Solution Structure of the C-Terminal Single-Stranded DNA-Binding Domain of *Escherichia coli* Topoisomerase I†

Liping Yu,‡ Chang-Xi Zhu,§ Yuk-Ching Tse-Dinh,§ and Stephen W. Fesik*,‡

NMR Research, D-47G, AP10, Abbott Laboratories, 100 Abbott Park Road, Abbott Park, Illinois 60064, and Department of Biochemistry and Molecular Biology, New York Medical College, Valhalla, New York 10595

Received March 2, 1995; Revised Manuscript Received April 17, 1995⊗

ABSTRACT: *Escherichia coli* DNA topoisomerase I catalyzes the interconversion of different topological forms of DNA. In this paper we describe NMR studies of a 14K C-terminal fragment of this enzyme that binds preferentially to single-stranded DNA and enhances the enzyme's ability to relax negatively supercoiled DNA under high salt conditions. The ^{1}H, ^{13}C, and ^{15}N resonances of the protein were assigned from a number of heteronuclear multidimensional NMR experiments, and the three-dimensional structure of the protein was determined from a total of 2188 NMR-derived restraints. The root-mean-square deviation about the mean coordinate positions for residues 13−120 is 0.68 ± 0.11 Å for the backbone atoms and 1.09 ± 0.09 Å for all heavy atoms. The overall fold, which consists of two four-stranded β-sheets separated by two helices, differs from other DNA- and RNA-binding proteins such as gene 5, cold shock protein, and hnRNP C. From an analysis of the changes in chemical shift upon the addition of single-stranded DNA, the location of the oligonucleotide binding site was determined. The binding site consists of a β-sheet containing positively charged and aromatic amino acids and, in spite of its different structure, is similar to that found in other proteins that bind single-stranded oligonucleotides.

Escherichia coli DNA topoisomerase I is a 97K enzyme that catalyzes the interconversion of different topological forms of DNA by cleaving single-stranded DNA, passing single- or double-stranded DNA through the cleavage site, and rejoining the broken strands (Wang, 1985). The enzyme consists of three domains: a 67K N-terminal domain responsible for performing the DNA cleavage reaction, a zinc-binding domain which is required for relaxing supercoiled DNA, and a C-terminal DNA-binding domain. The 14K C-terminal fragment enhances DNA binding and improves the enzyme's ability to relax negatively supercoiled DNA under high salt conditions (Beran-Steed et al., 1989; Tse-Dinh, 1994). As with the intact enzyme, the C-terminal fragment interacts with both single- and double-stranded DNA but prefers to bind to single-stranded oligonucleotides (Beran-Steed et al., 1989; Tse-Dinh, 1994). Recently, the X-ray crystal structure of the 67K N-terminal fragment of *E. coli* topoisomerase I has been reported (Lima et al., 1994). However, no structural information is available for the rest of the enzyme.

A considerable amount is known about the three-dimensional structures and DNA-binding interactions of the individual protein motifs that bind to double-stranded DNA [for review, see Pabo and Sauer (1992)]. In contrast, much less structural information has been obtained on single-stranded DNA-binding proteins. Here we describe the three-dimensional structure of the C-terminal DNA-binding domain of *E. coli* DNA topoisomerase I in solution as determined by heteronuclear multidimensional NMR[1] spectroscopy. In addition, from an analysis of the chemical shift changes of the protein upon the addition of single-stranded DNA, we describe the regions of the C-terminal fragment of topoisomerase I that bind to single-stranded DNA.

MATERIALS AND METHODS

Sample Preparation. The 14K fragment of *E. coli* DNA topoisomerase I (residues 745−865) was cloned, expressed, and purified as previously described (Zhu et al., 1995). Uniformly ^{15}N- and ^{15}N-,^{13}C-labeled proteins were prepared for the NMR experiments by growing bacteria that overexpress the C-terminal domain of *E. coli* topoisomerase I in a minimal medium containing $^{15}NH_4Cl$ with or without [U-^{13}C]glucose. Selectively ^{15}N-Lys- and ^{15}N-Leu-labeled proteins were prepared using previously reported procedures (Hoffman & Spicer, 1991). The protein was purified by phosphocellulose P11 ion exchange, ssDNA agarose affinity, and reverse-phase HPLC columns followed by refolding by dialysis from 6 M urea at 4 °C. The samples were concentrated to 3 mM in a solution containing 5 mM dithiothreitol-d_{10} and 200 mM potassium phosphate at pH 6.5. The protein was found to be very stable in this buffer. No precipitation was observed, and the NMR spectrum of this protein was the same even after several months at 30 °C. Although structural characterization on the original proteolytic fragment was not carried out due to the difficulties in obtaining sufficient quantities, the refolded protein showed

† The coordinates for the final structures have been deposited with the Brookhaven Protein Data Bank under the filename 1YUA.

* To whom correspondence should be addressed.
‡ Abbott Laboratories.
§ New York Medical College.

⊗ Abstract published in *Advance ACS Abstracts,* June 1, 1995.

[1] Abbreviations: NMR, nuclear magnetic resonance; HSQC, heteronuclear single-quantum coherence; TPPI, time-proportional phase incrementation; 2D, two-dimensional; 3D, three-dimensional; TOCSY, total correlation spectroscopy; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; HPLC, high-performance liquid chromatography; HMQC, heteronuclear multiple-quantum coherence.

0006-2960/95/0434-7622$09.00/0

identical ssDNA-binding affinity as the original proteolytic fragment.

NMR Spectroscopy. All NMR spectra were acquired at 30 °C on a Bruker AMX 600 MHz NMR spectrometer. Quadrature detection in the indirectly detected dimensions was accomplished using the States−TPPI methods (Marion et al., 1989a). NMR spectra were processed and analyzed using in-house written software on Silicon Graphics computers.

