Skolkovo Institute of Science and Technology

REGULATION OF BACTERIAL GENOME TOPOLOGY BY TOPOISOMERASES

Doctoral Thesis

by

Dmitry Sutormin

DOCTORAL PROGRAM IN LIFE SCIENCES

Supervisor

Professor Konstantin Severinov

Moscow – 2022

I hereby declare that the work presented in this thesis was carried out by myself at Skolkovo Institute of Science and Technology, Moscow, except where due acknowledgement is made, and has not been submitted for any other degree.

Dmitry Sutormin
Prof. Konstantin Severinov

Abstract

Topoisomerases are essential and ubiquitous enzymes that regulate supercoiling of nucleic acids by introducing temporary breaks in DNA and then religating them. Information about the distribution of topoisomerase activity across genomes is scarce, and it is generally unclear how a topoisomerase chooses a site to be cleaved.

We have developed Topo-Seq, a method that allows genome-wide strand-specific mapping of topoisomerase cleavage sites with single-base precision. We applied Topo-Seq to study three major topoisomerases of *Escherichia coli* – TopoI, DNA gyrase, and TopoIV. We found that while TopoI is enriched upstream of active transcription units (TUs), where transcription-induced negative supercoiling is high, it is lacking in downstream regions, where positive supercoiling is accumulated. The opposite enrichment pattern is found for the gyrase, fully consistent with the Liu & Wang twin-domain model predictions. Consistently, gyrase cleavage was found enhanced in downstream regions of highly expressed rRNA operons in *Caulobacter crescentus* genome. Genome-wide sampling of *E. coli* gyrase activity by Topo-Seq in stationary growth phase and in synchronously replicating cells revealed that gyrase is associated with regions of transcription and replication, respectively, supporting the major role of this enzyme in relaxation of positive supercoiling arising during these processes. Similarly to gyrase, the activity of TopoIV was found to be higher in regions downstream of TUs, indicating that this topoisomerase may also be involved in relaxation of positive supercoils.

The supreme resolution of the Topo-Seq method allowed us to identify topoisomerase cleavage motifs. We found that TopoI cleaves AT-rich sequences between T and A, when there is a C in the -4 position with respect to the cleavage site. For DNA gyrase we detected a long motif that reflects the wrapping of DNA around the CTDs of the GyrA subunits. The TopoIV motif was shorter and lacked the signs of DNA wrapping, indicating that DNA segment to be transferred through the temporary DNA break belongs to a different DNA region for TopoIV. Taken together our data indicate that topoisomerases identify optimal DNA sites by sensing both DNA topology and local sequence patterns.

Publications

1. **Sutormin D**., Rubanova N., Logacheva M., Ghilarov D., Severinov K. Single-nucleotide-resolution mapping of DNA gyrase cleavage sites across the *Escherichia coli* genome. **Nucleic Acids Research**, Vol. 47(3), p.1373–1388, 2019
2. **Sutormin D**., Galivondzhyan A., Polkhovskiy A., Kamalyan S., Severinov K., Dubiley S. Diversity and functions of type II topoisomerases. **Acta Naturae**, Vol. 13(1), p.59-75, 2021
3. **Sutormin D**., Galivondzhyan A., Musharova O., Travin D., Rusanova A., Obraztsova K., Borukhov S., Severinov K. Interaction Between Transcribing RNA Polymerase and Topoisomerase I Prevents R-loop Formation in *E. coli*. **Nature Communications***,* Vol. 13, 4524, 2022
4. Rusanova A., Fedorchuk V., Toshchakov S., Dubiley S., **Sutormin D**. An interplay between viruses and bacteria associated with marine sponges from the White Sea revealed by metagenomics. **Life**, Vol. 12(1), p.1-25, 2022

Conferences

- **Sutormin D**., Ghilarov D., Rubanova N., Guo M., Laub M., Severinov K. Whole genome DNA gyrase cleavage sites identification with single-nucleotide resolution. Chromosome architecture and topological stress workshop, National University of Andalusia, Baeza, Spain, 2018. **Poster presentation**
- **Sutormin D**. Role of topoisomerases in regulation of bacterial genome topology. Biology - science of XXI century 2019, Pushino, Russia, 2019. **Plenary talk**
- **Sutormin D**., Galivondzhyan A., Severinov K. Topoisomerase I binding sites identification in the *E. coli* genome. OpenBio, 2019, ISBN 978-5-4437-0957-4
- **Sutormin D**., Galivondzhyan A., Severinov K. Role of DNA topoisomerases in bacterial genome maintenance. Global Young Scientists Summit, Singapore, 2020. **Poster presentation**
- **Sutormin D**., Severinov K. DNA gyrase re-positions on *E. coli* genome sites in accordance with transcription profile in different growth phases. International Conference Lomonosov, Moscow, Russia, 2020. **Oral presentation**
- **Sutormin D**., Galivondzhyan A., Severinov K. Identification of binding and cleavage sites of topoisomerase I in the *E. coli* genome with ChIP-Seq and Topo-Seq. International Conference Lomonosov, Moscow, Russia, 2021. **Oral presentation**. *The talk was selected as the best talk of the Molecular Biology track of the conference*
- **Sutormin D**., Ghilarov D., Logacheva M., Severinov K. Identification of topoisomerase cleavage sites with a single-base precision using Topo-Seq. EMBO Workshop DNA Topology in genomic transactions, virtual, 2021. **Poster presentation**. *The poster was selected as the best poster of the workshop*

Acknowledgements

A deepest gratitude, first of all, should be expressed to my supervisor Konstantin Severinov, who once gave a *carte blanche* for my research project and constantly shares his scientific experience. The work would not be possible without Dmitry Ghilarov who mentored me a lot at the early stages of this scientific journey. I would like to acknowledge Natalia Rubanova and Vasiliy Sitnik for their support in NGS data analysis. Special thank should be addressed to Maria Logacheva not only for high quality DNA sequencing, but also for wise handling of complicated samples. I also acknowledge Olga Musharova for her help with sequencing of some samples. Particularly, I would like to pay respect to Ekaterina Savitskaya, who sadly passed away, for her help and sharing the expertise in NGS library preparations.

I am grateful to Svetlana Dubiley for fruitful discussions, for her every day help and support at the IGB RAS. My best regards to Sergey Borukhov who always inspires me during our long and fruitful discussions. Many thanks to Marina Serebryakova for excellent mass spectrometry analysis. I would like to thank my friend, Dmitrii Travin, for his invaluable help with some experiments and stimulating discussions. I am grateful to Aleksandra Galitsyna, who introduced me to and performed the Hi-C data analysis. Many thanks to Anastasiia Rusanova, who helped me with microscopy.

I would like to express my deepest gratitude to my colleagues from our Topo-team – to Alina Galivondzhyan, Sofia Kamalyan, and Alex Polkhovskiy. It is a pleasure and honor to work with you, and, I feel, we have a lot more fascinating puzzles to solve in the future.

I am grateful to Mike Laub and Monica Guo, who once hosted me at MIT and gave me a great opportunity to study Hi-C and set up experiments with *C. crescentus*. Many thanks to Yuk-Ching Tse Dinh and Marc Drolet for stimulating discussions and sharing some *E. coli* strains. I am extremely grateful to the whole TOPO Community for a friendly and inspiring atmosphere, fascinating talks, help and support.

Table of Contents

Abbreviations

A, C, G, T – adenine, cytosine, guanine, thymine, respectively

ACN – acetonitrile

BIME – bacterial interspersed mosaic elements

bp – base pair

CFU – colony forming unit

cSDR – constitutive stable DNA replication

CTD – C-terminal domain

DMSO – dimethyl sulfoxide

DNA – deoxyribonucleic acid

dNTPs – deoxynucleotides

DS, US – downstream and upstream regions of TUs, respectively

Ec (*E. coli*) – bacteria *Escherichia coli*; Msm (*M. smegmatis*) – bacteria *Mycobacterium smegmatis*; Mtb (*M. tuberculosis*) – bacteria *Mycobacterium tuberculosis*; Sp (*S. pneumoniae)* – bacteria *Streptococcus pneumoniae*

EMSA – electrophoretic mobility shift assay

EtBr – ethidium bromide

FE – fold enrichment

GCS – gyrase cleavage site

HETU, LETU – transcription units with high and low levels of expression, respectively: 200 TUs each group

IR – intergenic region

IPTG – isopropyl β-D-1-thiogalactopyranoside

kb, Mb – kilo- and megabase pairs, respectively

kDa – kiloDalton

LB – Luria-Bertani broth

LUCA – last universal common ancestor

MIC – minimal inhibitory concentration

MST – microscale thermophoresis

N3E, N5E – number of 3’ and 5’ read ends per genomic position, respectively

NGS – next generation sequencing

nt – nucleotide

OD – optical density

ORF – open reading frame

PAAG – polyacrylamide gel

PAGE – polyacrylamide gel electrophoresis

PBS – phosphate-buffered saline

PCR – polymerase chain reaction

PPM – position probability matrix

PWM – Position Weight Matrix

RBS – ribosome binding site

Rif – RNAP inhibitor rifampicin

RNA – ribonucleic acid

RNAP – RNA polymerase

RNase H – ribonuclease H

SDS – sodium dodecyl sulfate

SGS – strong gyrase binding/cleavage site

TAD – topologically associating domain

TCS – topoisomerase I cleavage site

TF – transcription factor

TFA – trifluoroacetic acid

TopoI – topoisomerase I

TSS – transcription start site

TU – transcription unit

v/v – volume/volume

w/v – weight/volume

There is something very strange and unaccountable about a tow-line. You roll it up with as much patience and care as you would take to fold up a new pair of trousers, and five minutes afterwards, when you pick it up, it is one ghastly, soul-revolting tangle. <…> I firmly believe that if you took an average tow-line, and stretched it out straight across the middle of a field, and then turned your back on it for thirty seconds, that, when you looked round again, you would find that it had got itself altogether in a heap in the middle of the field, and had twisted itself up, and tied itself into knots, and lost its two ends, and become all loops.

Jerome K Jerome, "Three Men in a Boat (to say nothing of the dog)"

Chapter 1. Literature Review

Literature Review presented here is based on a published review paper prepared by the author. All sections of the review were written by the author, except for the Topoisomerase VI section, which was prepared in a joint effort with Alina Galivondzhyan*. Text and some figures have undergone minor revisions to include discoveries that emerged after the publication.*

Sutormin D., Galivondzhyan A., Polkhovskiy A., Kamalyan S., Severinov K., Dubiley S. Diversity and functions of type II topoisomerases. **Acta Naturae**, Vol. 13(1), p.59-75, 2021

The DNA double helix provides a simple and elegant way to store and copy genetic information. However, processes requiring the DNA helix strands separation, such as transcription and replication, induce a topological side effect — supercoiling of the molecule. Topoisomerases comprise a specific group of enzymes that disentangle the topological challenges associated with DNA supercoiling. They relax DNA supercoils and resolve catenanes and knots. Here, we review the catalytic cycles, evolution, diversity, and functional roles of topoisomerases in prokaryotes.

1.1 DNA topology

The topological state of DNA and the level of its supercoiling are described using the linking number concept (Lk) [1]. If one thinks about one strand of a covalently closed circular DNA molecule as the edge of an imaginary surface, then the linking number of DNA strands is the number of intersections of this surface by the second DNA strand, with allowance for the sign of this intersection (**Figure 1A**). Lk does

not depend on molecule deformations and can only be changed by cleavage, passage, and religation of DNA strands (**Figure 1A**) [2].

$Lk^0 = N/h$ (1)

$\Delta Lk = Lk - Lk^0$ (2)

$Lk = Tw + Wr$ (3)

$\Delta Lk = \Delta Tw + \Delta Wr$ (4)

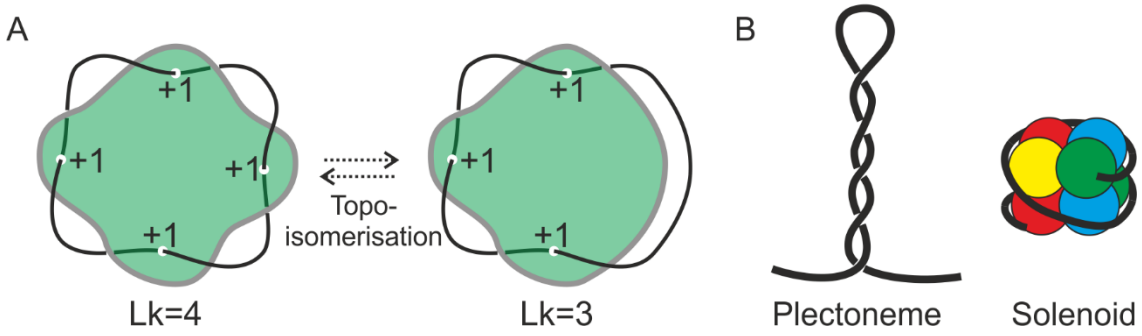


Figure 1. DNA topology. (**A**) Linking number of a circular DNA molecule and changes in the linking number resulting from strand cleavage and transfer. (**B**) Spatial structures, plectoneme and solenoid, arising from DNA supercoiling.

For a relaxed DNA molecule, the theoretical linking number (Lk^0) can be calculated as a ratio between DNA length in base pairs (N) and period of DNA (h = 10.5 bp/turn for the canonical B-form) (1). Lk of DNA molecules isolated from living organisms can either exceed Lk^0 ($\Delta Lk > 0$, a positively supercoiled molecule) or be less than Lk^0 ($\Delta Lk < 0$, a negatively supercoiled molecules) (2). Lk is the sum of two geometrical parameters of the double helix, called the twist (Tw) and the writhe (Wr) (3). The twist is defined as the number of times DNA chains turn around each other along the double helix axis, while the writhe is a measure of supercoiling of the DNA axis [3]. When Lk is different from Lk^0, supercoiling is partitioned between the twist and writhe (4), which can interconvert to each other. For example, according to electron microscopy of plasmids, the writhe and twist account for 75% and 25% of DNA supercoiling, respectively [3]. In nature, supercoiled DNA in the form of writhe stably exists in two forms: plectoneme (a higher order double helix, unconstrained writhe) and a solenoid (a higher order single helix, which is typical of DNA wrapped around a protein, constrained writhe) (**Figure 1B**). A more detailed and comprehensive discussion of DNA topology may be found, for example, in the book “DNA Topology” by Bates & Maxwell, 2005 [3].

1.2 Structure, evolution, and catalytic mechanism of type II topoisomerases

Special enzymes, topoisomerases, regulate the level of DNA supercoiling and resolve knots and catenanes [4,5]. According to their structure, homology, and catalytic mechanism, topoisomerases are usually divided into type I and type II [4]. Type I topoisomerases introduce a single-strand DNA break (nick) and alter the supercoiling state of a molecule either by rotating the DNA duplex around the intact second strand (class IB, change Lk of the molecule by an integer number per catalytic event) or by passing the intact strand through the nick (class IA, change Lk by ± 1 per catalytic event). Type II topoisomerases cleave both strands in a DNA fragment, termed the G-segment, and pass the second duplex, the T-segment, through this break, hydrolyzing at least one ATP molecule in the process (hydrolysis of two molecules is proposed, but has not been directly shown yet) (**Figure 3**) [6–8]. This is topologically equivalent to a change in Lk by ±2 [9]. Below, we will analyze the diversity, mechanisms, and physiological role of the enzymes with a focus on type II enzymes.

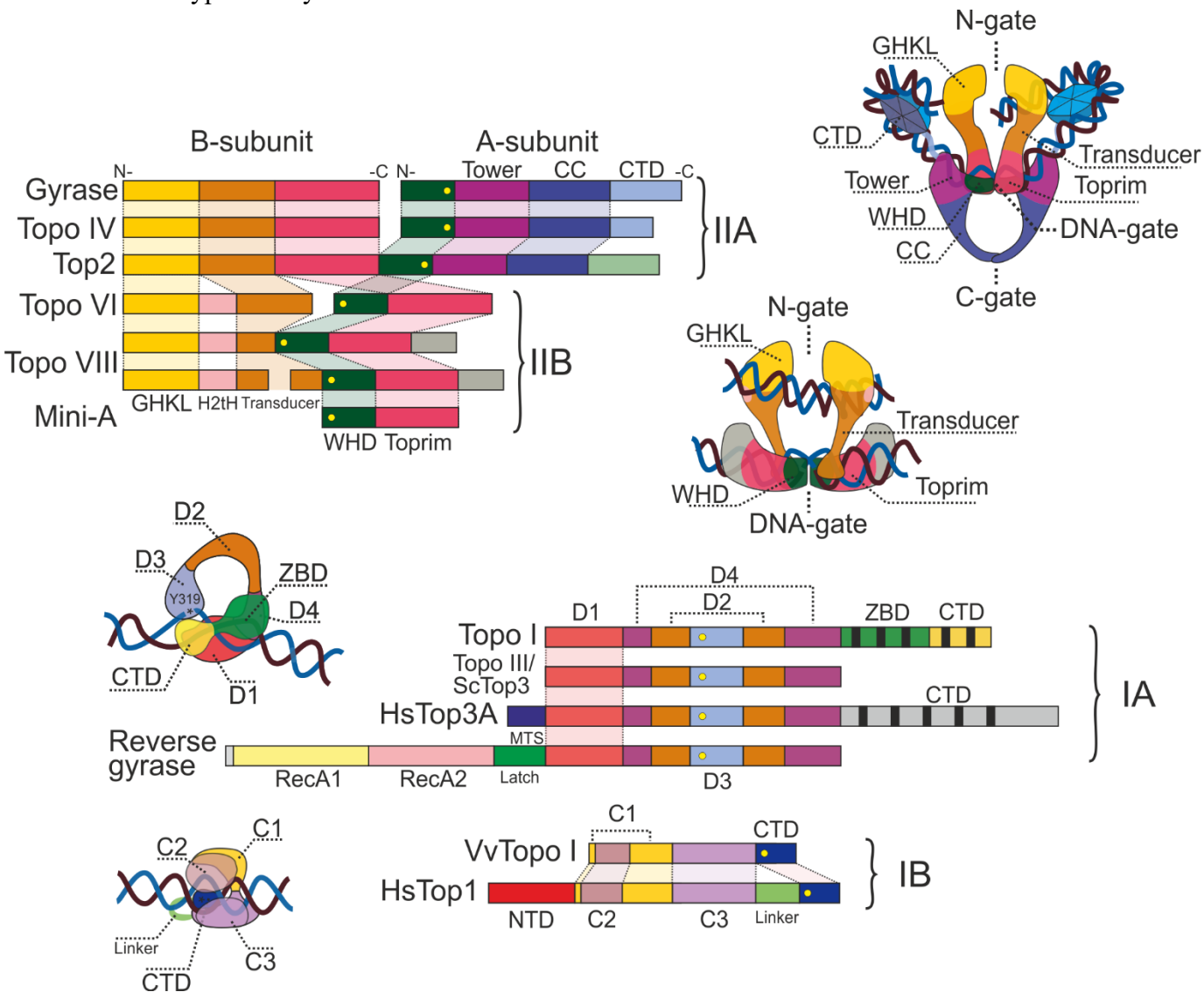


Figure 2. Types of topoisomerases and representative structures. Top – variants of the enzyme domain architecture (left) and domain organization of type IIA (DNA gyrase) and IIB (Topo VI) topoisomerases (right). Bottom – same for type IA and type IB topoisomerases. Homologous domains are shown by identical colors. The catalytic tyrosine residue responsible for DNA cleavage is depicted by a yellow circle.

Type II topoisomerases are found in organisms of all domains of life and are encoded in all but a few extremely reduced sequenced genomes of cellular organisms [10,11]. In all studied cases, type II topoisomerases have been shown to be necessary for transcription, replication, and segregation of chromosomes during cell division.

On the basis of their structure and catalytic cycle features, type II topoisomerases are subdivided into two classes: IIA and IIB (**Figures 2, 3**) [4]. Topoisomerases can be either heterotetramers consisting of two B and two A subunits or homodimers in which the B and A subunits are combined into a single polypeptide [10]. The topoisomerase subunits have dimerization interfaces, referred to as gates. The conserved ATP-hydrolysis GHKL (**G**yrase, **H**sp90, **H**istidine Kinase, **M**utL) domain [12] forms the **N-gate**, and the Toprim and WHD (**Topo**isomerase/Primase and **W**inged-**h**elix **d**omain) domains form the **DNA-gate** [13]. The G-segment of DNA binds to the DNA-gate region of the enzyme and is cleaved by active site tyrosyl residues of the WHD domains [14]. The third dimerization interface (**C-gate**), formed by the **c**oiled-**c**oil (CC) domain, is present only in type IIA enzymes (**Figure 2**) [15]. The **C**-**t**erminal **d**omains (CTD) are located either at the C-termini of A-subunits or at the end of fused polypeptides. The CTD determines the specificity of topoisomerases IIA to DNA structures (supercoils or crossovers), interacts with other proteins, and, in eukaryotes, is subject to post-translational modifications regulating activity of the enzyme [16–18].

At the first stage of the catalytic cycle, topoisomerase IIA is believed to bind the G-segment of DNA in the DNA-gate region [19]. The binding causes DNA bending, which is probably the basis of topological scanning of DNA by the enzyme: topoisomerase preferentially binds to supercoiled regions of the molecule that are either already bent or can be easily bent due to the energy of supercoiling [20–22]. Next, the T-segment of DNA is trapped between the GHKL domains and the DNA-gate. Binding of two ATP molecules to ATPase centers leads to dimerization of the GHKL domains, closure of the N-gate, and secure capture of the T-segment [23]. N-gate closure further enhances bending of G-segment, which together with hydrolysis of the first ATP molecule to ADP triggers cleavage of the G-segment by catalytic site tyrosyl residues of the WHDs and opens the DNA-gate, which results in the T-segment passage through the break to the protein cavity at the C-gate [7,13,24–26]. To stabilize the double-stranded break, hydroxyl groups of the tyrosyl

residues remain linked to the DNA 5'-ends by phosphodiester bonds. G-segment cleavage and opening of DNA-gate stabilize C-gate in a closed state, which is thought to further protect the transient DNA break. Closure of the DNA-gate and ligation of the G-segment leads to opening of the C-gate and releases the T-segment from the enzymatic complex [14]. C-gate opening is considered to be transient, probably, to secure the topoisomerase complex [27]. That is no consensus regarding the timing and the necessity of the second ATP molecule hydrolysis for catalysis [28]. The release of ADP molecules, which have low affinity for active centers, leads to the opening of the N-gate and transition of the enzyme to its original state (**Figure 3**) [23].

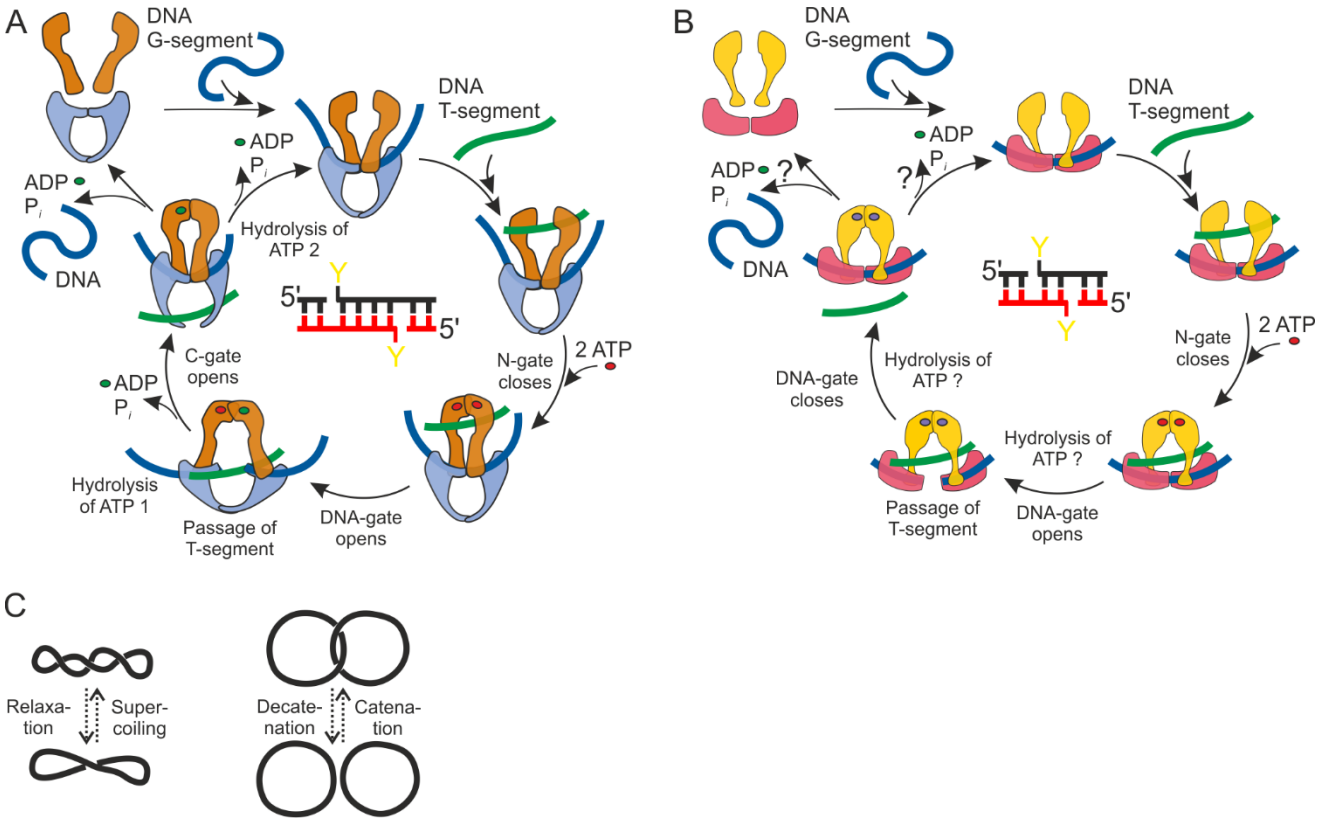


Figure 3. Catalytic cycles of topoisomerases IIA (**A**) and IIB (**B**) and the effect of topoisomerase activity on DNA topology (**C**). The scheme shows the following steps: binding of the DNA G-segment (blue) and T-segment (green); binding and hydrolysis of ATP molecules (ATP – red circle, ADP – green circle; if the bound nucleotide state (ATP/ADP) is unknown, it is depicted by a purple circle); cleavage and ligation of the G-segment and passage of the T-segment through the enzymatic complex. A scheme for G-segment cleavage is shown in the center of each cycle (Y – catalytic tyrosine residue of the WHD). Type II topoisomerases not only change DNA supercoiling but also unlink (decatenate) or link (catenate) DNA molecules.

Binding of ATP molecules is believed to be necessary for the unidirectional passage of the T-segment, since this segment is incapable of leaving the enzyme through the N-gate until both ATP molecules are hydrolyzed [24]. It should be noted that the role of ATP hydrolysis in segment passage has not been fully elucidated. According to one of the existing models, sequential hydrolysis of two ATP molecules promotes the T-segment passage by induced conformational rearrangements [28,29]. According to another model, the hydrolysis is required only for "restarting" the enzyme and trapping a new T-segment [30]. For example, in the presence of ADPNP, a non-hydrolyzable ATP analogue, topoisomerase is able to perform one act of T-segment passage, leaving the enzyme in an inactive state with closed N-gate [31]. According to recent single-molecule studies using magnetic tweezers, ATP hydrolysis is important both for accelerating the T-segment passage and for "restarting" the gyrase [7]. An alternative explanation considers ATP binding and GHKL domain dimerization as safeguards necessary to stabilize two halves of the enzymatic complex and to prevent the formation of double-strand breaks during T-segment transfer due to accidental dissociation of the two enzyme halves [8].

The catalytic mechanism of type IIB topoisomerases is considered to be similar to that of type IIA topoisomerases (**Figure 3B**) [32–34]. However, due to the absence of the C-gate, the T-segment immediately leaves the enzymatic complex after passing through the DNA-gate and the break in the G-segment [32]. In type IIB topoisomerases tyrosyl residues of WHDs are located on different secondary structure elements compared to the homologous domains of type IIA enzymes. When cleaving the G-segment of DNA, they generate two-nucleotide 5'-overhanging ends instead of four-nucleotide overhangs characteristic of type IIA topoisomerases [35,36]. The G-segment cleavage was shown to depend on ATP binding for IIB enzymes. This is considered necessary for stabilization of the complex and of the temporary double-stranded break [8,33].

The evolutionary relationships within type IIA and IIB topoisomerase groups and between these groups remain a matter of debate. Only a few evolutionary events can be reliably traced, for example, the duplication of a type IIA topoisomerase gene in the common ancestor of eubacteria which led to the emergence of DNA gyrase and topoisomerase IV (TopoIV), two enzymes with specific functions. Similarly, a duplication in the ancestor of vertebrates resulted in the emergence of Top2α and Top2β. Horizontal transfer of gyrase genes from different bacterial groups to Euryarchaeota and reverse transfer of Topo VI genes were also described. Bacterial gyrase found in Archaeplastida is likely inherited from chloroplasts during establishing of primary endosymbiosis [10] (**Figure 4**). There is no consensus for more ancient events of topoisomerase evolution.

1.3 Bacterial topoisomerases

Free-living fast-growing bacteria, such as *Escherichia coli*, *Caulobacter crescentus*, and *Bacillus subtilis*, usually possess multiple topoisomerases, including type I topoisomerases I and III as well as the type II, class IIA DNA gyrase and topoisomerase IV [4,37–39]. Slow-growing bacteria (e.g., *Mycobacterium tuberculosis*) or symbiotic/parasitic bacteria with reduced genomes (e.g., *Helicobacter pylori*), in contrast, often have the minimal essential set of one type I (topoisomerase I) and one type II (DNA gyrase) enzymes [40,41]. The genomes of several endosymbiotic bacteria, for example *Hodgkinia cicadicola* and *Tremblaya princeps* lack topoisomerase II genes or, like *Carsonella rudii*, encode only one subunit [42–44]. These organisms have extremely reduced (139–160 kb) genomes.

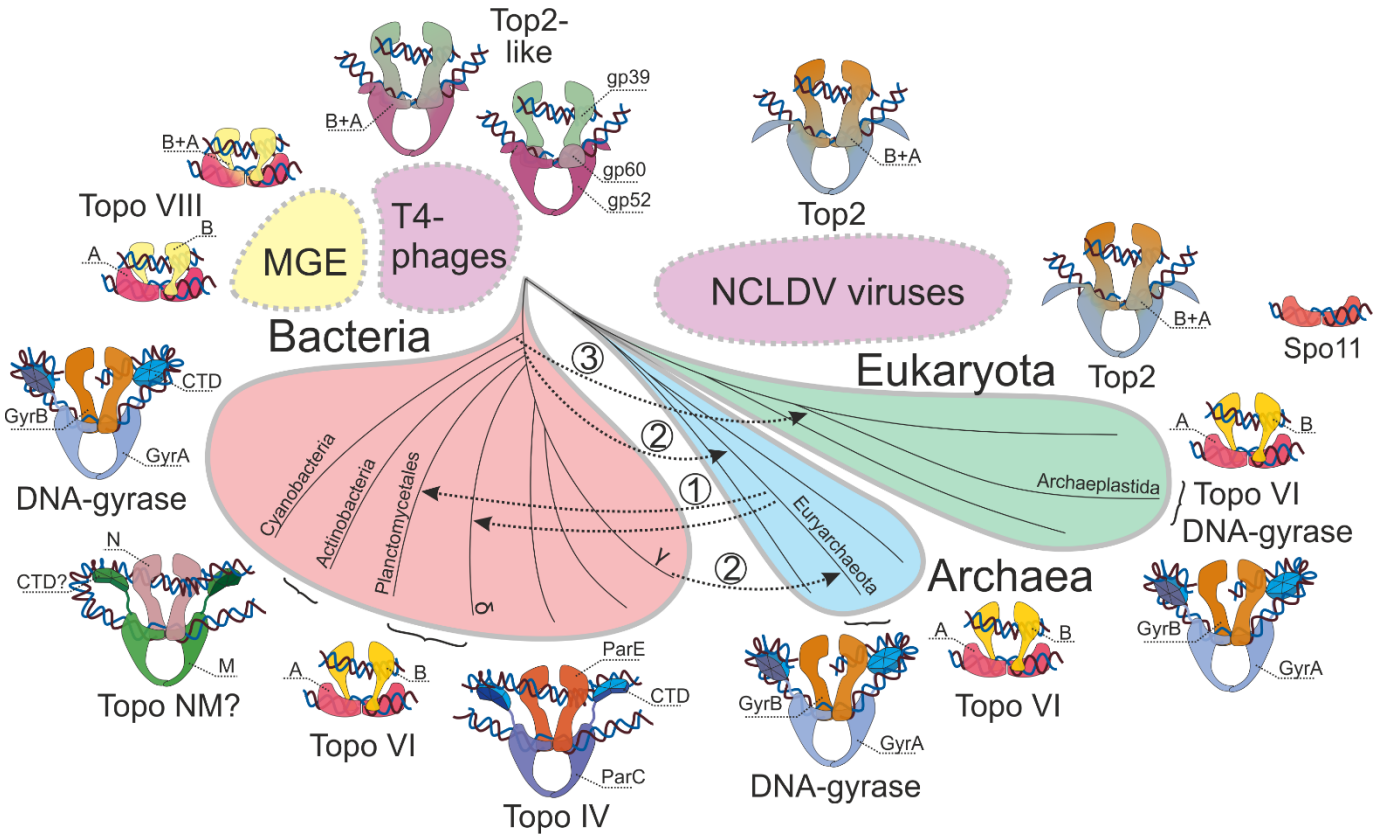


Figure 4. **Diversity of type II topoisomerases in different domains of life**. The tree of life is shown and the presence of type II topoisomerases is indicated. The following horizontal gene transfers are shown: 1 – a transfer of Topo VI genes from Euryarchaeota to Deltaproteobacteria and Planctomycetales bacteria; 2 – a transfer of gyrase genes from Bacteria to Euryarchaeota; 3 – a transfer of gyrase genes from Bacteria to Archaeplastida. MGE – mobile genetic elements.

The DNA gyrase and topoisomerase IV are the targets for many antibiotics that, according to their mechanism of action, may be divided into two groups: poisons and catalytic inhibitors. Poisons stabilize the

intermediate covalent complex of topoisomerase with the DNA G-segment. Accidental dissociation of enzyme subunits from such complex (for example induced by the collision with the replisome or RNA polymerase) causes double-stranded DNA breaks and leads to cell death if unrepaired. Catalytic inhibitors do not cause DNA breaks but inhibit the enzymatic activity, for example, by binding to the ATPase center of the GHKL domain and competing with ATP [45,46].

Quinolone and fluoroquinolone drugs (ciprofloxacin, levofloxacin, etc.), which are commonly used in clinical practice, are topoisomerase poisons [45,47]. Structural studies have shown that movement of divalent metal ions (most often magnesium) in the topoisomerase catalytic center is necessary for DNA cleavage and ligation. Gyrase poisons stabilize a metal ion in a position that promotes DNA cleavage, but not the sealing of the break [48,49]. The latter fact explains the effects of the most prevalent gyrase mutations leading to antibiotic resistance. The conserved serine and glutamine residues of the WHD were found to coordinate water molecules and magnesium ions necessary for the binding of fluoroquinolones [48]. Replacing at least one of these residues with a non-polar moiety leads to poison resistance [50]. Described drugs belongs to a wider group of 'interfacial inhibitors', which target macromolecular interfaces and typically stabilize normally transient intermediates [51,52]. According to this point of view, quinolone and fluoroquinolone drugs sterically prevent ligation of DNA break by intercalating between DNA nucleotides to be sealed with a phosphodiester bond.

Classical catalytic inhibitors are aminocoumarin compounds (novobiocin and coumermycin A1). They compete with ATP for the interaction with the ATPase center [45,53]. Inhibition of gyrase activity leads to inhibition of replication and transcription and cell division arrest. Due to low solubility and toxicity to humans, aminocoumarins are not used in the clinical practice, but found application in veterinary medicine [46].

The spread of antibiotic resistance necessitates searches for new antibacterial drugs; several new classes of topoisomerase inhibitors are currently in clinical trials [46,54,55].

1.3.1 DNA gyrase

Bacterial DNA gyrases are conserved enzymes (**Figure 5A**) able to induce negative supercoiling using the energy of ATP hydrolysis, as demonstrated in *in vitro* experiments for enzymes from *E. coli*, *B. subtilis*, *C. crescentus*, *M. tuberculosis*, and many other bacteria. In addition, DNA gyrases effectively relaxe positive supercoils and are capable of decatenating circular DNA molecules [40,56–59]. The *gyrA* and *gyrB* genes encoding the enzyme subunits are essential and inhibitors that reduce gyrase activity

significantly decrease cell viability [60–63]. Gyrase inhibition induces a similar phenotype in different bacteria: elongated cells incapable of dividing [63,64].

Gyrase maintains negative supercoiling of the genome, facilitating the initiation of transcription and replication. It also relaxes positive supercoils in front of elongating polymerases. Early ChIP-chip (immunoprecipitation of protein-bound DNA and its subsequent analysis on a chip to determine protein binding sites) experiments with *E. coli* revealed a positive correlation between gyrase binding and transcription level [65]. Similarly, the results of ChIP-Seq (immunoprecipitation of protein-bound DNA and its subsequent sequencing to determine protein binding sites) experiments with *M. tuberculosis* gyrase indicate preferential binding of the enzyme to transcriptionally active regions [66]. In *C. crescentus*, suppression of the *gapR* gene expression inhibits initiation and elongation of replication and increases sensitivity of cells to gyrase inhibitors. *In vitro* experiments have shown that the GapR protein preferentially binds to positively supercoiled DNA and interacts with the gyrase, increasing its ability to relax positive supercoils. Probably, GapR recruits the gyrase to positive supercoils formed in front of the moving replication complex, facilitating their relaxation and thus stimulating replication [58]. Single-molecule experiments have shown that in the absence of gyrase transcription on topologically constrained DNA molecules quickly slows down and eventually stops due to accumulation of positive supercoiling (**Figure 5B**). The binding of gyrase to such molecules results in rapid restoration of the normal transcription rate (transcriptional burst) [67].

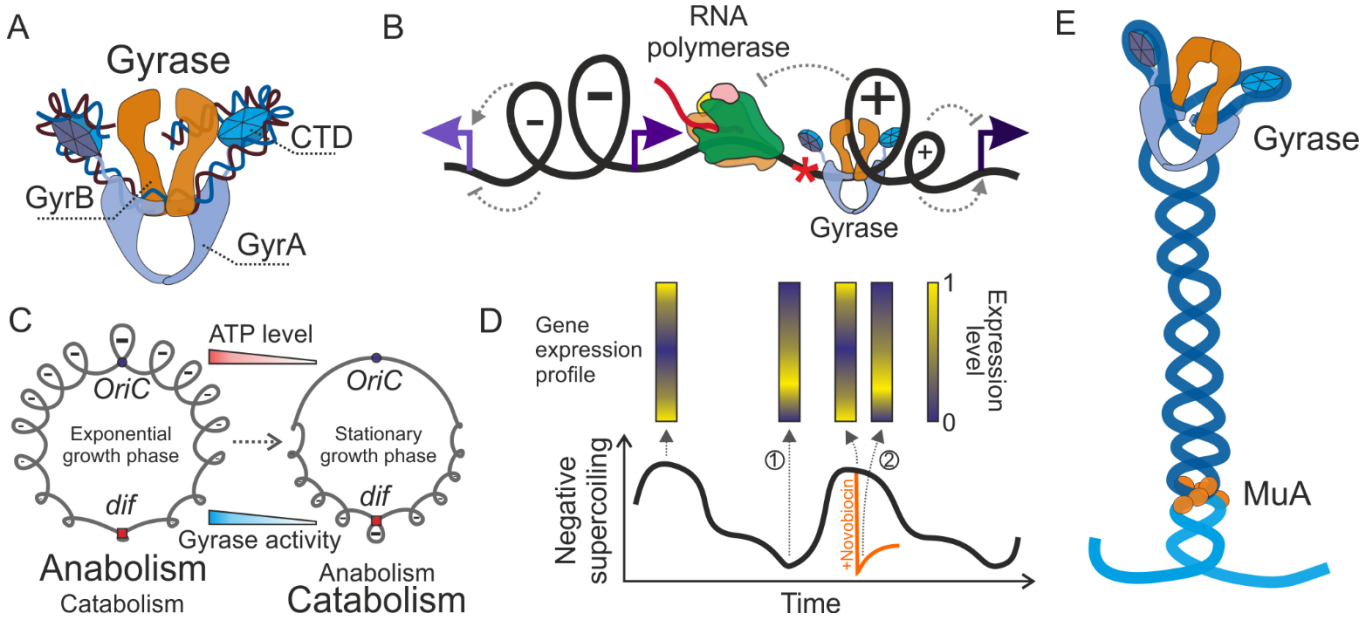


Figure 5. **DNA gyrase and its functions**. (**A**) Structure of a DNA gyrase complex with DNA. (**B**) a twin-domain model illustrating positive supercoiling upstream of the elongating RNA polymerase and negative

supercoiling downstream of it [68]. Co-transcriptional positive and negative supercoiling moves along the DNA molecule and influences initiation of transcription from adjacent promoters (indicated by arrows). Depending on the promoter, the effect can be either activating or inhibiting. The DNA gyrase promotes transcription elongation through relaxation of positive supercoiling ahead of RNA polymerase. (**C**) Changes in genome supercoiling during *E. coli* culture transition from the exponential to stationary growth phase promote switching of the cell from a mainly anabolic to catabolic physiological state [69]. *OriC* – origin of replication, *dif* – site recognized by XerC/XerD recombinases. (**D**) Circadian oscillations of the *Synechococcus elongatus* genome supercoiling level (at the bottom) correlate with changes in the gene transcriptional profile (at the top). A sharp decrease in the genome supercoiling level (indicated by the orange arrow) in the presence of the DNA gyrase inhibitor novobiocin causes a rapid change in the transcriptional profile (2), making it similar to the profile of bacteria in the physiologically relaxed genome state (1) [70]. (**E**) DNA gyrase is essential for the spatial organization of the Mu prophage and its transposition. The prophage DNA is shown in dark blue, and the bacterial genome DNA is in light blue.

In addition to its ability to relax positive supercoiling in front of elongating RNA polymerases, by introduction of negative supercoiling, the gyrase can both activate and suppress transcription initiation [71]. Up to half of *E. coli* genes were found to respond to genome relaxation by changing their transcription level [72,73]. Ontological analysis of *E. coli* genes sensitive to supercoiling revealed that the products of genes responding to relaxation of negative supercoils by increasing their transcription level are preferentially involved in catabolic reactions (for example, Krebs cycle enzymes). These genes are located closer to the terminus of replication. In contrast, genes that require negative supercoiling for initiation of their transcription are predominantly associated with anabolic processes (synthesis of amino acids and nucleotides) and are located closer to the region of replication origin [73,74]. According to a model, during active growth of *E. coli* culture, the activity of DNA gyrase generates a negative supercoiling gradient in the genome, with the maximal and minimal levels being in the replication origin and the terminus regions, respectively. This leads to predominant expression of genes involved in the anabolic process, promoting cell growth and division. Depletion of nutrients in the stationary phase decreases the ATP concentration, which reduces the DNA gyrase activity. This decreases the genome supercoiling level and, in combination with other factors, inverts the gradient of chromosome supercoiling, resulting in predominant expression of genes involved in catabolic processes [69]. It was hypothesized that *E. coli* uses supercoiling to globally modulate gene transcription upon starvation [74–76] (**Figure 5C**).

Promoters of the *E. coli gyrA*, *gyrB*, and *topA* genes that encode gyrase and topoisomerase I subunits are highly sensitive to supercoiling. The *gyrA* and *gyrB* transcription is activated upon genome relaxation,

while *topA* is better transcribed upon enhancement of negative supercoiling [77,78]. This enables mutually regulated synthesis of two topoisomerases with opposite activities and provides a homeostat for genome-wide supercoiling level control [79,80]. Similar mechanisms are operational in *S. coelicolor* and *C. crescentus* [61,81].

The supercoiling level in *Salmonella typhimurium* is believed to regulate the transition from anaerobic metabolism to aerobic respiration [82]. In *H. pylori*, negative supercoiling is an important regulator of flagellar synthesis [83]. Circadian oscillations of DNA supercoiling in the cyanobacterium *Synechococcus elongatus* correlate with specific changes in gene transcription and relaxation of negative supercoiling by the addition of DNA gyrase inhibitor novobiocin leads to a rapid change in gene transcription pattern, mimicking the changes observed during the circadian cycle (**Figure 5D**) [70]. Overall, these data allow one to consider supercoiling as a global transcription factor and show that the structure of regulatory regions has evolved to allow specific responses to this factor [69,71,74].

A number of studies indicated that gyrase and gyrase-induced negative supercoiling are involved in the spatial organization of bacterial genomes. For example, *in vivo* fluoroquinolone induces cleavage of *E. coli* genomic DNA by the gyrase into 50–100 kb fragments, which roughly corresponds to the length of supercoiled chromosome domains [84–86]. Activity of DNA gyrase at a high-affinity site located at the center of the bacteriophage Mu prophage was shown to cause a local increase in negative supercoiling, leading to plectonemic compaction of the chromosome region with the prophage. This brings prophage termini into proximity with each other and promotes their recombination by the MuA transposase [87,88] (**Figure 5E**). Similarly, excessive negative supercoiling accumulated in *E. coli* cells with a mutation in topoisomerase I is believed to lead to chromosome compaction [89]. As shown by Hi-C experiments (a method for determining the chromosome conformation) in *C. crescentus*, gyrase inhibition by novobiocin, on the contrary, makes the spatial structure of the chromosome more diffuse – it lowers the frequency of short-range contacts (20-200 kb) and blurs borders of topologically associated domains (TADs) [90]. Similar results were recently obtained for *E. coli* [91]. Further research is needed to elucidate the role of supercoiling in the regulation of spatial organization of prokaryotic genomes.

Gyrase genes have been found in members of several Euryarchaeota groups [11] (**Figure 4**). Like bacterial gyrase, the archaeal enzyme is sensitive to coumarins and quinolones [92–94]. *In vitro* experiments have shown that *Thermoplasma acidophilum* gyrase has a typical spectrum of activities: it relaxes positive supercoils, introduces negative supercoils, and decatenates circular DNA molecules [94]. Inhibition of gyrase activity by the addition of novobiocin to *Halobacterium halobium* cells leads to inhibition of DNA replication and significant decrease in levels of transcription and translation [92]. Thus, the archaeal gyrases

are believed to perform functions typical of bacterial homologues: relaxation of positive supercoils formed during transcription and replication as well as decatenation of linked DNA molecules during cell division.

1.3.2 Topoisomerase IV

In vitro experiments have demonstrated that despite their structural similarity TopoIV and gyrases have different spectra of activities. TopoIV is able to effectively relax positive supercoils. Negative supercoils are relaxed at a much slower rate. Unlike the gyrase, TopoIV cannot introduce excessive negative supercoiling [58,59,95]. At the same time, TopoIV is an efficient decatenase that separates interlinked circular DNA molecules much better than the gyrase [96–100]. Accordingly, TopoIV but not gyrase is capable of resolving knotted DNA molecules *in vivo* [101]. It is hypothesized that these differences are related to the structures of CTD domains in the GyrA subunit and in the homologous ParC subunit of TopoIV (**Figure 6B**). The gyrase CTD enables wrapping of DNA around the enzyme, such that DNA located *in cis* and close to the G-segment of DNA serves as a T-segment, which allows the introduction of negative supercoils in one DNA molecule [7,102]. The TopoIV CTD does not bend the G-segment; instead, it traps as a T-segment remote DNA sites or *in trans* DNA molecules. Since the T-segment must be perpendicular to the enzyme bound G-segment, catenanes are effectively recognized and resolved [95,99,103] (**Figures 5A**, **6A**). A comprehensive discussion of mechanistic and structural differences between the two topoisomerases can be found in a recent review [104].

Like the gyrase, TopoIV is necessary for bacterial division. Mutations in the *parC* and *parE* TopoIV subunits genes or inhibition of the enzyme activity by drugs causes the development of the so-called *par* phenotype in different bacteria. The *par* phenotype is characterized by elongated cells that are not capable of division and contain an increased amount of unsegregated DNA [37,105–108]. However, the lack of TopoIV activity does not interfere with chromosome replication and termination in *E. coli* [106,107]. The biochemical properties of the enzyme suggest that the main function of TopoIV in the cell is to resolve pre-catenanes during replication (intersections between sister DNA molecules arising from replisome rotation) and to separate catenanes of circular molecules upon the completion of replication [107,109]. According to this hypothesis, TopoIV is not essential for *Streptomyces* with its linear chromosomes but is important for the maintenance of circular plasmids [39]. Yet, *E. coli* cells with artificial linear chromosomes exhibit the *par* phenotype upon TopoIV inactivation. This may indicate the importance of the early removal of pre-catenanes and knots along the entire length of the replicating chromosome [110]. An increase in the TopoIV expression level leads to accelerated DNA segregation during division of *E. coli* cells [107].

