The Structural Basis for Substrate Specificity in DNA Topoisomerase IV

Kevin D. Corbett, Allyn J. Schoeffler, Nathan D. Thomsen and James M. Berger*

Department of Molecular and Cell Biology
237 Hildebrand Hall #3206
University of California
Berkeley, Berkeley, CA 94720-3206, USA

Most bacteria possess two type IIA topoisomerases, DNA gyrase and topo IV, that together help manage chromosome integrity and topology. Gyrase primarily introduces negative supercoils into DNA, an activity mediated by the C-terminal domain of its DNA binding subunit (GyrA). Although closely related to gyrase, topo IV preferentially decatenates DNA and relaxes positive supercoils. Here we report the structure of the full-length Escherichia coli ParC dimer at 3.0 Å resolution. The N-terminal DNA binding region of ParC is highly similar to that of GyrA, but the ParC dimer adopts a markedly different conformation. The C-terminal domain (CTD) of ParC is revealed to be a degenerate form of the homologous GyrA CTD, and is anchored to the top of the N-terminal domains in a configuration different from that thought to occur in gyrase. Biochemical assays show that the ParC CTD controls the substrate specificity of topo IV, likely by capturing DNA segments of certain crossover geometries. This work delineates strong mechanistic parallels between topo IV and gyrase, while explaining how structural differences between the two enzyme families have led to distinct activity profiles. These findings in turn explain how the structures and functions of bacterial type IIA topoisomerases have evolved to meet specific needs of different bacterial families for the control of chromosome superstructure.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: β-pinwheel; ATPase; supercoiling; decatenation; DNA binding

Introduction

All organisms invest significant resources toward preserving the informational and structural integrity of their genomes. Many cellular transactions involving DNA, including replication, transcription, and recombination, alter chromosome topology through supercoiling, knotting, and catenation. These structures, when left unresolved, can stall replication and transcription, generate double-strand DNA breaks, or impair the partitioning of replicated DNA to daughter cells. Type II topoisomerases circumvent many types of topological problems, transporting one double-helical DNA segment through a transient, enzyme-mediated break in another to modulate DNA superhelicity and unlink tangled chromosomes. Due to the importance of these activities, type II topoisomerases are essential for cell viability and are found throughout all cellular domains of life. Type II topoisomerases (topos) can be placed into two subfamilies based on sequence and structural features. Type IIA topos, the most common class, are found throughout eukaryotes, bacteria, and some archaea. The simpler type IIB topos are restricted to archaea and higher plants. Bacterial type IIA topos are A₂B₂ heterotetramers with three subunit interfaces that alternately open and close in response to ATP binding and hydrolysis to effect DNA transport. Their reaction cycle begins when one DNA duplex, termed the "gate" or G-segment, is bound by the A-subunits of the enzyme. Next, a second duplex, termed the "transfer" or T-segment, is captured when the B-subunits bind ATP and dimerize. This event triggers cleavage of the G-segment, and the subsequent hydrolysis of ATP mediates passage of the T-segment through the gap opened in the G-segment. Following transport, the G-segment is resealed and the T-segment is released from the enzyme.

Abbreviations used: CTD, C-terminal domain; NTD, N-terminal domain; topo, topoisomerase; SAD/MAD, single and multi-wavelength anomalous diffraction; CAP, catabolite activator protein.

E-mail address of the corresponding author: jmberger@berkeley.edu

doi:10.1016/j.jmb.2005.06.029


Available online at www.sciencedirect.com
Although this general reaction is common to all type IIA topoisomerases, bacteria have unique requirements that have necessitated further adaptation and specialization of their enzymes. For example, mesophilic bacteria tightly regulate the superhelical density of their genomes to maintain a specific level of negative supercoiling. Negative supercoiling creates strain that favors unwinding of the DNA duplex and is important for many cellular processes that require access to single-stranded DNA, including transcription and the initiation of DNA replication. In addition, closed circular bacterial genomes become catenated, or topologically linked, during genome replication, necessitating their separation before cell division. To meet their unique needs, most bacteria possess two specialized type IIA topois (Figure 1(a)): DNA gyrase, encoded by the gyrA and gyrB genes; and topo IV, encoded by parC and parE. Gyrase catalyzes the formation of negative supercoils and also can relax positive supercoils (a topologically equivalent reaction), but is inefficient at unknotting and decatenation. Gyrase’s negative supercoiling capability is mediated by a domain at the C terminus of its A-subunit, termed the GyrA CTD, that is not shared with eukaryotic type IIA topois. The GyrA CTD is a compact circular domain that can bend short DNA segments up to 180° and can constrain positive writhe in larger DNAs. In the context of the gyrase holoenzyme, the GyrA CTD wraps the DNA flanking a bound G-segment into a local positive-handed crossover, and supplies this DNA in cis to the enzyme as a T-segment. Strand passage inverts this positive crossover into a negative crossover, thereby introducing negative supercoils.

The functional specialization of gyrase prevents the enzyme from effectively catalyzing reactions that involve the capture of a T-segment in trans. As a consequence, decatenation and unknotting reactions are left to topo IV in most bacteria. Consistent with these primary tasks, topo IV localizes to sites of DNA replication and also interacts directly with the FtsK/Xer chromosome segregation machinery. Topo IV efficiently relaxes positive supercoils, such as those formed in front of a DNA replication fork, but is markedly less active on negatively supercoiled DNAs. This selectivity keeps topo IV from relaxing the negative supercoils introduced by gyrase. While the substrate specificity of topo IV is clearly advantageous for bacteria, its selectivity poses a dilemma: how can the enzyme detect and respond to the global topology of a substrate DNA when its small size limits it to sampling only local DNA crossover geometries?

Recent studies have provided valuable insights into how topo IV discriminates between different DNA topologies. Single-molecule experiments using braided DNAs have shown that the enzyme recognizes the local crossing geometry of two DNA duplexes and specifically acts on “left-handed” crossovers, juxtapositions that are found some 25-fold more frequently in positively supercoiled DNA (Figure 1(b)). This ratio agrees well with the ~20-fold preference for positive supercoils measured using bulk biochemical methods.
addition, several experiments have shown that the geometry preference is likely enforced after G-segment binding, suggesting that T-segment recognition is the step in which specificity is imposed.36–38

Recently, topo IV has been found to possess a domain in its ParC subunit that is structurally similar to the GyrA CTD and shares the ability to bind and bend DNA in vitro.27,39,40 We have taken a structural and biochemical approach to understand the contribution of this domain (the ParC CTD) to the unique activities of topo IV. Our structural analysis of Escherichia coli ParC shows that the CTD is a degenerate form of the GyrA CTD, and that it is positioned relative to the central DNA-binding site so as to interact favorably with incoming T-segments of a certain geometry. Relaxation/decatenation and DNA binding assays using both full-length and CTD-truncated enzymes suggest that the ParC CTD does not influence G-segment binding significantly, but instead acts as a T-segment recruiting element to control substrate specificity. Synthesis of our results with previous findings provides a physical mechanism for the preferential action of topo IV on positively supercoiled and catenated DNAs, and reveals how this enzyme has evolved to act on specific chromosome topologies to support the critical processes of genome replication and segregation in bacteria.

Results

Structure of ParC

To better understand the physical basis for topoisomerase IV function, we undertook a series of structural and mechanistic studies of the ParC subunit and the ParC2E2 holoenzyme. We first crystallized and solved the structure of the isolated E. coli ParC CTD (residues 497–752) to a resolution of 1.7 Å using a combination of single and multiwavelength anomalous diffraction (SAD/MAD) techniques. This structure was refined to an R-factor of 18.4% and a free R-factor of 21.6% with good stereochemistry (Table 1). We next designed a construct of E. coli ParC (ParC27) optimized for crystallization consisting of amino acid residues 27–742, which represents the full-length subunit minus small protease-sensitive regions (10–26 residues) at its N and C termini (H. Hiasa, personal communication). We crystallized and solved the structure of ParC27 to 3.0 Å resolution by molecular replacement, using the previously solved structure of ParC CTD.

