pair of outlying α helices (C and D) that are connected in the shape of a V (see Fig. 1a,b). Domain I is formed by two antiparallel α helices (A and

Table 2. Refinement statistics.

Crystal form	Orthorhombic	Tetragonal
Resolution limits (Å)	8–1.9	7–3.0
R-factor (%)*	22.2	20.4
R_{free} (%)[†]	29.0	#
Number of amino acid residues	220	214
Number of water molecules	232	**
Rms bond length (Å)[‡]	0.014	0.015
Rms bond angle (°)[‡]	1.73	3.0
Average B-factor (Å²)[§]	27.0	20.0

*R-factor is based on 90% of the native data set used in refinement. [†]R_{free} is the R-factor based on 10% of the native data withheld for the cross-validation test as implemented in X-PLOR. [‡]Root mean square deviation from ideal values for bond lengths and angles. [§]Average B-factor of all atoms in the structure. #No R_{free} evaluation was done for the tetragonal data because the data set was incomplete. **No water molecules were modeled into the tetragonal crystal form because of the limiting resolution of native data.

B) lying against one side of a three-stranded antiparallel β sheet (S2, S3 ,S4). Helices A and B make hydrophobic contacts with a long β-hairpin-like loop (CL) that protrudes from domain II. Both the N and C termini of the fragment reside in domain II, which is made of two β sheets forming a distorted β-barrel-like structure. The 'elbow' of the V-shaped α helices (residues 252–254) connecting domains I and II is disordered in both crystal forms, suggesting an inherent flexibility. The PPI segment connecting domains I and II is proline-rich and forms a left-handed polyproline helix of the type usually found in mobile regions on protein surfaces [11]. The presence of the disordered elbow and the polyproline helix suggests that there may be some flexibility in the 26 kDa fragment and in the relative disposition of domains I and II. This is supported by the observation that domain I is virtually identical in both crystal forms (rmsd for Cα atoms of 0.41 Å), but domain II is more variable (rmsd for Cα atoms of 0.76 Å). Superimposing both domains together reveals that the V-shaped helices differ the most (rmsd for Cα atoms of 1.6 Å).

The spatial disposition of domains I, II, and the CL loop creates a concave surface with a curvature of approximately 30 Å diameter on one side of the platform of the 26 kDa fragment, a curvature close to that observed in other proteins that are known, or thought, to interact with double-stranded DNA, such as *E. coli* DNA topoisomerase I [7], the DNA polymerase processivity factors from *E. coli* and *S. cerevisiae* [12,13], and the TATA-binding protein [14,15]. Additionally, the concave side has a total of 15 positively charged amino acids, 11 of which are highly conserved. These positive charges create a strong positive electrostatic potential [16] that is absent from the rest of the protein fragment (Fig. 2).

Photocrosslinking of the 26 kDa fragment to DNA

The 26 kDa fragment does not exhibit significant DNA-binding activity in filter-binding or band-shift assays (data

Table 1. Data collection and MIR statistics.

| Crystal	Native I	Hg	Pt	Native II	Native III
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P42_12$
Temperature	–1°C			–170°C	RT
Dimensions (Å)					
a	48.06			47.20	153.9
b	53.58			53.07	153.9
c	118.01			116.89	39.16
Resolution range (Å)	34.2–2.8	26.3–2.5	26.5–4.0	39.0–1.85	42.6–2.94
Reflections					
unique	7521	10722	2060	24425	8859
total	17921	62572	6596	127777	20584
Completeness (%)	96.6 (95.5)**	98.1 (89.5)	72.5 (74.9)	95.1 (80.1)	80.3 (21.4)
R_{sym} (%)*	6.3 (10.5)	6.4 (15.4)	5.0 (4.8)	6.4 (27.8)	7.8 (34.1)
MFID[†]	–	36.5	11.9	–	–
MIRAS analysis (20–2.8 Å)					
Number of sites	–	1	2	–	–
Phasing power[‡]					
acentric	–	1.2	0.9	–	–
centric	–	1.0	0.7	–	–
R_c[§]	–	0.83	0.90	–	–
Mean FOM#					
acentric	0.34				
centric	0.42				

*$R_{sym}=\Sigma$ I–<I> $|/\Sigma$I, where I=observed intensity, and <I>=average intensity obtained from multiple measurements. [†]MFID (mean fractional isomorphous difference)=Σ $|F_{ph}|-|Fp|$ $|/\Sigma|F_p|$, where $|F_p|$ and $|F_{ph}|$ are the protein and heavy-atom derivative structure factor amplitudes, respectively. [‡]Phasing power=rms($|F_h|$/E), where $|F_h|$=heavy-atom structure-factor amplitude and E=residual lack of closure error. [§]$R_c=\Sigma$ $||F_{h(obs)}|-|F_{h(calc)}||/\Sigma|F_{h(obs)}|$ for centric reflections where $|F_{h(obs)}|$ and $|F_{h(calc)}|$ are the observed and calculated heavy-atom structure-factor amplitudes, respectively. #FOM, figure of merit. **Numbers in parentheses designate values for the highest resolution shell.