DNA gyrase and type I topoisomerases are supposed to substitute TopoIV for decatenation in bacteria lacking this topoisomerase. For example, the *M. tuberculosis* gyrase is an efficient decatenase. ChIP-Seq experiments demonstrated that mycobacterial gyrase (in *M. tuberculosis*) and Topoisomerase I (in *M. smegmatis*) are significantly enriched at the chromosomal replication terminus region, which suggests that they substitute TopoIV [40,66,111,112]. However, no such enrichment was observed for the *E. coli* gyrase [113]. The involvement of *H. pylori* gyrase in chromosome segregation is indirectly confirmed by the fact that bacteria with deletion of the *xerH* gene, which encodes recombinase involved in resolution of chromosome dimers and, possibly, decatenation, are more sensitive to the gyrase inhibitor ciprofloxacin [106,114].

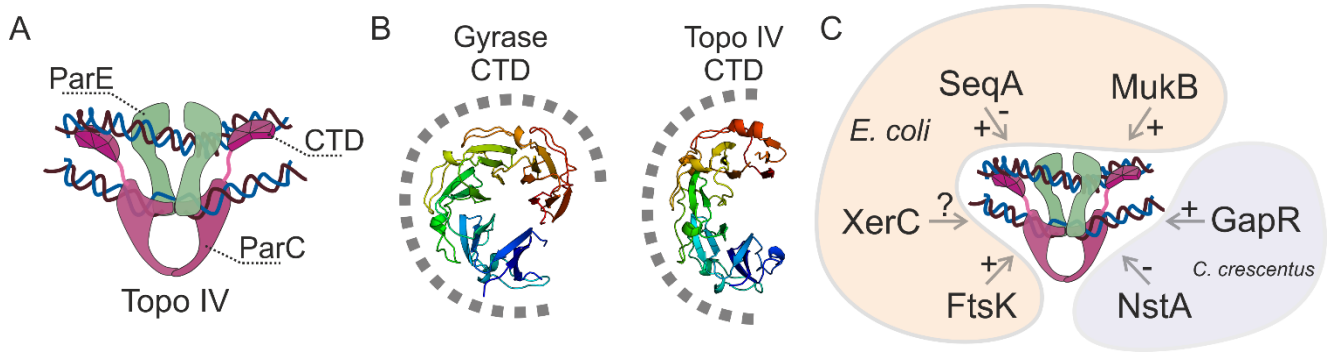


Figure 6. **Topoisomerase IV structure and interactions with other proteins**. (**A**) Hypothetical structure of the TopoIV complex with DNA. (**B**) Comparison of the GyrA CTD (PDB ID: **1zi0**) and TopoIV ParC CTD (PDB ID: **1zvt**) structures. A putative position of DNA is shown as a dashed line. (**C**) Proteins interacting with TopoIV. The effect of each protein on TopoIV activity is depicted as "+" (activation), "–" (inhibition), or "?" (interaction is not confirmed).

The ability of TopoIV to relax positive supercoils [59,95], suggests that it may cooperate with the DNA gyrase in removal of positive supercoils formed during transcription and replication [58,115] (**Figure 7A**). For example, treatment of *E. coli* cells with the RNA polymerase inhibitor rifampicin was found to reduce both gyrase and TopoIV activities, at least in some regions of the genome [85,116]. Interestingly, an increase in the copy number of *parC* and *parE* genes is a common suppressor mutation associated with deletion of topoisomerase I gene in *E. coli* and *B. subtilis*. In this case, TopoIV is believed to compensate for the loss of topoisomerase I and perform its function by removing negative supercoiling [38,105,117].

TopoIV interacts with a number of proteins that have completely different functions and structures but are involved in the organization and separation of replicated chromosomes. In *E. coli*, these are the SeqA protein that binds to hemimethylated GATC sites behind the moving replisome [118,119], the MukBEF

SMC (structural maintenance of chromosome) [120,121], the DNA translocase FtsK [122], and, probably, the XerC recombinase [116,123] (**Figure 6C**). The *C. crescentus* TopoIV interacts with GapR and NstA. These proteins have opposite effects on the enzyme - GapR stimulates enzyme activity, while NstA suppresses it [58,124].

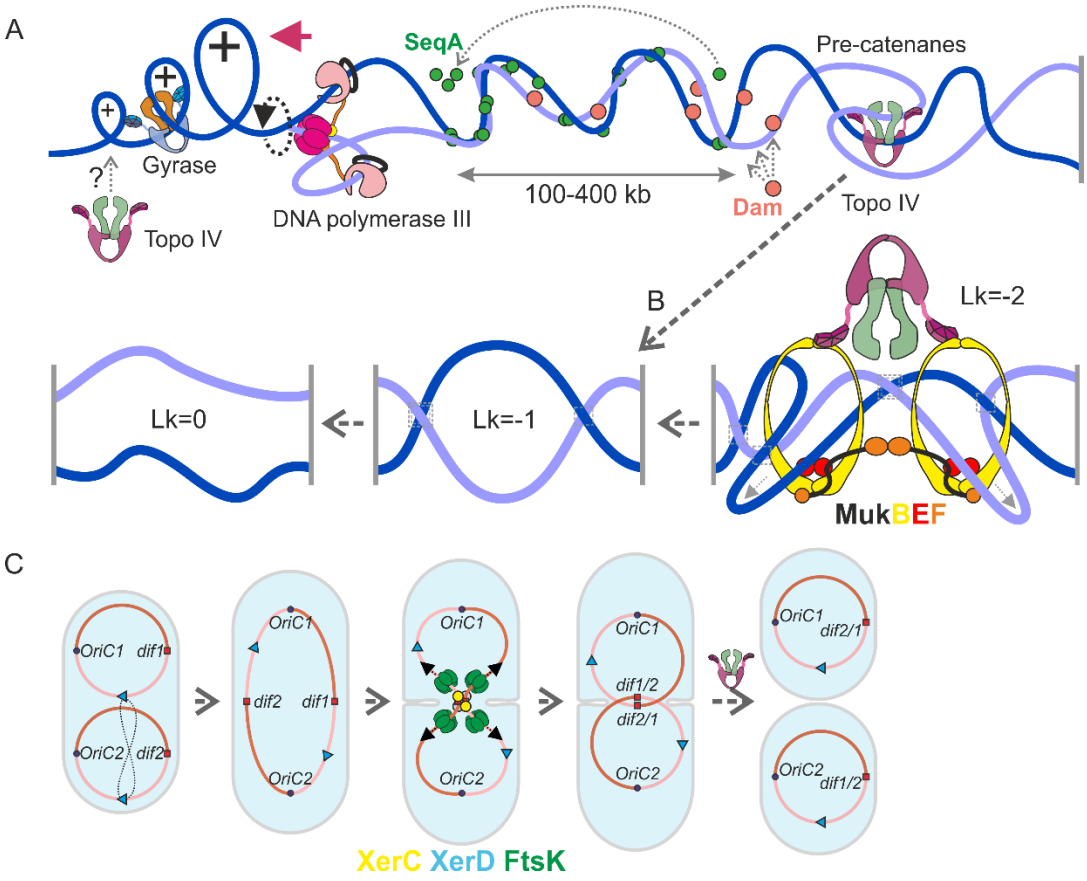


Figure 7. Topoisomerase IV role in decatenation. (**A**) Topological effects associated with DNA replication. Positive supercoils formed in front of the moving replisome are relaxed by the DNA gyrase and, presumably, TopoIV. Accumulation of DNA supercoiling leads to replisome rotation, thereby producing pre-catenanes. In *E. coli*, the SeqA protein binds to hemimethylated GATC sites of newly replicated DNA molecules. Dam methylates the GATC sites and displaces SeqA, so the SeqA concentration gradient extends 100–400 kb behind the replisome and moves together with it. TopoIV cannot interact with SeqA-bound DNA regions, which explains the temporary cohesion of daughter chromosomes during replication in *E. coli*; however, when all GATC sites are methylated and SeqA is no longer associated with DNA, topoisomerase removes pre-catenanes, enabling daughter chromosome separation [119]. (**B**) The proposed structure of the MukBEF-TopoIV complex and a hypothetical mechanism of selective resolution of precatenates. (**C**) A model of chromosome dimer resolution by the FtsK-XerC/D-TopoIV ensemble.

In vitro, low concentrations of *E. coli* SeqA stimulate relaxation and decatenation activities of TopoIV and have no effects on DNA gyrase and Topo I, while in higher concentrations it inhibits TopoIV activities. Authors suggest that SeqA can attract TopoIV to the newly replicated DNA downstream of replication fork to resolve pre-catenanes [118]. On the other hand, by *in vivo* microscopy it was demonstrated that SeqA may inhibit TopoIV activity on freshly replicated DNA and mediates temporary cohesion of these loci. Especially long cohesion is observed at specific regions known as snap-loci where concentration of SeqA is increased (**Figure 7A**). SeqA depletion impairs proper segregation of replicated chromosomes, which indicates that temporary cohesion is needed for this process [119]. To investigate how SeqA modulates TopoIV binding and activity on freshly replicated DNA *in vivo*, a time-series of TopoIV ChIP-Seq and Topo-Seq experiments can be conducted on synchronously replicating cells. If SeqA impairs TopoIV binding to DNA, there will be a depletion of TopoIV enrichment observed with both methods for 100-400 kb downstream of a replisome (where SeqA is bound [119,125]), in contrast, if TopoIV binding is not affected but TopoIV activity is inhibited, only Topo-Seq signal will be decreased.

SMC (structural maintenance of chromosomes) complex MukBEF is essential for *E. coli* rapid growth. Depletion of any gene of the complex results in a temperature-sensitive phenotype; the mutants do not grow at 37ºC in a rich medium, but survive at 25ºC in a minimal medium [89]. It was demonstrated that ParC subunit of TopoIV interacts by CTD with MukBEF complex [120,121,126]. Initial results demonstrated that in excess MukB stimulates the relaxation and decatenation activities of TopoIV *in vitro* [120,121]. However, recent biochemical experiments with a equimolar ratio between TopoIV and MukBEF (which recapitulates conditions in a cell) indicate that within the TopoIV-MukBEF complex catalytic activities of enzymes are mutually suppressed, presumably to stimulate DNA condensation by the super-complex [127]. *In vivo*, TopoIV and the *E. coli* SMC complex MukBEF form clusters consisting of ~15 topoisomerase molecules and ~16 SMC complexes [128,129]. These clusters colocalize with replication origins, determine their position in the cell, and are necessary for segregation of the origins of daughter chromosomes during division [128,130,131]. The proposed structure and functions of MukBEF-TopoIV super-complex are shown in **Figure 7B**. In *C. crescentus* TopoIV is also required for the correct movement of one of the origins to the opposite cell pole [108], however, *C. crescentus* lacks MukBEF and has a different type of SMC complex [132].

Ter region of *E. coli* chromosome (known as Ter macrodomain) is defined by a MatP protein that interacts with two dozen of *matS* sites inside the region. Depletion of MatP leads to *par*-phenotype, also in the mutant cells segregation of Ter regions is faster [133,134]. MatP is likely a master organizer of Ter macrodomain and regulates the final stages of chromosome segregation. First, MatP was shown to interact

with MukBEF and stimulates its unloading from DNA in Ter [134–136]. Second, MatP and TopoIV compete for binding with MukBEF [137]. Therefore, MatP displaces both complexes from Ter, presumably to stimulate the turnover of the complexes and increase the pool of enzymes for organization of the origin regions.

At the center of Ter region, a *dif* site is located, which is recognized by XerC/XerD recombinase [138]. The recombinase is needed to resolve chromosome dimers and FtsK DNA translocase, that interacts with the recombinase complex, facilitates the process of resolution [139,140]. According to the model, TopoIV within the recombinase complex is needed to release the catenated circular molecules upon recombination (**Figure 7C**).

Recent chromosome conformation capture experiments in *E. coli* demonstrated that inactivation of TopoIV activity results in the increase of mid-range contacts in all chromosome macrodomains except the Ter macrodomain, where the decrease in the contacts frequency was observed. A hallmarks of TopoIV inactivation were long-range contacts between *dif* region and other chromosomal regions (so called "butterfly wings" structure) lasting for the entire chromosome. Authors proposed that this structure reflects unrelaxed catenation of sister chromosomes in the absence of the topoisomerase [91]. I speculate, that "butterfly wings" reflects the threading of chromosomal DNA by FtsK translocase to resolve chromosome dimers. Continuous DNA translocation explains why these contacts do not follow the genomic distance law. If so, translocation by FtsK might be limited to the Ter macrodomain in *wild-type* cells, which is supported by earlier observations [141,142]. Probably, FtsK activity is required at the last stages of chromosome dimers resolution, when the majority of pre-catenanes is already resolved by TopoIV outside of Ter region.

1.3.3 Topoisomerase NM

A unique type II topoisomerase, called TopoNM, was discovered in *M. smegmatis* [143]. It consists of two subunits (TopoN and TopoM) homologous to the ParE/GyrB and ParC/GyrA subunits of topoisomerase IV and gyrase, respectively. According to phylogenetic analysis of amino acid sequences, TopoNM is distant from all known type IIA topoisomerases, which indicates early divergence of enzyme genes. The significant divergence from other topoisomerases II and the absence of TopoNM in other, even related, bacteria may indicate the viral origin of the enzyme. TopoNM has reduced sensitivity to fluoroquinolones and coumarins. The enzyme relaxes positive and negative supercoils and decatenates circular DNA molecules, which is typical of type II topoisomerases. A unique property of TopoNM is the ability to introduce positive supercoils into relaxed plasmids [143]. Besides TopoNM, only reverse gyrase

– a type I topoisomerase – is capable of introducing positive supercoils using the energy of ATP hydrolysis [144]. Neither the mechanism of positive supercoiling by TopoNM nore the functions of this enzyme are known.

An unusual system for protection against mobile genetic elements was found in *M. smegmatis*. It consists of genes encoding a cohesin-like complex that prevents effective transformation of bacteria with plasmids [145,146]. TopoNM may be part of this defense system like some bacterial topoisomerases interact with cohesins [120,121,147].

1.3.4 Topoisomerase VI and other IIB topoisomerases

Topoisomerase VI (Topo VI) is a heterotetrameric topoisomerase was first found in the hyperthermophilic archaeon *Sulfolobus shibatae* [148] and later in most other archaea, except for some members of the Thermoplasmatales group in which it is replaced by the DNA gyrase [11]. *In vitro*, archaeal Topo VI can relax both positive (more efficiently) and negative supercoils and exhibits a decatenation activity [33,149,150]. The letter activity was recently shown preferential by biochemical and single-molecule experiments [150]. Sequence similarity between IIA and IIB topoisomerases is rather low. Additionally, the catalytic tyrosine residues of WHDs are located on non-homologous secondary structure elements in the two groups [33,34,151] (**Figure 2**). Despite these, the catalytic mechanism of Topo VI is supposed to be similar to that of type IIA topoisomerases, which is based on biochemical and structural analyses (**Figure 3B**).

The physiological role of Topo VI has not been established. Activities of the enzyme demonstrated *in vitro* and the fact that Topo VI can be replaced with DNA gyrase indicate that the topoisomerase may be involved in decatenation of replicated chromosomes and in relaxation of supercoils formed during transcription and replication [149]. The expression level of Topo VI in the archaeon *S. islandicus* was found to increase 7 h after elevating the cultivation temperature above the optimal one. Probably, Topo VI compensates for an increase in reverse gyrase activity under these conditions [152].

Besides Topo VI, several other sub-families inside type IIB topoisomerases were recently identified. Genes of topoisomerases with predicted domains similar to the Topo VI domains are found in some archaeal and bacterial plasmids as well as in integrated mobile genetic elements. Typically, such genes encode a fusion of proteins similar to TopoIV topoisomerase subunits B and A. The topoisomerases encoded by these genes are allocated into a separate group of type IIB topoisomerases and are referred to as “Topo VIII” [153,154]. Several Topo VIII were shown to relax supercoiled plasmids and decatenate circular DNA

molecules *in vitro* [153]. Recently, a new group of proteins homologous to the A-subunit of Topo VIII was identified; they are called Mini-A due to their relatively small size (**Figure 2**) [154]. The function of these topoisomerases is unknown. Probably, they help to maintain plasmids and promote their propagation in host cells.

1.3.5 Topoisomerase I

Topoisomerase I of *Escherichia coli* (EcTopoI, encoded by the *topA* gene) belongs to the A class of type I topoisomerases [4]. EcTopoI relaxes only negatively supercoiled DNA and is thought to maintain the steady-state level of supercoiling by compensating the activity of another topoisomerase – the DNA gyrase, which introduces negative supercoiling utilizing the energy of ATP hydrolysis (Menzel and Gellert, 1978; Tse-Dinh, 1985; Liu *et al.*, 2017). Deletion of *topA* leads to rapid accumulation of suppressor mutations, mostly in genes encoding the DNA gyrase subunits. By reducing gyrase activity, these mutations balance the level of DNA supercoiling inside the cell [158,159]. Amplification of a chromosomal region containing the *parC* and *parE* genes encoding TopoIV is also frequently reported in *topA* null mutants [117,160,161]. Conversely, *topA* deletions complement growth and replication defects of temperature-sensitive (Ts) *gyrB* mutants at non-permissive temperatures [162].

Hypernegative supercoiling (a level of negative supercoiling which is much higher than in wild type cells) is a hallmark of *topA* mutants resulting from uncompensated gyrase activity [163]. It was observed that hypernegative supercoiling is interconnected with formation and stabilization of R-loops - RNA-DNA heteroduplexes formed when nascent transcripts anneal to the template DNA strand upstream of the transcribing RNA polymerase (RNAP) [163–165]. First, R-loops have been recently detected in *topA* mutants (exhibiting hypernegative supercoiling) by dot-blots assays with an RNA:DNA hybrid-specific antibody [166]. Second, R-loops formation during *in vitro* transcription was found to be dependent on negative supercoiling of a template plasmid and it was increased for hypernegatively supercoiled substrate [163,165]. Reciprocally, formation of a hypernegatively supercoiled plasmid form was reproduced during *in vitro* transcription in the presence of gyrase and it was dependent on R-loops formation, as treatment with RNase A or RNase HI (specifically removes R-loops) had a suppressive effect [163,164,167]. To explain the interdependence of hyernegatively supercoiling and R-loops formation in cells lacking topoisomerase I, an autocatalytic model was proposed. According to the model, elongating RNAP triggers this process, as it introduces negative supercoiling upstream and positive supercoiling downstream (a substrate for gyrase) and produces nascent RNA which can form an R-loop. Gyrase converts positive supercoiling into negative, and this promotes nucleation of an R-loop. Since one DNA strand is unpaired in the R-loop, a local hub

accumulating excessive negative supercoiling is created, masking negative supercoiling and leaving the nearby DNA relaxed, which was observed *in vitro* using transcription assay [165]. Same time, RNAP continues elongation and produces more negative and positive supercoiling, the latter of which is relaxed by gyrase. Excessive negative supercoiling leads to R-loops propagation and absorption of more negative supercoiling, therefore preparing a substrate for gyrase together with RNAP. A positive feedback loop is created, leading to further accumulation of R-loops and additional negative supercoiling (negative supercoiling → R-loops propagation → absorbtion of negative supercoiling → increase gyrase activity → excessive negative supercoiling) (**Figure 8**) [168]. Actually, in the presence of nascent transcripts as a source of RNA component of R-loops, formation of hypernegatively supercoiled plasmid requires only gyrase activity and can be observed without active transcription [167]. Should be noted that the outlined model and connected body of data indicate that in normal cells R-loops may play a role in supercoiling management as efficient supercoiling accumulators [169].

As a consequence of hypernegative supercoiling in *topA* mutants, degradation of nascent transcripts was observed likely due to RNase HI activity, which leads to rapid growth arrest [163]. Overexpression of RNase HI, an enzyme which degrades RNA in the R-loops [170,171] and should thus break the feedback loop, was reported to partially suppress the negative effects of a *topA* deletion [172,173], although this finding was disputed [174], while deletion of the RNase HI *rnhA* gene exacerbates the *topA* null phenotype [174,175]. Observed positive effect of RNase HI activity is contradictory considering its plausible role in degradation of nascent transcripts in these mutants (see above). Probably, when overexpressed, RNase HI prevents from stabilization of long R-loops breaking the positive feedback loop at the beginning and allows normal transcription by sacrificing some portion of transcripts. Stabilized R-loops can also prime *oriC*-independent replication – a phenomenon called "constitutive stable DNA replication" (cSDR) initially observed in cells lacking RNase HI. It was demonstrated that cells lacking type-I topoisomerases also exhibit cSDR, which is suppressed by overexpression of RNase HI [117,176]. Together, these data indicate that hypernegative supercoiling is the likely cause of severe growth defects of non-suppressed *topA* mutants [177]. Recently, by analyzing the viability of *topA* mutants in different genetic backgrounds, it was proposed

that bacterial lethality caused by the lack of the topoisomerase is due to RNAP backtracking, which in turn, leads to R-loops accumulation and transcription-replication conflicts [178].

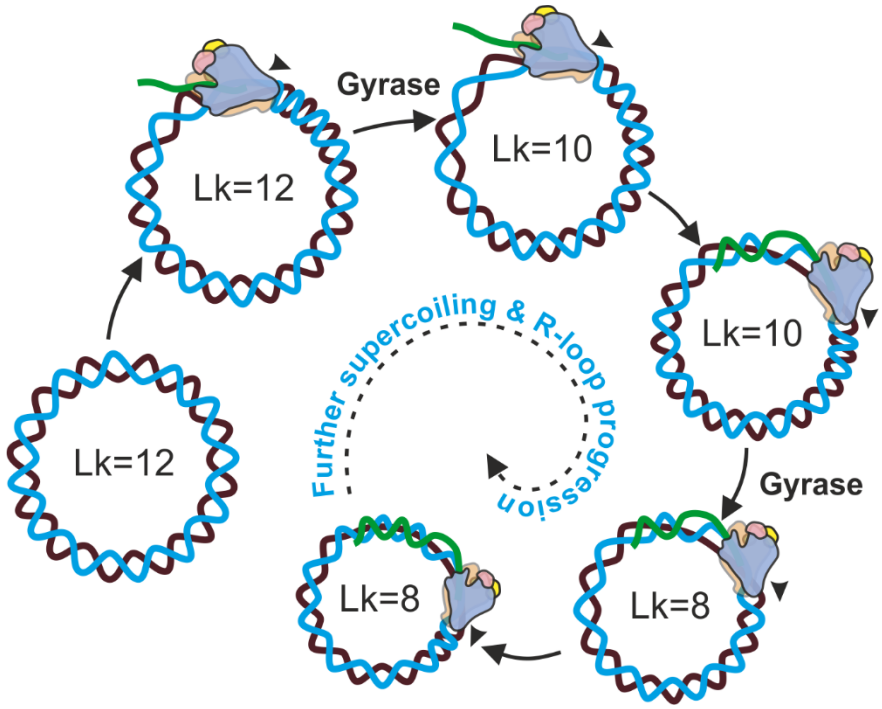


Figure 8. **A positive feedback cycle leads to hypernegative supercoiling of plasmid DNA and R-loops accumulation when transcription proceeds in the presence of DNA gyrase and the absence of TopoI**. RNAP is shown as a cartoonish protein complex. Direction of RNAP elongation is shown with an arrow. Nascent RNA is shown in green. Owervinding of RNA with a blue DNA strand (template strand) indicates formation of an R-loop. Decreased and increased density of DNA turns behind and ahead of elongating RNAP corresponds to negative and positive supercoiling, respectively. Gyrase indicates negative supercoiling by gyrase. Lk, linking number. Decreasing linking number corresponds to increasing level of negative supercoiling of a plasmid over time.

The EcTopoI was shown to bind RNAP, and the interaction was mapped to the C-terminal portion of EcTopoI and the β' subunit of RNAP [179]. The RNAP:TopoI interaction was also reported for mycobacteria [180] and *Streptococcus pneumoniae* [181]. It was hypothesized that association with RNAP allows TopoI to rapidly relax negative DNA supercoils forming behind the elongating RNAP, thereby preventing R-loop formation [179,182]. The chromosomal distribution of EcTopoI is currently unknown, although some sequence preferences have been reported *in vitro* [183,184]. Recently, genome-wide distribution of TopoI from *M. tuberculosis* (MtTopoI), *M. smegmatis* (MsTopoI), and *S. pneumoniae* (SpTopoI) was investigated using ChIP-Seq [66,112,181]. In all cases, the topoisomerase was shown to

associate with actively transcribed genes, with particular enrichment upstream of RNAP peaks at promoter regions. These findings agree with the twin-domain model proposed by Liu and Wang [68,185], but do not necessarily imply direct TopoI-RNAP association.

1.3.6 Topoisomerase III

Together with Topo I, bacteria typically have another type-IA topoisomerase called topoisomerase III (Topo III, encoded by *topB*). Topo III shares an N-terminal domain with Topo I, but lacks Zn-binding and C-terminal domains characteristic for Topo I. *In vitro*, Topo III was shown to have a range of activities including supercoiling relaxation, decatenation and resolution of replication intermediates. First, Topo III is capable of relaxation of negative supercoils when a DNA substrate has an unpaired or melted region. Single-molecule experiments demonstrated that Topo III performs fast and processive relaxation runs with long pauses between the runs [186,187]. Second, Topo III was shown to efficiently decatenate DNA braids comprised of nicked DNAs with bulges *in vitro* [188]. In another single-molecule experiment, Topo III was shown to have a comparable decatenation activity with TopoIV on DNA substrates with a single-stranded gapped region [99]. Consistently, treatment of replication intermediates of a plasmid containing *OriC* and two *ter* sites with Topo III resulted in resolution of precatenanes [189]. Last but not least, *in vitro* in cooperation with RecQ helicase and SSB, Topo III was able to resolve partially replicated circular molecules releasing sister molecules with single-stranded regions [190,191]. Of note, eukaryotic homologs of Topo III typically interact and cooperate with RecQ-like helicases to resolve double Holliday junctions formed from excessive recombination [192]. Presumable cooperation observed between Topo III and RecQ in bacteria implies that it is ubiquitous and, probably, evolutionary universal. Should be noted, that suggested cooperation is lacking strong evidence *in vivo* and was disputed by genetic experiments [193–196]. Topo III also acts as an RNA topoisomerase *in vitro*, able to cleave and reseal RNA strand while moving another one through the temporary gap. However, the physiological role of this activity is obscure [197,198].

Despite the extensive repertoire of activities detected *in vitro*, the cellular role of Topo III is enigmatic. *TopB-null* cells are lacking clear growth defects and have a phenotype of a wild-type. However, on a *parC1215* background (thermosensitive TopoIV, caused lethality at high temperatures), *topB* overexpression was found suppressive at non-permissive temperatures, indicating a supportive role of Topo III in DNA decatenation and DNA segregation [189]. Correspondingly, a *parC1215ΔtopB* double mutant (and similar double mutants by TopoIV and Topo III) exhibits severe growth defects and filamentation and can grow only at lower temperatures [196]. Via genetic experiments it was also demonstrated that Topo III

is required for DNA segregation when DNA gyrase activity is reduced [162]. A colocalization between replisome and Topo III was observed by microscopy and the interaction between DnaX complex and Topo III was also supported *in vitro*. Moreover, fine analysis of cell division demonstrated that DNA segregation in *topB-null* mutants is less organized and delayed [195]. Together, a corpus of data indicates that Topo III is likely involved in DNA segregation by resolution of precatenanes close to the replisome and/or relaxation of excessive supercoiling in support of other bacterial topoisomerases.

Chapter 2. Project Rationale and Objectives

DNA topology, local and global, impacts many if not all aspects of cell physiology, including transcription, replication, recombination, DNA compaction and post-replicative segregation. DNA topoisomerases are the major regulators of DNA topology. In recent years, DNA binding and/or activity profiles were obtained for several *E. coli* (DNA gyrase ChIP-chip data [65], TopoIV ChIP-Seq and cleavage data [116]), *Mycobacterium* (DNA gyrase ChIP-Seq data [66], TopoI ChIP-Seq and cleavage data [112]), *Streptococcus* (TopoI ChIP-Seq data [181]) topoisomerases. Overall data indicates that TopoI and DNA gyrase are associated with active transcription, in agreement with the Liu & Wang twin-domain model [68]. However, high-resolution topoisomerase activity data is lacking even for *E. coli* with more than a century history as a model organism. Comprehensive whole-genome analysis of topoisomerase activity in various conditions is required to test *in vitro* observations, assess the role of topoisomerases in resolution of supercoils generated by transcription and replication, and find binding and cleavage DNA motifs of the enzymes.

The aim of this project was to build high-resolution and whole-genome activity profiles for the main topoisomerases of *E. coli* (DNA gyrase, TopoIV, Topo I) and investigate the interplay of these enzymes with transcription and replication, their role in topology regulation *in vivo*. To achieve the aim, first, a suitable method, capable of detection of topoisomerase-induced DNA cleavage, was developed (Chapter 4.1, the Topo-Seq method). Using this instrument, then, whole-genome activities of DNA gyrase (Chapters 4.2-4.13), TopoIV (Chapter 4.14), and TopoI (Chapters 4.15-4.20) were investigated.

Chapter 3. Materials and Methods

3.1 Buffers

Strep-Tactin lysis buffer 1: 50 mM Tris HCl pH 8.0, 150 mM NaCl

Strep-Tactin lysis buffer 2: 100 mM Tris HCl pH 8.0, 150 mM NaCl, 1 mM EDTA

Strep-Tactin elution buffer 1: 50 mM Tris HCl pH 8.0, 150 mM NaCl, 2.5 mM desthiobiotin

Strep-Tactin elution buffer 2: 50 mM Tris HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA, 2.5 mM desthiobiotin

TALON lysis buffer: 50 mM Tris HCl pH 8.0, 150 mM NaCl, 2.5 mM imidazole

TALON elution buffer: 50 mM Tris HCl pH 8.0, 150 mM NaCl, 300 mM imidazole

AMpure binding buffer: 10 mM Tris-HCl pH 8.0, 1 mM EDTA, 250 mM NaCl, 20% w/v PEG-8000, 0.05% v/v Tween 20

TopoI storage buffer:10 mM Tris-HCl pH 7.5, 50 mM KCl, 0.1 mM EDTA, 1 mM 2-mercaptoethanol

10x TopoI reaction buffer: 100 mM Tris-HCl pH 8.0, 500 mM NaCl, 60 mM $MgCl_2$

5x TopoI Stop-buffer: 50 mM EDTA pH 8.0, 0.5% w/v bromophenol blue, 50% v/v glycerol

Gyrase reaction buffer: 35 mM Tris-HCl pH 7.5, 24 mM KCl, 4 mM $MgCl_2$, 0.1 mg/ml BSA, 6.5% v/v glycerol, 2 mM DTT, 1.8 mM spermidine, 1 mM ATP

3x Laemmli buffer with urea: 10% v/v glycerol, 4% w/v SDS, 4 mM EDTA pH 8.0, 8 M urea, 0.01% w/v bromophenol blue

RNase III reaction buffer: 10 mM Tris-HCl pH 7.9, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT

Transfer buffer: 48 mM Tris, 39 mM glycine, 0.037% w/v SDS, 20% v/v ethanol

TBST buffer: 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% v/v Tween 20

TAE buffer: 40 mM Tris-acetate pH 8.3, 1 mM EDTA

TES buffer: 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 250 mM NaCl

TESS buffer: 10 mM Tris-HCl pH 7.5, 250 mM NaCl, 1 mM EDTA, 0.1% v/v Tween 20, 0.05% w/v SDS

TE buffer: 10 mM Tris-HCl pH 7.5, 1 mM EDTA

3.2 Bacterial strains and plasmids

Full lists of bacterial strains and plasmids used in this study can be found, respectively, in Tables 1 and 2.

Table 1. Bacterial strains used in this study

Strain	Description	Source
E. coli DY330	W3110 *lacU169 gal490 cI857* Δ(*cro-bioA*)	Purchased from Horizon Discovery Biosciences
E. coli DY330 *topA-SPA*	W3110 *lacU169 gal490 cI857* Δ(*cro-bioA*) *topA-SPA*	Purchased from Horizon Discovery Biosciences
E. coli DY330 *gyrA-SPA*	W3110 *lacU169 gal490 cI857* Δ(*cro-bioA*) *gyrA-SPA*	Purchased from Horizon Discovery Biosciences
E. coli DY330 *gyrA-SPA MuSGS*	W3110 *lacU169 gal490 cI857* Δ(*cro-bioA*) *gyrA-SPA (dcuC-crcA)::(cat-MuSGS)*	This study
E. coli DY330 *parC-SPA*	W3110 *lacU169 gal490 cI857* Δ(*cro-bioA*) *parC-SPA*	Purchased from Horizon Discovery Biosciences
E. coli DY330 *parC-SPA gyrAS83L*	W3110 *lacU169 gal490 cI857* Δ(*cro-bioA*) *parC-SPA gyrAS83L*	This study
E. coli DY330 *rpoC-TAP*	W3110 *lacU169 gal490 cI857* Δ(*cro-bioA*) *rpoC*-TAP	Purchased from Horizon Discovery Biosciences
E. coli BW25113	*E. coli* K12 F– Δ(*araD-araB*)567 Δ*lacZ*4787(::*rrnB*-3) λ– *rph*-1 Δ(*rhaD-rhaB*)568 *hsdR*514	Laboratory stock
E. coli BW25113 *topA-Δ11kDa*	*E. coli* K12 F– Δ(*araD-araB*)567 Δ*lacZ*4787(::*rrnB*-3) λ– *rph*-1 Δ(*rhaD-rhaB*)568 *hsdR*514	This study
E. coli BW25113 *topA-Δ14kDa*	*E. coli* K12 F– Δ(*araD-araB*)567 Δ*lacZ*4787(::*rrnB*-3) λ– *rph*-1 Δ(*rhaD-rhaB*)568 *hsdR*514	This study
E. coli BW25113 *topA-Δ30kDa*	*E. coli* K12 F– Δ(*araD-araB*)567 Δ*lacZ*4787(::*rrnB*-3) λ– *rph*-1 Δ(*rhaD-rhaB*)568 *hsdR*514	This study
E. coli CSH50 *λsfiA::lacZ*	(Δ*lac-pro*) *ara rpsL thi*	[199]
E. coli NEB5α	*E. coli* K12 *huA*2 (*argF-lacZ*)*U*169 *phoA glnV*44 80 (*lacZ*)*M*15 *gyrA*96 *recA*1 *relA*1 *endA*1 *thi*-1 *hsdR*17	NEB
E. coli BL21(DE3)	*E. coli* B F– *omp*T *hsdSB* (rB–, mB–) *gal dcm* (DE3)	Laboratory stock

C. crescentus CB15N *gyrA-3xFLAG*		Constructed by Monica Guo

Table 2. Plasmids used in this study

Plasmid	Description	Source
pBAD-mcbABCDEFG	Plasmid containing a full operon required for microcin B17 production, *araBAD* promoter,	Gift of Dr. M. Metelev, [200]
pMP1000	Contains SGS from Mu phage genome, Amp^R	Gift of Dr. P. Higgins, [201]
pBR322	Tc^R, Amp^R	Laboratory stock
pKD3	Contains *cat* gene flanked by FRT sites, used as a source of a resistance cassette for recombineering	Gift of Dr. K. Datsenko, [202]
pKD4	Contains *kan* gene flanked by FRT sites, used as a source of a resistance cassette for recombineering	Gift of Dr. K. Datsenko, [202]
pKD46	pKD with λ RED systems' *exo*, *beta*, *gam* genes under the control of *araBAD* promoter	Gift of Dr. K. Datsenko, [202]
pCA24 *gfp*	Contains *gfp* gene insert under T5-lac promoter, Cm^R	ASKA collection, [203]
pCA24 *topA*	Contains *topA* gene insert under T5-lac promoter, Cm^R	ASKA collection, [203]
pET28	T7 promoter, Amp^R	Laboratory stock
pET28 topA_strepII	Contains *topA* gene fused with strepII encoding sequence (C-terminal), T7 promoter, Amp^R	This study
pCA24 14kDa CTD	Contains fragment encoding CTD of EcTopoI under T5-lac promoter, Cm^R	Alina Galivondzhyan
pCA24 topA Y319F	Contains fragment encoding catalytically inactive EcTopoI mutant under T5-lac promoter, Cm^R	This study
pBAD33	*araBAD* promoter, Cm^R	Laboratory stock
pBAD33 topA-strepII	Contains *topA* gene fused with strepII encoding sequence (C-terminal), *araBAD* promoter, Cm^R	This study
pBAD33 topA(G116S/M320V)	Contains fragment encoding EcTopoI double mutant under *araBAD* promoter, Cm^R	This study

pBAD30	*araBAD* promoter, Amp^R	Laboratory stock
pBAD30 rnhA	Contains *rnhA* gene with native RBS under *araBAD* promoter, Amp^R	This study
pBAD30 rnhAD10N	Contains fragment encoding catalytically inactive RNAse I mutant under *araBAD* promoter, Amp^R	This study

3.3 Cultivation of bacteria

Bacteria were grown in LB or 2YT (where indicated) rich media supplemented with the corresponding antibiotics when required. MacConkey lactose plates (1.2% agar) were prepared according to the manufacturer's instructions (BD Difco).

3.4 Microcin B17 purification

E. coli BW25113 was transformed with *pBAD-mcbABCDEFG* plasmid and night culture was inoculated with one isolated colony. 1 liter of 2YT media was inoculated with 1/100 volume of the overnight culture. When the culture reached OD_{600}=0.6-0.8, *mcb* operon expression was induced by addition of arabinose up to 1 mM. Cultivation was continued for 18-20 hours on 37°C with shaking on 180 rpm. Cells were pelleted by centrifugation, then resuspended in 40 ml of 100 mM acetic acid/1 mM EDTA and boiled for 15 min. The clarified supernatant was applied onto 1g C18 HyperSep cartridge (Thermo Scientific) pre-equilibrated with 0.1% trifluoroacetic acid (TFA). The cartridge was extensively washed with 0.1% TFA followed by 10% acetonitrile (ACN) in 0.1% TFA. The microcin B17-containing fraction was eluted in 30 ml 30% ACN in 0.1% TFA and vacuum dried (GeneVac). The resulting precipitate was dissolved in dimethyl sulfoxide (DMSO) and applied onto Phenomenex Luna C18 high-performance liquid chromatography (HPLC) column (pre-equilibrated with 0.1% TFA) in 10% DMSO/0.1% TFA. Elution was performed with linear gradient of ACN (from 0% to 50% ACN in 30 min) in 0.1% TFA. Microcin B17 was eluted between 12-16 min, individual peaks were collected. Fractions obtained were merged and dried *in vacuo*. Lyophilized powder was dissolved in DMSO and stored at -20°C. Concentration of the microcin B17 was determined spectroscopically as described previously [204].

3.5 *E. coli* DY330 *gyrA-SPA MuSGS* strain construction

349-bp DNA fragment containing strong gyrase binding site from bacteriophage Mu (Mu SGS) was amplified from pMP1000 plasmid [201] (a gift of P. Higgins) using Mu_G_F and Mu_G_R primers. *cat*

gene was amplified from pKD3 plasmid [202] with G_cat_Mu and cat_Mu_R primers (**Supplementary Table 1**). Fragments were joined by overlap PCR using G_cat_Mu and Mu_G_R primers [201]. Resulting cassette was inserted into intergenic region between *dcuC* and *crcA* genes using recombination techniques described elsewhere [205]. Successful insertion was confirmed by PCR and whole genome sequencing.

3.6 Minimal inhibitory concentration (MIC) measurement

MICs were determined by microdilution method in 96-well plates in liquid LB according to CLSI guidelines (M07-A10). Inoculum suspensions were prepared by dilution of night cultures grown in LB up to 1-5*10^5 colony forming units (CFU)/ml. Plates were incubated 18-24 hours on 37°C, MIC recorded as minimal concentration of antibiotic that completely inhibit bacterial growth.

3.7 ChIP with different DNA gyrase poisons as stabilizing agents (DNA gyrase Topo-Seq)

1 ml of overnight culture was prepared for *E. coli* DY330 *gyrA-SPA* or *E. coli* DY330 *gyrA-SPA MuSGS* by inoculating 2YT medium supplemented with antibiotics (kanamycin 50 μg/ml for DY330 *gyrA-SPA* and kanamycin 50 μg/ml, chloramphenicol 15 μg/ml for DY330 *gyrA-SPA MuSGS*) with cells from a single colony. 2 ml of starter culture was cultivated at 32°C with shaking (180 rpm), then 1 ml was inoculated into 100 ml of 2YT without antibiotics and cultivation was continued at 37°C and shaking until culture reaching mid-logphase (OD_{600}=0.6-0.8). At this point the culture was bisected and DNA gyrase poison (see drugs and concentrations in **Table 3**) was added to the first half (+A samples), while the second served as a control (-A samples).

Table 3. Gyrase poisons and their concentrations used for Topo-Seq

Poison	**Abbreviation**	**Concentration**	**Biological replicates #**
Ciprofloxacin	Cfx	0.9 μM	3
		10 μM	3
Oxolinic acid	Oxo	120 μM	3
Microcin B17	Micro	10 μM	2
		50 μM	1

Cultures (+A and -A) were incubated at 37°C with shaking for additional 15 min, then cells were pelleted by centrifugation at 10°C (4500g) and resuspended in 10 ml of TES buffer (10 mM Tris-Cl pH 7.5, 1 mM

EDTA, 250 mM NaCl). Washing procedure was repeated twice to remove all traces of culturing medium. Washed pellets were resuspended in 1 ml of TESS buffer (10 mM Tris-Cl pH 7.5, 1 mM EDTA, 250 mM NaCl, 0.02% SDS, 0.2% Tween-20) with addition of protease inhibitors cocktail (cOmplete ultra EDTA free, Roche) and RNase A (Thermo Scientific). Resulting suspensions were sonicated with parameters optimized to obtain DNA fragments between 200 and 700 bp (SONOPULS HD 3100). Lysates were diluted with 1 ml of TES buffer and 100 µl of ANTI-FLAG® M2 affinity gel (Sigma-Aldrich) was added. Immunoprecipitation was performed for 1.5-2 hours at room temperature with moderate mixing, then affinity gel was washed 4 times (two times with 1 ml of TESS buffer, once with 1 ml of TES buffer, and once 1 ml of TE buffer). For proteolysis, affinity gel obtained after the last wash step was diluted with TES buffer up to 200 µl, proteinase K (Sigma-Aldrich) was added (0.5 mg/ml) and samples were incubated at 55°C for at least 3 hours. After this step samples were centrifuged (2 minutes, 2000g at room temperature) and DNA was extracted from resulting supernatant with phenol/chlorophorm method followed by precipitation in ethanol. Mock controls (-IP) were made both for +A and for -A: for this, 100 µl aliquots of lysates obtained after sonication were deproteinized and DNA was purified as described before. The procedure described gives a quartet of samples (+A+IP, +A-IP, -A+IP, -A-IP), where +A-IP, -A+IP, and -A-IP serves as controls for gyrase poison action and immunoprecipitation.

Sequencing libraries were prepared with Accel NGS 1S kit (Swift Bioscience) from DNA obtained on step I procedure, according to the manufacturer's protocol. Library preparation and sequencing were performed by Dr. M. Logacheva at A.N.Belozersky Research Institute of Physico-Chemical Biology MSU. Sequencing was performed on Illumina NextSeq platform (150 bp paired-end reads). Combination of gyrase poison-mediated ChIP procedure (Step I) and specific sequencing libraries preparation step (Step II) was named by us a Topo-Seq technique. For each antibiotic Topo-Seq was performed in triplicate.
Reads were aligned to the *E. coli* W3110 *MuSGS* genome (*E. coli* W3110 genome with the insertion of *cat*-MuSGS cassette may be downloaded from GEO: GSE95567) using BWA-MEM [206]. BAM files were prepared with Samtools [207] and visualized in IGV [208].

For each position in the genome a number of 3'-ends (N3E) and 5'-ends were counted based on reads alignments stored in SAM file. Obtained values were divided by the total amount of reads aligned and multiplied by the lowest value across samples forming the quartet. Additional normalization was performed to get rid of the bias in the coverage depth across the genome: due to active replication, there is a significant difference in the total amount of DNA between origin of replication and terminator area. For this purpose, N3E values of +A+IP sample were divided by corresponding N3Es of +A-IP control and N3Es of –A+IP sample were divided by N3Es of –A-IP control (all samples originate from the same quartet). N3Es of +A-

IP and –A-IP controls were preliminarily smoothed using 200 kb sliding window. In resulting pairs of samples (+A+IP_norm and –A+IP_norm) gyrase cleavage sites (GCSs) were called if values in *i* and *i+5* positions in +A+IP_norm sample both exceed the right confident interval value calculated based on the appropriate values in -A+IP sample (Audic & Claverie statistical test from [209], p-value<0.05). As Topo-Seq was performed for each antibiotic in triplicate, GCS was called reliable if it was identified in at least two biological replicas. Only reliable GCSs sets were used for further analysis.

3.8 qPCR validation of Topo-Seq (Topo-qPCR)

qPCR was performed to estimate the enrichment of DNA at specific loci after Step I (MuSGS, rRNA A DS, *ccmH*, and rRNA A US) and validate data obtained with Topo-Seq (primers listed in **Supplementary Table 1**). Step I of the procedure followed by qPCR was named Topo-qPCR.

3.9 Gyrase Topo-Seq and Topo-qPCR with cells treated with RNAP inhibitor rifampicin

Topo-Seq and Topo-qPCR experiments were performed as described above with the only exception that *E. coli* DY330 *gyrA-SPA* or *E. coli* DY330 *gyrA-SPA MuSGS* cells were pretreated with RNAP inhibitor rifampicin (Rif) (final concentration 122 μM) for 15 minutes to stop transcription [210] before addition of DNA gyrase poison (10 μM Cfx). Topo-Seq was performed in triplicate.

3.10 Gyrase Topo-Seq with *E. coli* DY330 *gyrA-SPA* carrying pBR322 plasmid

Gyrase Topo-Seq experiments with *E. coli* DY330 *gyrA-SPA* pBR322 were performed as described above except that growth medium was supplemented with ampicillin 100 μg/ml and cultures were treated with increased, 100 μM, concentration of Cfx to trap the gyrase cleavage complexes. Topo-Seq was performed in triplicate.

3.11 Gyrase Topo-Seq with *E. coli* DY330 *gyrA-SPA* at different growth stages

Gyrase Topo-Seq experiments with *E. coli* DY330 *gyrA-SPA MuSGS* were performed as described above except cultures were grown to OD_{600}~6 (early stationary phase culture – ESP, approximately 6 h after inoculation of the starter culture) and OD_{600}~10 (stationary phase culture – SP, approximately 10 h after inoculation of the starter culture). To trap gyrase cleavage complexes, Cfx was added to concentration 10 μM. Topo-Seq experiments were performed in triplicates.

3.12 Time-resolved gyrase Topo-Seq with synchronized *E. coli* DY330 *gyrA-SPA*

To prepare synchronously replicating and dividing *E. coli* cells, a stationary-phase method was used [211]. A colony of *E. coli* DY330 *gyrA-SPA Mu SGS* was inoculated to 5 mL LB supplemented with 30 μg/mL chloramphenicol, and was cultivated overnight at 37ºC, 180 rpm. 0.5 mL of this starter culture was diluted 100x with LB and cultivation proceeded for 6 h (OD_{600}~6, ESP culture). At this time-point (named -3 min) 5 mL of culture was collected and treated with 10 μM Cfx for 15 min. Cells were collected by centrifugation and washed with 10 mL of TES buffer. Obtained cell pellet was snap-frozen in $N_{2(liq.)}$ and stored at -20ºC (+Cfx sample). Simultaneously with the aliquot collection, another, 1 mL, aliquot was taken and SDS was added to it to a concentration 1%. Cells were rapidly lysed at room temperature and the lysate was transferred to -20ºC (-Cfx control). The ESP culture was diluted 10x with LB and 50 mL aliquots were collected and immediately treated with 10 μM Cfx after 0, 5, 10, 15, 20, 25, 30 min after dilution. 15 min treatment with Cfx was immediately followed by centrifugation at 4ºC and washing with 10 mL of TES. Obtained cell pellet was snap-frozen in $N_{2(liq.)}$ (+Cfx samples). -Cfx controls for the time-points were processed as described above. All samples were rapidly transferred to -20ºC and stored overnight.

Next day, +Cfx samples were thawed at room temperature and processed as described above for DNA gyrase Topo-Seq using different gyrase poisons resulting in +Cfx-IP and +Cfx+IP DNA samples. -Cfx controls were treated with proteinase K (Sigma-Aldrich) (0.5 mg/ml) by incubation at 55°C for 2 hours and DNA was extracted with phenol/chlorophorm method followed by precipitation in ethanol resulting in -Cfx-IP method. Two biological replicates of the experiment were conducted.