Table 1. Data collection, refinement and stereochemistry

<table>
<thead>
<tr>
<th>Data collection</th>
<th>ParC CTD native</th>
<th>ParC CTD SeMet SAD</th>
<th>ParC CTD SeMet MAD peak</th>
<th>ParC CTD SeMet MAD remote</th>
<th>ParC27 native</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution (Å)</td>
<td>30–1.70</td>
<td>30–2.0</td>
<td>50–2.0</td>
<td>50–2.0</td>
<td>20–3.0</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.1271</td>
<td>0.9796</td>
<td>0.9796</td>
<td>1.0199</td>
<td>1.1157</td>
</tr>
<tr>
<td>Space group</td>
<td>P1</td>
<td>P1</td>
<td>P1</td>
<td>P1</td>
<td>P2,2,2</td>
</tr>
<tr>
<td>Unit cell dimensions (a, b, c (Å))</td>
<td>40.91, 50.49, 72.76</td>
<td>41.04, 50.48, 72.87</td>
<td>40.99, 50.53, 72.85</td>
<td>40.99, 50.53, 72.85</td>
<td>257.99, 62.14, 64.00</td>
</tr>
<tr>
<td>Unit cell angles (a, β, γ (deg.))</td>
<td>86.11, 86.91, 70.57</td>
<td>86.37, 86.84, 70.78</td>
<td>86.35, 86.89, 70.71</td>
<td>86.35, 86.89, 70.71</td>
<td>90, 90, 90</td>
</tr>
<tr>
<td>I/α (last shell)</td>
<td>22.0 (4.4)</td>
<td>15.3 (3.5)</td>
<td>12.8 (4.7)</td>
<td>14.4 (5.7)</td>
<td>9.5 (2.5)</td>
</tr>
<tr>
<td>Ksym (last shell) (%)</td>
<td>0.066 (0.257)</td>
<td>0.053 (0.220)</td>
<td>0.043 (0.144)</td>
<td>0.037 (0.120)</td>
<td>0.134 (0.387)</td>
</tr>
<tr>
<td>Completeness (last shell) (%)</td>
<td>96.4 (94.1)</td>
<td>96.7 (91.9)</td>
<td>98.0 (96.9)</td>
<td>97.8 (96.8)</td>
<td>89.7 (89.1)</td>
</tr>
<tr>
<td>No. of reflections</td>
<td>470,798</td>
<td>273,319</td>
<td>181,740</td>
<td>180,901</td>
<td>217,286</td>
</tr>
<tr>
<td>Unique</td>
<td>57,955</td>
<td>72,264</td>
<td>36,477</td>
<td>36,416</td>
<td>19,320</td>
</tr>
<tr>
<td>No. of sites</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>–</td>
</tr>
<tr>
<td>Refinement</td>
<td>ParC CTD</td>
<td>ParC27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>20–1.7</td>
<td>20–3.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of reflections</td>
<td>54,999</td>
<td>18,167</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working</td>
<td>52,062</td>
<td>17,175</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Free (% total)</td>
<td>2937 (5%)</td>
<td>992 (5%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rwork (last shell) (%)</td>
<td>18.35 (20.6)</td>
<td>24.04 (35.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rfree (last shell) (%)</td>
<td>21.55 (23.7)</td>
<td>29.58 (37.2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structure and stereochemistry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of atoms</td>
<td>4091</td>
<td>5402</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>3711</td>
<td>5367</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>380</td>
<td>35</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r.m.s.d. bond lengths (Å)</td>
<td>0.011</td>
<td>0.007</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r.m.s.d. bond angles (deg.)</td>
<td>1.339</td>
<td>1.087</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a Rsym = \sum \| I_i \| - \langle I \rangle \sum \langle I \rangle$, where $I_i$ is the intensity measurement for reflection $i$ and $\langle I \rangle$ is the mean intensity for multiply recorded reflections.

*b Rwork,free = \| F_{obs} - | F_{calc} | \| / F_{obs}$, where the working and free R-factors are calculated using the working and free reflection sets, respectively. The free reflections were held aside throughout refinement.
of the *E. coli* GyrA core DNA binding/cleavage domains\(^4\) and our structure of the ParC CTD. This structure was refined to a final R-factor of 24.0% and a free R-factor of 29.6%, and exhibits good stereochemistry (Table 1).

The ParC\(^{27}\) fragment consists of two major regions. The N-terminal region (NTD; residues 28–480) is composed of three domains and adopts a tertiary and quaternary structure highly similar to the equivalent region of GyrA\(^4\) (Figure 2(a)). The first domain, consisting of residues 28–158, contains a helix-turn-helix motif similar to that of the catabolite activator protein (CAP) and contains the active site residues essential for DNA cleavage, Arg119 and Tyr120. The second domain (residues 159–340), termed the “tower,” adopts an extended bi-lobed \(\alpha/\beta\) structure that packs against the CAP domain, providing structural support and contributing to the primary DNA binding site. The third domain is a compact \(\alpha\)-helical bundle connected by long \(\alpha\)-helices to the tower domain and the C-terminal domain. This domain makes up the primary dimer interface of all type IIA topoisomerase structures solved to date.\(^4\)–\(^6\) In our structure of ParC\(^{27}\), this dimer interface is recapitulated through a crystallographic 2-fold axis (Figure 2(a)).

Three structures of the type IIA topo DNA binding/cleavage domains, one from *E. coli* GyrA and two from *Saccharomyces cerevisiae* topo II, have revealed that these proteins can adopt a range of conformational states. In GyrA, the CAP domains of a dimer contact each other, placing the two active-site tyrosine residues in position to cleave the two

![Figure 2](image-url)
strands of a DNA duplex. By contrast, the CAP domains are seen to move apart from one another in both topo II structures, opening a gap between these elements that separates the two active sites by up to 30 Å. Remarkably, ParC adopts a more open conformation than seen in any of these structures, separating its active-site tyrosine residues by over 40 Å (Figures 2(a) and 3(a)). This state likely mimics a conformation of the enzyme accessed during strand passage, when the two ends of a cleaved G-segment must be separated to allow the passage of a Y-segment. The conformational change from closed to open occurs solely through flexion of the two α-helices that connect the CAP and tower domains to the dimerization domain (Figure 3(b)).

The structure of the ParC DNA binding/cleavage domains also differs from that of GyrA in two loops situated near the active site. A region containing residues 102–124 of ParC, including the active site residues Arg119 and Tyr120, is significantly rearranged in ParC when compared to GyrA (Figure 3(c)). The region also has high B-factors, indicating structural flexibility. This difference could reflect functional differences between DNA gyrase and topo IV, but more likely reflects the fact that the CAP domains are separated in our structure of ParC, whereas in GyrA they are dimerized and thus may become more structurally rigid. Residues 55–64 of ParC, which are equivalent to residues 58–67 of GyrA, are also disordered in our structure, probably because the dimer is opened and the adjacent 102–124 loop is rearranged. It should be noted that in our crystals of ParC, symmetry-related molecules pack against the rearranged 102–124 loop, and probably further contribute to the observed local conformational rearrangements.

The second major region of ParC, the CTD, comprises residues 500–742. This element is connected by a well-ordered linker (residues 481–499) to the N-terminal region (Figure 2(b)). As anticipated, the CTD adopts a β-pinwheel fold, which was first identified in the B. burgdorferi GyrA CTD and later observed in the B. stearothermophilus ParC CTD and the E. coli GyrA CTD (Figure 4). The architecture of the β-pinwheel fold is reminiscent of a β-propeller, but the Greek key-like topology of its four-stranded β-sheet “blades” is distinct from the antiparallel hairpin sheets found in β-propellers. One hallmark of the β-pinwheel fold is that the outer-most strand of each blade associates with the inner strands from the previous blade, creating an interlocking structure that holds the blades together (Figure 4(a) and (d)).

Two structures of the GyrA CTD, from B. burgdorferi and E. coli, have shown that this domain is made up of six blades that pack into a circular structure. Interestingly, one difference between these structures is that the “GyrA box,” the motif that holds the B. burgdorferi GyrA CTD in a closed ring by locking blade 1 onto blade 6, is disordered in the E. coli domain, allowing this CTD to adopt a spiral shape. A distinguishing feature of all ParC CTDs is that they universally lack the GyrA box motif (Figure 4(b) and (d)). As a consequence, the B. stearothermophilus ParC CTD was also observed to adopt an open, spiral conformation highly similar to that of the E. coli GyrA CTD. Upon solving the structure of the E. coli ParC CTD, we observe that it too adopts an open C-shaped structure, although the pitch of the spiral in this domain is markedly reduced compared to the other spiral-shaped CTD structures, and it also lacks one entire blade (Figure 4; Supplementary Data, Figure S1). In the GyrA CTD, the spiral shape of the E. coli domain is thought to impart a directional bias to the domain’s DNA wrapping ability to enhance supercoiling activity. It is as yet unknown how this phenomenon might affect the properties of topoisomerase IV. Despite these structural differences, all GyrA and ParC CTDs share a key feature: a positively charged outer rim thought to comprise a DNA binding/surfacing region (Figure 4(c)).

