DNA TOPOISOMERASES: Structure, Function, and Mechanism

James J. Champoux

Department of Microbiology, School of Medicine, University of Washington, Seattle, Washington 98195–7242; e-mail: champoux@u.washington.edu

Key Words   Supercoiling, replication, gyrase, decatenation, catalysis

■ **Abstract**   DNA topoisomerases solve the topological problems associated with DNA replication, transcription, recombination, and chromatin remodeling by introducing temporary single- or double-strand breaks in the DNA. In addition, these enzymes fine-tune the steady-state level of DNA supercoiling both to facilitate protein interactions with the DNA and to prevent excessive supercoiling that is deleterious. In recent years, the crystal structures of a number of topoisomerase fragments, representing nearly all the known classes of enzymes, have been solved. These structures provide remarkable insights into the mechanisms of these enzymes and complement previous conclusions based on biochemical analyses. Surprisingly, despite little or no sequence homology, both type IA and type IIA topoisomerases from prokaryotes and the type IIA enzymes from eukaryotes share structural folds that appear to reflect functional motifs within critical regions of the enzymes. The type IB enzymes are structurally distinct from all other known topoisomerases but are similar to a class of enzymes referred to as tyrosine recombinases. The structural themes common to all topoisomerases include hinged clamps that open and close to bind DNA, the presence of DNA binding cavities for temporary storage of DNA segments, and the coupling of protein conformational changes to DNA rotation or DNA movement. For the type II topoisomerases, the binding and hydrolysis of ATP further modulate conformational changes in the enzymes to effect changes in DNA topology.

CONTENTS

369

INTRODUCTION

DNA topoisomerases are marvelous molecular machines that manage the topo-
logical state of the DNA in the cell. These enzymes accomplish this feat by either
passing one strand of the DNA through a break in the opposing strand (type I
subfamily) or by passing a region of duplex from the same or a different molecule
through a double-stranded gap generated in a DNA (type II subfamily). Topoiso-
merases are known that relax only negative supercoils, that relax supercoils of both
signs, or that introduce either negative (bacterial DNA gyrase) or positive super-
coils into the DNA (reverse gyrase). Besides altering the supercoiling of a closed
DNA domain, the strand passing activities of topoisomerases can promote the cate-
nation and decatenation of circular DNAs or the disentanglement of intertwined
linear chromosomes.

 The fundamental need for topoisomerases derives from the double helical struc-
ture of DNA. For most processes that must access the information stored in the
DNA, the two strands of the helix must separate either temporarily, as in tran-
scription or recombination, or permanently, as in replication. The circular na-
ture of bacterial chromosomes and the large size of eukaryotic DNA molecules
preclude a simple winding solution to many of the topological problems that
accompany the DNA transactions that occur in the cell (1). Thus, during DNA
replication, the two strands of the DNA must become completely unlinked by
topoisomerases, and during transcription, the translocating RNA polymerase gen-
erates supercoiling tension in the DNA that must be relaxed (1, 2). The association
of DNA with histones and other proteins introduces supercoiling that requires
relaxation by topoisomerases. In addition, transcription from some promoters
in bacteria requires a minimal level of negative supercoiling, but too much su-
percoiling of either sign is disastrous (3–6). In all cells, completely replicated
chromosomes must be untangled by DNA topoisomerases before partitioning and
cell division can occur. These examples illustrate that not only does DNA struc-
ture lead to topological predicaments that must be solved by topoisomerases, but
that the topological state of the DNA itself must be fine-tuned to optimize DNA
function.

The immense interest in topoisomerases in recent years derives not only from the recognition of their crucial role in managing DNA topology, but also from two major advances in the field. First, a wide variety of topoisomerase-targeted drugs have been identified, many of which generate cytotoxic lesions by trapping the enzymes in covalent complexes on the DNA. These topoisomerase poisons include both antimicrobials and anticancer chemotherapeutics, some of which are currently in widespread clinical use. The extensive literature concerned with drugs that target topoisomerases has been the subject of several recent reviews and is not covered here (7–11). Second, the crystal structures of numerous topoisomerase fragments have been published in the last few years that complement an extensive biochemical literature and provide key insights into how these machines work. The theme for this review is how protein structure contributes to our understanding of the function, catalysis, and mechanism of DNA topoisomerases. Owing in part to space limitations and in part to the recent publication of excellent reviews on the type II enzymes (1, 12–15), the scope of this review is broader for the type I topoisomerases than for the type II enzymes. In view of the expanding literature in the field, citations are restricted to recent work and the most pertinent of the older papers; the reader is referred to the excellent earlier comprehensive review of topoisomerases for the background literature (16).

CLASSIFICATION OF TOPOISOMERASES

DNA cleavage by all topoisomerases is accompanied by the formation of a transient phosphodiester bond between a tyrosine residue in the protein and one of the ends of the broken strand. DNA topology can be modified during the lifetime of the covalent intermediate, and the enzyme is released as the DNA is religated. Those enzymes that cleave only one strand of the DNA are defined as type I and are further classified as either type IA subfamily members if the protein link is to a $5'$ phosphate (formerly called type I-$5'$) or type IB subfamily members if the protein is attached to a $3'$ phosphate (formerly called type I-$3'$). Topoisomerases that cleave both strands to generate a staggered double-strand break are grouped together in the type II subfamily of topoisomerases. Further division of the subfamilies is based on structural considerations. Table 1 lists representatives of the various subfamilies of both prokaryotic and eukaryotic topoisomerases. The eubacterial topoisomerases I and III and the α and β forms of the mammalian topoisomerases II and III are apparently examples of paralogues that arose by gene duplication (17). Sequence alignments for the various topoisomerase families have been published previously (18).

The recent discovery of a novel type II enzyme from the hyperthermophilic archaeon *Sulfolobus shibatae* (19, 20) prompted the division of the type II topoisomerases into the type IIA and type IIB subfamilies, with the *S. shibatae* topoisomerase VI as the prototype of the IIB subfamily. Although monomeric type IA reverse gyrases have been found in hyperthermophilic representatives of both

TABLE 1 Classification of Topoisomerases

Topoisomerase[a]	Subfamily type	Subunit structure	Size(s) (aa)[b]
Eubacterial DNA topoisomerase I (*E. coli*)	IA	Monomer	865
Eubacterial DNA topoisomerase III (*E. coli*)	IA	Monomer	653
Yeast DNA topoisomerase III (*S. cerevisiae*)	IA	Monomer	656
Mammalian DNA topoisomerase IIIα (human)	IA	Monomer	1001
Mammalian DNA topoisomerase IIIβ (human)	IA	Monomer	862
Eubacterial and archaeal reverse DNA gyrase (*Sulfolobus acidocaldarius*)	IA	Monomer	1247
Eubacterial reverse gyrase (*Methanopyrus kandleri*)[c]	IA	Heterodimer	A, 358 B, 1221
Eukaryotic DNA topoisomerase I (human)	IB	Monomer	765
Poxvirus DNA topoisomerase (vaccinia)	IB	Monomer	314
Hyperthermophilic eubacterial DNA topoisomerase V (*Methanopyrus kandleri*)[d]	IB	Monomer	—[e]
Eubacterial DNA gyrase (*E. coli*)	IIA	A$_2$B$_2$ hetero-tetramer	GyrA, 875 GyrB, 804
Eubacterial DNA topoisomerase IV (*E. coli*)	IIA	C$_2$E$_2$ hetero-tetramer	ParC, 752 ParE, 630
Yeast DNA topoisomerase II (*S. cerevisiae*)	IIA	Homodimer	1428
Mammalian DNA topoisomerase IIα (human)	IIA	Homodimer	1531
Mammalian DNA topoisomerase IIβ (human)	IIA	Homodimer	1626
Archaeal DNA topoisomerase VI (*Sulfolobus shibatae*)	IIB	A$_2$B$_2$ hetero-tetramer	A, 389 B, 530

[a]The source of the most extensively studied family member is given in parentheses. The top portion of the table lists the type I topoisomerases; the bottom portion the type II enzymes.

[b]The subunit sizes are those corresponding to the most extensively studied family member.

[c]Included as the only known reverse gyrase with a heterodimeric structure.

[d]Only known representative at present. Probably present in other hyperthermophilic eubacteria.

[e]Gene not yet cloned; purified protein has a molecular size of 110 kDa.

prokaryotic domains of life, the *Methanopyrus kandleri* reverse gyrase is included in the table because it is the only known example of a heterodimeric type IA enzyme (21). The availability of complete genome sequences for many bacteria and yeast indicates that no additional homologues of the known enzymes are likely to be discovered. It remains to be seen whether new and distinct subclasses of topoisomerases will be identified in the future.

Besides those topoisomerases listed in Table 1, a number of viral or plasmid-encoded topoisomerases have been described through the years. Thus, in addition

to the well-studied vaccinia virus type IB topoisomerase, type II DNA topoisomerases have been described for bacteriophage T4 (22), African swine fever virus (ASF) (23), and paramecium bursaria chlorella virus 1 (PBCV-1) (24). The PBCV-1 topoisomerase is the smallest known type II enzyme to date with a molecular weight of 120,000. Plasmid-encoded type IA topoisomerases have been identified in a number of Gram-negative as well as Gram-positive bacterial genera [reference (25) and references cited therein] that have features characteristic of eubacterial topoisomerases I and III. Their function remains obscure, although in at least one case it has been shown that a specific substrate for the plasmid-encoded topoisomerase is generated during the initiation of plasmid DNA replication by DNA polymerase I (25).

CELLULAR ROLES OF TOPOISOMERASES

As a backdrop to the discussion to follow, the cellular roles of the various topoisomerases are reviewed, especially as they relate to enzyme function and substrate specificity. The roles of topoisomerases in various cells types are the subject of three recent reviews (26–28).

Eubacteria

Escherichia coli   The four DNA topoisomerases found in *E. coli* consist of the two type IA enzymes, DNA topoisomerases I and III, and the two type IIA enzymes, DNA gyrase and DNA topoisomerase IV. Although some overlap of function has been shown genetically, each of the DNA topoisomerases appears optimized to carry out its own particular set of topological manipulations. DNA gyrase is the only known topoisomerase able to generate negative supercoiling at the expense of ATP hydrolysis and is responsible for global generation of negative supercoils in the bacterial chromosome. Such global supercoiling in combination with the activity of the *E. coli* Muk proteins is essential for chromosome condensation leading to proper chromosome partitioning at cell division (29, 30). The recent realization that besides its decatenating role in replication, topoisomerase IV also relaxes negative supercoils in the cell (32), implicates topoisomerase IV along with topoisomerase I (27) as activities that prevent excessive negative supercoiling by DNA gyrase. Together, topoisomerases I and IV along with DNA gyrase set the steady-state level of negative supercoiling that is required for the initiation of replication and for transcription from at least some promoters (32). Transcription itself generates positive supercoils ahead of and negative supercoils behind the translocating RNA polymerase that are rapidly resolved by DNA gyrase and DNA topoisomerase I, respectively.

Fork movement during replication of a circular DNA can generate topological changes in both the unreplicated region ahead of the fork and in the already replicated region behind the fork. Early after initiation, movement of a replication fork causes overwinding of the DNA in the unreplicated region of the theta

intermediate, and the resulting positive supercoils are rapidly removed by DNA gyrase (28). However, as replication proceeds the overwinding in the unreplicated region can diffuse back into the already-replicated region to cause interwinding of the two daughter duplexes (33). Peter et al (34) refer to these interwindings as precatenanes because if not removed, they cause the two daughter molecules to be catenated at the end of replication. Recent evidence confirms that excess helical windings generated by replication can be distributed both in front of and behind the replication fork (34). Therefore, fork movement during replication can be maintained not only by DNA gyrase relaxing positive supercoils in front of the fork but also through the "unwinding" of the daughter duplexes behind the replication fork. Of the two type II enzymes in *E. coli*, topoisomerase IV is much more effective at decatenating DNA than at relaxing positive supercoils whereas the converse is true for DNA gyrase (28, 35, 36). This functional difference between the two enzymes (37) suggests that topoisomerase IV is primarily responsible for unlinking the precatenanes that occur behind the replication fork as well as the catenanes that are generated at the end of replication by the failure of DNA gyrase to remove the last of the positive supercoils when the two replication forks converge (37).

Surprisingly, topoisomerase III supports fork movement on a circular DNA in vitro despite its inability to relax positive supercoils (38). This observation is best explained by the ability of topoisomerase III, but not topoisomerase I, to "decatenate" the precatenanes behind the fork by acting at nicks or gaps that are present in the replicating DNA (38–40). Thus, topoisomerases III and IV appear to have overlapping functions. In vivo, it seems likely that precatenane unlinking is carried out by topoisomerase III acting in the region just behind the replication fork where the template for discontinuous synthesis remains single-stranded and by topoisomerase IV acting in more distal regions where the daughter DNAs are completely duplex. Under normal conditions where topoisomerase I is acting to prevent excess negative supercoiling, topoisomerase III, which requires a hyper-negatively supercoiled substrate, is probably not involved in supercoil relaxation in the cell (39).

Recently an interaction has been demonstrated between *E. coli* topoisomerase III and the *E. coli* RecQ helicase in which the helicase produces a substrate for catenation of duplex molecules by the topoisomerase (41). This finding is especially interesting in view of the finding that human topoisomerases III act in concert with human homologues of the bacterial RecQ helicase (see below).

Other Eubacteria   An analysis of the genome sequences of the 17 mesophilic eubacterial organisms for which a complete annotated sequence is available at the time of this writing reveals that several bacteria (*Haemophilus influenzae*, *Bacillus subtilis*, and *Xylella fastidiosa*) have the same topoisomerase complement as *E. coli*. It seems likely that these enzymes share the same relative distribution of functions discussed above. However, given the apparent functional redundancy of topoisomerases III and IV mentioned above, it is not surprising to find a number of mesophilic eubacteria that appear to lack a homologue of topoisomerase III

but that still possess the other three topoisomerases. This group includes *Borrelia burgdorferi*, *Chlamydophila pneumoniae*, *Chlamydia trachomatis*, *Mycoplasma pneumoniae*, *Mycoplasma genitalium*, *Neisseria meningitidis*, *Rickettsia prowazekii*, *Synechocystis PCC6803*, and *Ureaplasma urealyticum*. Apparently in these cases topoisomerase IV is solely responsible for unlinking precatenanes as well as daughter molecule catenanes.

Based on the current level of genome annotation, a number of mesophiles (*Campylobacter jejuni*, *Deinococcus radiodurans*, *Mycobacterium tuberculosis*, and *Treponema pallidum*) appear to possess only two topoisomerase genes: one is homologous to the gene encoding *E. coli* DNA gyrase and a second to the *E. coli* topoisomerase I *topA* gene. This configuration probably represents the minimal topoisomerase makeup for a bacterial cell. In these bacteria, the DNA gyrase as the sole type II enzyme would negatively supercoil the DNA, relax positive supercoils associated with transcription and replication, and decatenate replicated DNA. In turn, topoisomerase I would function to relax negative supercoils associated with transcription and to prevent excessive negative supercoiling by DNA gyrase.

Complete genome sequences are available for two eubacterial hyperthermophiles, *Aquifex aeolicus* and *Thermotoga maritima*. Like all of the hyperthermophilic bacteria, both organisms code for a reverse gyrase that uses the energy from ATP hydrolysis to introduce positive supercoils into the chromosomal DNA. The exact role of reverse gyrase in hyperthermophiles remains uncertain, but it seems likely that positive supercoiling of the DNA functions to counteract the helix unwinding and strand separation effects of growth at high temperature (42). In addition to reverse gyrase, both of these organisms have genes coding for DNA gyrase and for a homologue of the *E. coli* topoisomerase type IA enzymes. All three topoisomerases have been characterized biochemically in *Thermotoga maritima* (43–45); the type IA activity was found to be more similar to *E. coli* topoisomerase I than to topoisomerase III and thus probably functions to relax negative supercoils associated with transcription. DNA gyrase, the only type II enzyme activity, is likely important as a decatenating activity during DNA replication, and may also function to oppose reverse gyrase and prevent excessive positive supercoiling.

Topoisomerase V, originally identified in the hyperthermophile *Methanopyrus kandleri* (46, 47), remains the only known example of a type IB topoisomerase in bacteria. *M. kandleri* possesses, in addition to topoisomerase V, a novel heterodimeric reverse gyrase (21), a type II topoisomerase, and probably also a type IA topoisomerase. How these topoisomerases collaborate to manage the topological state of the DNA in this hyperthermophilic methanogen remains unknown.