Fig. 1. Crystal structure of the 26 kDa *S. cerevisiae* DNA topoisomerase I fragment. **(a)** Ribbon diagram showing the overall architecture of the fragment. The two domains are labeled as well as the secondary structural elements. α helices are denoted A to D, β strands S1 to S10. The polyproline helix joining domains I and II is labeled PPI; the loop extending from domain II into domain I is labeled CL. (Drawn using MOLSCRIPT [32].) **(b)** Stereo drawing of the Cα backbone trace of the fragment. The view is perpendicular to (a), with the two V-shaped α helices pointing away from the viewer, and with every 20th residue numbered. The dotted line denotes a break in electron density for residues 252–254.
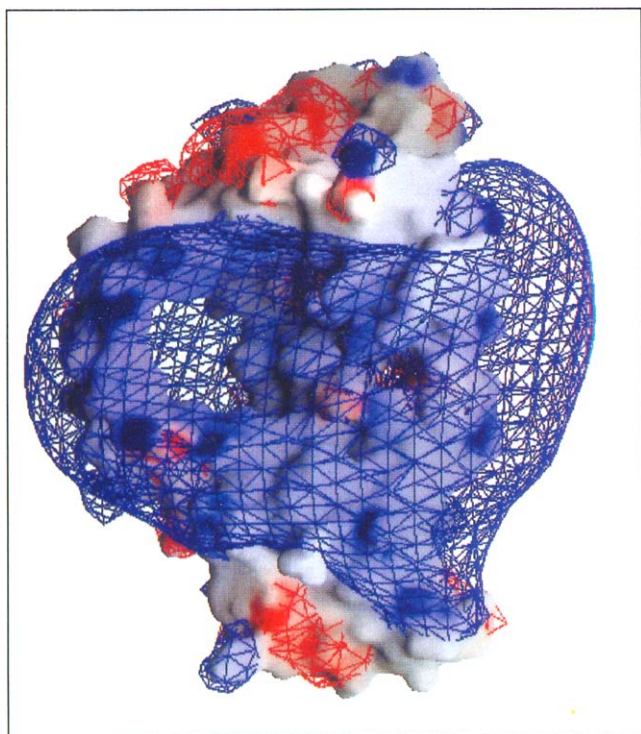

Fig. 2. Electrostatic potential map of the fragment. The surface of the protein is shown with positively and negatively charged residues shown in blue and red respectively. The grid surrounding the fragment corresponds to contours of positive (blue) and negative (red) potential at ± 2 $K_b T/e$ (where K_b is the Boltzmann constant, T is the temperature, and e is the charge of the electron). This figure clearly shows the positive potential region around one face of the protein. (Figure drawn with GRASP [33].)

not shown), yet this portion of eukaryotic DNA topoisomerase I appears be be required for enzymatic activity. To test the possibility that this domain is involved in DNA binding in the intact enzyme, photocrosslinking was carried out between a ^{32}P-labeled DNA fragment containing a photoactivatable base (see the Materials and methods section for details) [17] and a catalytically active yeast enzyme immunotagged at its N terminus. The product was first digested with staphylococcal nuclease to remove the bulk of the DNA, partially trypsinized, precipitated with antibodies specific to the N-terminal immunotag, and analyzed by SDS-PAGE. Figure 3a shows that the C terminus of the smallest ^{32}P-labeled immunotagged fragment is near Cys232, which is located at the beginning of α helix C and is well within the 26 kDa fragment. Thus, in the functional enzyme, the 26 kDa domain appears to contact DNA.
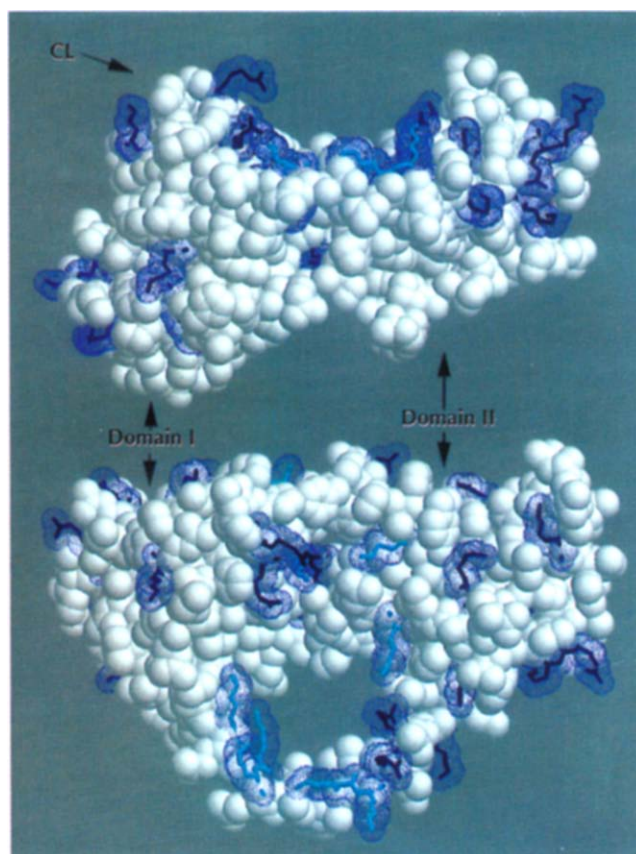
Protein footprinting