The two-dimensional (2D) $^1H-^{15}N$ heteronuclear single-quantum coherence (HSQC) (Bodenhausen & Ruben, 1980) spectra of uniformly ^{15}N-labeled and selectively ^{15}N-labeled proteins were recorded using a spin-lock purge pulse for water suppression (Messerle et al., 1989).

The 3D HNCA experiment (Kay et al., 1990) and 3D HN-(CO)CA experiment (Bax & Ikura, 1991) were acquired with 55 × 32 × 1024 complex points using sweep widths of 1946 Hz (^{15}N, t_1), 4529 Hz ($^{13}C^{\alpha}$, t_2), and 10 000 Hz (^1H, t_3), respectively. A total of 32 and 16 scans per increment were used for the HNCA and HN(CO)CA experiments, respectively.

The 3D C(CO)NH-TOCSY experiment (Montelione et al., 1992; Logan et al., 1992; Grzesiek et al., 1993) was recorded with 55 × 40 × 1024 complex points using sweep widths of 1946 Hz (^{15}N, t_1), 9058 Hz (^{13}C, t_2), and 10 000 Hz (^1H, t_3), respectively. A total of 32 scans per increment and a mixing time of 10 ms were used. The 3D CBCANH experiment (Grzesiek & Bax, 1992) was acquired with 48 scans per increment and 40 × 40 × 1024 complex points using identical sweep widths as in the C(CO)NH-TOCSYexperiment.

The 3D HNCO experiment (Kay et al., 1990) and HN-(CA)CO experiment (Clubb et al., 1992) were recorded with 47 × 40 × 1024 complex points using sweep widths of 1946 Hz (^{15}N, t_1), 2264 Hz ($^{13}C'$, t_2), and 10 000 Hz (^1H, t_3), respectively. A total of 16 and 48 scans per increment were used in the HNCO and HN(CA)CO experiments, respectively. The 3D HCACO (Grzesiek et al., 1993) was acquired using sweep widths of 2264 Hz ($^{13}C'$, t_1), 4529 Hz ($^{13}C^{\alpha}$, t_2), and 10 000 Hz (^1H, t_3) and 45 × 30 × 1024 complex points with 32 scans per increment.

The 3D HCCH-TOCSY experiment (Bax et al., 1990) for the assignment of aliphatic side chains was collected using sweep widths of 10 570 Hz (^{13}C, t_1), 4545 Hz (^1H, t_2), and 10 000 Hz (^1H, t_3) over 64 × 128 × 1024 complex points with 16 scans per increment. A mixing time using a DIPSI-2 pulse sequence of 17.3 ms was employed. The ^{13}C carrier was set at 40 ppm. For assignment of aromatic ring spin systems, a similar HCCH-TOCSY experiment was performed with the carrier set at 125 ppm and with a mixing time of 7 ms. A total of 34 × 48 × 1024 complex points were collected with 16 scans per increment and using sweep widths of 5285 Hz (^{13}C, t_1), 3000 Hz (^1H, t_2), and 10 000 Hz (^1H, t_3), respectively. The 2D CCH-COSY experiment (Pelton et al., 1993) was acquired with the carbon carrier set at 85 ppm and with 256 scans per increment using a sweep widths of 20 660 Hz (^{13}C, t_1) and 10 000 Hz (^1H, t_2).

The 3D ^{15}N-edited NOESY-HSQC experiment (Fesik & Zuiderweg, 1988; Marion et al., 1989b) was performed with a mixing time of 100 ms. The data were collected as 55 × 96 x 1024 complex points over sweep widths of 2136 Hz (^{15}N, t_1), 7692 Hz (^1H, t_2), and 10 000 Hz (^1H, t_3) in 32 scans per increment. The 3D ^{15}N-edited TOCSY-HSQC experiment (Marion et al., 1989c) and 3D HNHB experiment (Archer et al., 1991) were collected using spectrometer conditions identical to the ^{15}N-edited NOESY experiment. A 26-ms clean MLEV mixing sequence (Griesinger et al., 1988) was used in the TOCSY-HSQC experiment to effect $^1H-^1H$ magnetization transfer.

The 3D ^{13}C-edited NOESY-HMQC experiment (Fesik & Zuiderweg, 1988) was collected with 100-ms mixing time. Gradients were used to suppress artifacts and the residual water signal (Bax & Pochapsky, 1992). The carbon carrier was placed at 40 ppm. The data were collected as 54 × 140 × 1024 complex points using sweep widths of 10 570 Hz (^{13}C, t_1), 6250 Hz (^1H, t_2), and 10 000 Hz (^1H, t_3) and with 16 scans per increment.

The 2D HMQC-J experiment (Kay & Bax, 1990) and 3D HNHA experiment (Vuister & Bax, 1993) were acquired to obtain $^3J_{HN,H^{\alpha}}$ coupling constants. The HNHA experiment was collected in 32 scans per increment with 34 × 64 × 1024 complex points using sweep widths of 1946 Hz (^{15}N, t_1), 6250 Hz (^1H, t_2), and 10 000 Hz (^1H, t_3), respectively.