Archaebacteria

The archaebacteria for which complete genome sequences with annotations are available are the hyperthermophiles *Aeropyrum pernix*, *Archaeoglobus fulgidus*, *Methanococcus jannaschii*, *Pyrococcus horikoshii*, and *Pyrococcus abyssi*. With one exception, these organisms share the same ensemble of topoisomerases: a

reverse gyrase, a type IA topoisomerase, and topoisomerase VI. Topoisomerase VI may be exclusive to the archaea. In addition to these three enzymes, *Archaeoglobus fulgidus* also appears to possess a typical DNA gyrase. While functional assignments can only be provisional in the absence of a detailed characterization of the enzymes and the topological state of the chromosomes in these organisms, it seems likely that the functions will parallel those described above for the hyperthermophilic eubacteria except that topoisomerase VI is the type II enzyme responsible for decatenating replication intermediates (48). The observation that purified topoisomerase VI can relax both positive as well as negative supercoils (19, 20) suggests that it may also prevent excessive positive supercoiling generated by transcription and/or reverse gyrase. Consistent with its presence exclusively in the hyperthermophilic bacteria, reverse gyrase is not found in the mesophilic members of the archaeal domain (49).

Yeasts

Saccharomyces cerevisiae and *Schizosaccharomyces pombe* code for three topoisomerases that belong to the three of the four known subfamilies. In both organisms, topoisomerase I (IB subfamily) is dispensable for growth, but topoisomerase II (IIA subfamily) is absolutely required to "decatenate" linked chromosomes and prepare chromosomes for segregation at mitosis (16, 50, 51). Unlike DNA gyrase, topoisomerase II cannot introduce negative supercoils but can relax supercoils of both signs, and is therefore able to substitute for topoisomerase I in its absence. In the wild-type situation, topoisomerase I probably provides the swivels for fork movement during DNA replication by acting in the unreplicated region between converging replication forks. Topoisomerase I also removes both the negative and positive supercoils associated with transcription. It is likely that topoisomerase II is involved in these processes as well, but the relative contributions of the two enzymes to replication and transcription remain unclear. Similar to bacteria, yeasts poorly tolerate excessive negative supercoiling, and in the absence of either topoisomerase I or II, mitotic recombination occurs at repeated sequences within the genome (6).

Although topoisomerase III (type IA subfamily) is dispensable for growth in *S. cerevisiae*, in its absence, the cells grow slowly, exhibit elevated levels of mitotic recombination, and fail to sporulate owing to a defect in meiotic recombination (5, 52). On the other hand, topoisomerase III is required in *S. pombe* for sustained cell division (53, 54). Despite the different phenotypes of mutants lacking topoisomerase III, in both organisms it appears that there is a functional association between topoisomerase III and the Sgs1 helicase (or its homologue in *S. pombe*), an enzyme that belongs to the RecQ family of helicases (52–56). It has been suggested that the interaction between these two proteins may prevent the formation of lethal intermediates during DNA recombination and perhaps DNA replication (52, 56, 57). The targets for the combined activity of these two enzymes remain unknown.

Higher Eukaryotes

Like the yeasts, all higher eukaryotes contain a single topoisomerase I enzyme that plays a major role in supporting fork movement during replication and in relaxing transcription-related supercoils. Topoisomerase I is indispensable during development and probably also during cell division (58, 59). The realization that chromosome condensation during mitosis involves the generation of right-handed solenoidal supercoils by proteins belonging to the SMC family (29, 60) means that topoisomerase I would also be required at this stage of the cell cycle to relax compensatory negative supercoils.

Unlike the yeasts, most higher eukaryotes (*Drosophila* being the known exception) appear to contain two type IIA isoforms termed topoisomerases IIα and IIβ. Topoisomerase IIα (*Drosophila* topoisomerase II) is essential and found in all cell types; the activity is responsible for unlinking intertwined daughter duplexes during DNA replication and likely also contributes to DNA relaxation during transcription and other cellular processes. Although topoisomerase IIβ is dispensable for cell division (61, 62), mice with a homozygous defect at the *top2β* locus die at birth with abnormal neural development (63). This result may implicate topoisomerase IIβ in suppressing recombination or supporting transcription specifically in early developing neurons.

Higher eukaryotes also contain two isoforms of topoisomerase III. In mice, the topoisomerase IIIα isoform is required early during embryogenesis and has been hypothesized to assist in the unlinking of precatenanes during cellular DNA replication (64). Although no information is yet available concerning the role of the mammalian topoisomerase IIIβ, preliminary evidence indicates that the *Drosophila* topoisomerase IIIβ may not be essential for cell viability or development (65). Based on the findings in yeast, it was suggested that topoisomerases IIIα and IIIβ might interact with one or more of the mammalian homologues of the *SGS1* gene (64). Two such genes are *BLM* and *WRN*, which code for helicases that, when mutated, cause Bloom's syndrome and Werner's syndrome, respectively (66, 67). Support for such a view comes from the recent finding that human topoisomerase IIIα physically interacts with the BLM protein (64) and that another member of the human RecQ family, RecQ5β, interacts with both topoisomerases IIIα and IIIβ (68). Thus the situation in humans parallels that observed for yeasts and suggests that whatever their molecular functions, the topoisomerase III isoforms probably act in concert with a member of the RecQ family of helicases.

TYPE IA DNA TOPOISOMERASES

General Features

The topoisomerases belonging to the type IA subfamily share the following properties (16): (a) They are all monomeric with the exception of the *Methanopyrus kandleri* reverse gyrase (21). (b) Cleavage of a DNA strand is accompanied by

covalent attachment of one of the DNA ends to the enzyme through a $5'$ phosphodiester bond to the active site tyrosine. (c) All require Mg(II) for the DNA relaxation activity. (d) Plasmids containing negative, but not positive, supercoils are substrates for the relaxation reaction. (e) The relaxation of negative supercoils does not go to completion. (f) Consistent with the last two points, all enzymes in this subfamily require an exposed single-stranded region within the substrate DNA (69). (g) Relaxation of the DNA changes the linking number in steps of one. (h) Where the process has been examined, in addition to the ability to relax negative supercoils, these enzymes can catalyze the knotting, unknotting, and interlinking of single-stranded circles as well as the knotting, unknotting, catenation, and decatenation of gapped or nicked duplex DNA circles.

Members of the type IA subfamily of enzymes can be distinguished by two properties that likely reflect their roles in the cell. First, the enzymes vary with respect to the degree of negative supercoiling required for a DNA to be an effective substrate in vitro. For example, the bacterial topoisomerases I will relax native plasmid DNAs isolated directly from bacteria, whereas the *E. coli* and eukaryotic topoisomerases III (α and β) characterized to date require a hypernegatively supercoiled plasmid DNA substrate (39, 65, 70). The degree of negative supercoiling required reflects the extent to which the enzymes can facilitate duplex opening by binding preferentially to single strands (16). Apparently, both the initial superhelix density as well as the final endpoint are important in vivo. Thus, in the case of *E. coli* topoisomerase I, which functions to prevent excess negative supercoiling, the enyzme does not act until some threshold has been exceeded and then only relaxes the DNA back to the supercoiling range that is optimal for gene function. Similarly, given the substrate specificity of the topoisomerases III, and the apparent collaboration between these enzymes and helicases, it appears that the topoisomerases only function when the helicases have generated extensive duplex unwinding.

A second and possibly related property that distinguishes the members of the type IA subfamily is the relative effectiveness of supercoil relaxation versus DNA catenation or decatenation. *E. coli* topoisomerase III is much more effective at decatenation than topoisomerase I, consistent with its support role in precatenane unlinking during DNA replication (39, 40). Conversely, *E. coli* topoisomerase I is relatively ineffective at decatenation and is therefore unlikely to facilitate fork movement during replication. Detailed information concerning the relaxation versus decatenation activities of the eukaryotic topoisomerases III is lacking, but their strong preference for single-stranded DNA is consistent with a role in unlinking precatenanes during replication (64).

Although considerable sequence diversity is observed among the known members of the type IA subfamily, certain features are discernible that appear to identify particular subclasses within the subfamily (65). As the prototype of this subfamily, the 865 amino acid *E. coli* topoisomerase I can be divided into three domains (Figure 1). The first 582 N-terminal amino acids correspond to a core "cleavage/strand passage" domain containing the active site tyrosine at position 319. Expression of a 596 amino acid N-terminal fragment of *E. coli* topoisomerase I
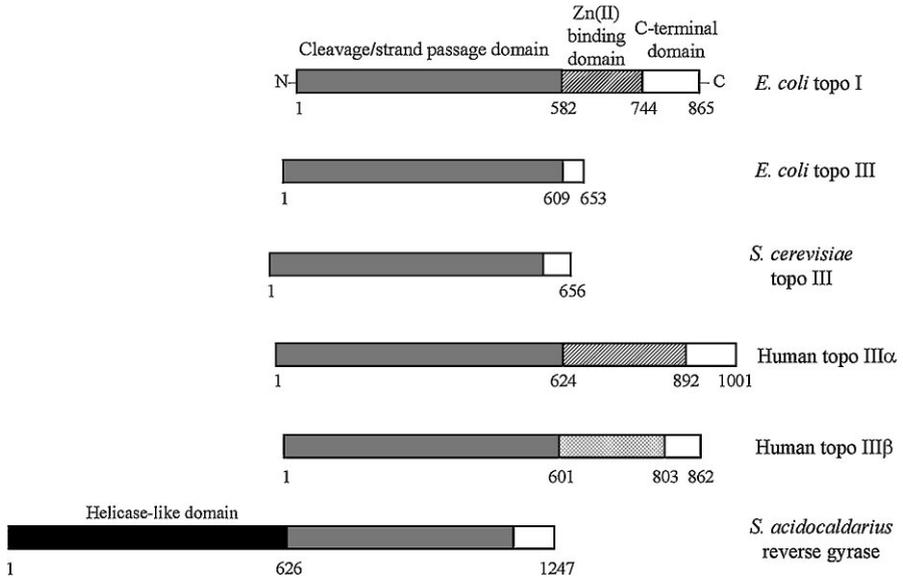
Figure 1 Domain structure of type IA topoisomerases. The sequences of the type IA topoisomerases from the indicated organisms were aligned based on homology to the cleavage/strand passage domain of *E. coli* topoisomerase I, and the domains are drawn approximately to scale (18, 65, 70). The boxes corresponding to the various domains are coded as follows: gray, cleavage/strand passage domain; black, "helicase-like domain" of reverse gyrase; unfilled, basic C-terminal domain; diagonal striping, *E. coli*-like Zn(II) binding domain; stippling, Zn(II) binding domain found in human and *Drosophila* (not shown) topoisomerases IIIβ (65). Although all of the type IA enzymes possess a basic C-terminal domain, only *E. coli* topoisomerase I and human topoisomerase IIIα exhibit significant sequence homology throughout this region. The domain boundaries lacking numerical labels are only approximate.

yields a protein that retains the ability to cleave a single-stranded oligonucleotide but is unable to relax plasmid DNA (71). This catalytic domain is followed by a Zn(II)-binding domain consisting of 162 amino acids that contains three tetracysteine motifs (72, 73). An N-terminal fragment containing the first of the three tetracysteine motifs still binds Zn(II) and is active for relaxation indicating that this region of the protein is required for the strand passage reaction (74, 75). Although the C-terminal-most 121 amino acids are dispensable for activity in vitro, this domain of the protein is rich in basic amino acids and contributes to substrate binding with a preference for single-stranded DNA (76–78). A recent analysis based on sequence comparisons reveals that the C-terminal domain is related to the adjacent Zn(II) binding domain and contains two regions homologous to a tetracysteine motif that have lost the ability to bind Zn(II) (79).

How do other members of the type IA subfamily compare to the prototypic *E. coli* enzyme (Figure 1)? Notably, both yeast and *E. coli* topoisomerases III lack a

Zn(II) binding domain altogether. Human topoisomerase IIIα has a Zn(II) binding domain with significant homology to that found in *E. coli*, but only the first tetra-cysteine motif found to be essential in *E. coli* is present (70). The topoisomerases IIIβ from a phylogenetically diverse group that extends from *Drosophila* to humans all possess a highly conserved set of eight CXXC motifs that are distinct from the tetracysteine motifs found in the bacterial topoisomerases I and that could correspond to four zinc fingers (65). Although distinct in sequence and probably structure, the Zn(II) binding domains in the different classes likely share the same DNA binding function. The reverse gyrase may also possess one or more Zn(II) binding regions (27, 79, 80). As can be seen from the alignments in Figure 1, members of the type IA subfamily consistently possess a relatively basic C-terminal domain that is not conserved and presumably contributes to DNA binding.
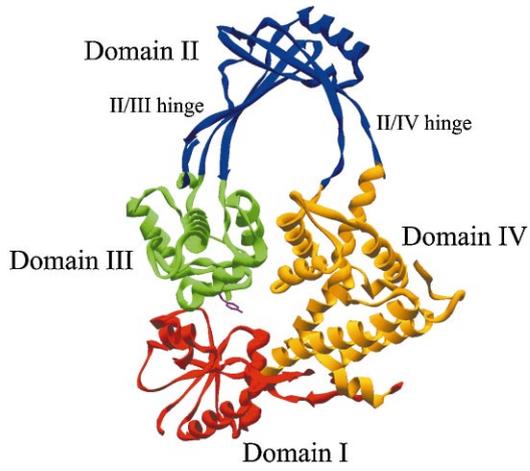
The remainder of this section on type IA topoisomerases considers the structures and specific properties of the well-studied bacterial members of this subfamily.
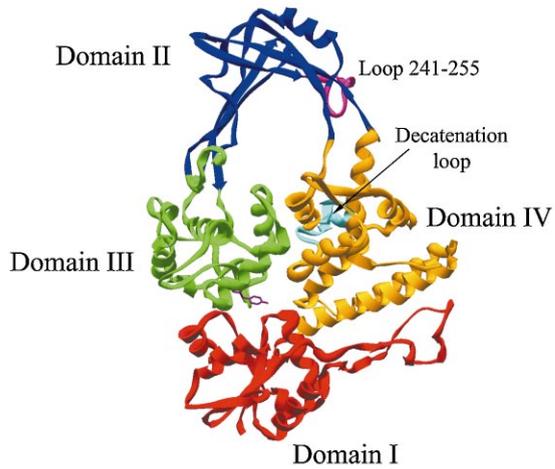
E. coli DNA Topoisomerase I

Structure   The determination of the crystal structure of a 67 kDa N-terminal fragment of *E. coli* topoisomerase I to 2.2 Å (81) represented a major advance in our understanding of how these enzymes function. The 590 amino acid N-terminal fragment observed in the structure corresponds to the cleavage/strand passage domain (Figure 1), and although it is inactive for DNA relaxation, it retains the ability to cleave a single-stranded oligonucleotide (71). The protein assumes the configuration of a flattened torus with a positively charged hole large enough to accommodate double-stranded DNA (Figure 2*a*). The remarkable arrangement of the four domains of the protein is consistent with the conformational plasticity that is likely required for DNA binding and relaxation. Domain I is composed of four criss-crossed β-strands surrounded by four α-helices that are arranged in a Rossmann fold, a structure previously found to be associated with proteins that bind nucleotides and enzymes that catalyze phosphotransfer reactions (82, 83). From domain I, the polypeptide chain travels through domain IV, which is mostly

Figure 2   Crystal structures of *E. coli* topoisomerases I and III. The toroidal structure of *E. coli* topoisomerase I (aa 32–509, pdb entry 1ECL) (panel *a*) and topoisomerase III (aa 1–609, pdb entry 1D6M) (panel *b*) are shown looking directly through the hole. For *E. coli* topoisomerase I (81), domains I (residues 32–63, 72–157), II (214–278, 406–433, 438–475), III (279–405, 434–437), and IV (64–71, 158–213, 476–590) are colored red, blue, green, and orange, respectively. The domains for *E. coli* topoisomerase III follow the same convention as described for topoisomerase I (81). The two hinge points between domain II and the other domains are labeled II/III hinge and II/IV hinge in panel *a*. The active site tyrosines are shown in ball and stick and colored magenta. For *E. coli* topoisomerase III in panel *b*, Loop 241–255 is colored magenta and the decatenation loop is shown in cyan. The drawings in this and subsequent structure figures were created using Swiss-Pdb Viewer software (v3.51) (Glaxo Wellcome Experimental Research) (85a).

a

Domain II

II/III hinge

II/IV hinge

Domain IV

Domain III

Domain I

E. coli topoisomerase I

b

Domain II

Loop 241-255

Decatenation loop

Domain IV

Domain III

Domain I

E. coli topoisomerase III

α-helical, up through domain II on the top before descending to form the α-helical domain III. The chain then reverses direction and returns to domain IV, traversing again through the β-strands that are laced together within domain II. The active site Tyr319 is located within domain III, but is buried at the interface between domains III and I in a region which structurally resembles the fold of the DNA-binding region of the *E. coli* catabolite activator protein (CAP). Importantly, although domain III is in intimate contact with domain I, it is tethered to domain IV by the arch of domain II that contains two possible hinge regions to allow domain III to lift away from domain I and provide a channel for DNA entry into and out of the central cavity (see below). Although not visible in the view shown in Figure 2*a*, a large cleft extending from the lower region of domain IV into the top of domain I is believed to provide a binding groove for the single strand to be cleaved. The preference for a cytosine base in the cleaved strand 4 nucleotides 5′ of the cleavage site (84, 85) is likely determined by a base-specific contact involving the cleft region of the Rossmann fold in domain I.