To better define the nature of the interaction between the 26 kDa fragment and DNA, we mapped the positions of lysine residues that are most probably involved in binding DNA and thereby protected from modification by citraconic anhydride [18]. Figure 3b shows that lysines 245, 248, 249, 256 and 257 in α helices C and D, 286, 301, 304 and 306 in the CL loop, and 357 in domain II are protected. The protected lysines are present on one side of the 26 kDa fragment (shown in Fig. 4) in both domains I and II and the V-shaped α helices. The asymmetric distribution of lysine residues confers a high positive electrostatic potential [16], appropriate for interaction with negatively charged DNA.
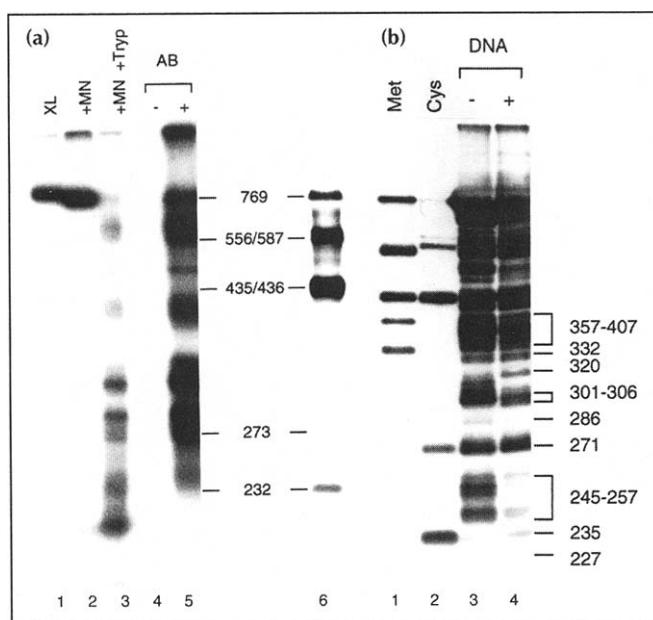
Fig. 3. Binding of the 26 kDa fragment to DNA. **(a)** Photo-crosslinking of DNA and catalytically active yeast DNA topo-isomerase I immunotagged at its N terminus. Lane 1, photocrosslinking reaction products; lane 2, reaction products digested with staphylococcal nuclease; lane 3, nuclease-treated products after partial digestion with trypsin; lane 4, products of mock immunoprecipitation in which the monoclonal antibody against the N-terminal immunotag was omitted; lane 5, immuno-precipitated peptides using monoclonal antibody directed against the N-terminal epitope; lane 6, size markers (see Materials and methods for details). **(b)** Detection of specific lysines in yeast DNA topoisomerase I that are protected against citracony-lation by DNA. Lanes 1 and 2 contained samples serving as length markers; lane 1, immunotagged yeast DNA topoisomerase I subjected to complete acetylation of lysine residues and partial cleavage by cyanogen bromide, and lane 2 the same protein subjected to complete acetylation and partial cleavage by 2-nitro-5-thiocyanatobenzoic acid. Lanes 3 and 4 contained immunotagged yeast DNA topoisomerase I footprinted by the two-step lysine-modification method [18] in the absence and presence of DNA, as indicated at the top of the lanes.

Fig. 4. Footprinted lysine residues. Orthogonal views of the van der Waals surface of the 26 kDa fragment in which all lysine and arginine residues are in blue and have blue dotted surfaces, and the side chains of lysines shown to interact with DNA by the protection experiments are in cyan with blue dotted surfaces. These views show that the concave surface is lined with positively charged residues.

