PROTEINS: Structure, Function, and Genetics 6:231–239 (1989)

Peptide Sequencing and Site-Directed Mutagenesis Identify Tyrosine-319 as the Active Site Tyrosine of *Escherichia coli* DNA Topoisomerase I

Richard M. Lynn and James C. Wang
Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138

ABSTRACT **Tyrosine 319 of *E. coli* topoisomerase I is shown to be the active site tyrosine that becomes covalently attached to a DNA 5′ phosphoryl group during the transient breakage of a DNA internucleotide bond by the enzyme. The tyrosine was mapped by trapping the covalent complex between the DNA and DNA topoisomerase I, digesting the complex exhaustively with trypsin, and sequencing the DNA-linked tryptic peptide. Site-directed mutagenesis converting Tyr-319 to a serine or phenylalanine completely inactivates the enzyme. The structure of the enzyme and its catalysis of DNA strand breakage, passage, and rejoining are discussed in terms of the available information.**

Key words: phosphotyrosine linkage, protein–DNA transesterification, enzyme mechanism, DNA–protein covalent complex

INTRODUCTION

A hallmark of all DNA topoisomerases is the formation of an enzyme–DNA covalent intermediate during their catalysis of DNA strand breakage and rejoining through two successive transesterification reactions (for recent reviews, see 1–3). For each of the archetypical enzymes, including eubacterial and eukaryotic DNA topoisomerases I and II, an active site tyrosine residue plays a key role in these reactions. In the first reaction, the phenolic oxygen of this residue forms a covalent link with an internucleotide phosphate, severing a DNA phosphodiester bond and leaving a hydroxyl group on a deoxyribose. In the second reaction, a reversal of the first occurs, restoring the DNA phosphodiester linkage and freeing the enzyme tyrosyl group from the DNA.

The positions of the DNA-linked tyrosines in the enzyme polypeptides can be identified by trapping the DNA–protein covalent intermediates with protein denaturants or with unusual DNA substrates. There is strong evidence that such trapped intermediates are indeed the true intermediates in the normal reactions catalyzed by the enzymes; hence the DNA-linked tyrosines in the trapped complexes are the active site tyrosines.[4,5] For *Escherichia coli* DNA topoisomerase II (DNA gyrase) and its evolutionarily related eukaryotic counterpart *Saccharomyces cerevisiae* DNA topoisomerase II, direct sequencing of the DNA-linked peptides following tryptic digestion of the protein–DNA covalent intermediates has identified tyrosine-122 of the gyrase A-subunit and tyrosine-783 of the single-subunit eukaryotic enzyme as the active site tyrosines.[6,7] By comparing amino acid sequences of other topoisomerases related to these enzymes, it is readily deduced that tyrosine-116 of phage T4 gene 52 protein, one of the three subunits of T4 DNA topoisomerase, tyrosine-123 of the gyrase A-subunit of *Bacillus subtilis*, and tyrosines 785, 781, and 804 of *Drosophila melanogaster*, *Schizosaccharomyces pombe*, and human DNA topoisomerase II, respectively, are the active site tyrosines in these enzymes.[8–12]

In all of the examples cited in the above paragraph, the topoisomerases are classified as type II DNA topoisomerases. These enzymes cleave a pair of DNA strands of a double-stranded DNA in concert, opening a transient gate in the DNA, and mediate the passage of another double-stranded DNA segment through such a gate (see the reviews cited and 13–15). In contrast, the type I DNA topoisomerases transiently cleave one DNA strand at a time in their catalysis of DNA topoisomerization. All organisms examined to date contain one or more type I topoisomerase.[1,2] The major eubacterial type I enzyme, for which *E. coli* DNA topoisomerase I is the most extensively studied example, shares essentially no sequence similarity with eukaryotic DNA topoisomerase I, the major eukaryotic type I enzyme.[16]

In this communication, we report the identification of Tyr-319 as the active site tyrosine of *E. coli* DNA topoisomerase I by direct peptide sequencing and by site-directed mutagenesis of the gene *topA* encoding the enzyme. Mechanistic implications of this identification and the possibility of developing antibiotics targeting the bacterial type I DNA topoisomerases are discussed.

Received March 7, 1989; accepted April 10, 1989.

Address reprint requests to James C. Wang, Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, MA 02138.

EXPERIMENTAL PROCEDURES

Materials

Some of the materials used and their sources are as follows: Brij-58, phenylmethylsulfonylflouride (PMSF) and lysozyme were purchased from Sigma; Polymin P, isopropyl-1-thio-β-D-galactoside (IPTG) and single-stranded DNA agarose were from Bethesda Research Labs; proteinase K and sodium dodecyl sulfate (SDS) were from Boehringer-Manheim Biochemicals; trypsin (TPCK-treated), phosphocellulose, site-directed mutagenesis kit, and Spherogel TSK-TM HPLC column were purchased from Worthington, Whatman, Amersham, and Beckman, respectively. Single-stranded M13mp19 phage was purified from 500 ml of infected culture of *E. coli* strain JM101 by the polyethylene glycol precipitation procedure.[17] The purified phage was extracted several times with phenol and phenol-chloroform, and the aqueous layer containing the DNA was dialyzed exhaustively against 10 mM Tris, pH 8.0, 0.1 mM Na_3EDTA. *E. coli* strain RB968, a phage T1-resistant W3110 derivative with an up-mutation in the lacI promoter, was obtained from Dr. R. Brent (Harvard Medical School).

Purification of *E. coli* Topoisomerase I

Overproduction of *E. coli* DNA topoisomerase I was achieved in strain RB968 cells transformed with pJW312-Sal, in which the coding sequence of the gene *topA* is expressed from the inducible lac promoter. This plasmid differs from pJW312 described in Zumstein and Wang[18] only by an inconsequential change of a *Bgl*II site upstream of the lac promoter to a *Sal*I site (by linker insertion). A typical preparation of the topoisomerase is described below; unless stated otherwise, processing of cells and protein fractions was carried out at 4°C.

Ten liters of cells was grown in rich broth at 37°C to an optical density of 0.6 at 595 nm (measured in a Cary 118 spectrophotometer). Then, IPTG was added to 1 mM to induce transcription of the *topA* gene from the lac promoter. Cells were harvested after 3 hours by centrifugation and were resuspended in 50 mM Tris–HCl, pH 8.0, 25% sucrose. One milliliter of the suspension buffer was added for each gram of wet-packed cells. The mixture (240 ml) was immediately frozen in an ethanol–dry ice bath and stored at −70°C.