The structure of a 14-kDa fragment of *E. coli* topoisomerase I corresponding to the C-terminal DNA binding region of the protein (see Figure 1) has been solved by NMR (86). The single-stranded DNA binding region is localized to a positively charged β-sheet within the structure. The spatial relationship of this domain to the remainder of the protein shown in Figure 2*a* is unknown, but the fact that it is dispensable for cleavage and DNA relaxation suggests that it lies adjacent to domain IV and simply provides an extended single-stranded DNA binding region.

Mechanism of DNA Relaxation   To achieve the strand passage event that is required for either DNA relaxation or any of the catenation/decatenation reactions catalyzed by topoisomerase I, it was proposed that the enzyme cleaves a single strand of DNA and by holding onto both ends at the site of the break, bridges the gap through which the intact strand is passed. This mechanism has been referred to as the enzyme-bridging model for DNA relaxation (87–89). The crystal structure of the 67-kDa fragment of the *E. coli* enzyme suggests how such a reaction occurs (81), and Figure 3 shows the proposed model in schematic form. Since the active site Tyr319 is buried in the structure, it is proposed that as the single-stranded DNA binds to the cleft, domain III undergoes a conformational adjustment to place the nucleophilic O-4 oxygen of the tyrosine side chain in position to attack the scissile phosphate. After cleavage, the active site Tyr319 is covalently bound to the 5′ phosphate on one end of the cleaved strand, and the other end is proposed to occupy a nucleotide binding site at the end of the cleft in the Rossmann fold of domain I (81, 90). Immediately upon cleavage, domain III, which is holding onto the 5′ end of the broken strand, lifts away from domain I (Figure 3*b*) to create a gap through which is passed either the intact strand (Figure 3*c*) or the duplex region of a DNA to be catenated. It was originally proposed that the hinge for this conformational change was located between domains II and IV (II/IV hinge in Figure 2*a*), but the recent solution of the crystal structure of a fragment comprising domains II and III (91) indicates that the β-strands joining domains II and III (II/III hinge in Figure 2*a*)

have an unusual amount of flexibility and potentially also contribute to the hinging required to open the "gate" in the cleaved strand. Once the intact strand has moved into the hole of the torus, the clamp closes, and the cleaved strand is religated (Figure 3*d*). The protein must then open and close a second time to release the passed strand (Figure 3*e* and 3*f*) to complete the cycle. Once reset, the enzyme can dissociate from the DNA or act processively to carry out another cycle of strand passage.

Since this series of reactions occurs without an external energy source such as ATP, the protein conformational changes that accompany each step must be driven by DNA interactions with the protein. Although Mg(II) is not required for cleavage, it is required for DNA relaxation (84, 92–94). Recently, the highly conserved acidic triad Asp111, Asp113, and Glu115 within the Rossmann-like fold has been shown to bind Mg(II), and this binding is required for at least some of the conformational changes that accompany the relaxation reaction (93). The negative supercoiling of the substrate DNA provides directionality to the overall process and likely facilitates protein-DNA interactions.

Although the proposed model for DNA relaxation by DNA topoisomerase I is appealing and almost certainly correct in most aspects, the role of the Zn(II) binding domain that is required for DNA relaxation remains unknown. At present, the location of this region in relation to the 67-kDa catalytic fragment remains open to conjecture. Since the Zn(II) domain is absolutely required for the relaxation and decatenation/catenation reactions, it is tempting to propose that, by binding to the DNA, it facilitates DNA movement into and out of the central hole. However, if this hypothesis is correct, then it is puzzling how such movement is accomplished by the related bacterial and yeast topoisomerases III that lack a Zn(II) binding domain.

Mechanism of Catalysis   The proximity of a number of acidic residues to the active site Tyr319 originally suggested that bound Mg(II) might be involved in the catalytic mechanism (81). However, this hypothesis appears to be ruled out by the absence of a Mg(II) requirement for cleavage and probably also for religation (94). Although mutation of a number of other residues near the active site tyrosine to alanines had little effect on the cleavage reaction, changing Glu9 to alanine abolished cleavage whereas a change to glutamine at this position had little effect (94). This result may implicate the side chain of Glu9 in catalysis, perhaps through an interaction with the 3′ bridging oxygen of the scissile phosphate.

A more detailed description of the catalytic mechanism requires a determination of the amino acid side chains involved in stabilization of the transition state as well as whether general acid-base catalysis is implicated in the reaction (81, 94). Such a description is confounded by the lack of a structure containing bound DNA and by the likelihood that a conformational change occurs after DNA binding and before cleavage, which changes the positions of residues proximal to the active site Tyr319. However, one of the crystal forms recently solved for a fragment comprising domains II and III reveals a conformation that may approximate this form of the enzyme (91). In this structure, the guanidinium group of Arg321,
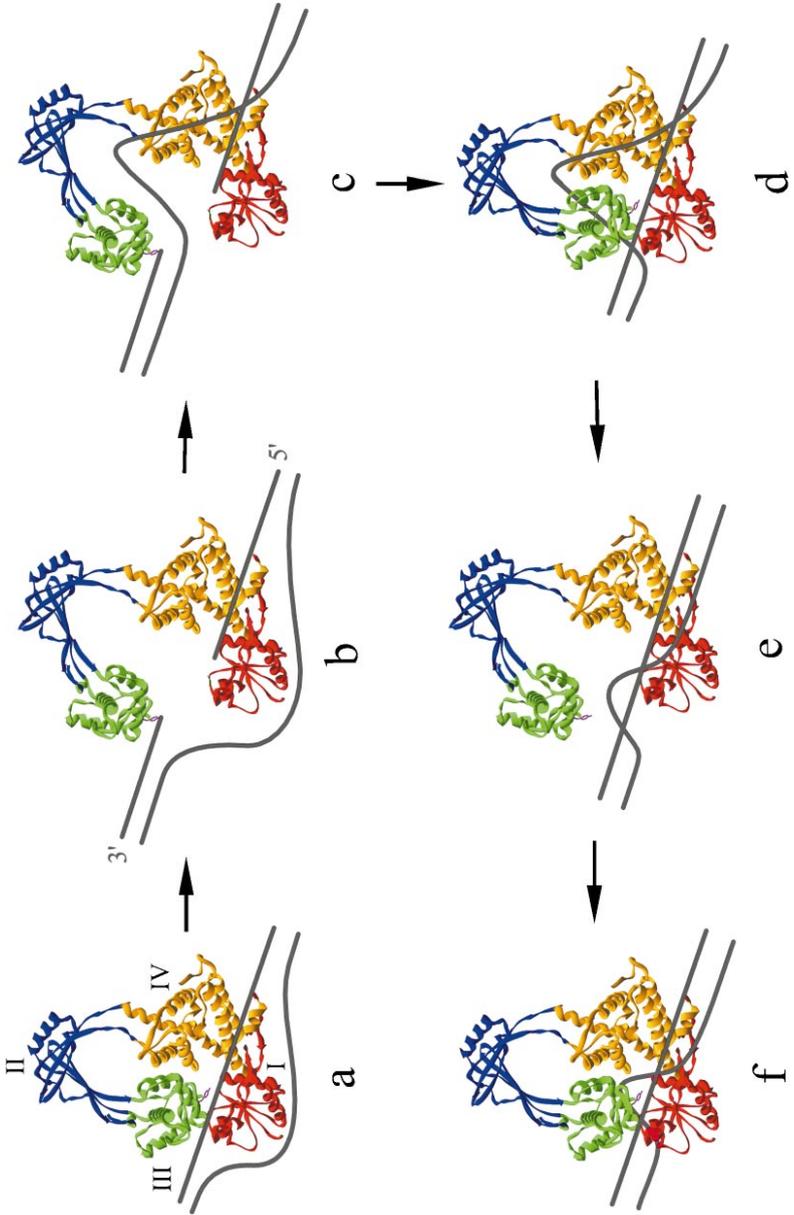
which is ∼4.5 Å away from Tyr319 in the 67-kDa closed structure, has moved to within hydrogen-bonding distance of the oxygen atom of the Tyr319 hydroxyl. This observation led to the suggestion that the positively charged arginine side chain might promote nucleophilic attack of the tyrosine O-4 atom on the scissile phosphate by stabilizing the phenolate anion. An earlier finding that Arg321 can be replaced by lysine, but only poorly by alanine, led to the same proposal for the role of Arg321 in catalysis (94). Interestingly, an invariant arginine is similarly positioned near the active site tyrosines of human topoisomerase I (type IB) and the type II topoisomerases (see below).

E. coli DNA Topoisomerase III

The first 609 N-terminal amino acids of *E. coli* topoisomerase III are homologous to the first 581 amino acids of *E. coli* topoisomerase I, with 26% of the residues being identical (Figure 1). The C-terminal domains of the two proteins lack any homology, and topoisomerase III is missing the Zn(II) binding domain altogether. However, as with topoisomerase I, the C-terminal domain of topoisomerase III (44 amino acids) contributes to DNA binding (78), and since the C-terminal domains of the two enzymes are distinct, this region of topoisomerase III might contribute to the powerful decatenating potential of the enzyme (39). Besides its more potent decatenating activity, *E. coli* topoisomerase III is also distinct from topoisomerase I in its ability to cleave and decatenate RNA molecules (95). Whether this latter activity has a physiological role remains to be determined.

The crystal structure of the entire cleavage/strand passage domain of *E. coli* topoisomerase III protein (Figure 2*b*) (96) exhibits a remarkable resemblance to the structure of the homologous 67-kDa fragment (first 590 amino acids) of *E. coli* topoisomerase I. Topoisomerase III is toroidal in shape with four distinct domains

---

Figure 3   Proposed mechanism of relaxation by *E. coli* topoisomerase I. Shown schematically is a series of proposed steps for the relaxation of one turn of a negatively supercoiled plasmid DNA. The same mechanism would apply in the case of decatenation or knotting by substituting a DNA segment from another molecule or from another region of the same molecule for the intact strand in the figure, respectively. The two strands of the DNA are shown as dark gray lines (not to scale). The four domains of the protein are labeled in panel *a* and colored as described for Figure 2*a*. The strand to be cleaved is shown bound to the surface of the protein at the approximate location of the large cleft, and its polarity is indicated in panel *b*. In panels *a–d*, the length of the intact strand that is passed through the open gate is exaggerated to simplify the drawing. The conformational change that occurs upon binding the substrate to expose the active site Tyr319 has not been characterized and is not shown. The protein is proposed to oscillate between a closed conformation (panels *a*, *d*, and *f*) and an open conformation (panels *b*, *c*, and *e*) that provides access of the DNA to the central hole. For the purposes of generating the figure, the open conformation was modeled by allowing movement at both the II/III and II/IV hinges shown in Figure 2*a*. See text for detailed description of each step of the cycle.

that closely parallel those found in topoisomerase I. Most notably, the structures are very similar within the CAP region of domain III that contains the active site tyrosine residues and surrounding amino acids. The cleft where the single-stranded substrate has been proposed to bind appears a little deeper and better defined in topoisomerase III. Despite the obvious similarities, the presence of short deletions and insertions of amino acids results in subtle differences in the shapes of the four domains. In addition, the orientation of the four domains relative to each other differs somewhat between the two proteins. Two insertions that contribute unique structures to topoisomerase III are referred to as loop 241–255 and the decatenation loop; these loops are highlighted in Figure 2*b*.

Although the functional significance of loop 241–255 remains unknown, a variant of *E. coli* topoisomerase III was recently engineered in which the 17 amino acid decatenation loop was removed and replaced by the homologous region from *E. coli* topoisomerase I (97). Although this change reduced the relaxation activity of the enzyme by approximately 20-fold, the mutant enzyme completely lacked decatenating activity on either a gapped, multiply catenated circular dimer substrate or with replication intermediates generated in an in vitro plasmid replication system. The decatenation loop is rich in basic amino acids and is located at the edge of the hole in the torus, extending away from the body of the protein (Figure 2*b*). These properties suggest a possible direct interaction with DNA. Since singly catenated, but not multiply catenated dimers can be unlinked by topoisomerase I (88), whereas topoisomerase III can unlink both types of catenanes, the decatenation loop apparently endows topoisomerase III with the ability to act processively and carry out multiple decatenation reactions. A model has been proposed to explain the processive decatenating activity of *E. coli* topoisomerase III in which the decatenation loop promotes the entry and binding of a duplex DNA segment inside the cavity of the closed form of the enzyme prior to the binding of the single-stranded segment (97). This complex then binds to a single-stranded region of a gapped circle, and after cleavage, the duplex DNA segment is passed out through the break. The decatenation loop would then promote rebinding to double-stranded DNA, leading to yet another cycle of decatenation. Note that the order of binding of the two DNA segments involved in the reaction in this model is the reverse of that proposed above for *E. coli* topoisomerase I. It has been pointed out that although bacterial topoisomerases III possess the basic decatenation loop, yeast topoisomerase III as well as both the α and β isoforms of higher eukaryotic topoisomerases III resemble *E. coli* topoisomerase I and lack this region of the protein (97). It remains to be seen whether the eukaryotic enzymes are like *E. coli* topoisomerase I and lack the ability to efficiently carry out multiple decatenations or whether some other region of the protein or an accessory protein substitutes for the decatenation loop.

Reverse Gyrase

Although reverse gyrase was originally discovered in a hyperthermophilic archaeon (98), the activity has been found in hyperthermophilic and some thermophilic eubacteria as well as in all hyperthermophilic archaea (49, 99), and it

clearly represents an important adaptation to growth at high temperatures. The finding that reverse gyrase makes a single-strand break and attaches to the 5′ end of the broken strand (100–102) led to the realization that it is a type IA rather than a type II enzyme. The similarity to the bacterial type IA enzymes was extended by functional studies (102, 103) and by the observation that the cleavage specificity of the enzyme resembles that of *E. coli* topoisomerase I (104, 105). These findings indicate that reverse gyrase must catalyze ATP-dependent positive supercoiling by a mechanism distinct from the duplex DNA transport mechanism described below for DNA gyrase (106).

The cloning of the first reverse gyrase gene from *Sulfolobus acidocaldarius* provided a critical clue to understanding the mechanism of this class of enzymes (80). As expected, the protein shows a clearcut homology with other bacterial type IA enzymes, but the region of similarity is limited to the C-terminal half of the protein (Figure 1). The N-terminal half of the protein contains an ATP-binding site and eight motifs characteristic of the DEAD family of DNA and RNA helicases (80). The novel presence of separate topoisomerase and helicase domains within the same protein molecule suggests a model for the mechanism of reverse gyrase in which the helicase domain unwinds the helix as it translocates along the DNA, generating positive supercoils in front and negative supercoils behind. In turn, the topoisomerase domain, restricted to relaxing only negative supercoils, relaxes the DNA behind the translocating helicase to leave the DNA with net positive supercoiling. Support for a helicase-like component in reverse gyrase is the finding that the binding of reverse gyrase to a nicked circular DNA at high molar ratios of enzyme to DNA results in the unwinding of the helix even in the absence of ATP (102).

As a possible test of the above model, the topoisomerase and "helicase" domains of *Sulfolobus acidocaldarius* reverse gyrase were separately expressed and individually examined for relaxation and helicase functions (107). Although the topoisomerase fragment was capable of DNA relaxation, no helicase activity could be detected in either the intact protein or in the isolated putative helicase fragment. Moreover, the isolated helicase fragment was incapable of unwinding DNA at high enzyme to DNA ratios. Unfortunately the topoisomerase fragment alone was not tested in the unwinding assay. However, reverse gyrase activity and the unwinding activity could be reconstituted by mixing the two fragments together. Together, these results suggest either that the original hypothesis described above is correct and the putative helicase activity depends critically on some portion of the topoisomerase half of the protein, or that ATP hydrolysis is not directly coupled to strand separation but serves some other function (107).

The heterodimeric reverse gyrase from *Methanopyrus kandleri* displays an unexpected distribution of functions between the two subunits (Table 1) (21). As with other reverse gyrases, there is a helicase-like domain that is contained entirely within the larger B subunit and begins at its N-terminus. However, the B subunit also contains a portion of a type IA topoisomerase region corresponding to domain I and the first half of domain IV found in *E. coli* topoisomerase I (Figure 2*a*). This region is followed by a domain of unknown function that could be involved in

promoting dimerization. The smaller A subunit begins with a region correspond-
ing to domain II of *E. coli* topoisomerase I and extends through to the second
half of domain IV. Curiously, the break point within the topoisomerase half of
the molecule occurs at the point hypothesized to represent one of the two hinges
within *E. coli* topoisomerase I (II/IV hinge) (Figure 2*a*).