DNA sequencing and initial data analysis were performed as described above, except that for GCS-calling procedure -Cfx-IP sample substituted both -Cfx-IP and -Cfx+IP. To perform time-series analysis, all +Cfx-IP and +Cfx-IP tracks were normalized by average coverage depth (named, N_tracks). Normalized tracks were smoothed by Fourier filtering (20 first harmonics were retained) (named, NSF_tracks). Locations of replisomes at a time-point *t+1* were identified as global maximum and global minimum of the function 1 for the counter-clockwise and clockwise replisomes, respectively.

$$\frac{\mathrm{d}(\mathrm{NSF\_+Cfx-IP}(t+1)/\ \mathrm{NSF\_+Cfx-IP}(t))}{\mathrm{dx}} \quad \textbf{(1)}$$

Locations of progressing gyrase enrichment at a time-point *t+1* was identified as two global maxima of function 2 – one for the left replichore and one for the right replichore.

$$\frac{\mathrm{N\_+Cfx+IP}(t+1)/\mathrm{NSF\_+Cfx-IP}(t+1)}{\mathrm{N\_+Cfx+IP}(t)/\mathrm{NSF\_+Cfx-IP}(t)} \quad \textbf{(2)}$$

Replication rates were derived from the locations of replisomes at different time-points.

3.13 Gyrase motif identification

DNA sequences were extracted from *E. coli* W3110 *MuSGS* genome as a 130 bp vicinity of GCSs' positions identified for exponentially growing cells. Therefore, all the sequences are centered relative to the DNA gyrase cleavage sites and sequences sets can be processed as multiple alignments: nucleotide frequencies were counted within formed columns giving position probability matrix (PPM) visualized with Python Matplotlib package [212]. The degenerate GC% motif was obtained similarly by calculating the frequency of G or C in the columns of the alignment. Logos were calculated with WebLogo for the same sets of DNA sequences [213]. Motif visualization and Logos were made for each Topo-Seq condition independently. "Combined" gyrase motif was constructed using sequences obtained from Cfx, Micro and Oxo Topo-Seq experiments. For each antibiotic top 732 GCSs having the highest N3E values were taken, resulting in a set containing 1828 sequences that was used for PPM and Position Weight Matrix (PWM) construction. Antibiotic-specific bias was removed for the most antibiotic-influenced positions (0-3 bp) in the intermediate PPM by changing corresponding values with a baseline frequency of nucleotides observed in a *E. coli* W3110 genome. To find potential DNA gyrase binding sites, sequences of interest were scanned with final PWM in forward and reverse-complement forms. For the particular position final score was specified as a maximum between values obtained for both strands. PWM construction and sequences scanning were performed with Bio.motifs from Biopython package [214].

3.14 3D modelling

DNA model (B-form, 10.7 bp/turn, bend angle 210°) was constructed for 43 bp fragment of the consensus sequence (-63:-21 or 24:66 regions, corresponds to periodic areas) using 3D-DART web service [215]. The model was manually docked with the structure of *E. coli* DNA gyrase CTD (1zi0 [216]) in PyMOL (PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC).

3.15 Genome editing, introduction of the gyrA-S83L mutation

The *gyrA-S83L* mutation conferring resistance to fluoroquinolone antibiotics [217], was introduced to the *E. coli* DY330 *parC-SPA* strain by recombineering using the template oligonucleotide complementary to the lagging DNA strand [218]. A culture of *E. coli* DY330 *parC-SPA* harboring the pKD46 plasmid was prepared for recombineering as described in [202] and 50 µL aliquot of cells (~$6*10^8$ cells) was electroporated with 0.2 µL of 50 µM chemically synthesized 70-mer oligonucleotide (~$6*10^{12}$ molecules)

gyrA_S83L (**Supplementary Table 1**). Fluoroquinolone-resistant mutants were selected on LB plates containing 500 nM ciprofloxacin. To verify the mutation acquisition, a fragment of *gyrA* gene was PCR amplified using gyrA_F and gyrA_R (**Supplementary Table 1**) primers and subjected to Sanger sequencing.

3.16 TopoIV Topo-Seq and data analysis

Purification and sequencing of DNA fragments covalently attached to the poisoned TopoIV (Topo-Seq) was performed with *E. coli* DY330 *parC-SPA* and *E. coli* DY330 *parC-SPA gyrA-S83L* as described previously for *E. coli* DNA gyrase [219]. Briefly, ciprofloxacin (final concentration 10 μM) was added to the 50 mL of exponentially growing cell culture in LB medium when OD_{600}~0.6, and the culture was incubated for 15 min with rotation at 37°C. Cells were harvested by centrifugation and washed twice with TES buffer (10 mM Tris–HCl pH 7.5, 1 mM EDTA, 250 mM NaCl). Then, cells were resuspended in 1 mL of TES and disrupted by sonication (DNA fragmentation range 200-800 bp). Lysate was cleared by centrifugation and incubated with 100 μL of ANTI-FLAG M2 affinity gel (Sigma-Aldrich). After immunoprecipitation, the resin was treated with Proteinase K (Sigma-Aldrich) in 200 μL of TES and DNA fragments were purified with AMPure XP magnetic beads from the supernatant (Beckman Coulter). Non-treated control was processed identically, excepting that ciprofloxacin was not added. For mock control, DNA was extracted from 100 μL of cleared lysate using the GeneJET DNA purification kit (Thermo Scientific).

NGS libraries were prepared using Accel NGS 1S kit (Swift Bioscience). DNA sequencing was performed with Illumina HiSeq 4000 in a 150+150 bp paired-end mode. Library preparation and sequencing were performed at Skoltech Genomics Core Facility. Topo-Seq was performed in triplicate.

Raw reads were filtered with Trimmomatic [220] and then aligned to the *E. coli* W3110 MuSGS genome (the genome may be obtained from GEO: GSE95567) using BWA-MEM [206]. SAM, BAM, and bed files were prepared with Samtools [207]. The number of DNA fragments' 3'-ends (N3E) was calculated per position, based on the read alignments stored in SAM files. Normalization of N3E tracks and identification of Topoisomerase Cleavage Sites (TCSs) was performed as described previously [219]. The resultant tracks were further analyzed using custom Python scripts (https://github.com/sutormin94/TopoIV_Topo-Seq_experiment).

3.17 Cloning of EcTopoI and EcTopoI 14kDa CTD, construction of pCA24 topA Y319F plasmid

The *topA* gene was PCR-amplified from genomic DNA extracted from *E. coli DY330* (Genomic DNA extraction kit, ThermoFisher Scientific). To remove the NcoI restriction site at the end of the *topA* gene,

two overlapping fragments were generated using the following primers: topA_NcoI_fw + NcoI_mut_rev, NcoI_mut_fw + topA_HindIII_strepII_rev (for primers see **Supplementary Table 1**). The two fragments were joined by overlap extension PCR using topA_NcoI_fw and topA_HindIII_strepII_rev primers. The resulting PCR product was cloned into pET28 at NcoI and HindIII sites, giving pET28 topA-strep. The C-terminal StrepII tag was introduced with topA_HindIII_strepII_rev primer.

3.18 Construction of pBAD33 topA-strepII and pBAD33 topA(G116S/M320V)-strepII plasmids

To generate *topA* double-mutant (G116S/M320V), three overlapping fragments were generated by PCR of pET28 topA-strep plasmid using three primer pairs: topA_XbaI_RBS_fw+topA_G116S_out_rev, topA_G116S_in_fw+topA_M320V_in_rev, and topA_M320V_out_fw+topA_strepII_HindIII_rev. The fragments were fused by overlap extension PCR using primers topA_XbaI_RBS_fw+topA_strepII_HindIII_rev (**Supplementary Table 1**). The final amplicon treated with DpnI was cloned into pBAD33 at XbaI and HindIII sites, resulting in pBAD33 topA(G116S/M320V)-strepII.

To construct pBAD33 topA-strep, a plasmid backbone of pBAD33 topA(G116S/M320V)-strepII was obtained by NdeI and HindIII and ligated with a PCR-product (topA_NdeI_fw+topA_HindIII_strepII_rev primers, see **Supplementary Table 1**) digested with the same restriction enzymes and DpnI.

3.19 Toxicity assay of EcTopoI G116S/M320V

E. coli DY330 *topA-SPA* strain harboring pBAD33 topA(G116S/M320V)-strepII plasmid was grown in LB supplemented with 0.5% glucose. At OD_{600}~0.2 the culture was divided and 10 mM Ara was added to one half (+Ara 10 mM), while the second half served as a non-induced control (-Ara). To construct growth curves, OD_{600} of the cultures were measured every 15 min. The experiment was performed in triplicate.

3.20 SOS-response detection and quantification

E. coli CSH50 *λsfiA::lacZ* reporter strain was used for SOS-response detection [199]. For 14kDa CTD overexpression-induced SOS-response detection, the strain was transformed with pCA24 14kDa CTD. Non-transformed strain and strain transformed with pCA24 GFP served as controls. Transformants were grown on MacConkey agar supplemented with 1% lactose for 4-6 h. SOS-response induction leads to *lacZ* expression, followed by active lactose digestion and acidification of the medium, which is monitored by color change – red color indicates acidification and SOS-response. Intensity of a red color adjusted for the

background level was qantified and used as a proxy for SOS-response levels. The assay was performed in triplicate.

For EcTopoI G116S/M320V double-mutant overexpression-induced SOS-response detection, the reporter strain was transformed with pBAD33 EcTopoI G116S/M320V. Non-transformed strain and strains transformed with pBAD33 or pBAD33 EcTopoI served as controls. The strains were grown on MacConkey agar supplemented with 1% lactose and 1 mM Ara for 4-6 h.

3.21 Strand-specific EcTopoI Topo-Seq and data analysis

E. coli DY330 *topA-SPA* cells transformed with pBAD33 topA(G116S/M320V)-strepII plasmid were grown at 37ºC in LB supplemented with chloramphenicol (34 μg/mL) and 0.5% glucose until OD_{600}=0.4. The 100 mL culture was then divided: one half was induced by adding arabinose to 10 mM (+Ara), and the other half served as a non-induced control (-Ara). 30 min after the induction, cells were harvested by centrifugation (3000 g), the cell pellet was frozen in liquid N_2 and stored at -80ºC until further processing. The cell pellet was resuspended in 1 mL of Strep-Tactin lysis buffer (50 mM Tris HCl pH 8.0, 150 mM NaCl) containing protease inhibitors (cOmplete ULTRA, Sigma-Aldrich) and RNase A (0.1 mg/ml, Thermo Scientific). Cells were disrupted by sonication as described for EcTopoI ChIP-Seq. Lysates were clarified by centrifugation at 8000 g for 5 min at 4°C, and the resulting supernatant was used for further analysis.
For the preparation of Input DNA, the initial lysate (100 μL) was treated with proteinase K (Thermo-Fisher Scientific) and decrosslinked by incubation at 55°C for 4 h. Input DNA was purified using a DNA clean-up kit (Thermo Fisher Scientific), and the DNA fragmentation range was assessed by electrophoresis.
For immunoprecipitation of EcTopoI-DNA cleavage complexes (+Ara+IP and -Ara+IP samples), 900 μL of the lysate was mixed with 80 μL of Strep-Tactin Superflow Plus affinity resin (Qiagen) pre-equilibrated with Strep-Tactin lysis buffer containing 0.05% SDS. After 1 h of incubation at 25ºC the resin was washed 3x with Strep-Tactin lysis buffer, and the complexes were eluted with 100 μL of Elution buffer (50 mM Tris HCl pH 8.0, 150 mM NaCl, 2.5 mM desthiobiotin). A 20 μL-aliquot of the eluate was analyzed by SDS-PAGE. The remaining 80 μL of the eluate was treated with proteinase K (Thermo-Fisher Scientific) overnight at 50ºC. The IP-DNA samples were purified using AMPure XP magnetic beads (Beckman Coulter). All Topo-Seq experiments were performed in triplicates.
NGS libraries were prepared using a strand-specific Accel NGS 1S kit (Swift Bioscience) suitable for damaged DNA. DNA sequencing was performed by Illumina NextSeq 150+150 bp paired-end protocol. Library preparation and sequencing were performed at Skoltech Genomics Core Facility.

Reads were prepared and mapped to the reference genome as described above for WGS and ChIP-Seq sequencing data. The number of DNA fragments' 3'-ends was calculated per position (N3E) separately for forward and reverse strands, based on the read alignments stored in SAM files. The tracks were scaled by the total number of aligned reads to normalize the coverage across samples, and the biological replicates were averaged. After that, the -IP tracks (+Ara-IP and -Ara-IP) were subtracted from the +IP tracks (+Ara+IP and -Ara+IP) tracks strand-wise, resulting in +Ara and -Ara tracks, respectively. Finally, the -Ara tracks were subtracted from +Ara tracks strand-wise to obtain the enriched signal. The resultant tracks were further analyzed using custom Python scripts (https://github.com/sutormin94/TopoI_Topo-Seq).

3.22 Identification of EcTopoI cleavage sites (TCSs)

Separately for forward and reverse strands, the number of DNA fragments 3'-ends was calculated per position (N3E) based on read alignments stored in SAM files (giving N3E_F and N3E_R tracks respectively). The tracks were scaled by the total number of aligned reads to get normalized coverage across samples and biological replicates were averaged. After that -IP tracks (+Ara-IP and -Ara-IP) were subtracted from corresponding +IP tracks (+Ara+IP and -Ara+IP respectively) strand-wise resulting in +Ara and -Ara tracks. TCSs were detected as sites having a signal higher than threshold value 15 (values 10 and 20 were also tested and given similar results in respect to identified DNA motif).

3.23 Purification of EcTopoI

15 mL of a night culture of *E. coli* BL21(DE3) pET28 topA_strep in LB, supplemented with kanamycin 50 µg/mL, was added to 1.5 L of LB medium and cultivated at 37°C and 180 rpm. At OD_{600}~0.6 the culture was induced with IPTG (final concentration 0.4 mM) and cultivation continued for 4 h. The cells were pelleted (5000 rpm, 10 min, 4°C) and frozen in liquid nitrogen. Subsequently, the pellet was thawed on ice and resuspended in 5 mL of Lysis Buffer (100 mM Tris HCl pH 8.0, 150 mM NaCl, 1 mM EDTA) supplemented with 1 mM PMSF, lysed using incubation with 2 mg/mL lysozyme (1 h, 4°C) followed by sonication for 10 min. The obtained lysate was cleared by centrifugation (14000 rpm, 20 min, 4°C). Strep-tagged EcTopoI was purified on StrepTrap HP column (GE Healthcare). Briefly, the lysate was loaded on the 5 mL column pre-equilibrated with Lysis Buffer and the column was washed with 50 mL of Lysis buffer. Then proteins bound were eluted using 10 mL of Elution Buffer (50 mM Tris HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA, 2.5 mM desthiobiotin). The eluted proteins were analyzed using electrophoresis in 11% PAA gel, the bands were visualized with Instant Blue protein stain and subjected to protein identification via MS. 10 mL of the eluate was dialyzed overnight at 4°C against 1.5 L of the Storage buffer (10 mM Tris-HCl pH

7.5, 50 mM KCl, 0.1 mM EDTA, 1 mM 2-mercaptoethanol) and then concentrated 5x on Amicon Ultra-15 filter (Millipore) by centrifugation (3500 g, 20 min, 4ºC). Glycerol was added to 30% w/v and the sample was stored at -20ºC. The activity of the enzyme was confirmed by relaxation assay with a negatively supercoiled pBR322 plasmid.

3.24 Relaxation assay for EcTopoI

2 µL (330 ng/µL) of negatively supercoiled pMP1000 plasmid extracted from *E. coli* DH5α cells was combined with 2 µL of 10x Reaction buffer (100 mM Tris-HCl pH 8.0, 500 mM NaCl, 60 mM $MgCl_2$), different amount of EcTopoI (0, 50, 100, 200, 300, 400 ng of the enzyme with concentration 170 ng/µL), and milli-Q H_2O up to 20 µL. The mixture was incubated 30 min at 37ºC and the reaction was stopped by the addition of 5 µL of 5x Stop-buffer (50 mM EDTA pH 8.0, 0.5% bromophenol blue, 50% glycerol). Relaxation products were separated by electrophoresis in 1% agarose gel supplemented with 2 µg/mL chloroquine. After separation, for DNA visualization gel was stained with ethidium bromide.

3.25 Electrophoretic mobility shift assay (EMSA)

DNA fragments of *dps*, *potF* or *nuoN* were PCR-amplified from *E. coli* DY330 genomic DNA (for primers, see **Supplementary Table 1**) and purified by GeneJET Gel Extraction and DNA cleanup micro kit (PCR cleanup protocol, ThermoFisher). For a DNA-binding assay, 2.85 pmoles of DNA fragment was mixed with 0-16 pmoles of purified EcTopoI in 20 µL of Binding buffer-1 [221]. Reactions were incubated at 37ºC for 10 min, and the samples were separated in 10% PAGE in TAE buffer at room temperature at 100 V. For DNA visualization, the gel was stained with ethidium bromide.

3.26 *E. coli* RNA-Seq and data analysis

Total RNA was extracted from 2 mL of *E. coli* DY330 culture exponentially growing in LB to OD_{600}=0.6 or from 1 mL of a pre-stationary culture at OD_{600}=6 or from 0.5 mL of a stationary culture at OD_{600}=10 using ExtractRNA reagent (Evrogen). RNA samples were treated with DNase I (ThermoFisher Scientific) and purified by RNAClean XP beads (Beckman Coulter). Sequencing libraries were prepared without rRNA depletion using NEBNext Ultra II Directional RNA Library kit (NEB) with the following modifications: 10 min of fragmentation and 10 PCR cycles. The libraries were sequenced on HiSeq4000 instrument (Illumina, USA) with 50 bp long reads protocol. Library preparation and sequencing were performed at Skoltech Genomics Core Facility. RNA-Seq was performed in triplicate.

Reads were prepared and mapped to the reference genome as described above for WGS and ChIP-Seq. RSeQC package was used for FPKM and genes expression level calculations [222].

3.27 Meta-gene analysis

To produce scaled meta-gene plots, fold enrichment of a protein of interest or another signal (coverage depth, GC%, etc.) was extracted in vicinity of transcription units (15 kb upstream, TU body, 15 kb downstream) in concordance to their orientation. Regions were scaled to have the same number of positions (5000 bp) by omitting of randomly chosen points (if a region is longer than 5000) or by random duplication of points (if a region is shorter than 5000). Data extraction and scaling were performed with *FE_over_US_GB_DS.py* custom script. The resulting scaled arrays were averaged by position, smoothed with averaging sliding window 200bp and plotted with *Plot_signal_over_transcription_units.py* custom script. Zoom-in meta-gene plots representing the proximity of transcription start and transcription and sites (+300 bp:-200 bp and -200 bp:+300 bp, respectively) were produced without smoothing.

Normalization of ChIP-Seq and Topo-Seq data for the construction of joint metagene plots with several types of data or with data from different experiments was performed as following. For each nucleotide position of a metagene plot the z-score transformation was applied – a genome average was subtracted from the signal values and the result was divided by a genome standard deviation.

For TopoI Topo-Seq data, in respect to the data strands-specificity, signals for coding and template strands of TUs in a forward orientation were taken from N3E_F and N3E_R tracks respectively; signals for coding and template strands of TUs in a reverse orientation were taken from N3E_R and N3E_F tracks respectively (see *FE_over_US_GB_DS_strand_specific_binning_and_statistics.py* script in the https://github.com/sutormin94/TopoI_Topo-Seq repository).

3.28 Data and code availability

Custom scripts written for data analysis and data visualization for this work can be freely obtained from GitHub repositories (**Supplementary Table 2**). NGS datasets generated and analyzed in this work are listed in **Supplementary Table 3**.

Chapter 4. Results and discussion

To understand the role of topoisomerases in regulation of genome topology and, particularly, in dissipation of DNA supercoiling, a method for detection of topoisomerase activity is needed. In this section I introduce such a method, named Topo-Seq, and describe findings obtained with it for key bacterial topoisomerases - DNA gyrase, TopoIV, and TopoI. With Topo-Seq, the activity of topoisomerases was linked to transcription and replication genome-wide and topoisomerase DNA motifs were characterized.

Results of sections 4.1-4.8 are published in: **Sutormin D.**, Rubanova N., Logacheva M., Ghilarov D., Severinov K. Single-nucleotide-resolution mapping of DNA gyrase cleavage sites across the *Escherichia coli* genome. **Nucleic Acids Research**, Vol. 47(3), p.1373–1388, 2019

4.1 Topo-Seq allows precise localization of gyrase cleavage sites

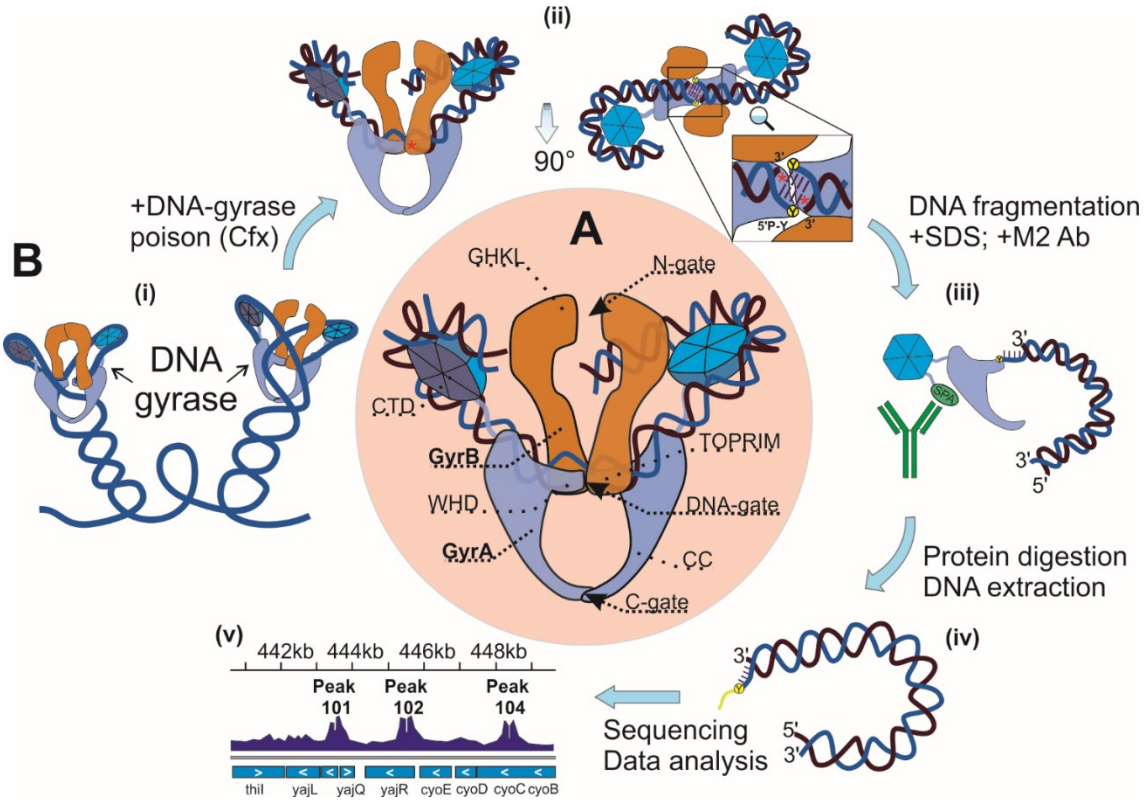
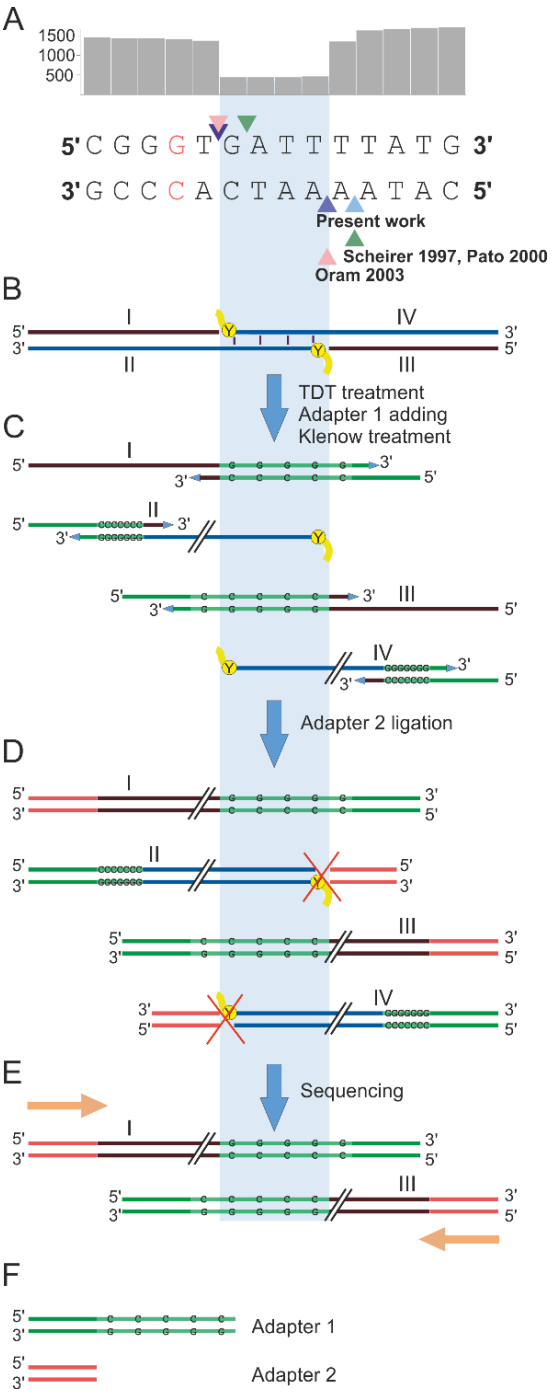


Figure 9. DNA gyrase structure and Topo-Seq procedure. (**A**) DNA gyrase structure with gyrase domains (GHKL, TOPRIM, WHD, CC, CTD) indicated. (**B**) Topo-Seq workflow: (i) Schematic illustration

of DNA gyrase on DNA plectonemes; (ii) A gyrase-DNA complex trapped with an inhibitor, here quinolone-like (red stars); (iii) C-terminal SPA tag is recognized by M2 antibodies and used to precipitate gyrase-bound fragments. For DNA fragment, polarity of strands is shown; (iv) Deproteinized DNA fragments have blocked 5'-ends; (v) Resulting signals visualized. Cfx – ciprofloxacin, SDS – sodium dodecyl sulfate, M2 Ab – monoclonal anti-FLAG antibody, SPA-tag – sequential peptide affinity tag



Many DNA gyrase (**Figure 9A**) inhibitors trap covalent cleavage complexes leading to accumulation of double-stranded breaks and, ultimately, cell death [223–226]. Purification of trapped complexes followed by deproteination and sequencing (or hybridization) of recovered DNA allows one to analyze the sites of gyrase cleavage (**Figure 9B**). Earlier such approach was widely used for mapping of particular cleavage sites [201,227–229] and for investigating relatively short DNA regions in detail [230–234]. Next-generation sequencing allowed genome-wide analysis: cleavage maps were generated for TopoIV in *E. coli* and Topo IIA in human cells [116,235]. As a consequence of trapping, several amino acid residues from topoisomerase active site remain linked to the 5'-ends of nucleic acid fragments after proteolytic treatment, resulting in poor adapter ligation efficiency [116]. To overcome this challenge, I developed a procedure, which we name Topo-Seq, that employs single-strand DNA sequencing protocol [236] to obtain information on DNA strands that remain free within the cleavage complexes. Topo-Seq allows us to get rid of modified DNA chains at the library preparation step (**Figure 10**).

Figure 10. DNA gyrase mechanism of action and sequencing libraries preparation protocol allows to identify cleavage sites with single-nucleotide resolution. (**A**) Coverage depth in the close vicinity of Mu SGS; cleavage site sequence is written under the depth track; known cleavage positions (with references) marked with colored

triangles. (**B**) Schematic view of double strand DNA fragment cleaved with DNA gyrase; 5'-ends blocked with peptides including catalytic Tyr^{122} (Y) labeled with yellow. (**C**) dsDNA is melted giving 4 chains. Terminal Deoxynucleotidyl Transferase (TDT) adds Gs to the free 3'-ends; single strand Adapter 1 aligns to the oligoG region with its oligoC track; Klenow fragment fills the gap to make dsDNA. (**D**) Double strand Adapter 2 ligates with free blunt ends. Ligation is inefficient when the end is blocked with peptide (in yellow). (**E**) Sequencing starts from Adapter 2 (arrows). (**F**) Legend for adapters

Thus, only free chains with 3'-ends directed towards the gyrase catalytic site are being sequenced, resulting in a specific structure of enrichment peaks at GCSs. Each of them should have a characteristic bimodal shape with a sharp 4-bp gap in the middle. Sequences to the left of the gap should align in a forward orientation with the GCS; their 5'-ends should vary, while 3'-ends should be identical and form the left "wall" of the central 4-bp gap. Sequences to the right of the gap should similarly align in reverse orientation.

To validate Topo-Seq, I constructed an *E. coli* strain with the Mu SGS strong gyrase binding site inserted into a non-essential region of the genome. Exponentially growing cells were treated with several DNA gyrase inhibitors - ciprofloxacin (Cfx), microcin B17 (Micro), and oxolinic acid (Oxo) and subjected to Topo-Seq. While a weak signal was observed in control untreated cells at Mu SGS, in the presence of the inhibitors there was a dramatic increase in the abundance of intermediate complexes, resulting in a strong signal (**Figure 11A**).

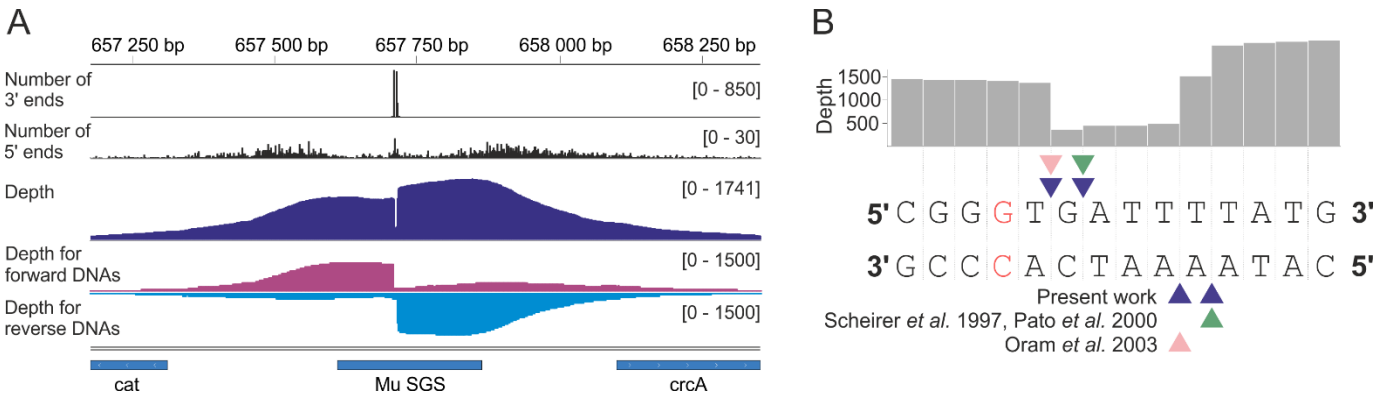


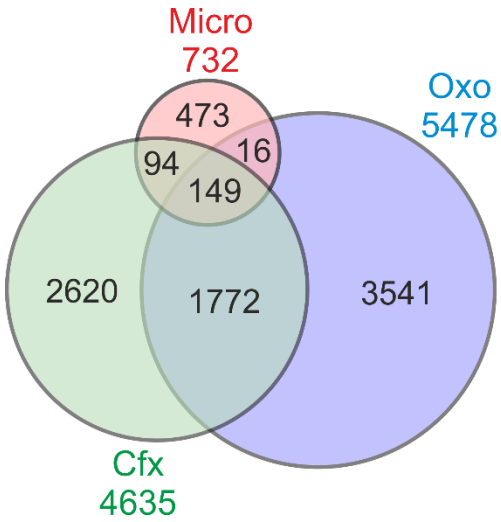
Figure 11. Signal structure at the strong gyrase binding site from bacteriophage Mu (Mu SGS). (A) Profiles of the number of 5'- and 3'-ends (N5E and N3E correspondingly) are shown in black. Total coverage depth (Depth) is in blue and coverages for DNA fragments that were aligned in forward and reverse orientations (F DNAs and R DNAs correspondingly) are in red and sky blue respectively. Tracks height (depth) are shown in brackets. The data visualized in IGV [208]. (**B**) Close-up of the cleavage site. Coverage depth around the site is shown as a grey track, local sequence lies below. Cleavage sites known from the existing literature are shown [201,229,237].

Concentrations of poisons used in Topo-Seq were chosen to be well above experimentally determined minimal inhibitory concentrations (MICs) for strains tested (300-625x for Cfx, 32-40x for Micro, and 120-188x for Oxo). The shape of a signal at Mu SGS fully matched the expectations based on gyrase catalytic mechanism and the sequencing protocol used. The positions of the 4-bp gap "walls" coincided with cleavage positions observed in previous biochemical studies of gyrase complexes trapped on Mu SGS *in vitro* [201,229,237] (**Figure 11B**). Thus, our procedure allows for accurate single-nucleotide identification of GCSs *in vivo.*

4.2 Thousands of DNA gyrase cleavage sites are distributed throughout the *E. coli* genome

I next investigated the global distribution of GCSs in cells treated with gyrase poisons. I used the hallmark 4-bp gaps between 3'-ends of Topo-Seq enriched DNA fragments to develop an automatic GCS-calling procedure. Plotting the number of 3'-ends (N3E) allowed us to globally identify pairs of enriched positions (gap walls) separated by 4-bp gaps. We interpret these signals as gyrase trapping sites; the heights of gap walls provide an estimate of the relative number of gyrase binding events that initiate the strand passage step at a particular site.

GCSs were detected as significantly enriched signals (statistical test from [209], p-value<0.0025) during the two-step normalization procedure and have passed additional filtering as being shared between at least two out of three biological replicas that were made for each gyrase inhibitor. In total, 4635 GCSs distributed throughout the genome were identified in Cfx-treated cells, 5478 in Oxo-treated, and 732 in Micro-treated cells. 41% of GCSs identified in the presence of Cfx are shared with Oxo set. The level of GCSs common for Micro and Cfx, or Micro and Oxo-treated cells is lower (33% and 23% of Micro GCSs, respectively), consistent with higher degree of similarity between Cfx and Oxo compared to a non-quinolone poison. We found that GCSs revealed simultaneously by several drugs tend to have higher N3E (**Figure 12**).



Set	N3E median	Score mean
Cfx only	7.6	2.3
Micro only	10.3	3.3
Oxo only	9.7	2.9
Cfx&Micro only	12.0	4.7
Cfx&Oxo only	13.3	3.9
Micro&Oxo only	11.5	4.6
Cfx&Micro&Oxo	20.1	5.5

Figure 12. **Cleavage sites simultaneously revealed by several poisons are stronger: have higher signals (N3E) and are high-scored**. Venn diagram representation of the relations between GCSs sets obtained with different gyrase poisons (on the left). Comparison of N3E medians and score means for different subsets of GSCs (on the right).

We also found that the number of GCSs depends on the concentration of the poison used: for example, using 30-50-fold excess of Cfx over MIC we got 50-250 GCSs, while 300-600-fold excess gave 6000-7000 GCSs. Similarly, a five-fold increase in Micro concentration resulted in the rise of the number of GCSs from 500-700 up to more than 3000. Observed proportionality indicates that increased drug concentration induces more stable cleavage complexes, which are detected by sequencing. Thus, the relatively low number of GCSs observed with Micro (732 GCSs at 30-40x MIC) is likely due to lack of drug saturation. Interestingly, compared to Oxo, Cfx is a much more effective cell growth inhibitor. Yet, judging by number of GCSs revealed by Topo-Seq, 120 μM Oxo (120-180xMIC) traps gyrase significantly more effectively than 10 μM Cfx (300-600xMIC). These observations seem to suggest that gyrase capture and subsequent cell death may be mediated by different processes. Indeed, differences in killing mechanisms between oxolinic acid and newer generations of quinolones have been reported [238].

A comparison of GCSs detected in strains with or without Mu SGS revealed no noticeable differences (apart from the expected strong gyrase enrichment at Mu SGS in the former strain) indicating that insertion of Mu SGS does not significantly affect genome-wide gyrase distribution.

When regions flanking Mu SGS and another strong GCS (site downstream of rRNA A operon) identified by our procedure were tested by Topo-qPCR (for Cfx), enrichment was detected, and extent of this enrichment corresponded to levels of Topo-Seq signals. In contrast, no enrichment was observed during

qPCR with primers specific for control sites where no gyrase cleavage was detected by Topo-Seq (Pearson's correlation: 0.94) (**Figure 13**). We conclude that our peak calling procedure is robust and reflects the *in vivo* positioning of gyrase intermediates trapped by the inhibitor.

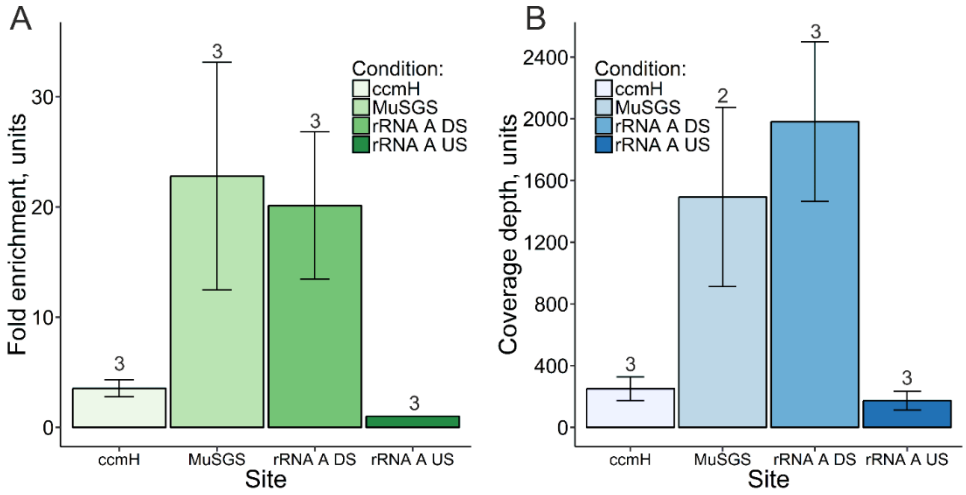


Figure 13. **Comparison of the results of Cfx-mediated Topo-qPCR and Topo-Seq**. Enrichment was estimated for four sites: Mu SGS, ccmH, rRNA A DS, and rRNA A US. Numbers above the error bars indicate the numbers of biological replicas, error bars constructed as +/- 2 standard errors. (**A**) Fold enrichment obtained by Topo-qPCR at genome sites indicated. (**B**) Average coverage depth for +Cfx+IP Topo-Seq track at genome sites indicated.

Topo-Seq data was compared with available for *E. coli* gyrase ChIP-chip data [65]. Despite the differences between the two approaches and the resolution of final data, we found a significant positive correlation between both datasets and a very similar distribution of signals over the *E. coli* genome (**Supplementary Figure 1**). Thus, gyrase binding and cleavage are likely to be connected, additionally validating the Topo-Seq methodology.

4.3 DNA gyrase has an extensive and degenerate binding motif

Single-nucleotide resolution GCS mapping allowed us to directly look for potential sequence preferences of the gyrase in antibiotic-treated cells. Aligning of sequences flanking the positions of GCSs revealed a significant deviation from random frequencies of nucleotides indicating the presence of a potential motif. When the obtained frequency matrix was converted to GC%, similar 130-bp degenerate pattern was observed for signals obtained from cells treated with each of the gyrase poisons tested (**Figure 14A**).

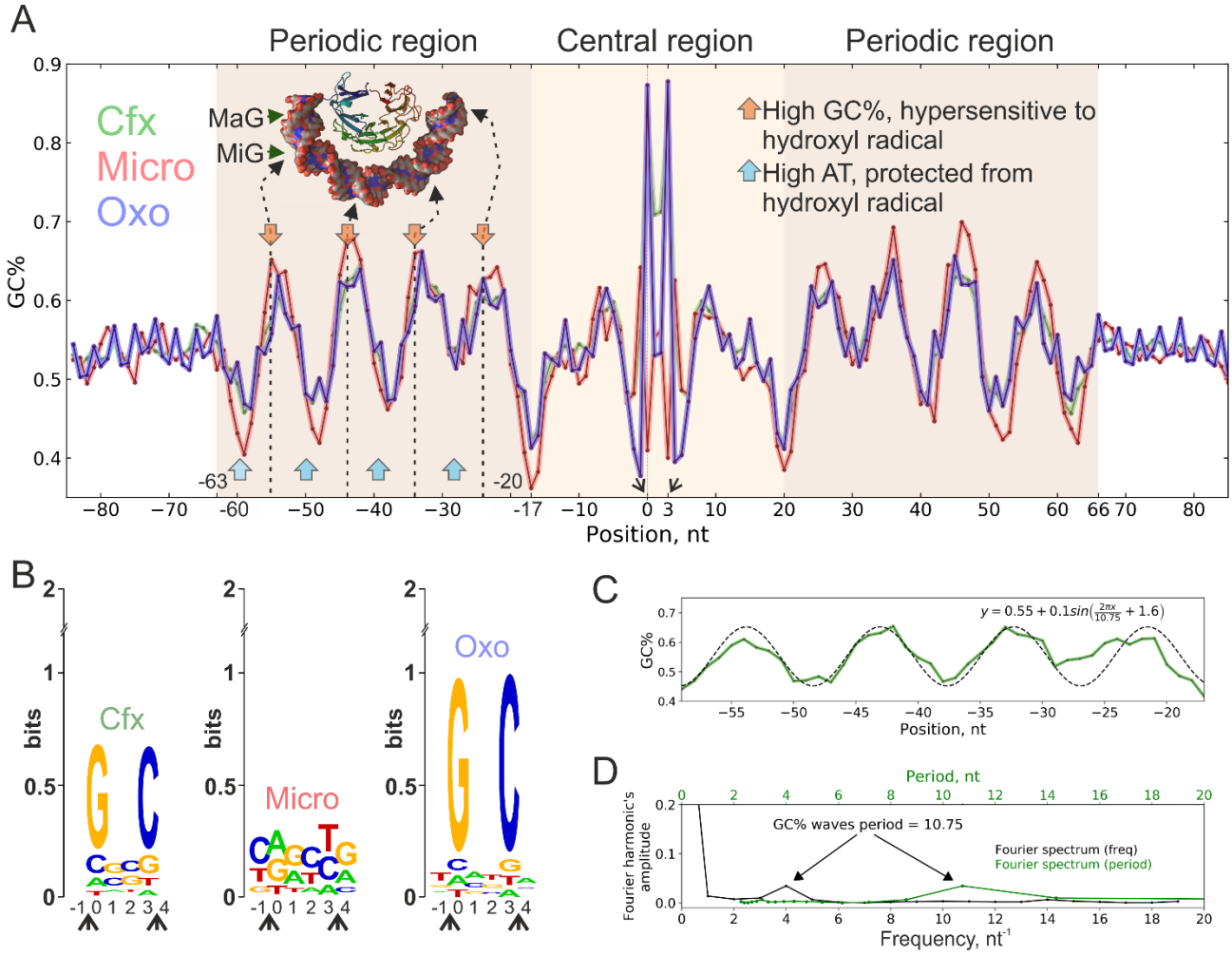


Figure 14. DNA gyrase has a binding motif revealed with a range of poisons: Cfx, Micro, and Oxo. (**A**) DNA sequences under GCSs were extracted and aligned; resulting motif, shown as a plot of GC content, has a central region (-16:19 nt) around the cleavage site, and two periodic regions (-63:-17 and 20:66 nt). Orange arrows indicate GC-rich and hydroxyl radical hypersensitive sites, blue arrows – AT-rich and protected regions. The structure of DNA-wrapped CTD is shown above the motif, dashed arrows point hypersensitive sites in DNA's minor groove. MaG – DNA major groove, MiG – DNA minor groove.

Hydroxyl radical sensitivity data is taken from [239]. (**B**) Logo representation of motif's central part around cleavage site. In the coordinates chosen, DNA gyrase cleaves forward chain between -1 and 0 and reverse chain between 3 and 4 bp (cleavage events are indicated with arrows). (**C**) Fitting of the motif GC% in the periodic region. (**C**) Fourier frequency spectrum for the gyrase motif periodic region. Local maximum of the spectra (4 nt^{-1} in 43 nt region or 10.75 bp) is marked with arrows.

The pattern is symmetrical with respect to the cleavage sites, which are located between -1/0 and +3/+4 positions of the motif. We refer to this pattern as the "gyrase motif". It consists of a central (from -16 to 19 bp) part containing the cleavage site and two flanking (from -63 to -17 bp and from 20 to 66 bp) regions, each with periodic (10.75 bp) changes in GC% resembling the binding pattern of eukaryotic nucleosomes [240] (see **Figure 14C**, **D**). Analysis of calculated DNA geometry for GCSs sequences using GBshape database [241] revealed noticeable deviations for such parameters as helix twist, propeller twist, and minor groove width. This observation may indicate that DNA recognition by the enzyme occurs through indirect readout [242].

The only significant differences between motifs obtained with different inhibitors were observed at the cleavage site (**Figure 14B**). GCS motifs obtained with Cfx and Oxo-treated cells confirmed the well-established tendency of quinolones to intercalate and facilitate scission before guanine residues [231,243,244]. In contrast, this pattern was not observed in the case of Micro, which likely has a different mode of interaction with gyrase and/or DNA [245], leading to a different consensus in the central part of GCS. This difference in cleavage preferences may explain a poor overlap between the Micro and quinolones GCSs sets (**Figure 12**). When sequences under the GCSs were screened for overrepresentation of known motifs with Tomtom [246], none were detected, suggesting that gyrase trapping is independent of other DNA binding proteins.

4.4 Gyrase activity correlates with sequence properties

To remove antibiotic-specific biases from the gyrase motif, the frequency values for central positions most affected by antibiotics (0-3 bp) were made equal to baseline nucleotide frequencies of the *E. coli* genome. The resulting position-weight matrix corresponding to a "combined" gyrase motif was used to scan the genome of bacteriophage Mu and the pBR322 and pSC101 plasmids. As expected, previously known strong gyrase sites from Mu, pSC101 and pBR322 [227,229,231] were identified as having the highest scores. When the genome of *E. coli* DY330 Mu SGS was scanned, Mu SGS had the 3rd highest score. 5 out of 13 highest scoring *E. coli* sites were among the GCSs identified by Topo-Seq, a highly significant (binomial test, p-value: 8.4e-14) overrepresentation. In addition, for 8 out of 13 highest scoring sites, there

was a total of 27 GCSs within 5 bp of their central regions providing independent support for the relevance of the motif (binomial test, p-value: 0) and suggesting some flexibility of the cleavage sites within the motif.

Analyzing the predictive power of gyrase motif score, we found that GCSs located in sequences with higher scores tend to have higher N3E values, an evidence that gyrase activity depends on DNA substrate sequence composition. Correlations between N3E and score are small, but statistically significant for all Topo-Seq experiments (Pearson's correlation: 0.22, 0.15 and 0.23 for Cfx, Micro and Oxo respectively with p-values<4.3e-5). N3E and score values for GCSs revealed by more than one poison are higher (**Figure 12**) meaning that 149 GCSs reproduced in experiments with all three drugs might be the most preferred gyrase sites in the genome.

4.5 Gyrase is attracted to the regions downstream of transcribed loci

Based on the "twin-domain model" and the fact that gyrase prefers positively supercoiled DNA to act upon, we expected the enzyme to be preferentially found downstream of transcribing RNAP [59,185]. Whole-genome analysis revealed that in general less GCSs associate with poorly transcribed transcription units (TUs) than with highly transcribed ones: 223 vs 552 GCSs for Cfx (t-test, p-value<2.5e-7), 35 vs 88 for Micro (t-test, p-value<1.6e-4), 339 vs 498 for Oxo (t-test, p-value<3.9e-3). The same conclusions were made with gene-centric analysis for poorly expressed and highly expressed genes: 569 vs 1128 GCSs for Cfx (t-test, p-value<4.9e-9), 88 vs 183 for Micro (t-test, p-value<3.0e-5), 815 vs 1113 for Oxo (t-test, p-value<1.7e-4) (**Figure 15 A-C**). Association and depletion of cleavage sites for highly and poorly transcribed TUs, respectively, is in line with observations made for *M. tuberculosis* gyrase [66]. We have screened the highly transcribed loci for their association with GCSs and found that cleavage sites significantly accumulate in extended downstream regions of active TUs (**Figure 15A-C**). The enrichment is most noticeable for rRNA operons, which have the highest transcription rate [247] (**Figure 15D**).