Model for protein–DNA interaction

Considered together, the biochemical and crystallographic data suggest a plausible model for interaction between the 26 kDa fragment and DNA. DNA can be docked onto the fragment such that the backbone phosphates are close to the footprinted lysines. One face of the DNA contacts two protein regions separated by about 25 Å (shown in Fig. 5a,b): one region includes domain II and the CL loop; the second region includes the distal part of the V-shaped α helices. In this model, the DNA lies in a region of positive electrostatic potential. Because domains I and II are joined by two potentially flexible regions (the V-shaped α helices and the polyproline helix) small interdomain movements may occur to improve the fit between the protein and DNA or to accommodate local structural variations in the DNA helix; relative displacements between the outlying V-helices and the concave platform may serve similar purposes.

Drug-resistant mutants

In the model shown, the CL loop is also in close proximity to DNA. Biochemical studies have identified two mutations in the N-terminal region of mammalian DNA topoisomerase I that diminish camptothecin sensitivity [19,20]. These mutations map to the CL loop in the *S. cerevisiae* protein (Phe293 and Gly295). The residues forming the CL loop are highly conserved between yeast and mammalian DNA topoisomerase I. Because camptothecin is known to stabilize the covalent intermediate between DNA and eukaryotic DNA topoisomerase I, the location of these mutations suggests that the CL loop might be near the active site and may participate in camptothecin binding as well as in DNA cleavage and rejoining.

Comparison with other type I enzymes

The very high amino acid sequence similarity (>50%) [9] among enzymes of the eukaryotic DNA topoisomerase I type implies a strong structural conservation. It is very likely that the two-domain architecture, the β hairpin loop (CL), the V-shaped α helices, and the asymmetric charge distribution are general characteristics of all enzymes of this subfamily rather than specific features of

Fig. 5. Model for the possible interaction of DNA with the 26 kDa fragment of *S. cerevisiae* DNA topoisomerase I based on structural and biochemical data. **(a)** A B-DNA molecule (in gold) was modeled to lie on the concave side of the protein and in the region of positive electrostatic potential. The 10 protected lysines (green), additional lysine and arginine residues in the region (blue) and Phe293 (one of the residues involved in camptothecin resistance) (red) are in close proximity to the modeled DNA. The lysines are shown in the conformations observed in the crystal structure; local adjustments are likely when the enzyme interacts with DNA. (Drawn using SETOR [34].) **(b)** Surface drawing of the protein showing the good fit of the DNA to the concave side of the structure. The ε-carbon and amino group of lysines that are protected from citraconylation by DNA are highlighted in green. Residues affecting camptothecin resistance when mutated are highlighted in red. (Drawn with GRASP [33].)
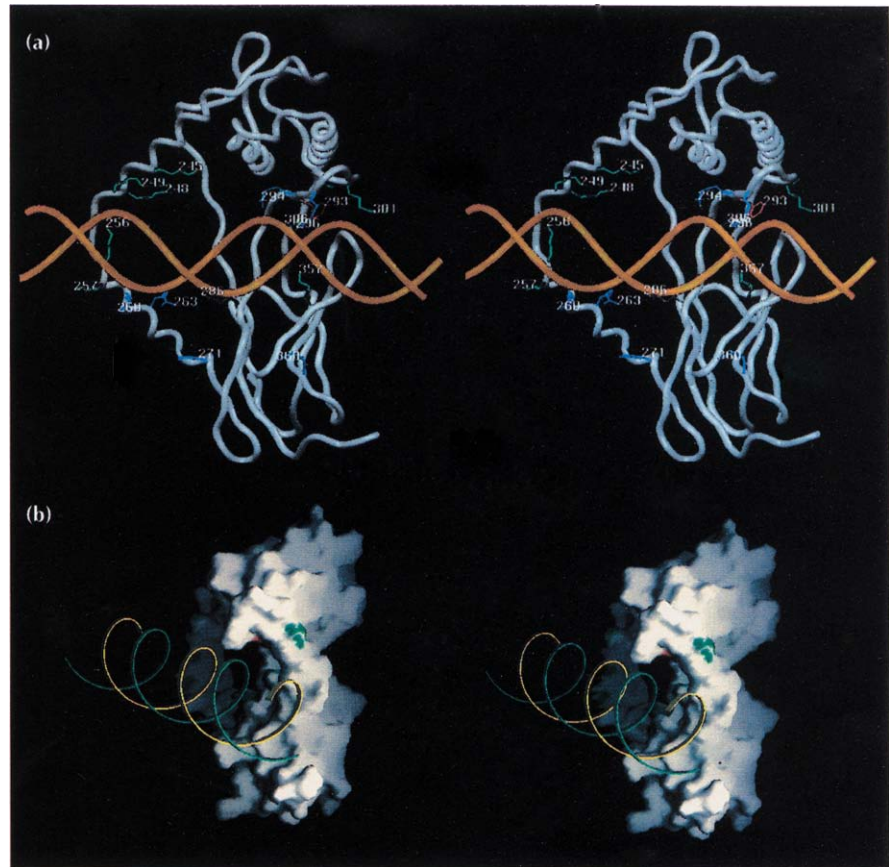
the yeast enzyme. The 26 kDa yeast topoisomerase I fragment shows no structural similarity to the 67 kDa N-terminal fragment of *E. coli* DNA topoisomerase I [7], a member of a separate subfamily of type I DNA topoisomerases. A comparison of the structure of the 26 kDa fragment with that of a 9 kDa N-terminal fragment of vaccinia virus topoisomerase [8] also shows that they are structurally distinct.