Structure Calculations. Structure calculations employed a distance geometry/simulated annealing (DG/SA) protocol (Nilges et al., 1988; Kuszewski et al., 1992) using the X-PLOR program (Brünger, 1992). A linear representation of the polypeptide chain was used as the initial structure for all calculations. A total of 2113 distance restraints were derived from ^{15}N- and ^{13}C-resolved 3D NOESY spectra acquired using a 100-ms mixing time. The distance restraints were given upper bounds of 2.7, 3.3, 4.0, and 5.0 Å on the basis of the measured NOE cross-peak intensity. An additional 0.5 Å was added to the upper distance limits for NOEs involving methyl protons. Lower bounds for all NOE-derived restraints were set to 1.8 Å. Of these 2113 restraints, 514 were intraresidue, 501 were sequential, 302 were between residues separated by less than five residues in the primary sequence, and 796 were from long-range NOEs. A correction for center averaging was applied where appropriate (Wüthrich et al., 1983). A total of 33 hydrogen bonds from an analysis of the amide exchange rates measured from a series of $^1H/^{15}N$ HSQC spectra recorded after the addition of ^2H_2O were included and given bounds of 1.8−2.3 Å (H → O) and 2.7−3.3 Å (N → O). In addition, 42 ϕ angle restraints from the $^3J_{HN,H^{\alpha}}$ coupling constants were measured in HNHA and HMQC-J experiments and were included in structure calculations. Restraints were included for those angles that exhibited $^3J_{HN,H^{\alpha}}$ coupling constants of >9.0 Hz (−120 ± 30°) or <5.5 Hz (−60 ± 30°) in both NMR experiments. Since $^3J_{HN,H^{\alpha}}$ values of <5.5 Hz may correspond to more than one ϕ angle, the ϕ angles that were determined from small coupling constants were used as restraints only if they appeared in the regions of α-helical secondary structure as determined from the NOE data.

DNA-Binding Experiments. The single-stranded DNAs used in the binding experiments, d(ACGACAGGCTAC), d(GGCTAC), and d(CTCC), were synthesized on a DNA synthesizer (Applied Biosystems) and purified by reverse-phase HPLC. For determining the DNA-binding affinity to the C-terminal domain of *E. coli* DNA topoisomerase I, a series of $^1H-^{15}N$ HSQC spectra were collected on uniformly ^{15}N-labeled protein at about 0.5 mM as a function of added DNA. A nonlinear least-squares fitting program was used to fit the chemical shift changes of the protein vs the

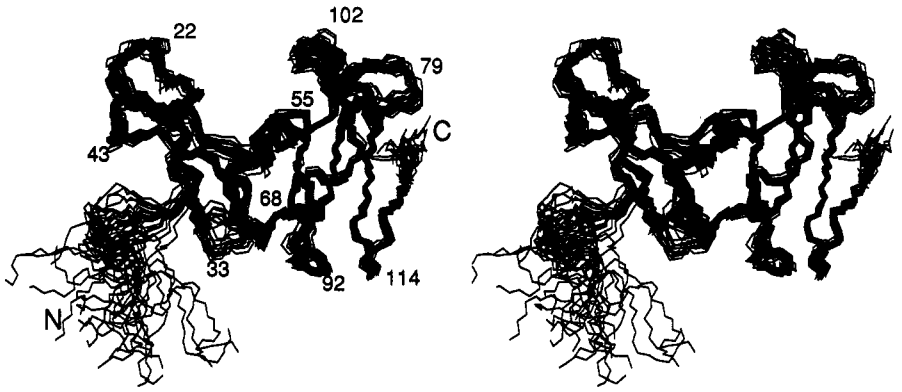

FIGURE 1: Stereoview of the backbone (N, C^{α}, C′) of 26 superimposed NMR-derived structures of the C-terminal single-stranded DNA-binding domain of *E. coli* topoisomerase I (residues 1−122).

concentration of DNA to obtain the disassociation constants (K_D).

RESULTS AND DISCUSSION

Assignments. The ^1H, ^{13}C, and ^{15}N resonances of the C-terminal domain of topoisomerase I were assigned from an analysis of several heteronuclear multidimensional NMR spectra of uniformly ^{15}N- and ^{15}N-,^{13}C-labeled protein (Bax & Grzesiek, 1993). In order to identify the adjacent amino acid spin systems, the ^1H and ^{15}N amide frequencies were correlated to the C^{α} and C^{β} frequencies of the i and $i-1$ amino acids from 3D HNCA, HN(CO)CA, CBCANH, and C(CO)NH-TOCSY experiments. The amides were assigned by amino acid type from HSQC spectra of selectively ^{15}N-Lys- and ^{15}N-Leu-labeled proteins and from a chemical shift analysis of the side-chain resonances obtained from 3D C(CO)NH-TOCSY, CBCANH, and HNHB experiments. In order to complete the backbone assignments, the H^{α} frequencies were correlated to the ^1H and ^{15}N signals of the amides from 3D ^{15}N-resolved TOCSY-HSQC and HNHA experiments. The backbone assignments were confirmed from the correlations observed in HCACO, HNCO, and HN(CA)CO experiments. The ^1H and ^{13}C signals of the side chains were assigned by correlating H^{α}/C^{α} signals of the backbone to the ^1H and ^{13}C signals of the side chains in a 3D HCCH-TOCSY experiment and confirmed from the correlations observed in 3D C(CO)NH-TOCSY, CBCANH, and HNHB experiments. The ^1H and ^{13}C signals of the aromatic rings were assigned using a 3D HCCH-TOCSY spectrum collected with the carbon carrier set at 125 ppm. These signals were linked to the H^{β} signals of the aromatic amino acids using a 2D CCH-COSY experiment and to the C^{α}/H^{α} and C^{β}/H^{β} from the NOE cross peaks observed in a ^{13}C-edited NOESY-HMQC spectrum. Stereospecific assignments of the Val and Leu methyl groups were obtained by biosynthetic fractional ^{13}C-labeling (Neri et al., 1989). The ^1H, ^{13}C, and ^{15}N assignments of the C-terminal domain of topoisomerase I are given as supplementary material. In general, the chemical shifts (H^{α}, C^{α}, C′, and C^{β}) of the C-terminal DNA-binding domain of topoisomerase I were consistent with the secondary structural elements (Wishart & Sykes, 1994) determined from the NOE data.