TYPE IB DNA TOPOISOMERASES

General Features

Three classes of enzymes make up the type IB subfamily of topoisomerases: the
topoisomerases I found in all eukaryotic cells, the poxvirus topoisomerases (vac-
cinia enzyme), and the prokaryotic topoisomerase V from *Methanopyrus kandleri*
(46, 47). Because so little is presently known about the third enzyme on the list, it
is not considered further here. As is discussed below, these topoisomerases share
structural and functional properties with the tyrosine recombinases that include the
bacteriophage P1 Cre, and *E. coli* XerD recombinases, and certain phage integrases
(108).

The members of the type IB subfamily of topoisomerases I share no sequence
or structural homology with other known topoisomerases and are functionally
distinct from the members of the type IA subfamily. Unlike the type IA enzymes,
the type IB subfamily members can relax both positive and negative supercoils,
and relaxation goes to completion. Consistent with these properties, there is no
requirement that the substrate DNA be at least partially single-stranded. The type
IB topoisomerases form a covalent intermediate in which the active site tyrosine
becomes attached to the 3′ phosphate end of the cleaved strand rather than the 5′
phosphate end found for the type IA enzymes. The enzymes contain no bound
metal ions, and DNA relaxation does not require Mg(II).

Human DNA Topoisomerase I

This section mainly focuses on the properties and structure of human DNA topoi-
somerase I, the best studied of the cellular type IB enzymes. Based on sequence
comparisons, it is likely that most features described here for the human enzyme
apply to the cellular topoisomerases I from other eukaryotic species as well.

Structure   Based on conservation of sequence, sensitivity to limited proteolysis,
hydrodynamic properties, and fragment reconstitution experiments, the 91-kDa
human topoisomerase I protein has been subdivided into four distinct domains
(109–112) (Figure 4). The N-terminal 214 amino acids of the human enzyme are
dispensable for relaxation activity in vitro and constitute a hydrophilic, unstruc-
tured, and highly protease-sensitive region of the protein (109, 110). Contained
with the N-terminal domain are four nuclear localization signals and sites for in-
teraction with other cellular proteins, including such proteins as nucleolin, SV40
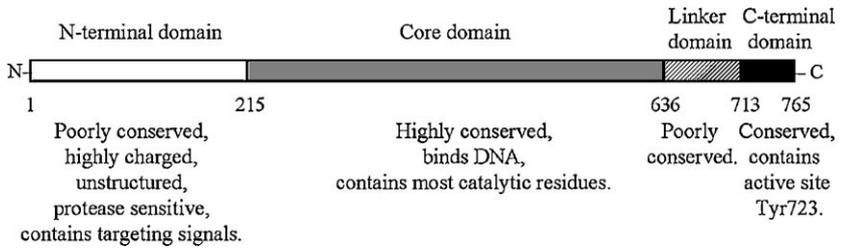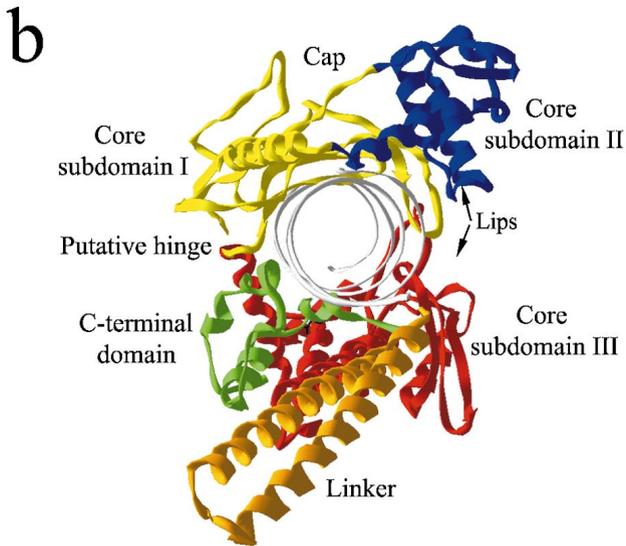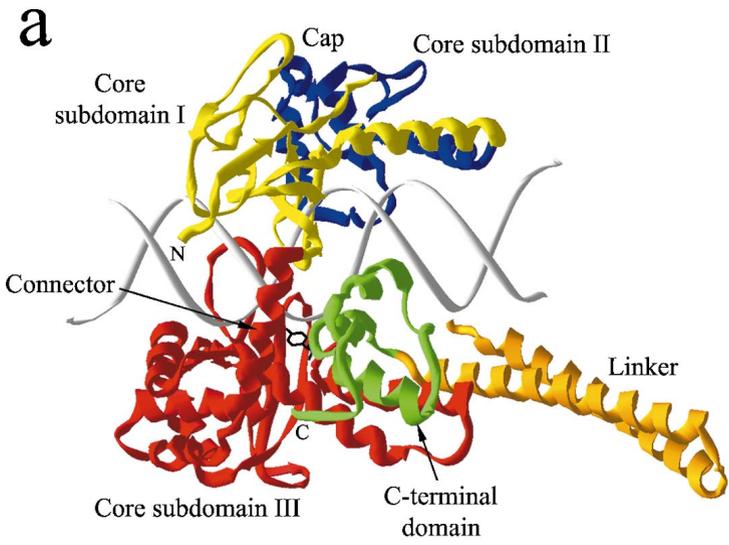
Figure 4 Domain structure of human topoisomerase I. Human topoisomerase I comprises an N-terminal domain (open box), a core domain (gray box), a linker domain (diagonally striped box), and a C-terminal domain (black box). The domain boundaries are based on sequence alignments, limited proteolysis studies, and the crystal structures of the protein. The solution of the crystal structure led to the division of the core domain into three subdomains (see Figure 5). The properties of the four domains are summarized in the lower portion of the figure.

T-antigen, certain transcription factors, p53, and the WRN protein (113–124). The N-terminal domain is followed by a highly conserved, 421 amino acid core domain that contains all of the catalytic residues except the active site tyrosine (125). A protease-sensitive and poorly conserved linker domain comprising 77 amino acids connects the core domain to the 53 amino acid C-terminal domain. An active form of the enzyme can be reconstituted by mixing together fragments corresponding approximately to the core domain (residues 175 to 659) and the C-terminal domain (residues 713 to 765), and thus the linker is dispensable for relaxation activity in vitro. The active site Tyr723 is found within the C-terminal domain.

The crystal structures of several forms of the human enzyme with both non-covalently and covalently bound DNA have been determined (125–128). These co-crystal structures represent the only examples to date of a topoisomerase containing bound DNA. Two views of the structure of human topoisomerase I non-covalently complexed with a 22 base pair DNA are shown in Figure 5. Although the crystals were grown with an N-terminal truncated active form of the protein missing the first 174 amino acids, the X-ray density was only interpretable beginning at residue 215, and thus the entire N-terminal domain is missing from the structure (see Figure 4). Recently the crystal structure of another form of the protein extends the structure back to amino acid residue 203 (not shown) (126). The bi-lobed protein clamps tightly around the DNA with contacts between the protein and the DNA phosphate backbone that extend over 14 base pairs. Most of the contacts are clustered around the five base pairs upstream (−5 to −1) of the cleavage site, which is defined as occurring between DNA residues −1 and +1. The domain structure of the protein, including the subdivision of the core domain into core subdomains I, II, and III, is shown by the color scheme in Figure 5. One of the lobes of the protein comprises core subdomains I and II (yellow and blue) and forms what has been referred to as the "cap" of the protein (125). The front

a

Cap
Core subdomain II
Core subdomain I
Connector
N
Linker
C
Core subdomain III
C-terminal domain

b

Cap
Core subdomain I
Core subdomain II
Putative hinge
Lips
C-terminal domain
Core subdomain III
Linker

end of the cap consists of a pair of long α-helices in a V-shaped configuration that had been seen earlier in the structure of a 26-kDa N-terminal fragment of yeast topoisomerase I (129). The other lobe forms a base that cradles the DNA and consists of core subdomain III (red) and the C-terminal domain (green). This second lobe is connected to the cap by a long α-helix labeled the "connector" in Figure 5a. On the side opposite to this α-helix are a pair of opposing loops called the "lips" that interact via six amino acids and one salt bridge to connect the cap to the base of the protein. Opening and closing of the protein clamp during DNA binding and release must involve the breaking of this interaction between the lips and the lifting of the cap away from the base as shown by the opposing arrows in Figure 5b. The hinge for this movement may be located at the top of the connector α-helix as shown in Figure 5b (putative hinge) (127). The dispensable 77 amino acid linker domain (orange) is a coiled-coil structure that protrudes conspicuously from the base of the protein. Despite the clustering of basic amino acids on the DNA proximal sides of both the linker and the V-shaped α-helices of the cap, neither region of the protein is in direct contact with the DNA in the structure. Although it was necessary to inactivate the enzyme by replacing the active site Tyr723 with phenylalanine to obtain crystals, by modeling a tyrosine in place of the phenylalanine in the structure, it can be seen that the tyrosine side chain is buried between core subdomain III and the C-terminal domain and is close to the scissile phosphate (Figure 5a).

Despite very little sequence homology, most of core subdomain III of human topoisomerase I (residues 440–614) superimposes structurally on the catalytic core region of a family of tyrosine recombinases that includes bacteriophages HP1 and λ integrases, and *E. coli* XerD and bacteriophage P1 Cre recombinases (125, 130–133). In addition, a region near the active site of human topoisomerase I (residues 717–724) appears to correspond structurally to the active site region of the recombinases. In retrospect, this structural similarity is not surprising in view of the many biochemical properties shared by the two classes of enzymes, including a lack of Mg(II) dependence and the formation of a covalent intermediate involving attachment of the enzyme to the 3′ phosphate of the cleaved strand. Most

Figure 5 Two views of the structure of human topoisomerase I. The structure of the enzyme (pdb entry 1A36) is viewed from the side with the DNA axis horizontally oriented (*a*) and looking down the axis of the DNA (*b*). Core subdomains I (residues 215–232 and 320–433), II (residues 233–319) and III (residues 434–633) are colored yellow, blue, and red, respectively. The linker (residues 641–712) and C-terminal domain (residues 713–765) are shown in orange and green, respectively. In panel *a*, the amino terminus (N) and carboxyl terminus (C) of the protein are indicated, and the active site tyrosine is shown in black ball and stick. The long α-helix that connects the cap to the base of the core is labeled "Connector" in panel *a*. In panel *b*, the "Lips" region where the protein opens during DNA binding and unbinding is shown with opposing arrows, and the hinge region at the top of the connector helix is labeled "Putative hinge."

strikingly, the architecture of the active sites for the tyrosine recombinases are very similar to what is found for human topoisomerase I (see below) (108, 125). Thus, the tyrosine recombinases and type IB topoisomerases possess functional cores that use the same chemistry to carry out cleavage and religation. The two classes of enzymes are distinguishable by what occurs during the lifetime of the cleaved intermediate. The topoisomerases permit topological changes in the DNA and restore the original DNA linkage, whereas the recombinases promote strand exchange and join the DNA ends to new partners.

Substrate Specificity   The substrate specificity of eukaryotic topoisomerases I has been characterized at both the nucleotide sequence level and at the level of DNA tertiary structure. By mapping cleavage sites using detergent to trap the covalent complex (134–136), the enzymes were found to nick the DNA with a preference for a combination of nucleotides that extends from positions −4 to −1. The preferred nucleotides in the scissile strand are 5′-(A/T)(G/C)(A/T)T-3′ with the enzyme covalently attached to the −1 T residue. Occasionally a C residue is found at the −1 position. A particularly strong cleavage site for all eukaryotic topoisomerases I was identified by Westergaard and his colleagues in the rDNA repeats found in *Tetrahymena pyriformis* (137), and this sequence, which has a T at the −1 position, formed the basis for designing the 22 base pair oligonucleotide used in the crystallographic studies (125).

The only base-specific contact between the protein and the DNA in the crystal structures involves a hydrogen-bond between Lys532 and the O-2 atom of the strongly preferred thymine base located at the −1 position. This base-specific contact would seem to provide a good explanation for the preference of a thymine at this location. The presence of a carbonyl oxygen at the same position in a cytosine base, which is also tolerated at this position, would be consistent with this conclusion. However, replacing Lys532 with an alanine drastically reduces the activity of the enzyme without detectably changing the cleavage specificity (H Interthal, J Champoux, unpublished). Apparently other protein-DNA interactions in addition to the Lys532 contact with the −1 base play an important role in cleavage site selection. Thus, the structural basis for the weak nucleotide sequence preference of the enzyme remains elusive.

A number of studies have indicated that eukaryotic topoisomerase I prefers supercoiled over relaxed plasmid DNAs as substrates (138, 139), and the use of a mutant form of the protein with phenylalanine instead of tyrosine at the active site (Y723F) showed that the enzyme has a strong preference for binding to supercoiled DNA over relaxed DNA (140). This binding property was localized to the core domain by showing that a fragment including amino acid residues 175 to 659 exhibited the same preference for supercoils. Since the enzyme preferentially binds to supercoils of either sign, it seems likely that the structural feature recognized in the DNA is a node where two duplexes cross (140). Indeed, topoisomerase I has been observed to associate with nodes by electron microscopy (141).

The association of human topoisomerase I with a node would appear to require two DNA binding sites on the protein. One way to accommodate this requirement would be for the enzyme to dimerize upon binding DNA (16). Although there is no evidence for dimerization, this possibility has not been ruled out. An alternative hypothesis is suggested by the crystal structure. The three-dimensional structure of core subdomain II was found to superimpose on the structure of the human Oct-1 POU homeodomain (125). Since Oct-1 belongs to a family of transcription factors that specifically bind DNA (142), this homeodomain-like region could represent a second DNA binding site on topoisomerase I. However the residues in Oct-1 that confer sequence specificity are not conserved in core subdomain II, so if this region of human topoisomerase I does bind DNA, it would seem to be a low-affinity, nonspecific interaction. Further experimentation will be required to test this hypothesis.

Mechanism of Catalysis   Nucleophilic attack by the O-4 oxygen of Tyr723 on the scissile phosphate breaks the DNA strand to generate a phosphodiester link between the tyrosine and the 3′ phosphate, releasing a 5′ hydroxyl. Two different orientations of key active site residues in relation to the scissile phosphate were observed in the crystal structures, suggesting the interesting possibility that two distinct stages along the catalytic pathway have been observed. In noncovalent complexes in which the DNA contains a thymine at the −1 position (−1T structure) (125, 127), Arg488 and Arg590 appear to hydrogen bond with one of the nonbridging oxygens of the scissile phosphate while His632 is hydrogen-bonded to the other nonbridging oxygen. In a structure where the −1 thymine has been replaced with cytosine (−1C structure), the phosphate group has rotated by 75° relative to its position in the −1T structure (126), leaving Arg488 and His632 hydrogen-bonded to the same two nonbridging oxygens, but now Lys532 rather than Arg590 is hydrogen-bonded to a nonbridging oxygen. This −1C configuration is shown in Figure 6 where it can be seen that Lys532 is also hydrogen-bonded to the carbonyl O-2 oxygen of the −1 cytosine base. In either orientation, it appears that Tyr723 is perfectly aligned for nucleophilic attack and that a triad of basic amino acids is positioned to stabilize the pentavalent transition state through interactions with the nonbridging oxygens of the scissile phosphate.