Biological implications

Eukaryotic DNA topoisomerase I is a ubiquitous enzyme that participates in transcription, replication and recombination [5]. It is also the target of the antitumor alkaloid camptothecin [6]. A combined crystallographic and biochemical analysis of a 26 kDa fragment (residues 135 to about 363) of yeast DNA topoisomerase I indicates that this domain is involved in DNA–protein complex formation and probably in camptothecin–enzyme–DNA interactions as well.

Interaction with DNA occurs on one surface of the fragment that has a strong positive electrostatic potential, suggesting the importance of electrostatic forces in the interaction. Two separate regions of the 26 kDa fragment are involved in DNA binding, and a comparison of the structure in two different crystal forms suggests that there is some conformational flexibility within and between the two DNA-binding regions. This flexibility may facilitate interactions between the enzyme and DNA substrates of different sequences or conformations.

Two amino acid residues have been identified in human DNA topoisomerase I (equivalent to Phe293 and Gly295 in the yeast 26 kDa fragment) that confer camptothecin resistance when mutated [19,20]. These residues map to a loop that constitutes part of the DNA-binding surface. Because the drug is thought to stabilize the covalent protein–DNA intermediate, it is plausible that this loop may be near the active site for DNA cleavage and rejoining.

Materials and methods
Protein expression, purification and crystallization
A 45 kDa recombinant protein consisting of the first 383 amino acids of yeast DNA topoisomerase I plus a nonapeptide immunotag at its N terminus was overexpressed in yeast and purified. The purified protein was subjected to partial proteolysis with subtilisin to yield a 35 kDa protein with Thr101 as its N terminus. This was further proteolyzed by chymotrypsin to yield a stable 26 kDa fragment with Glu135 at its N terminus. The 26 kDa protein was purified and stored at 4°C at a concentration of 10–20 mg ml⁻¹ in 10 mM Hepes (pH 7.5), 30 mM KCl. Orthorhombic crystals of the 26 kDa fragment were grown by vapor diffusion at 18°C against 20% PEG 3350, 0.1 M Tris-HCl (pH 8.0). Typical orthorhombic crystals are

$75 \times 75 \times 250$ μm^3 in size. Tetragonal crystals were grown from the same precipitant plus 0.4 M ammonium sulfate; macroseeding significantly improved the size of these crystals [21].

Data collection

Diffraction data from both crystal forms were collected using a Xentronics/Siemens area detector [22]. Orthorhombic crystals were cooled to $-1°C$ for data collection; data for the tetragonal crystals were collected at room temperature. All data were processed using the programs XDS and XSCALE [23,24]. High-resolution, cryo-cooled native data were collected at the Stanford Synchrotron Radiation Laboratory using the MAR imaging plate system. For cryo-cooling, single orthorhombic crystals were transferred over a period of 6 min to 18% sorbitol plus mother liquor in 6% steps. Crystals were scooped out using rayon loops and flash frozen in a $-170°C$ nitrogen stream. Data for cryo-cooled crystals were indexed and reduced using the program DENZO [25]. All subsequent data processing was done using the CCP4 suite [26]. Statistics are summarized in Table 1.

Heavy-atom derivatives and phasing

Orthorhombic crystals were soaked in 2 mM mercuric acetate for 8 days, and in 10 mM K_2PtCl_6 for 36 h. Several derivatives were tried but these two heavy-atom compounds gave the most interpretable difference Patterson maps. The position of the single mercury site was determined from difference Patterson and anomalous difference Patterson maps. The phases calculated from this derivative were used to locate the two platinum sites by difference Fourier synthesis. MLPHARE [27] was used for heavy-atom refinement and phasing.