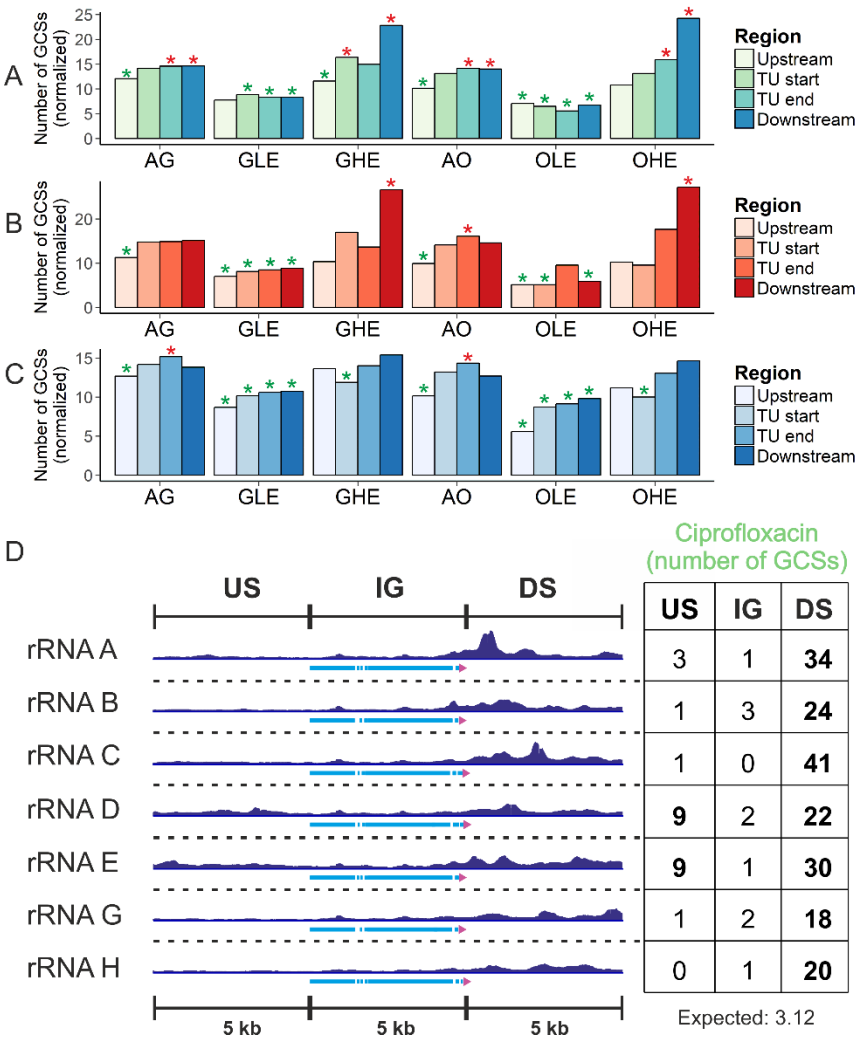


Figure 15. GCSs are associated with highly transcribed genes and accumulate at the ends of genes and downstream (DS) regions. (A) Ciprofloxacin-mediated Topo-Seq. AG – all genes (4240 genes), GLE – genes with low transcription level (379 genes with transcription level less than 0.95 units), GHE – genes with high transcription level (379 genes with transcription level more than 100 units), AO – all operons were analyzed (2330 operons), OLE – operons with low transcription level (194 operons with transcription level less than 1.36 units), OHE – operons with high transcription level (194 operons with transcription level more than 101 units). Upstream – upstream region, TU start – beginning of a transcription unit, TU end – end of a transcription unit, Downstream – downstream region. All regions have a constant length – 650 bp. Stars above the bars indicate statistically significant deviations in the number of GCSs (p-value<0.01) from the expected number. Green asterisks – regions with significantly lower GCSs than expected, red asterisks – significantly higher number of GCSs. (**B**) Micro Topo-Seq. (**C**) Oxo Topo-seq. (**D**) Enrichment of GCSs in downstream (DS) regions of rRNA operons revealed by Topo-Seq with Cfx. Coverage depth of the regions containing operons are shown on the left. Amounts of GCSs in upstream (US), operon bodies (IG) and DS are summarized in the table on the right. Numbers that are significantly exceeded expected values are in bold (binomial test, p-value<0.01).

To experimentally test whether transcription affects the gyrase distribution, we treated cells with RNAP inhibitor rifampicin (Rif) before Cfx-mediated gyrase trapping (the procedure further referred as RifCfx). While the overall number of cleavage sites dropped twofold after Rif treatment (4635 for Cfx vs 2355 for RifCfx, **Figure 16A**) their average strength (measured as N3E value) and sequence specificity (measured as gyrase binding score value) slightly increased (mean score 3.06 vs 3.49 for Cfx and RifCfx,

respectively). Rifampicin did not affect either overall shape of gyrase motif, or local cleavage properties characteristic of Cfx (**Figure 16B-C**).

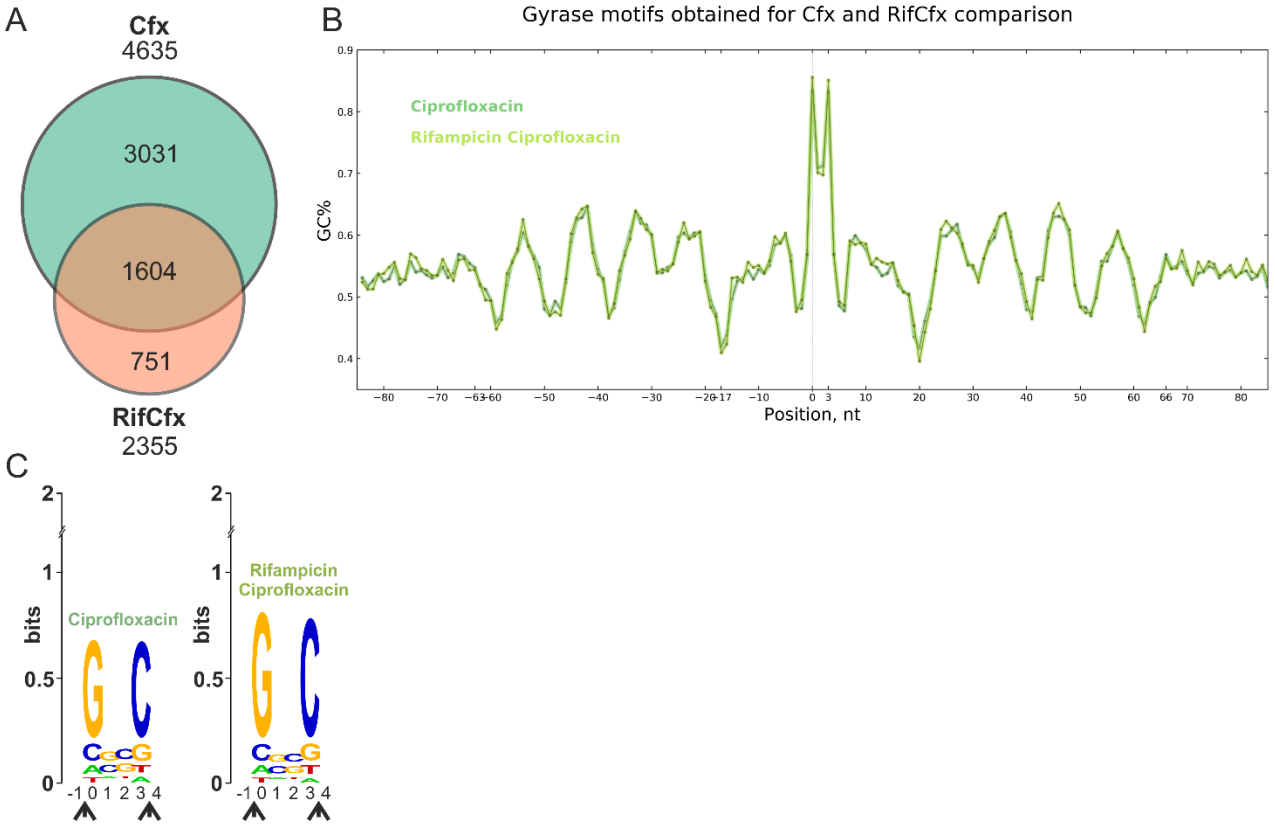


Figure 16. Effect of transcription inhibition by Rif on cleavage by gyrase. (**A**) Venn diagram showing overlap of Cfx and RifCfx GCSs sets. (**B**) DNA gyrase motif for Cfx (green) and RifCfx (light green). (**C**) Logo representation of motif's central part for Cfx and RifCfx.

Most notably, Rif eliminated both the accumulation of GCSs in the downstream regions of rRNA operons (**Figure 17B**, **C**) and avoidance of poorly transcribed TUs (**Figure 17A**). The GCSs which remained downstream of rRNA operons decreased their signals (**Figure 17D**, **E**). In contrast, signals from some other sites, potentially not directly related to transcription intensity, for example GCSs located inside repetitive sequences known as class 2 bacterial interspersed mosaic elements (BIME-2), were increased in the presence of Rif (**Figure 17E**).

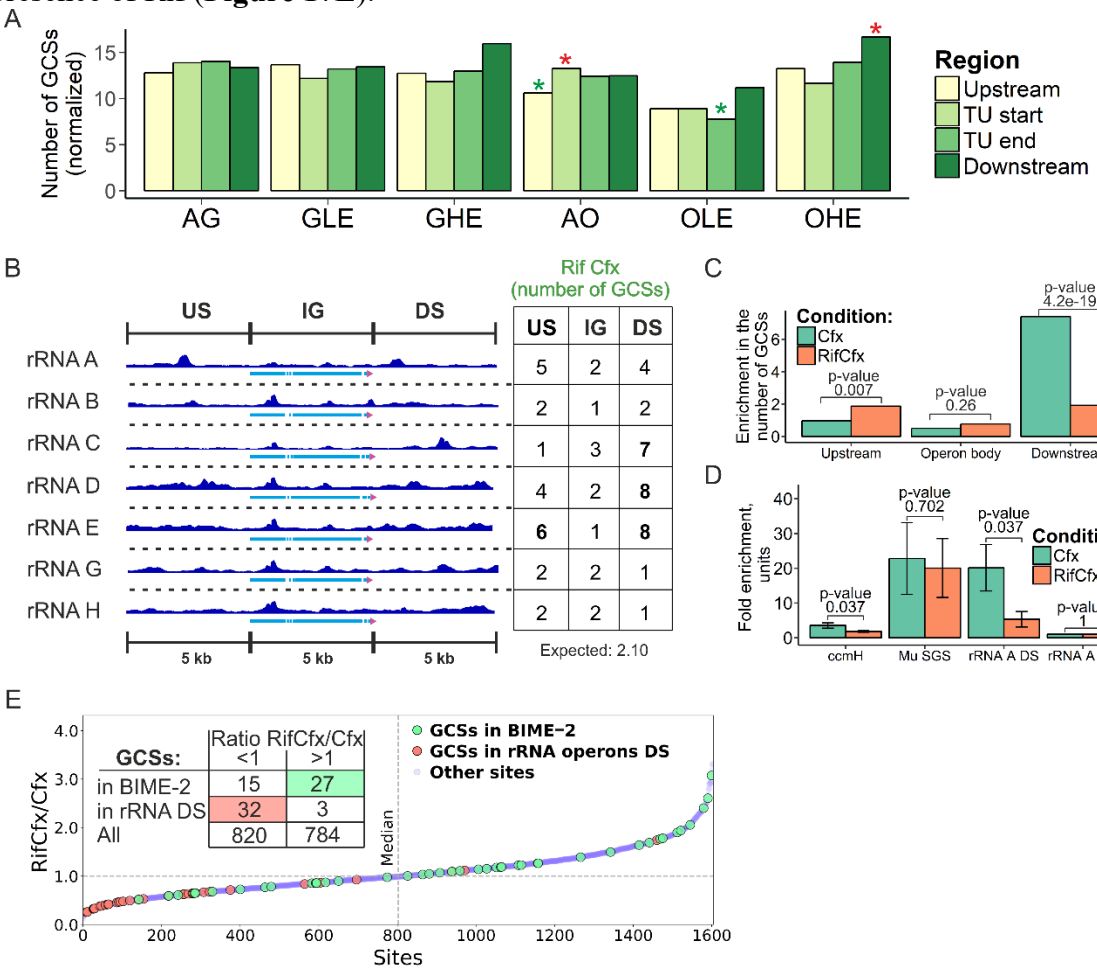


Figure 17. **Transcription facilitates gyrase activity**. (**A**) Number of RifCfx GCSs at Upstream, Downstream regions and at the beginning and ending segments of genes for different gene and operon sets, grouped by transcription levels (see Figure 7A for annotation). (**B**) Enrichment of GCSs in downstream (DS) regions of rRNA operons revealed by Topo-Seq with RifCfx. Coverage depth of the regions containing operons are shown on the left. See Figure 7D for annotation. (**C**) GCSs relocation from downstream regions of rRNA operons when transcription is inhibited with rifampicin. Plotted is enrichment in the number of GCSs which is a ratio of the number of observed GCSs to the number of expected GCSs (statistic - Fisher exact test). (**D**) Enrichment observed at several genomic sites by Topo-qPCR for Cfx and RifCfx (data for 3 replicas). Error bars constructed as +/- 2 standard errors, p-values for t-test for means are indicated above

the bar pairs. (**E**) GCSs shared between Cfx and RifCfx sorted by ratio of their N3E values. GCSs that fall in 5 kb downstream regions of rRNA operons (red dots) have a significant tendency to have lower N3E values when cells are treated with Rif (binomial test, p-value=0.015). On the other hand, GCSs that fall into REPs (green dots) reveal tendency to increase N3E under the same conditions (binomial test, p-value=4e-7).

Metagene analysis performed for different groups of transcription units (TUs) stratified by level of transcription revealed that gyrase is enriched downstream (DS) of active TUs and that enrichment level is roughly proportional to the transcription level. Gyrase enrichment is observed for more than 15 kb downstream of TUs presumably indicating a diffusion range for positive supercoiling over bacterial DNA. Consistently, treatment with Rif abolished gyrase enrichment confirming that gyrase is attracted by positive supercoiling (**Figure 18**).

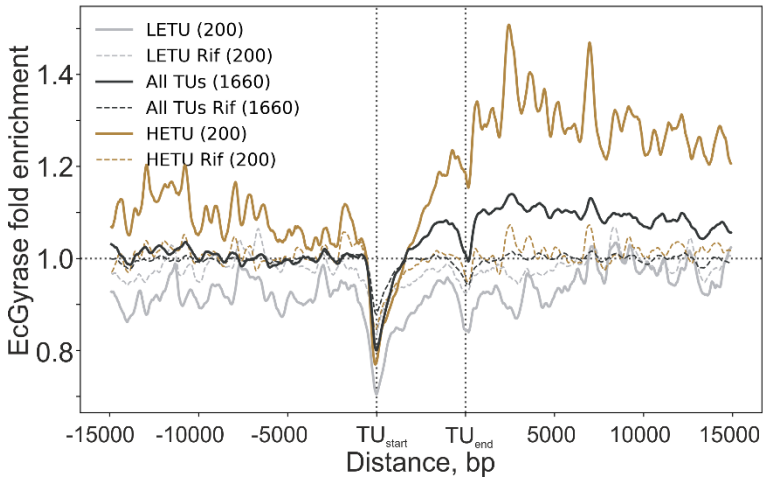


Figure 18. **Metagene plot showing DNA gyrase enrichment in TUs, their upstream and downstream regions**. Enrichment is shown for all TUs (black curve), highly-expressed (HETU, orange curve), and least-expressed (LETU, grey curve) sets. The number of TUs in each group is indicated in parentheses. Topo-Seq experiments with Cfx as a gyrase poison were used in the analysis. Data for gyrase under untreated and Rif-pretreated growth conditions are shown with solid and dashed curves, respectively. Gyrase fold enrichment is given relative to the input sample (+Cfx+IP/+Cfx-IP).

Overall, we conclude that transcription inhibition with rifampicin leads to significant relocation of active gyrase, suggesting that transcription is a strong factor that determines gyrase enrichment in downstream DNA.

4.6 GCSs are overrepresented in a subset of BIME-2 sequences

We found significant overrepresentation of GCSs in BIME-2 elements in full agreement with previous experiments [234] (binomial test, p-value: 0 for Cfx, 3e-4 for Micro, and 1e-9 for Oxo). In particular, BIME-2 located between *sucC* and *sucB* and between *tldD* and *yhdP* genes house multiple GCSs (6 and 35 GCSs for Cfx, respectively). Closer look revealed that gyrase much more frequently cleaves y-type than z2-type repetitive extragenic palindromes (REPs) that form BIME-2 in agreement with old observations [233] and especially prefers loops and non-complementary regions within the stems of cruciform-forming REPs (**Figure 19**).

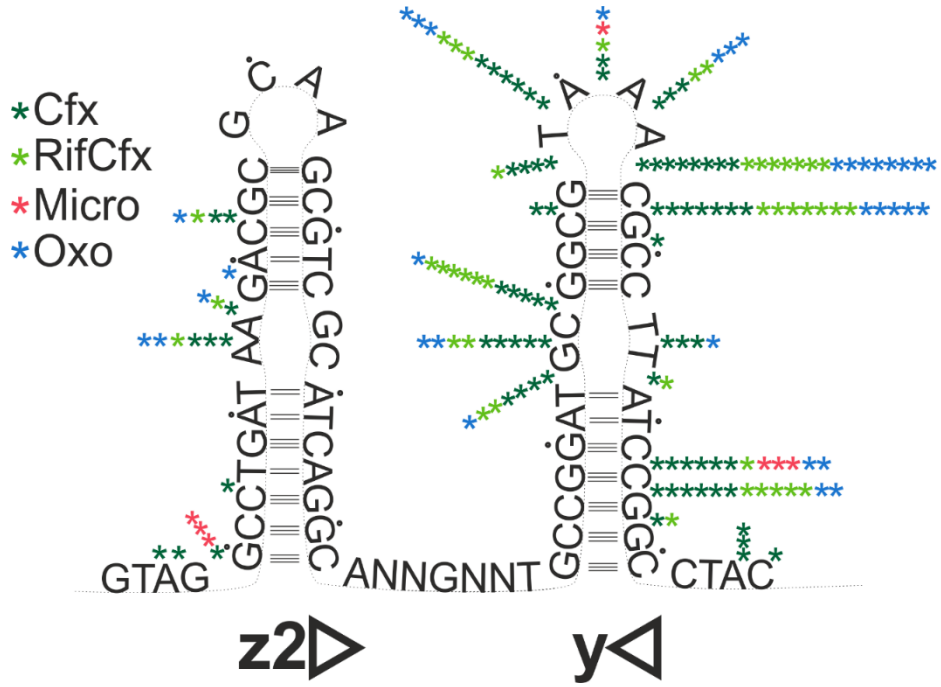


Figure 19. Frequencies of GCSs within schematic BIME-2 region that consists of one z2-type and one y-type REP. 13 BIME-2s with at least 3 GCSs were analyzed. REPs are shown as cruciforms, GCSs indicated as asterisks, color of asterisk corresponds to Topo-Seq condition when a GCS was observed.

The number of GCSs-containing REPs within particular BIME-2 elements considered by Espeli and Boccard [234] is also similar to the previously observed number of gyrase-generated cuts. Thus, DNA gyrase binds at least some BIME-2s *in vivo*. BIME-2 locate in intergenic regions and accumulation of GCSs in them could be due to positive supercoiling associated with transcription of adjacent genes. In this case one would expect the gyrase signal to decrease when transcription is inhibited by Rif: however, the opposite is observed (**Figure 17E**), indicating that BIME cleavage by gyrase is transcription independent. Correspondingly, BIME-2 have a higher mean score than expected (-1.3 vs -2.3 over the *E. coli* genome; t-test, p-value: 0), which may be a reason of their preferential binding by gyrase and recognition by gyrase in

Rif conditions. Thus, BIME-2 have properties of transcription-independent strong gyrase sites and might contribute to genome organization.

4.7 GCSs colocalize with MukB and avoid H-NS binding regions

Nucleoid-associated proteins (NAPs) contribute to genome organization and local topology. High-resolution maps of the binding sites of *E. coli* NAPs Fis, IHF, H-NS, MatP, and MukB are available [134,248,249]. Genome-wide cleavage data are also published for *E. coli* TopoIV, which is a close paralog of DNA gyrase [116]. We compared available datasets with our Topo-Seq data to find potential associations between gyrase activity and these proteins. We found that GCSs are slightly underrepresented at IHF sites and at Fis sites in Cfx (+/-Rif) and Oxo treated cells (binomial test, p-value<0.002) but not in Micro-treated cells. GCSs are strongly underrepresented at H-NS binding regions (binomial test, p-value<9.75e-23) for all Topo-Seq experiments. H-NS, a well-known transcription repressor, primarily associates with silent regions of the genome [249]. Hence, H-NS occupied areas are expected to lack transcription-mediated positive supercoiling. Sequences of H-NS occupied areas also score lower than the genome mean for gyrase motif (-4.4 vs -2.3, t-test p-value<2.2e-308). A combination of these two factors may jointly contribute to gyrase avoidance.

Due to similarities in structure and mechanism of TopoIV and gyrase action, one could expect a correspondence between their cleavage sites distribution. However, our analysis did not reveal any significant association, either positive or negative, between GCSs and TopoIV cleavage sites. TopoIV activity and distribution within a cell are thought to be positively connected to MukBEF – a complex involved in structural maintenance of chromosomes [250]. In turn, MukBEF is known to physically interact with the MatP protein [134]. MatP binds with high affinity to specific *matS* sites that concentrate in the Ter macrodomain [133], and displaces the MukBEF complex [134,135]. Together, the TopoIV-MukBEF-MatP system is thought to coordinate the proper timing of replicating chromosome unlinking and segregation [134]. We found that GCSs are overrepresented at MatP-occupied regions (binomial test, p-value<0.006) but not in the immediate vicinity of *matS* sites. Strikingly, the enrichment is even more pronounced for MukB sites (binomial test, p-value<3.9e-14). Moreover, the gyrase binding motif score of MukB binding regions is relatively high (-1.5 vs -2.3 genome average, t-test p-value≈0). Overall, we suggest that DNA gyrase likely acts independently of the MatP-MukBEF-TopoIV decatenation ensemble, but the MukBEF complex might synergize with the gyrase downstream of active TUs by stabilizing plectonemes induced by transcription-dependent positive supercoiling.

4.8 Lack of association between GCSs density, topologically associated domains (TADs), and sites of spontaneous mutations

DNA gyrase supercoiling activity might contribute to bacterial genome structuring [88,251]. We therefore looked for association of GCSs with TADs, as defined for *E. coli* MG1655 [135]. However, no significant over- or underrepresentation of GCSs were found either in TADs, the inter-TADs, or nearby the TADs borders. Gyrase sites on DNA can also be expected to have an increased rate of spontaneous mutations due to association with double-strand breaks introduced by gyrase. We compared the distribution of spontaneous mutations throughout the *E. coli* genome (for wild-type and mutagenic *mutL* cells [252]) and found no pronounced associations between mutations and GCSs.

4.9 Topo-Seq allows mapping of GCSs in plasmids *in vivo*

Well characterized strong gyrase sites are known for several plasmids including pSC101 and pBR322 [229]. To test whether Topo-Seq is capable for mapping GCSs in plasmids, I performed Topo-Seq with *E. coli* DY330 *gyrA-SPA* strain transformed with pBR322. To increase sensitivity, I treated cells with increased, 100 μM, concentration of Cfx.

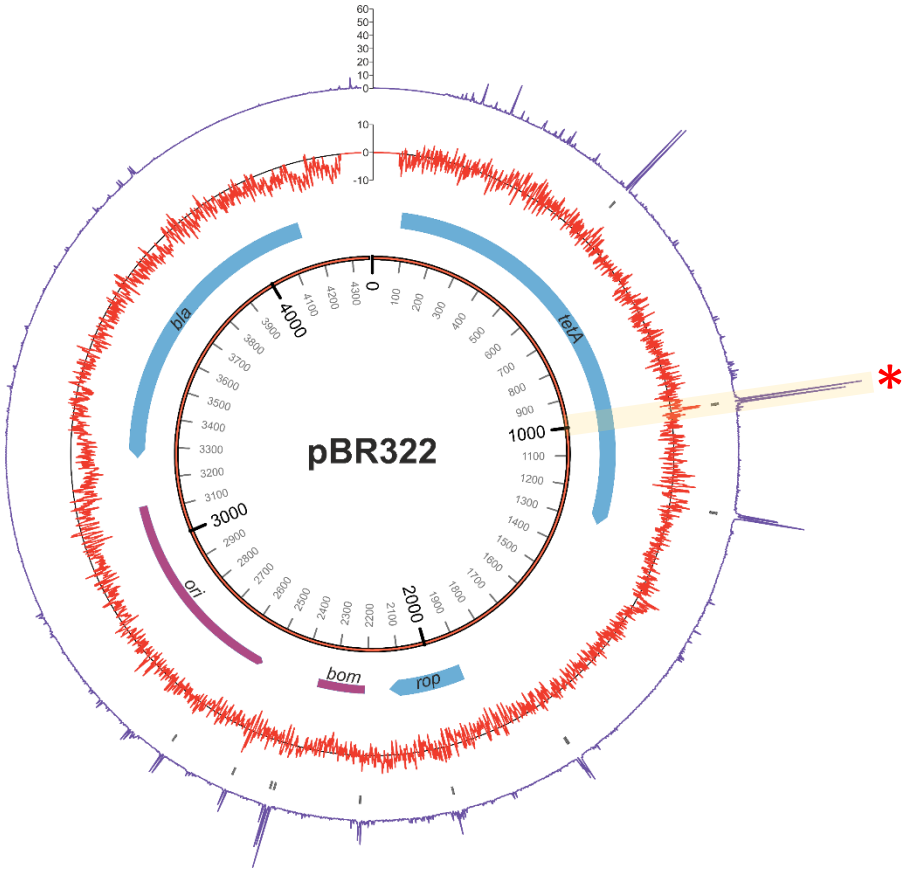


Figure 20. **GCSs revealed in pBR322 plasmid *in vivo***. Gyrase enrichment is shown in purple (outer circle). Sequence score weighted by gyrase binding/cleavage motif is shown in red. Detected GCSs marked with

grey boxes. GCS in the strong gyrase site with a top-scored sequence marked with a red asterisk. Map was constructed with Circos [253].

Topo-Seq revealed 12 GCSs in the pBR322 plasmid (**Figure 20**). As predicted by scanning pBR322 sequence with gyrase motif, known strong gyrase binding site with a high-score sequence (score 12.9) hosted a GCS with the highest enrichment level across all GCSs detected in the plasmid. Coordinate of this GCS matched the cleavage site identified earlier using oxolinic acid [231,232]. We conclude that Topo-Seq can be used for high-throughput mapping of GCSs in plasmids, and that strong gyrase sites can be predicted with a gyrase motif.

4.10 Gyrase activity follows transcription at different growth stages of a cell culture

To systematically track gyrase activity in different growth stages, I performed Topo-Seq using Cfx with exponentially growing *E. coli* DY330 *gyrA-SPA MuSGS* (EP, OD_{600}~0.6, 2 h after inoculation), stationary culture (SP, OD_{600}~10, 14 h after inoculation), and a culture transitioning from exponential to stationary stages (ESP, OD_{600}~6, 6 h after inoculation) (**Figure 21A**). To complement Topo-Seq data, transcription was assessed with RNA-Seq at selected time points.

Interestingly, for ESP and SP stages an increased number of GCSs was observed relative to EP stage and EP culture treated with Rif (**Figure 21B**). Comparison of GCSs sets indicated that ESP dataset is closer to SP than to EP (**Figure 21C**). Concordantly, by RNA-Seq, transcription levels of TUs at SP and ESP were well-clustered and differs from EP data (**Figure 21D**).

To investigate, whether changes in gyrase cleavage site localization at different growth stages reflect changes in a transcriptional profile, I stratified transcription units by changes in a transcription level and calculated the average number of GCSs in DS regions (5 kb). Enrichment of GCSs was increased for TUs with increased transcription level and decreased for TUs with decreased transcription level (p-value < 0.05, t-test), supporting the idea that gyrase has higher affinity to positively supercoiled DNA and is associated with active transcription (**Figure 21E**). For example, activation of *gad*-regulon genes (*gadC/B* and *gadA/gadE*) in ESP and SP stages coincides with a dramatic increase in gyrase-mediated cleavage downstream of these genes (**Figure 21F**).

Analysis of gyrase motifs reveled no differences among growth stages (**Figure 21G**) indicating that despite transcription-mediated alterations in GCSs positioning, gyrase binds and cleaves specific sequences “activated” by positive supercoiling.

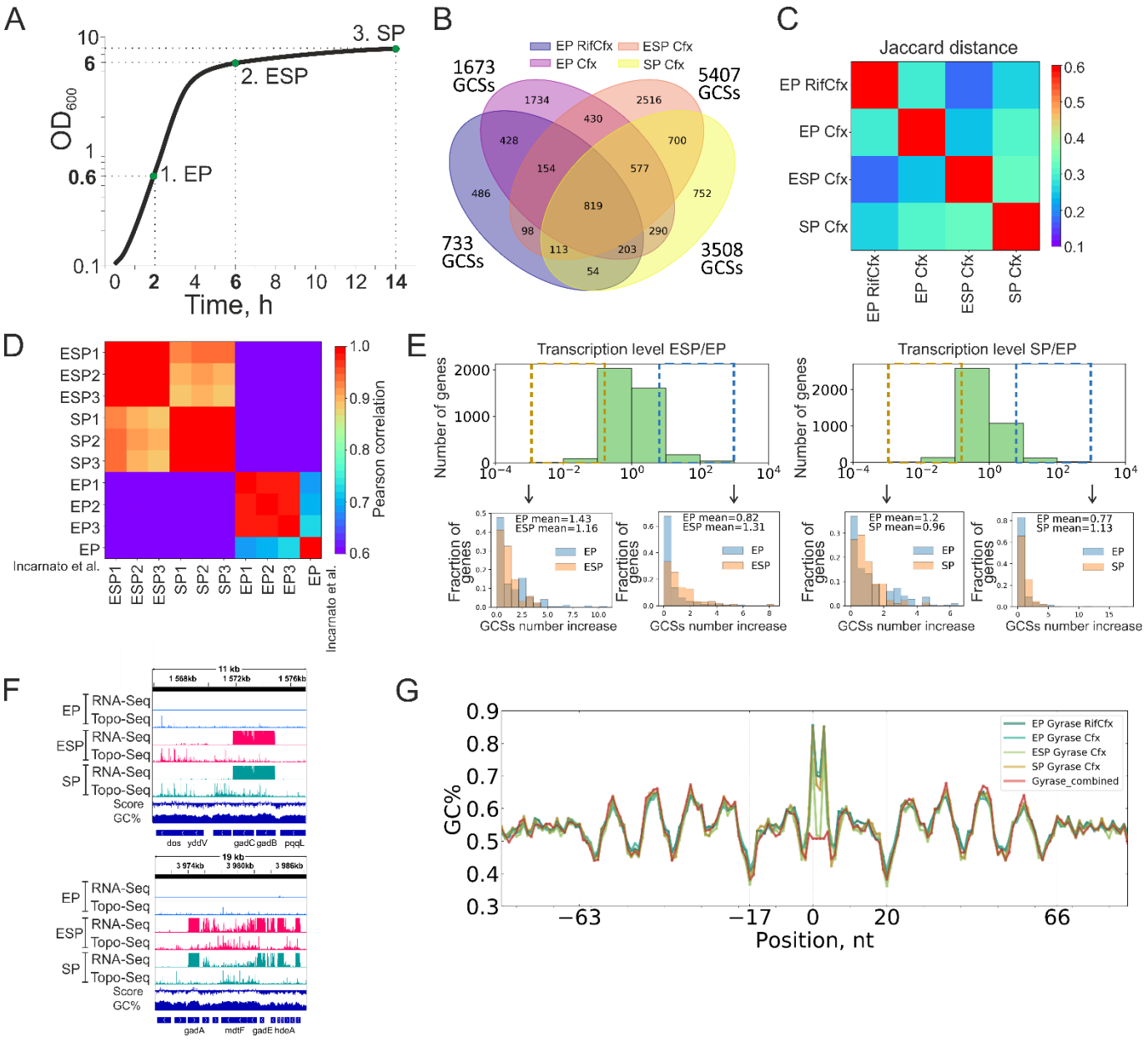
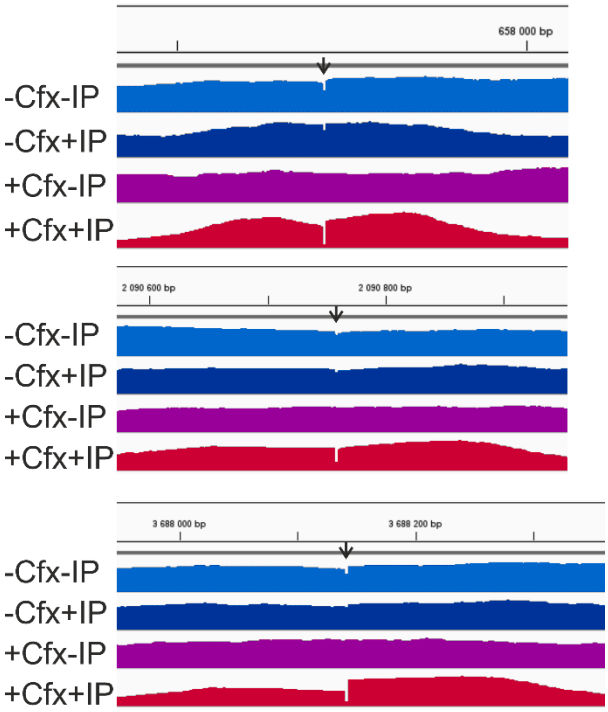


Figure 21. **DNA gyrase activity in different growth stages of *E. coli* culture**. (**A**) Time points at which *E. coli* cells were collected for Topo-Seq and RNA-Seq. EP – exponentially growing culture, SP – stationary culture, ESP - culture transitioning from exponential to stationary stages. (**B**) Venn diagram of GCSs sets for different growth stages. (**C**) Jaccard distances between GCSs sets. (**D**) Pearson correlations among RNA-Seq datasets (for genes). (**E**) Left - genes which have increased transcription levels at ESP relative to EP stage tend to have more GCSs in DS regions (5 kb), genes which have decreased transcription levels at ESP relative to EP stage tend to have less GCSs in DS regions (5 kb). Right – the same analysis but for SP and

EP datasets. (**F**) Representative genomic regions demonstrating changes in transcription levels and gyrase-induced cleavage between growth stages. (**G**) DNA gyrase motif for different culture growth stages.

4.11 Gyrase resides at strong genomic sites in a stationary phase culture



In Topo-Seq data for SP culture, I observed characteristic, gyrase-induced, cleavage of DNA in samples not treated with gyrase poison (**Figure 22**). Cleavage was much less intensive than for Cfx treated and enriched samples and detectable only for high-scored sites. Interestingly, no cleavage in non-treated samples was observed for exponentially growing culture. This indicates that gyrase naturally resides at some strong binding sites in stationary stage culture. I suggest that absense of replication and less intensive transcription in SP might increase gyrase residence time on DNA.

Figure 22. **Gyrase resides at strong binding sites in stationary phase**. Representative genomic regions with GCSs (cleavage sites marked with black arrows): 657709 [MuSGS] (score 13.3), 2090756 (score 7.2), 3688139 (score 9.1). All tracks comprising a Topo-Seq experiment are shown.

4.12 Gyrase follows replication forks as demonstrated by time-resolved Topo-Seq

Azamat Gafurov contributed to the results presented in this section. Azamat assisted with samples preparation during the time-course experiment.

As expected by twin-domain model, moving DNA-polymerase produces ahead of itself positive supercoiling during DNA replication [254]. DNA topoisomerases, gyrase and TopoIV, are responsible for supercoiling relaxation, promoting DNA replication [115]. To directly observe localization of gyrase activity in the *E. coli* genome during replication, I performed a time-resolved series of gyrase Topo-Seq experiments for synchronously replicating *E. coli* DY330 *gyrA-SPA Mu SGS* (**Figure 23A**). To synchronize cells' replication and division, a stationary-phase method was used. Briefly, replication was started in a pre-

stationary culture (OD_{600}~6) by dilution of the culture 10 times with fresh LB medium [211]. Samples were collected before replication initiation (-3 min time-point), when culture was diluted (0 min time-point) and every 5 min upon dilution (5-30 min time-points) (**Figure 23A**).

Analysis of sequencing data revealed that replication initiation in the cell population was not 100% synchronous. A fraction of cells (~15%) initiated replication in 10 min after dilution (early replicating cells) and after 30 min nearly all cells initiated replication of DNA (**Figure 23B**). By comparison of coverage depth for +Cfx-IP samples frame-by-frame, replisome locations in early replicating cells were inferred. A single replisome speed deduced from this data was 41.5 ± 1.9 kb/min or 690 ± 30 nt/s (± standard deviation) (**Figure 23C**), which is in agreement with earlier reports based on DNA combing - 653 ± 9 nt/s (± SEM) [255] and DNA-microarrays – 750 nt/s [115].

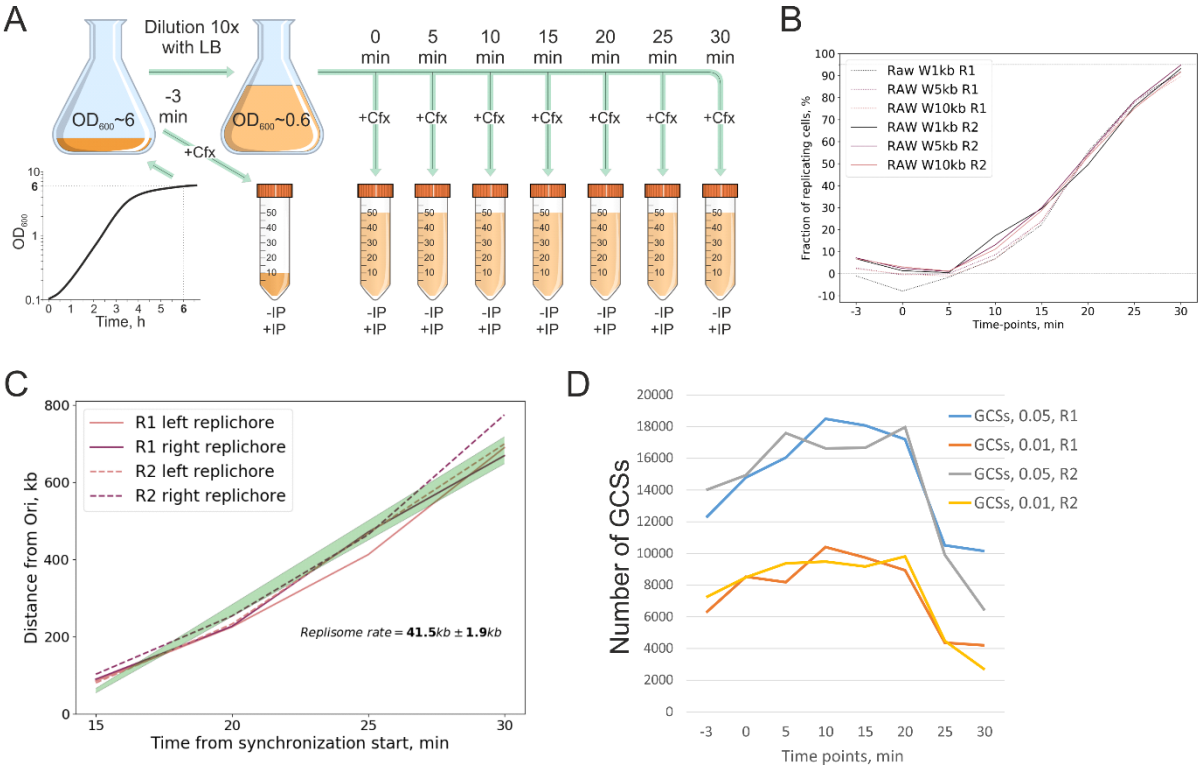


Figure 23. **Gyrase time-course Topo-Seq experiment**. (**A**) Overview of *E. coli* cells replication synchronization procedure and time-points collection for Topo-Seq. (**B**) Fraction of cells initiated replication at a given time-point. Data is shown for two biological replicates (R1 and R2) and for different track smoothing conditions (1, 5, 10 kb smoothing windows, respectively). (**C**) Determination of replication speed for two biological replicates (R1 and R2) and two replichores (left and right) separately. Green area marks the confidential interval for the approximation with a linear function. Linear regression equation is shown on the plot. (**D**) Number of GCSs called at different time-points. Data is shown for two biological replicates (R1 and R2) and for two GCSs calling conditions (0.01 and 0.05 statistical cutoffs, respectively).

GCSs calling revealed that number of cleavage sites increases for the first 10 min after replication was induced by dilution with fresh medium (when first cells initiate replication – **Figure 23B**) and decreases after 25 min (when over 80% of cells were already replicating DNA) (**Figure 23D**). I propose that replication initiation and/or progression of replisomes close to the origin region may be associated with increased gyrase activity likely via increased positive supercoiling generated. Interestingly, gyrase enrichment dramatically increases over the whole genome at a time-point 5 min and remains high at later time-points (**Figure 24A**). I suppose that in 5 min upon addition of a fresh medium, cells response to the flush of nutrients by increasing level of metabolism, produce more ATP and/or activate transcription, which both increase gyrase activity.

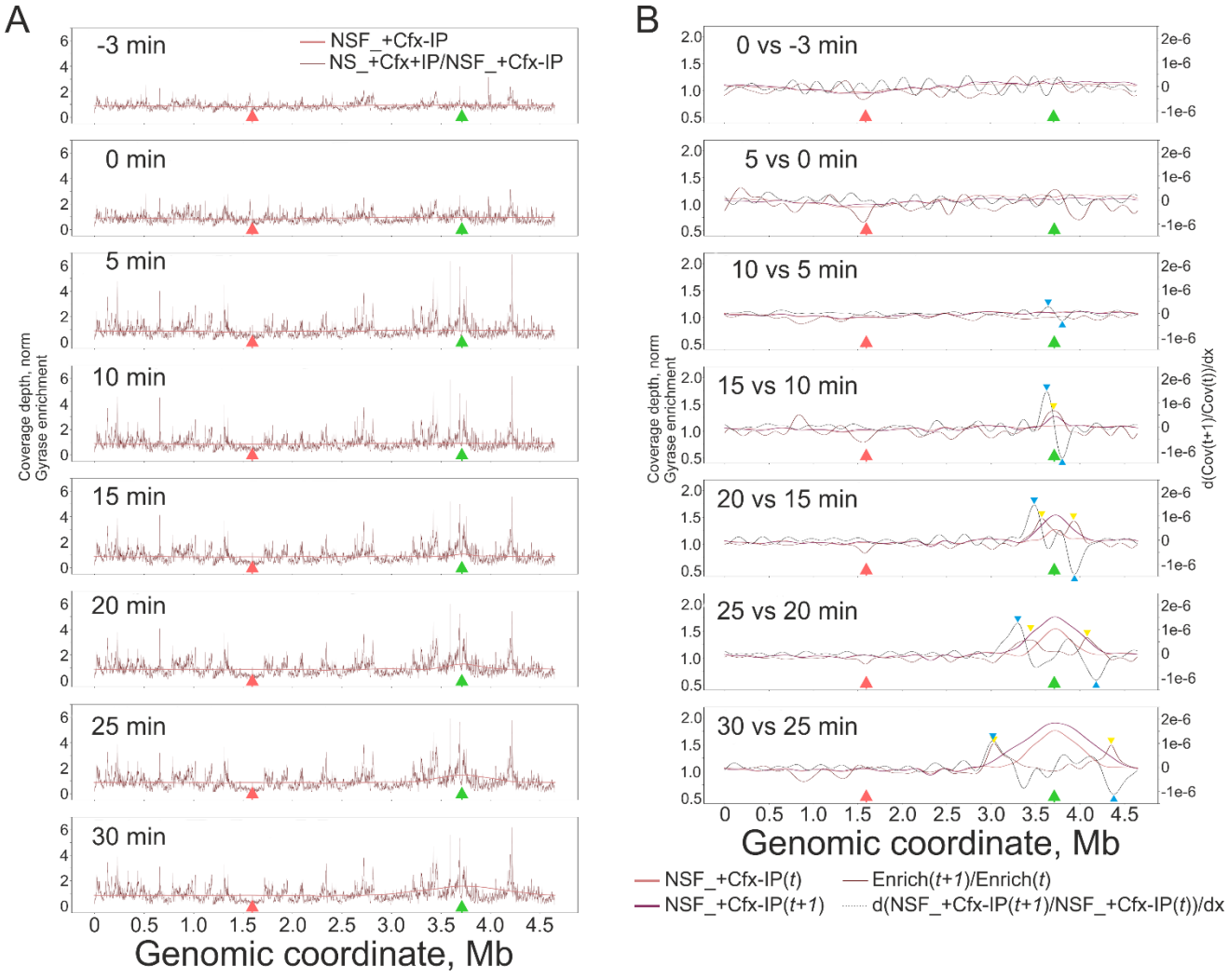


Figure 24. **Colocalization of gyrase activity and replication**. (**A**) Smoothed genome coverage depth (NSF_+Cfx-IP) and non-smoothed gyrase enrichment (NS_+Cfx+IP/NSF_+Cfx-IP) for all time-points of the experiment. Red and green arrows indicate replication terminator and origin, respectively. (**B**) Tracking

of the early initiated replisomes (blue triangles mark global minimum and maximum of the function 1 – see Material and Methods) and emerging gyrase enrichment (yellow triangles mark two local maxima of the function 2). Red and green arrows indicate replication terminator and origin, respectively.

Average distance between leading replisomes in early replicating cells and emerging replication-induced gyrase enrichment was -15 ± 30 kb (± standard deviation) (**Figure 24B**), indicating colocalization, which is in accordance with gyrase attraction by positive supercoiling generated by moving DNA polymerase. Recent studies involving single-molecule tracking microscopy of DNA gyrase revealed colocalization between replisome and DNA gyrase clusters comprising of ~12 molecules. Interestingly, average distance between replisome and a cluster was 135 ± 14 nm (± SEM) suggesting that gyrase operates at some distance from replisome, probably, to avoid collisions [256]. I conclude that by Topo-Seq data, DNA gyrase colocalizes with replisomes where, presumably, relaxes positive supercoiling generated by DNA unwinding.

4.13 Identification of GCSs in *Caulobacter crescentus* genome

Monica Guo contributed to the results presented in this section; Topo-Seq experiments were conducted at Mike Laub laboratory at MIT, during my visit in 2018.

To investigate localization of gyrase activity on a different background than *E. coli*, Topo-Seq was conducted with *C. crescentus* (class Alphaproteobacteria). First, chromosomal *gyrA* gene was fused with a sequence encoding 3xFLAG tag. Second, an efficient gyrase poison was selected to trap gyrase on DNA by MIC measurements on solid and in liquid media (**Figure 25A**). To increase solubility, Cfx stock solution was prepared in 0.1 mM HCl (30 mM Cfx, pH~4), and oxolinic acid stock solution was prepared in 5 mM NaOH (3 mM Oxo, pH~11.5). To control for toxicity of high and low pH, MICs of HCl and NaOH were also determined: 1.5 mM and >5 mM for HCl and NaOH, respectively. Cfx had a much lower MIC compared to Oxo (5.9 μM vs 95 μM, respectively) and was used in further Topo-Seq experiments in a concentration 1.5 mM (~256*MIC). Concentration of the acid in the 1.5 mM Cfx working solution is significantly lower than HCl MIC (5 μM vs 1.5 mM) and, therefore, likely does not introduce bias. Third, conditions of cells sonication were adjusted to reach the optimal level of DNA fragmentation for IP and sequencing library preparations **Figure 25B**.

Topo-Seq experiments performed in two biological replicates revealed 43 GCSs, which is much lower than is usually detected for *E. coli* (several thousand), likely indicating the insufficient concentration of Cfx. The amount of detected GCSs was not adequate to reveal the gyrase motif. However, the majority

of detected GCSs were located in the downstream vicinity of two rRNA operons in the *C. crescentus* genome (**Figure 25C**). This observation confirms the role of gyrase in the association with positive supercoiling, generated by transcription.