Although it has been hypothesized that the nucleophilic O-4 oxygen of Tyr723 could be activated by the proximity of a general base (89, 108, 143), none of the structures revealed an amino acid close enough to the tyrosine to play this role. However, a water molecule has been found hydrogen-bonded to Arg590 that is only 2.3 Å from the O-4 oxygen (not shown in Figure 6) (126) and therefore is in a position to act as a specific base and accept the proton as catalysis proceeds. There are two reasons for suspecting that the −1C structure represents a more advanced stage in catalysis than the −1T structure. First, the orientation of the amino acid side chains in the structure of the covalent complex that is produced on cleavage (125) is the same as that observed in the −1C structure (126). Second, after the 75° reorientation of the phosphate group on going from the −1T to the −1C
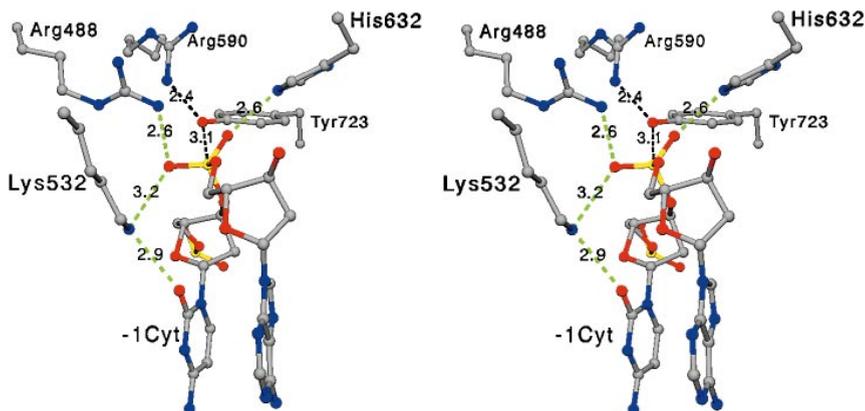
Figure 6 Stereoview of the amino acid side chains in the active site of human topoisomerase I. The three-dimensional relationship between the side chains of the active site residues Tyr723, Arg488, Arg590, His632, and Lys532 and the scissile phosphate is shown for the structure containing a cytosine at the −1 position in the DNA (−1Cyt) (pdb entry 1EJ9). Hydrogen bonds are shown as dashed green lines; other distances as dashed black lines. Bond distances are indicated in angstroms. Figure was generated with Swiss-Pdb Viewer software (v. 3.51) in combination with POV-Ray for Windows (v. 3.1).

structure, Arg590 is no longer within hydrogen-bonding distance of either nonbridging oxygen but is now brought into close proximity to the Tyr723 hydroxyl (2.4 Å) (Figure 6). Hence, analogous to the situation described above for Arg321 of *E. coli* topoisomerase I, Arg590 could facilitate nucleophilic attack by stabilizing the phenolate anion. Furthermore, a direct role for Lys532 in transition-state stabilization is consistent with the large effect on the rates of cleavage observed for the K532A mutation in the human enzyme (H Interthal, J Champoux, unpublished) and the corresponding K167A mutation in vaccinia topoisomerase (144).

It was originally proposed that His632, in addition to stabilizing the transition state through its interaction with a nonbridging oxygen, might also act as a general acid to protonate the leaving 5′ oxygen in the cleavage reaction (127). The findings that His632 can be replaced by glutamine with only a modest reduction in activity and that the H632Q mutation has the same pH profile as the wild-type enzyme rule out the possibility that His632 acts as a general acid (144a). A similar conclusion has been reached concerning the role of the corresponding His265 of vaccinia topoisomerase (145–147). Recent results indicate that Lys167 of vaccinia topoisomerase and the corresponding Lys532 in the human enzyme function instead as general acids to protonate the leaving 5′ oxygen during cleavage (148 and H Interthal, J Champoux, in preparation). The distance of Lys532 from the 5′ oxygen (∼4 Å) in the human enzyme is compatible with this suggestion and confirms that Lys532 may very well play a key role in catalysis by human topoisomerase I.

In principle, DNA religation could be mechanistically just the reverse of the cleavage reaction and the proximity of most of the same amino acids to the scissile phosphate in the crystal structure of the covalent complex is compatible with this possibility. However, the observation that the critical active site residue His632 is within a disordered region in the structure of the covalent complex (125) suggests that there could be some mechanistic differences between the cleavage and religation reactions. Although the close proximity of the attacking 5′ oxygen nucleophile to the scissile phosphate likely contributes significantly to the rate of religation, other nucleophiles such as water, hydrogen peroxide, and certain alcohols can replace the 5′ hydroxyl under some conditions (149, 150). Similar results have been reported for the vaccinia topoisomerase where there is also the suggestion that after mutating the active site tyrosine to histidine or glutamine, water may be able to participate in the cleavage reaction (146, 151, 152). These results indicate that the key to catalysis is the proper three-dimensional organization of the side chains of the active site residues around the scissile phosphate, leaving some flexibility with respect to the nature of the attacking nucleophiles.

Mechanism of DNA Relaxation   A schematic of one possible model for DNA relaxation is shown in Figure 7. The spatial arrangement of the DNA in the co-crystal structure of human topoisomerase I renders unlikely an enzyme-bridging model for DNA relaxation analogous to what has been proposed for *E. coli* topoisomerase I (see above). Instead, it appears that once the DNA has been cleaved, the helical duplex downstream of the cleavage site rotates to relieve any torsional stress within the substrate DNA (127). Attempts to model this rotation within the confines of the hollow interior of the protein indicate that if the protein were to remain clamped around the DNA as shown in Figure 5, the rotating DNA would likely contact and perhaps shift the position of both the cap and the linker. Thus, the term "controlled rotation" was coined to indicate that these structural domains of the protein likely hinder or impede the rotation reaction (127). An alternative hypothesis is diagrammed in Figure 7, where it is proposed that once cleavage has occurred (Figure 7*d*), the enzyme opens up to allow rotation and perhaps assumes a conformation approximating that proposed for the DNA-free form of the enzyme (Figure 7*a*). In this model, there may be little or no hindrance to DNA rotation (Figure 7*e*). These two models differ mainly in the extent to which the DNA interacts with the cap and linker elements of the structure during rotation. Either model could accommodate multiple rotation events before religation of the DNA. Although the number of rotations per cleavage is not known for the human enyzme, the vaccinia topoisomerase appears to allow, on the average, five rotations for each nicking-closing cycle (153).

Vaccinia DNA Topoisomerase

With a molecular size of 36 kDa (314 amino acids), vaccinia is the smallest known topoisomerase. Amino acid sequence similarity to cellular topoisomerases I and
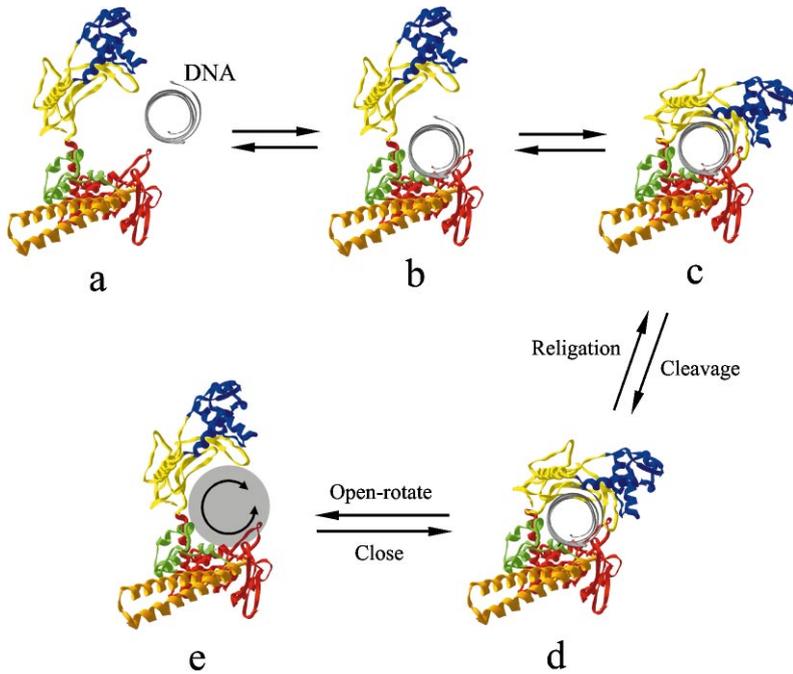
Figure 7 A possible mechanism of DNA relaxation by human topoisomerase I. The enzyme is viewed looking down the axis of the bound DNA. The enzyme is modeled into an open configuration for DNA binding by allowing upward movement at the hinge point identified in Figure 5*b*. Upon binding the DNA (*a* and *b*), the closed form of the enzyme catalyzes cleavage of the scissile strand (*c* and *d*). In this model, rotation of the DNA is proposed to occur in the open conformation of the enzyme (*e*), although the protein may not be as open as depicted here (see text). The cleaved DNA rotates about a series of bonds in the intact strand (*e*); the gray transparent circle depicts the approximate cross section circumscribed by the rotating DNA (127, 128). Multiple rotations are possible before the steps are reversed, and the DNA is either cleaved again or released.

the polarity of attachment to the cleaved strand clearly place the viral enzyme in the type IB subfamily. Interestingly, the vaccinia enzyme is more sequence-specific than the cellular enzyme with a pentanucleotide recognition sequence of 5′-(T/C)CCTT-3′. The recognition site specificity and the observation that the vaccinia enzyme can resolve Holliday junctions (154, 155) indicate an overlap in function with the related tyrosine recombinases and suggest that the viral topoisomerase may have multiple functions during the infection.

Structurally, vaccinia topoisomerase appears to be a pared-down cellular topoisomerase I (128). The crystal structures of two fragments that constitute virtually the entire vaccinia topoisomerase have been determined (156, 157). These two fragments correspond to the amino-terminal domain (amino acids 1–77) and the

catalytic domain (amino acids 81–314) of the enzyme. The structure of the catalytic domain is very similar to core subdomain III and a 19 amino acid region encompassing the active site of human topoisomerase I. This region of similarity is the same as that found between the human topoisomerase I and the tyrosine recombinases. When the vaccinia and human enzymes are superimposed, all of the active site residues align very well except for a significant displacement of the active site tyrosine in the vaccinia structure compared to the human enzyme (128). This displacement, which was also observed for the λ integrase and the XerD recombinase, suggests that DNA binding induces a local shift in the relative positions of the active site tyrosines and the other elements important for catalysis. Three β-strands found in the structure of the 9-kDa amino terminal fragment exhibit a structural similarity with a portion of core subdomain I of the human enzyme, indicating that the amino-terminal domain of vaccinia topoisomerase probably constitutes a cap that interacts with DNA analogous to the role of the cap in the human enzyme.

TYPE II DNA TOPOISOMERASES

General Features

The following properties are shared by all type II topoisomerases (16): (a) The dimeric enzymes bind duplex DNA and cleave the opposing strands with a four-base stagger (topoisomerase VI may generate a two-base stagger) (83). (b) Cleavage involves covalent attachment of each subunit of the dimer to the 5′ end of the DNA through a phosphotyrosine bond. (c) A conformational change pulls the two ends of the cleaved duplex DNA apart to create an opening in what is referred to as the gated or G-segment DNA. A second region of duplex DNA from either the same molecule (relaxation, knotting or unknotting) or a different molecule (catenation or decatenation), referred to as the transported or T-segment, is passed through the open DNA gate. This feature of the reaction explains why the linking number is changed in steps of two when the supercoiling of a circular DNA is changed (106, 158, 159). (d) The reactions require Mg(II), and ATP hydrolysis is required for enzyme turnover and rapid kinetics, although one cycle of relaxation or decatenation/catenation can occur in the presence of the nonhydrolyzable analog of ATP, ADPNP (5′-adenylyl-β,γ-imidodiphosphate) (160–162). (e) The crystal structures of several members, including the structurally distinct topoisomerase VI (see below), reveal that the active site tyrosines are situated in a helix-turn-helix (HTH) motif found within a domain that strongly resembles the DNA binding region of the *E. coli* catabolite activator protein (CAP). In addition, acidic residues within a Rossmann fold on the opposing protomer appear to collaborate with the HTH region of the CAP-like domain to assemble the active site for catalysis and may be involved in metal ion binding in some cases (83, 163, 164). This fold has also been referred to as a "toprim motif" because it is found in DNA primases as well as the topoisomerases and also in a number of other enzymes that catalyze

phosphotransfers or hydrolyze phosphodiester bonds (83, 165). Interestingly, both the CAP-like domain and the Rossmann/toprim fold of the type II enzymes are shared by the type IA, but not the type IB, subfamily of topoisomerases (see above) (163). (f) As with the type I enzymes, a highly conserved arginine residue is implicated in catalysis by its close proximity to the active site tyrosine (83, 166).

Several characteristics distinguish individual members of the type II family. All of the type II enzymes from both prokaryotic domains of life described to date contain two different subunits and are therefore heterotetrameric in structure, whereas the eukaryotic enzymes are homodimers (Table 1). Among all of the known type II enzymes, DNA gyrase stands alone as the only enzyme capable of using the energy from ATP hydrolysis to introduce negative supercoils into the DNA. Finally, different members of the type II family can be distinguished by their relative proficiency at DNA relaxation versus decatenation (or catenation), and this property likely reflects their specialized roles in the cell.

Type IIA DNA Topoisomerases

Structure   Alignment of the sequences of the type IIA subfamily based on their homology with *E. coli* DNA gyrase is shown in Figure 8*a* and reveals a three-domain structure. It can be seen that the B subunit of *E. coli* DNA gyrase (GyrB) corresponds to the ParE subunit of *E. coli* topoisomerase IV and to the N-terminal half of the eukaryotic enzymes, whereas the gyrase A subunit (GyrA) aligns with the ParC subunit and the C-terminal half of the eukaryotic enzymes. Except for an insertion of 170 amino acids near the C-terminus of GyrB, there is a strong correspondence between the sequences of the various type IIA enzymes up to the C-terminal tails, but the C-terminal domains are only similar in sequence for closely related species and then only for the first half of the tail. Similar to the N-terminal domains of type IB topoisomerases, the C-terminal tails of the type IIA enzymes are important for nuclear targeting and for interactions with other proteins (120, 167, 168). The active site tyrosines are located ∼120 amino acids from the N-terminus of GyrA or the B′-A′ boundary in the single subunit enzymes (Figure 8*a*). The ATP binding site is located within the first ∼400 amino acids of GyrB, ParE and the intact enzymes. The ATPase domain is followed by a region located in the C-terminal half of GyrB (referred to as B′) that mediates the interaction between the two different subunits in the heterotetramers and is necessary for communication between the two halves of the single-subunit enzymes (169). The region of the yeast enzyme that is homologous to the first 505 amino acids of GyrA is referred to as A′ and the combined B′-A′ region as the DNA binding/cleavage domain (Figure 8*a*). The 170 amino acid insertion in GyrB is required for activity in vivo, and a mutant protein deleted for this region exhibits a DNA binding defect in vitro (170). The region of the DNA gyrase that wraps ∼140 bp of the bound G-segment DNA into a right-handed supercoil to create the appropriate substrate for negative supercoiling is located in the C-terminal tail domain (1, 171, 172).
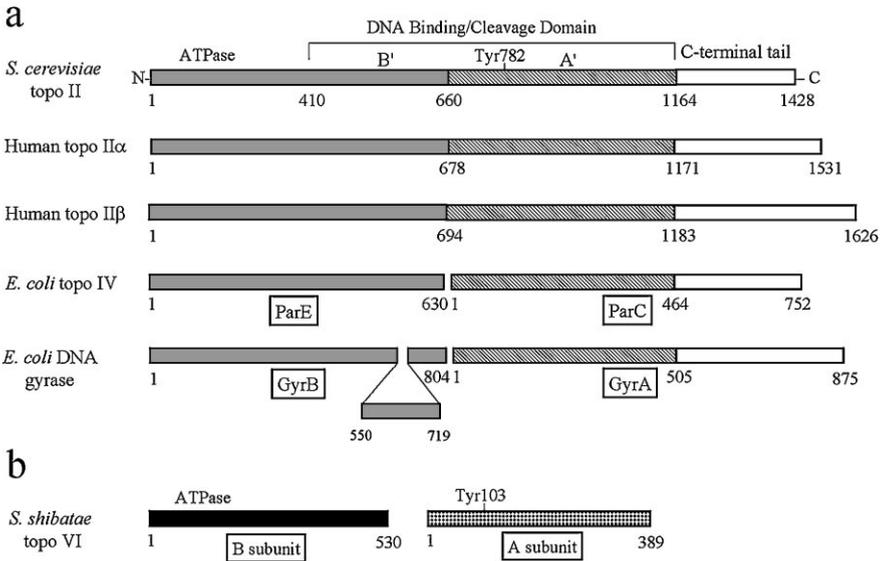
Figure 8 Sequence comparisons among type II topoisomerases. (*a*) The three-domain structure of the type IIA subfamily of topoisomerases is shown based on amino acid sequence homologies with *E. coli* DNA gyrase (18). In each case, the region or subunit that is homologous to the GyrB subunit (excluding the insertion from 550–719) is shown with sequence coordinates and a gray shaded box. The region homologous to the highly conserved first 505 amino acids of GyrA is depicted by a box with diagonal striping. For the *S. cerevisiae* enzyme, the C-terminal half of the GyrB-like region (residues 410–660) is referred to as B′ and the region homologous to GyrA as A′ (residues 661–1164); the combined B′-A′ regions constitute the DNA binding/cleavage domain containing the active site Tyr782. The N-terminal half of GyrB and the corresponding regions in the other type IIA enzymes contain the ATPase domain (and the DNA capture domain, not indicated). The C-terminal tail domains of the enzymes are depicted as open boxes. The ParC and ParE subunits of *E. coli* topoisomerase IV are also referred to as the A and B subunits, respectively, to denote their relationship to the gyrase subunits. (*b*) The subunit structure of *S. sulfolobus* topoisomerase VI, the prototype of the type IIB subfamily, is shown with the B and A subunits in black and stippled boxes, respectively, to indicate that they are structurally distinct from any of the domains in the type IIA subfamily. The positions of the ATPase region and the active site Tyr103 are indicated.