Model building and refinement of the orthorhombic crystal form

An MIR map from data between 10 Å and 2.8 Å was calculated and improved by several cycles of solvent flattening [26]. A Cα trace was built using the BONES option of O [28]; β strands and α helices were discernible in the starting map. The model was broken into four discontinuous polyalanine segments containing a total of 187 residues; all residues with ambiguous densities and connectivities were left out of this

initial model. Model refinement of the fragment was done with the program X-PLOR [29]. Five rounds of positional refinement were carried out against native data in the range of 10–2.8 Å. The starting R-factor was 44.5% and the R_{free} was 54.7%. The R_{free} evaluation test was used throughout the course of the refinement as a measure of the correctness of the model [30]. Combined model and MIR phases were calculated using SIGMAA [26] and new electron-density maps calculated. These combined maps were used to thoroughly rebuild the initial model. After several cycles of rebuilding and refinement, side chains and residues were added to the model, further reducing the R-factor to 31.2% and the R_{free} to 46.4%. At this stage there were 199 residues in total, in three segments. However, the model refined no further, probably due to the weakness of the native data set. A new native data set was collected at SSRL using cryo-cooled protein crystals. Rigid-body refinement of the protein model against the cryo-cooled data in the resolution range of 8–4 Å gave an initial R-factor of 31% and an R_{free} of 34.1% (the set of test reflections was different for each data set). This was followed by simulated annealing slow-cool refinement (5–2.8 Å) [29] which reduced the R-factor to 26.3% and increased the R_{free} to 36.3%. Extensive model rebuilding was done using $2F_o-F_c$ and F_o-F_c electron-density maps. Side chains and residues were added, and a total of 220 residues were finally modeled into the electron density. As the refinement started to converge, and once the R-factor was below 25%, water molecules were added using the water-pekpik option of O. The native data were corrected for anisotropic temperature factors using the 'baoverall' option of X-PLOR. The model was checked periodically using the program PROCHECK [31], and outliers fixed. Finally, individual atomic temperature factors were refined. The final model was confirmed by a series of simulated annealing omit maps (Fig. 6a), which were computed by sequentially omitting from the structure a sphere, every twentieth residue, of 8 Å. The final R-factor for data between 8–1.9 Å is 22.2% with an R_{free} of 29.0%. There are 220 residues in the model and 232 water molecules. The main-chain density in $2F_o-F_c$ maps, an example of which is shown in Figure 6, is broken at the intersection of the two V-shaped α helices, and therefore residues 252–254 have not been
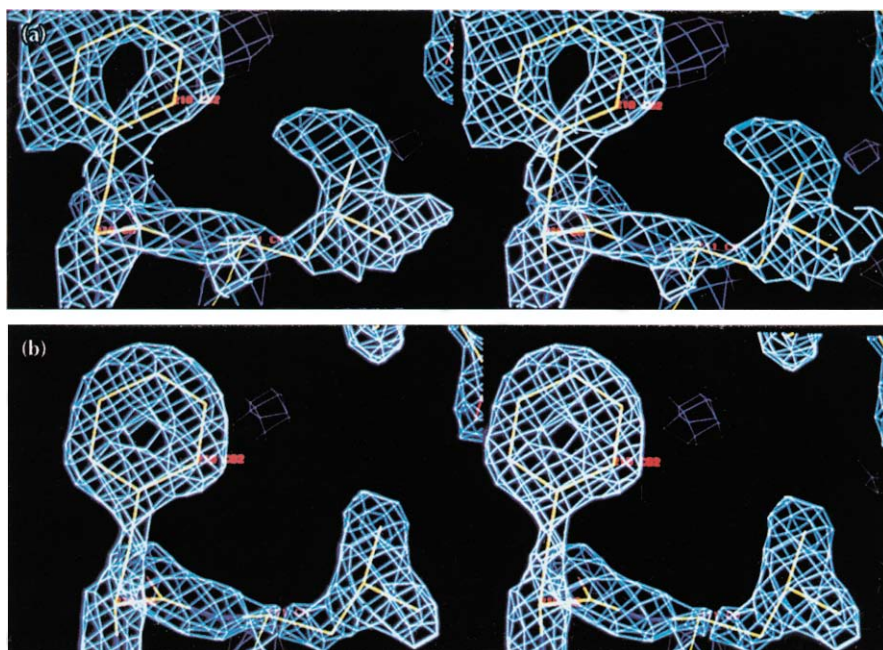
Fig. 6. Electron-density maps. **(a)** Simulated annealing omit map. The 1.9 Å map was calculated by removing all atoms inside an 8 Å sphere centered on residue 210. Similar maps were calculated covering the whole molecule to assess the quality of the structure. Contours correspond to the 0.95σ level. **(b)** A $2F_o-F_c$ map contoured at the 2.25σ level showing the region around residue 210.

modeled (dotted lines in Fig. 1b). The stereochemical quality of the final model is excellent, with an rmsd from ideality in bond lengths of 0.014 Å, and in angles of 1.73°. The average temperature factor for all atoms is 27 Å^2 and there are no outliers in the Ramachandran plot.