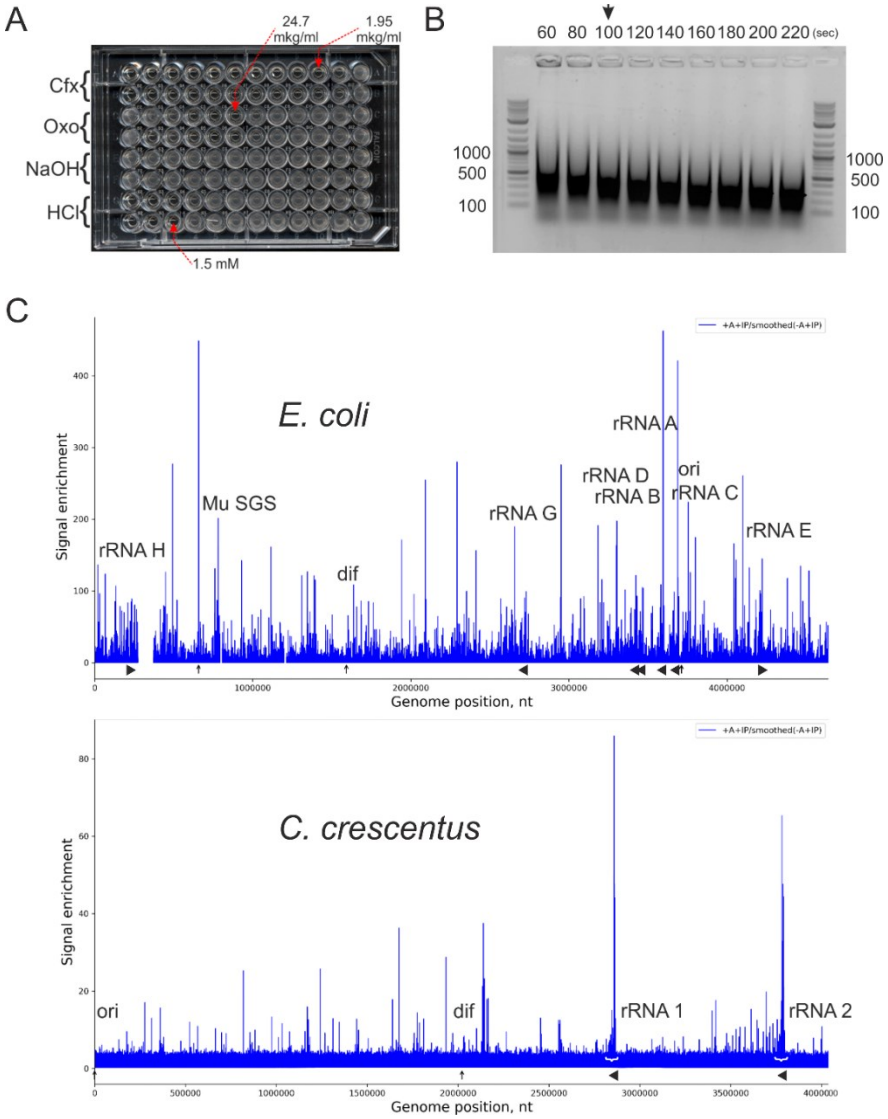


Figure 25. **Mapping gyrase cleavage sites in *C. crescentus* genome with Topo-Seq**. (**A**) MIC determination of Cfx, Oxo, sodium hydroxide and hydrochloric acid for *C. crescentus* by serial dilution method. Minimal concentrations for which growth inhibition was observed are marked with red arrows. (**B**) DNA fragmentation range as a function of sonication time (20 s pulse, 20 s relaxation, power 30%, Virsonic Virtis). (**C**) Gyrase enrichment is shown for a typical Cfx-based Topo-Seq experiment with *E. coli* (upper panel) and for *C. crescentus* (lower panel). Positions of *ori*, *dif* and Mu SGS sites are shown as vertical black arrows. Positions and orientations of rRNA operons are shown with black triangles. For *C. crescentus*, GCSs located in downstream regions of rRNA operons marked with white brackets.

4.14 Mapping of TopoIV cleavage sites in *E. coli* genome with a single-nucleotide resolution using Topo-Seq

TopoIV is another type IIA topoisomerase abundant in bacteria. Despite homology to gyrase, structural and mechanistic similarities, it is thought to be mainly involved in DNA decatenation and not in relaxation of supercoils [99,101]. Earlier, TopoIV binding and cleavage sites were determined in *E. coli* genome with ChIP-Seq and NorfIP [116]. The latter approach, similarly to Topo-Seq, relies on topoisomerase poison to trap cleavage complexes, enrich them and then sequence DNA attached to TopoIV. However, NorfIP was lacking a single-nucleotide resolution due to a library preparation step based on a standard ligation protocol. To overcome this limitation, I decided to apply Topo-Seq to map TopoIV cleavage sites in a genome of exponentially growing *E. coli* cells carrying chromosomal *parC* (encodes ParC subunit of TopoIV) fused with sequence encoding SPA tag.

Topoisomerase cleavage complexes were trapped with a TopoIV poison Cfx and ParC subunits were purified by SPA tag (**Figure 26A**, verified by MS), followed by strand-specific sequencing of covalently attached DNA fragments. With Topo-Seq, 356 TopoIV cleavage sites were identified in the *E. coli* genome.

Cfx inhibits both Type IIA bacterial topoisomerases - TopoIV and DNA gyrase – so, to increase the specificity of treatment, a mutation in a *gyrA* gene leading to the mutation S83L conferring resistance of DNA gyrase to fluoroquinolones was introduced by recombineering. In agreement with published data [217], the mutation led to the specific 32-fold increase of the minimal inhibitory concentration to Cfx (**Figure 26B**). On average, for the wild-type strain Topo-Seq detected ~8-times lower amount of cleavage sites than with Cfx-resistant strain, which might indicate that more Cfx molecules are available for binding to TopoIV when DNA gyrase is resistant to it (**Figure 26C**). Alternatively, GyrA S83L mutant could have a decreased DNA binding and/or supercoiling activity, which might lead to proportionally increased TopoIV activity detected with Topo-Seq. The latter hypothesis is supported by the observation that in *gyrA S83L* background DNA is less negatively supercoiled implying the reduced activity of the mutated protein [217]. However, other groups reported that supercoiling is not affected by this mutation and that this mutation has a neutral effect on cell fitness [257,258]. Activity of GyrA S83L-containing gyrase should be directly compared to the wild-type enzyme *in vitro* and amounts of gyrase, TopoIV, and TopoI should be compared between *wt* and mutated strains to test this hypothesis.

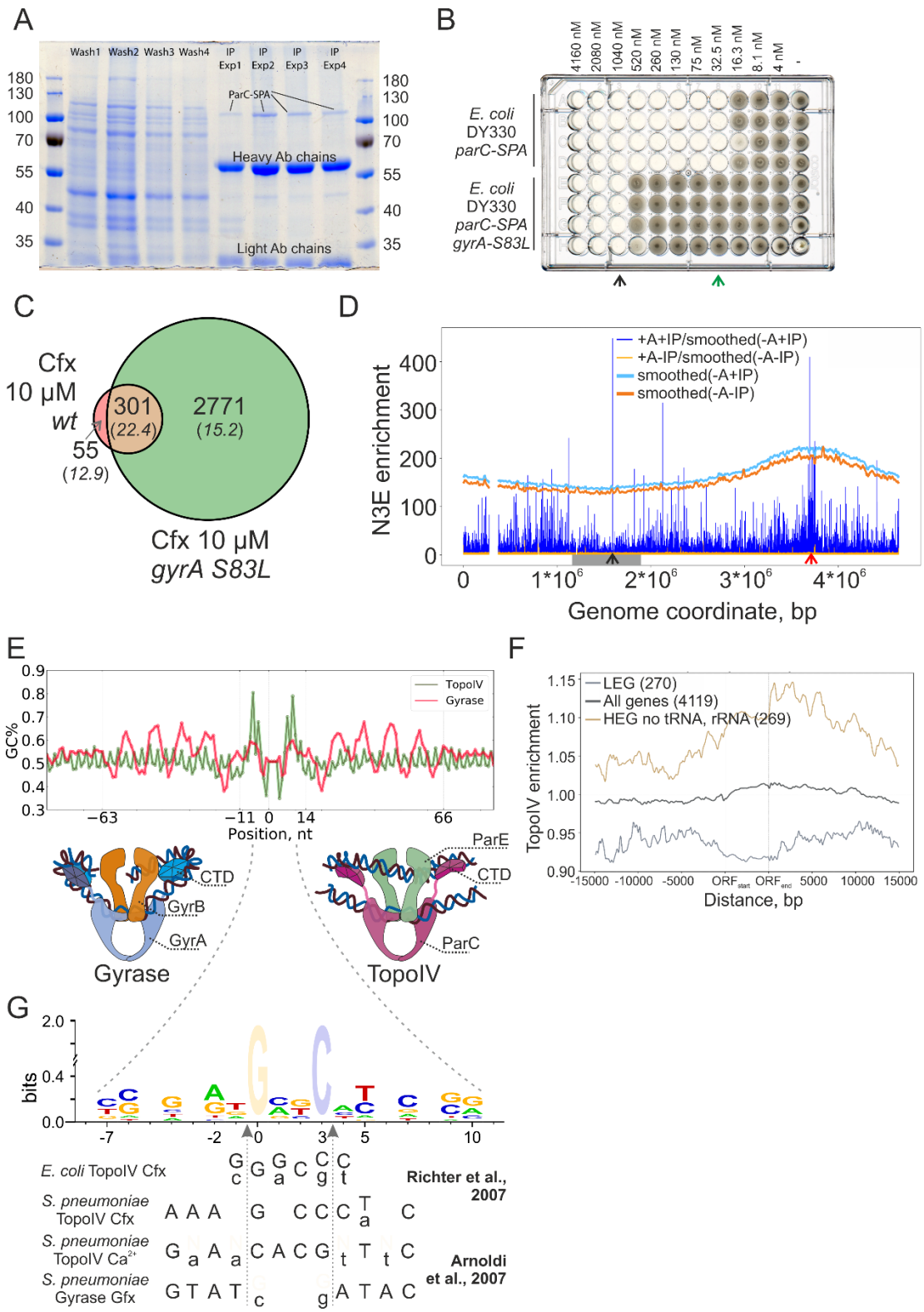


Figure 26. ***E. coli* TopoIV Topo-Seq**. (**A**) Purification of ParC-SPA subunits conjugated with DNA during Topo-Seq with M2-agarose affinity resin (Sigma Aldrich). Bands of protein marker are labeled with kDa. (**B**) MIC determination for DY330 *parC-SPA* strain and DY330 *parC-SPA gyrA-S83L* derivative. 4 biological replicates are shown. (**C**) Venn diagram representing overlap between GCS sets identified for *wt*

and *gyrA-S83L* strains. (**D**) TopoIV enrichment (N3E) over *E. coli* genome (blue). Black and red arrows mark positions of *dif* and *oriC* sites, respectively. Grey rectangle indicates *ter*-region of *E. coli* chromosome containing *matS* sites. (**E**) Comparison of nucleotide motifs found for gyrase and TopoIV with Topo-Seq. GC frequencies at positions 0 and +3 are manually set to the average GC frequency to remove ciprofloxacin-induced bias. Below the plot are the schematic representations of gyrase and TopoIV complexes with DNA. (**F**) Metagene plot of averaged TopoIV enrichment in upstream and downstream regions of genes (15 kb each) and inside genes. Enrichment is shown for all genes (black curve, 4119 genes), least-expressed genes (grey curve, LEG, 270 genes), and highly-expressed genes (orange curve, HEG, 269 genes). (**G**) Logo representation of TopoIV motif determined with Topo-Seq. Cfx-biased G and C located in 0 and +3 positions, respectively, are shown not to scale. Below the logo, cleavage consensus sequences identified in previous works for TopoIV and gyrase *in vitro* are shown. Preferred bases are shown in capital letters and non-preferred bases are shown in lowercase letters. Cfx – ciprofloxacin, gfx - gemifloxacin

Enrichment of TopoIV cleavage activity spans over the *E. coli* genome and has several remarkable peculiarities that differs it from gyrase enrichment (**Figure 26D**).

First, TopoIV has an increased number of cleavage sites with a high level of enrichment near the origin region. Interestingly, this phenomenon was not observed by ChIP-Seq and NorfIP earlier [116]. However, by super-resolution microscopy it was observed that ~15 TopoIV complexes are colocalized with ~16 MukBEF complexes in a cluster at replication origin [131,134,259]. I speculate that TopoIV may be involved in early steps of replication progression where its activity can be critical to resolve precatenanes when two replisomes are relatively close to each other.

Second, TopoIV enrichment is depleted in a *ter*-region (grey zone, **Figure 26D**), which was also detected by NorfIP. *Ter*-region contains 23 *matS* sites which are bound by a MatP protein. MatP is known to organize the Ter macrodomain by forming dimers and dimers of dimers, also it binds MukBEF and promotes its unloading from DNA [134]. Moreover, interaction of MatP and TopoIV with MukBEF was shown to be competitive [137]. So, the enrichment depletion in the *ter*-region is likely reflecting the displacement of MukBEF-TopoIV complexes by MatP. Should be noted, that TopoIV enrichment is symmetrically increased left and right to the *ter*-region, presumably indicating active resolution of precatenanes at the edges of the region. Why MukBEF-TopoIV complexes should be displaced from *ter*? It was proposed that unloading system based on *matS*-MatP is required to recycle SMC associated with DNA, and promote it uploading at replication origin [134]. I speculate that TopoIV depletion may prevent knotting and catenating of over-replicated chromosomes when a replisome did not stop between inner *ter*-sites *terA*

and *terC* and travels further, replicating an opposite replichore. Particularly, two pairs of *ter*-sites (B, C and A, D), which are considered as primary replisome traps, located inside the *ter*-region meaning that over-replicated regions are mainly restricted by this region [260].

Third, the strongest cleavage site detected with Topo-Seq (as well as with NorfIP) is located at the *dif*-site inside the *ter*-region (marked with black arrow, **Figure 26D**). The site is bound by XerC/XerD recombinase which resolves chromosome dimers and is known to bind TopoIV (by XerC). Existing model implies that a chromosome dimer is resolved as a catenane which, in turn, is released by TopoIV activity [116].

Single-nucleotide resolution of Topo-Seq allowed to identify the cleavage motif of TopoIV. As was for gyrase, in the positions 0 and +3 of TopoIV motif there were G and C nucleotides, respectively, which are the hallmark of Cfx-induced cleavage (removed from the motif in **Figure 26E**). The overall motif of TopoIV was much shorter and was lacking the periodic regions characteristic for gyrase motif. The motif resembles motif of eukaryotic Top2 [261]. The absence of periodic regions indicates that CTD of ParC does not interact with proximal extensions of DNA G-segments, but rather binds a different DNA region, a mechanism that favors resolution of precatenates and catenanes. Identified TopoIV motif was compared with cleavage motifs determined earlier with *in vitro* assays [243,244,262]. Interestingly, the motif resembles more motifs identified for *S. pneumoniae* TopoIV than one identified for *E. coli* enzyme (**Figure 26G**). Motifs for *S. pneumoniae* TopoIV were identified using drugs (ciprofloxacin and gemifloxacin - Gfx) and with Ca^{2+} (a method considered more native), therefore, nucleotides determined with Topo-Seq likely reflect native binding preferences of the enzyme. Interestingly, identified TopoIV motif resembles much of the *S. pneumoniae* gyrase (identified using gemifloxacin), indicating that the pattern can be conserved for type-II topoisomerases.

To investigate if TopoIV is somehow involved in relaxation of transcription-induced supercoiling, I performed a metagene analysis for TopoIV enrichment. Turned out, that TopoIV enrichment was dependent on the transcription level of selected genes – with the highest enrichment associated with downstream regions of highly expressed genes (where positive supercoiling is expected) (**Figure 26F**). This observation indicates that TopoIV, similarly to gyrase, is likely involved in relaxation of transcription-induced positive supercoiling *in vivo* which is in line with its *in vitro* preferences [59,95].

Overall, compared with ChIP-Seq and NorfIP, Topo-Seq demonstrated enhanced sensitivity and precision. With ChIP-Seq only ~10 peaks were identified spanning 200 bp each and with NorfIP 223 cleavage regions were repeatedly detected, while with Topo-Seq 1908 cleavage sites were detected (**Figure**

27A). Among the three approaches, only Topo-Seq allowed to map cleavage sites with a single-nucleotide resolution making possible to unambiguously detect a cleavage motif (**Figure 27B**).

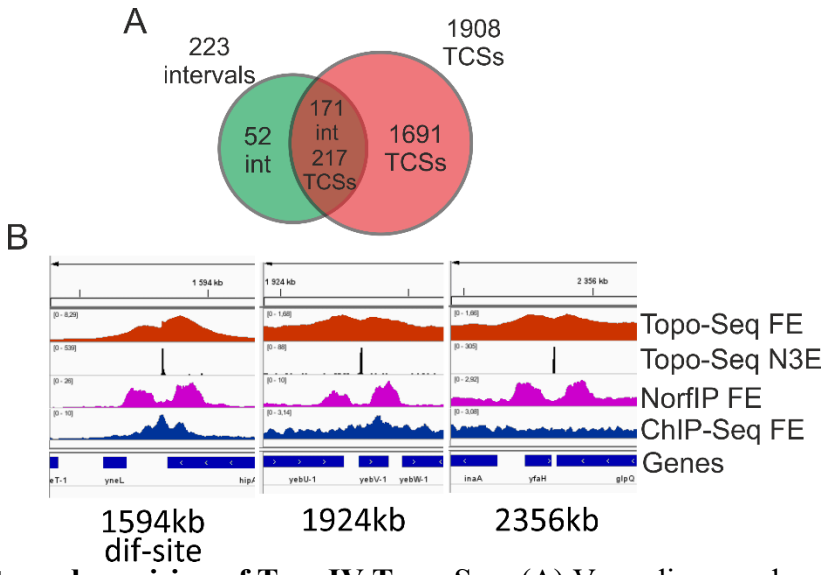


Figure 27. **Specificity and precision of TopoIV Topo-Seq**. (A) Venn diagram demonstrating the overlap between TopoIV cleavage sites identified with Topo-Seq (TCSs, 1908 sites) and NorfIP (intervals, 223 sites). (**B**) Three representative genomic regions containing TopoIV binding and cleavage sites. Red track is a TopoIV fold enrichment (FE, Topo-Seq data), black track is a TopoIV N3E (Topo-Seq data), pink and blue tracks are TopoIV fold enrichments generated by, respectively, NorfIP and ChIP-Seq (data is from [116]).

Short and weak motif observed for TopoIV rises a question which factors (besides the Cfx bias and a specialized *dif*-site) contribute to the cleavage specificity of TopoIV. Probably, other proteins interacting with TopoIV, such as SeqA and MukBEF, as well as topological factors, supercoiling and DNA crossings, together contribute to this.

Results of sections 4.15-4.8 are published in: **Sutormin D.**, Galivondzhyan A., Musharova O., Travin D., Rusanova A., Obraztsova K., Borukhov S., Severinov K. Interaction Between Transcribing RNA Polymerase and Topoisomerase I Prevents R-loop Formation in *E. coli*. **Nature Communications**, 2022. **ACCEPTED FOR PUBLICATION**

Alina Galivondzhyan and Ksenia Obraztsova contributed to the results of these sections.

4.15 EcTopoI is widely distributed over the *E. coli* genome, colocalized with RNAP, and enriched in regions with negative supercoiling, which was revealed by ChIP-Seq

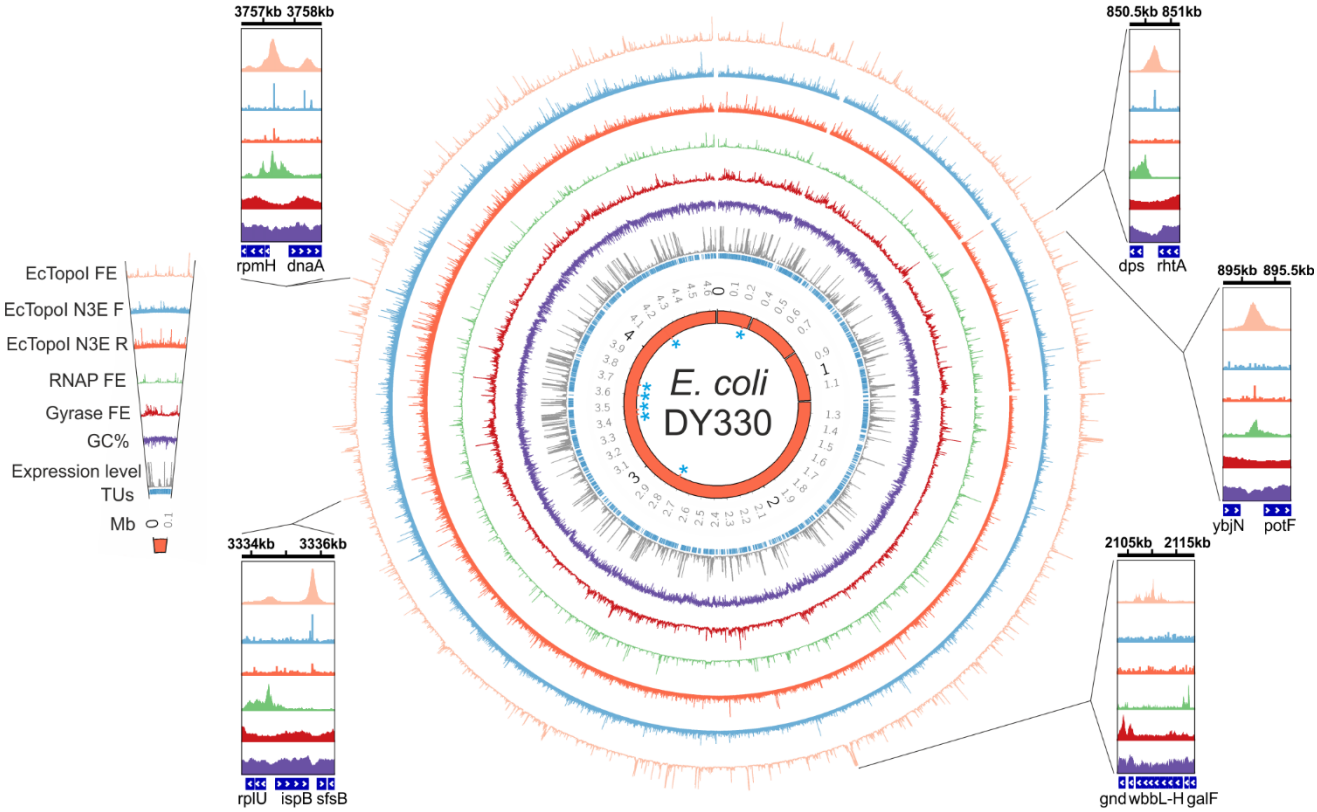


Figure 28. Distribution of EcTopoI, RNAP, and gyrase enrichment peaks over the *E. coli* chromosome. Circular maps show fold enrichment profiles of EcTopoI (ChIP-Seq, light orange; Topo-Seq, cyan and dark orange for two separate DNA strands), RNAP (ChIP-Seq, green), and DNA gyrase (Topo-Seq, dark red). Additionally, GC-content (%, purple) and mean expression levels (FPKM, RNA-Seq, grey) for annotated TUs (inner blue segments) are also shown. Blue asterisks indicate positions of rRNA operons on the innermost orange ring representing *E. coli* DY330 genome. The numbers on the outside of the orange ring indicate genome coordinates in megabase pairs (Mbs). Three gaps around ~0.3 Mb, ~0.8 Mb, and ~1.2 Mb correspond to deletions in the *E. coli* DY330 genome relative to the *E. coli* W3110 reference genome. Insets provide a zoom-in view at representative regions with high EcTopoI signals. Coordinates in kb are

indicated on top of each inset. For ChIP-Seq, fold enrichment is given relative to the input sample in all figures. The maps were constructed with the Circos tool [253], and the insets were prepared using IGV [208].

The topoisomerase I distribution along the *E. coli* chromosome was determined using ChIP-Seq with a DY330 strain derivative carrying a fusion of the *topA* gene with the SPA tag encoding sequence (**Figure 28**, orange track). Three biological replicas were made, showing good reproducibility between them (Pearson correlation > 0.6). Using the MACS2 analysis pipeline [263], we detected 403 significantly enriched regions (e-value < 0.001) present in all three replicas. EcTopoI peaks tend to have a lower GC-content than the genome average. Indeed, a positive correlation between peaks log-fold enrichment and the AT-content was observed (Spearmen correlation 0.36, p-value 2.3e-5). Furthermore, the peaks appeared to be uniformly distributed over the entire chromosome. Of note, there was no enrichment of the EcTopoI signal at the terminator region of chromosome replication, in contrast to observations made for *M. smegmatis* [112].

We next determined whether the EcTopoI ChIP-Seq signal overlaps with the RNAP signal, a result that might be expected based on the published data about the interaction between the two enzymes [179]. We performed a ChIP-Seq experiment with a DY330 strain derivative expressing TAP-tagged RpoC (β' RNAP subunit) (**Figure 28**, green track). The RpoC ChIP-Seq signal correlated well with the published ChIP-Seq obtained for RpoB (β RNAP subunit; Spearman correlation 0.59, p-value=2.4e-158; **Figure 29A**) [249] and transcription level (RNA-Seq performed with exponentially growing *E. coli* DY330, Spearman correlation 0.55, p-value 1.4e-133; **Figure 29B**). Overall, we found 3635 RpoC peaks with fold enrichment of at least 3, ~25% of which overlapped with earlier reported RpoB peaks (Monte-Carlo simulation with 10000 iterations, p-value<1e-308). 60% of topoisomerase peaks (243/403, Monte-Carlo simulation with 10000 iterations, p-value=4.9e-6) overlapped with the RpoC peaks (**Figure 29C**). Consistently, enrichment of RpoC is significantly higher within the EcTopoI-enriched regions compared to the outside regions (Welch t-test, p-value<1e-308). Reciprocally, enrichment of EcTopoI is significantly higher inside the RpoC-occupied regions than outside of these regions (Welch t-test, p-value<1e-308) (**Figure 29D**). Colocalization of the RpoB and EcTopoI signals was also observed with a publicly available RpoB ChIP-Seq dataset for *E. coli* MG1655 [249]. Overall, we conclude that EcTopoI is significantly colocalized with RNAP on the *E. coli* chromosome in exponentially growing cells.

The ChIP-Seq signal of EcTopoI was also generally proportional to transcript abundance, with the highest enrichment values observed for 200 most highly-expressed transcription units (HETUs, expression

level > 31 FPKM) and, particularly, for rRNA operons and their upstream regions (**Figures 29F**). In contrast, little or no EcTopoI enrichment was observed for 200 least expressed TUs (LETUs, expression level < 0.31 FPKM) (**Figure 29E**).

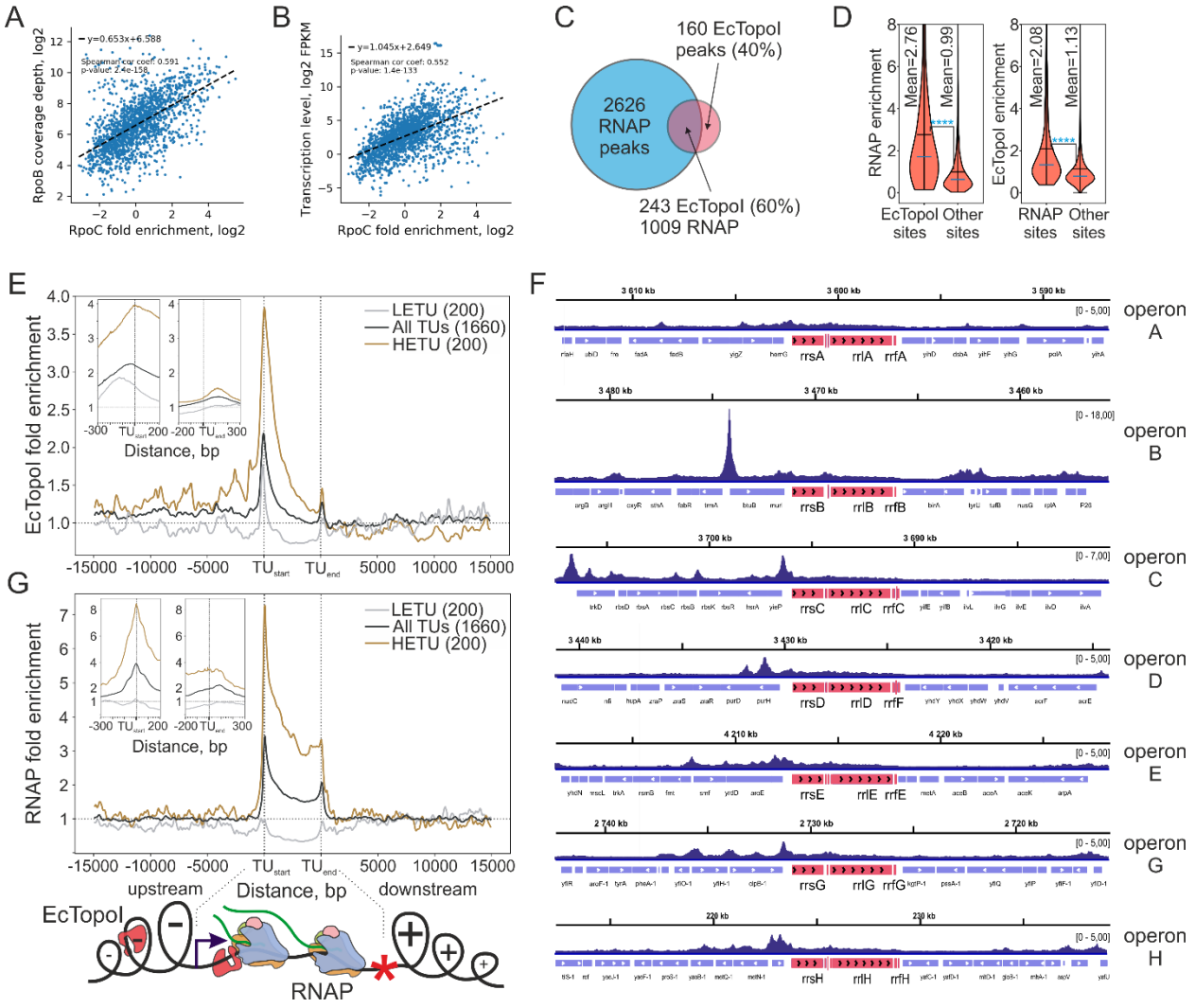


Figure 29. **EcTopoI and EcRNAP enrichment colocalization**. (**A**) Correlation between RpoC fold enrichment and RpoB signal for TUs. (**B**) Correlation between RpoC fold enrichment and transcription level of TUs. (**C**) Venn diagram represents an overlap between the EcTopoI peaks (403 total) and RNAP peaks (3635 total). (**D**) Violin plots of RNAP enrichment in EcTopoI peaks and outer regions (left), and EcTopoI enrichment in RNAP peaks and outer regions (right). The means and medians are indicated by black and blue lines, respectively. Statistically significant differences between means (t-test, p-value<<0.05) are indicated by asterisks. (**E**) Metagene plot of EcTopoI enrichment within TUs (middle), their upstream (left), and downstream (right) regions. Enrichment is shown for all TUs (black curve), highly-expressed (HETU, orange curve), and least-expressed (LETU, grey curve) sets. The number of TUs in each group is indicated

in parentheses. The two insets show zoom-in views of EcTopoI enrichment near transcription start (TU start) and termination (TU end) sites. (**F**) Snapshots from IGV genomic browser representing EcTopoI fold enrichment around rRNA operons for Rif-/CTD- conditions. rRNA operons are listed from A to H; genes comprising the operons are colored in red. For panels F and G ChIP-Seq fold enrichment is given relative to the input sample. (**G**) Top, metagene plot of RNAP enrichment within TUs, their upstream, and downstream regions. Analysis and groups of TUs are the same as in **E**. Bottom, graphical representation of the Liu & Wang twin-domain model [185] showing localization of RNAP and EcTopoI according to the metagene plots in **E** and **G**.

Next, we analyzed the enrichment of EcTopoI and RNAP within TUs, and in upstream and downstream regions (**Figures 29E**, **G**). Metagene analysis indicated colocalization of EcTopoI and RNAP within the TU bodies, with the highest enrichment for both enzymes near the transcription start sites (TSS). A decreasing RNAP enrichment gradient toward the ends of TUs, presumably caused by premature transcriptional termination [264,265], was observed. A gradient with a similar slope was also detected for EcTopoI enrichment, suggesting that EcTopoI either directly follows elongating RNAPs or physically associates with the enzyme.

EcTopoI accumulated upstream but was depleted downstream of TUs, a result that is consistent with the predictions of the Liu & Wang twin-domain model that posits accumulation of negative supercoiling (a substrate of TopoI) upstream, i.e., behind the elongating RNAP [185]. Excessive accumulation of EcTopoI could be tracked up to 12-15 kb upstream of TSS for HETUs (**Figure 29E**), suggesting that negative supercoiling diffuses over significant lengths of the *E. coli* chromosome. Interestingly, this range is significantly longer than that observed for eukaryotic chromatin, possibly, due to the absence of supercoiling-"buffering" nucleosomes [266,267]. Alternatively, the enrichment of EcTopoI in upstream regions might be mediated by transcription-induced chromatin remodeling, though we did not observe any significant skew in enrichment of Fis, HNS, MatP, and MukB nucleoid-associated proteins in these regions (data not shown). A small peak of EcTopoI and RNAP enrichment at TU ends may correspond to enrichment at promoter regions of closely packed adjacent genes or result from the physical association of the two enzymes at transcription termination sites. The roles of associated EcTopoI in transcription termination process, if any, could be in facilitation of DNA duplex restoration after the RNAP complex dissociation.

Overall, our observations strongly support the association of EcTopoI with RNAP at TSSs and within the TUs, as well as with negatively supercoiled DNA upstream of actively transcribed genes.

4.16 EcTopoI is not involved in chromosome decatenation in the Ter region

Topoisomerase I can catenate and decatenate ssDNA and dsDNA circles containing nicks *in vitro* [112,268,269]. Recently, an enrichment of TopoI activity was found at the Ter region of *M. smegmatis* chromosome [112]. This bacterium lacks classical decatenating topoisomerases, TopoIII and TopoIV, and, thus, TopoI is likely to be involved in chromosomal decatenation after replication. Despite the similarity in TopoI *in vitro* activities in both bacterial species, we did not observe any specific binding or cleavage by EcTopoI near Ter regions of the *E. coli* chromosome, indicating that this enzyme is not involved in chromosome decatenation.

4.17 The RNAP inhibitor rifampicin causes EcTopoI re-localization to promoter regions

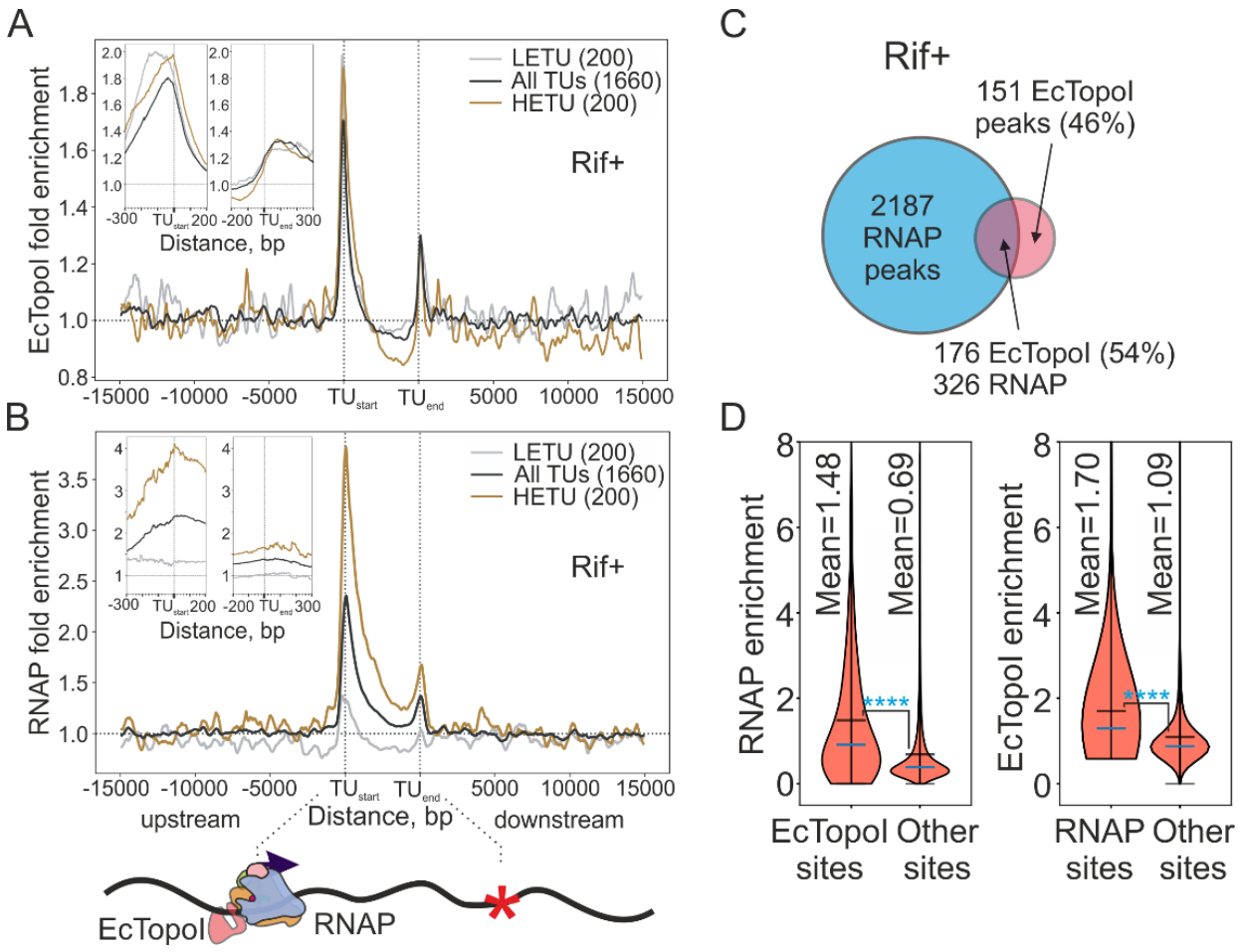


Figure 30. **Enrichment of EctopoI and RNAP upon Rif treatment**. (**A**) Metagene plot of EcTopoI enrichment within TUs, their upstream, and downstream regions for cells pretreated with Rif. (**B**) Top, metagene plot of RNAP enrichment within TUs, their upstream, and downstream regions for cells pretreated with Rif. Bottom, graphical representation showing localization RNAP and EcTopoI according to metagene

plots in **A** and **B**. (**C**) Venn diagram represents an overlap between the EcTopoI peaks (327 total) and RNAP peaks (2513 total) in cells pretreated with Rif before ChIP-Seq. (**D**) Violin plots of RNAP enrichment in EcTopoI peaks and outer regions (left), and of EcTopoI enrichment in RNAP peaks and outer regions (right) for Rif-treated cells (an RpoC ChIP-chip dataset for Rif-treated cells was taken from [270]).

If EcTopoI interacts with RNAP, it should redistribute to promoters upon the treatment with rifampicin (Rif), an inhibitor that prevents RNAP escape into elongation [270,271]. In addition, if EcTopoI association with extended regions upstream of TUs is driven by excessive transcription-generated negative supercoiling, Rif treatment should abolish this association. To test these predictions, we performed EcTopoI ChIP-Seq in cells treated with Rif prior to formaldehyde fixation. According to metagene analysis, EcTopoI enrichment along the lengths of HETUs bodies disappeared in Rif-treated samples, reaching values below the background (**Figure 30A**). This was consistent with the disappearance of elongating RNAP from TU bodies in Rif-treated samples (**Figures 30B**; see [270]). Association of EcTopoI with upstream regions of HETUs was also abolished upon Rif treatment (**Figures 30A**).

Yet, the enrichment of EcTopoI at promoter regions of HETUs remained, although at lower levels compared to the untreated control. This decrease may be caused by the dissipation of transcription-induced negative supercoiling and/or by a more uniform redistribution of RNAP holoenzymes across promoters (since high-affinity promoters cannot be occupied by more than one RNAP molecule, remaining molecules become trapped by Rif at weaker promoters). The latter scenario is supported by the observation that in Rif-treated samples, enrichment of both EcTopoI and RNAP is increased at LETU promoters (**Figures 30A**, **B**). Be that as it may, EcTopoI and RNAP remained colocalized in Rif-treated cells (Welch t-test, p-value<1e-308, **Figures 30C**, **D**), sharing a significant number of enrichment peaks (Monte-Carlo simulation with 10000 iterations, p-value<1e-308). Consistent with the re-localization of EcTopoI to promoters, EcTopoI peaks found by MACS2 in Rif-treated cells were narrower (median width 311 bp) and more AT-rich (43.5% GC) than peaks in untreated samples. Overall, these results further support the EcTopoI interaction with elongating RNAP, promoter initiation complexes, and regions upstream of transcribed genes.

4.18 EcTopoI is recruited to chromosomal regions with excessive negative supercoiling surrounded by topological barriers

I examined EcTopoI distribution in 1529 *E. coli* intergenic regions (IRs) in more detail. I observed significant EcTopoI enrichment at IRs flanked by highly-transcribed genes and/or having a high level of RNAP enrichment (**Supplementary Figures 2A**, **B**). Irrespective of RNAP enrichment/transcriptional

activity, high levels of EcTopoI enrichment were found at IRs that *i*) were located between divergently transcribed genes (**Supplementary Figures 2C**), *ii*) harbored transcription factor-binding sites (**Supplementary Figure 2D**), and *iii*) were flanked by genes coding for membrane proteins (**Supplementary Figures 2E**). Consistently, IRs that fulfilled all three criteria and were located between highly transcribed genes had the highest level of EcTopoI enrichment (**Figures 31A**).

The meta-intergene analysis revealed that IRs flanked by divergent genes exhibit, on average, much higher EcTopoI signal than those located between convergent genes (**Figure 32A**). Based on these results, we propose that EcTopoI is preferentially recruited to regions with excessive negative supercoiling stabilized by local topological barriers (see representative examples in **Figure 31B**). These barriers may be generated by divergent transcription from highly complex promoters and by transertion process (a coupled transcription/translation/polypeptide chain translocation into the cell membrane) [272] (**Figure 31C**).

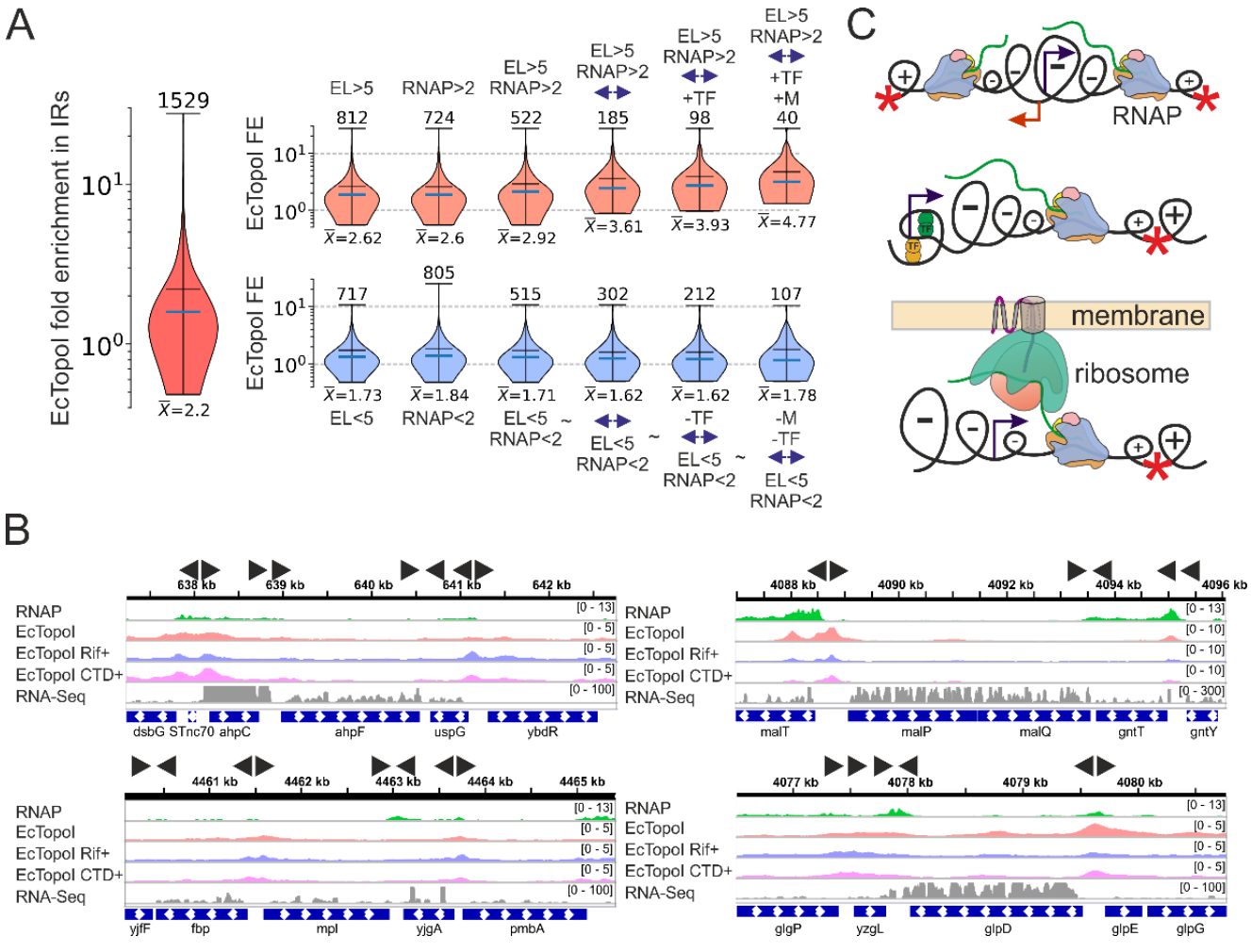


Figure 31. Topological barriers attract EcTopoI to intergenic regions. (**A**) Cumulative effect of transcription and local topological barriers on EcTopoI fold enrichment (FE) in IRs. Violin plots show the

contribution of positive (top row) and negative (bottom row) factors to the increased EcTopoI enrichment in IR, including the expression level of adjacent genes (EL), fold enrichment of RNAP (RNAP), presence of divergently orientated adjacent genes (double arrows), annotated sites for transcription factors (TF), and membrane localization of proteins encoded by flanking genes (M). The leftmost plot shows an overall EcTopoI fold enrichment in 1529 IRs. Positively associated features are: high transcription level and high level of RNAP enrichment (EL>5 FPKM, RNAP>2 fold enrichment units), ←→ - divergently transcribed genes, +TF – at least one annotated site of a transcription factor is present in an IR, M – at least one gene flanking IR encodes a membrane protein. Negatively associated features are: low level of transcription and RNAP enrichment (EL<5, RNAP<2), ~←→ - genes separated by an IR are NOT in a divergent orientation, -TF – no annotated sites of transcription factors binding are present, -M – genes flanking an IR do not encode membrane proteins. Means are shown with horizontal black lines and numeric values are shown below each violin-plot, medians are shown as blue horizontal lines; vertical axes are log-scaled. (**B**) Representative genomic regions showing the increased enrichment of EcTopoI in divergent IRs upstream of highly transcribed TUs and decreased enrichment in convergent IRs. Data for four representative genomic regions are shown. Enrichment of RNAP is shown in green; EcTopoI - in red; EcTopoI in Rif+ condition - in blue; EcTopoI in CTD+ condition - in magenta. RNA-Seq data are shown in grey. Data scaling is shown in parenthesis for each genomic track. Locations of IRs with specific enrichment patterns and directions of transcription of flanking genes is shown by black horizontal arrows. (**C**) Graphical representation of topological borders that trap negative supercoiling: divergent orientation of genes, complexly organized promoters, transertion.

4.19 EcTopoI and DNA gyrase have mutually exclusive localization on the *E. coli* chromosome

EcTopoI and DNA gyrase have opposite binding preferences and activities: while EcTopoI is attracted to and relaxes negative supercoils, DNA gyrase is attracted to and removes positive supercoils [59,155,186,219]. Comparison of ChIP-Seq data for EcTopoI and Topo-Seq data for DNA gyrase [219] directly demonstrates that *in vivo* gyrase enrichment is significantly lower in regions occupied by EcTopoI and *vice versa* (Welch t-test, p-value<1e-308, **Figure 32B**).