During the catalytic cycle of gyrase, the GyrA CTD is thought to cycle between two different locations: an “upper” position where the CTD lies near the tower domain and interacts with the DNA flanking a G-segment, and a “lower” position near the connector α-helices that could represent either a pre-DNA-binding or post-strand passage state. In our structure of ParC, the CTD resides at the very top of the tower domain in a location analogous to the upper position observed for the GyrA CTD (Figure 2(a)). The linker attaching the N and C-terminal domains in ParC is well-ordered in refined 2Fo–Fc and simulated annealing omit electron density maps (Figure 2(b)), and it wraps around the tower domain for most of its length. Four residues between the NT and the CTD (residues 496–499) are ordered in the structure, but are entirely solvent-exposed and do not pack against either domain. Two hydrophobic residues (Met489 and Met494) just upstream of this accessible segment are conserved in ParC orthologs and dock into a hydrophobic patch on the NT and CTD. This type of hydrophobic “ball joint,” especially involving flexible methionine residues, has been proposed to allow a degree of plasticity in protein–protein interfaces while maintaining a tight association between the interacting partners. Together, these features likely permit the CTD to rotate with respect to the N-terminal region, but constrain it in a location analogous to the upper position of the GyrA CTD.

**Supercoil relaxation and decatenation by topo IV + ParC CTD**

To determine the role of the ParC CTD in supercoil relaxation and decatenation, we reconstituted topo IV ParC-C2E2 heterotetramers using either wild-type or CTD-truncated (residues 2–482) ParC. We confirmed holoenzyme formation using analytical gel filtration (see Materials and Methods), and tested the activity of the reconstituted enzymes on a number of different substrates. Comparison of
Figure 3 (legend on next page)
the activities of wild-type and CTD-truncated topo IV on negatively supercoiled DNA revealed a \( \sim 10 \)-fold reduction in relaxation activity upon the removal of the ParC CTD (Figure 5(a)).

Since topo IV has been shown to relax positive supercoils \( \sim 20 \) times more efficiently than negative supercoils,\(^{35} \) we next tested the effect of the ParC CTD truncation on the relaxation of positively supercoiled DNA (Figure 5(b)). We observe that the wild-type enzyme is indeed much more active (\( \sim 15 \)-fold) on the positively supercoiled substrate when compared to its activity on negative supercoils. Interestingly, this robust relaxation activity drops about 100-fold upon ablation of the ParC CTD.

**Figure 3.** Comparison of GyrA and ParC NTD conformations. (a) The N-terminal regions of GyrA\(^{41} \) (gray) and ParC (blue) are presented side-by-side. Shown in red are a pair of \( \alpha \)-helices that connect the dimerization domain to the CAP and tower domains. Top-down views show that a twisting motion accompanies the separation of the CAP domains (arrows). (b) Stereo view of an overlay of one dimerization domain from GyrA (gray) and ParC (red), showing the flexion of the connector helices that leads to the global conformational differences observed between the two structures. Not shown in this panel is an insertion in the dimerization domain (residues 413–451) specific to GyrA. (c) Top-down view of the active site of ParC (purple) overlaid with that of GyrA (gray and cyan, dimer mate in gray surface). The reconfigured loop in the CAP domain (residues 102–124 of ParC) is shown with active-site arginine and tyrosine residues (labeled) in stick representation.

**Figure 4.** Structure of the ParC CTD. (a) Top view of the \( E. \ coli \) ParC C-terminal domain. Blades 1–5 are labeled and colored purple, blue, green, yellow, and orange, respectively. (b) Overlay of blade 1 of the \( E. \ coli \) ParC CTD (purple) with that of the \( B. \ burgdorferi \) GyrA CTD\(^{27} \) (gray), showing that the GyrA box motif is deleted in ParC. (c) Side and top views of the electrostatic surface of the ParC CTD, showing the curved, positively charged outer surface that likely comprises its DNA binding surface.\(^{27} \) (d) Structural relationships between the different bacterial type IIA topo CTDs. CTDs with an intact GyrA box motif can adopt open or closed conformations, potentially depending on the sequence of the GyrA box (Supplementary Data, Figure S2). Truncation of the GyrA box (purple) from the six-bladed GyrA CTD\(^{27,28} \) (left) gives rise to the open six-bladed CTD found in some ParC orthologs\(^{39} \) (middle). The further loss of one full blade (shown here as blade six, in red) gives rise to the open five-bladed ParC CTD of \( E. \ coli \) and other Proteobacteria (right) (also see Figure 8).
Comparison of the activity of the CTD-truncated enzyme on positively and negatively supercoiled DNAs reveals that without the ParC CTD, the enzyme relaxes supercoils of either polarity with similar efficiencies (Figure 5(d)). Topo IV’s decatenation activity is also dramatically affected by removal of the ParC CTD. Figure 5(c) shows the activity of wild-type and CTD-truncated topo IV on kinetoplast DNA, a system of relaxed, highly catenated 2.5 kb DNA minicircles classically used in decatenation assays. The kinetoplast DNA network is too large to enter the gel, so unreacted substrate remains in the wells and decatenated product appears as a single band. As with positively supercoiled DNA, truncation of the ParC CTD results in a ~100-fold reduction in decatenation activity by topo IV.

Overall, these data show that while wild-type topo IV is a robust decatenase and relaxes positive supercoils much more efficiently than negative supercoils, removal of the ParC CTD reduces the enzyme’s activity on all substrates to a similar “baseline” level. Strikingly, the CTD enforces substrate specificity by activating the enzyme up to 100-fold on its preferred substrates (positively supercoiled and catenated DNAs), while activating the enzyme to a lesser degree on negatively supercoiled DNA (Figure 5(d)). Time-course experiments performed with reconstituted enzymes held at a fixed concentration showed similar results for all assays (data not shown). Taken together, these data indicate that the ParC CTD is an important structural determinant of topo IV’s specificity for catenanes and positively supercoiled DNA.

Figure 5. DNA relaxation and decatenation by wild-type and CTD-truncated topo IV. Reconstituted wild-type or CTD-truncated topo IV was incubated at various concentrations (indicated above each lane, in nM topo IV) with negatively supercoiled plasmid (a), positively supercoiled plasmid (b), or kDNA (c). The positions of negatively and positively supercoiled DNA are indicated on the left-hand side of each panel by SC, and the distribution of relaxed topoisomers is indicated by R. In the decatenation assay (c), the unreacted kDNA network (K, left side) does not enter the gel, and remains in the wells. Fully decatenated mini-circles are indicated by MC, and incompletely decatenated products are indicated with asterisks (*). (d) Quantification of enzyme activities from (a)–(c). The specific activities of the two enzymes on each substrate were quantified (see Materials and Methods), normalized to that of CTD-truncated topo IV on negatively supercoiled DNA, and plotted on a log scale.

DNA binding by topo IV +/− ParC CTD

To ensure that the reduction in topo IV activity...
arising from removal of the ParC CTD did not result simply from a loss in DNA binding affinity, we next performed DNA binding assays with full-length and CTD-truncated topo IV. As shown in Figure 6, the full-length and CTD-truncated enzymes bind DNA with very similar affinities in filter-binding experiments ($K_d = 119 \pm 8$ nM for wild-type, 79 ($\pm 22$) nM for CTD-truncated). This finding shows that removal of the ParC CTD does not exert its effects on activity simply by weakening the enzyme’s ability to associate with DNA. It should be noted that while the DNA-binding affinities we observe are lower than those reported for wild-type topo IV, this difference likely arises from the use of chloride instead of glutamate as the predominant anion in our assays. This change was necessary to overcome modest aggregation of the CTD-truncated enzyme at low ionic strength in the presence of glutamate (data not shown).