After thawing on ice, stock solutions of Na_3EDTA, 2-mercaptoethanol, and lysozyme were added to the mixture to final concentrations of 20 mM, 10 mM, and 200 μg/ml, respectively, and the mixture was incubated at 4°C for 15 minutes. KCl was then added to 0.15 M and the mixture was distributed into centrifuge tubes for the Beckman 60.1 Ti rotor, each containing a measured volume of the detergent Brij-58 to achieve a final detergent concentration of 0.5%. After mixing, the tubes were incubated at 4°C for 15 minutes to lyse the cells and then spun for 60 minutes at 60,000 rpm in a Beckman 60.1 Ti rotor.

A 10% Polymin P (pH 7.5) solution was added to the cleared lysate (Fraction I, 190 ml) to a final concentration of 0.7%. The mixture was spun for 10 minutes at 10,000 rpm. Under these conditions DNA topoisomerase I remains in the supernatant. Ammonium sulfate was then added to 70% saturation. After centrifuging for 25 minutes at 12,000 rpm, the supernatant was discarded and the pellet was resuspended in Buffer I, which contained 20 mM potassium phosphate, pH 7.4, 1 mM Na_3EDTA, 10% (w/v) glycerol, 10 mM 2-mercaptoethanol, and 0.1 mM freshly prepared PMSF. The solution was dialyzed for 2 hours against two changes of Buffer I + 0.1 M KCl. The conductivity of the solution was adjusted by the addition of Buffer I to that of Buffer I + 0.1 M KCl and the fraction (Fraction II, 225 ml) was applied to a phosphocellulose column (30 cm × 9.6 cm^2), which had been preequilibrated with Buffer I + 0.1 M KCl. After washing the column with 500 ml of Buffer I + 0.1 M KCl, the column was eluted with a 2 liter linear gradient of Buffer I containing 0.1–0.5 M KCl. DNA topoisomerase I eluted approximately midway through the gradient.

The pooled phosphocellulose fraction (Fraction III, 150 mg protein in 285 ml) contained DNA topoisomerase I with a purity greater than 95%, and was used in all of the experiments described here.

Cleavage of Single-Stranded DNA by DNA Topoisomerase I

Each sample contained 4 μg of single-stranded M13mp19 DNA in 20 μl of 10 mM Tris–HCl, pH 7.8, 120 mM KCl, 10 mM $MgCl_2$. The DNA was incubated with varying amounts of the enzyme, ranging from a protein:DNA molar ratio of 1 to several hundred. After 10 minutes at 37°C, 2 μl of 2 M NaOH was added to each sample and incubation was continued for 1 minute more. Each reaction mixture was then neutralized by the addition of 4 μl of 1.8 M Tris–HCl and 0.2 M Tris base. DNA and DNA topoisomerase I covalently linked to DNA were precipitated by adding in succession 200 μl of 0.2 M ammonium acetate, 10 mM Na_3EDTA, 50 μg/ml tRNA, and 600 μl of ethanol. Upon centrifugation, the pellet from each sample was resuspended in 100 μl of 10 mM Tris, pH 7.8, 5 mM Na_3EDTA, 0.5% SDS, and treated with 100 μg/ml proteinase K for 1 hour at 37°C. The DNA and DNA-linked protein in each sample were again precipitated as before with ethanol, washed with 70% ethanol, dried, and resuspended in 5 μl of 10 mM Tris, pH 8.0, 2 mg/ml RNase A and incubated at 37°C for 15 minutes to remove tRNA. The digest was mixed with 155 μl of 50 mM NaOH, 1 mM EDTA, 2.5% Ficoll (type 400), 0.025% bromocresol green, and incubated at 55°C for 5 minutes. The extent of DNA cleavage by topoisomerase in the samples was examined by loading portions of

the samples in separate lanes of a 1% alkaline-agarose gel. Electrophoresis was for 16 hours at a field of 1 V/cm.

Preparation of the Covalent Complex and Treatment With Trypsin

Three hundred micrograms of single-stranded M13mp19DNA was incubated with 3 mg of *E. coli* DNA topoisomerase I in 1.5 ml of 10 mM Tris–HCl, pH 7.8, 120 mM KCl, 10 mM $MgCl_2$. After 10 minutes at 37°C, 80 μl of 2 M NaOH was added to trap the covalent complex. Neutralization of the solution and ethanol precipitation of the covalent complex were carried out by the successive addition of 160 μl 1 M HCl, 30 μl 1 M Tris–HCl, pH 7.8, 430 μl 1.4 M sodium acetate, and 5 ml of ethanol. This mixture was kept in a freezer overnight. After centrifugation and washing the pellet with 70% ethanol and drying, 150 μl of 10 mM Tris, pH 8.0, 1% SDS were added and the tube was warmed to 60°C for 30 minutes to solubilize the pellet. The solution was diluted to 3 ml with 10 mM Tris, pH 7.8, 3 mM $CaCl_2$. Trypsin (300 μg) was then added and the mixture was incubated at 37°C. At 3 hour intervals 100 μg of trypsin was added. After 9 hours total, the mixture was extracted once with phenol and twice with phenol:chloroform:isoamyl alcohol (50:49:1 by volume) and the DNA and DNA-linked peptides in the aqueous phase were precipitated by the addition of 0.1 volume of 3 M sodium acetate and 2.5 volumes of ethanol. The pellet was dissolved in 2 ml of 10 mM Tris, pH 7.8, 10 mM $CaCl_2$ and treated again with 100 μg of trypsin for 8 hours. The mixture was phenol extracted and alcohol precipitated as above. After washing with 70% ethanol and drying, the pellet was resuspended in 0.3 M sodium acetate. Insoluble matter was removed by centrifugation and the supernatant was loaded onto a Spherogel TSK-TM HPLC column equilibrated with 0.3 M sodium acetate. Absorbance of the effluent was monitored at 254 nm. A peak corresponding to material excluded from the column was collected and precipitated by the addition of 2.5 volumes of ethanol. After precipitation, approximately 200 μg of DNA was recovered and resuspended in water for amino acid sequence analysis, using an Applied Biosystems model 477A Protein sequenator connected on-line to a model 120A HPLC. Phenylthiohydantoin derivatives of the amino acid peaks were assigned without resort to any sequence information of the protein.

Site-Directed Mutagenesis of Tyrosine-319 of *E. coli* Topoisomerase I

The starting material for the mutagenesis was plasmid pJW312-*topA874H*, a pJW312 derivative with a *Hin*dII dodecamer linker inserted at the *Alu*I site at position 2191 of the *E. coli topA* DNA sequence.[18] The *Hin*dII–*Eco*RI fragment spanning bases 2188 to 3404 of the *E. coli topA* coding sequence was cloned in between the *Hin*dIII and *Eco*RI sites of the double-stranded replicating form of M13mp18, and single-stranded phage DNA containing the insert was prepared. Mutagenesis was accomplished by the oligonucleotide directed method of Nakamaye and Eckstein,[19] using a kit supplied by Amersham corporation. Two oligonucleotides, 5′GGCTATATCACTTCCATGCG3′ and 5′GGCTATATCACTTTCATGCG3′, were used to change Tyr-319 to a serine (replacing an A at position 2273 of the published sequence with a C) or a phenylalanine (replacing an A at 2273 with a T). The oligonucleotides were separately annealed to the single-stranded DNA; subsequent steps were carried out following the protocol of the supplier of the mutagenesis kit. Phage plaques were purified and DNA samples were prepared from several individual isolates. The desired mutations were identified by nucleotide sequencing.