Crystal structures have been solved for fragments of both subunits of *E. coli* DNA gyrase and a large fragment of *S. cerevisiae* topoisomerase II (Figures 9*a* and 10) (166, 173–176). Although a complete structure of an intact functional type IIA topoisomerase is lacking, the available structures provide a near-complete view of how these amazing proteins are assembled and function. The N-terminal half-fragment of the GyrB subunit was crystallized in the presence of the nonhydrolyzable ATP analog ADPNP to promote dimer formation (173, 175), and the structure is shown in Figure 9*a*. The N-terminal ATPase domains in this fragment sit above the C-terminal DNA capture domain that forms a cavity just large enough
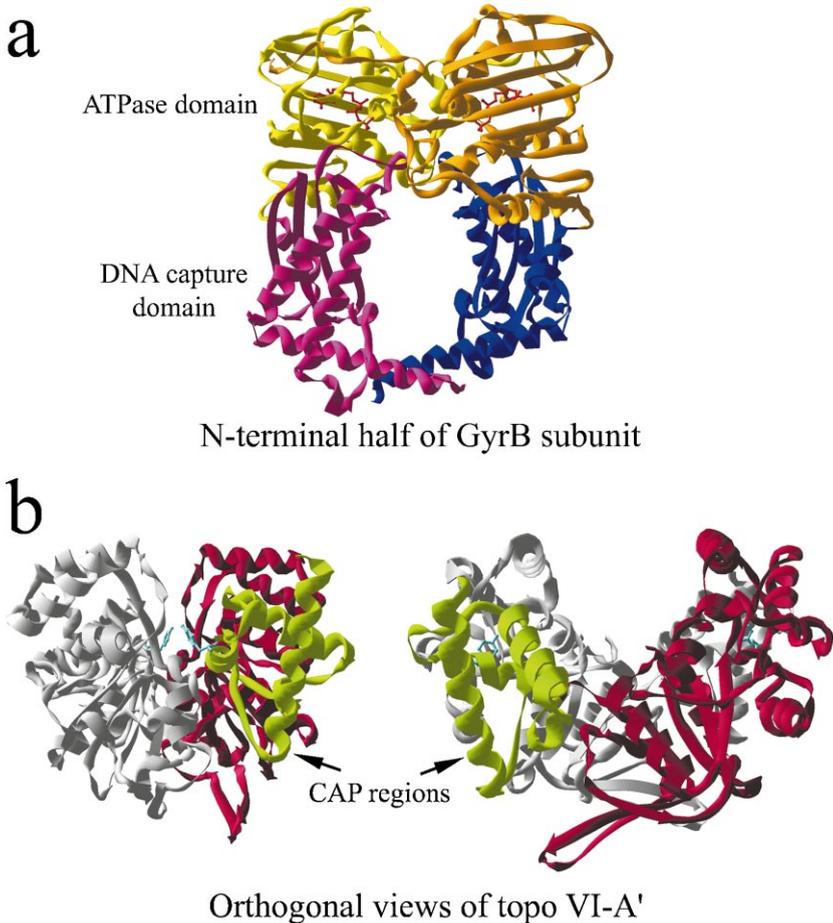
Figure 9   Structures of fragments of DNA gyrase B subunit and DNA topoisomerase VI. (*a*) The dimeric N-terminal half of the *E. coli* DNA gyrase B subunit (pdb entry 1EI1, Y5S mutant) (175) is shown with the ATPase domains (residues 1–220) (monomers colored yellow and orange) oriented above the DNA capture domain (residues 221–392) (monomers colored magenta and blue) to form an opening for the DNA. The bound ADPNP is colored red. (*b*) Orthogonal views of the structure of *Methanococcus jannaschii* DNA topoisomerase VI fragment missing the first 68 amino acids (VI-A′) (pdb entry 1D3Y). The view on the left looks down the dimer interface with the active site tyrosines shown in ball and stick and colored cyan. The view on the right faces one of the monomer subunits and shows the DNA binding cleft. In one of the subunits, the CAP domain (residues 71–140) is colored light green and the other domain containing the Rossmann fold (residues 147–369) is shown in red; the other subunit in both views is shaded gray.
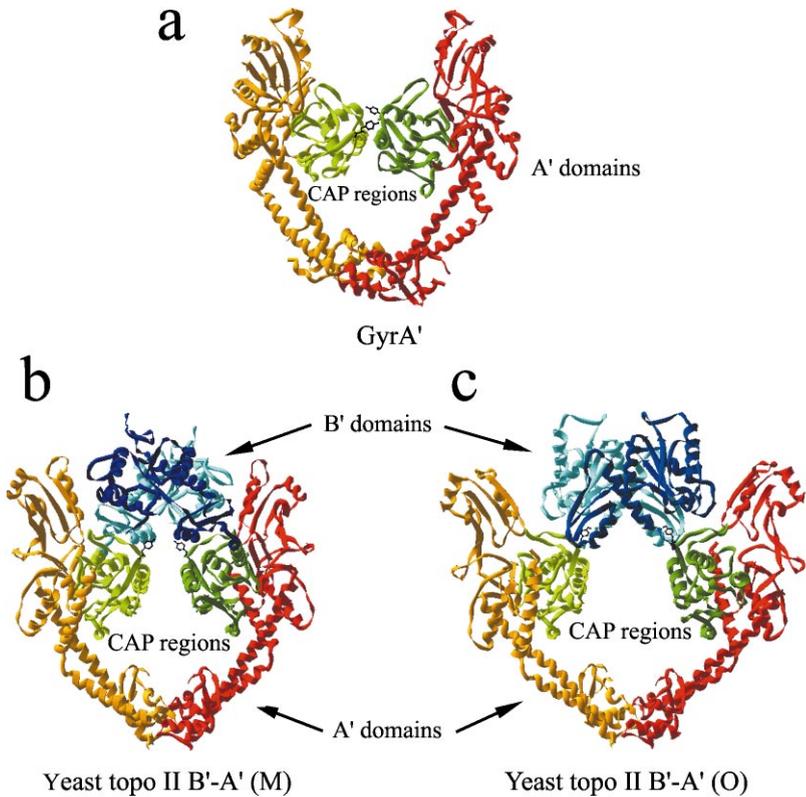
Figure 10   Structures of *E. coli* DNA gyrase subunit A fragment and yeast topoisomerase II regions B′-A′. (*a*) The GyrA′ fragment (residues 30–522) (pdb entry 1AB4) is shown facing the large DNA-binding cavity. The CAP regions of the two monomers of GyrA′ are colored light and dark green with the remainder of the two A′ fragment monomers shown in orange and red. The active site tyrosines are shown in black ball and stick. (*b*) and (*c*) The dimeric structures of fragments of *S. cerevisiae* DNA topoisomerase II are missing the first 419 amino acids. The two structures differ with respect to the orientation of the B′-CAP region relative to the remainder of the molecule. In (*b*) (pdb entry 1BJT), the "gate" is opened midway (M) between the configuration shown in (*a*) for DNA gyrase A′ (designated C in text) and that shown in (*c*) for the open (O) (pdb entry 1BGW) form of the yeast topoisomerase B′-A′ fragment. The color schemes in (*b*) and (*c*) are identical. The B′ domains (residues 420–633) of the two monomers are shown in cyan and dark blue. The CAP regions of the A′ domains of the two monomers (residues 683–819) are shown in light and dark green with the remainder of the A′ domains (residues 820–1178) of the two monomers shown in orange and red. The active site tyrosines are shown in black ball and stick.

(∼20 Å) to accommodate a DNA molecule. The lower extensions of the DNA capture domains are connected to the second half of the GyrB protein that is not present in the crystal structure (yeast topoisomerase B′ region) and are likely held apart rather than in the attached configuration apparent in Figure 9*a*.

The structure of the N-terminal half of *E. coli* GyrA (174) is shown in Figure 10*a*; this fragment is called GyrA′ because it corresponds to the A′ region of yeast topoisomerase II (Figure 8). Although the GyrA′ fragment lacks the C-terminal tail domain required for negative supercoiling, in combination with GyrB, it will carry out ATP-dependent relaxation of supercoils (171). The protein encloses a hole that is ∼30 Å in diameter with two long coiled-coil structures connecting the top portion with the tightly closed bottom. The active site tyrosines are located in CAP-like folds near the N-termini of the GyrA′ fragments (light and dark green in Figure 10*a*). A large cleft that runs across the top of the two CAP regions and beside the upward extensions of the molecule is the putative DNA binding site (see Figure 11). This orientation of the DNA relative to the protein places the active site tyrosines close enough to the DNA that with a modest distortion of the DNA helix they are in a position to make the double-strand staggered cut required for duplex DNA transport (174). Based on the crystal structure of the yeast enzyme and electron micrographs of the homologous yeast and human enzymes (166, 177, 178), it is likely that the entire GyrB subunit sits on top of the GyrA subunit with the B′ domains nestled between the upward extensions of the A subunit (see below).

The structures of two different crystal forms of a large yeast topoisomerase II fragment containing the B′-A′ regions of the protein have been solved. These structures not only provide a glimpse of the structure of the B′ domain missing from the gyrase structures but also reveal alternative conformations that could account for the gate-opening that must precede DNA transport (Figure 10*b* and *c*) (166, 176). The overall architecture of the (B′-A′)$_2$ yeast structures is strikingly similar to that of the GyrA′ dimer, with the presence of a large opening (∼50 Å in this case) bounded by the CAP-like regions containing the active site tyrosines and the connecting antiparallel α-helices. Rossmann-like folds are found in the B′ domains of the protein that rest atop the A′ structures and contain amino acids that contribute to the active site region (166, 179, 180).

Despite these similarities, the individual subdomains clearly assume different orientations in the two yeast structures relative to the GyrA′ structure. Most noticeable is the rotation of the CAP regions (light and dark green in Figure 10*b* and *c*) such that the active site tyrosines are forced apart to different extents in the three structures. In the first of the yeast structures to be solved (166), the lateral distance between the active site tyrosines along what would be the axis of the bound DNA (27 Å), as well as the opening between the CAP regions, is sufficient to easily accommodate a passing duplex DNA (Figure 10*c*), and it was proposed that this structure represents the open DNA gate conformation. In the second structure (Figure 10*b*) (176), the opening is midway between the tightly closed configuration observed for the gyrase (designated C) and that observed in
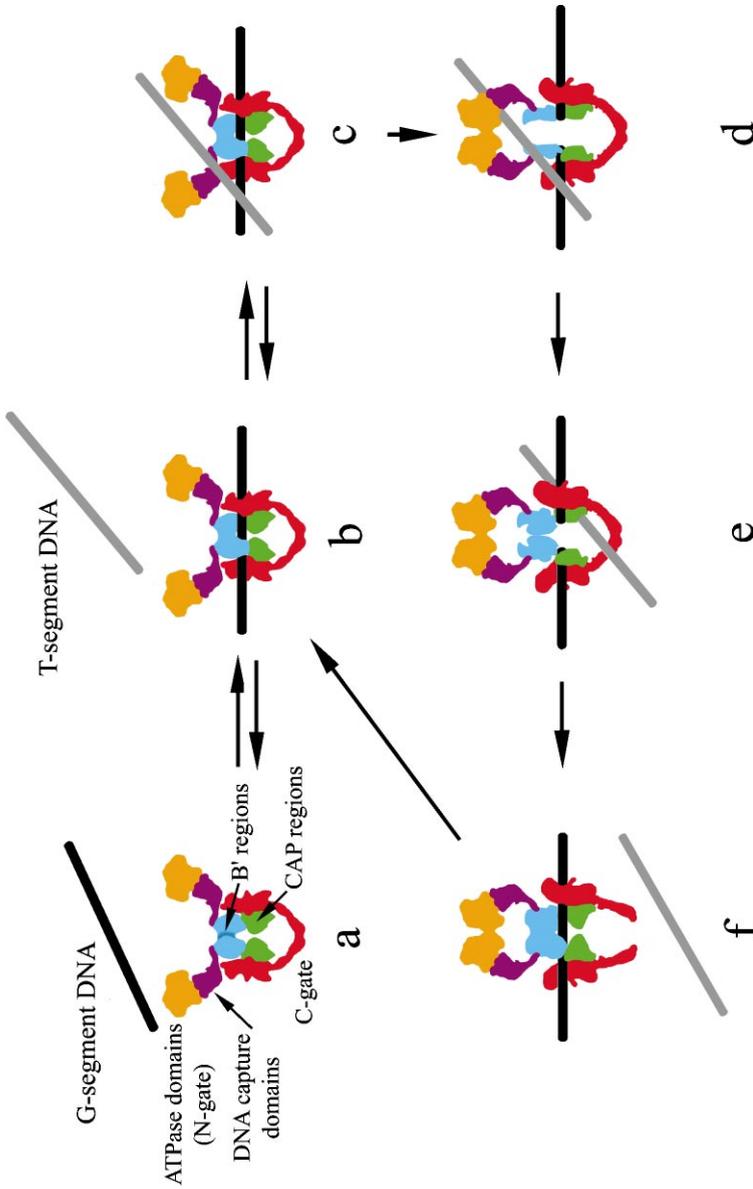
Figure 11 Proposed mechanism for DNA transport by type IIA topoisomerases. A schematic model of the type IIA topoisomerase is shown with the following color scheme: ATPase domains, orange; DNA capture domains, purple; B′ domains corresponding to the C-terminal half of GyrB, cyan; CAP regions, green; and the remainder of the A′ domains, red. The N-gate is shown in the open configuration in panels a–e and is closed in panels d–f. The C-gate is closed in panels a–e and is open in panel f. The G-segment and T-segment DNAs are shown as black and gray rods, respectively (DNA not drawn to scale). See text for an explanation of the individual steps of the DNA transport reaction.

the first yeast structure. The two yeast structures are designated as O and M to reflect the completely open and intermediate size of this channel, respectively (176). The presence of these different conformations implies the existence of considerable flexibility at the interface between the B′ and the A′ domains and the presence of a hinge at the dimer interface between the two coiled-coil regions at the bottom of the structure. It is likely that these three different structures capture different conformational states along the pathway of the enzyme-catalyzed DNA transport reaction.

Mechanism of the DNA Transport Reaction Although many of the details concerning the mechanism of DNA transport by type IIA topoisomerases require further investigation, the biochemical evidence in combination with the available structural information (166, 173–178, 181) provides a useful framework for building molecular models (1, 182). What follows is an overview of the mechanism; for a detailed description of the experiments, see review by Wang (1).

A type IIA topoisomerase structure lacking only the C-terminal tail domain can be approximated by piecing together the N-terminal half of GyrB (Figure 9a) (173, 175) and the different structural conformations described for fragments of yeast topoisomerase II and the A subunit of DNA gyrase (166, 174, 176) (Figure 10) as shown schematically in Figure 11. For the single subunit enzymes, the C-terminal tail is dispensable for cleavage and religation as well as for DNA transport (1), and thus the resulting hypothetical dimeric structure represents a functional topoisomerase. Such an arrangement generates a molecule containing a high-affinity DNA binding site for the DNA to be cleaved (G-segment) and two gated cavities for entrapping the T-segment DNA before, during, and after its transit through the DNA gate (Figure 11d and e). The two-gated structure fits nicely with the biochemical data that provide strong support for such a mechanism of DNA transport (160, 183–185). The cavity that sits above the bound G-segment (as diagrammed) is formed from the ATPase and the DNA capture domains (Figure 9a) with the ATPase domain constituting what is referred to as the N-gate (near to the N-terminus of the intact protein). Below the bound G-segment DNA lies the second cavity that accepts the transported DNA after it passes through the DNA gate (Figure 10). The gate on this lower cavity opens to allow the T-segment DNA to escape (Figure 11f) and is referred to as the exit gate or C-gate (near the C-terminus of the protein). It is clear that the protein must undergo a series of conformational changes during the course of the overall reaction. Such changes are likely coordinated by, as yet poorly understood, interactions with the DNA substrate as well as by the binding of ATP and the release of ATP hydrolysis products (182, 186, 187).