Interestingly, the fractional saturation of the DNA-binding curve for CTD-truncated topo IV is reproducibly about half that of the wild-type enzyme (maximum fraction DNA bound = 0.52 $\pm$ 0.01 for wild-type, 0.26 $\pm$ 0.02 for CTD-truncated) (Figure 6). A concern with these data is that neither protein bound DNA at levels approaching full saturation. This again may be due to the use of chloride instead of glutamate in our assays, which could decrease the half-life of the protein–DNA interaction and allow DNA to be washed off the membrane more readily. As such, it is possible that the lower fractional saturation achieved with CTD-truncated topo IV is due to an increased dissociation rate compared to the wild-type enzyme. Alternatively, since the CTD-truncated enzyme consistently binds half as much DNA as the wild-type enzyme, it may be that wild-type topo IV is able to bind two DNA segments, while the CTD-truncated enzyme has lost the ability to bind one of these segments. Since both the wild-type and CTD-truncated topo IV constructs retain the primary G-segment binding site, the simplest interpretation of these results is that the ParC CTD makes up a T-segment binding site, and that this site is lost upon CTD truncation. This interpretation is supported by our relaxation and decatenation data, which suggests that the ParC CTD is involved in T-segment selection. In addition, our results agree with footprinting data showing that topo IV protects a region of the G-segment corresponding only to the primary DNA binding site, implying that no other elements of the enzyme are involved in G-segment binding.38

**Discussion**

**The ParC CTD is a DNA geometry sensor**

The data presented here allow us to explain how structural differences between two paralogous type IIA topoisomerases in *E. coli* lead to distinct functional profiles. In gyrase, it is thought that the GyrA CTD binds 40–50 base-pairs of DNA flanking a bound G-segment and imposes a “U-turn” on the DNA so that it wraps around the enzyme to be used as a T-segment in cis.25,27,28,51,52 (Figure 7(a)). This juxtaposition creates a positive crossover that is converted into a negative crossover upon strand passage, resulting in the introduction of negative supercoils into substrate DNA. Interestingly, the GyrA CTD appears to be only loosely tethered to the N-terminal region, and may adopt at least two positional states: an “upper” orientation for wrapping DNA prior to strand passage,33 and a “lower” position that may represent a resting or post-strand passage state.33 It has been proposed that the GyrA CTD may be able to cycle between these conformations during the strand passage cycle, and that this motion could be important for shuttling a T-segment through the enzyme.45

While topo IV has been shown to act specifically on positive crossovers and catenated DNAs, how the enzyme discriminates between different substrates is not understood. Recent experiments have shown that the substrate specificity of topo IV derives from a preference for binding incoming T-segments of a specific geometry, so-called “left-handed crossovers,” that occur more often in positively supercoiled DNA.36,37 (Figure 1(b)). DNA footprinting studies on gyrase and topo IV have indicated that despite their close homology, topo IV binds only a short (~30 bp) region of G-segment DNA,38 in contrast to the extended 120–150 bp footprint of gyrase.51,53,54 Topo IV binds
supercoiled DNA more tightly than linear, but does not distinguish positively from negatively supercoiled DNA at this step, indicating that G-segment binding is not sufficient to distinguish crossover geometry. Together, these results have suggested that the substrate specificity of topo IV may arise from the capture of incoming T-segments with particular crossing angles. Our current studies strongly support this model. Structural and biochemical data demonstrate that the ParC CTD is the major determinant for substrate selection by topo IV, and that this behavior likely arises from direct interactions with the T-segment. In the absence of the ParC CTD, topo IV acts on different substrates at a single low “baseline” level (Figure 5). With the CTD, however, the activity of topo IV is enhanced on all substrates, particularly on positively supercoiled and catenated DNAs.

This behavior can be explained by a scheme in which the ParC CTD not only helps recruit T-segments in trans, but also selects for crossover geometries common in positively supercoiled or catenated substrates (Figure 7(b) and (d)).

Our crystallographic studies provide a physical framework with which to understand this behavior. The structure of ParC27 shows that the ParC CTD is anchored to the top of the N-terminal region, where they may aid the binding of a T-segment and favor interactions with crossovers of the correct left-handed geometry (see Figure 1(b)).
most catenated DNAs should have no bias in crossover angle, a recent computational study by Buck et al. proposed that type IIA topoisomerases may specifically recognize catenated and knotted DNA crossings by virtue of the two DNA segments curving toward each other in a “hooked” juxtaposition. The authors propose that the ability of type IIA topoisomerases to bend a bound G-segment “upward” toward an incoming T-segment allows them to specifically recognize crossovers in a hooked geometry. With respect to topo IV, hooked geometries may be a general feature of DNA crossovers that this enzyme has evolved to recognize. For example, the two DNA duplexes in plectonemic supercoils curve toward each other, and a similar geometry is likely found in the positive-handed precatenanes that occur behind a DNA replication fork (Figure 1(a)). The relative positioning of the N and C-terminal domains of ParC, together with the CTD’s ability to bend DNA, could allow the enzyme to specifically bind T-segments bent toward the bound G-segment (Figure 7(d)). Thus, topo IV may recognize both DNA segments of a hooked juxtaposition, in contrast to canonical type II topos, which specifically recognize only one. Consistent with this possibility, topo IV is significantly more effective at recognizing and resolving knotted and catenated DNA crossovers than eukaryotic type IIA topos, which lack a homologous C-terminal domain.

Given the structural similarity and evolutionary relatedness between topo IV and DNA gyrase, it might be expected that the actions of the GyrA and ParC CTDs parallel one another to a certain extent. Indeed, our models invoke highly similar roles for these two domains, particularly when considering the different activities of these enzymes, and the fact that the DNA crossover imposed by the GyrA CTD likely adopts a geometry very much like that recognized by the ParC CTD (Figure 7(a) and (b)). It should be noted, however, that instead of imposing proper crossover geometry as per GyrA, ParC can only recognize it. In addition, the ParC CTD does not bind or bend DNA as effectively as its counterpart in gyrase, and thus may not be able to stabilize the DNA wrap required to negatively supercoil DNA. Together, the two enzymes represent an elegant example of how modest structural change can significantly impact enzymatic and biological function.

**Evolution of bacterial type IIA topoisomerases**

Of the two paralogous type IIA topoisomerases found in *E. coli*, topo IV has sometimes been considered to possess the more “canonical” enzyme activity and cellular function. By contrast, DNA gyrase has usually been thought of as a specialized enzyme evolved primarily to supercoil DNA. Recent studies, however, have clearly shown that topo IV is as functionally specialized as DNA gyrase. This information, along with emerging phylogenetic and structural data on both DNA gyrase and topo IV, has begun to paint a new picture of how bacterial type IIA topoisomerases likely evolved.

With the *E. coli* ParC CTD structure presented here, there now exist four distinct views of β-pinwheel domains. These structures, along with phylogenetic analyses and secondary structure prediction (Figure 8), indicate that all GyrA CTDs possess six full blades and an intact GyrA box sequence, which may or may not close the β-pinwheel into a circular structure (Supplementary Data, Figure S2), and likely also plays a role in DNA binding by the domain. The evolution of topo IV from this framework seems to have coincided with the progressive loss or alteration of structural features from this core domain, concomitant with functional modification of the enzyme. The first step in this sequence was likely the truncation of the GyrA box as observed in the *B. stearothermophilus* ParC CTD structure, which may have compromised the domain’s DNA wrapping function to some extent, in addition to enforcing the open spiral conformation. Functionally, this change likely abolished the supercoiling activity of the enzyme and allowed it to select for the binding and passage of DNA duplexes in trans. The open six-bladed ParC CTD architecture is found throughout the Firmicutes, encompassing the Bacilli and Mollicutes, and also in the Actinobacteria (Figure 8). In addition to the loss of the GyrA box, the ParC CTDs of the Proteobacteria are missing one entire blade, resulting in the five-bladed, open β-pinwheel typified by the ParC CTD of *E. coli*. Still other groups of bacteria appear to have gained more blades (Clostridia), lost most or all of the ParC CTD (the Spirochetes and Chlamydiales), or lack topo IV entirely (the Bacteroidetes/Chlorobi group, ε-Proteobacteria, and most Mycobacteria) (Figure 8). The deepest branches, possibly representing ancestral enzymes, contain CTDs whose lineage is somewhat ambiguous, but may be more GyrA-like. These incremental changes within the ParC family, together with wider distribution and higher conservation of gyrase than topo IV, indicate that ParC CTDs are degenerate forms of the GyrA CTD and that modern bacterial type IIA topoisomerases likely evolved from a gyrase-like enzyme. Interestingly, these findings echo the conclusions of an earlier study of bacterial type IIA topos, which found similar relationships by comparing the primary DNA binding regions in the NTDs of GyrA and ParC proteins.