Replicating forms of the two M13 clones were then prepared and the *Bst*XI to *Acc*I fragment from bases 2232 to 2797 was excised from each and used to substitute for the corresponding wild-type *topA* fragment in pJW312-Sal. DNA was prepared from each clone, and the *Bst*XI–*Acc*I fragment containing the mutated stretch was sequenced again to confirm the desired substitutions.

Activity Assays and Complementation Tests

DNA topoisomerase I activity assays were performed using extracts from transformants of strain DM800 (Δ*topA* Δ*cysB*).[20] DM800 was first transformed with the plasmid pACYC184-I^{Q1} in which an 1100 bp *lacI* gene with an upmutation in the promoter[21] is cloned into the *Eco*RI site of pACYC184, a pBR322-compatible high-copy number plasmid conferring tetracycline resistance.[22] Cells harboring the *lacI* plasmid were next transformed with pJW312-Sal or its derivatives containing mutations in the *topA* gene. Transformants were grown up in Luria broth containing cysteine, tetracycline, and ampicillin. The expression of DNA topoisomerase I in each culture was induced by the addition of 10 mM IPTG when the optical density of the culture reached 0.6 at 595 nm. After 1.5 more hours of growth, 5 ml of cells from each culture was pelleted by centrifugation and then resuspended in 100 μl of 40 mM Tris, pH 8.0, 25% sucrose, 1 mM EDTA, and 1 mg/ml lysozyme. After incubation on ice for 2 minutes, 100 μl of 10 mM Tris, pH 8.0, 0.9% Brij-58, and 60 mM $MgCl_2$ were added and incubation was continued for 2 minutes more. Next, 10 μl of 5 M NaCl was added and after 2 minutes on ice the mixture was centrifuged in a microcentrifuge for 4 minutes. For activity assays, 2 μl of this extract or appropriate dilutions of it was incubated with 18 μl of reaction mixture containing 30 mM Tris, pH 8.0, 80 mM NaCl, 4 mM $MgCl_2$, 40 μg/ml gelatin, and 22 μg/ml supercoiled pHC624 DNA. Extracts were di-

luted with 1.3× reaction mixture without DNA. The assays were incubated at 37°C for 15 minutes and then terminated by extraction with phenol:chloroform:isoamyl alcohol (50:49:1 by volume). The aqueous phase of each assay was mixed with a gel electrophoresis loading buffer containing Ficoll and tracking dyes and a portion of it was loaded on a 0.7% agarose gel. Electrophoresis was typically run for 15 hours at 1 V/cm.

Complementation tests were done in strain AS17, which has a temperature sensitive *topA* phenotype, and does not grow at 42°C unless DNA topoisomerase I activity is supplied[18,23] (see also 24). AS17 cells were first transformed with pMK16-I^{Q1}, a plasmid that confers kanamycin resistance and overproduces the lac repressor.[21] Cells harboring the *lacI* plasmid were transformed with the various *topA* plasmids to be tested for activity. Transformants were grown up at 30°C, and suspensions containing individual colonies were diluted and then plated on selective plates containing ampicillin and kanamycin. Plates containing equal numbers of cells were incubated at 30°C and 42°C overnight, and the number of colonies on each plate was counted.

RESULTS

Formation of a Protein–DNA Covalent Complex and Direct Sequencing of the DNA-Linked Tryptic Peptide Identifies Tyr-319 as the Active Site Tyrosine of *E. coli* DNA Topoisomerase I

Several hundred picomoles of the protein–DNA covalent complex is needed in order to identify an active site tyrosine by first forming a protein–DNA covalent complex and then directly sequencing the DNA-linked tryptic peptide.[6,7] To estimate the yield of the protein–DNA covalent complex, reaction mixtures each containing 4 μg of single-stranded M13mp19 DNA (7250 nucleotides per molecule) and varying amounts of *E. coli* DNA topoisomerase I were incubated and treated with NaOH to induce the cleavage of the DNA and the formation of the covalent complex. As described in detail in the Experimental section, each NaOH treated sample was neutralized, digested with proteinase K, and run on an alkaline gel to determine the extent of cleavage of the DNA in that sample. Because the cleavage of each DNA internucleotide bond is accompanied by the formation of a DNA–protein covalent bond, the yield of the covalent complex is directly related to the extent of DNA cleavage.

Figure 1 depicts the result of such an experiment. Lane 1 contains *Hin*dIII digested phage λ DNA as size markers. Lanes 2 and 3 contain untreated DNA and DNA mock-treated by the omission of DNA topoisomerase I in the reaction mixture. In either case, the intact ring form of single-stranded M13mp19 DNA is the major species and it migrates slower than the minor linear DNA band. In lanes 4–8, the molar ratios of DNA topoisomerase I to M13mp19 DNA during incubation and NaOH treatment were estimated to be 1.2, 13, 33, 150, and 300, respectively, based on the known molecular weights of the protein and the DNA. At a protein to DNA ratio such that a significant fraction *f* of the DNA rings remains intact, the average number of DNA phosphodiester bonds broken or the average number of protein–DNA covalent links formed per DNA ring, denoted by $\langle i \rangle$, can be estimated from the