Structure Determination. The three-dimensional structure of the C-terminal DNA-binding domain of topoisomerase I was determined using a distance geometry/simulated annealing protocol (Nilges et al., 1988; Kuszewski et al., 1992) from a total of 2188 NMR-derived restraints including 2113 NOE-derived distance restraints, 33 hydrogen bonds, and 42 ϕ angular restraints. A superposition of the 26 low-energy NMR structures is shown in Figure 1. The structure statistics are given in Table 1. As shown in Figure 1, the structure of the backbone is well-defined by the NMR data except for the first 12 residues at the N-terminus and the last two residues at the C-terminus. The atomic root-mean-square deviation (rmsd) about the mean coordinate positions for residues 13−120 is 0.68 ± 0.11 Å for the backbone atoms and 1.09 ± 0.09 Å for all heavy atoms. If one loop (residues 100−107) and one turn (residues 32−35) are excluded, the rmsd drops even further (Table 1). Moreover, the C-terminal

Table 1: Structural Statistics and Root-Mean-Square Deviation for 26 Structures of the C-Terminal Domain of *E. coli* DNA Topoisomerase I[a]

structural statistics	⟨SA⟩	$\langle\overline{SA}\rangle_r$
X-PLOR energies (kcal/mol)		
E_{total}	336 ± 7	325
E_{vdw}[b]	19 ± 3	16
E_{cdih}[c]	0.5 ± 0.5	0.6
E_{NOE}[d]	27 ± 3	23
rmsd from idealized geometry		
bonds (Å)	0.00 ± 0.00	0.00
angles (deg)	0.64 ± 0.00	0.63
impropers (deg)	0.57 ± 0.03	0.57
Cartesian coordinate rmsd (Å)	backbone	all non-H
⟨SA⟩ vs $\langle\overline{SA}\rangle$[e]	0.68 ± 0.11	1.09 ± 0.09
⟨SA⟩ vs $\langle\overline{SA}\rangle$[f]	0.51 ± 0.06	1.01 ± 0.05
⟨SA⟩ vs $\langle\overline{SA}\rangle$[g]	0.39 ± 0.13	0.89 ± 0.12

[a] ⟨SA⟩ is the ensemble of 26 final X-PLOR structures generated from the DG/SA protocol (Nilges et al., 1988; Kuszewski et al., 1992) as described in the X-PLOR manual (Brünger, 1992). $\langle\overline{SA}\rangle$ is the Cartesian coordinates obtained by averaging ⟨SA⟩ following a least-squares superposition of the backbone heavy atoms for residues 13−120. $\langle\overline{SA}\rangle_r$ is the energy-minimized averaged Cartesian coordinates. [b] The X-PLOR F_{repel} function was used to simulate van der Waals interaction with atomic radii set to 0.8 times their CHARMM values (Brooks et al., 1983). [c] Torsional restraints were applied to 42 ϕ angles with bounds of $-120 \pm 30°$ for those angles with $^3J_{HN,H^{\alpha}}$ coupling constants >9.0 Hz and $-60 \pm 30°$ for coupling constants <5.5 Hz. Restraints for the latter were only applied in the α-helical region. Force constants of 200 kcal mol^{-1} rad^{-2} were used for all torsional restraints. No torsional restraints were violated >3° in any of the final structures. [d] A total of 2146 NMR-derived distance restraints including 2113 NOE-derived distance restraints and 33 hydrogen bonds were applied with a square well potential and a force constant of 50 kcal mol^{-1} $Å^{-2}$. No distance restraint was violated by >0.3 Å in any of the final structures. [e] rmsd for residues 13−120. [f] rmsd for the residues (13−75) located in the top half of the molecule, excluding the residues in the β-turn (32−35). [g] rmsd for the residues (76−120) located in the bottom half of the molecule, excluding the residues in the loop (100−107).

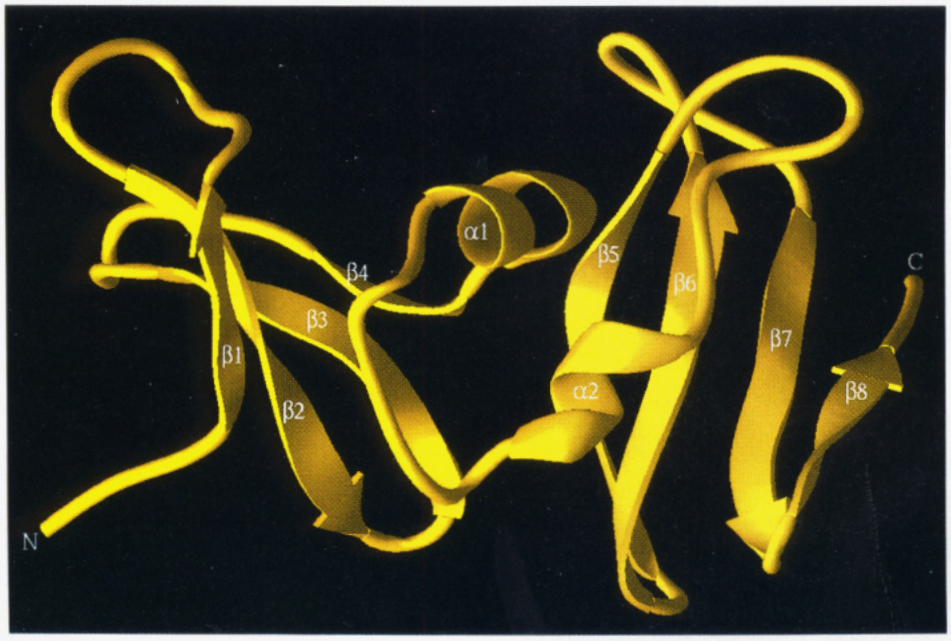

FIGURE 2: Ribbon plot (Carson, 1987) depicting the averaged minimized NMR structure (residues 9−122) which consists of eight β-strands (β1, 13−16; β2, 26−32; β3, 35−40; β4, 49−51; β5, 84−90; β6, 93−99; β7, 108−113; β8, 116−119) and two α-helices (α1, 53−59; α2, 69−73).

β-sheet (β5−β8) is slightly better defined than the other regions with a rmsd of 0.23 ± 0.11 Å for the backbone atoms and 0.80 ± 0.12 Å for all heavy atoms.