Refinement of the tetragonal crystal form

The tetragonal crystal form of the same fragment was solved primarily to confirm the model derived from the orthorhombic crystal form. This was done using an intermediate model of the orthorhombic crystal form. The initial solution for the rotation and translation functions gave an R-factor of 48.3% and a correlation coefficient of 0.47 using the program AMoRe [26]. Rigid-body refinement of this solution reduced the R-factor to 46.1% and gave a correlation coefficient of 0.55. At this stage the lattice packing of the model was checked, and a few residues from the N and C termini which could sterically clash with symmetry-related molecules were removed. Several cycles of positional refinement using X-PLOR were done in the resolution range of 7–3 Å, without modeling any water molecules, leading to an R-factor of 20.4%. The coordinates of the orthorhombic crystal model are being deposited in the Brookhaven Protein Data Bank.

Photocrosslinking

DNA with one 20 nucleotide strand and one 14 nucleotide strand recessed at its 3′-end was used as the substrate for repair synthesis by the Klenow fragment of *E. coli* DNA polymerase I. Triphosphates of a photoreactive nucleotide 5-[*N*-(*p*-azidobenzoyl)-3-aminoallyl]-deoxyuridine [17] and α-^{32}P-deoxyadenosine were used in the repair of the 5′-T$_5$A overhang. A catalytically active yeast DNA topoisomerase I derivative, with an 11-residue phage T7 gene 10 peptide fused to amino acids 139–769 of yeast DNA topoisomerase I, was overproduced and purified from *E. coli*. The immunotagged enzyme (18 μg) and labeled DNA (1.6 μg) were incubated in 100 μl of a buffer containing 20 mM Tris-HCl, pH 8.0, 50 mM KCl, 1 mM 2-mercaptoethanol, 10% (v/v) glycerol, and 100 μg of bovine serum albumin as a carrier protein. Following UV irradiation, CaCl$_2$ was added to a final concentration of 1 mM and the mixture was digested with 1 unit of staphylococcal nuclease (Boehringer Mannheim, Indianapolis, IN) at room temperature for 10 min. Nuclease digestion was terminated by the addition of EGTA to 10 mM. Following incubation with trypsin (10 ng) for 20 min at 30°C, the partially trypsinized products were immunoprecipitated using a monoclonal antibody directed against the T7 gene 10 epitope (Novagen, Madison, WI). Immunoprecipitated polypeptides that became radiolabeled by their crosslinking to DNA were visualized by autoradiography following SDS-PAGE. Aliquots of the reaction mixtures were sampled at various steps for comparison with the immunoprecipitated final products. Size markers (lane 6, Fig. 3a) were prepared by partial cleavage of immunotagged yeast DNA topoisomerase I with 2-nitro-5-thiocyanatobenzoic acid (which cleaves at cysteine residues), followed by fractionation on the same polyacrylamide gel, and immunochemical visualization.

Protein footprinting

The chemical footprinting reactions employing citraconic anhydride, *N*-hydroxysuccinimide, and endoproteinase Lys-C, were carried out using the immunotagged yeast DNA topoisomerase I as previously described [18] except that, after treatment with citraconic anhydride, the reaction mixture was diluted with three volumes of 8 M guanidine hydrochloride before the acetylation reaction.