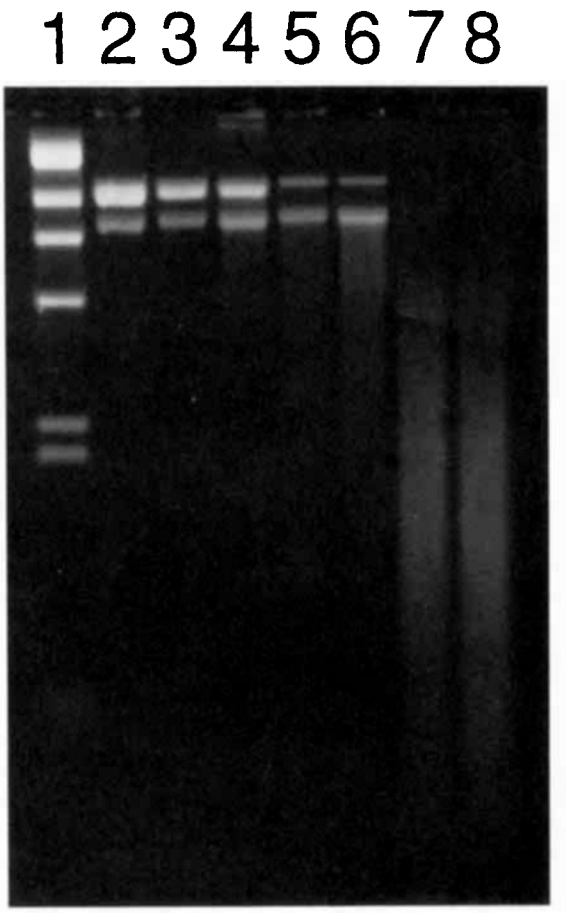

Fig. 1. Cleavage of single-stranded M13mp19 DNA by *E. coli* DNA topoisomerase I. Reaction mixtures each containing 4 μg of single-stranded M13mp19 DNA and varying amount of *E. coli* DNA topoisomerase I in a total volume of 20 μl were incubated at 37°C for 10 minutes, and NaOH was added to the mixtures to a final concentration of 50 mM to form the enzyme–DNA covalent complex. Following neutralization, ethanol precipitation, and proteinase K treatment, each sample was reprecipitated with alcohol and a portion of it was examined by alkali-agarose gel electrophoresis (see the Experimental section for details). Lane 1, *Hin*dIII restriction fragments of phage λ DNA used as size markers; lane 2, 0.4 μg of untreated input DNA. The amounts of *E. coli* DNA topoisomerase I in the reaction mixtures run in lanes 3–8 were lane 3, 0 μg; lane 4, 0.2 μg; lane 5, 2.1 μg; lane 6, 5.3 μg; lane 7, 25 μg; lane 8, 50 μg. Approximately 0.3 μg of DNA was loaded in each of the lanes 3–6, and approximately 0.75 μg was loaded in each of the lanes 7 and 8. The gel was stained with ethidium bromide after neutralization and photographed over a UV light source.

Poisson relation $f = e^{-\langle i \rangle}$. For the sample run in lane 5, densitometer tracings of the negative of the photograph shown indicate that about 35% of the DNA rings remained intact. Thus $\langle i \rangle$ is estimated to be about 1 for this sample. Since there were 13 enzyme molecules per M13mp19 DNA, the yield of the covalent complex is estimated to be about 1/13 or 8%. At much higher protein to DNA ratios, however, the yield might be lower. For the sample shown in lane 8, an 8% yield would predict a value of $\langle i \rangle$ of 300 × 0.08 or 24, and the average size of the DNA after the cleavage reaction would be 7250/24 or 300 nucleotides. The size distribution of the DNA shown in lane 8 suggests that the average size might be significantly above 300. However, it is also plausible that the gel intensity pattern tends to underrepresent shorter DNA fragments. Any clustering of the topoisomerase binding sites on the DNA might also lead to an underestimate of the yield.

Because of technical complications of peptide sequencing in the presence of a large excess of DNA,[7] it was necessary to carry out the cleavage reaction at a high enzyme to DNA ratio. A mixture containing 3 mg (30 nmol) of *E. coli* DNA topoisomerase I and 300 μg of M13mp19 DNA (120 pmol of DNA or 900 nmol of nucleotides) was used to form the covalent complex, as described in detail in the Experimental section. Following precipitation with ethanol and resuspension, the covalent protein–DNA complex as well as any free protein that might have been precipitated in ethanol was exhaustively digested with trypsin, and the covalent complex was purified by phenol extraction, ethanol precipitation, and chromatography through a large-pore HPLC gel filtration column with a high-molecular-weight exclusion limit. The DNA eluted from this column in a very broad peak; only the portion that eluted in the void volume of the column was collected. The fractions collected eluted well before large proteins such as bovine serum albumin (66 kDa). Thus, the presence of contaminating noncovalently attached peptides in the sample used in protein sequencing is highly unlikely.

Amino acid sequencing of the collected material showed that the amino acids from the first two cycles of Edman degradation were isoleucine followed by threoine. In the third cycle there were no assignable amino acid peaks. The fourth, fifth, and sixth cycles contained methionine, arginine, and "blank," respectively. The same sequence could be read at a lower level delayed by one cycle; this "shadow-casting" phenomenon was observed previously with the covalent complexes of *E. coli* DNA gyrase and yeast DNA topoisomerase II, and reflects incomplete reaction during each Edman cycle.[6,7]

The observed sequence Ile-Thr-(blank)-Met-Arg-(blank) matches perfectly to the expected pattern if Tyr-319 is covalently linked to the DNA. The amino acid sequence of *E. coli* DNA topoisomerase I starting at residue 317 is Ile-Thr-Tyr-Met-Arg-. Because of the covalent attachment of Tyr 319 to the DNA, one would not observe a recognizable phenylthiohydantoin derivative at this position. The match between the amino acid sequencing results and the published sequence from Ile-317 to Arg-321 implies that trypsin cleaved after residues 316 and 321. Residue 321 is an arginine, and is thus an expected trypsin cleavage site. Reisude 316 is a tyrosine. Because of the very high level of trypsin used during the digestion of the DNA-linked topoisomerase, it could not be ruled out that a trace amount of contaminating chymotrypsin was present in the TPCK-treated trypsin. Trypsin cleavage after tyrosine is known, however[25,26]; the specificity of this reaction is not well understood. Aside from the stretch Ile-317-Thr-Tyr-Met-Arg-321, no other sequences in *E. coli* DNA topoisomerase I match even two of the residues.

Mutagenesis of Tyr-319 of *E. coli* Topoisomerase I to Ser or Phe Inactivates the Enzyme

The biochemical result that Tyr-319 is the active site tyrosine of *E. coli* DNA topoisomerase I is supported further by the use of oligonucleotide-directed mutagenesis. Two mutated *topA* genes were constructed: in *topA-319S,* a serine codon replaces that for Tyr-319; in *topA-319F,* the replacement is by a phenylalanine codon.

The properties of the mutant enzymes were tested in *E. coli* strain DM800 (Δ*topA*) transformed with a lac repressor overproducing plasmid and a derivative of pJW312 in which *topA-319S* or *topA-319F* is expressed from the lac promoter. The various transformants were grown up, induced with IPTG, and lysed. The cell extracts were then tested for the presence of *E. coli* DNA topoisomerase I activity, as assayed by its relaxation of negatively supercoiled DNA (See Experimental Procedures for details).

Figure 2 shows that in extracts of cells expressing either *topA-319S* (lane 4) or *topA-319F* (lane 5), no relaxation activity is detectable. An extract from DM800 cells expressing the wild-type *topA* gene, on the other hand, showed relaxation activity even when the extract was diluted 200-fold (lane 3).

The results shown in Figure 3 show that the absence of DNA topoisomerase I activity in extracts of cells expressing the two mutated *topA* genes is not due to proteolysis of the mutant proteins. Fractionation of proteins in the various extracts by SDS–polyacrylamide gel electrophoresis, followed by immunostaining with antibodies specific to *E. coli* DNA topoisomerase I, shows similar patterns of protein bands in extracts of cells expressing the wild-type and the mutant *topA* genes.