Description of Structure. The structure of the C-terminal fragment of *E. coli* topoisomerase I consists of two four-stranded antiparallel β-sheets that are separated by two α-helices (Figure 2). The molecule exhibits pseudo C_2 symmetry centered between the two helices (Figure 2). A hydrophobic core is formed between one side of the N-terminal β-sheet (residues V13, L15, L18, F27, L29, V36, and L38) and the helix−loop−helix (residues L57, F60, L68, and L71) (Figure 3). Several hydrophobic residues (P14, P16, Y26, V28, F37, and F43) are also found on the other side of the N-terminal β-sheet together with R30, R47, and R50. A second hydrophobic core is formed by the C-terminal β-sheet (V86, V97, W108, Y112, and W117) and the two helices (V54 and Y70) (Figure 4). There are 11 proline residues in this protein. However, only one cis peptide bond is observed between F43 and P44, which forms part of the loop connecting strands 3 and 4. The F43 aromatic ring is pointed toward the DNA binding cleft (see below). The two helices in the middle of the protein cross at an angle of 123°, which is similar to the interhelical angle observed in the classical helix−turn−helix DNA-binding motif (Pabo & Sauer, 1992). However, the second helix in the C-terminal fragment of topoisomerase I is located in a lower position and is shorter (five residues) compared to the helix−turn−helix motif commonly found in other DNA-binding proteins (Pabo & Sauer, 1992).

Binding to Single-Stranded DNA. The binding affinity of the C-terminal fragment of topoisomerase I to single-stranded DNA of various lengths was measured from the changes in the ^1H and ^{15}N amide chemical shifts of the protein as a function of added DNA. Under low salt conditions (10 mM potassium phosphate, pH = 6.5), the 12-mer (ACGACAG-GCTAC) exhibited a slightly higher affinity (K_D = 4 μM) for the protein compared to the 6-mer (GGCTAC) (K_D = 14 μM), but both oligonucleotides bind much more tightly

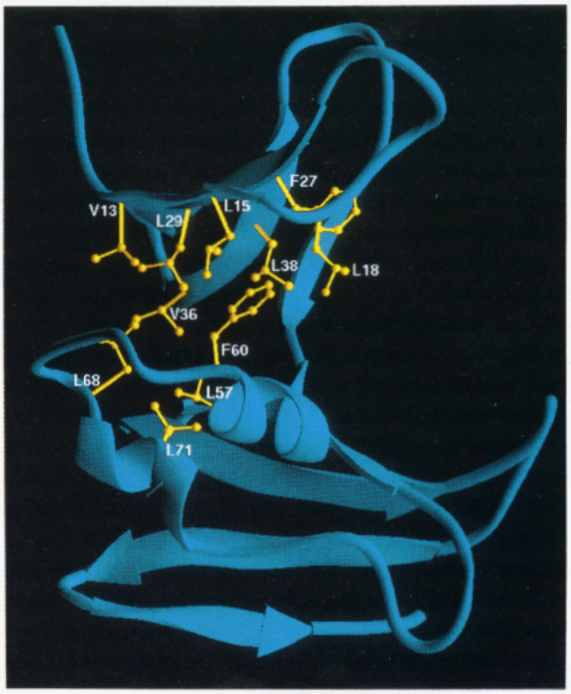

FIGURE 3: Ribbon plot of the C-terminal fragment of *E. coli* topoisomerase I (residues 9−122). The residues that form the N-terminal hydrophobic core are indicated in gold.

than the 4-mer (CTCC) (K_D = 950 μM) (Table 2). These results are consistent with a previous report of a site size of 4.5 nucleotides per protein and UV-crosslinking studies in which binding affinities of 1.2 and 0.35 μM were obtained for $p(dT)_8$ and $p(dT)_{16}$, respectively (Zhu et al., 1995).

The amino acid residues of the protein that experience the largest amide chemical shift changes upon the addition of single-stranded DNA were predominantly located on one side of the first β-sheet (Figures 5 and 6). This portion of the protein contains positively charged residues (R30, R47, and R50) which could form favorable interactions with the negatively charged phosphates of DNA. The side chains of

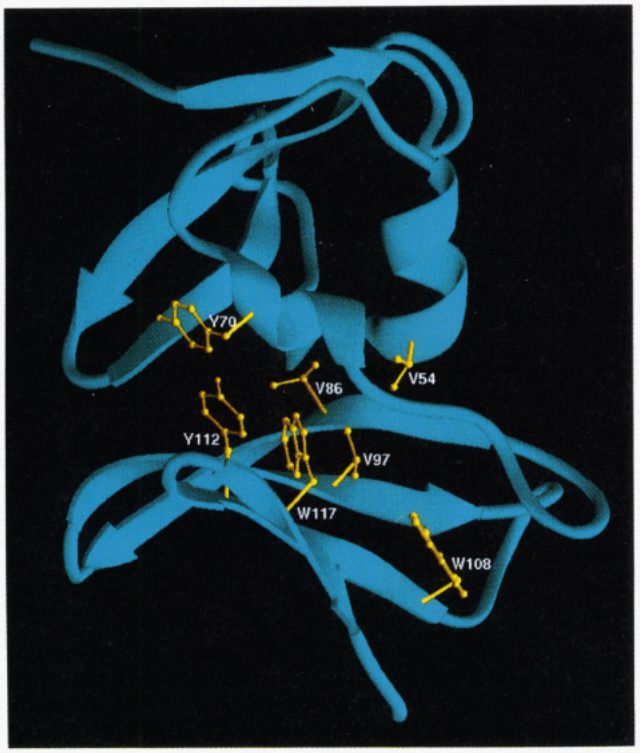

FIGURE 4: Ribbon plot of the C-terminal fragment of *E. coli* topoisomerase I (residues 9−122). The residues that form the C-terminal hydrophobic core are indicated in gold.