While EcTopoI is enriched upstream of HETUs (where transcription-induced negative supercoiling should be high) and depleted in downstream regions (where positive supercoiling should be accumulated), the gyrase enrichment shows the opposite pattern (**Figure 32C**). I used Psora-Seq and GapR-Seq data recently published and available for *E. coli* to localize enrichment of topoisomerases with, respectively, regions of negative and positive supercoiling genome-wide. A signal of negative supercoiling revealed by

Psora-Seq [273] matches the enrichment of EcTopoI, indicating that EcTopoI accumulation upstream of active TUs indeed colocalizes with increased negative supercoiling. Concordantly, a signal of positive supercoiling revealed by GapR-Seq [274] matches the enrichment of DNA gyrase in regions downstream of active TUs (**Figure 32C**). Both gyrase enrichment downstream of TUs and EcTopoI enrichment upstream of TUs positively correlate with transcription activity and are abolished by Rif (**Figure 18**). Furthermore, while EcTopoI is particularly enriched in IRs flanked by divergent genes (see above) where cumulative negative supercoiling is expected, the DNA gyrase signal is the highest for IRs between convergent genes (where cumulative positive supercoiling is expected) (**Figure 32A**), in line with observations made in *M. tuberculosis* [66]. Together, these data indicate that EcTopoI and gyrase have opposing patterns of distribution genome-wide, fully consistent with the predictions of and validating the Liu & Wang model [185].

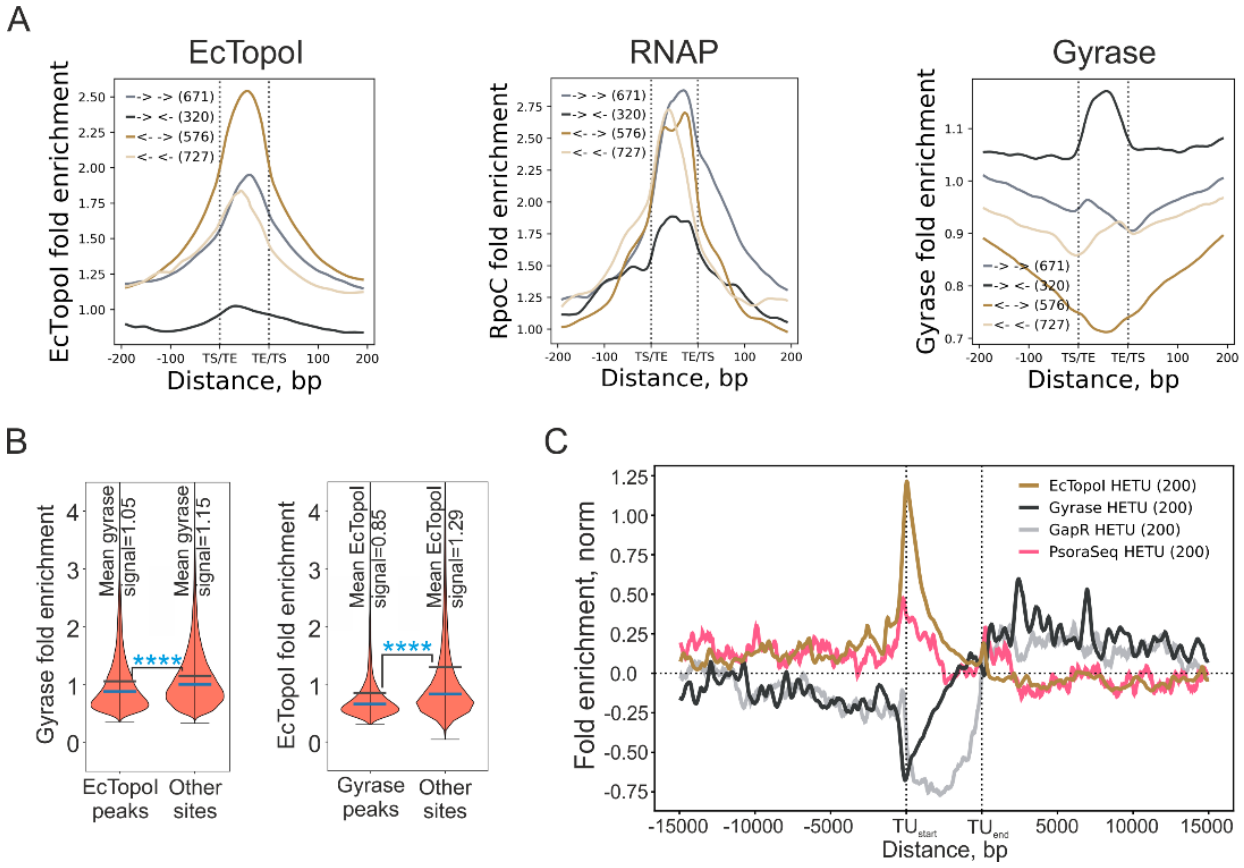


Figure 32. **EcTopoI and DNA gyrase mutual exclusion**. (**A**) Meta-intergene plots of EcTopoI (on the left), RNAP (in the center), and gyrase (on the right) enrichments in IRs. IRs were classified according to the orientation of flanking genes. The number of regions comprising a group is indicated in brackets. (**B**) Violin plots of gyrase enrichment in EcTopoI peaks and outside of these regions (left) and of EcTopoI enrichment in gyrase peaks and outside of these regions (right). Mean and median are indicated by black

and blue lines, respectively. The statistically significant difference between means (t-test, p-value<<0.05) is indicated by asterisks. (**C**) Comparison of EcTopoI ChIP-Seq data (CTD-/Rif- conditions), gyrase Topo-Seq data (experiments with ciprofloxacin [219]), GapR-Seq [274], and Psora-Seq [273]. Metagene analysis performed for the HETU set of TUs. For ChIP-Seq, fold enrichment is given relative to the input sample.

4.20 Mapping of TopoI cleavage sites in *E. coli* with a single-nucleotide resolution with Topo-Seq

The DNA topoisomerases binding and cleavage sites may not completely overlap (El Sayyed *et al.*, 2016; Baranello *et al.*, 2017). To identify EcTopoI cleavage sites (TCSs) *in vivo* genome-wide, a Topo-Seq was conducted. Due to the unavailability of efficient poisons (like ciprofloxacin for gyrase) for EcTopoI, I constructed EcTopoI G116S/M320V, an "intrinsically-poisoned" double mutant that forms stable covalent complexes with DNA (Cheng, Sorokin and Tse-Dinh, 2008). As expected, continuous production of EcTopoI G116S/M320V from a plasmid led to growth inhibition (**Figure 33A**) and SOS-response (**Figure 33B**). EcTopoI G116S/M320V was transiently (30 min) expressed in *E. coli* DY330, and the trapped protein-DNA cleavage complexes were purified through a C-terminal StrepII tag fused with the mutant topoisomerase (**Figure 33C**). Expression of EcTopoI G116S/M320V had no apparent adverse effect on cell culture growth in the course of the experiment (**Figure 33A**). Topoisomerase-associated DNA fragments were isolated and subjected to strand-specific sequencing of ssDNA using the Accel NGS 1S kit, and the reads were mapped to the reference genome. The number of 3'-ends (N3E) was counted for every genomic position strand-specifically. Since EcTopoI forms a covalent intermediate with the 5'-end of a single-stranded break it introduces and leaves the 3'-end unmodified, an increase in the N3E should mark a TCS. A total of 262 TCSs were identified in the *E. coli* genome (125 on the forward and 137 on the reverse strand). The TCSs determined by Topo-Seq, which identifies the sites of EcTopoI activity with single-base precision, significantly overlap with EcTopoI peaks detected by ChIP-Seq (Monte-Carlo simulation with 10000 iterations, p-value 3.5e-13). Interestingly, several chromosomal regions with increased EcTopoI binding (as evidenced by ChIP-Seq and ChIP-qPCR data shown in **Figure 33F**) and enhanced cleavage (as evidenced by Topo-Seq, **Figure 33D**) also demonstrated high affinity to purified EcTopoI *in vitro*, revealing sequence specificity of the enzyme (**Figure 33E**).

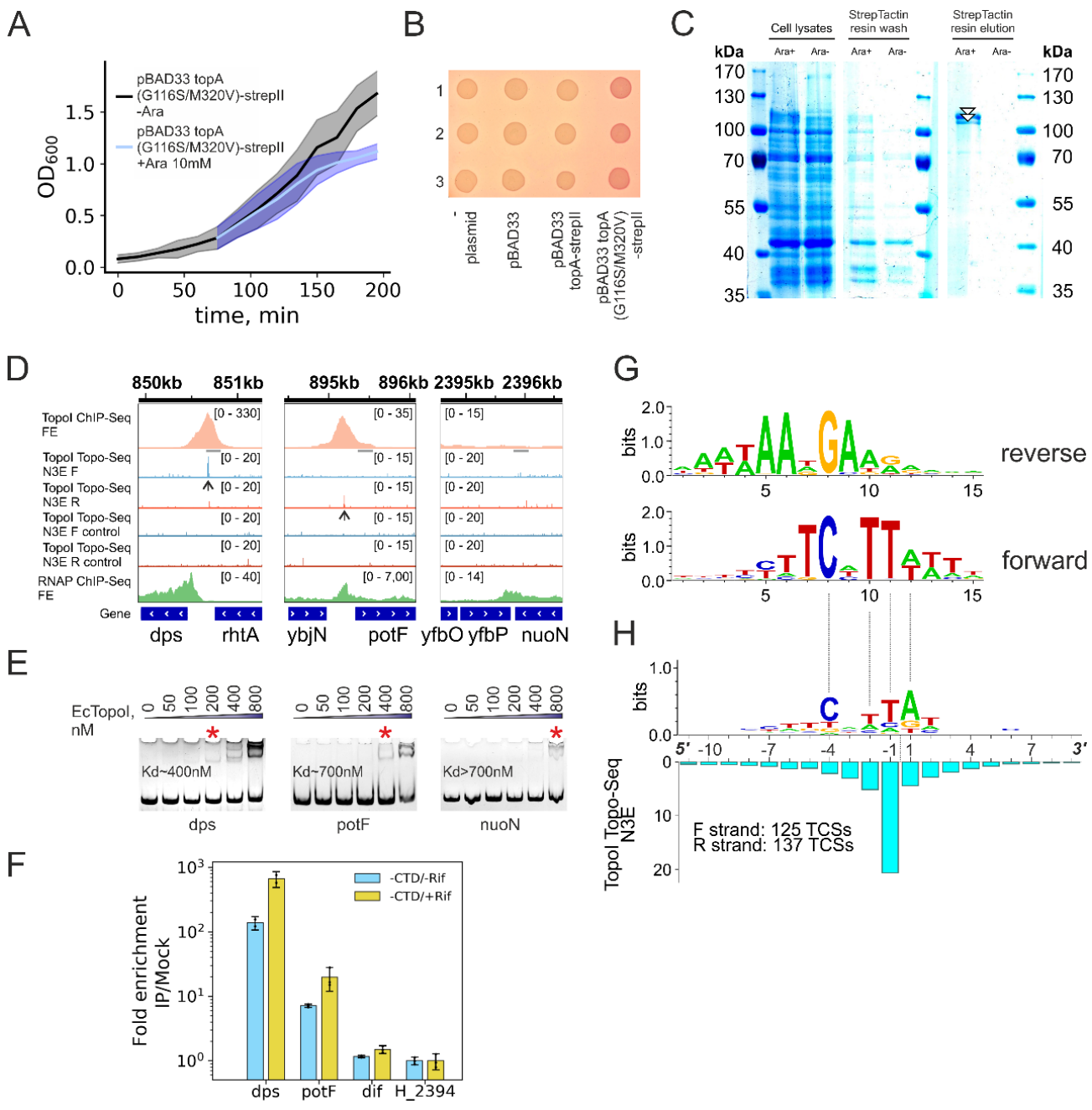


Figure 33. **Identification of EcTopoI cleavage sites by Topo-Seq**. (**A**) Effect of EcTopoI G116S/M320V overexpression (from pBAD33 topA(G116S/M320V)-strepII) on *E. coli* DY330 *topA-SPA* growth. Shaded area represents 0.95 confidence intervals for the mean of three biological replicas. (**B**) SOS-response is induced by EcTopoI G116S/M320V overexpression from pBAD33 plasmid. *E. coli* CSH50 λ*sfiA::lacZ* [199] reporter strains (empty, harboring empty pBAD33, pBAD33 EcTopoI or pBAD33 EcTopoI G116S/M320V) were grown on MacConkey indicator medium. Red color indicates a SOS-response. The assay is shown for three replicates. (**C**) Overexpression and purification of EcTopoI G116S/M320V during

the Topo-Seq procedure. Bands corresponding to EcTopoI G116S/M320V (verified by MS) are marked with white triangles. (**D**) Representative regions of the *E. coli* chromosome demonstrate the matching of EcTopoI peaks identified by ChIP-Seq and EcTopoI TCSs obtained by Topo-Seq (*dps*, *potF*) and a region lacking EcTopoI binding and activity (*nuoN*). EcTopoI cleavage activity is shown strand-specifically for forward and reverse strands separately. As a control, a non-induced culture was used. Amplified regions used for ChIP-qPCR and affinity measurements are indicated with grey rectangles. (**E**) The affinity of purified EcTopoI to three amplified regions of the genome indicated in panel **D** assessed by EMSA: *nuoN* (control, low ChIP-Seq and Topo-Seq signals), *potF*, and *dps* (two highest peaks identified by ChIP-Seq and strong TCSs). Red stars mark the lowest concentration of EcTopoI at which a gel-shift was detected. (**F**) ChIP-qPCR validation of ChIP-Seq peaks near the *dps* and *potF* genes. The *dps* and *potF* regions exhibit dramatically increased fold enrichment signals compared to the two control regions – *dif* and H2394. (**G**) A logo representing a binding motif of EcTopoI identified by ChIPMunk [275] in sequences under the ChIP-Seq peaks. The motif is shown in both orientations. (**H**) EcTopoI cleavage motif identified by alignment of TCSs (cleavage is introduced between nucleotides at positions -1 and 1 and is indicated by a dashed line). The N3E signal is plotted below the motif.

To further characterize sequence specificity of EcTopoI, I used ChIPMunk [275] to find overrepresented motifs in EcTopoI ChIP-Seq peaks sequences. A strong motif was detected in more than 90% of enrichment peaks. The motif was AT-rich, strongly asymmetric, with a conserved central TCNTTA/T part (**Figure 33G**) and limited similarity to known DNA motifs in *E. coli*. Next, to test whether EcTopoI has a specific cleavage motif, I aligned sequences around the established TCSs. The identified cleavage motif was very similar to the binding motif identified using ChIP-Seq. As shown in **Figure 33H**, EcTopoI preferentially cleaves a TA dinucleotide located 4 nt downstream of the conserved position of C. Earlier *in vitro* experiments demonstrated that type-IA topoisomerases, including EcTopoI, specifically recognize a C residue and cleave DNA 4 nt downstream [221,276–279]. Likely, this requirement is characteristic for the entire protein family. The results extend these observations and show that a C in a specific context is required for EcTopoI cleavage *in vivo*.

The DNA-binding and cleavage activities of EcTopoI were compared by metagene analysis. Both signals were increased over the background upstream of active TUs, where the GapR-Seq signal is significantly depleted, indicating the attraction of EcTopoI by transcription-generated negative supercoiling followed by relaxation of bound DNA (**Figures 34A** and **32C** for GapR-Seq data). Intriguingly, compared to regions upstream of TUs, cleavage activity of EcTopoI was significantly lower at promoters and within

the TU bodies (**Figure 34B**), i.e., at sites where the formation of complexes with RNAP is expected. Indeed, the overlap of TCSs with RNAP peaks is decreased (Monte-Carlo simulation with 10000 iterations, p-value 1.6e-2; **Figure 34C**).

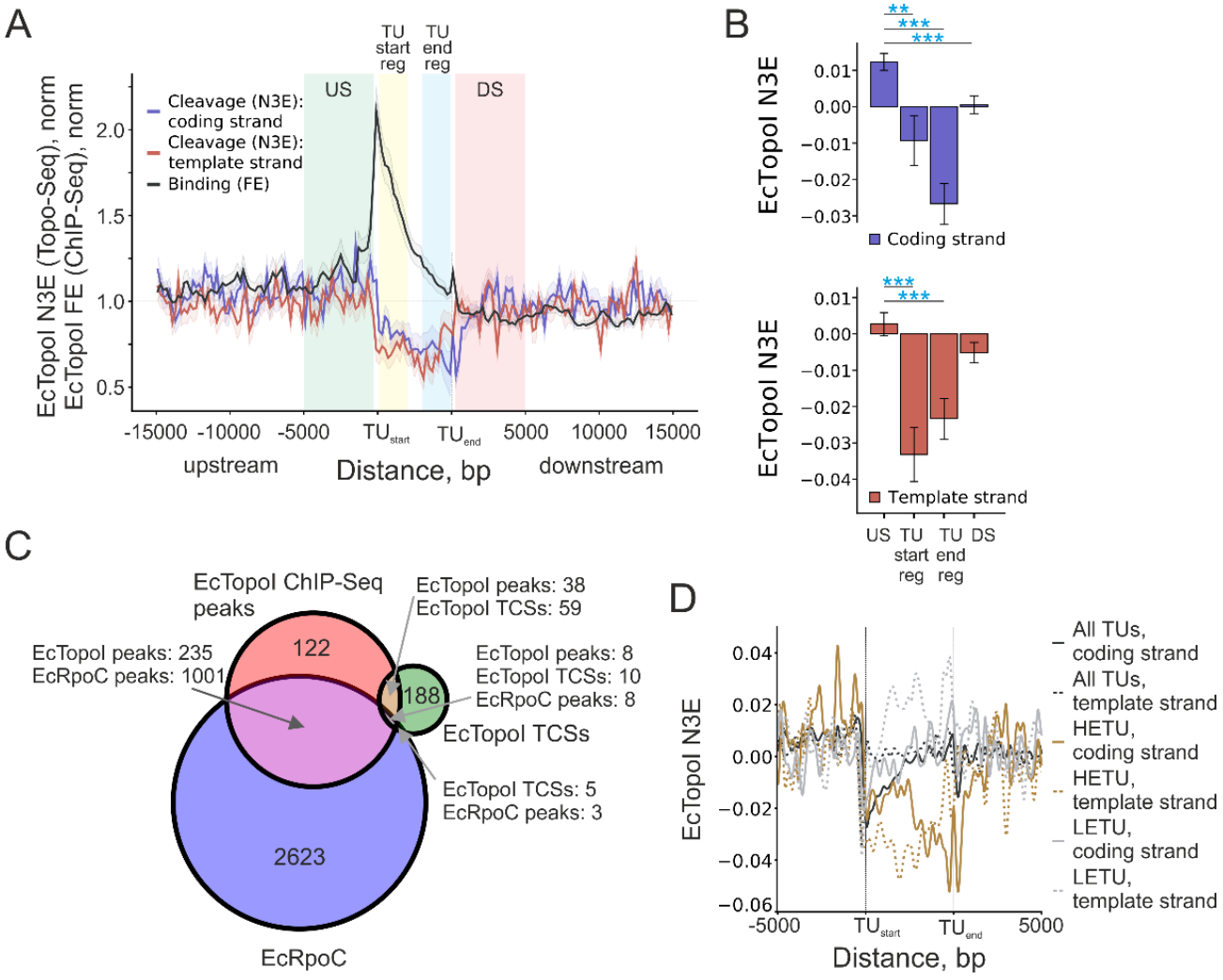


Figure 34. **EcTopoI cleavage activity and transcription**. (**A**) Comparison of EcTopoI ChIP-Seq data (untreated condition, black curve) and EcTopoI Topo-Seq data. EcTopoI Topo-Seq signal is strand-specific and shown for coding (blue curve) and template (red curve) strands separately. Confidence bands around the mean metagene signal are represented by ±SEM. Metagene analysis was performed for the HETU set. Regions used for further quantification of enrichment in panel B (US, TU start region, TU end region, and DS) are shown by colored areas on the plot. (**B**) Comparison of EcTopoI cleavage signal in different regions relative to HETUs. Differences in the signal were tested using the Welch t-test. Significance is indicated by asterisks. Error bars represent ±SEM. (**C**) Correspondence between EcTopoI ChIP-Seq peaks, topoisomerase cleavage sites (TCSs) identified with Topo-Seq, and EcRpoC ChIP-Seq peaks. The Venn diagram demonstrates the number of overlapping peaks and TCSs. (**D**) A metagene plot comparing the

cleavage activity (N3E) of EcTopoI for all TUs, LETU, and HETU subsets. Cleavage is shown for coding and template strands separately.

Overall, the data for EcTopoI resemble the activity pattern of eukaryotic Top1, which remains inactive in complex with RNAPII until it is triggered by an RNAPII stalled by torsional stress [280]. A similar, independently evolved mechanism could be at work for bacterial topoisomerases. A naturally occurring modification of EcTopoI, N^{ε}-acetylation of lysines, was reported to reduce the enzyme's activity *in vivo* [281]. I speculate that this modification can regulate the activity of EcTopoI when it's in complex with RNAP. I predict that EcTopoI activity is inhibited by acetylation at promoters and at the beginning of TU bodies, and is activated by deacetylation (probably, by the CobB protein) at the end of TUs. Deacetylation may be triggered by conformational changes within the RNAP:EcTopoI complex caused by RNAP stalling or by extensive torsional stress (e.g., excessive negative supercoiling). Interestingly, while the cleavage is decreased in HETU bodies, it is increased within inside LETUs, particularly towards their ends (**Figure 34D**). I speculate that in HETUs, where EcTopoI remains inactive, RNAP molecules move in convoy and mutually annihilate positive and negative supercoils generated by polymerases elongating consequently [282,283]. In contrast, EcTopoI activity is needed to remove excessive negative supercoiling generated by individual RNAP molecules in LETUs. Alternatively, the apparent absence of the TCSs within active TUs may be explained by the activity of transcription-coupled DNA repair pathways which hypothetically can remove such complexes *in situ* [284,285]. If true, RNAP:EcTopoI (G116S M320V) complexes, stalled due to the catalytic activity of the "intrinsically poisoned" EcTopoI double mutant, can be recognized by Mfd or NusA-UvrD and be removed by nucleotide-excision repair pathway.
I conclude, that for the optimal activity of EcTopoI, the binding/cleavage motif should be "activated" by melting of DNA, excessive negative supercoiling upstream of TUs, or, perhaps, by hypothetical signaling such as deacetylation of EcTopoI when in complex with RNAP at the end of TUs.

4.21 *E. coli* and *Mycobacterium* rely on different versions of the twin-domain model

Analysis of the genome-wide distribution of mycobacterial topoisomerases (DNA gyrase and TopoI) and RNAP reported by the Nagaraja and co-workers showed co-localization of all three enzymes (Ahmed *et al.*, 2017; Rani and Nagaraja, 2018). I re-examined the published ChIP-Seq data. Indeed, a significant colocalization between RNAP and gyrase, and RNAP and TopoI was observed (**Supplementary Figure 3A-D**). Surprisingly, in *Mycobacterium*, both topoisomerases are enriched within the TU bodies with the highest signal near the transcription start sites. In contrast to *E. coli,* there is no evidence for supercoiling diffusing away from the TUs (**Figures 35A**, **B**). These patterns define two possible variations of the original

Liu & Wang's scheme: an "open" model for *E. coli* in which supercoiling domains extend on DNA bi-directionally over a substantial distance from transcribing RNAP and a "closed" model for *Mycobacterium* where the supercoiling domains are trapped within RNAP-topoisomerase I/gyrase complex and cannot escape. Likely, *S. pneumoniae* has also an "open" model, because EcTopoI enrichment is observed for a long upstream region of TUs (**Supplementary Figure 3E**, based on data published in [181]). Possibly, both mycobacterial topoisomerases form a relatively stable complex with RNAP, in which they fully relax the supercoils generated within TUs during transcription and, therefore, their activity is not needed in adjacent regions.

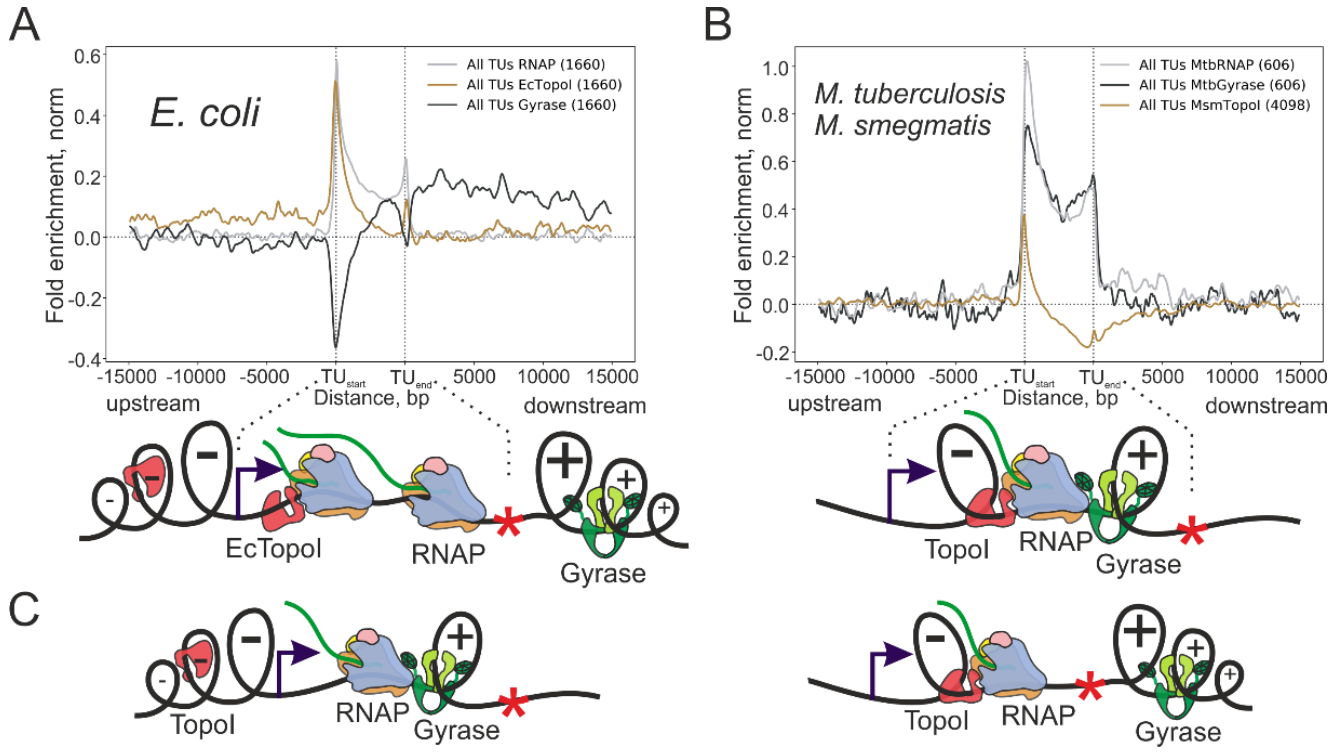


Figure 35. Possible variations of the twin-domain model. Average normalized enrichment of TopoI, DNA gyrase, and RNAP over transcription units of *E. coli* (**A**, "open" model) and *Mycobacterium* (**B**, "closed" model). Graphical representation of twin-domain sub-models is shown below. ChIP-Seq data for *M. tuberculosis* MtbRNAP, MtbGyrase, and *M. smegmatis* MsmTopoI was taken from publicly available datasets [66,112,286]. (**C**) Other "semi-open" hypothetical variations of the twin-domain model, based on the interaction of key topoisomerases (TopoI, DNA gyrase) with RNAP and their activity within a complex. For ChIP-Seq, fold enrichment is given relative to the input sample.

Hypothetically, “semi-open” models may also exist: i) when gyrase is highly-active in complex with RNAP while TopoI does not interact with RNAP, allowing negative supercoils to diffuse freely upstream of TUs; ii) alternatively, when only TopoI is active within the RNAP complex, allowing positive supercoils to diffuse downstream of TUs where they are relaxed ahead of RNAP by free gyrase (**Figure 35C**).

Additional variations of these models may also be possible depending on the balance of topoisomerase and RNAP activities. If topoisomerase acts faster than the rate at which RNAPs generate supercoiling, diffusion of supercoiling will be limited due to rapid relaxation by topoisomerases. Following this logic, EcTopoI in complex with RNAP may allow a portion of unconstrained supercoiling to diffuse upstream, where it is subsequently relaxed by free topoisomerase. Our observation that EcTopoI does not actively cleave DNA in TUs when associated with RNAP is consistent with this hypothesis. In contrast, for *Mycobacterium* interactions between enzymes are likely playing a critical role in the establishing of the “closed” model as M RNAP elongation rate and M DNA gyrase turnover rate are proportionally slower than for *E. coli* enzymes (approximately 10 times slower) [56,111,287]. Of note, MtTopoI relaxation activity is also lower than EcTopoI activity, but to less extent [278].

Chapter 5. Conclusions

In this thesis, a Topo-Seq method for genome-wide and high-resolution identification of topoisomerase cleavage sites was developed. Using this method, I mapped sites of activity for three major *E. coli* topoisomerases – DNA gyrase, TopoIV, and Topo I.

For DNA gyrase, several thousand cleavage sites distributed throughout the *E. coli* genome were identified. The sites were enriched in downstream regions of highly transcribed transcription units, where positive supercoiling is accumulated, in agreement with Liu & Wang twin-domain model. Treatment with RNAP inhibitor rifampicin abolishes this pattern, indicating that gyrase accumulation is supercoiling and transcription dependent. Consistently, gyrase cleavage sites were found accumulated downstream of rRNA operons in *C. crescentus* genome. Topo-Seq with stationary phase *E. coli* culture revealed that gyrase cleavage sites are re-distributed in the *E. coli* genome following the changes in transcriptional profile, again demonstrating the attraction of DNA gyrase by positive supercoiling *in vivo*. A time-resolved series of Topo-Seq experiments with synchronously replicating cells demonstrated that DNA gyrase is colocalized with moving replisome. The obtained data indicates that gyrase is involved in relaxation of positive supercoiling generated by transcription and replication.

Using ChIP-Seq and Topo-Seq, binding and cleavage sites of topoisomerase I were mapped genome-wide. Enrichment of the topoisomerase was significantly increased in the upstream regions of highly transcribed transcription units, where according to the twin domain model, negative supercoiling is accumulated. Consistently, TopoI enrichment was also increased in intergenic regions, where accumulation of negative supercoiling is expected. TopoI enrichment is transcription specific, as treatment with rifampicin removes accumulation in the upstream regions. The data indicates that TopoI is attracted by negative supercoiling generated by transcription.

Several thousand of TopoIV cleavage sites were identified in the *E. coli* genome with Topo-Seq. In accordance with previous data, the strongest cleavage site was identified at *dif*-site, supporting the role of the topoisomerase in decatenation of sister chromosomes. At the same time, cleavage activity of TopoIV was significantly reduced in the Ter macrodomain, where complexes between TopoIV and MukBEF SMC are disassembled by the action of MatP. Interestingly, TopoIV cleavage activity was dramatically increased nearby oriC (a novel observation). Presumably, this enrichment highlights the TopoIV activity at early stages of DNA replication, probably, when TopoIV is organized with MukBEF in multimeric complexes colocalized with *oriC*. However, the role of TopoIV in organization and spatial positioning of the origin

regions during replication is still unknown. I speculate that accumulated and spatially organized activity of TopoIV within such complexes is required for fast and efficient decatenation of origin regions, which further stimulates the proper positioning of origins and initiates the segregation of replicating chromosomes. TopoIV enrichment was found increased in the downstream regions of highly transcribed transcription units (similar to DNA gyrase), indicating that *in vivo* it is involved in the relaxation of transcription-induced positive supercoiling.

Taken together, generated data indicates that TopoI relaxes negative supercoiling generated by transcription, while DNA gyrase and TopoIV relax positive supercoiling. The enrichments of the topoisomerases can be tracked 12-15 kb upstream or downstream of transcription units which gives a diffusion range for unconstrained supercoiling in the *E. coli* genome. We, therefore, named this variation of the twin-domain model "open", an opposite to the "closed" variation observed for *Mycobacterium*. In the "closed" model, all topoisomerases are colocalized with RNAP within transcription units and there is no enrichment upstream or downstream. "Closed" model may be characteristic for bacteria having strong interactions between the topoisomerases and RNAP and/or when the catalytic activity of the topoisomerases exceeds the rate of supercoiling generation. In both cases supercoiling will be efficiently relaxed *in situ* and its diffusion is minimized.

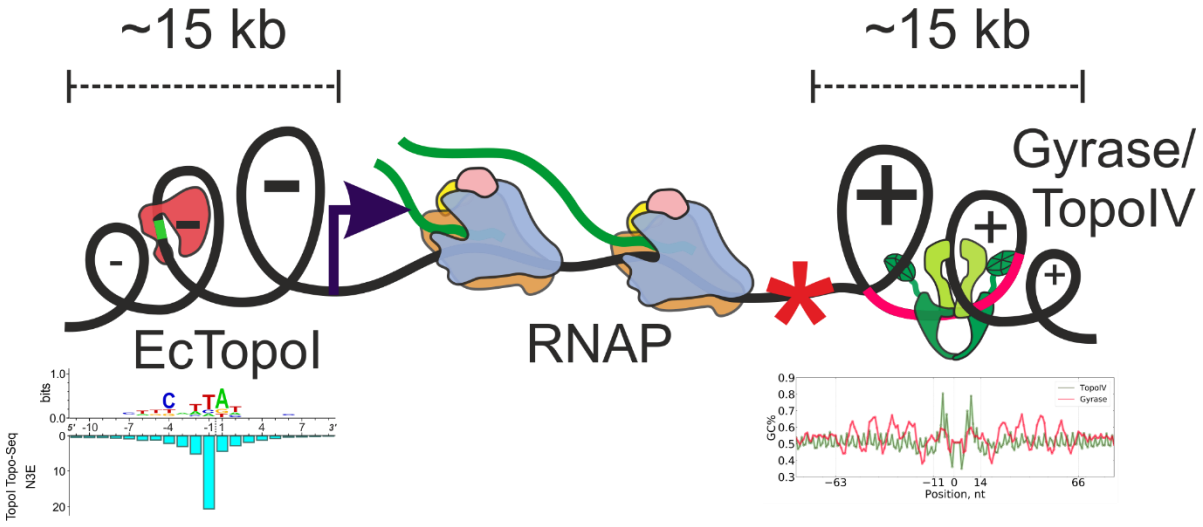


Figure 36. A scheme illustrating genomic factors attracting DNA topoisomerases: supercoiling (negative in the upstream regions of TUs for TopoI, positive in downstream regions for DNA gyrase and TopoIV) and local DNA motifs. In *E. coli* genome unconstrained transcription-induced supercoiling can diffuse over distances of up to 15 kb.

Single-nucleotide resolution of Topo-Seq allows us to build cleavage motifs for the studied topoisomerases. For DNA gyrase a long (~120 nt) symmetric and degenerate motif was observed. The motif has two periodic regions with a period corresponding to the period of DNA in a B-form. Periodic regions reflect wrapping of DNA around the gyrase CTDs. The central region of the motif was biased with the effects of topoisomerase poisons used in Topo-Seq, however, by comparison of obtained cleavage motifs for different drugs, an unbiased motif was constructed. Scanning of DNA sequences with gyrase motif revealed that known strong gyrase sites or strong cleavage sites identified by Topo-Seq all have a high similarity to the motif identified. As a validation, with Topo-Seq, a high-scored gyrase site was confirmed in the pBR322 plasmid. Interestingly, background gyrase cleavage was observed without any poison at several strong sites for the stationary phase *E. coli* culture, indicating that gyrase may reside for a long time at optimal substrates *in vivo*. In contrast to gyrase, TopoIV cleavage motif was lacking periodic regions, supporting that TopoIV does not wrap the DNA G-segment, but rather transfer an unrelated DNA region during catalysis. For *E. coli* TopoI, a short (~15 nt), asymmetric, AT-rich binding/cleavage motif was found with both ChIP-Seq and Topo-Seq. In accordance with the previous studies, in the -4 position to the cleavage site, a conserved C nucleotide is positioned. The AT-richness of the observed motif probably reflects the thermodynamic advantage of such sequences, as for catalysis the topoisomerase first needs to melt DNA duplex. A single-nucleotide resolution achieved with Topo-Seq, and its high sensitivity make the method superior over other ChIP-based methods applied for bacterial topoisomerases up to date.

Our data demonstrate that the topoisomerases choose their sites by sensing both DNA topology (supercoiling of a particular sign) and local sequence properties (motifs) (**Figure 36**). Optimal sequences (e.g., strong gyrase sites like Mu SGS) may be recognized *per se*. Sub-optimal sequences, which comprise the majority of topo-sites, however, should be “activated” by supercoiling to be efficiently recognized and cleaved by a topoisomerase.

Bibliography

1. Crick, F.H.C. Linking Numbers and Nucleosomes. *Proc. Natl. Acad. Sci. U. S. A.* **1976**, *73*, 2639–2643, doi:10.1073/pnas.73.8.2639.

2. Mirkin, S. DNA Topology: Fundamentals. *Encycl. life Sci.* 2001.

3. Bates, A.; Maxwell, A. *DNA Topology*; Second Edi.; Oxford University Press, 2005; ISBN 019856709X.

4. Maxwell, A.; Bush, N.G.; Evans-Roberts, K. DNA Topoisomerases. *EcoSal Plus* **2015**, *6*, doi:10.1128/ecosalplus.ESP-0010-2014.

5. Seol, Y.; Neuman, K.C. The Dynamic Interplay between DNA Topoisomerases and DNA Topology. *Biophys. Rev.* **2016**, *8*, 101–111, doi:10.1007/s12551-016-0240-8.

6. Bizard, A.H.; Hickson, I.D. The Many Lives of Type IA Topoisomerases. *J. Biol. Chem.* **2020**, *295*, 7138–7153, doi:10.1074/jbc.REV120.008286.

7. Basu, A.; Hobson, M.; Lebel, P.; Fernandes, L.E.; Tretter, E.M.; Berger, J.M.; Bryant, Z. Dynamic Coupling between Conformations and Nucleotide States in DNA Gyrase Article. *Nat. Chem. Biol.* **2018**, *14*, 565–574, doi:10.1038/s41589-018-0037-0.

8. Bates, A.D.; Berger, J.M.; Maxwell, A. The Ancestral Role of ATP Hydrolysis in Type II Topoisomerases: Prevention of DNA Double-Strand Breaks. *Nucleic Acids Res.* **2011**, *39*, 6327–6339, doi:10.1093/nar/gkr258.

9. Brown, P.O.; Cozzarelli, N.R. A Sign Inversion Mechanism for Enzymatic Supercoiling of DNA. *Science* **1979**, *206*, 1081–1083, doi:10.1126/science.227059.

10. Forterre, P.; Simonetta, G.; Daniele, G.; Serre, M.-C. Origin and Evolution of DNA Topoisomerases. *Biochimie* **2007**, *89*, 427–446, doi:10.1016/j.biochi.2006.12.009.

11. Forterre, P.; Gadelle, D. Phylogenomics of DNA Topoisomerases : Their Origin and Putative Roles in the Emergence of Modern Organisms. *Nucleic Acids Res.* **2009**, *37*, 679–692, doi:10.1093/nar/gkp032.

12. Dutta, R.; Inouye, M. GHKL, an Emergent ATPase/Kinase Superfamily Rinku Dutta And. *TIBS* **2000**, *25*, 24–28.

13. Broeck, A. Vanden; Lotz, C.; Ortiz, J.; Lamour, V. Cryo-EM Structure of the Complete E. Coli DNA Gyrase Nucleoprotein Complex. *Nat. Commun.* **2019**, *10*, 1–12, doi:10.1038/s41467-019-12914-y.

14. Schmidt, B.H.; Burgin, A.B.; Deweese, J.E.; Osheroff, N.; Berger, J.M. A Novel and Unified Two-Metal Mechanism for DNA Cleavage by Type II and IA Topoisomerases Bryan. *Nature* **2010**, *465*, 641–644, doi:10.1038/nature08974.A.

15. Berger, J.M.; Gambling, S.; Harrison, S.; Wang, J.C. Structure and Mechanism of DNA Topoisomerase II. *Nature* **1996**, *379*, 225–232.

16. Clarke, D.J.; Azuma, Y. Non-Catalytic Roles of the Topoisomerase IIα C-Terminal Domain. *Int. J. Mol. Sci.* **2017**, *18*, doi:10.3390/ijms18112438.

17. Vos, S.M.; Lee, I.; Berger, J.M. Distinct Regions of the Escherichia Coli ParC C-Terminal Domain Are Required for Substrate Discrimination by Topoisomerase IV. *J. Mol. Biol.* **2014**, *425*,

doi:10.1016/j.jmb.2013.04.033.Distinct.

18. Mcclendon, A.K.; Gentry, A.C.; Dickey, J.S.; Brinch, M.; Bendsen, S.; Andersen, A.H.; Osheroff, N. Bimodal Recognition of DNA Geometry by Human Topoisomerase IIα: Preferential Relaxation of Positively Supercoiled DNA Requires Elements in the C-Terminal Domain. *Biochemistry* **2009**, *47*, 13169–13178, doi:10.1021/bi800453h.Bimodal.

19. Morais Cabral, J.H.; Jackson, A.P.; Smith, C. V.; Shikotra, N.; Maxwell, A.; Liddington, R.C. Crystal Structure of the Breakage-Reunion Domain of DNA Gyrase. *Nature* **1997**, *388*, 903–906, doi:10.1038/42294.

20. Vologodskii, a V; Zhang, W.; Rybenkov, V. V; Podtelezhnikov, a a; Subramanian, D.; Griffith, J.D.; Cozzarelli, N.R. Mechanism of Topology Simplification by Type II DNA Topoisomerases. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 3045–3049, doi:10.1073/pnas.061029098.

21. Dong, K.C.; Berger, J.M. Structural Basis for Gate-DNA Recognition and Bending by Type IIA Topoisomerases. *Nature* **2007**, *450*, 1201–1205, doi:10.1038/nature06396.

22. Vologodskii, A. Theoretical Models of DNA Topology Simplification by Type IIA DNA Topoisomerases. *Nucleic Acids Res.* **2009**, *37*, 3125–3133, doi:10.1093/nar/gkp250.

23. Gubaev, A.; Klostermeier, D. DNA-Induced Narrowing of the Gyrase N-Gate Coordinates T-Segment Capture and Strand Passage. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 14085–14090, doi:10.1073/pnas.1102100108.

24. Roca, J. The Path of the DNA along the Dimer Interface of Topoisomerase II. *J. Biol. Chem.* **2004**, *279*, 25783–25788, doi:10.1074/jbc.M402555200.

25. Gubaev, A.; Klostermeier, D. The Mechanism of Negative DNA Supercoiling: A Cascade of DNA-Induced Conformational Changes Prepares Gyrase for Strand Passage. *DNA Repair (Amst).* **2014**, *16*, 130–141, doi:10.1016/j.dnarep.2014.06.006.

26. Lee, S.; Jung, S.-R.; Heo, K.; Byl, J.A.W.; Deweese, J.E.; Osheroff, N.; Hohng, S. DNA Cleavage and Opening Reactions of Human Topoisomerase IIα Are Regulated via Mg2+ - Mediated Dynamic Bending of Gate-DNA. *PNAS* **2012**, *109*, 1–6, doi:10.1073/pnas.1115704109.

27. Rudolph, M.G.; Klostermeier, D. Mapping the Spectrum of Conformational States of the DNA- and C-Gates in Bacillus Subtilis Gyrase. *J. Mol. Biol.* **2013**, *425*, 2632–2640, doi:10.1016/j.jmb.2013.04.010.

28. Baird, C.L.; Harkins, T.T.; Morris, S.K.; Lindsley, J.E. Topoisomerase II Drives DNA Transport by Hydrolyzing One ATP. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 13685–13690, doi:10.1073/pnas.96.24.13685.

29. Basu, A.; Parente, A.C.; Bryant, Z. Structural Dynamics and Mechanochemical Coupling in DNA Gyrase. *J. Mol. Biol.* **2016**, *428*, 1833–1845, doi:10.1016/j.jmb.2016.03.016.

30. Sugino, A.; Higgins, N.P.; Brown, P.O.; Peebles, C.L.; Cozzarelli, N.R. Energy Coupling in DNA Gyrase and the Mechanism of Action of Novobiocin. *Proc. Natl. Acad. Sci. U. S. A.* **1978**, *75*, 4838–4842, doi:10.1073/pnas.75.10.4838.

31. Bates, A.D.; O'Dea, M.H.; Gellert, M. Energy Coupling in Escherichia Coli DNA Gyrase: The Relationship between Nucleotide Binding, Strand Passage, and DNA Supercoiling. *Biochemistry* **1996**, *35*, 1408–1416, doi:10.1021/bi952433y.

32. Corbett, K.D.; Berger, J.M. Structure of the Topoisomerase VI-B Subunit: Implications for Type II Topoisomerase Mechanism and Evolution. *EMBO J.* **2003**, *22*, 151–163, doi:10.1093/emboj/cdg008.

33. Corbett, K.D.; Benedetti, P.; Berger, J.M. Holoenzyme Assembly and ATP-Mediated Conformational Dynamics of Topoisomerase VI. *Nat. Struct. Mol. Biol.* **2007**, *14*, 611–619, doi:10.1038/nsmb1264.

34. Graille, M.; Cladie, L.; Durand, D.; Lecointe, F.; Gadelle, D.; Quevillon-Cheruel, S.; Vachette, P.; Forterre, P.; Tilbeurgh, H. Van Article Crystal Structure of an Intact Type II DNA Topoisomerase : Insights into DNA Transfer Mechanisms. *Structure* **2008**, *16*, 360–370, doi:10.1016/j.str.2007.12.020.

35. Buhler, C.; Lebbink, J.H.G.; Bocs, C.; Ladenstein, R.; Forterre, P. DNA Topoisomerase VI Generates ATP-Dependent Double-Strand Breaks with Two-Nucleotide Overhangs. *J. Biol. Chem.* **2001**, *276*, 37215–37222, doi:10.1074/jbc.M101823200.

36. Morrison, A.; Cozzarelli, N.R. Site-Specific Cleavage of DNA by E. Coli DNA Gyrase. *Cell* **1979**, *17*, 175–184, doi:10.1016/0092-8674(79)90305-2.

37. Ward, D.; Newton, A. Requirement of Topoisomerase IV ParC and ParE Genes for Cell Cycle Progression and Developmental Regulation in Caulobacter Crescentus. *Mol. Microbiol.* **1997**, *26*, 897–910, doi:10.1046/j.1365-2958.1997.6242005.x.

38. Reuß, D.R.; Faßhauer, P.; Mroch, P.J.; Ul-Haq, I.; Koo, B.M.; Pöhlein, A.; Gross, C.A.; Daniel, R.; Brantl, S.; Stülke, J. Topoisomerase IV Can Functionally Replace All Type 1A Topoisomerases in Bacillus Subtilis. *Nucleic Acids Res.* **2019**, *47*, 5231–5242, doi:10.1093/nar/gkz260.

39. Huang, T.W.; Hsu, C.C.; Yang, H.Y.; Chen, C.W. Topoisomerase IV Is Required for Partitioning of Circular Chromosomes but Not Linear Chromosomes in Streptomyces. *Nucleic Acids Res.* **2013**, *41*, 10403–10413, doi:10.1093/nar/gkt757.

40. Aubry, A.; Mark Fisher, L.; Jarlier, V.; Cambau, E. First Functional Characterization of a Singly Expressed Bacterial Type II Topoisomerase: The Enzyme from Mycobacterium Tuberculosis. *Biochem. Biophys. Res. Commun.* **2006**, *348*, 158–165, doi:10.1016/j.bbrc.2006.07.017.

41. Ambur, O.H.; Davidsen, T.; Frye, S.A.; Balasingham, S. V.; Lagesen, K.; Rognes, T.; Tønjum, T. Genome Dynamics in Major Bacterial Pathogens. *FEMS Microbiol. Rev.* **2009**, *33*, 453–470, doi:10.1111/j.1574-6976.2009.00173.x.

42. McCutcheon, J.P.; McDonald, B.R.; Moran, N.A. Origin of an Alternative Genetic Code in the Extremely Small and GC-Rich Genome of a Bacterial Symbiont. *PLoS Genet.* **2009**, *5*, doi:10.1371/journal.pgen.1000565.

43. López-Madrigal, S.; Latorre, A.; Porcar, M.; Moya, A.; Gil, R. Complete Genome Sequence of "Candidatus Tremblaya Princeps" Strain PCVAL, an Intriguing Translational Machine below the Living-Cell Status. *J. Bacteriol.* **2011**, *193*, 5587–5588, doi:10.1128/JB.05749-11.

44. Tamames, J.; Gil, R.; Latorre, A.; Peretó, J.; Silva, F.J.; Moya, A. The Frontier between Cell and Organelle: Genome Analysis of Candidatus Carsonella Ruddii. *BMC Evol. Biol.* **2007**, *7*, 1–7, doi:10.1186/1471-2148-7-181.