The lack of DNA relaxation activity in the enzymes encoded by the *topA* mutants *319S* and *319F* is also reflected by the inability of the mutated *topA*

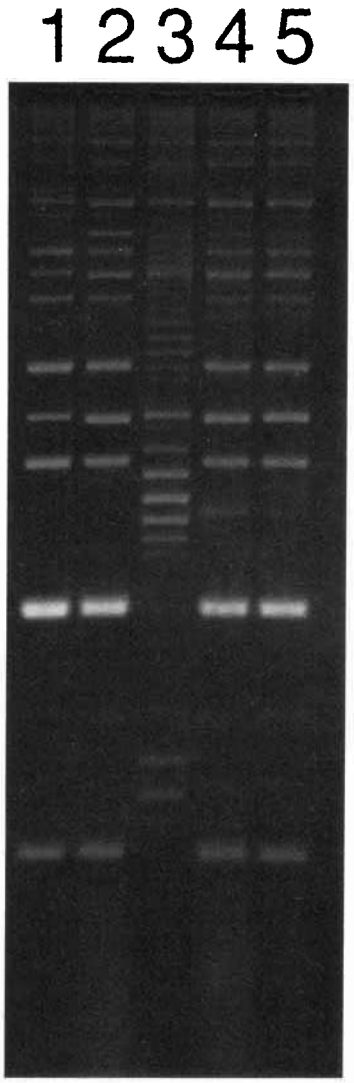

Fig. 2. Relaxation of negatively supercoiled pHC624 DNA by extracts of DM800 (*ΔtopA ΔcysB*) cells. Assay of *E. coli* DNA topoisomerase I activity in extracts of DM800 (*ΔtopA ΔcysB*) cells harboring various *topA* plasmids has been described in detail in the Experimental section. Cell extracts added to the incubation mixtures run in lanes 2–5 were from: lane 2, untransformed DM800 (*ΔtopA ΔcysB*); lane 3, DM800 transformed with pJCW312, which makes wild-type enzyme. This extract was diluted 1 to 200; lane 4, DM800 transformed with pJCW312-319S; lane 5, DM800 cells transformed with pJCW-319P.

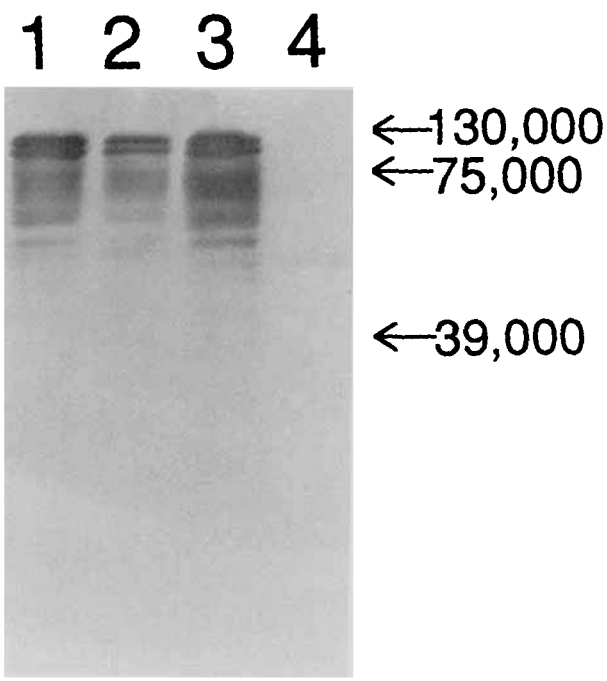

Fig. 3. Immunochemical detection of *E. coli* DNA topoisomerase I in extracts of transformed and untransformed strain DM800 cells. Extracts from untransformed DM800 cells (lane 4), or DM800 transformed with plasmids expressing a wildtype or mutant *topA* gene (lanes 1–3), were prepared as described in the Experimental section. Fifteen microliters of each extract was mixed with an equal volume of 0.12 M Tris–HCl, pH 6.8, 4% SDS, 10% 2-mercaptoethanol, 20% glycerol, and the mixture was loaded in a sample well of an SDS–10% polyacrylamide gel after heating at 100°C for 10 minutes. Following electrophoresis, the protein bands in the gel were blot-transferred to a nitrocellulose sheet and the gel-blot was treated with rabbit antibodies against *E. coli* DNA topoisomerase I. The bound rabbit antibodies were detected with alkaline phosphatase-conjugated goat antibodies specific to rabbit immunoglobulin. Lanes 1–3 contained extracts from DM800 transformed with pJW312-319P, pJW312-319S, and pJCW312, respectively; lane 4 contained cell extract from untransformed DM800 cells.

genes to complement the temperature-sensitive *topA* in strain AS17[18,22] (see also 24). As shown in Table I, whereas transformation of AS17 with pJW312 containing the wild-type *topA* gene restored completely viability at 42°C, transformation with the same plasmid in which the wild-type *topA* is replaced by either *topA-319S* or *topA-319F* had no effect on cell viability at the higher temperature.

TABLE I. Complementation of Strain AS17 [topA$_{am}$ pLLL1 (supD^{ts})] Cells by Plasmid-Borne *topA*

Transforming plasmid	Cell viability*
pJW312	1.0
pJW312-319S	<0.002
pJW312-319P	<0.002
None	<0.002

*Cell viability was measured by the ratio of the number of colonies at 42°C to that at 30°C.

DISCUSSION

E. coli DNA topoisomerase I is encoded by an open reading frame of 864 amino acids in the *topA* gene, of which the first f-Met is probably removed post-translationally to give a protein with a calculated molecular weight of 97,413.[16]

Biochemical and genetic studies of the enzyme have suggested several structural features. Near the amino-terminus, there is a potential "helix-turn-helix" motif between Ala-120 and Phe-139, which might be involved in the binding of the enzyme to

double-stranded DNA.[16] In the carboxyl-terminal region, three motifs, each containing four cysteine residues, are thought to be the binding sites of the three Zn(II) ions known to be associated with each mole of the enzyme.[27] Deletion of the region beyond the putative Zn-binding motifs has little effect on the activity of the enzyme, as evidenced by studies of a deletion mutant *2277ter*.[18] Deletion of all three Zn-binding motifs inactivates the *E. coli* enzyme,[18,27] as well as the homologous enzyme from *Klebsiella aerogenes* (S. Lynch, personal communication). These putative Zn-binding motifs might also be involved in DNA binding[27] (Y.-C. Tse-Dinh, personal communication).