Table 2: Single-Stranded DNA Binding Affinity for the C-Terminal Domain of *E. coli* DNA Topoisomerase I

DNA	salt (mM)[a]	K_D (μM)[b]
5′-ACGACAGGCTAC-3′	10	4 ± 2
	200	43 ± 18
5′-GGCTAC-3′	10	14 ± 5
5′-CTCC-3′	10	950 ± 180

[a] Potassium phosphate was used, and the pH of the sample is 6.5. [b] standard deviations were obtained from the calculated K_D values of the seven amides which shift significantly upon binding of single-stranded DNA.

several aromatic residues (Y26, F37, and F43) are also located on this side of the first β-sheet and are poised for undergoing stacking interactions with the bases of single-stranded DNA. Although the second β-sheet is structurally similar to the first and also contains positively charged and aromatic amino acids (Figure 6), single-stranded DNA does not bind to this site, suggesting that the proper positioning of these side chains is critical for forming a favorable interaction with oligonucleotides.

Electrostatic interactions must contribute to the binding of the topoisomerase fragment to single-stranded DNA as evidenced by the decrease in binding observed with increasing concentrations of salt (Table 2). However, base-stacking interactions must also be important, since a significant amount of binding affinity (K_D = 43 μM) is retained for the 12-mer oligonucleotide even at a salt concentration of 200 mM. These results are consistent with the spectral changes observed in UV spectra upon complex formation (Zhu et al., 1995) and help to explain the role of this portion of the enzyme for enhancing the relaxation of supercoiled DNA under high salt conditions (Beran-Steed & Tse-Dinh, 1989).

Comparison to RNA- and Single-Stranded DNA-Binding Proteins. In contrast to the large amount of structural

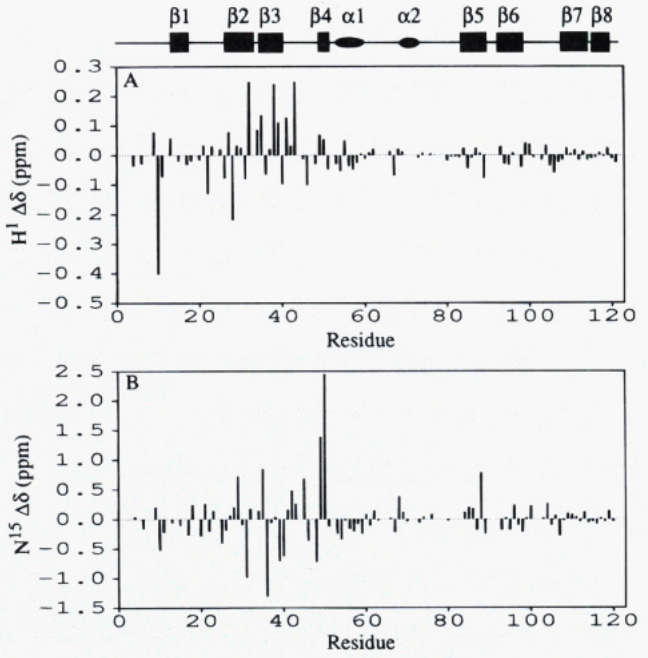

FIGURE 5: Chemical shift changes ($\delta_{free} - \delta_{complexed}$) of the (A) ^{1}H and (B) ^{15}N signals of the backbone amides of the uniformly ^{15}N-labeled C-terminal topoisomerase fragment upon the addition of the 6-mer ssDNA (5′-GGCTAC-3′) under low salt conditions (10 mM potassium phosphate at pH = 6.5).

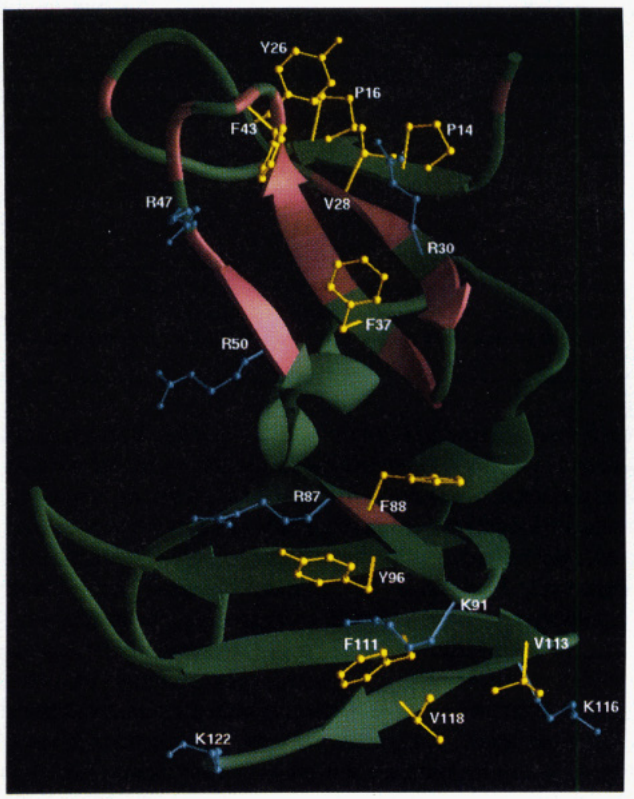

FIGURE 6: Ribbon plot of the C-terminal fragment of *E. coli* topoisomerase I (residues 9−122). The side chains of selected positively charged (blue) and hydrophobic amino acids (gold) are shown. The ribbon is colored rose for those residues in which the ^{15}N or ^{1}H chemical shifts of the backbone amides change by more than |0.5| or |0.1| ppm, respectively, upon the addition of the 6-mer ssDNA. The chemical shift changes upon the addition of the 12-mer and 4-mer oligonucleotides are very similar to those shown here for the 6-mer.

information available on the individual protein motifs that bind to double-stranded DNA (Pabo & Sauer, 1992),