45. Collin, F.; Karkare, S.; Maxwell, A. Exploiting Bacterial DNA Gyrase as a Drug Target: Current State and Perspectives. *Appl. Microbiol. Biotechnol.* **2011**, *92*, 479–497, doi:10.1007/s00253-011-

3557-z.

46. Dighe, S.N.; Collet, T.A. Recent Advances in DNA Gyrase-Targeted Antimicrobial Agents. *Eur. J. Med. Chem.* **2020**, *199*, 112326, doi:10.1016/j.ejmech.2020.112326.

47. Hooper, D.C.; Jacoby, G.A. Topoisomerase Inhibitors: Fluoroquinolone Mechanisms of Action and Resistance. *Cold Spring Harb. Perspect. Med.* **2016**, *6*, 1–22, doi:10.1101/cshperspect.a025320.

48. Wohlkonig, A.; Chan, P.F.; Fosberry, A.P.; Homes, P.; Huang, J.; Kranz, M.; Leydon, V.R.; Miles, T.J.; Pearson, N.D.; Perera, R.L.; et al. Structural Basis of Quinolone Inhibition of Type IIA Topoisomerases and Target-Mediated Resistance. *Nat. Struct. Mol. Biol.* **2010**, *17*, 1152–1153, doi:10.1038/nsmb.1892.

49. Bax, B.D.; Chan, P.F.; Eggleston, D.S.; Fosberry, A.; Gentry, D.R.; Gorrec, F.; Giordano, I.; Hann, M.M.; Hennessy, A.; Hibbs, M.; et al. Type IIA Topoisomerase Inhibition by a New Class of Antibacterial Agents. *Nature* **2010**, *466*, 935–940, doi:10.1038/nature09197.

50. Aldred, K.J.; McPherson, S.A.; Turnbough, C.L.; Kerns, R.J.; Osheroff, N. Topoisomerase IV-Quinolone Interactions Are Mediated through a Water-Metal Ion Bridge: Mechanistic Basis of Quinolone Resistance. *Nucleic Acids Res.* **2013**, *41*, 4628–4639, doi:10.1093/nar/gkt124.

51. Pommier, Y.; Marchand, C. Interfacial Inhibitors: Targeting Macromolecular Complexes. *Nat. Rev. Drug Discov.* **2012**, *11*, 25–34, doi:10.1038/nrd3404.

52. Pommier, Y.; Kiselev, E.; Marchand, C.; Pommier, Y.; Kiselev, E.; Marchand, C. Interfacial Inhibitors. *Bioorg. Med. Chem. Lett.* **2015**, *25*, 3961–3965.

53. Alt, S.; Mitchenall, L.A.; Maxwell, A.; Heide, L. Inhibition of DNA Gyrase and DNA Topoisomerase IV of Staphylococcus Aureus and Escherichia Coli by Aminocoumarin Antibiotics. *J. Antimicrob. Chemother.* **2011**, *66*, 2061–2069, doi:10.1093/jac/dkr247.

54. Butler, M.S.; Paterson, D.L. Antibiotics in the Clinical Pipeline in October 2019. *J. Antibiot. (Tokyo).* **2020**, *2019*, 329–364, doi:10.1038/s41429-020-0291-8.

55. Lewis, K. The Science of Antibiotic Discovery. *Cell* **2020**, *181*, 29–45, doi:10.1016/j.cell.2020.02.056.

56. Weidlich, D.; Klostermeier, D. Functional Interactions between Gyrase Subunits Are Optimized in a Species-Specific Manner. *J. Biol. Chem.* **2020**, *295*, 2299–2312, doi:10.1074/jbc.RA119.010245.

57. Rovinskiy, N.S.; Agbleke, A.A.; Chesnokova, O.N.; Patrick Higgins, N. Supercoil Levels in E. Coli and Salmonella Chromosomes Are Regulated by the C-Terminal 35–38 Amino Acids of GyrA. *Microorganisms* **2019**, *7*, doi:10.3390/microorganisms7030081.

58. Guo, M.S.; Haakonsen, D.L.; Zeng, W.; Schumacher, M.A.; Laub, M.T. A Bacterial Chromosome Structuring Protein Binds Overtwisted DNA to Stimulate Type II Topoisomerases and Enable DNA Replication. *Cell* **2018**, *175*, 583-597.e23, doi:10.1016/j.cell.2018.08.029.

59. Ashley, R.E.; Dittmore, A.; McPherson, S.A.; Turnbough, C.L.; Neuman, K.C.; Osheroff, N. Activities of Gyrase and Topoisomerase IV on Positively Supercoiled DNA. *Nucleic Acids Res.* **2017**, *45*, 9611–9624, doi:10.1093/nar/gkx649.

60. Grompone, G.; Ehrlich, S.D.; Michel, B. Replication Restart in GyrB Escherichia Coli Mutants. *Mol. Microbiol.* **2003**, *48*, 845–854, doi:10.1046/j.1365-2958.2003.03480.x.

61. Rizzo, M.F.; Shapiro, L.; Gober, J. Asymmetric Expression of the Gyrase B Gene from the Replication-Competent Chromosome in the Caulobacter Crescentus Predivisional Cell. *J. Bacteriol.* **1993**, *175*, 6970–6981, doi:10.1128/jb.175.21.6970-6981.1993.

62. Ogasawara, N.; Seiki, M.; Yoshikawa, H. Effect of Novobiocin on Initiation of DNA Replication in Bacillus Subtilis [26]. *Nature* 1979, *281*, 702–704.

63. Guha, S.; Udupa, S.; Ahmed, W.; Nagaraja, V. Rewired Downregulation of DNA Gyrase Impacts Cell Division, Expression of Topology Modulators, and Transcription in Mycobacterium Smegmatis. *J. Mol. Biol.* **2018**, *430*, 4986–5001, doi:10.1016/j.jmb.2018.10.001.

64. Nakanishi, A.; Oshida, T.; Matsushita, T.; Imajoh-Ohmi, S.; Ohnuki, T. Identification of DNA Gyrase Inhibitor (GyrI) in Escherichia Coli. *J. Biol. Chem.* **1998**, *273*, 1933–1938, doi:10.1074/jbc.273.4.1933.

65. Jeong, K.S.; Ahn, J.; Khodursky, A.B. Spatial Patterns of Transcriptional Activity in the Chromosome of Escherichia Coli. *Genome Biol.* **2004**, *5*, R86, doi:10.1186/gb-2004-5-11-r86.

66. Ahmed, W.; Sala, C.; Hegde, S.R.; Jha, R.K.; Cole, S.T.; Nagaraja, V. Transcription Facilitated Genome-Wide Recruitment of Topoisomerase I and DNA Gyrase. *PLoS Genet.* **2017**, *13*, e1006754, doi:10.1371/journal.pgen.1006754.

67. Chong, S.; Chen, C.; Ge, H.; Xie, X.S. Mechanism of Transcriptional Bursting in Bacteria. *Cell* **2015**, *158*, 314–326, doi:10.1016/j.cell.2014.05.038.Mechanism.

68. Wu, H.Y.; Shyy, S.; Wang, J.C.; Liu, L.F. Transcription Generates Positively and Negatively Supercoiled Domains in the Template. *Cell* **1988**, *53*, 433–440, doi:10.1016/0092-8674(88)90163-8.

69. Travers, A.; Muskhelishvili, G. DNA Supercoiling - a Global Transcriptional Regulator for Enterobacterial Growth? *Nat. Rev. Microbiol.* **2005**, *3*, 157–169, doi:10.1038/nrmicro1088.

70. Vijayan, V.; Zuzow, R.; O'Shea, E.K. Oscillations in Supercoiling Drive Circadian Gene Expression in Cyanobacteria. *Proc. Natl. Acad. Sci.* **2009**, *106*, 22564–22568, doi:10.1073/pnas.0912673106.

71. Dorman, C.J. DNA Supercoiling and Transcription in Bacteria: A Two-Way Street. *BMC Mol. cell Biol.* **2019**, *20*, 26, doi:10.1186/s12860-019-0211-6.

72. Peter, B.J.; Arsuaga, J.; Breier, A.M.; Khodursky, A.B.; Brown, P.O.; Cozzarelli, N.R. Genomic Transcriptional Response to Loss of Chromosomal Supercoiling in Escherichia Coli. *Genome Biol.* **2004**, *5*, R87, doi:10.1186/gb-2004-5-11-r87.

73. Blot, N.; Mavathur, R.; Geertz, M.; Travers, A.; Muskhelishvili, G. Homeostatic Regulation of Supercoiling Sensitivity Coordinates Transcription of the Bacterial Genome. *EMBO Rep.* **2006**, *7*, 710–715, doi:10.1038/sj.embor.7400729.

74. Sobetzko, P.; Travers, A.; Muskhelishvili, G. Gene Order and Chromosome Dynamics Coordinate Spatiotemporal Gene Expression during the Bacterial Growth Cycle. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, E42-50, doi:10.1073/pnas.1108229109.

75. Martis B., S.; Forquet, R.; Reverchon, S.; Nasser, W.; Meyer, S. DNA Supercoiling: An Ancestral Regulator of Gene Expression in Pathogenic Bacteria? *Comput. Struct. Biotechnol. J.* **2019**, *17*, 1047–1055, doi:10.1016/j.csbj.2019.07.013.

76. Muskhelishvili, G.; Forquet, R.; Reverchon, S.; Nasser, W.; Meyer, S. Coherent Domains of Transcription Coordinate Gene Expression during Bacterial Growth and Adaptation. *Microorganisms* **2019**, *7*, 10–15, doi:10.3390/microorganisms7120694.

77. Pruss, G.J.; Drlicat, K. DNA Supercoiling and Pmkatyotic Transcription Minireview. **1988**, *56*, 1969.

78. Ahmed, W.; Menon, S.; D. N. B. Karthik, P. V.; Nagaraja, V. Autoregulation of Topoisomerase I Expression by Supercoiling Sensitive Transcription. *Nucleic Acids Res.* **2015**, *44*, gkv1088, doi:10.1093/nar/gkv1088.

79. Menzel, R.; Gellert, M. Regulation of the Genes for E. Coli DNA Gyrase: Homeostatic Control of DNA Supercoiling. *Cell* **1983**, *34*, 105–113.

80. Drlica, K. Control of Bacterial DNA Supercoiling. *Mol. Microbiol.* **1992**, *6*, 425–433, doi:10.1111/j.1365-2958.1992.tb01486.x.

81. Szafran, M.J.; Gongerowska, M.; Gutkowski, P.; Zakrzewska-czerwin, J. The Coordinated Positive Regulation of Topoisomerase Genes Maintains Topological Homeostasis in Streptomyces Coelicolor. *J. Bacteriol.* **2016**, *198*, 3016–3028, doi:10.1128/JB.00530-16.Editor.

82. Yamamoto, N.; Droffner, M.L. Mechanisms Determining Aerobic or Anaerobic Growth in the Facultative Anaerobe Salmonella Typhimurium. *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82*, 2077–2081, doi:10.1073/pnas.82.7.2077.

83. Ye, F.; Brauer, T.; Niehus, E.; Drlica, K.; Josenhans, C.; Suerbaum, S. Flagellar and Global Gene Regulation in Helicobacter Pylori Modulated by Changes in DNA Supercoiling. *Int. J. Med. Microbiol.* **2007**, *297*, 65–81, doi:10.1016/j.ijmm.2006.11.006.

84. Hsu, Y.H.; Chung, M.W.; Li, T.K. Distribution of Gyrase and Topoisomerase IV on Bacterial Nucleoid: Implications for Nucleoid Organization. *Nucleic Acids Res.* **2006**, *34*, 3128–3138, doi:10.1093/nar/gkl392.

85. Condemine, G.; Smith, C.L. Transcription Regulates Oxolinic Acid-Induced DNA Gyrase Cleavage at Specific Sites on the E.Coli Chromosome. *Nucleic Acids Res.* **1990**, *18*, 7389–7396, doi:10.1093/nar/18.24.7389.

86. Postow, L.; Hardy, C.D.; Arsuaga, J.; Cozzarelli, N.R. Topological Domain Structure of the Escherichia Coli Chromosome. *Genes Dev.* **2004**, *18*, 1766–1779, doi:10.1101/gad.1207504.Fortunately.

87. Pato, M.L.; Karlok, M.; Wall, C.; Higgins, N.P. Characterization of Mu Prophage Lacking the Central Strong Gyrase Binding Site: Localization of the Block in Replication. *J. Bacteriol.* **1995**, *177*, 5937–5942.

88. Saha, R.P.; Lou, Z.; Meng, L.; Harshey, R.M. Transposable Prophage Mu Is Organized as a Stable Chromosomal Domain of E. Coli. *PLoS Genet.* **2013**, *9*, e1003902, doi:10.1371/journal.pgen.1003902.

89. Sawitzke, J.A.; Austin, S. Suppression of Chromosome Segregation Defects of Escherichia Coli Muk Mutants by Mutations in Topoisomerase I. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 1671–1676, doi:10.1073/pnas.030528397.

90. Le, T.B.; Imakaev, M.V.; Mirny, L.A.; Laub, M.T. High-Resolution Mapping of the Spatial

Organization of a Bacterial Chromosome. *Science (80-.).* **2014**, *342*, 731–734, doi:10.1126/science.1242059.High-resolution.

91. Conin, B.; Billault-Chaumartin, I.; Sayyed, H. El; Du, N.Q.; Cockram, C.; Koszul, R.; Espeli, O. Extended Sister-Chromosome Catenation Leads to Massive Reorganization of the E . Coli Genome. *Nucleic Acids Res.* **2022**, *50*, 2635–2650.

92. Sioud, M.; Possot, O.; Elie, C.; Sibold, L.; Forterre, P. Coumarin and Quinolone Action in Archaebacteria: Evidence for the Presence of a DNA Gyrase-like Enzyme. *J. Bacteriol.* **1988**, *170*, 946–953, doi:10.1128/jb.170.2.946-953.1988.

93. Holmes, M.L.; Dyall-Smith, M.L. Mutations in DNA Gyrase Result in Novobiocin Resistance in Halophilic Archaebacteria. *J. Bacteriol.* **1991**, *173*, 642–648, doi:10.1128/jb.173.2.642-648.1991.

94. Yamashiro, K.; Yamagishi, A. Characterization of the DNA Gyrase from the Thermoacidophilic Archaeon Thermoplasma Acidophilum. *J. Bacteriol.* **2005**, *187*, 8531–8536, doi:10.1128/JB.187.24.8531-8536.2005.

95. Crisona, N.J.; Strick, T.R.; Bensimon, D.; Croquette, V.; Cozzarelli, N.R. Preferential Relaxation of Positively Supercoiled DNA by E. Coli Topoisomerase IV in Single-Molecule and Ensemble Measurements. *Genes Dev.* **2000**, *14*, 2881–2892, doi:10.1101/gad.838900.

96. Peng, H.; Marians, K.J. Decatenation Activity of Topoisomerase-IV during OriC and PBR322 DNA Replication Invitro. *Proc Natl Acad Sci USA* **1993**, *90*, 8571–8575.

97. Ullsperger, C.; Cozzarelli, N.R. Contrasting Enzymatic Activities of Topoisomerase IV and DNA Gyrase from Escherichia Coli. *J. Biol. Chem.* **1996**, *271*, 31549–31555, doi:10.1074/jbc.271.49.31549.

98. Hiasa, H.; DiGate, R.J.; Marians, K.J. Decatenating Activity of Escherichia Coli DNA Gyrase and Topoisomerases I and III during OriC and PBR322 DNA Replication in Vitro. *J. Biol. Chem.* **1994**, *269*, 2093–2099.

99. Seol, Y.; Hardin, A.H.; Strub, M.P.; Charvin, G.; Neuman, K.C. Comparison of DNA Decatenation by Escherichia Coli Topoisomerase IV and Topoisomerase III: Implications for Non-Equilibrium Topology Simplification. *Nucleic Acids Res.* **2013**, *41*, 4640–4649, doi:10.1093/nar/gkt136.

100. Blanche, F.; Cameron, B.; Bernard, F.X.; Maton, L.; Manse, B.; Ferrero, L.; Ratet, N.; Lecoq, C.; Goniot, A.; Bisch, D.; et al. Differential Behaviors of Staphylococcus Aureus and Escherichia Coli Type II DNA Topoisomerases. *Antimicrob. Agents Chemother.* **1996**, *40*, 2714–2720, doi:10.1128/aac.40.12.2714.

101. Deibler, R.; Rahmati, S.; Zechiedrich, E.L. Topoisomerase IV, Alone, Unknots DNA in E. Coli. *Genes Dev.* **2001**, *15*, 748–761, doi:10.1101/gad.872301.DNA.

102. Kampranis, S.C.; Maxwell, A. Conversion of DNA Gyrase into a Conventional Type II Topoisomerase. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 14416–14421, doi:10.1073/pnas.93.25.14416.

103. Hsieh, T.J.; Farh, L.; Huang, W.M.; Chan, N.L. Structure of the Topoisomerase IV C-Terminal Domain: A Broken ??-Propeller Implies a Role as Geometry Facilitator in Catalysis. *J. Biol. Chem.* **2004**, *279*, 55587–55593, doi:10.1074/jbc.M408934200.

104. Hirsch, J.; Klostermeier, D. What Makes a Type IIA Topoisomerase a Gyrase or a Topo IV?

Nucleic Acids Res. **2021**, *49*, 6027–6042.

105. Kato, J. ichi; Nishimura, Y.; Imamura, R.; Niki, H.; Hiraga, S.; Suzuki, H.; Grainge, I.; Bregu, M.; Vazquez, M.; Sivanathan, V.; et al. New Topoisomerase Essential for Chromosome Segregation in E. Coli. *EMBO J.* **1990**, *26*, 4228–4238, doi:10.1038/sj.emboj.7601849.

106. Grainge, I.; Bregu, M.; Vazquez, M.; Sivanathan, V.; Ip, S.C.Y.; Sherratt, D.J. Unlinking Chromosome Catenanes in Vivo by Site-Specific Recombination. *EMBO J.* **2007**, *26*, 4228–4238, doi:10.1038/sj.emboj.7601849.

107. Wang, X.; Reyes-Lamothe, R.; Sherratt, D.J. Modulation of Escherichia Coli Sister Chromosome Cohesion by Topoisomerase IV. *Genes Dev.* **2008**, *22*, 2426–2433, doi:10.1101/gad.487508.

108. Wang, S.C.; Shapiro, L. The Topoisomerase IV ParC Subunit Colocalizes with the Caulobacter Replisome and Is Required for Polar Localization of Replication Origins. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 9251–9256, doi:10.1073/pnas.0402567101.

109. Zechiedrich, E.L.; Khodursky, A.B.; Cozzarelli, N.R. Topoisomerase IV, Not Gyrase, Decatenates Products of Site-Specific Recombination in Escherichia Coli. *Genes Dev.* **1997**, *11*, 2580–2592, doi:10.1101/gad.11.19.2580.

110. Cui, T.; Moro-oka, N.; Ohsumi, K.; Kodama, K.; Ohshima, T.; Ogasawara, N.; Mori, H.; Wanner, B.; Niki, H.; Horiuchi, T. Escherichia Coli with a Linear Genome. *EMBO Rep.* **2007**, *8*, 181–187, doi:10.1038/sj.embor.7400880.

111. Manjunatha, U.H.; Dalal, M.; Chatterji, M.; Radha, D.R.; Visweswariah, S.S.; Nagaraja, V. Functional Characterisation of Mycobacterial DNA Gyrase: An Efficient Decatenase. *Nucleic Acids Res.* **2002**, *30*, 2144–2153, doi:10.1093/nar/30.10.2144.

112. Rani, P.; Nagaraja, V. Genome-Wide Mapping of Topoisomerase I Activity Sites Reveal Its Role in Chromosome Segregation. *Nucleic Acids Res.* **2018**, *47*, 1416–1427, doi:10.1093/nar/gky1271.

113. Sutormin, D.; Rubanova, N.; Logacheva, M.; Ghilarov, D.; Severinov, K. Single-Nucleotide-Resolution Mapping of DNA Gyrase Cleavage Sites across the Escherichia Coli Genome. *Nucleic Acids Res.* **2019**, *47*, 1–16, doi:10.1093/nar/gky1222.

114. Debowski, A.W.; Carnoy, C.; Verbrugghe, P.; Nilsson, H.O.; Gauntlett, J.C.; Fulurija, A.; Camilleri, T.; Berg, D.E.; Marshall, B.J.; Benghezal, M. Xer Recombinase and Genome Integrity in Helicobacter Pylori, a Pathogen without Topoisomerase IV. *PLoS One* **2012**, *7*, doi:10.1371/journal.pone.0033310.

115. Khodursky, A.B.; Peter, B.J.; Schmid, M.B.; DeRisi, J.; Botstein, D.; Brown, P.O.; Cozzarelli, N.R. Analysis of Topoisomerase Function in Bacterial Replication Fork Movement: Use of DNA Microarrays. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 9419–9424, doi:10.1073/pnas.97.17.9419.

116. El Sayyed, H.; Le Chat, L.; Lebailly, E.; Vickridge, E.; Pages, C.; Cornet, F.; Cosentino Lagomarsino, M.; Espéli, O. Mapping Topoisomerase IV Binding and Activity Sites on the E. Coli Genome. *PLoS Genet.* **2016**, *12*, 1–22, doi:10.1371/journal.pgen.1006025.

117. Brochu, J.; Vlachos-Breton, E.; Sutherland, S.; Martel, M.; Drolet, M. Topoisomerases I and III Inhibit R-Loop Formation to Prevent Unregulated Replication in the Chromosomal Ter Region of Escherichia Coli. *PLoS Genet.* **2018**, *9*, 1–25.

118. Kang, S.; Han, J.S.; Park, J.H.; Skarstad, K.; Hwang, D.S. SeqA Protein Stimulates the Relaxing

and Decatenating Activities of Topoisomerase IV. *J. Biol. Chem.* **2003**, *278*, 48779–48785, doi:10.1074/jbc.M308843200.

119. Joshi, M.C.; Magnan, D.; Montminy, T.P.; Lies, M.; Stepankiw, N.; Bates, D. Regulation of Sister Chromosome Cohesion by the Replication Fork Tracking Protein SeqA. *PLoS Genet.* **2013**, *9*, doi:10.1371/journal.pgen.1003673.

120. Li, Y.; Stewart, N.K.; Berger, A.J.; Vos, S.; Schoeffler, A.J.; Berger, J.M.; Chait, B.T.; Oakley, M.G. Escherichia Coli Condensin MukB Stimulates Topoisomerase IV Activity by a Direct Physical Interaction. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 18832–18837, doi:10.1073/pnas.1008678107.

121. Hayama, R.; Marians, K.J. Physical and Functional Interaction between the Condensin MukB and the Decatenase Topoisomerase IV in Escherichia Coli. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 18826–18831, doi:10.1073/pnas.1008140107.

122. Espeli, O.; Lee, C.; Marians, K.J. A Physical and Functional Interaction between Escherichia Coli FtsK and Topoisomerase IV. *J. Biol. Chem.* **2003**, *278*, 44639–44644, doi:10.1074/jbc.M308926200.

123. Hojgaard, A.; Szerlong, H.; Tabor, C.; Kuempel, P. Norfloxacin-Induced DNA Cleavage Occurs at the Dif Resolvase Locus in Escherichia Coli and Is the Result of Interaction with Topoisomerase IV. *Mol. Microbiol.* **1999**, *33*, 1027–1036, doi:10.1046/j.1365-2958.1999.01545.x.

124. Narayanan, S.; Janakiraman, B.; Kumar, L.; Radhakrishnan, S.K. A Cell Cycle-Controlled Redox Switch Regulates the Topoisomerase IV Activity. *Genes Dev.* **2015**, *29*, 1175–1187, doi:10.1101/gad.257030.114.

125. Helgesen, E.; Fossum-Raunehaug, S.; Sætre, F.; Schink, K.O.; Skarstad, K. Dynamic Escherichia Coli SeqA Complexes Organize the Newly Replicated DNA at a Considerable Distance from the Replisome. *Nucleic Acids Res.* **2015**, *43*, 2730–2743, doi:10.1093/nar/gkv146.

126. Vos, S.M.; Stewart, N.K.; Oakley, M.G.; Berger, J.M. Structural Basis for the MukB-Topoisomerase IV Interaction and Its Functional Implications in Vivo. *EMBO J.* **2013**, *32*, 2950–2962, doi:10.1038/emboj.2013.218.

127. Kumar, R.; Bahng, S.; Marians, K.J. The MukB-Topoisomerase IV Interaction Mutually Suppresses Their Catalytic Activities. *Nucleic Acids Res.* **2022**, *50*, 2621–2634.

128. Zawadzki, P.; Stracy, M.; Ginda, K.; Zawadzka, K.; Lesterlin, C.; Kapanidis, A.N.; Sherratt, D.J. The Localization and Action of Topoisomerase IV in Escherichia Coli Chromosome Segregation Is Coordinated by the SMC Complex, MukBEF. *Cell Rep.* **2015**, *13*, 2587–2596, doi:10.1016/j.celrep.2015.11.034.

129. Badrinarayanan, A.; Reyes-Lamothe, R.; Uphoff, S.; Leake, M.C.; Sherratt, D.J. In Vivo Architecture and Action of Bacterial Structural Maintenance of Chromosome Proteins. *Science (80-.).* **2012**, *338*, 528–531, doi:10.1126/science.1227126.

130. Badrinarayanan, A.; Lesterlin, C.; Reyes-Lamothe, R.; Sherratt, D. The Escherichia Coli SMC Complex, MukBEF, Shapes Nucleoid Organization Independently of DNA Replication. *J. Bacteriol.* **2012**, *194*, 4669–4676, doi:10.1128/JB.00957-12.

131. Nicolas, E.; Upton, A.L.; Uphoff, S.; Henry, O.; Badrinarayanan, A.; Sherratt, D. The SMC Complex MukBEF Recruits Topoisomerase IV to the Origin of Replication Region in Live

Escherichia Coli. *MBio* **2014**, *5*, 1–10, doi:10.1128/mBio.01001-13.

132. Tran, N.T.; Laub, M.T.; Le, T.B.K. SMC Progressively Aligns Chromosomal Arms in Caulobacter Crescentus but Is Antagonized by Convergent Transcription. *Cell Rep.* **2017**, *20*, 2057–2071, doi:10.1016/j.celrep.2017.08.026.

133. Mercier, R.; Petit, M.A.; Schbath, S.; Robin, S.; El Karoui, M.; Boccard, F.; Espéli, O. The MatP/MatS Site-Specific System Organizes the Terminus Region of the E. Coli Chromosome into a Macrodomain. *Cell* **2008**, *135*, 475–485, doi:10.1016/j.cell.2008.08.031.

134. Nolivos, S.; Upton, A.L.; Badrinarayanan, A.; Müller, J.; Zawadzka, K.; Wiktor, J.; Gill, A.; Arciszewska, L.; Nicolas, E.; Sherratt, D. MatP Regulates the Coordinated Action of Topoisomerase IV and MukBEF in Chromosome Segregation. *Nat. Commun.* **2016**, *7*, doi:10.1038/ncomms10466.

135. Lioy, V.S.; Cournac, A.; Marbouty, M.; Duigou, S.; Mozziconacci, J.; Espéli, O.; Boccard, F.; Koszul, R. Multiscale Structuring of the E. Coli Chromosome by Nucleoid-Associated and Condensin Proteins. *Cell* **2018**, *172*, 771-783.e18, doi:10.1016/j.cell.2017.12.027.

136. Mäkelä, J.; Sherratt, D.J. Organization of the Escherichia Coli Chromosome by a MukBEF Axial Core. *Mol. Cell* **2020**, *78*, 250-260.e5, doi:10.1016/j.molcel.2020.02.003.

137. Fisher, G.L.M.; Bolla, J.R.; Rajasekar, K. V; Mäkelä, J.; Rachel, B.; Zhou, M.; P, P.J.; Stracy, M.; Robinson, C. V; Arciszewska, L.K.; et al. Competitive Binding of MatP and Topoisomerase IV to the MukB Hinge Domain. *Elife* **2021**, doi:10.7554/eLife.70444.

138. Blakely, G.; May, G.; McCulloch, R.; Arciszewska, L.K.; Burke, M.; Lovett, S.T.; Sherratt, D.J. Two Related Recombinases Are Required for Site-Specific Recombination at *Dif* and *Cer* in *E. Coli* K12. *Cell* **1993**, *75*, 351–361.

139. Keller, A.N.; Xin, Y.; Boer, S.; Reinhardt, J.; Baker, R.; Arciszewska, L.K.; Lewis, P.J.; Sherratt, D.J.; Löwe, J.; Grainge, I. Activation of Xer-Recombination at Dif: Structural Basis of the FtsKγ-XerD Interaction. *Sci. Rep.* **2016**, *6*, 1–12, doi:10.1038/srep33357.

140. Sivanathan, V.; Emerson, J.E.; Pages, C.; Cornet, F.; Sherratt, D.J.; Arciszewska, L.K. KOPS-Guided DNA Translocation by FtsK Safeguards Escherichia Coli Chromosome Segregation. *Mol. Microbiol.* **2009**, *71*, 1031–1042, doi:10.1111/j.1365-2958.2008.06586.x.

141. Deghorain, M.; Pages, C.; Meile, J.-C.; Stouf, M.; Capiaux, H.; Mercier, R.; Lesterlin, C.; Hallet, B.; Cornet, F. A Defined Terminal Region of the E . Coli Chromosome Shows Late Segregation and High FtsK Activity. *PLoS One* **2011**, *6*, 1–8, doi:10.1371/journal.pone.0022164.

142. Stouf, M.; Meile, J.-C.; Cornet, F. FtsK Actively Segregates Sister Chromosomes in Escherichia Coli. *PNAS* **2013**, *110*, 1–6, doi:10.1073/pnas.1304080110.

143. Jain, P.; Nagaraja, V. An Atypical Type II Topoisomerase from Mycobacterium Smegmatis with Positive Supercoiling Activity. *Mol. Microbiol.* **2005**, *58*, 1392–1405, doi:10.1111/j.1365-2958.2005.04908.x.

144. Rudolph, M.G.; Del Toro Duany, Y.; Jungblut, S.P.; Ganguly, A.; Klostermeier, D. Crystal Structures of Thermotoga Maritima Reverse Gyrase: Inferences for the Mechanism of Positive DNA Supercoiling. *Nucleic Acids Res.* **2013**, *41*, 1058–1070, doi:10.1093/nar/gks1073.

145. Panas, M.W.; Jain, P.; Yang, H.; Mitra, S.; Biswas, D.; Wattam, A.R.; Letvin, N.L.; Jacobs, W.R.

Noncanonical SMC Protein in Mycobacterium Smegmatis Restricts Maintenance of Mycobacterium Fortuitum Plasmids. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 13264–13271, doi:10.1073/pnas.1414207111.

146. Doron, S.; Melamed, S.; Ofir, G.; Leavitt, A.; Lopatina, A.; Keren, M.; Amitai, G.; Sorek, R. Systematic Discovery of Antiphage Defense Systems in the Microbial Pangenome. *Science (80-.).* **2018**, *359*, 0–12, doi:10.1126/science.aar4120.

147. Tadesse, S.; Mascarenhas, J.; Kösters, B.; Hasilik, A.; Graumann, P.L. Genetic Interaction of the SMC Complex with Topoisomerase IV in Bacillus Subtilis. *Microbiology* **2005**, *151*, 3729–3737, doi:10.1099/mic.0.28234-0.

148. Bergerat, A.; De Massy, B.; Gadelle, D.; Varoutas, P.C.; Nicolas, A.; Forterre, P. An Atypical Topoisomerase II from Archaea with Implications for Meiotic Recombination. *Nature* **1997**, *386*, 414–417, doi:10.1038/386414a0.

149. Wendorff, T.J.; Berger, J.M. Topoisomerase VI Senses and Exploits Both DNA Crossings and Bends to Facilitate Strand Passage. *Elife* **2018**, *7*, 1–35, doi:10.7554/eLife.31724.

150. McKie, S.J.; Desai, P.R.; Seol, Y.; Allen, A.M.B.; Maxwell, A.; Neuman, K.C. Topoisomerase VI Is a Chirally-Selective, Preferential DNA Decatenase. *Elife* **2022**, 1–30.

151. Visone, V.; Vettone, A.; Serpe, M.; Valenti, A.; Perugino, G.; Rossi, M.; Ciaramella, M. Chromatin Structure and Dynamics in Hot Environments: Architectural Proteins and DNA Topoisomerases of Thermophilic Archaea. *Int. J. Mol. Sci.* **2014**, *15*, 17162–17187, doi:10.3390/ijms150917162.

152. Lopez-Garcia, P.; Forterre, P. Control of DNA Topology during Thermal Stress in Hyperthermophilic Archaea : DNA Topoisomerase Levels , Activities and Induced Thermotolerance during Heat and Cold Shock in Sulfolobus. *Mol. Microbiol.* **1999**, *33*, 766–777.

153. Gadelle, D.; Krupovic, M.; Raymann, K.; Mayer, C.; Forterre, P. DNA Topoisomerase VIII : A Novel Subfamily of Type IIB Topoisomerases Encoded by Free or Integrated Plasmids in Archaea and Bacteria. *Nucleic Acids Res.* **2014**, *42*, 8578–8591, doi:10.1093/nar/gku568.

154. Takahashi, T.S.; Cunha, V. Da; Krupovic, M.; Mayer, C.; Forterre, P.; Gadelle, D. Expanding the Type IIB DNA Topoisomerase Family : Identification of New Topoisomerase and Topoisomerase-like Proteins in Mobile Genetic Elements. *Nucleic Acids Res.* **2020**, *2*, 1–13, doi:10.1093/nargab/lqz021.

155. Liu, Y.; Berrido, A.M.; Hua, Z.C.; Tse-Dinh, Y.C.; Leng, F. Biochemical and Biophysical Properties of Positively Supercoiled DNA. *Biophys. Chem.* **2017**, *230*, 68–73, doi:10.1016/j.bpc.2017.08.008.

156. Menzel, R.; Gellert, M. Regulation of the Genes for E . Coli DNA Gyrase: Homeostatic Control of DNA Supercoiling. *Cell* **1983**, *34*, 105–113.

157. Tse-Dinh, Y.-C. Regulation of the Escherichia Coli DNA Topoisomerase I by DNA Supercoiling. *Nucleic Acids Res.* **1985**, *13*, 4751–4763.

158. Pruss, G.J.; Manes, S.H.; Drlica, K. Escherichia Coli DNA Topoisomerase I Mutants: Increased Supercoiling Is Corrected by Mutations near Gyrase Genes. *Cell* **1982**, *31*, 35–42.

159. DiNardo, S.; Voelkel, K.A.; Sternglanz, R. Escherichia Coli DNA Topoisomerase I Mutants Have Compensatory Mutations in DNA Gyrase Genes. *Cell* **1982**, *31*, 43–51.

160. Dorman, C.J.; Lynch, A.S.; Bhriain, N.N.; Higgins, C.F. DNA Supercoiling in Escherichia Coli: TopA Mutations Can Be Suppressed by DNA Amplifications Involving the TolC Locus. *Mol. Microbiol.* **1989**, *3*, 531–540, doi:10.1111/j.1365-2958.1989.tb00199.x.

161. Kato, J.; Nishimura, Y.; Imamura, R.; Niki, H.; Hiraga, S.; Suzuki, H. New Topoisomerase Essential for Chromosome Segregation in E. Coli. *Cell* **1990**, *63*, 393–404, doi:10.1016/0092-8674(90)90172-B.

162. Usongo, V.; Tanguay, C.; Nolent, F.; Bessong, J.E.; Drolet, M. Interplay between Type 1A Topoisomerases and Gyrase in Chromosome Segregation in Escherichia Coli. *J. Bacteriol.* **2013**, *195*, 1758–1768, doi:10.1128/JB.02001-12.

163. Baaklini, I.; Usongo, V.; Nolent, F.; Sanscartier, P.; Hraiky, C.; Drlica, K.; Drolet, M. Hypernegative Supercoiling Inhibits Growth by Causing RNA Degradation. *J. Bacteriol.* **2008**, *190*, 7346–7356, doi:10.1128/JB.00680-08.

164. Masse, E.; Drolet, M. Escherichia Coli DNA Topoisomerase I Inhibits R-Loop Formation by Relaxing Transcription-Induced Negative Supercoiling. *J. Biochem.* **1999**, *274*, 16659–16664.

165. Stolz, R.; Sulthana, S.; Hartono, S.R.; Malig, M.; Benham, C.J.; Chedin, F. Interplay between DNA Sequence and Negative Superhelicity Drives R-Loop Structures. *Proc Natl Acad Sci USA* **2019**, *116*, 6260–6269, doi:10.1073/pnas.1819476116.

166. Brochu, J.; Vlachos-Breton, E.; Drolet, M. Bacterial Type 1A Topoisomerases Maintain the Stability of the Genome by Preventing and Dealing with R-Loop-and Nucleotide Excision Repair-Dependent Topological Stress. *Biorxiv Prepr.* **2021**, doi:https://doi.org/10.1101/2021.07.10.451908.

167. Drolet, M.; Bi, X.; Liu, L.F. Hypernegative Supercoiling of the DNA Template during Transcription Elongation in Vitro. *J. Biol. Chem.* **1994**, *269*, 2068–2074, doi:10.1016/S0021-9258(17)42136-3.

168. Drolet, M. Growth Inhibition Mediated by Excess Negative Supercoiling: The Interplay between Transcription Elongation, R-Loop Formation and DNA Topology. *Mol. Microbiol.* **2006**, *59*, 723–730, doi:10.1111/j.1365-2958.2005.05006.x.

169. Chedin, F.; Benham, C.J. Emerging Roles for R-Loop Structures in the Management of Topological Stress. *J. Biol. Chem.* **2020**, *295*, 4684–4695, doi:10.1074/jbc.REV119.006364.

170. Kogoma, T. Stable DNA Replication: Interplay between DNA Replication, Homologous Recombination, and Transcription. *Microbiol. Mol. Biol. Rev.* **1997**, *61*, 212–238, doi:10.1128/.61.2.212-238.1997.

171. Drolet, M.; Brochu, J. R-Loop-Dependent Replication and Genomic Instability in Bacteria. *DNA Repair (Amst).* **2019**, *84*, 102693, doi:10.1016/j.dnarep.2019.102693.

172. Drolet, M.; Phoenix, P.; Menzelt, R.; Masse, E.; Liu, L.F.; Crouchs, R.J. Overexpression of RNase H Partially Complements the Growth Defect of an Escherichia Coli Delta TopA Mutant : R-Loop Formation Is a Major Problem in the Absence of DNA Topoisomerase I. *Proc Natl Acad Sci USA* **1995**, *92*, 3526–3530.

173. Hraiky, C.; Raymond, M.A.; Drolet, M. RNase H Overproduction Corrects a Defect at the Level of Transcription Elongation during RRNA Synthesis in the Absence of DNA Topoisomerase I in Escherichia Coli. *J. Biol. Chem.* **2000**, *275*, 11257–11263, doi:10.1074/jbc.275.15.11257.

174. Stockum, A.; Lloyd, R.G.; Rudolph, C.J. On the Viability of Escherichia Coli Cells Lacking DNA Topoisomerase I. *BMC Microbiol.* **2012**, *12*, 26, doi:10.1186/1471-2180-12-26.

175. Usongo, V.; Nolent, F.; Sanscartier, P.; Tanguay, C.; Broccoli, S.; Baaklini, I.; Drlica, K.; Drolet, M. Depletion of RNase HI Activity in Escherichia Coli Lacking DNA Topoisomerase I Leads to Defects in DNA Supercoiling and Segregation. *Mol. Microbiol.* **2008**, *69*, 968–981, doi:10.1111/j.1365-2958.2008.06334.x.

176. Martel, M.; Balleydier, A.; Sauriol, A.; Drolet, M. Constitutive Stable DNA Replication in Escherichia Coli Cells Lacking Type 1A Topoisomerase Activity. *DNA Repair (Amst).* **2015**, *35*, 37–47, doi:10.1016/j.dnarep.2015.08.004.

177. Baaklini, I.; Hraiky, C.; Rallu, F.; Tse-Dinh, Y.C.; Drolet, M. RNase HI Overproduction Is Required for Efficient Full-Length RNA Synthesis in the Absence of Topoisomerase I in Escherichia Coli. *Mol. Microbiol.* **2004**, *54*, 198–211, doi:10.1111/j.1365-2958.2004.04258.x.

178. Leela, J.K.; Raghunathan, N.; Gowrishankar, J. Topoisomerase I Essentiality, DnaA-Independent Chromosomal Replication, and Transcription-Replication Conflict in Escherichia Coli. *J. Bacteriol.* **2021**, *203*, 1–13.

179. Cheng, B.; Zhu, C.; Ji, C.; Ahumada, A.; Tse-Dinh, Y.-C. Direct Interaction between Escherichia Coli RNA Polymerase and the Zinc Ribbon Domains of DNA Topoisomerase I. *J. Biol. Chem.* **2003**, *278*, 30705–30710, doi:10.1074/jbc.M303403200.

180. Banda, S.; Cao, N.; Tse-Dinh, Y.-C. Distinct Mechanism Evolved for Mycobacterial RNA Polymerase and Topoisomerase I Protein-Protein Interaction. *J. Mol. Biol.* **2017**, *429*, 2931–2942, doi:10.1016/j.jmb.2017.08.011.Distinct.

181. Ferrandiz, M.-J.; Hernandez, P.; de la Campa, A.G. Genome-Wide Proximity between RNA Polymerase and DNA Topoisomerase I Supports Transcription in Streptococcus Pneumoniae. *PLoS Genet.* **2021**, *4*, 1–21, doi:10.1371/journal.pgen.1009542.

182. Yang, J.; Annamalai, T.; Cheng, B.; Banda, S.; Tyagi, R.; Tse-Dinh, Y.-C. Antimicrobial Susceptibility and SOS-Dependent Increase in Mutation Frequency Are Impacted by Escherichia Coli Topoisomerase I. *Antimicrob. Agents Chemother.* **2015**, *59*, 6195–6202, doi:10.1128/AAC.00855-15.

183. Kirkegaard, K.; Pflugfelder, G.; Wang, J.C. The Cleavage of DNA by Type-I DNA Topoisomerases. *Cold Spring Harb Symp Quant Biol* **1984**, *49*, 411–419.

184. Tse, Y.; Kirkegaard, K.; Wang, J.C. Covalent Bonds between Protein and DNA: Formation of Phosphotyrosine Linkage between Certain DNA Topoisomerases and DNA. *J. Biol. Chem.* **1980**, *255*, 5560–5565.

185. Liu, L.F.; Wang, J.C. Supercoiling of the DNA Template during Transcription. *Proc Natl Acad Sci USA* **1987**, *84*, 7024–7027, doi:10.1073/pnas.84.20.7024.

186. Terekhova, K.; Gunn, K.H.; Marko, J.F.; Mondragón, A. Bacterial Topoisomerase I and Topoisomerase III Relax Supercoiled DNA via Distinct Pathways. *Nucleic Acids Res.* **2012**, *40*, 10432–10440, doi:10.1093/nar/gks780.

187. Terekhova, K.; Marko, J.F.; Mondragón, A. Studies of Bacterial Topoisomerases I and III at the Single Molecule Level. *Biochem. Soc. Trans.* **2014**, *41*, 571–575, doi:10.1042/BST20120297.Studies.

188. Terekhova, K.; Marko, J.F.; Mondragon, A. Single-Molecule Analysis Uncovers the Difference between the Kinetics of DNA Decatenation by Bacterial Topoisomerases I and III. *Nucleic Acids Res.* **2014**, *42*, 11657–11667, doi:10.1093/nar/gku785.

189. Nurse, P.; Levine, C.; Hassing, H.; Marians, K.J. Topoisomerase III Can Serve as the Cellular Decatenase in Escherichia Coli. *J. Biol. Chem.* **2003**, *278*, 8653–8660, doi:10.1074/jbc.M211211200.

190. Suski, C.; Marians, K.J. Resolution of Converging Replication Forks by RecQ and Topoisomerase III. *Mol. Cell* **2009**, *30*, 779–789.

191. Harmon, F.G.; Brockman, J.P.; Kowalczykowski, S.C. RecQ Helicase Stimulates Both DNA Catenation and Changes in DNA Topology by Topoisomerase III. *J. Biol. Chem.* **2003**, *278*, 42668–42678, doi:10.1074/jbc.M302994200.

192. Bizard, A.H.; Hickson, I.D. The Dissolution of Double Holliday Junctions. *Cold Spring Harb. Perspect. Biol.* **2014**, *6*, 1–14.

193. Lopez, C.R.; Yang, S.; Deibler, R.W.; Ray, S.A.; Pennington, J.M.; Digate, R.J.; Hastings, P.J.; Rosenberg, S.M.; Zechiedrich, E.L. A Role for Topoisomerase III in a Recombination Pathway Alternative to RuvABC. *Mol. Microbiol.* **2005**, *58*, 80–101, doi:10.1111/j.1365-2958.2005.04812.x.

194. Usongo, V.; Drolet, M. Roles of Type 1A Topoisomerases in Genome Maintenance in Escherichia Coli. *PLoS Genet.* **2014**, *10*, 1–18, doi:10.1371/journal.pgen.1004543.

195. Lee, C.M.; Wang, G.; Pertsinidis, A.; Marians, K.J. Topoisomerase III Acts at the Replication Fork to Remove Precatenanes. *J. Bacteriol.* **2019**, *201*, 1–13, doi:10.1128/JB.00563-18.

196. Perez-Cheeks, B.A.; Lee, C.; Hayama, R.; Marians, K.J. A Role for Topoisomerase III in Escherichia Coli Chromosome Segregation. *Mol. Microbiol.* **2012**, *86*, 1007–1022, doi:10.1111/mmi.12039.A.

197. Wang, H.; Di Gate, R.J.; Seeman, N.C. An RNA Topoisomerase. *Proc Natl Acad Sci USA* **1996**, *93*, 9477–9482.

198. Ahmad, M.; Xue, Y.; Lee, S.K.; Martindale, J.L.; Shen, W.; Li, W.; Zou, S.; Ciaramella, M.; Debat, H.; Nadal, M.; et al. RNA Topoisomerase Is Prevalent in All Domains of Life and Associates with Polyribosomes in Animals. *Nucleic Acids Res.* **2016**, *44*, 6335–6349, doi:10.1093/nar/gkw508.

199. Allali, N.; Afif, H.; Couturier, M.; Van Melderen, L. The Highly Conserved TldD and TldE Proteins of Escherichia Coli Are Involved in Microcin B17 Processing and in CcdA Degradation. *J. Bacteriol.* **2002**, *184*, 3224–3231, doi:10.1128/JB.184.12.3224-3231.2002.

200. Ghilarov, D.; Serebryakova, M.; Stevenson, C.E.M.; Hearnshaw, S.; Volkov, D.; Maxwell, A.; Lawson, D.M.; Severinov, K. The Origins of Specificity in the Microcin-Processing Protease TldD/E. *Struct. Des.* **2017**, *25*, 1549-1561.e5, doi:10.1016/j.str.2017.08.006.

201. Scheirer, K.E.; Higgins, N.P. The DNA Cleavage Reaction of DNA Gyrase. *J. Biol. Chem.* **1997**, *272*, 27202–27209.

202. Datsenko, K.A.; Wanner, B.L. One-Step Inactivation of Chromosomal Genes in Escherichia Coli K-12 Using PCR Products. *Proc Natl Acad Sci USA* **2000**, *97*, 6640–6645.

203. Kitagawa, M.; Ara, T.; Arifuzzaman, M.; Ioka-Nakamichi, T.; Inamoto, E.; Toyonaga, H.; Mori, H. Complete Set of ORF Clones of Escherichia Coli ASKA Library (A Complete Set of E. Coli K-12

ORF Archive): Unique Resources for Biological Research. *DNA Res.* **2005**, *12*, 291–299, doi:10.1093/dnares/dsi012.

204. Roy, R.S.; Kelleher, N.L.; Milne, J.C.; Walsh, C.T. In Vivo Processing and Antibiotic Activity of Microcin B17 Analogs with Varying Ring Content and Altered Bisheterocyclic Sites. *Chem. Biol.* **1999**, *6*, 305–318, doi:10.1016/S1074-5521(99)80076-3.

205. Yu, D.; Ellis, H.M.; Lee, E.-C.; Jenkins, N.A.; Copeland, N.G.; Court, D.L. An Efficient Recombination System for Chromosome Engineering in Escherichia Coli. *Proc. Natl. Acad. Sci.* **2000**, *97*, 5978–5983, doi:10.1073/pnas.100127597.

206. Li, H.; Durbin, R. Fast and Accurate Long-Read Alignment with Burrows-Wheeler Transform. *Bioinformatics* **2010**, *26*, 589–595, doi:10.1093/bioinformatics/btp698.

207. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The Sequence Alignment/Map Format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079, doi:10.1093/bioinformatics/btp352.