Additional information on the structure of the enzyme has been gleaned from studies of how the enzyme interacts with DNA. Catalysis by bacterial DNA topoisomerase I appears to be preceded by the unpairing of a short stretch of its double-stranded DNA substrate.[28-32] Nuclease and chemical "footprinting" experiments with heteroduplex DNA molecules containing short single-stranded loops indicate that although the enzyme cleaves within the single-stranded region, interactions between the enzyme and the DNA extend beyond the single-stranded region on both sides, spanning a total length of approximately 30 base pairs.[32]

In view of the above information about the enzyme and the way it interacts with its DNA substrate, the location of the active site tyrosine, Tyr-319, in between the two putative DNA binding elements is probably significant. The enzyme most likely consists of at least three domains: the central one contains the active site pocket inside which the transient breakage of a single DNA strand and the crossing of the complementary strand through this break occur, and two flanking domains which interact with the DNA outside the region of DNA strand breakage, crossing, and rejoining. The presence of distinct domains is supported by the recent finding that there are two proteinase-sensitive sites which demarcate the middle domain and the other two putative domains. The native protein is readily cleaved by chymotrypsin after Phe-214 to give a 35,000 Da fragment (A. Mondragon, personal communication). The various features of *E. coli* DNA topoisomerase I discussed above are summarized in Figure 4, in which the effects of C-terminal deletions and four amino acid insertions at precise locations in the protein are also indicated.[18]

Despite the lack of three-dimensional structural data, several inferences regarding the chemical environment of the active site tyrosine can be made. As evidenced by studies with oligonucleotides,[5,33] DNA strand cleavage and rejoining are not necessarily coupled to the passage of another DNA strand through the transient break. The known minimal lengths of oligonucleotides that are required for the cleavage reaction [7 for oligo(dA) and longer for others], and the positions of cleavage in these oligomers (at least four nucleotides from the 3′ ends), suggest that there is a cleft centering around Tyr-319 for the binding of a DNA strand approximately eight nucleotides in length. During catalysis, Tyr-319 attacks at the middle point of the bound DNA segment or upstream of it (the 5′ to 3′ direction of the DNA strand to be cleaved is taken as the downstream direction).

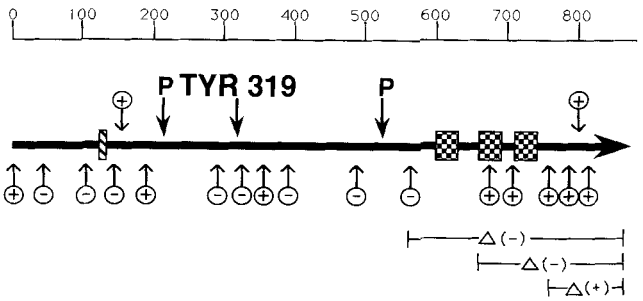

Fig. 4. A sketch of the polypeptide chain of *E. coli* DNA topoisomerase I. The scale at the top indicates the number of amino acids from the amino terminus. The position of the active site tyrosine (319) is indicated. Checkerboard regions indicate putative zinc-binding motifs, and the striped box indicates a putative helix-turn-helix motif. The two P's denote protease-sensitive sites that demarcate a central domain containing Tyr-319. The circles with arrows point to the sites of four amino acid insertions in the protein which were created by linker mutagenesis; the scope of several carbosyl-terminal deletions is indicated by the bracketed triangles. Plus (+) signs indicate those insertions or C-terminal deletions which do not abolish activity; minus (−) signs indicate those which do.

Furthermore, as reported in Depew et al.[29] and Liu and Wang,[30] eubacterial DNA topoisomerase I can form a salt-stable and alkali-cleavable complex with single-stranded or negatively supercoiled DNA. This complex is characterized by its stability in concentrated NaCl or CsCl solution, and by the cleavage of the DNA strand in the complex when exposed to alkali. Strikingly, even though the complex is stable in concentrated salt solutions, the addition of Mg(II) to the concentrated salt solutions, in the presence or absence of excess EDTA, dissociates the complex. From these results we infer that there are two very different modes of interactions between the enzyme and its substrate. In one mode, nonionic interactions are involved. The resistance of the protein–DNA complex to high salt suggests that hydrophobic interactions are likely. In recent crystallographic studies of DNA bound to the 3′-5′ exonuclease site of *E. coli* DNA polymerase I, for example, hydrophobic interactions between protein side chains and single-strand DNA bases are evident.[34]

Upon the addition of Mg(II), a major change occurs in the mode of interactions between the enzyme and the DNA: the enzyme–DNA complex can compete effectively with EDTA for the binding of one or more Mg(II), and a number of nonionic interactions between the enzyme and the DNA are presumably eliminated. We believe that these two modes of in-

teractions are intimately related to the DNA strand passage and rejoining reactions.[29,30]

Hydrogen bonds between bacterial DNA topoisomerase I and DNA bases are also likely. In the mapping of the cleavage sites of DNA topoisomerase I, a cytosine (or thymine in a few cases) is present four nucleotides upstream of the point of cleavage[35,36]; this pyrimidine is probably hydrogen-bonded to a protein side chain group. In the case of bovine pancreatic RNase, H-bonding between a pyrimidine in the substrate and a threonine in the enzyme is important in determining the sequence specificity of the enzyme.[37]

Trans-esterification between nucleic acids and protein moieties (tyrosine, serine, and threonine) occurs in a number of systems besides those involving DNA topoisomerases. Examples include the DNA topoisomerase/DNA strand-transferases that are involved in site-specific recombination or viral DNA replication, and the polio RNA virus protein VPg (or its predecessor) that can form a covalent link with a uridine in the RNA. Among those forming phosphotyrosine linkages, biochemical and mutational analyses have identified the active site tyrosines of phage λ integrase and its homologs,[38] phage ϕX174 gene A protein,[26] and polio VPg.[39] Among those forming phosphoserine linkages, Ser-10 of the enzyme resolvase of bacterial transposons Tn3 and γδ,[40] and Ser-232 of the phage ϕ29 terminal protein have been implicated as the active site residues.[41] In the absence of detailed biochemical and three-dimensional structural data, it is difficult to extract the common features around the active sites. A comparison of a large number of site-specific recombinases that share sequence homologies with λ integrase, including bacterial transposases and recombinases encoded by plasmids of several yeasts, indicates that within a region of 40 amino acids there are three perfectly conserved residues corresponding to His-308, Arg-311, and Tyr-342 of λ integrase, of which the last is the active site tyrosine.[38,42] The presence of these common residues hints that catalysis of DNA strand breakage and rejoining by DNA topoisomerases and strand-transferases may resemble the actions of nucleases (see for example, 34,43). In these reactions, activation of the internucleotide phosphate group is presumably achieved by interactions with protein side chains and/or divalent metal ions. The trans-esterification reactions catalyzed by the DNA topoisomerases and their related enzymes differ from the nucleolytic actions of the nucleases in that a tyrosyl or a sugar hydroxyl group is respectively the nucleophile in the two sequential trans-esterification reactions; in the nuclease-catalyzed reactions, a water hydroxyl is the nucleophile. It is plausible that the nucleophiles in the trans-esterification reactions are also activated by general base catalysis, as in the case of nuclease actions.