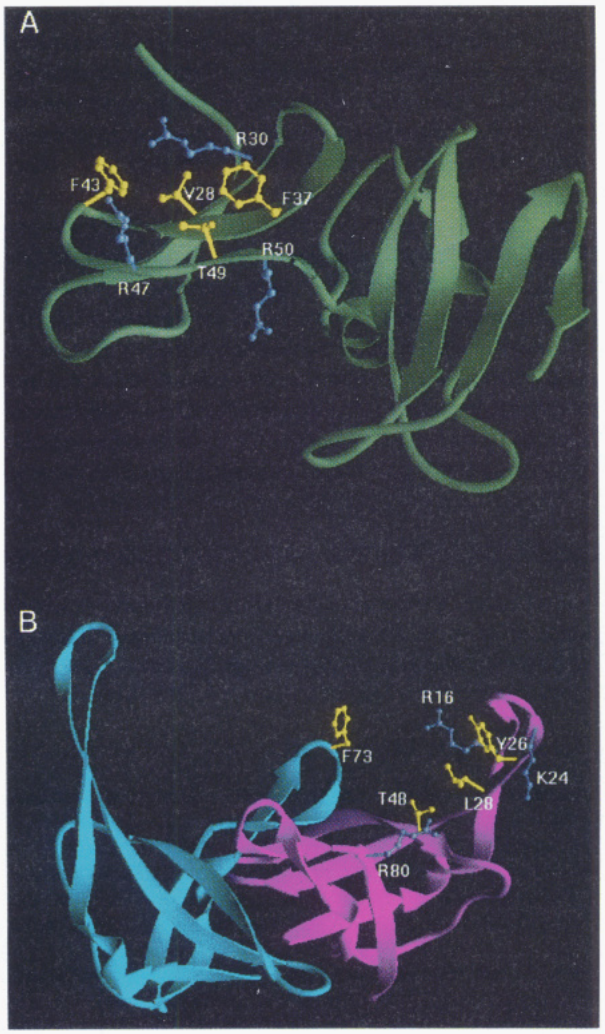

FIGURE 7: Ribbon plots comparing the ssDNA-binding sites of (A) the C-terminal single-stranded DNA-binding domain of *E. coli* topoisomerase I (residues 9−122) and (B) the gene 5 protein of bacteriophage f1. The residues in common in the ssDNA-binding cleft are indicated. The hydrophobic residues are in gold, and the positively charged residues are in blue. For the gene 5 protein, R16C, Y26C, F73L, and R80V mutants are inactive, suggesting that these surface residues are important for binding to single-stranded DNA (Stassen et al., 1992; Terwilliger et al., 1994).

relatively little is known about the structures and binding characteristics of proteins that interact with single-stranded DNA. Gene 5, the most well-studied single-stranded DNA-binding protein, adopts a five-stranded β-barrel in solution (Folkers et al., 1994) and in the crystalline state (Skinner et al., 1994) and binds to single-stranded DNA as a dimer in a cleft formed between the loops of the β-sheets (Folkers et al., 1993) (Figure 7B). The cold shock protein which also binds to single-stranded DNA adopts a five-stranded β-barrel (Schindelin et al., 1993, 1994; Schnuchel et al., 1993; Newkirk et al., 1994) that is structurally similar to gene 5 and the proposed oligonucleotide/oligosaccharide binding fold (Murzin, 1993). The nucleic acid-binding domain of transcriptional elongation factor TFIIS consists of a three-stranded β-sheet and a zinc-binding site (Qian et al., 1993). The overall architecture of the C-terminal fragment of topoisomerase I is different from all of these proteins but more closely resembles the structure of RNA-binding proteins such as U1 snRNP A, hnRNP A1, and hnRNP C that contain a four-stranded antiparallel β-sheet and two α-helices (Nagai et al., 1990; Hoffman et al., 1991; Görlach et al, 1992; Wittekind et al., 1992; Garrett et al., 1994). However, unlike these RNA-binding proteins in which the α-helices are interspersed among the β-strands in a $\beta\alpha\beta-\beta\alpha\beta$ motif, the topoisomerase C-terminal fragment has a contiguous antiparallel β-sheet followed by two α-helices. Furthermore, the angles between the helices in the two proteins are very different. Despite these differences in structure, however, the oligonucleotide-binding site in the topoisomerase fragment, a β-sheet containing positively charged and aromatic residues, is similar to that of all the RNA- and single-stranded DNA-binding proteins characterized to date as illustrated in the comparison with the gene 5 protein (Figure 7). Thus, this binding motif which is formed from a variety of protein architectures is highly conserved and plays an important role in the molecular recognition of single-stranded nucleic acids.

SUPPLEMENTARY MATERIAL AVAILABLE

One table containing the ^{1}H, ^{13}C, and ^{15}N chemical shifts of the C-terminal fragment of *E. coli* topoisomerase I (10 pages). Ordering information is given on any current masthead page.

REFERENCES

Archer, S. J., Ikura, M., Torchia, D. A., & Bax, A. (1991) *J. Magn. Reson. 95,* 636−641.

Bax, A., & Ikura, M. (1991) *J. Biomol. NMR 1,* 99−104.

Bax, A., & Pochapsky, S. (1992) *J. Magn. Reson. 99,* 638−643.

Bax, A., & Grzesiek, S. (1993) *Acc. Chem. Res. 26,* 131−138.

Bax, A., Clore, G. M., & Gronenborn, A. M. (1990) *J. Magn. Reson. 88,* 425−431.

Beran-Steed, R. K., & Tse-Dinh, Y.-C. (1989) *Proteins: Struct. Funct., Genet. 6,* 249−258.

Bodenhausen, G., & Ruben, D. J. (1980) *Chem. Phys. Lett. 69,* 185−189.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4,* 187−217.

Brünger, A. T. (1992) *X-PLOR 3.1 Manual,* Yale University Press, Yale University, New Haven, CT.

Carson, M. (1987) *J. Mol. Graphics 5,* 103−106.

Clubb, R. T., Thanabal, V., & Wagner, G. (1992) *J. Magn. Reson. 97,* 213−217.