In the absence of bound DNA, but in the presence of ATP, the N-gate apparently cycles between the open and closed states (1, 188). Binding of ATP to one of the ATPase domains is believed to alter the conformation of the domain leading to closing of the N-gate and binding of the second ATP molecule (188, 189). ATP hydrolysis with release of ADP is required to reopen the gate. The first step of a

productive reaction is the binding of a G-segment DNA to the high-affinity binding site located across the tops of the two CAP regions of the dimer (Figure 11*a*). This binding requires an open N-gate and also an opening between the two B′ regions of the protein. Once the DNA is bound, the CAP regions of the protein remain in a closed conformation (C) analogous to the GyrA′ structure (Figure 10*a*), and the two B′ regions are also most likely in close proximity at the dimer interface. Although the active site tyrosines are now near the scissile phosphates, they are not yet positioned for the cleavage reaction (174). After the G-segment DNA is bound, a conformational change leads to an increase in the rate of N-gate cycling between the open and closed states (188, 190). For each cycle without T-segment trapping, it is believed that the G-segment is transiently cleaved, but this cleavage is not followed by opening of the DNA gate (191). However, once ATP binding leads to closure with a T-segment trapped in the DNA capture domain (Figure 11*c* and *d*), a series of highly concerted steps are set in motion to rapidly transfer the T-segment through an open gate in the G-segment. T-segment trapping is likely facilitated by weak binding between the DNA and the interior of the DNA capture domain and is also influenced by the topology of the DNA (1).

The closure of the N-gate with a trapped T-segment leads to cleavage of the G-segment followed by a conformational change in the protein that pulls the two CAP regions apart to open the DNA gate (Figure 11*d*). The T-segment is efficiently transported through the open DNA gate to enter the lower cavity (Figure 11*d* and *e* and also Figure 10*c*). Although all of these steps can occur in the presence of a nonhydrolyzable analog of ATP (ADPNP), ATP significantly speeds up the reaction, and it is likely that a conformational change that is coupled to ATP hydrolysis occurs after cleavage and helps "squeeze" the T-segment through the open gate (182, 186, 187, 192). Once DNA transport has occurred, closure of the DNA gate and subsequent religation of the DNA is believed to force open the C-gate to facilitate release of the T-segment from the enzyme (Figure 11*f*). The cycle is completed by the rapid closure of the C-gate followed presumably by hydrolysis of the second ATP. The release of the products of ATP hydrolysis opens the N-gate and readies the enzyme for potentially another cycle of DNA transport.

Negative Supercoiling by DNA Gyrase   Although all type II topoisomerases require ATP hydrolysis for efficient relaxation of supercoils, DNA gyrase is capable of coupling the energy of ATP hydrolysis to the generation of negative supercoils in a plasmid DNA. This property of the enzyme requires the DNA binding region found in the C-terminal domain of the A subunit, since removal of this region produces an enzyme capable of relaxing DNA but incapable of introducing negative supercoils into the DNA (171).

There are two keys to the negative supercoiling reaction catalyzed by DNA gyrase. The first is that the DNA bound to the C-terminal domain is wrapped around the protein with a right-handed writhe. The resulting constrained positive supercoil is associated with approximately 140 bp of DNA (1). The second key to the reaction relates to the resulting spatial relationship between the enzyme-bound

G-segment DNA and the T-segment DNA trapped in the DNA capture domain. The two crossing segments of DNA are referred to as a DNA node that can be either positive or negative in sign, depending on how the two DNA regions of the same molecule lie across each other (1). For all of the type IIA topoisomerases except DNA gyrase, the sign of the node formed by the bound G- and T-segments is determined by the sense of the supercoiling of the substrate DNA; negatively supercoiled DNAs bind to produce a negative node, and positively supercoiled DNAs bind to produce a positive node. In either case, inversion of the node by the DNA transport reaction relaxes the DNA by canceling out two of the supercoils (106). On the other hand, for DNA gyrase the right-handed wrap of the DNA around the enzyme must dictate that the nodal arrangement of the G- and T-segments leading to a productive transport reaction is nearly always a positive one (1). Thus, one plausible explanation for the directionality of DNA transfer by DNA gyrase is that the C-terminal region delivers the two DNA segments to the enzyme in such a way that only a positive node can be formed prior to cleavage.

Based on an analysis of the topology of the bound DNA at various stages of the reaction, Kampranis et al (192) have proposed a slightly different model for the supercoiling preference of DNA gyrase. They suggest that two regions of a DNA can potentially associate with the DNA gyrase to produce a node of either sign. The bound T-segment DNA promotes opening of the DNA gate to set up an "on-enzyme" equilibrium in which the T-segment distributes itself either before or after the DNA gate depending on the topology of the DNA (192). For a positively supercoiled substrate, which is a good substrate for gyrase, the equilibrium position for the T-segment is beyond the gate so that with closure and release two negative supercoils are introduced that offset two of the positive supercoils. For a negatively supercoiled DNA, the equilibrium favors a position for the T-segment before the DNA gate, and therefore a much reduced chance for any change in the supercoiling of the DNA. The right-handed wrap of the DNA around the enzyme must play a critical role in determining the equilibrium position of the two DNA segments.

Type IIB DNA Topoisomerases

Structure   The prototype of the type IIB subfamily is a type II enzyme identified in the archaeon *Sulfolobus shibatae* and called DNA topoisomerase VI (20); a similar enzyme is found in all archaea sequenced to date. The enzyme is an A_2B_2 heterotetramer with the A and B subunits containing 389 and 530 amino acids, respectively (Figure 8*b*) (19, 193). A search for sequence similarities between the *S. shibatae* topoisomerase VI and the type IIA topoisomerases reveals a weak homology between the N-terminal region of the B subunit, the ATPase domains of the type IIA enzymes, and other ATP binding proteins. Moreover, an isolated B subunit binds an ATP analog (20). The sequence of the A subunit is distinct from any other topoisomerases but displays an intriguing homology to the Spo11

protein from *S. cerevisiae*, which is involved in cleaving DNA to initiate homologous recombination during meiosis (194, 195). Thus, it is likely that, similar to the eubacterial type IIA enzymes (DNA gyrase and topoisomerase IV), the B subunit contains the ATPase domain, and the A subunit is the DNA binding/cleavage subunit (Figure 8*b*).

The crystal structure of a *Methanococcus jannaschii* DNA topoisomerase VI A subunit fragment reported recently (83) confirms these subunit designations and provides a number of key insights into how this novel type II enzyme functions (Figure 9*b*). The fragment consists of amino acids 69 to the end of the protein at residue 389 and is referred to as topoisomerase VI-A′. Each protomer of the dimeric structure can be divided into two clearly defined domains that overall are distinct from the three-dimensional structures of the type IIA enzymes. The N-terminal domain (green, Figure 9*b*) contains the putative active site Tyr103 and is folded into a structure resembling a CAP-like fold, while the remainder of the protein (red, Figure 9*b*) contains a central subdomain that possesses a Rossmann-like toprim fold. The toprim fold contains a tightly coordinated Mg(II) in the structure, possibly implicating the divalent cation in catalysis (83). Thus, despite the lack of any substantial sequence homology between the A′ subunit of topoisomerase VI and any other topoisomerase, topoisomerase VI contains a CAP region that resembles the CAP-like fold found in the A′ fragment of type IIA topoisomerases and in domain III of the type IA enzymes, and a Rossmann-like fold that is found in the B′ domains of the type IIA enzymes and in domain I of the type IA topoisomerases (Figures 2 and 10) (83, 163).

Mechanism The positively charged central groove that runs across the dimer interface shown in Figure 9*b* (*right*) possesses the correct dimensions to function as the binding site for the G-segment DNA. Moreover, after a slight rotation of the CAP regions, the active site tyrosines are juxtaposed with the putative scissile phosphates in the DNA (83). Cleavage followed by a separation of the protomers at the dimer interface would be expected to create a DNA gate for the transport reaction. By comparison with the type IIA enzymes, the most striking feature of the topoisomerase VI A subunit structure is the lack of a cavity below the putative DNA gate to accommodate the T-segment DNA after it has been transported. It will be interesting to determine what biochemical properties correlate with the lack of an exit gate.

The spatial relationship of the ATPase-containing B subunit to the A subunit is unknown, but in the absence of an exit gate, it seems quite likely that the dimeric B domains not only provide a DNA capture domain but also form the bridge that holds the two A′ protomers together when the gate opens (83). The homologous yeast Spo11 protein is involved in producing double-strand breaks that initiate the process of meiotic recombination, but unlike a topoisomerase, the enzyme does not rejoin the two ends and is instead cleaved away from the DNA (194, 195). Thus, the Spo11 protein functionally behaves like a type II topoisomerase that is unable to close the gate after cleavage. One possible explanation for the properties of Spo11

is that Spo11 lacks a functional analog of the topoisomerase VI B subunit bridge and cannot prevent dissociation of the A subunits after cleavage. Interestingly, budding yeast lack a homologue of the topoisomerase VI B subunit. However, the isolated *S. shibatae* topoisomerase VI A subunit is unable to cleave DNA and only does so in the presence of the B subunit (19). If the above hypothesis is correct, then it seems likely that the cleavage activity of Spo11 is regulated by association with one or more accessory proteins, none of which provide a bridge between the two A subunits. Any combination of the nine other proteins required in yeast for producing double-strand breaks during meiosis could function in this regard (196).

CONCLUDING REMARKS

Despite enormous progress in the last few years, many issues concerning the functioning of topoisomerases remain unresolved. Several questions regarding the structural basis of enzyme specificity remain elusive and address issues integral to the in vivo functions of these enzymes. What dictates the preference of most DNA topoisomerases for a particular DNA topology? What determines whether a type IA or a type IIA topoisomerase is more proficient at DNA relaxation or at DNA catenation/decatenation? What structural features affect the extent of negative supercoiling required for a DNA to serve as a good substrate for a type IA enzyme? The relationship between topoisomerases and helicases is an intriguing one and is most dramatically illustrated by the chimeric reverse gyrases. What is the basis for the interaction between bacterial or eukaryotic topoisomerases III, as well as type IB topoisomerases with members of the RecQ family of helicases? What are the cellular targets and functions of these interacting pairs? The availability of crystal structures for enzyme fragments simply whets one's appetite for more structural information, including complete structures with bound DNA for members of both subfamilies. Such information will provide crucial insights into the protein-protein, protein-DNA, and protein-ATP interactions that coordinate conformational changes to DNA transport or rotation. A complete understanding of the mechanisms of catalysis and topoisomerization will continue to require integrating structural information with biochemical experimentation. Ultimately, the goal of creating dynamic models of topoisomerase function may inspire new approaches and continue to challenge the imagination of those working on these remarkable machines.

Visit the Annual Reviews home page at www.AnnualReviews.org

LITERATURE CITED

1. Wang JC. 1998. *Q. Rev. Biophys.* 31:107–44
2. Wu HY, Shyy SH, Wang JC, Liu LF. 1988. *Cell* 53:433–40
3. Kaguni JM, Kornberg A. 1984. *Biol. Chem.* 259:8578–83
4. Masse E, Drolet M. 1999. *J. Biol. Chem.* 274:16659–64
5. Wallis JW, Chrebet G, Brodsky G, Rolfe M, Rothstein R. 1989. *Cell* 58:409–19
6. Wang JC, Caron PR, Kim RA. 1990. *Cell* 62:403–6
7. Rothenberg ML. 1997. *Ann. Oncol.* 8:837–55
8. Pommier Y. 1998. *Biochimie* 80:255–70
9. Wang HK, Morris-Natschke SL, Lee KH. 1997. *Med. Res. Rev.* 17:367–425
10. Maxwell A. 1999. *Biochem. Soc. Trans.* 27:48–53
11. Maxwell A. 1997. *Trends Microbiol.* 5:102–9
12. Berger JM, Wang JC. 1996. *Curr. Opin. Struct. Biol.* 6:84–90
13. Berger JM. 1998. *Curr. Opin. Struct. Biol.* 8:26–32
14. Berger JM. 1998. *Biochim. Biophys. Acta* 1400:3–18
15. Wigley DB. 1996. *Structure* 4:117–20
16. Wang JC. 1996. *Annu. Rev. Biochem.* 65:635–92
17. Sng JH, Heaton VJ, Bell M, Maini P, Austin CA, Fisher LM. 1999. *Biochim. Biophys. Acta* 1444:395–406
18. Caron PR, Wang JC. 1994. *Adv. Pharmacol.* 29B:271–97
19. Buhler C, Gadelle D, Forterre P, Wang JC, Bergerat A. 1998. *Nucleic Acids Res.* 26:5157–62
20. Bergerat A, de Massy B, Gadelle D, Varoutas PC, Nicolas A, Forterre P. 1997. *Nature* 386:414–17
21. Krah R, Kozyavkin SA, Slesarev AI, Gellert M. 1996. *Proc. Natl. Acad. Sci. USA* 93:106–10
22. Liu LF, Liu CC, Alberts BM. 1979. *Nature* 281:456–61
23. Garcia-Beato R, Freije JM, Lopez-Otin C, Blasco R, Viñuela E, Salas ML. 1992. *Virology* 188:938–47
24. Lavrukhin OV, Fortune JM, Wood TG, Burbank DE, Van Etten JL, et al. 2000. *J. Biol. Chem.* 275:6915–21
25. Bidnenko V, Ehrlich SD, Janniere L. 1998. *Mol. Microbiol.* 28:1005–16
26. Nitiss JL. 1998. *Biochim. Biophys. Acta* 1400:63–81
27. Tse-Dinh YC. 1998. *Biochim. Biophys. Acta* 1400:19–27
28. Levine C, Hiasa H, Marians KJ. 1998. *Biochim. Biophys. Acta* 1400:29–43
29. Holmes VF, Cozzarelli NR. 2000. *Proc. Natl. Acad. Sci. USA* 97:1322–24
30. Sawitzke JA, Austin S. 2000. *Proc. Natl. Acad. Sci. USA* 97:1671–76
31. Deleted in proof.
32. Zechiedrich EL, Khodursky AB, Bachellier S, Schneider R, Chen D, et al. 2000. *J. Biol. Chem.* 275:8103–13
33. Champoux JJ, Been MD. 1980. In *Mechanistic Studies of DNA Replication and Genetic Recombination*, ed. B Alberts, pp. 809–15. New York: Academic
34. Peter BJ, Ullsperger C, Hiasa H, Marians KJ, Cozzarelli NR. 1998. *Cell* 94:819–27
35. Hiasa H, Marians KJ. 1996. *J. Biol. Chem.* 271:21529–35
36. Ullsperger C, Cozzarelli NR. 1996. *J. Biol. Chem.* 271:31549–55
37. Zechiedrich EL, Cozzarelli NR. 1995. *Genes Dev.* 9:2859–69
38. Hiasa H, Marians KJ. 1994. *J. Biol. Chem.* 269:32655–59
39. DiGate RJ, Marians KJ. 1988. *J. Biol. Chem.* 263:13366–73

40. Hiasa H, DiGate RJ, Marians KJ. 1994. *J. Biol. Chem.* 269:2093–99

41. Harmon FG, DiGate RJ, Kowalczykowski SC. 1999. *Mol. Cell* 3:611–20

42. Forterre P, Bergerat A, Lopez-Garcia P. 1996. *FEMS Microbiol. Rev.* 18:237–48

43. Guipaud O, Marguet E, Noll KM, de la Tour CB, Forterre P. 1997. *Proc. Natl. Acad. Sci. USA* 94:10606–11

44. de la Tour CB, Kaltoum H, Portemer C, Confalonieri F, Huber R, Duguet M. 1995. *Biochim. Biophys. Acta* 1264:279–83

45. de la Tour CB, Portemer C, Kaltoum H, Duguet M. 1998. *J. Bacteriol.* 180:274–81

46. Slesarev AI, Lake JA, Stetter KO, Gellert M, Kozyavkin SA. 1994. *J. Biol. Chem.* 269:3295–303

47. Slesarev AI, Stetter KO, Lake JA, Gellert M, Krah R, Kozyavkin SA. 1993. *Nature* 364:735–37

48. Wang JC. 1997. *Nature* 386:329–31

49. de la Tour CB, Portemer C, Nadal M, Stetter KO, Forterre P, Duguet M. 1990. *J. Bacteriol.* 172:6803–8