Finally, we note that the DNA topoisomerases are important cellular targets of antibiotics and anticancer drugs. It has been known for some time that the antibiotic nalidixic acid and its analogs act by interfering with the DNA strand rejoining step catalyzed by DNA gyrase.[13,14] More recently, a number of antitumor therapeutics have been shown to target the eukaryotic counterpart of bacterial gyrase, eukaryotic DNA topoisomerase II (reviewed in 44–48). A cytotoxic drug, camptothecin, which acts on eukaryotic DNA topoisomerase I by interfering with the DNA rejoining step catalyzed by the enzyme, has also been identified.[49] The last example in particular suggests that bacterial DNA topoisomerase I could be a target of therapeutic agents. In contrast to the type II DNA topoisomerases (DNA gyrase and the structurally and evolutionarily related eukaryotic DNA topoisomerase II) which are essential and are therefore logical targets of cytotoxic drugs, eukaryotic DNA topoisomerase I is nonessential, but has nevertheless been shown to be the sole cellular target of a cytotoxic drug.[50,51] Thus, as information about the active site, protein structure and mechanism of prokaryotic type I topoisomerases accumulates, it should be possible to design bactericidal drugs that interfere with the DNA rejoining step catalyzed by this enzyme, in spite of its dispensability.

ACKNOWLEDGMENTS

We thank Drs. Steve Worland and Mary-Ann Bjornsti for helpful discussions and Drs. Alfonso Mondragon and Yuk-ching Tse-Dinh for sharing their unpublished results with us. This work was supported by a grant from the U. S. Public Health Service (GM 24544). We dedicate this work in memory of Dr. Irving Sigal, whose introduction to molecular biology in 1981 started with the subcloning of the *E. coli topA* gene for sequencing.

REFERENCES

1. Wang, J. C. DNA topoisomerases. Annu. Rev. Biochem. 54:665–697, 1985.
2. Wang, J. C. Recent studies of DNA topoisomerases. Biochim. Biophys. Acta 909:1–9, 1987.
3. Vosberg, H.-P. DNA topoisomerases: Enzymes that control DNA conformation. Curr. Top. Microbiol. Immunol. 114: 19–102, 1985.
4. Been, M. D., Champoux, J. J. DNA breakage and closure by rat liver type I topoisomerase: Separation of the half-reactions by using a single-stranded DNA substrate. Proc. Natl. Acad. Sci. U.S.A. 78:2883–2887, 1981.
5. Tse-Dinh, Y-C. Uncoupling of the DNA breaking and rejoining steps of *Escherichia coli* type I DNA topoisomerase: Demonstration of an active covalent protein-DNA complex. J. Biol. Chem. 261:10931–10935, 1986.
6. Horowitz, D. S., Wang, J. C. Mapping the active site tyrosine of *Escherichia coli* DNA gyrase. J. Biol. Chem. 262: 5339–5344, 1987.
7. Worland, S. T., Wang, J. C. Inducible overexpression, purification and active site mapping of DNA topoisomerase II from the yeast *Saccharomyces cerevisae*. J. Biol. Chem. 264:4412–4416, 1989.
8. Huang, W. M. The 52-protein subunit of T4 DNA topoisomerase is homologous to the GyrA-protein of gyrase. Nucleic Acids Res. 14:7379–7390, 1986.
9. Moriya, S., Ogasawara, N., Yoshikawa, H. Structure and function of the region of the replication origin of the *Ba-*