Fesik, S. W., & Zuiderweg, E. R. P. (1988) *J. Magn. Reson. 78,* 588−593.

Folkers, P. J. M., van Duynhoven, J. P. M., van Lieshout, H. T. M., Harmsen, B. J. M., van Boom, J. H., Tesser, G. I., Konings, R. N. H., & Hilbers, C. W. (1993) *Biochemistry 32,* 9407−9416.

Folkers, P. J. M., Nilges, M., Folmer, R. H. A., Konings, R. N. H., & Hilbers, C. W. (1994) *J. Mol. Biol. 236,* 229−246.

Garrett, D. S., Lodi, P. J., Shamoo, Y., Williams, K. R., Clore, G. M., & Gronenborn, A. M. (1994) *Biochemistry 33,* 2852−2858.

Görlach, M., Wittekind, M., Beckman, R. A., Mueller, L., & Dreyfuss, G. (1992) *EMBO J. 11,* 3289−3295.

Griesinger, C., Otting, G., Wüthrich, K., & Ernst, R. R. (1988) *J. Am. Chem. Soc. 110,* 7870−7872.

Grzesiek, S., & Bax, A. (1992) *J. Magn. Reson. 99,* 201−207.

Grzesiek, S., & Bax, A. (1993) *J. Magn. Reson., Ser. B 102,* 103−106.

Grzesiek, S., Anglister, J., & Bax, A. (1993) *J. Magn. Reson., Ser. B 101,* 114−119.

Hoffman, D. W., & Spicer, L. D. (1991) *Techniques in Protein Chemistry II,* pp 409−416, Academic Press, New York.

Hoffman, D. W., Query, C. C., Golden, B. L., White, S. W., & Keene, J. D. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88,* 2495−2499.

Kay, L. E., & Bax, A. (1990) *J. Magn. Reson. 86,* 110−126.

Kay, L. E., Ikura, M., Tschudin, R., & Bax, A. (1990) *J. Magn. Reson. 89,* 496−514.

Kuszewski, J., Nilges, M., & Brünger, A. T. (1992) *J. Biomol. NMR 2*, 33−56.

Lima, C. D., Wang, J. C., & Mondragón, A. (1994) *Nature 367*, 138−146.

Logan, T. M., Olejniczak, E. T., Xu, R. X., & Fesik, S. W. (1992) *FEBS Lett. 314*, 413−418.

Marion, D., Ikura, M., Tschudin, R., & Bax, A. (1989a) *J. Magn. Reson. 85*, 393−399.

Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., & Bax, A. (1989b) *J. Am. Chem. Soc. 111*, 1515−1517.

Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., & Clore, G. M. (1989c) *Biochemistry 28*, 6150−6156.

Messerle, B. A., Wider, G., Otting, G., Weber, C., & Wüthrich, K. (1989) *J. Magn. Reson. 85*, 608−613.

Montelione, G. T., Lyons, B. A., Emerson, S. D., & Tashiro, M. (1992) *J. Am. Chem. Soc. 114*, 10974−10975.

Murzin, A. G. (1993) *EMBO J. 12*, 861−867.

Nagai, K., Oubridge, C., Jessen, T. H., Li, J., & Evans, P. R. (1990) *Nature 348*, 515−520.

Neri, D., Szyperski, T., Otting, G., Senn, H., & Wüthrich, K. (1989) *Biochemistry 28*, 7510−7516.

Newkirk, K., Feng, W., Jiang, W., Tejero, R., Emerson, S. D., Inouye, M., & Montelione, G. T. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 5114−5118.

Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988) *FEBS Lett. 229*, 317−324.

Pabo, C. O., & Sauer, R. T. (1992) *Annu. Rev. Biochem. 61*, 1053−1095.

Pelton, J. G., Torchia, D. A., Meadow, N. D., & Roseman, S. (1993) *Protein Sci. 2*, 543−558.

Qian, X., Jeon, C., Yoon, H., Agarwal, K., & Weiss, M. A. (1993) *Nature 365*, 277−279.

Schindelin, H., Marahiel, M. A., & Heinemann, U. (1993) *Nature 364*, 164−168.

Schindelin, H., Jiang, W., Inouye, M., & Heinemann, U. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 5119−5123.

Schnuchel, A., Wiltscheck, R., Czisch, M., Herrier, M., Willimsky, G., Graumann, P., Marahiel., M. A., & Holak, T. A. (1993) *Nature 364*, 169−171.

Skinner, M. M., Zhang, H., Leschnitzer, D. H., Guan, Y., Bellamy, H., Sweet, R. M., Gray, C. W., Konings, R. N. H., Wang, A. H.-J., & Terwilliger, T. C. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 2071−2075.

Stassen, A. P. M., Zaman, G. J. R., van Deursen, J. M. A., Schoenmakers, J. G. G., & Konings, R. N. H. (1992) *Eur. J. Biochem. 204*, 1003−1014.

Terwilliger, T. C., Zabin, H. B., Horvath, M. P., Sandberg, W. S., & Schlunk, P. M. (1994) *J. Mol. Biol. 236*, 556−571.

Tse-Dinh, Y.-C. (1994) *Adv. Pharmacol. 29A*, 21−37.

Vuister, G. W., & Bax, A. (1993) *J. Am. Chem. Soc. 115*, 7772−777.

Wang, J. C. (1985) *Annu. Rev. Biochem. 54*, 665−697.

Wishart, D. S., & Sykes, B. D. (1994) *J. Biomol. NMR 4*, 171−180.

Wittekind, M., Görlach, M., Friedrichs, M., Dreyfuss, G., & Mueller, L. (1992) *Biochemistry 31*, 6254−6265.

Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949−961.

Zhu, C.-X., Samuel, M., Pound, A., Ahumada, A., & Tse-Dinh, Y.-C. (1995) *Biochem. Mol. Biol. Int. 35*, 375−385.

BI950481C