The surprising structural variety in the β-pinwheel domains of different bacterial type IIA topos, together with the significant functional effects that appear to result from these structural changes, implies that there may exist a wide range of enzymes with activity profiles distinct from those of *E. coli* gyrase and topo IV. For example, there may be instances where the division of labor between the two enzymes is distinct from those found in *E. coli*, and some bacteria likely possess a single type IIA topoisomerase able to perform all functions...
Figure 8. Phylogenetic tree of full-length GyrA and ParC sequences. ParC sequences are shown in blue, GyrA sequences in red, and more ambiguous genes from *Aquifex aeolicus* and *Thermatoga maritima*, as well as from Cyanobacteria (which each possess two very similar GyrA proteins, one with a truncated GyrA box) are shown in black. Note that the branch lengths are longer overall for ParC versus GyrA, and that different subfamilies of ParC proteins are better defined by deep branches. Families of GyrA/ParC proteins are noted at the perimeter, along with representations of the probable CTD architecture in each family. The locations of the four CTD structures are noted with stars (*B. stearothermophilus* ParC was not included in the alignment, since the full-length ParC sequence corresponding to the structure is not available; however, its closest sequence relative (*B. subtilis*) is noted[27,39]). It is currently unknown how various GyrA C-terminal domains partition between the closed, *B. burgdorferi*-like and the open, *E. coli*-like architectures, but it is likely that the open form is more prevalent (see Figure S2). CTD architectures in those families for which structures are not available were inferred from inspection of the sequences and PSIPRED secondary structure predictions of subfamily sequence alignments[50]. Protein sequences used for the alignment can be found in Supplementary Data, Table S1.
attributed to gyrase and topo IV. In support of these ideas, a recent study found that DNA gyrase from *Mycobacterium smegmatis* possesses significant decatenation activity in addition to its supercoiling capability. This behavior may reflect the need for a minimal level of decatenation function in the *Mycobacteria*, a group that largely lacks topo IV. Interestingly, the recently sequenced genome of *M. smegmatis* revealed that this organism, the only “rapid growing” mycobacterium sequenced to date, is also the only sequenced mycobacterium to contain a topo IV homolog (The Institute for Genome Research). Thus, it may be that growth rate is one important factor determining an organism’s topoisomerase needs: the faster a cell divides, the more it may require a dedicated decatenating enzyme (topo IV) to avoid chromosome partitioning defects.

**Conclusion**

It is becoming evident that bacterial type IIA topoisomerases are highly specialized to satisfy specific cellular needs. The GyrA and ParC C-terminal domains have emerged as important elements that help define these enzymes’ unique activity profiles, and minor alterations to CTD structure or positioning can generate profound functional changes. This work shows how the *E. coli* ParC CTD helps control the substrate specificity of topo IV, and reveals that this mechanism exhibits significant parallels and contrasts with that of the paralogous DNA gyrase. Our findings suggest that the classic functional definitions of DNA gyrase and topo IV may be too restrictive, and that there may be a continuum of functions among bacterial type IIA topoisomerases. Further study of these enzymes from a range of bacteria will illuminate how type IIA topoisomerases are fine-tuned to meet specific needs for regulating chromosome superstructure during DNA replication, repair, and segregation.

**Materials and Methods**

**Cloning and protein purification**

 Constructs containing residues 497–752 (ParC CTD) or 27–742 (ParC CTD) of *E. coli* ParC were amplified from genomic DNA (ATCC) and cloned into pET28b behind an N-terminal, tobacco etch virus (TEV) protease-cleavable His<sub>6</sub> tag. The full-length *E. coli* parE and *parC* genes, as well as a CTD-truncated *parC* (ParC NTD) (residues 2–482) were also cloned into the same vector.

 Proteins were overexpressed in *E. coli* BL21-Codon-Plus(DE3)-RIL cells (Stratagene) by inducing with 0.5 mM isopropyl-1-thiogalactopyranoside (IPTG) at 37°C for four hours. Most cell cultures were harvested by centrifugation, resuspended in buffer A (20 mM Heps (pH 7.5), 10% (v/v) glycerol, 2 mM β-mercaptoethanol) plus 800 mM NaCl, 10 mM imidazole, 50 μg/ml of lysozyme, and protease inhibitors, and frozen drop-wise into liquid nitrogen. Cells expressing the ParC NTD were harvested and resuspended in buffer B (50 mM Tris–HCl (pH 7.5), 20% glycerol, 2 mM β-mercaptoethanol) plus 800 mM NaCl, 10 mM imidazole, 50 μg/ml of lysozyme, and protease inhibitors, and frozen.

 For protein purification, cells were sonicated and centrifuged, and the clarified lysate passed over a Ni<sup>2+</sup>-affinity column (Amersham). Peak fractions were pooled, concentrated and incubated overnight at 4°C with His<sub>6</sub>-tagged TEV protease<sup>66</sup> using a ratio of 1:50–1:10 (w/w) TEV protease/protein. This mixture was again passed over a Ni<sup>2+</sup>-affinity column to remove uncleaved protein, the cleaved His<sub>6</sub>-tag, and TEV protease. Proteins were then run over S200 or S300 gel filtration columns (Amersham Biosciences), except for the ParC NTD, which was further purified over a HiTrap Q ion-exchange column (Amersham Biosciences). Proteins used for crystallography were concentrated by ultrafiltration and kept at 4°C. Proteins used for biochemical assays were flash-frozen in aliquots and stored at −80°C. Purification of selenomethionine-labeled ParC CTD, prepared by the method described by Van Duyne et al.,<sup>67</sup> was performed as for native protein, with the addition of 1 mM Tris(2-carboxyethyl)phosphine (TCEP) (Fluka) in the gel-filtration step and thereafter.

 To confirm that the two ParC constructs formed heterotetramers with similar efficiency, equimolar amounts of ParC and ParE subunits were mixed, incubated 30 minutes on ice, and passed over an analytical gel filtration column (Superdex-200 HR 10/30; Amersham Biosciences) (data not shown).

**Crystallization, data collection, and structure solution**

 For crystallization of the *E. coli* ParC CTD, purified protein (15–25 mg/ml) was dialyzed overnight at 4°C against 20 mM Heps (pH 7.5) and 100 mM NaCl (plus 0.5 mM TCEP for selenomethionine-labeled protein). Crystals were grown at 19°C or 4°C in hanging drop format by mixing 1 μl of protein with 1 μl of well solution (100 mM sodium acetate (pH 4.5), 50 mM ammonium acetate, and 20% (w/v) polyethylene glycol (PEG) 4000). For harvesting, a cryoprotectant solution containing well solution plus 25% ethylene glycol was added directly to the drop, and the crystals were immediately looped and flash-frozen in liquid nitrogen.

 Crystals of ParC<sub>27</sub> were grown in hanging drop format by mixing 1 μl of purified protein (20 mg/ml) in buffer A plus 200 mM NaCl with 1 μl well solution (20 mM Heps (pH 7.5), 10% glycerol, 100 mM NaCl, 4% PEG 6000, and 8% (v/v) 1,3-butane diol). For harvesting, a cryoprotectant solution containing well solution plus 20% PEG 400 was added directly to the drop, and the crystals were immediately looped and flash-frozen in liquid nitrogen.

 All datasets were collected on Beamline 8.3.1 at the Advanced Light Source at Lawrence Berkeley National Laboratory. Data were indexed and reduced with HKL2000<sup>68</sup> or ELVES<sup>64</sup> using MOSFLM<sup>65</sup>. For the *E. coli* ParC CTD, phasing was performed using single and multi-wavelength anomalous diffraction (SAD/MAD) methods with selenomethionine-substituted crystals. Selenium sites were first located using SOLVE<sup>69</sup> from a single-wavelength dataset (SeMet SAD; Table 1). These sites were supplied to MLPHARE<sup>65</sup> along with a
multi-wavelength dataset (SeMet MAD Peak and Remote; Table 1) for phase calculation and refinement, and density modification was performed using DM to produce initial electron density maps. A preliminary model of the ParC CTD was built with RESOLVE and manual rebuilding was performed with O. Refinement was carried out using a Refmac/ARP procedure to automatically place ordered water molecules, followed by TLS refinement in Refmac5. The final model consists of two chains of amino acid residues 497–742 of the ParC CTD; residue 694 of chain A and residues 743–752 of both chains are missing from the model. A total of 94.6% of non-glycine residues are in the most favored regions of Ramachandran space, and none are in disallowed regions.