50. Holm C, Goto T, Wang JC, Botstein D. 1985. *Cell* 41:553–63

51. Uemura T, Ohkura H, Adachi Y, Morino K, Shiozaki K, Yanagida M. 1987. *Cell* 50:917–25

52. Gangloff S, de Massy B, Arthur L, Rothstein R, Fabre F. 1999. *EMBO J.* 18:1701–11

53. Goodwin A, Wang SW, Toda T, Norbury C, Hickson ID. 1999. *Nucleic Acids Res.* 27:4050–58

54. Maftahi M, Han CS, Langston LD, Hope JC, Zigouras N, Freyer GA. 1999. *Nucleic Acids Res.* 27:4715–24

55. Gangloff S, McDonald JP, Bendixen C, Arthur L, Rothstein R. 1994. *Mol. Cell. Biol.* 14:8391–98

56. Bennett RJ, Noirot-Gros MF, Wang JC. 2000. *J. Biol. Chem.* 275:26898–905

57. Kim RA, Caron PR, Wang JC. 1995. *Proc. Natl. Acad. Sci. USA* 92:2667–71

58. Lee MP, Brown SD, Chen A, Hsieh TS. 1993. *Proc. Natl. Acad. Sci. USA* 90:6656–60

59. Morham SG, Kluckman KD, Voulomanos N, Smithies O. 1996. *Mol. Cell. Biol.* 16:6804–9

60. Kimura K, Rybenkov VV, Crisona NJ, Hirano T, Cozzarelli NR. 1999. *Cell* 98:239–48

61. Chen M, Beck WT. 1995. *Oncol. Res.* 7:103–11

62. Dereuddre S, Frey S, Delaporte C, Jacquemin-Sablon A. 1995. *Biochim. Biophys. Acta* 1264:178–82

63. Yang X, Li W, Prescott ED, Burden SJ, Wang JC. 2000. *Science* 287:131–34

64. Li W, Wang JC. 1998. *Proc. Natl. Acad. Sci. USA* 95:1010–13

65. Wilson TM, Chen AD, Hsieh T. 2000. *J. Biol. Chem.* 275:1533–40

66. Yu CE, Oshima J, Fu YH, Wijsman EM, Hisama F, et at. 1996. *Science* 272:258–62

67. Ellis NA, Groden J, Ye TZ, Straughen J, Lennon DJ, et al. 1995. *Cell* 83:655–66

68. Shimamoto A, Nishikawa K, Kitao S, Furuichi Y. 2000. *Nucleic Acids Res.* 28:1647–55

69. Kirkegaard K, Wang JC. 1985. *J. Mol. Biol.* 185:625–37

70. Hanai R, Caron PR, Wang JC. 1996. *Proc. Natl. Acad. Sci. USA* 93:3653–57

71. Lima CD, Wang JC, Mondragon A. 1993. *J. Mol. Biol.* 232:1213–16

72. Zhu CX, Tse-Dinh YC. 1994. *Biochem. Mol. Biol. Int.* 33:195–204

73. Ahumada A, Tse-Dinh YC. 1998. *Biochem. Biophys. Res. Commun.* 251:509–14

74. Tse-Dinh YC. 1991. *J. Biol. Chem.* 266:14317–20

75. Zumstein L, Wang JC. 1986. *J. Mol. Biol.* 191:333–40

76. Beran-Steed RK, Tse-Dinh YC. 1989. *Proteins* 6:249–58

77. Zhu CX, Samuel M, Pound A, Ahumada A, Tse-Dinh YC. 1995. *Biochem. Mol. Biol. Int.* 35:375–85

78. Zhang HL, Malpure S, Li Z, Hiasa H, DiGate RJ. 1996. *J. Biol. Chem.* 271:9039–45

79. Grishin NV. 2000. *J. Mol. Biol.* 299:1165–77

80. Confalonieri F, Elie C, Nadal M, de La Tour C, Forterre P, Duguet M. 1993. *Proc. Natl. Acad. Sci. USA* 90:4753–57

81. Lima CD, Wang JC, Mondrag'on A. 1994. *Nature* 367:138–46

82. Grau UM, Trommer WE, Rossmann MG. 1981. *J. Mol. Biol.* 151:289–307

83. Nichols MD, DeAngelis K, Keck JL, Berger JM. 1999. *EMBO J.* 18:6177–88

84. Tse YC, Kirkegaard K, Wang JC. 1980. *J. Biol. Chem.* 255:5560–65

85. Dean FB, Cozzarelli NR. 1985. *J. Biol. Chem.* 260:4984–94

85a. Guex N, Peitsch MC. 1997. *Electrophoresis* 18:2714–23

86. Yu L, Zhu CX, Tse-Dinh YC, Fesik SW. 1995. *Biochemistry* 34:7622–28

87. Brown PO, Cozzarelli NR. 1981. *Proc. Natl. Acad. Sci. USA* 78:843–47

88. Tse Y, Wang JC. 1980. *Cell* 22:269–76

89. Champoux JJ. 1990. In *DNA Topology and Its Biological Effects*, ed. JC Wang, NR Cozzarelli, pp. 217–42. Cold Spring Harbor, NY: Cold Spring Harbor Lab. Press

90. Feinberg H, Changela A, Mondragon A. 1999. *Nat. Struct. Biol.* 6:961–68

91. Feinberg H, Lima CD, Mondragon A. 1999. *Nat. Struct. Biol.* 6:918–22

92. Zhu CX, Roche CJ, Tse-Dinh YC. 1997. *J. Biol. Chem.* 272:16206–10

93. Zhu CX, Tse-Dinh YC. 2000. *J. Biol. Chem.* 275:5318–22

94. Chen SJ, Wang JC. 1998. *J. Biol. Chem.* 273:6050–56

95. DiGate RJ, Marians KJ. 1992. *J. Biol. Chem.* 267:20532–35

96. Mondragon A, DiGate R. 1999. *Struct. Fold. Des.* 7:1373–83

97. Li Z, Mondragon A, Hiasa H, Marians KJ, DiGate RJ. 2000. *Mol. Microbiol.* 35:888–95

98. Kikuchi A, Asai K. 1984. *Nature* 309:677–81

99. de la Tour CB, Portemer C, Huber R, Forterre P, Duguet M. 1991. *J. Bacteriol.* 173:3921–23

100. Forterre P, Mirambeau G, Jaxel C, Nadal M, Duguet M. 1985. *EMBO J.* 4:2123–28

101. Nakasu S, Kikuchi A. 1985. *EMBO J.* 4:2705–10

102. Jaxel C, Nadal M, Mirambeau G, Forterre P, Takahashi M, Duguet M. 1989. *EMBO J.* 8:3135–39

103. Nadal M, Couderc E, Duguet M, Jaxel C. 1994. *J. Biol. Chem.* 269:5255–63

104. Jaxel C, Duguet M, Nadal M. 1999. *Eur. J. Biochem.* 260:103–11

105. Kovalsky OI, Kozyavkin SA, Slesarev AI. 1990. *Nucleic Acids Res.* 18:2801–5

106. Brown PO, Cozzarelli NR. 1979. *Science* 206:1081–83

107. Declais AC, Marsault J, Confalonieri F, de la Tour CB, Duguet M. 2000. *J. Biol. Chem.* 275:19498–504

108. Sherratt DJ, Wigley DB. 1998. *Cell* 93:149–52

109. Stewart L, Ireton GC, Champoux JJ. 1996. *J. Biol. Chem.* 271:7602–8

110. Stewart L, Ireton GC, Parker LH, Madden KR, Champoux JJ. 1996. *J. Biol. Chem.* 271:7593–601

111. Stewart L, Ireton GC, Champoux JJ. 1997. *J. Mol. Biol.* 269:355–72

112. Champoux JJ. 1998. *Prog. Nucleic Acid Res. Mol. Biol.* 60:111–32

113. Bharti AK, Olson MO, Kufe DW, Rubin EH. 1996. *J. Biol. Chem.* 271:P1993–97

114. Albor A, Kaku S, Kulesz-Martin M. 1998. *Cancer Res.* 58:2091–94

115. Gobert C, Skladanowski A, Larsen AK. 1999. *Proc. Natl. Acad. Sci. USA* 96:10355–60

116. Gobert C, Bracco L, Rossi F, Olivier M, Tazi J, et al. 1996. *Biochemistry* 35:5778–86

117. Simmons DT, Melendy T, Usher D, Stillman B. 1996. *Virology* 222:365–74

118. Lebel M, Spillare EA, Harris CC, Leder P. 1999. *J. Biol. Chem.* 274:37795–99

119. Haluska P Jr, Saleem A, Edwards TK, Rubin EH. 1998. *Nucleic Acids Res.* 26:1841–47

120. Shaiu WL, Hu T, Hsieh TS. 1999. *Pac. Symp. Biocomput.*, pp. 578–89

121. Merino A, Madden KR, Lane WS, Champoux JJ, Reinberg D. 1993. *Nature* 365:227–32

122. Kretzschmar M, Meisterernst M, Roeder RG. 1993. *Proc. Natl. Acad. Sci. USA* 90:11508–12

123. Shykind BM, Kim J, Stewart L, Champoux JJ, Sharp PA. 1997. *Genes Dev.* 11:397–407

124. Wang Z, Roeder RG. 1998. *Mol. Cell* 1:749–57

125. Redinbo MR, Stewart L, Kuhn P, Champoux JJ, Hol WG. 1998. *Science* 279:1504–13

126. Redinbo MR, Champoux JJ, Hol WG. 2000. *Biochemistry* 39:6832–40

127. Stewart L, Redinbo MR, Qiu X, Hol WG, Champoux JJ. 1998. *Science* 279:1534–41

128. Redinbo MR, Champoux JJ, Hol WG. 1999. *Curr. Opin. Struct. Biol.* 9:29–36

129. Lue N, Sharma A, Mondragon A, Wang JC. 1995. *Structure* 3:1315–22

130. Subramanya HS, Arciszewska LK, Baker RA, Bird LE, Sherratt DJ, Wigley DB. 1997. *EMBO J.* 16:5178–87

131. Guo F, Gopaul DN, van Duyne GD. 1997. *Nature* 389:40–46

132. Kwon HJ, Tirumalai R, Landy A, Ellenberger T. 1997. *Science* 276:126–31

133. Hickman AB, Waninger S, Scocca JJ, Dyda F. 1997. *Cell* 89:227–37

134. Tanizawa A, Kohn KW, Pommier Y. 1993. *Nucleic Acids Res.* 21:5157–66

135. Bonven BJ, Gocke E, Westergaard O. 1985. *Cell* 41:541–51

136. Been MD, Burgess RR, Champoux JJ. 1984. *Nucleic Acids Res.* 12:3097–114

137. Andersen AH, Gocke E, Bonven BJ, Nielsen OF, Westergaard O. 1985. *Nucleic Acids Res.* 13:1543–57

138. Muller MT. 1985. *Biochim. Biophys. Acta* 824:263–67

139. Camilloni G, Di Martino E, Di Mauro E, Caserta M. 1989. *Proc. Natl. Acad. Sci. USA* 86:3080–84

140. Madden KR, Stewart L, Champoux JJ. 1995. *EMBO J.* 14:5399–409

141. Zechiedrich EL, Osheroff N. 1990. *EMBO J.* 9:4555–62

142. Klemm JD, Rould MA, Aurora R, Herr W, Pabo CO. 1994. *Cell* 77:21–32

143. Stivers JT, Shuman S, Mildvan AS. 1994. *Biochemistry* 33:15449–58

144. Wittschieben J, Shuman S. 1997. *Nucleic Acids Res.* 25:3001–8

144a. Yang Z, Champoux JJ. 2001. *J. Biol. Chem.* 276:677–85

145. Petersen BO, Shuman S. 1997. *J. Biol. Chem.* 272:3891–96

146. Krogh BO, Shuman S. 2000. *Biochemistry* 39:6422–32

147. Stivers JT, Jagadeesh GJ, Nawrot B, Stec WJ, Shuman S. 2000. *Biochemistry* 39:5561–72

148. Krogh BO, Shuman S. 2000. *Mol. Cell* 5:1035–41

149. Lisby M, Krogh BO, Boege F, Westergaard O, Knudsen BR. 1998. *Biochemistry* 37:10815–27

150. Christiansen K, Knudsen BR, Westergaard O. 1994. *J. Biol. Chem.* 269:11367–73

151. Petersen BO, Shuman S. 1997. *Nucleic Acids Res.* 25:2091–97

152. Wittschieben J, Petersen BO, Shuman S. 1998. *Nucleic Acids Res.* 26:490–96

153. Stivers JT, Harris TK, Mildvan AS. 1997. *Biochemistry* 36:5212–22

154. Sekiguchi J, Cheng C, Shuman S. 2000. *Nucleic Acids Res.* 28:2658–63

155. Sekiguchi J, Seeman NC, Shuman S. 1996. *Proc. Natl. Acad. Sci. USA* 93:785–89

156. Sharma A, Hanai R, Mondragon A. 1994. *Structure* 2:767–77

157. Cheng C, Kussie P, Pavletich N, Shuman S. 1998. *Cell* 92:841–50

158. Mizuuchi K, Fisher LM, O'Dea MH, Gellert M. 1980. *Proc. Natl. Acad. Sci. USA* 77:1847–51

159. Liu LF, Liu CC, Alberts BM. 1980. *Cell* 19:697–707

160. Roca J, Wang JC. 1994. *Cell* 77:609–16

161. Sugino A, Higgins NP, Brown PO, Peebles CL, Cozzarelli NR. 1978. *Proc. Natl. Acad. Sci. USA* 75:4838–42

162. Osheroff N, Shelton ER, Brutlag DL. 1983. *J. Biol. Chem.* 258:9536–43

163. Berger JM, Fass D, Wang JC, Harrison SC. 1998. *Proc. Natl. Acad. Sci. USA* 95:7876–81

164. Liu QY, Wang JC. 1999. *Proc. Natl. Acad. Sci. USA* 96:881–86

165. Aravind L, Leipe DD, Koonin EV. 1998. *Nucleic Acids Res.* 26:4205–13

166. Berger JM, Gamblin SJ, Harrison SC, Wang JC. 1996. *Nature* 379:225–32

167. Adachi N, Miyaike M, Kato S, Kanamaru R, Koyama H, Kikuchi A. 1997. *Nucleic Acids Res.* 25:3135–42

168. Mirski SE, Gerlach JH, Cole SP. 1999. *Exp. Cell Res.* 251:329–39

169. Bjergbaek L, Kingma P, Nielsen IS, Wang Y, Westergaard O, et al. 2000. *J. Biol. Chem.* 275:13041–48

170. Chatterji M, Unniraman S, Maxwell A, Nagaraja V. 2000. *J. Biol. Chem.* 275:22888–94

171. Kampranis SC, Maxwell A. 1996. *Proc. Natl. Acad. Sci. USA* 93:14416–21

172. Reece RJ, Maxwell A. 1991. *Nucleic Acids Res.* 19:1399–405

173. Wigley DB, Davies GJ, Dodson EJ, Maxwell A, Dodson G. 1991. *Nature* 351:624–29

174. Cabral JHM, Jackson AP, Smith CV, Shikotra N, Maxwell A, Liddington RC. 1997. *Nature* 388:903–6

175. Brino L, Urzhumtsev A, Mousli M, Bronner C, Mitschler A, et al. 2000. *J. Biol. Chem.* 275:9468–75

176. Fass D, Bogden CE, Berger JM. 1999. *Nat. Struct. Biol.* 6:322–26

177. Schultz P, Olland S, Oudet P, Hancock R. 1996. *Proc. Natl. Acad. Sci. USA* 93:5936–40

178. Benedetti P, Silvestri A, Fiorani P, Wang JC. 1997. *J. Biol. Chem.* 272:12132–37

179. Brown PO, Peebles CL, Cozzarelli NR. 1979. *Proc. Natl. Acad. Sci. USA* 76:6110–19

180. Gellert M, Fisher LM, O'Dea MH. 1979. *Proc. Natl. Acad. Sci. USA* 76:6289–93

181. Kirchhausen T, Wang JC, Harrison SC. 1985. *Cell* 41:933–43

182. Baird CL, Harkins TT, Morris SK, Lindsley JE. 1999. *Proc. Natl. Acad. Sci. USA* 96:13685–90

183. Roca J, Berger JM, Harrison SC, Wang JC. 1996. *Proc. Natl. Acad. Sci. USA* 93:4057–62

184. Williams NL, Maxwell A. 1999. *Biochemistry* 38:13502–11

185. Williams NL, Maxwell A. 1999. *Biochemistry* 38:14157–64

186. Morris SK, Lindsley JE. 1999. *J. Biol. Chem.* 274:30690–96

187. Morris SK, Harkins TT, Tennyson RB, Lindsley JE. 1999. *J. Biol. Chem.* 274:3446–52

188. Lindsley JE, Wang JC. 1993. *J. Biol. Chem.* 268:8096–104

189. Ali JA, Orphanides G, Maxwell A. 1995. *Biochemistry* 34:9801–8

190. Maxwell A, Gellert M. 1984. *J. Biol. Chem.* 259:14472–80

191. Roca J, Wang JC. 1992. *Cell* 71:833–40

192. Kampranis SC, Bates AD, Maxwell A. 1999. *Proc. Natl. Acad. Sci. USA* 96:8414–19

193. Bergerat A, Gadelle D, Forterre P. 1994. *J. Biol. Chem.* 269:27663–69

194. Keeney S, Kleckner N. 1995. *Proc. Natl. Acad. Sci. USA* 92:11274–78

195. Keeney S, Giroux CN, Kleckner N. 1997. *Cell* 88:375–84

196. Haber JE. 1997. *Cell* 89:163–66