cillus subtilis chromosome. III. Nucleotide sequence of some 10,000 base pairs in the origin region. Nucleic Acids Res. 13:2251–2265, 1986.
10. Wyckoff, E., Natalie, D., Nolan, J., Lee, M., Hsieh, T.-S. Structure of the Drosophila DNA topoisomerase II gene: Nucleotide sequence and homology among topoisomerases II. J. Mol. Biol. 205:1–13, 1989.
11. Uemura, T., Morikawa, K., Yanagida, M. The nucleotide sequence of the fission yeast topoisomerase II gene: Structural and functional relationships to other DNA topoisomerases. EMBO J. 5:2355–2361, 1986.
12. Tsai-Pflugfelder, M., Liu, L. F., Liu, A., Twewy, K. M., Whang-Peng, J., Knutsen, T., Huebner, K., Croce, C. M., Wang, J. C. Cloning and sequencing of cDNA encoding human DNA topoisomerase II and localization of the gene to the chromosome region 17q21-22. Proc. Natl. Acad. Sci. U.S.A. 85:7177–7181, 1988.
13. Cozzarelli, N.R. DNA gyrase and the supercoiling of DNA. Science 207:953–960, 1980.
14. Gellert, M. DNA topoisomerases. Annu. Rev. Biochem. 50: 879–910, 1981.
15. Liu, L.F. DNA topoisomerases—enzymes that catalyze the breaking and rejoining of DNA. CRC Crit. Rev. Biochem. 15:1–24, 1983.
16. Tse-Dinh, Y-C., Wang, J. C. Complete nucleotide sequence of the *topA* gene encoding *Escherichia coli* DNA topoisomerase I. J. Mol. Biol. 191:312–331, 1986.
17. Messing, J. New M13 vectors for cloning. Methods Enzymol. 101:20–78, 1983.
18. Zumstein, L., Wang, J. C. Probing the structural domains and functions *in vivo* of *Escherichia coli* DNA topoisomerase I by mutagenisis. J. Mol. Biol. 191:333–340, 1986.
19. Nakamaye, K., Eckstein, F. Inhibition of restriction endonuclease Nci I cleavage by phosphorthioate groups and its application to oligonucleotide-directed mutagenisis. Nucl Acids Res. 14:9679–9698, 1986.
20. Sternglanz, R., DiNardo, S., Voelkel, K. A., Nishimura, Y., Hirota, Y., Becherer, K., Zumstein, L., Wang, J. C. Mutations in the gene coding for *Escherichia coli* DNA topoisomerase I affect transcription and transposition. Proc. Natl. Acad. Sci. U.S.A. 78:2747–2751, 1981.
21. Giaever, G.N., Snyder, L., Wang, J.C. DNA supercoiling *in vivo*. Biophys. Chem. 29:7–15, 1988.
22. Rose, R. E. The nucleotide sequence of pACYC184. Nucleic Acids Res. 16:355, 1988.
23. Stankiewicz, A. F., Depew, R. E. A conditional lethal mutation in the *topA* gene of *Escherichia coli*. 83rd Annual Meeting of the American Society of Microbiology, March 6–11, 1983, New Orleans, p. 122, abstr. H102.
24. Bjornsti, M.-A., Wang, J. C. Expression of yeast DNA topoisomerase I can complement a conditional-lethal DNA topoisomerase I mutation in *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. 84:8971–8975, 1987.
25. Keil-Dlouha, V., Zylber, N., Imhoff, J-M., Tong, N-T., Keil, B. Proteolytic activity of pseudotrypsin. FEBS Lett. 16: 291–295, 1971.
26. van Mansfeld, A. D. M., van Teeffelen, H. A. A. M., Baas, P. D., Jansz, H. S. Two juxtaposed tyrosyl-OH groups participate in ϕX174 gene A protein catalysed cleavage and ligation of DNA. Nucleic Acids Res. 14:4229–4238, 1986.
27. Tse-Dinh, Y.-C., Beran-Steed, R. K. *Escherichia coli* DNA topoisomerase I is a zinc metalloprotein with three repetitive zinc-binding domains J. Biol. Chem. 31:15857–15859, 1988.
28. Wang, J. C. Interaction between DNA and an *Escherichia coli* protein ω. J. Mol. Biol. 55:523–533, 1971.
29. Depew, R. E., Liu, L. F., Wang, J. C. Interaction between DNA and *Escherichia coli* protein ω: Formation of a complex between single-stranded DNA and ω protein. J. Biol. Chem. 253:511–518, 1978.
30. Liu, L. F., Wang, J. C. Interaction between DNA and *Escherichia coli* DNA topoisomerase I: Formation of complexes between the protein and superhelical and nonsuperhelical duplex DNAs. J. Biol. Chem. 254:11082–11088, 1979.
31. Kirkegaard, K., Pflugfelder, G., Wang, J. C. The Cleavage of DNA by type I DNA topoisomerases. In: "Recombination at the DNA Level." Cold Spring Harbor Symposia on Quantitative Biology. Vol. XLIX. Cold Spring Harbor, New York: The Cold Spring Harbor Laboratory, 1984: 411–419.
32. Kirkegaard, K., Wang, J. C. Bacterial DNA topoisomerase I can relax positively supercoiled DNA containing a single-stranded loop. J. Mol. Biol. 185:625–637, 1985.
33. Tse-Dinh, Y-C., McCarron, B. G. H., Arentzen, R. and Chowdry, V. Mechanistic study of *E. coli* DNA topoisomerase I: Cleavage of oligonucleotides. Nucleic Acids Res. 11:8691–8701, 1986.
34. Freemont, P. S., Friedman, J. M., Beese, L. S., Sanderson, M. R., Steitz, T. A. Cocrystal structure of an editing complex of Klenow fragment with DNA. Proc. Natl. Acad. Sci. U.S.A. 85:8924–8928, 1988.
35. Tse, Y-C., Kirkegaard, K., Wang, J. C. Covalent bonds between protein and DNA: Formation of phosphotyrosine linkage between certain DNA topoisomerases and DNA. J. Biol. Chem. 255:5560–5565, 1980.
36. Dean, F., Krasnow, M. A., Otter, R., Matzuk, M. M., Spengler, S. J., Cozzarelli, N. R. *Escherichia coli* type-1 topoisomerases: Identification, mechanism, and role in recombination. In: "Structures of DNA." Cold Spring Harbor Symposia on Quantitative Biology. Volume XLVII-Part II. Cold Spring Harbor, New York: The Cold Spring Harbor Laboratory, 1983: 769–777.
37. Blackburn, P., Moore, S. Pancreatic ribonuclease. In: "Nucleic Acids. The Enzymes," Vol. 15B. Boyer, P. (ed.). New York: Academic Press, 1982: 317–433.
38. Pargellis, C. A., Nunes-Duby, S. E., Moitoso de Vargas, L., Landy, A. Suicide recombination substrates yield covalent λ integrase-DNA complexes and lead to identification of the active site tyrosine. J. Biol. Chem. 263:7678–7685, 1988.
39. Ambros, V., Baltimore, D. Protein is linked to the 5′ end of poliovirus RNA by a phosphodiester linkage to tyrosine. J. Biol. Chem. 253:5263–5266, 1978.
40. Hatfull, G. F., Grindley, N. D. F. Analysis of λδ resolvase mutants *in vitro:* Evidence for an interaction between serine-10 of resolvase and site I of *Res*. Proc. Natl. Acad. Sci. U.S.A. 83:5429–5433, 1986.
41. Hermoso, J. M., Mendez, E., Soriano, F., Salas, M. Location of the serine residue involved in the linkage between the terminal protein and the DNA of phage ϕ29. Nucleic Acids Res. 13:7715–7728, 1985.
42. Argos, P., Landy, A., Abremski, K., Egan, J. B., Haggard-Ljungquist, E., Hoess, R. H., Kahn, M. L., Kalionis, B., Narayana, S. V. L., Pierson, L. S., Sternberg, N., Leong, J. M. The integrase family of site-specific recombinases: Regional similarities and global diversity. EMBO J. 5:433–440, 1986.
43. Fersht, A. In: "Enzyme Structure and Mechanism." New York: W. H. Freeman, 1984: 426–433.
44. Ross, W. E. DNA topoisomerases as targets for cancer therapy. Biochem. Pharm. 34:4191–4195, 1985.
45. Chen, G. L., Liu, L. F. DNA topoisomerases as targets in cancer chemotherapy. In: "Annual Report of Medicinal Chemistry," Vol. 21. Bailey, D. M. (ed.). New York: Academic Press, 1986: 257–262.
46. Zwelling, L. A., Silberman, L., Estey E. Intercalator-induced, topoisomerase II-mediated DNA cleavage and its modification by antineoplastic antimetabolites. Int. J. Radiat. Oncol. Biol. Phys. 12:1041–1047, 1986.
47. Potmesil, M., Ross, W. E. eds. Proceedings of the First Conference on DNA topoisomerases in Cancer Chemotherapy. NCI Monographs Vol. 4. Washington, D.C.: National Cancer Institute, 1987: 1–133.
48. Drlica, K., Franco, R. J. Inhibitors of DNA topoisomerases. Biochemistry 27:2253–2259, 1988.
49. Hsiang, Y.-H., Hertzberg, R., Hecht, S., Liu, L. F. Camptothecin induces protein-linked DNA breaks via mammalian DNA topoisomerase I. J. Biol. Chem. 260:14873–14878, 1985.
50. Nitiss, J., Wang, J. C. DNA topoisomerase-targeting antitumor drugs can be studied in yeast. Proc. Natl. Acad. Sci. U.S.A. 85:7501–7505, 1988.
51. Eng, W.-K., Faucette, L., Johnson, R.K., Sternglanz, R. Evidence that DNA topoisomerase I is necessary for the cytotoxic effects of camptothecin. Mol. Pharm. 34:755–760, 1988.