For ParC27, molecular replacement was performed at 3.0 Å resolution using PHASER, searching with poly- merine models of the ParC CTD and the CAP and tower domains (residues 30–362) of E. coli DNA gyrase A. The remaining domains of E. coli GyrA (residues 363–522) were placed manually into initial maps produced by PHASER. Solvent flattening was performed with DM and prime and switch phasing was performed with RESOLVE to produce electron density maps suitable for manual rebuilding. Multiple rounds of simulated annealing refinement in CNS and manual rebuilding reduced the free R-factor to ~34%. TLS refinement in Refmac5 and manual placement of ordered water molecules yielded a final R-factor of 24.0% and a free R-factor of 29.6% (Table 1). The final model consists of residues 28–740 of ParC, with disordered loop residues 55–64, 516–530, and 565–567 missing from the model. A total of 86.5% of non-glycine residues are in the most favored regions of Ramachandran space, and none are in disallowed regions. Electrostatic surfaces were calculated with APBS. All molecular Figures were produced with PyMOL.

**DNA relaxation/decatenation assays**

Plasmid pSG483, a 3 kb derivative of pBluescript II SK (Stratagene), was used for supercoiled DNA substrates. Negatively supercoiled plasmid was purified from E. coli XL1 Blue cells using CsCl gradients. Positively supercoiled DNA was made using A. fulgidus reverse gyrase. Kinetoplast DNA (kDNA) from Crithidia fasciculata was purchased from Topogen. DNA relaxation or decatena- tion assays (20 μl) were performed in a buffer containing 50 mM Tris-HCl (pH 7.8), 6 mM MgCl2, 10 mM DTT, 1 mM ATP, 20 mM KCl, 1 mM spermidine, 100 μg/ml bovine serum albumin (BSA), with 300 ng supercoiled DNA or kDNA. Varying amounts of reconstituted wild-type or CTD-truncated topo IV were added to reaction mixtures, and reactions were allowed to proceed for 30 minutes at 30°C. Reactions were stopped by the addition of SDS (1% final) and EDTA (1 mM final). Stopped reactions were analyzed by electrophoresis through 1.2% (w/v) agarose gels with 0.5X TBE running buffer. Gels were run at 2 V/cm for 12–18 hours, stained with ethidium bromide (EtBr), and visualized by UV illumination. To corroborate the findings from enzyme titrations, time-course assays were run with constant amounts of protein (data not shown).

To determine relative activities of topo IV enzymes on different substrates, we quantified the amounts of supercoiled and relaxed DNA in each lane using ImageJ. First we divided each lane into two zones corresponding to “supercoiled” and “relaxed,” with the split halfway between the most supercoiled and relaxed species. Next, we corrected for background staining and the different staining intensity of supercoiled versus relaxed DNA, and plotted the percent of relaxed/decatenated DNA versus enzyme concentration. Fitting these points allowed us to estimate the amount of enzyme needed to achieve “half-relaxation,” and thus determine a rough measure of specific activity. These measures are to some extent approximate, but should be internally consistent for the purposes of comparison. For the graph in Figure 5(d), all specific activities are normalized to that of CTD-truncated topo IV on negatively supercoiled DNA.

**Topo IV DNA binding assays**

Purified negatively supercoiled pSG483 was linearized with BamHI (NEB), radiolabeled with T4 polynucleotide kinase in the presence of [γ-32P]ATP (3000 Ci/mmol), and separated from unincorporated nucleotides using a spin column (BioRad; Biogel P-30). Filler binding was carried out using a Schleicher and Schuell 96-well MiniFold dot-blot apparatus. Whatman filter paper and nitrocellulose and Nytran membranes (Schleicher and Schuell) were wet briefly in water, then soaked in wash buffer (25 mM Tris–HCl (pH 7.5), 20 mM NaCl, 10 mM MgCl2, 1 mM EDTA, 10 mM β-mercaptoethanol). Two Whatman filters, the Nytran membrane, and the nitrocellulose membrane were placed, in that order, on the MiniFold apparatus.

Binding reactions (160 μl) were carried out in binding buffer (25 mM Tris–HCl (pH 7.6 at 30°C), 13 mM NaCl, 6 mM KCl, 10 mM MgCl2, 2 mM DTT, 50 μg/ml BSA). Varying amounts of reconstituted topo IV were incubated with 0.5 nM 32P-labeled linearized pSG483 for 30 minutes at 30°C, then three 50 μl aliquots were added to separate 350 μl aliquots of binding buffer. Reaction mixtures were passed through the filters, then each well was washed three times with 450 μl of wash buffer and membranes were removed from the apparatus and dried. Dried filters were exposed to a BioRad imaging screen, then scanned using a Typhoon 8600 imager (Amersham Biosciences). The amounts of DNA bound at each spot were quantified by densitometry (ImageQuant). Total DNA amounts for each well were determined by summing the densities of spots on the nitrocellulose (protein–DNA complex) and Nytran (free DNA) filters. The fraction of DNA bound to protein was plotted versus protein (heterotetramer) concentration using KaleidaGraph (Synergy Software) and fit to an independent-sites binding model:

\[
\tau = \frac{z_M (|P|_T + |DNA|_T + K_d(app)) - ((|P|_T + |DNA|_T + K_d(app))^2 - 4|P|_T|DNA|_T)^{0.5}}{2|DNA|_T} 
\]

where \(z\) is the percentage of DNA bound, \(z_M\) is the percentage of DNA bound at saturation, \(|P|_T\) is the total protein concentration, \(|DNA|_T\) is the total concentration of DNA and \(K_d(app)\) is the apparent dissociation constant.

**Sequence alignments**

All annotated bacterial GyrA and ParC sequences (as of March 2004) were downloaded from TIGR. In cases

† www.tigr.org
where multiple annotations disagreed in terms of gene length, the TIGR annotated version was used. Sequences were aligned using CLUSTALX, and an unrooted phylogenetic tree was generated using DRAWTREE from the PHYLIP software package.

Coordinates

The coordinates of the two structures have been deposited in the RCSB Protein Data Bank under accession numbers 1ZVT (ParC CTD) and 1ZVU (ParC27).

Acknowledgements

The authors thank J. Holton, J. Tanamachi, and G. Meigs for assistance at ALS Beamline 8.3.1; S. Gradia for supplying plasmid pSG483; A. McClendon and N. Osheroff for assistance with the preparation of positively supercoiled DNA; and D. Classen, J. Erzberger, N. Crisona, N. Cozzarelli, and H. Hiasa for helpful discussions and valuable assistance. K.D.C. & A.J.S. acknowledge support from NSF Graduate Research Fellowships, and assistance. K.D.C. & A.J.S. acknowledge support from the NIH (CA 77373).

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2005.06.029

References


Edited by J. Doudna

(Received 27 April 2005; received in revised form 9 June 2005; accepted 10 June 2005)
Available online 28 June 2005
**Supplemental Information**

**Figure S1.** Structural comparison of the four β-pinwheel domains solved to date. Comparison of the helical pitch of the *B. burgdorferi* GyrA CTD (gray), *E. coli* GyrA CTD (green), *B. stearothermophilus* ParC CTD (red), and the *E. coli* ParC CTD (blue). Blade 1 of each domain is shown in yellow. All four domains were aligned on common β-sheet elements of blade 1 (average pairwise Cα r.m.s.d. 0.99 Å). The *B. burgdorferi* GyrA CTD is a closed circular domain and has almost no helical rise about its perimeter. In contrast, the *E. coli* GyrA CTD and the *B. stearothermophilus* ParC CTDs show a distinct right-handed helical rise. The *E. coli* ParC CTD shows a similar helical rise, but the effect is much less pronounced than in the previous two instances.