References

1. Wang, J. (1985). DNA topoisomerases. *Annu. Rev. Biochem.* **54**, 665–697.
2. Champoux, J.J. (1990). Type I topoisomerase mechanisms. In *DNA Topology and Its Biological Effects.* (Cozzarelli, N.R. & Wang, J.C., eds), pp. 217–242, Cold Spring Harbor Laboratory Press, New York.
3. Shuman, S. & Moss, B. (1987). Identification of a vaccinia gene encoding a type I DNA topoisomerase. *Proc. Natl. Acad. Sci. USA* **84**, 7478–7482.
4. Slesarev, A.I., Lake, J.A., Stetter, K.O., Gellert, M. & Kozyavkin, S.A. (1994). Purification and characterization of DNA topoisomerase V. *J. Biol. Chem.* **69**, 295–303.
5. Wang, J. (1991). DNA topoisomerases: why so many? *J. Biol. Chem.* **266**, 6659–6662.
6. Liu, L.F. (1989). DNA topoisomerase poisons as antitumor drugs. *Annu. Rev. Biochem.* **58**, 351–375.
7. Lima, C.D., Wang, J.C. & Mondragón, A. (1994). Three-dimensional structure of the 67K N-terminal fragment of *E. coli* DNA topoisomerase I. *Nature* **367**, 138–146.
8. Sharma, A., Hanai, R. & Mondragón, A. (1994). Crystal structure of the amino-terminal fragment of vaccinia virus DNA topoisomerase I at 1.6 Å resolution. *Structure* **2**, 767–777.
9. Caron, P.R. & Wang, J.C. (1994). Alignment of primary sequences of DNA topoisomerases. In *DNA Topoisomerases and Their Applications in Pharmacology. Adv. Pharmacol.* (Liu, L.F., ed), pp. 271–291, Academic Press, Boca Raton, FL.
10. Bjornsti, M.A. & Wang, J.C. (1987). Expression of yeast DNA topoisomerase I can complement a conditional-lethal DNA topoisomerase I mutation in *Escherichia coli. Proc. Natl. Acad. Sci. USA* **84**, 8971–8975.
11. Adzhubei, A.A. & Sternberg, M.J.E. (1993). Left-handed polyproline II helices commonly occur in globular proteins. *J. Mol. Biol.* **229**, 472–493.
12. Kong, X.-P., Onrust, R., O'Donnell, M. & Kuriyan, J. (1992). Three-dimensional structure of the β subunit of *E. coli* polymerase III holoenzyme: a sliding clamp. *Cell* **69**, 425–437.
13. Talluru, S.R., Kong, X.-P., Gary, S., Burgers, P.M. & Kuriyan, J. (1994). Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell* **79**, 1233–1243.
14. Nikolov, D.B., *et al.*, & Burley, S.K. (1992). Crystal structure of TFIID TATA-box binding protein. *Nature* **360**, 40–46.
15. Kim, Y., Geiger, J.H., Hahn, S. & Sigler, P.B. (1993). Crystal structure of a yeast TBP/TATA-box complex. *Nature* **365**, 512–520.
16. Nicholls, A. & Honig, B. (1991). A rapid finite difference algorithm, utilizing successive over-relaxation to solve the Poisson–Boltzmann equation. *J. Comput. Chem.* **12**, 435–445.
17. Bartholomew, B., Kassavetis, G.A., Braun, B.R. & Geiduschek, E.P. (1990). The subunit structure of *Saccharomyces cerevisiae* transcription factor IIIC probed with a novel photocrosslinking reagent. *EMBO J.* **9**, 2197–2205.
18. Hanai, R. & Wang, J.C. (1994). Protein footprinting by the combined use of reversible and irreversible lysine modifications. *Proc. Natl. Acad. Sci. USA* **91**, 11904–11908.
19. Benedetti, P., Fiorani, P., Campuani, L. & Wang, J.C. (1993). Camptothecin resistance from a single mutation changing glycine 363 of human DNA topoisomerase I to cysteine. *Cancer Res.* **53**, 4343–4348.
20. Rubin, E., Pantazis, P., Bharti, A., Toppmeyer, D., Giovanella, B. & Kufe, D. (1994). Identification of a mutant human topoisomerase I with intact catalytic activity and resistance to 9-nitro-camptothecin. *J. Biol. Chem.* **269**, 2433–2439.
21. Thaller, C., Eichele, G., Weaver, L.H., Wilson, E., Karlsson, R. & Jansonius, J.N. (1985). Diffraction methods for biological macromolecules: seeding enlargement and repeated seeding. *Methods Enzymol.* **114**, 132–135.
22. Blum, M., Metcalf, P., Harrison, S.C. & Wiley, D.C. (1987). A system for collection and on-line integration of X-ray diffraction data from a multiwire area detector. *J. Appl. Cryst.* **20**, 235–242.
23. Kabsch, W. (1988). Automatic indexing of rotation diffraction patterns. *J. Appl. Cryst.* **21**, 67–71.

24. Kabsch, W. (1988). Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector. *J. Appl. Cryst.* **21**, 916–924.

25. Otwinowski, Z. (1993). DENZO. In *Data Collection and Processing.* (Sawyer, L., Isaacs, N. & Bailey, S., eds), pp. 556–562, SERC Daresbury Laboratory, Warrington, UK.

26. Collaborative Computational Project, Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst.* D **50**, 760–763.

27. Otwinowski, Z. (1991). Maximum likelihood refinement of heavy atom parameters. In *Isomorphous Replacement and Anomalous Scattering.* pp. 80–86, SERC Daresbury Laboratory, Warrington, UK.

28. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst.* A **47**, 110–119.

29. Brünger, A.T., Kuriyan, J. & Karplus, M. (1987). Crystallographic R factor refinement by molecular dynamics. *Science* **235**, 458–460.

30. Brünger, A.T. (1992). The free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472–474.

31. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291.

32. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946–950.

33. Nicholls, A.J. (1993). *GRASP Manual.* Columbia University, New York.

34. Evans, S.V. (1993). SETOR: hardware-lighted three dimensional solid model representations of macromolecules. *J. Mol. Graphics* **11**, 134–138.