208. Thorvaldsdóttir, H.; Robinson, J.T.; Mesirov, J.P. Integrative Genomics Viewer (IGV): High-Performance Genomics Data Visualization and Exploration. *Brief. Bioinform.* **2013**, *14*, 178–192, doi:10.1093/bib/bbs017.

209. Audic, S.; Claverie, J.M. The Significance of Digital Gene Expression Profiles. *Genome Res.* **1997**, *7*, 986–995, doi:10.1101/gr.7.10.986.

210. Reid, P.; Speyer, J. Rifampicin Inhibition of Ribonucleic Acid and Protein Synthesis in Normal and Ethylenediaminetetraacetic Acid-Treated Escherichia Coli. *J. Bacteriol.* **1970**, *104*, 376–389.

211. Cutler, R.G.; Evans, J.E. Synchronization of Bacteria by a Stationary-Phase Method. *J. Bacteriol.* **1966**, *91*, 469–476.

212. Hunter, J.D. MATPLOTLIB : A 2D GRAPHICS ENVIRONMENT. *Sci. Program.* **2007**, *May-June*, 90–95, doi:10.1109/MCSE.2007.55.

213. Crooks, G.; Hon, G.; Chandonia, J.; Brenner, S. WebLogo: A Sequence Logo Generator. *Genome Res* **2004**, *14*, 1188–1190, doi:10.1101/gr.849004.1.

214. Cock, P.J.A.; Antao, T.; Chang, J.T.; Chapman, B.A.; Cox, C.J.; Dalke, A.; Friedberg, I.; Hamelryck, T.; Kauff, F.; Wilczynski, B.; et al. Biopython: Freely Available Python Tools for Computational Molecular Biology and Bioinformatics. *Bioinformatics* **2009**, *25*, 1422–1423, doi:10.1093/bioinformatics/btp163.

215. Dijk, M. Van; Bonvin, A.M.J.J. 3D-DART : A DNA Structure Modelling Server. **2009**, *37*, 235–239, doi:10.1093/nar/gkp287.

216. Ruthenburg, A.J.; Graybosch, D.M.; Huetsch, J.C.; Verdine, G.L. A Superhelical Spiral in the Escherichia Coli DNA Gyrase A C-Terminal Domain Imparts Unidirectional Supercoiling Bias. *J. Biol. Chem.* **2005**, *280*, 26177–26184, doi:10.1074/jbc.M502838200.

217. Bagel, S.; Hullen, V.; Wiedemann, B.; Heisig, P. Impact of GyrA and ParC Mutations on Quinolone Resistance, Doubling Time, and Supercoiling Degree of Escherichia Coli. *Antimicrob. Agents Chemother.* **1999**, *43*, 868–875.

218. Sawitzke, J.A.; Costantino, N.; Li, X.; Thomason, L.C.; Bubunenko, M.; Court, C.; Court, D.L. Probing Cellular Processes with Oligo-Mediated Recombination and Using the Knowledge Gained

to Optimize Recombineering. *J. Mol. Biol.* **2011**, *407*, 45–59, doi:10.1016/j.jmb.2011.01.030.

219. Sutormin, D.; Rubanova, N.; Logacheva, M.; Ghilarov, D.; Severinov, K. Single-Nucleotide-Resolution Mapping of DNA Gyrase Cleavage Sites across the Escherichia Coli Genome. *Nucleic Acids Res.* **2019**, *47*, 1–16, doi:10.1093/nar/gky1222.

220. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* **2014**, *30*, 2114–2120, doi:10.1093/bioinformatics/btu170.

221. Narula, G.; Tse-Dinh, Y.C. Residues of E. Coli Topoisomerase i Conserved for Interaction with a Specific Cytosine Base to Facilitate DNA Cleavage. *Nucleic Acids Res.* **2012**, *40*, 9233–9243, doi:10.1093/nar/gks688.

222. Wang, L.; Wang, S.; Li, W. RSeQC: Quality Control of RNA-Seq Experiments. *Bioinformatics* **2012**, *28*, 2184–2185, doi:10.1093/bioinformatics/bts356.

223. Fàbrega, A.; Madurga, S.; Giralt, E.; Vila, J. Mechanism of Action of and Resistance to Quinolones. *Microb. Biotechnol.* **2009**, *2*, 40–61, doi:10.1111/j.1751-7915.2008.00063.x.

224. Collin, F.; Karkare, S.; Maxwell, A. Exploiting Bacterial DNA Gyrase as a Drug Target: Current State and Perspectives. *Appl. Microbiol. Biotechnol.* **2011**, *92*, 479–497, doi:10.1007/s00253-011-3557-z.

225. Hooper, D.C.; Jacoby, G.A. Topoisomerase Inhibitors : Fluoroquinolone Mechanisms of Action and Resistance. *Cold Spring Harb. Perspect. Med.* **2016**, 1–22.

226. Heddle, J.G.; Blance, S.J.; Zamble, D.B.; Hollfelder, F.; Miller, D.A.; Wentzell, L.M.; Walsh, C.T.; Maxwell, A. The Antibiotic Microcin B17 Is a DNA Gyrase Poison: Characterisation of the Mode of Inhibition. *J. Mol. Biol.* **2001**, *307*, 1223–34, doi:10.1006/jmbi.2001.4562.

227. Wahle, E.; Kornberg, A. The Partition Locus of Plasmid PSC101 Is a Specific Binding Site for DNA Gyrase. *EMBO J.* **1988**, *7*, 1889–1895.

228. Pato, M.L.; Howe, M.M.; Higgins, N.P. A DNA Gyrase-Binding Site at the Center of the Bacteriophage Mu Genome Is Required for Efficient Replicative Transposition. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*, 8716–8720, doi:10.1073/pnas.87.22.8716.

229. Oram, M.; Howells, A.J.; Maxwell, A.; Pato, M.L. A Biochemical Analysis of the Interaction of DNA Gyrase with the Bacteriophage Mu, PSC101 and PBR322 Strong Gyrase Sites: The Role of DNA Sequence in Modulating Gyrase Supercoiling and Biological Activity. *Mol. Microbiol.* **2003**, *50*, 333–347, doi:10.1046/j.1365-2958.2003.03690.x.

230. Franco, R.J.; Drlica, K. DNA Gyrase on the Bacterial Chromosome. Oxolinic Acid-Induced DNA Cleavage in the DnaA-GyrB Region. *J. Mol. Biol.* **1988**, *201*, 229–233, doi:10.1016/0022-2836(88)90449-4.

231. Lockshon, D.; Morris, D.R. Sites of Reaction of Escherichia Coli DNA Gyrase on PBR322 in Vivo as Revealed by Oxolinic Acid-Induced Plasmid Linearization. *J. Mol. Biol.* **1985**, *181*, 63–74, doi:10.1016/0022-2836(85)90324-9.

232. O'Connor, M.B.; Malamy, M.H. Mapping of DNA Gyrase Cleavage Sites in Vivo Oxolinic Acid Induced Cleavages in Plasmid PBR322. *J. Mol. Biol.* **1985**, *181*, 545–550, doi:10.1016/0022-2836(85)90426-7.

233. Yang, Y.; Ames, G.F. DNA Gyrase Binds to the Family of Prokaryotic Repetitive Extragenic

Palindromic Sequences. *Proc. Natl. Acad. Sci. USA* **1988**, 8850–8854.

234. Espeli, O.; Boccard, F. In Vivo Cleavage of Escherichia Coli BIME-2 Repeats by DNA Gyrase: Genetic Characterization of the Target and Identification of the Cut Site. *Mol. Microbiol.* **1997**, *26*, 767–777.

235. Yu, X.; Davenport, J.W.; Urtishak, K.A.; Carillo, M.L.; Gosai, S.J.; Kolaris, C.P.; Byl, J.A.W.; Rappaport, E.F.; Osheroff, N.; Gregory, B.D.; et al. Genome-Wide TOP2A DNA Cleavage Is Biased towards Translocated and Highly Transcribed Loci. *Genome Res.* **2017**, doi:10.1101/gr.211615.116.

236. Tin, M.M.Y.; Economo, E.P.; Mikheyev, A.S. Sequencing Degraded DNA from Non-Destructively Sampled Museum Specimens for RAD-Tagging and Low-Coverage Shotgun Phylogenetics. *PLoS One* **2014**, *9*, doi:10.1371/journal.pone.0096793.

237. Pato, M.L.; Banerjee, M. Genetic Analysis of the Strong Gyrase Site (SGS) of Bacteriophage Mu: Localization of Determinants Required for Promoting Mu Replication. *Mol. Microbiol.* **2000**, *37*, 800–810, doi:10.1046/j.1365-2958.2000.02042.x.

238. Drlica, K.; Malik, M.; Kerns, R.J.; Zhao, X. Quinolone-Mediated Bacterial Death. *Antimicrob. Agents Chemother.* **2008**, *52*, 385–392, doi:10.1128/AAC.01617-06.

239. Orphanides, G.; Maxwell, A.; Le, L. Evidence for a Conformational Change in the DNA Gyrase - DNA Complex from Hydroxyl Radical Footprinting. *Nucleic Acids Res.* **1994**, *22*, 1567–1575.

240. Satchwell, S.C.; Drew, H.R.; Travers, A.A. Sequence Periodicities in Chicken Nucleosome Core DNA. *J. Mol. Biol.* **1986**, *191*, 659–675, doi:10.1016/0022-2836(86)90452-3.

241. Chiu, T.P.; Yang, L.; Zhou, T.; Main, B.J.; Parker, S.C.J.; Nuzhdin, S. V.; Tullius, T.D.; Rohs, R. GBshape: A Genome Browser Database for DNA Shape Annotations. *Nucleic Acids Res.* **2015**, *43*, D103–D109, doi:10.1093/nar/gku977.

242. Koudelka, G.B.; Mauro, S.A.; Ciubotaru, M. Indirect Readout of DNA Sequence by Proteins : The Roles of DNA Sequence - Dependent Intrinsic and Extrinsic Forces I . Introduction. *Prog. Nucleic Acid Res. Mol. Biol.* **2006**, *81*, 143–177, doi:10.1016/S0079-6603(06)81004-4.

243. Richter, S.N.; Giaretta, G.; Comuzzi, V.; Leo, E.; Mitchenall, L.A.; Fisher, L.M.; Maxwell, A.; Palumbo, M. Hot-Spot Consensus of Fluoroquinolone-Mediated DNA Cleavage by Gram-Negative and Gram-Positive Type II DNA Topoisomerases. *Nucleic Acids Res.* **2007**, *35*, 6075–6085, doi:10.1093/nar/gkm653.

244. Leo, E.; Gould, K.A.; Pan, X.S.; Capranico, G.; Sanderson, M.R.; Palumbo, M.; Fisher, L.M. Novel Symmetric and Asymmetric DNA Scission Determinants for Streptococcus Pneumoniae Topoisomerase IV and Gyrase Are Clustered at the DNA Breakage Site. *J. Biol. Chem.* **2005**, *280*, 14252–14263, doi:10.1074/jbc.M500156200.

245. Pierrat, O.A.; Maxwell, A. Evidence for the Role of DNA Strand Passage in the Mechanism of Action of Microcin B17 on DNA Gyrase. *Biochemistry* **2005**, *44*, 4204–4215, doi:10.1021/bi0478751.

246. Gupta, S.; Stamatoyannopoulos, J.A.; Bailey, T.L.; Noble, W.S. Quantifying Similarity between Motifs. *Genome Biol.* **2007**, *8*, doi:10.1186/gb-2007-8-2-r24.

247. Gourse, R.L.; Gaal, T.; Bartlett, M.S.; Appleman, J.A.; Ross, W. RRNA TRANSCRIPTION AND

GROWTH RATE–DEPENDENT REGULATION OF RIBOSOME SYNTHESIS IN *ESCHERICHIA COLI*. *Annu. Rev. Microbiol.* **1996**, *50*, 645–677, doi:10.1146/annurev.micro.50.1.645.

248. Prieto, A.I.; Kahramanoglou, C.; Ali, R.M.; Fraser, G.M.; Seshasayee, A.S.N.; Luscombe, N.M. Genomic Analysis of DNA Binding and Gene Regulation by Homologous Nucleoid-Associated Proteins IHF and HU in Escherichia Coli K12. *Nucleic Acids Res.* **2012**, *40*, 3524–3537, doi:10.1093/nar/gkr1236.

249. Kahramanoglou, C.; Seshasayee, A.S.N.; Prieto, A.I.; Ibberson, D.; Schmidt, S.; Zimmermann, J.; Benes, V.; Fraser, G.M.; Luscombe, N.M. Direct and Indirect Effects of H-NS and Fis on Global Gene Expression Control in Escherichia Coli. *Nucleic Acids Res.* **2011**, *39*, 2073–2091, doi:10.1093/nar/gkq934.

250. Zawadzki, P.; Stracy, M.; Ginda, K.; Zawadzka, K.; Lesterlin, C.; Kapanidis, A.N.; Sherratt, D.J. The Localization and Action of Topoisomerase IV in Escherichia Coli Chromosome Segregation Is Coordinated by the SMC Complex , MukBEF Article The Localization and Action of Topoisomerase IV in Escherichia Coli Chromosome Segregation Is Coordinated by The. *CellReports* **2016**, *13*, 2587–2596, doi:10.1016/j.celrep.2015.11.034.

251. Deng, S.; Stein, R. a.; Higgins, N.P. Organization of Supercoil Domains and Their Reorganization by Transcription. *Mol. Microbiol.* **2005**, *57*, 1511–1521, doi:10.1111/j.1365-2958.2005.04796.x.

252. Foster, P.L.; Lee, H.; Popodi, E.; Townes, J.P.; Tang, H. Determinants of Spontaneous Mutation in the Bacterium Escherichia Coli as Revealed by Whole-Genome Sequencing. *PNAS* **2015**, *3*, E5990–E5999, doi:10.1073/pnas.1512136112.

253. Krzywinski, M.; Schein, J.; Birol, I.; Connors, J.; Gascoyne, R.; Horsman, D.; Jones, S.J.; Marra, M.A. Circos: An Information Aesthetic for Comparative Genomics. *Genome Res.* **2009**, *19*, 1639–1645, doi:10.1101/gr.092759.109.

254. Peter, B.J.; Ullsperger, C.; Hiasa, H.; Marians, K.J.; Cozzarelli, N.R. The Structure of Supercoiled Intermediates in DNA Replication. *Cell* **1998**, *94*, 819–827, doi:10.1016/S0092-8674(00)81740-7.

255. Pham, T.M.; Tan, K.W.; Sakumura, Y.; Okumura, K.; Maki, H.; Akiyama, M.T. A Single-Molecule Approach to DNA Replication in Escherichia Coli Cells Demonstrated That DNA Polymerase III Is a Major Determinant of Fork Speed. *Mol. Micro* **2013**, *90*, 584–596, doi:10.1111/mmi.12386.

256. Stracy, M.; Wollman, A.J.M.; Kaja, E.; Gapinski, J.; Lee, J.E.; Leek, V.A.; McKie, S.J.; Mitchenall, L.A.; Maxwell, A.; Sherratt, D.J.; et al. Single-Molecule Imaging of DNA Gyrase Activity in Living Escherichia Coli. *Nucleic Acids Res.* **2019**, *47*, 210–220, doi:10.1093/nar/gky1143.

257. Aleixandre, V.; Urios, A.; Herrera, G.; Blanco, M. New Escherichia Coli GyrA and GyrB Mutations Which Have a Graded Effect on DNA Supercoiling. *Mol Gen Genet* **1989**, *219*, 306–312.

258. Marcusson, L.L.; Frimodt-Møller, N.; Hughes, D. Interplay in the Selection of Fluoroquinolone Resistance and Bacterial Fitness. *PLoS Pathog.* **2009**, *5*, doi:10.1371/journal.ppat.1000541.

259. Zawadzki, P.; Stracy, M.; Ginda, K.; Zawadzka, K.; Lesterlin, C.; Kapanidis, A.N.; Sherratt, D.J.; Nolivos, S.; Upton, A.L.; Badrinarayanan, A.; et al. The SMC Complex MukBEF Recruits Topoisomerase IV to the Origin of Replication Region in Live Escherichia Coli. *Proc. Natl. Acad.*

Sci. U. S. A. **2010**, *107*, 16921–16932, doi:10.1128/mBio.01001-13.

260. Syeda, A.H.; Dimude, J.U.; Skovgaard, O.; Rudolph, C.J. Too Much of a Good Thing : How Ectopic DNA Replication Affects Bacterial Replication Dynamics. *Front. Microbiol.* **2020**, *11*, 1–21, doi:10.3389/fmicb.2020.00534.

261. Gittens, W.H.; Johnson, D.J.; Allison, R.M.; Cooper, T.J.; Thomas, H.; Neale, M.J. A Nucleotide Resolution Map of Top2-Linked DNA Breaks in the Yeast and Human Genome. *Nat. Commun.* **2019**, *10*, 1–16, doi:10.1038/s41467-019-12802-5.

262. Arnoldi, E.; Pan, X.; Fisher, M. Functional Determinants of Gate-DNA Selection and Cleavage by Bacterial Type II Topoisomerases. *Nucleic Acids Res.* **2013**, *41*, 9411–9423, doi:10.1093/nar/gkt696.

263. Zhang, Y.; Liu, T.; Meyer, C.A.; Eeckhoute, J.; Johnson, D.S.; Bernstein, B.E.; Nussbaum, C.; Myers, R.M.; Brown, M.; Li, W.; et al. Model-Based Analysis of ChIP-Seq (MACS). *Genome Biol.* **2008**, *9*, R137.1-R137.9, doi:10.1186/gb-2008-9-9-r137.

264. De Smit, M.H.; Verlaan, P.W.G.; Van Duin, J.; Pleij, C.W.A. Intracistronic Transcriptional Polarity Enhances Translational Repression: A New Role for Rho. *Mol. Microbiol.* **2008**, *69*, 1278–1289, doi:10.1111/j.1365-2958.2008.06360.x.

265. Zhu, M.; Mori, M.; Hwa, T.; Dai, X. Disruption of Transcription–Translation Coordination in Escherichia Coli Leads to Premature Transcriptional Termination. *Nat. Microbiol.* **2019**, *4*, 2347–2356, doi:10.1038/s41564-019-0543-1.

266. Kouzine, F.; Gupta, A.; Baranello, L.; Wojtowicz, D.; Ben-Aissa, K.; Liu, J.; Przytycka, T.M.; Levens, D. Transcription-Dependent Dynamic Supercoiling Is a Short-Range Genomic Force. *Nat. Struct. Mol. Biol.* **2013**, doi:10.1038/nsmb.2517.

267. Kouzine, F.; Sanford, S.; Elisha-feil, Z.; Levens, D. The Functional Response of Upstream DNA to Dynamic Supercoiling in Vivo. *Nat. Struct. Mol. Biol.* **2008**, *15*, 146–154, doi:10.1038/nsmb.1372.

268. Tse, Y.; Wang, J. E. Coli and M. Luteus DNA Topoisomerase I Can Catalyze Catenation or Decatenation of Double-Stranded DNA Rings. *Cell* **1980**, *22*, 269–276.

269. Bhaduri, T.; Bagui, T.K.; Sikder, D.; Nagaraja, V. DNA Topoisomerase I from Mycobacterium Smegmatis. An Enzyme with Distinct Features. *J. Biol. Chem.* **1998**, *273*, 13925–13932, doi:10.1074/jbc.273.22.13925.

270. Mooney, R.A.; Davis, S.E.; Peters, J.M.; Rowland, J.L.; Ansari, A.Z.; Landick, R. Regulator Trafficking on Bacterial Transcription Units in Vivo. *Mol. Cell* **2009**, *33*, 97–108, doi:10.1016/j.molcel.2008.12.021.

271. Campbell, E.A.; Korzheva, N.; Mustaev, A.; Murakami, K.; Nair, S.; Goldfarb, A.; Darst, S.A. Structural Mechanism for Rifampicin Inhibition of Bacterial RNA Polymerase. *Cell* **2001**, *104*, 901–912, doi:10.1016/S0092-8674(01)00286-0.

272. Woldringh, C.L. The Role of Co-Transcriptional Translation and Protein Translocation (Transertion) in Bacterial Chromosome Segregation. *Mol. Microbiol.* **2002**, *45*, 17–29, doi:10.1046/j.1365-2958.2002.02993.x.

273. Visser, B.J.; Sharma, S.; Chen, P.J.; Mcmullin, A.B.; Bates, M.L.; Bates, D. Psoralen Mapping Reveals a Bacterial Genome Supercoiling Landscape Dominated by Transcription. *Nucleic Acids*

Res. **2022**, *1*, 1–14, doi:https://doi.org/10.1093/nar/gkac244.

274. Guo, M.S.; Kawamura, R.; Littlehale, M.; Marko, J.F.; Laub, M.T. High-Resolution, Genome-Wide Mapping of Positive Supercoiling in Chromosomes. *Elife* **2021**, *10*, doi:10.7554/eLife.67236.

275. Kulakovskiy, I. V.; Boeva, V.A.; Favorov, A. V.; Makeev, V.J. Deep and Wide Digging for Binding Motifs in ChIP-Seq Data. *Bioinformatics* **2010**, *26*, 2622–2623, doi:10.1093/bioinformatics/btq488.

276. Zhang, Z.; Cheng, B.; Tse-Dinh, Y.-C. Crystal Structure of a Covalent Intermediate in DNA Cleavage and Rejoining by Escherichia Coli DNA Topoisomerase I. *Proc Natl Acad Sci USA* **2011**, *108*, 3–8, doi:10.1073/pnas.1100300108.

277. Kovalsky, O.I.; Kozyavkin, S.A.; Slesarev, A.I. Archaebacterial Reverse Gyrase Cleavage-Site Specificity Is Similar to That of Eubacterial DNA Topoisomerases I. *Nucleic Acids Res.* **1990**, *18*, 2801–2806, doi:10.1093/nar/18.9.2801.

278. Annamalai, T.; Dani, N.; Cheng, B.; Tse-Dinh, Y.C. Analysis of DNA Relaxation and Cleavage Activities of Recombinant Mycobacterium Tuberculosis DNA Topoisomerase I from a New Expression and Purification Protocol. *BMC Biochem.* **2009**, *10*, 8–15, doi:10.1186/1471-2091-10-18.

279. Tse, Y.-C.; Kirkegaard, K.; Wang, J.C. Covalent Bonds between Protein and DNA. *J. Biol. Chem.* **1980**, *255*, 5560–5565.

280. Baranello, L.; Wojtowicz, D.; Cui, K.; Devaiah, B.N.; Chung, H.-J.; Chan-salis, K.Y.; Guha, R.; Wilson, K.; Zhang, X.; Piotrowski, J.; et al. RNA Polymerase II Regulates Topoisomerase 1 Activity to Favor Efficient Transcription. *Cell* **2017**, *165*, 357–371, doi:10.1016/j.cell.2016.02.036.RNA.

281. Zhou, Q.; Zhou, Y.N.; Jin, D.J.; Tse-Dinh, Y.C. Deacetylation of Topoisomerase I Is an Important Physiological Function of E. Coli CobB. *Nucleic Acids Res.* **2017**, *45*, 5349–5358, doi:10.1093/nar/gkx250.

282. Kim, S.; Beltran, B.; Irnov, I.; Jacobs-Wagner, C. Long-Distance Cooperative and Antagonistic RNA Polymerase Dynamics via DNA Supercoiling. *Cell* **2019**, *179*, 106-119.e16, doi:10.1016/j.cell.2019.08.033.

283. Klindziuk, A.; Kolomeisky, A. Long-Range Supercoiling-Mediated RNA Polymerase Cooperation in Transcription. *J. Phys. Chem. B* **2021**, *4*, doi:10.1021/acs.jpcb.1c01859.

284. Park, J.; Marr, M.T.; Roberts, J.W. E . Coli Transcription Repair Coupling Factor (Mfd Protein) Rescues Arrested Complexes by Promoting Forward Translocation. *Cell* **2002**, *109*, 757–767.

285. Martinez, B.; Bharati, B.K.; Epshtein, V.; Nudler, E. Pervasive Transcription-Coupled DNA Repair in E. Coli. *Nat. Commun.* **2022**, *13*, 1–10, doi:10.1038/s41467-022-28871-y.

286. Uplekar, S.; Rougemont, J.; Cole, S.T.; Sala, C. High-Resolution Transcriptome and Genome-Wide Dynamics of RNA Polymerase and NusA in Mycobacterium Tuberculosis. *Nucleic Acids Res.* **2013**, *41*, 961–977, doi:10.1093/nar/gks1260.

287. Harshey, R.M.; Ramakrishnan, T. Rate of Ribonucleic Acid Chain Growth in Mycobacterium Tuberculosis H37Rv. *J. Bacteriol.* **1977**, *129*, 616–622.

288. Landick, R.; Krek, A.; Glickman, M.S.; Socci, N.D.; Stallings, C.L. Genome-Wide Mapping of the

Distribution of CarD, RNAP ΣA, and RNAP β on the Mycobacterium Smegmatis Chromosome Using Chromatin Immunoprecipitation Sequencing. *Genomics Data* **2014**, *2*, 110–113, doi:10.1016/j.gdata.2014.05.012.

289. Feng, S.; Liu, Y.; Liang, W.; El-Sayed Ahmed, M.A.E.G.; Zhao, Z.; Shen, C.; Roberts, A.P.; Liang, L.; Liao, L.; Zhong, Z.; et al. Involvement of Transcription Elongation Factor GreA in Mycobacterium Viability, Antibiotic Susceptibility, and Intracellular Fitness. *Front. Microbiol.* **2020**, *11*, 1–15, doi:10.3389/fmicb.2020.00413.

290. Li, X.; Mei, H.; Chen, F.; Tang, Q.; Yu, Z.; Cao, X.; Andongma, B.T.; Chou, S.H.; He, J. Transcriptome Landscape of Mycobacterium Smegmatis. *Front. Microbiol.* **2017**, *8*, 1–16, doi:10.3389/fmicb.2017.02505.

291. Santos-Zavaleta, A.; Sánchez-Pérez, M.; Salgado, H.; Velázquez-Ramírez, D.A.; Gama-Castro, S.; Tierrafría, V.H.; Busby, S.J.W.; Aquino, P.; Fang, X.; Palsson, B.O.; et al. A Unified Resource for Transcriptional Regulation in Escherichia Coli K-12 Incorporating High-Throughput-Generated Binding Data into RegulonDB Version 10.0. *BMC Biol.* **2018**, *16*, 1–12.

Supplementary Materials

Supplementary Table 1. Primers and oligonucleotides used in this study

Primer	Primer sequence	Description
Mu_1_F	CGACGTTGTAAAACGACGGCC	Targeting MuSGS site with qPCR
Mu_1_R	CGCGCCACGGGATTTTTAAACC	
L_2394_F	CCAGTACCAACAGTGCAGAGATC	Targeting *ccmH* site with qPCR
L_2394_R	GGTGAGCCTGTATCTTCACGC	
H_3594_F	ATTGAAGATTCACTGGCGGATTACC	Targeting rRNA A DS site with qPCR
H_3594_R	CTCACCAGCTTCCTGCTGC	
L_3602_F	CCAGAAAGTCGGGGTAAAGAGC	Targeting rRNA A US site with qPCR
L_3602_R	CTTTACCAGTATGAACCCGGTGG	
G_cat_Mu	CATAATAATGTGTGCAAATTGATGGTGGATTGATGCAAATTT GTTAATTGGAATAGGAACTTCATTTAAATGGCGCG	*cat* gene amplification from pKD3 plasmid
cat_Mu_R	TTACCCAACTTAATCGCCTTGCAGCACATCCTACCTGTGACG GAAGATCACTTCG	
Mu_G_F	TCTGCGAAGTGATCTTCCGTCACAGGTAGGATGTGCTGCAAG GCGATTAAGTTGG	MuSGS site amplification from pMP1000 plasmid
Mu_G_R	AAAGAAAAACCAAGTCTTTTTTGATGACAAAATGCAATCAA G GAAAAATGGGCGGTACACCCGATATGG	
topA_NcoI_fw	GTAAGGTGACCATGGGTAAAGCTCTTG	
NcoI_mut_rev	GGCTAAAGCGAACCACGGTCTTATTACC	
NcoI_mut_fw	GGTAATAAGACCGTGGTTCGCTTTAGCC	
topA_HindIII_strepII_rev	atataaagcttttaTTTTTCGAACTGCGGGTGGCTCCAcgcgctttttttccttca acccatttgccatc	
topA_NdeI_fw	GTAAGGTGCatATGGGTAAAGCTCTTGTCATCG	
topA_strepII_HindIII_rev	atatcaagcttTTATTTTTCGAACTGCGGGTGGCTCCAC	
TopoA_14_kDa_CTD_BamHI_fw	GGATCGCATCACCATCACCATCACGGATCCGGCGAAGTGGC A CCACCG	
TopoA_CTD_HindIII_rev	aacaggagtccaagctcagctaattaagctTTATTTTTTTCCTTCAACCCATTT G CC	
pCA24_core_NcoI_rev	AATCTAACCATGGATCCGTGATGGTGATGGTGATGC	
pCA24_core_HindIII_fw	AAACTGGGGCACAGATGAAGCTTAATTAGC	
topA_Y319F_fw	TATATCACTTTCATGCGTACCGACTCCACTAACC	
topA_Y319F_for	AGTCGGTACGCATGAAAGTGATATAGCCTG	
topA_XbaI_RBS_fw	Tatataactctagaaataattttgtttaactttaagaaggagatataccatgggcagcagccatcatcat cat catcacagcagcggcctggtgccgcgcggcagccataTGGGTAAAGCTCTTGTCA T CG	
topA_G116S_out_rev	ATGCAATGGCTTCAGATTCGCGGTCAAGGTCGG	
topA_G116S_in_for	TTGACCGCGAATCTGAAGCCATTGCATGGCACC	
topA_M320V_in_rev	AGTCGGTACGCACGTAAGTGATATAGCCTGCTTCATAC	
topA_M320V_out_for	ATCACTTACGTGCGTACCGACTCCACTAACC	
topA_strepII_HindIII_rev	atatcaagcttTTATTTTTCGAACTGCGGGTGGCTCCAC	
dps_850_F	CGGTGTAGAGGAAAAGTAGCGAGAAATTCTGC	
dps_850_R	ATGGGGTCTACGCTGACAGTACG	
potF_895_F	CGGGAGAAAGTTCTCTTTTCTTACACCG	
potF_895_R	GCGGTCATATCTCTTTCCTTCTGAAAGTTCG	
L_2394_F	CCAGTACCAACAGTGCAGAGATC	

L_2395_R	GGTGAGCCTGTATCTTCACGC	
topA_delta_topA66_kanR_F	CACCACCGAAAGAAGATCCGGTGCCATTACCTGAGCTGCCGT GCGAAAAATAATAGTGATCGTGTAGGCTGGAGCTGCTTC	
topA_delta_14kDa_kanR_F	GTAAATACATGGCCTGCACCAACGAAGAGTGTAAAAACACAC GTAAGATTTTACGTAACTAATAGTGAATCGTGTAGGCTGGAGC TGCTTC	
topA_delta_30kDa_kanR_F	AGTTAGATAAAGCTGAAAAAGATCCGGAAGAGGGTGGTATGC GCCCGAACTAATAGTGAATCGTGTAGGCTGGAGCTGCTTC	
topA_SPA_kanR_cysB_R	ATGCAATAAAAAAGGGCCGCTTTCGCGACCCTTTGTTTATA AAAACCTGACAGAATTAAAGGTGAAGAAGTAAAACATATG AATATCCTCC	
BW25113_topA_wt_F	GCCGCTTTCGCGACCCTTTG	
BW25113_topA_wt_R	CGACTGGCTGGTCAGCATTTTATGTTGATGG	
BW25113_topA_delta11_SPA_F	GAGCTGCCGTGCGAAAAATCCATG	
BW25113_SPA_qPCR_R	CAAGTGCCCCGGAGGATGAGATTTTC	
BW25113_topA_delta14_SPA_F	ACACACGTAAGATTCTGCGCAATTCCATG	
BW25113_SPA_qPCR_R	CAAGTGCCCCGGAGGATGAGATTTTC	
gyrA_S83L	GCGCCATGCGGACGATCGTGTCATAGACCGCCAGATCACCA TGGGGATGGTATTTACCGATTACGTCACC	
gyrA_F	ATGAGCGACCTTGCGAGAGAAATTACACCG	
gyrA_R	CCGGAATTTTTTCCGTGCCGTC	

Supplementary Table 2. Links to GitHub repositories with code written for data analysis and data visualization for this work

Dataset	Link
E. coli DNA gyrase Topo-Seq	https://github.com/sutormin94/Gyrase_Topo-Seq
E. coli DNA gyrase time-course Topo-Seq	https://github.com/sutormin94/Gyrase_time-course-experiment
E. coli TopoIV Topo-Seq	https://github.com/sutormin94/TopoIV_Topo-Seq_experiment
E. coli TopoI Topo-Seq	https://github.com/sutormin94/TopoI_Topo-Seq
E. coli RNA-Seq	https://github.com/sutormin94/E_coli_RNA-Seq_analysis

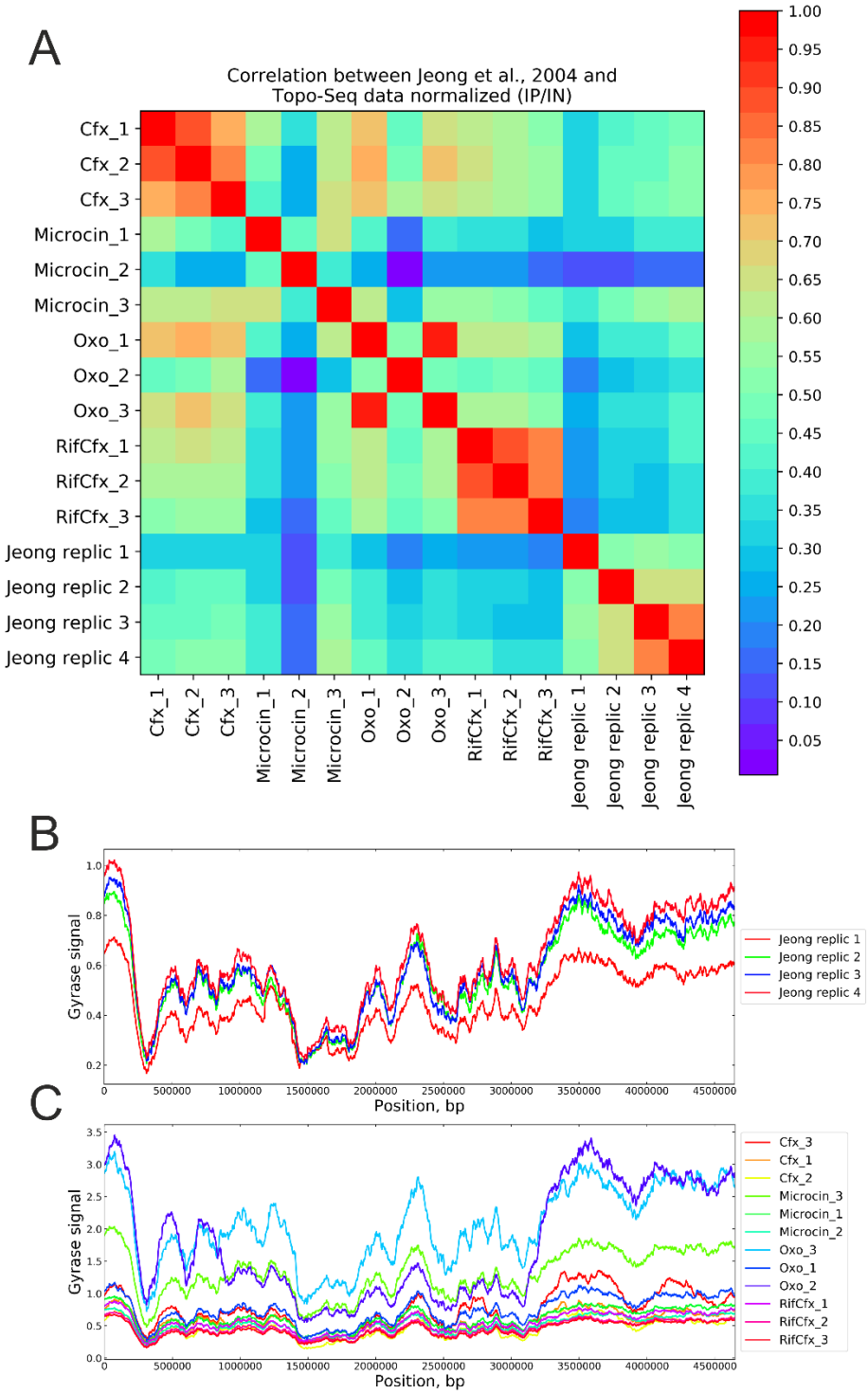
Supplementary Table 3. NGS datasets generated, analyzed, and re-analyzed in this work

Dataset	Dataset ID and links	Description	Study
E. coli DY330 DNA gyrase Topo-Seq	GSM3273139-GSM3273144	Topo-Seq performed for exponentially growing culture in LB. Cfx, Oxo, and Micro were used to trap covalent intermediate complexes between gyrase and DNA. Performed in triplicate.	This work, [113]
E. coli DY330 DNA gyrase Topo-Seq for cells pre-treated with Rif	GSM3273145-GSM3273150	Topo-Seq performed for exponentially growing culture in LB. Culture was pre-treated with 100 µg/ml Rif for 20 min before adding the Cfx. Performed in triplicate.	This work, [113]
E. coli DY330 pBR322 DNA gyrase Topo-Seq	Not published	Topo-Seq performed for exponentially growing culture in LB. Covalent complexes trapped with Cfx. Performed in triplicate.	This work
E. coli DY330 DNA gyrase time-course Topo-Seq	Not published	Time-resolved DNA gyrase Topo-Seq on synchronized *E. coli* cells. Covalent complexes trapped with Cfx. Performed in duplicate.	This work
E. coli DY330 DNA gyrase Topo-Seq for different growth phases	Not published	DNA gyrase Topo-Seq conducted with exponentially growing, stationary, and transition to stationary cultures. Covalent complexes trapped with Cfx.	This work
Caulobacter crescentus DNA gyrase Topo-Seq	Not published	DNA gyrase Topo-Seq conducted with exponentially growing culture. Covalent complexes trapped with Cfx.	This work
E. coli DY330 TopoIV Topo-Seq	Not published	Topo-Seq performed for exponentially growing culture in LB. Covalent complexes trapped with Cfx. Performed in triplicate.	This work
E. coli DY330 *gyrAS83L* TopoIV Topo-Seq	Not published	Topo-Seq performed for exponentially growing culture in LB. Covalent complexes trapped with Cfx. Performed in triplicate.	This work

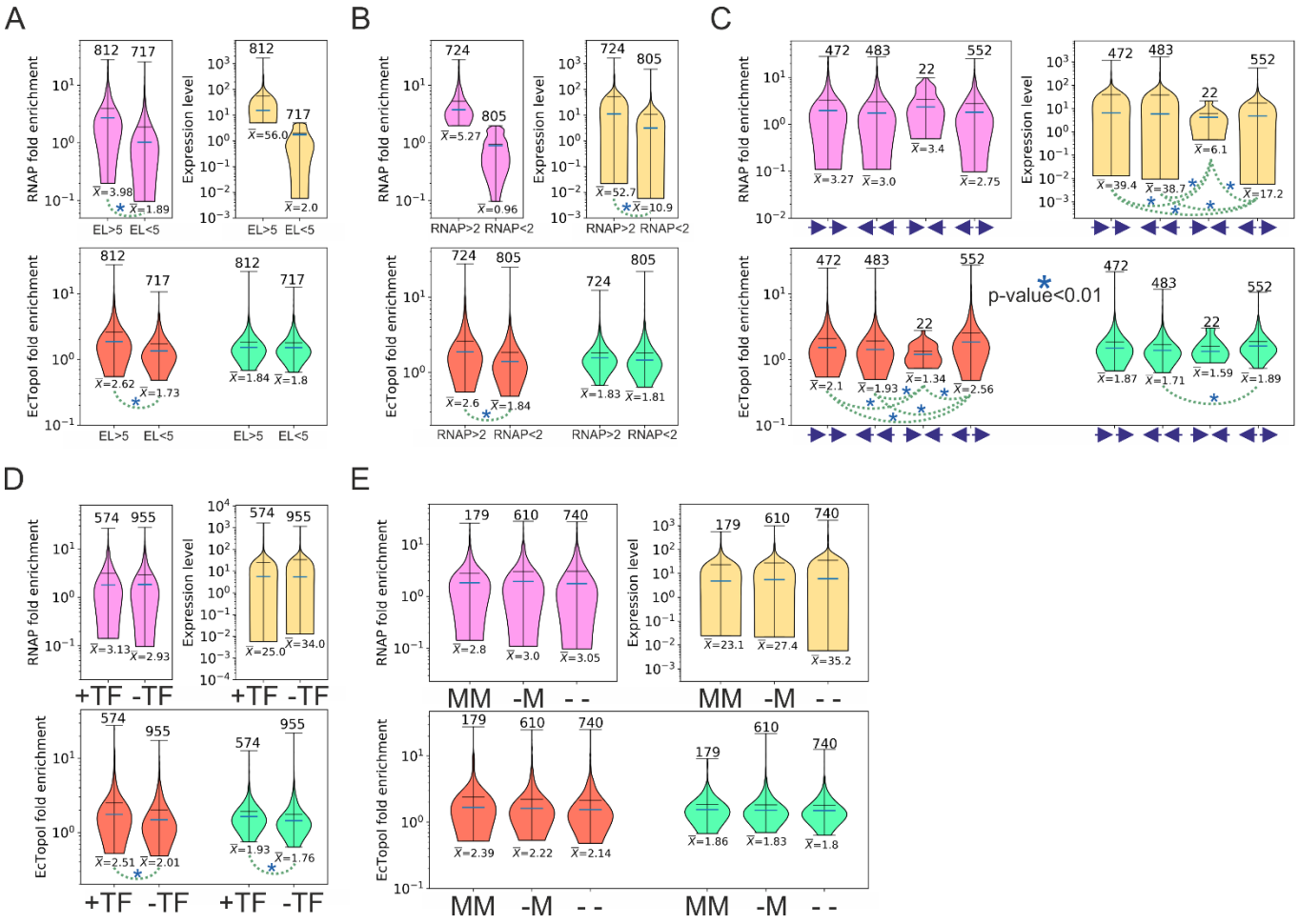
E. coli DY330 EcTopoI Topo-Seq	GSM5529907-GSM5529918	Topo-Seq performed for exponentially growing culture in LB. EcTopoI G116S/M320V mutant was used to obtain covalent intermediates. Performed in triplicate.	This work
E. coli DY330 RNA-Seq	GSM5509255-GSM5509263	Total RNA-Seq for exponentially growing, stationary, and transition to stationary cultures in LB. Performed in triplicate.	This work
E. coli DY330 RpoC ChIP-Seq	GSM5538289, GSM5538290	ChIP-Seq performed for exponentially growing culture in LB. Single replicate.	This work
E. coli DY330 EcTopoI ChIP-Seq	GSM5514277-GSM5514282	ChIP-Seq performed for exponentially growing culture in LB. Performed in triplicate.	This work
E. coli DY330 EcTopoI ChIP-Seq with cells pre-treated with Rif	GSM5514289-GSM5514294	ChIP-Seq performed for exponentially growing culture in LB. Culture was pre-treated with 100 µg/ml rifampicin for 20 min before crosslinking. Performed in triplicate.	This work
E. coli MG1655 RpoB ChIP-Seq	GSM613808	ChIP-Seq performed for exponentially growing culture in LB. Single replicate.	[249]
E. coli MG1655 RpoC ChIP-chip for cells pre-treated with Rif	GSM351003	ChIP-chip performed for exponentially growing culture in M9 medium. The culture was pre-treated with Rif before the experiment.	[270]
E. coli MG1655 GapR-Seq	GSM4628311-GSM4628314	ChIP-Seq performed for exponentially growing aerobic culture in LB.	[274]
E. coli MG1655 Psora-Seq	SRR17974444-SRR17974449, SRR17974460, SRR17974471, SRR17974480-SRR17974483	ChIP-Seq performed for exponentially growing aerobic culture in LB.	[273]
Mycobacterium tuberculosis DNA gyrase ChIP-Seq	GSM2538162	ChIP-Seq performed on exponentially growing culture of MtbRa cells.	[66]
Mycobacterium tuberculosis RNAP ChIP-Seq	GSM1003214, GSM1003215	ChIP-Seq performed on exponentially growing culture.	[286]

Mycobacterium tuberculosis RNA-Seq	GSM1003224, GSM1003225	RNA-Seq performed for exponentially growing culture.	[286]
Mycobacterium smegmatis RNAP ChIP-Seq	GSM1171544, GSM1171545	ChIP-Seq performed on exponentially growing culture.	[288]
Mycobacterium smegmatis TopoI ChIP-Seq	SRX4970107	ChIP-Seq performed on exponentially growing culture.	[112]
Mycobacterium smegmatis Mock DNA for ChIP-Seq normalization	GSM4274349, GSM4274350	Mock DNA data	[289]
Mycobacterium smegmatis RNA-Seq	GSM2756262, GSM2756263	RNA-Seq performed for exponentially growing culture.	[290]
Streptococcus pneumoniae TopoI and RNAP ChIP-Seq	SRR12427932, SRR12427936, SRR12427942	ChIP-Seq performed for exponentially growing culture.	[181]

Supplementary Figure 1. **Correlation and correspondence between gyrase binding data (Jeong et al., 2004** [65]**) and cleavage data revealed with Topo-Seq data**. (**A**) Heatmap represents the Pearson correlation coefficients were calculated between 4 replicas of Jeong et al., 2004 gyrase binding data and all replicas of normalized Topo-Seq data (IP/Input). (**B**) Binding of DNA gyrase to the *E. coli W3110* chromosome according to Jeong et al., 2004. Smoothed with a 200 kb sliding window. (**C**) Gyrase enrichment (IP/Input) over the *E. coli W3110* genome from Topo-Seq experiments, smoothed with a 200 kb sliding window.



Supplementary Figure 2. **Factors associated with high or low EcTopoI enrichment signal at intergenic regions (IRs).** (**A**) Effect of a high-level expression of adjacent genes. IRs were classified as regions associated with a high level of expression if the cumulative expression of adjacent genes was higher than 5 FPKM units (**EL>5**) and as regions associated with a low level of expression if the cumulative expression of adjacent genes was lower than 5 FPKM units (**EL<5**). Then, for each group of IRs, RNAP fold enrichment (**pink**), the expression level of adjacent genes (**yellow**), and fold enrichment of EcTopoI in Rif-/CTD- (**red**) and Rif+/CTD- (**green**) conditions were identified. Violin plots demonstrate the distribution of the parameter's values; the vertical axis is log-scaled. The sample size is indicated above the plots, means are marked with a horizontal black line and indicated below with ($\overline{X}$), medians are marked as a blue horizontal line. If present, statistically significant differences between means are visualized with dashed lines with stars (p-value<0.01, Welch t-test). (**B**) Violin plots showing the effect of RNAP with high and low enrichments at IR. IRs were classified as regions with a high RNAP signal if the fold enrichment was higher than 2 (**RNAP>2**) or as regions with a low RNAP signal if the fold enrichment was low (**RNAP<2**). The color coding and denotations are the same as in **A**. (**C**) Violin plots for IRs classified by the orientation of adjacent genes: genes can be in the same orientation on the positive or negative strand (→→ or ←←), divergent (←→), or convergent (→←). (**D**) Violin plots for IRs containing (**+TF**) and not containing (**-TF**) annotated binding sites for transcription factors (based on data from RegulonDB [291]). (**E**) Violin plots for IRs classified by the presence of adjacent genes encoding membrane-localized protein: both genes encode membrane protein (**MM**), one gene encodes membrane protein (**M-**), and neither gene encodes a membrane protein (--). The subcellular localization of *E. coli* proteins data was taken from EcoCyc database.



Supplementary Figure 3. **Colocalization of TopoI and RNAP and DNA gyrase and RNAP in mycobacteria; TopoI enrichment in *S. pneumoniae***. (**A**) The Venn diagram represents the number of overlapping RNAP and TopoI peaks in *M. smegmatis*. (**B**) Enrichment of MsmRNAP in regions occupied by MsmTopoI peaks and enrichment of MsmTopoI in regions occupied by MsmRNAP peaks. (**C**) The Venn diagram represents the number of overlapping RNAP and gyrase peaks in *M. tuberculosis*. (**D**) Enrichment of MtbGyrase in regions occupied by MtbRNAP peaks and enrichment of MtbRNAP in regions occupied by MtbGyrase peaks. (**E**) Metagene plot represents enrichment of SpTopoI and SpRNAP over a set of all *S. pneumoniae* TUs. ChIP-Seq data is taken from [181]. For all panels ChIP-Seq fold enrichment is given relative to the input sample.