**Figure S2.** The closed-circular conformation of the *B. burgdorferi* GyrA CTD is likely rare or unique. (a) A section of the sequence alignment used to create Figure 8 (showing a subset of GyrA sequences only) shows that the GyrA box motif (consensus sequence Q++GG+G, where + is a positively charged residue: arginine or lysine) of *B. burgdorferi* (highlighted in yellow) is highly divergent compared to that of *E. coli* (highlighted in green). In particular, the threonine residue in the third position is not shared with any other bacterium. (b) Detailed view of the GyrA box in the *B. burgdorferi* GyrA CTD structure. Blade 1 is shown in stick view with each residue of the GyrA box labeled, while the rest of the domain is shown in surface representation. The γ-methyl group of Thr531 docks into a small hydrophobic depression in blade six (colored yellow, including residues Ile765, Thr777, Val782, and Val794). This interaction may allow the *B. burgdorferi* GyrA CTD to form a closed circle, while those of most other bacteria may be open with disordered GyrA box motifs. It remains to be determined how many bacterial GyrA CTDs adopt each possible conformation.
Table S1 – GyrA and ParC sequences used in Figures 8 and S2

<table>
<thead>
<tr>
<th>Actinobacteria</th>
<th>ParC - TC0031</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actinobacteridae</td>
<td><em>Chlamydia pneumoniae</em> CWL029</td>
</tr>
<tr>
<td>Actinomycetales</td>
<td><em>GyrA</em> - CP04274</td>
</tr>
<tr>
<td><em>Corynebacterium glutamicum</em> ATCC 13032</td>
<td><em>ParC</em> - CP01761</td>
</tr>
<tr>
<td><em>GyrA</em> - NT01CG0012</td>
<td><em>Chlamydia pneumoniae</em> AR39</td>
</tr>
<tr>
<td><em>Corynebacterium efficiens</em> YS-314</td>
<td><em>GyrA</em> - CP0485</td>
</tr>
<tr>
<td><em>Mycobacterium tuberculosis</em> H37Rv</td>
<td><em>ParC</em> - CP0030</td>
</tr>
<tr>
<td><em>GyrA</em> - NT02MT0007</td>
<td><em>Chlamydia pneumoniae</em> J138</td>
</tr>
<tr>
<td><em>Mycobacterium tuberculosis</em> CDC1551</td>
<td>(no annotated GyrA or ParC)</td>
</tr>
<tr>
<td><em>GyrA</em> - MT0006</td>
<td><em>Chlamydophila caviae</em> GPIC</td>
</tr>
<tr>
<td><em>Mycobacterium leprae</em> TN</td>
<td><em>GyrA</em> - CC.A00507</td>
</tr>
<tr>
<td><em>GyrA</em> - NT01ML0005</td>
<td></td>
</tr>
<tr>
<td><em>Mycobacterium bovis</em> subsp. bovis AF2122/97</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - Mb0006</td>
<td></td>
</tr>
<tr>
<td><em>Tropheryma whipplei</em> str. Twist</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - TW006</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> (ParC) - TW491</td>
<td></td>
</tr>
<tr>
<td><em>Tropheryma whipplei</em> TW08/27</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - TW0006</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> (ParC) - TW0254</td>
<td></td>
</tr>
<tr>
<td><em>Streptomyces coelicolor</em> A3(2)</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - SCO3873</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> (ParC) - SCO5836</td>
<td></td>
</tr>
<tr>
<td><em>Streptomyces avermitilis</em> MA-4680</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - SAV4322</td>
<td></td>
</tr>
<tr>
<td><em>ParC</em> - SAV2423</td>
<td></td>
</tr>
<tr>
<td><em>Bifidobacteriales</em></td>
<td></td>
</tr>
<tr>
<td><em>Bifidobacterium longum</em> NCC2705</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - BL0634</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> (ParC) - NT01BL1519</td>
<td></td>
</tr>
<tr>
<td><em>Bacteria</em></td>
<td></td>
</tr>
<tr>
<td><em>Aquificae</em></td>
<td></td>
</tr>
<tr>
<td><em>Aquificales</em></td>
<td></td>
</tr>
<tr>
<td><em>Aquificaceae</em></td>
<td></td>
</tr>
<tr>
<td><em>Aquifex aeolicus</em> VF5</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - aq_980</td>
<td></td>
</tr>
<tr>
<td><em>Bacteroides</em></td>
<td></td>
</tr>
<tr>
<td><em>Bacteroidales</em></td>
<td></td>
</tr>
<tr>
<td><em>Bacteroides thetaiotaomicron</em> VPI-5482</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - BT0899</td>
<td></td>
</tr>
<tr>
<td><em>Porphyromonas gingivalis</em> W83</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - PG1386</td>
<td></td>
</tr>
<tr>
<td><em>Bacteroidetes/Chlorobi group</em></td>
<td></td>
</tr>
<tr>
<td><em>Bacteroides (class)</em></td>
<td></td>
</tr>
<tr>
<td><em>Prevotella intermedia</em> 17</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> (ParC) - PINA0121</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - PINA1342</td>
<td></td>
</tr>
<tr>
<td><em>Chlamydiae</em></td>
<td></td>
</tr>
<tr>
<td><em>Chlamydiaceae</em></td>
<td></td>
</tr>
<tr>
<td><em>Chlamydia trachomatis</em> serovar D</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - CT189</td>
<td></td>
</tr>
<tr>
<td><em>ParC</em> - CT660</td>
<td></td>
</tr>
<tr>
<td><em>Chlamydia muridarum</em> strain Nigg</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - TC0461</td>
<td></td>
</tr>
<tr>
<td><em>Chlorobi</em></td>
<td></td>
</tr>
<tr>
<td><em>Chlororhabdus</em></td>
<td></td>
</tr>
<tr>
<td><em>Chlororhabdaceae</em></td>
<td></td>
</tr>
<tr>
<td><em>Chlorobacteriales</em></td>
<td></td>
</tr>
<tr>
<td><em>Chlorobacterium tepidum</em> TLS</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - CT0141</td>
<td></td>
</tr>
<tr>
<td><em>Cyanobacteria</em></td>
<td></td>
</tr>
<tr>
<td><em>Chroococcales</em></td>
<td></td>
</tr>
<tr>
<td><em>Synechocystis</em></td>
<td></td>
</tr>
<tr>
<td><em>Synechocystis</em> sp. PCC6803</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - sll1941</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - NT01SS3061</td>
<td></td>
</tr>
<tr>
<td><em>Thermosynechococcus</em></td>
<td></td>
</tr>
<tr>
<td><em>Thermosynechococcus elongatus</em> BP-1</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - tlr1369</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - tll1617</td>
<td></td>
</tr>
<tr>
<td><em>Nostocales</em></td>
<td></td>
</tr>
<tr>
<td><em>Nostocaceae</em></td>
<td></td>
</tr>
<tr>
<td><em>Nostoc</em> sp. PCC 7120</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - al10441</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - al10860</td>
<td></td>
</tr>
<tr>
<td><em>Deinococcus-Thermus</em></td>
<td></td>
</tr>
<tr>
<td><em>Deinococci</em></td>
<td></td>
</tr>
<tr>
<td><em>Deinococcaceae</em></td>
<td></td>
</tr>
<tr>
<td><em>Deinococcus radiodurans</em> R1</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - DR1913</td>
<td></td>
</tr>
<tr>
<td><em>Firmicutes</em></td>
<td></td>
</tr>
<tr>
<td><em>Bacillales</em></td>
<td></td>
</tr>
<tr>
<td><em>Bacillaceae</em></td>
<td></td>
</tr>
<tr>
<td><em>Bacillus subtilis</em> 168</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - Bsu0007</td>
<td></td>
</tr>
<tr>
<td><em>ParC/GrlA</em> - Bsu1810</td>
<td></td>
</tr>
<tr>
<td><em>Bacillus halodurans</em> C-125</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - BHa_gyrA</td>
<td></td>
</tr>
<tr>
<td><em>Bacillus anthracis</em> Ames</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - BA0006</td>
<td></td>
</tr>
<tr>
<td><em>ParC</em> - BA3565</td>
<td></td>
</tr>
<tr>
<td><em>Bacillus cereus</em> ATCC14579</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - BC0006</td>
<td></td>
</tr>
<tr>
<td><em>Oceanobacillus iheyensis</em> HTE831</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - OB0007</td>
<td></td>
</tr>
<tr>
<td><em>ParC/GrlA</em> - OB1694</td>
<td></td>
</tr>
<tr>
<td><em>Listeriaceae</em></td>
<td></td>
</tr>
<tr>
<td><em>Listeria innocua</em> CLIP 11262</td>
<td></td>
</tr>
<tr>
<td><em>GyrA</em> - NT01LI0007</td>
<td></td>
</tr>
<tr>
<td><em>ParC</em> - NT01LI1412</td>
<td></td>
</tr>
<tr>
<td><em>Listeria monocytogenes</em> EGD-e</td>
<td></td>
</tr>
</tbody>
</table>
GyrA - NT01LM0007  
ParC - NT01LM1386  

Staphylococcus  
*S. aureus* COL  
GyrA - SA0006  
ParC - SA1390  

*S. aureus* N315  
GyrA - SA_2_0006  
ParC - SA_2_1189  

*S. aureus* Mu50  
GyrA - SAV0006  
ParC - SAV1340  

*S. aureus* MW2  
GyrA - MW0006  
ParC - MW1412  

*S. epidermidis* ATCC 12228  
GyrA - SE0005  
ParC - SE1037  

*S. aureus* Michigan VRSA  
no annotated GyrA or ParC genes  

Clostridia  
Clostridiales  
*C. acetobutylicum* ATCC824  
GyrA - CAC0007  
**GyrA (ParC)** - CAC1628  

*C. perfringens* 13  
GyrA - CPE0007  
**GyrA (ParC)** - CPE2068  

*C. tetani* E88  
GyrA - CTC00090  
**GyrA (ParC)** - CTC01766  

Thermoanaerobacteriales  
*Thermoanaerobacter tengcongensis* MB4(T)  
GyrA - TTE0011  

Lactobacillales  
Enterococcaceae  
*Enterococcus faecalis* V583  
GyrA - EF0006  
ParC - EF1614  

Lactobacillaceae  
*Lactobacillus planta*r*um* WCFS1  
GyrA - lp_0007  
ParC - lp_1839  

Streptococcaceae  
*Lactococcus lactis* subsp. lactis IL1403  
GyrA - Lla_gyrA  
ParC - Lla_parC  

*Streptococcus pyogenes* TIGR4  
GyrA - SP1219  
ParC - SP0855  

*Streptococcus pneumoniae* R6  
GyrA - spr0757  
ParC - spr1099  

*Streptococcus pyogenes* MGAS8232  
GyrA - spyM18_1112  
ParC - spyM18_0968  

*Streptococcus pyogenes* SF370 serotype M1  
GyrA - NT01SP1022  
ParC - NT01SP0801  

*Streptococcus agalactiae* 2603V/R  
GyrA - SAG0960  
ParC - SAG1153  

*Streptococcus pneumoniae* MGAS315  
GyrA - SpyM3_0810  
ParC - SpyM3_0625  

*Streptococcus mutans* UA159  
GyrA - SMU.1114  
ParC - SMU.1204  

*Streptococcus agalactiae* NEM316  
GyrA - Sag_gyrA  

*Streptococcus pyogenes* SSI  
GyrA - SPs1009  
ParC - SPs1228  

Mollicutes  
Mycoplasmataceae  
*Mycoplasma genitalium* G-37  
GyrA - MG004  
ParC - MG204  

*Mycoplasma pneumoniae* M129  
GyrA - MPN004  
ParC - MPN123  

*Ureaplasma urealyticum* parvum biovar serovar 3  
GyrA - UU082  
ParC - UU467  

*Mycoplasma pulmonis* UAB CTIP  
GyrA - MYPU_1470  
ParC - MYPU_3730  

*Mycoplasma penetrans* HF-2  
GyrA - MYPE40  
ParC - MYPE6470  

*Mycoplasma gallisepticum* strain R  
GyrA - MGA_0612  
ParC - MGA_0056  

Fusobacteria  
Fusobacteriales  
Fusobacteriaceae  
*Fusobacterium nucleatum* ATCC 25586  
GyrA - FN2125  

Planctomycetes  
Planctomycetacia  
Planctomycetales  
*Pirellula* sp strain 1  
**GyrA (ParC)** - RB3465  
GyrA - RB8841  

Proteobacteria  
α-Proteobacteria  
Caulobacteriales  
*Caulobacter crescentus* CB15  
GyrA - CC1580  
ParC - CC1566  

Rhizobiales  
*Bradyrhizobium japonicum* USDA 110  
GyrA - bll4696  
**GyrA (ParC)** - bll5057  

*Brucella suis* 1330  
GyrA - BR1097  
ParC - BR0754  

*Brucella melitensis* 16M  
GyrA - NT01BM1124  
ParC - NT01BM1522
Mesorhizobium loti MAFF303099
  GyrA - mll0732
  ParC - SMc01761

Sinorhizobium meliloti 1021
  GyrA - SMc01231
  ParC - SMc01761

Agrobacterium tumefaciens C58 Cereon
  GyrA - AGR_C_2778
  ParC - AGR_C_2144

Agrobacterium tumefaciens C58 UWash
  GyrA - NT02AT1837
  ParC - Atu1158

Rickettsia prowazekii Madrid E
  GyrA - RP206
  ParC - RP067

Rickettsia conorii Malish 7
  GyrA - RC0273
  ParC - RC0097

Wolbachia pipiensis wMel
  GyrA - WD1202

β-Proteobacteria
Neisseriales
  Neisseria meningitidis MC58
    GyrA - NMB1384
    ParC - NMB1605
  Neisseria meningitidis serogroup A Z2491
    GyrA - NMA1600
    ParC - NT01NM1859

Nitrosomonadales
  Nitrosomonas europaea ATCC 19718
    GyrA - NE0332

Ralstonia solanacearum GMI1000
  GyrA - NT01RS0957

ε-Proteobacteria
Campylobacteriae
  Campylobacter jejuni NCTC 11168
    GyrA - Cj1027c
  Helicobacter hepaticus ATCC 51449
    GyrA - HH1633
  Helicobacter pylori 26695
    GyrA - HP0701
  Helicobacter pylori 999
    GyrA - jhp0641

Wolinella succinogenes DSMZ 1740
  GyrA - WS0365

γ-Proteobacteria
Alteromonadales
  Shewanella oneidensis MR-1
    GyrA - SO2411
    ParC - SO3897

Enterobacteriales
  Buchnera aphidicola (Baizongia pistaciae)
    GyrA - bpp169
  Buchnera aphidicola Sg
    GyrA - BUsg174
  Escherichia coli K12-MG1655
    GyrA - b2231
    ParC - b3019
  Escherichia coli O157:H7 EDL933
    GyrA - Z3484

  ParC - Z4373

Escherichia coli O157:H7 VT2-Sakai
  GyrA - ECs3114
  ParC - ECs3903

Escherichia coli CFT073
  GyrA - c2773
  ParC - c3760

Salmonella typhimurium LT2 SGSC1412
  GyrA - STM2272
  ParC - STM3174

Salmonella enterica serovar Typhi CT18
  GyrA - STY2499
  ParC - STY3351

Salmonella enterica serovar Typhi Ty2
  GyrA - t0592
  ParC - t3095

Shigella flexneri 2a 2457T
  GyrA - S2444
  ParC - S3267

Shigella flexneri 2a str. 301
  GyrA - Shf_gyrA
  ParC - Shf_parC

Wigglesworthia glossinidia brevipalpis
  GyrA - Wbr0314

Yersinia pestis CO92
  GyrA - YPO1216
  ParC - YPO0671

Yersinia pestis KIM
  GyrA - y2972
  ParC - y3507

Buchnera sp. APS
  GyrA - BU180

Legionellaceae group
  Coxiella burnetii RSA 493
    GyrA - CBU0524
    ParC - CBU1866

Pasteurellaceae
  Pasteurella multocida PM70
    GyrA - PM0841
    ParC - PM0369

Pasteurellales
  Haemophilus influenzae KW20 Rd
    GyrA - HH1264
    ParC - H11529

Pseudomonadaeae
  Pseudomonas aeruginosa PAO1
    GyrA - PA3168

Pseudomonas putida KT2440
  GyrA - PP1767
  ParC - PP4912

Pseudomonas syringae DC3000
  GyrA - PSPT01745
  ParC - PSPT04960

Vibrionales
  Vibrio cholerae El Tor N16961
    GyrA - VC1258
    ParC - VC2430

  Vibrio parahaemolyticus RIMD 2210633
    GyrA - VP1932
    ParC - VP0431

  Vibrio vulnificus CMCP6
    GyrA - VV13038
ParC - VV10605
Xanthomonadales
Xylella fastidiosa Temecula1
  GyrA - NT02XF2589
  ParC - PD0594
Xylella fastidiosa 9a5c
  GyrA - NT01XF2866
  ParC - XF1353
Xanthomonas group
Xanthomonas campestris pv. campestris ATCC33913
  GyrA - XCC1574
  ParC - XCC1395
Xanthomonas axonopodis pv. citri 306
  GyrA - XAC1631
  ParC - NT01XA1962
δ subdivision
Desulfuromonas group
  Geobacter sulfurreducens PCA
    GyrA - GSU0004
Spirochaetes

GyrA and ParC genes are listed as annotated in the TIGR database (www.tigr.org). Those annotations that are crossed out are likely mis-assigned in the database, and are reassigned here based on how each sequence segregated in our analysis.