DNA in a twist? How topoisomerases solve topological problems in DNA

Natassja G. Bush, Monica Agarwal, Sara R. Henderson, Nidda F. Waraich and Anthony Maxwell (John Innes Centre, UK)

DNA topoisomerases have been described as 'the magicians of the DNA world', somehow allowing DNA strands to pass through each other. These ingenious enzymes are both essential and potentially dangerous, as any interruption in their DNA breakage-reunion reactions can lead to chromosome breaks and cell death. This has led to their development as important targets for antibacterial and anti-cancer chemotherapy.

DNA topology and the importance of topoisomerases

Watson and Crick, upon proposing the double-helical structure of DNA, noted that the two strands of the helix would have to untwist in order to separate. They also suggested that this untwisting and separating of the two strands would lead to entanglements. While the structure of the DNA double helix is ideally suited to its function– the opening of the helix permitting access to the bases that encode the genetic information – it does indeed lead to topological problems. This process of helix unwinding, which occurs in DNA replication and transcription, leads to DNA supercoiling, where the DNA helix is coiled about its own axis (Figure 1). The DNA topoisomerases have evolved to deal with problems of over-winding (and under-winding) of the DNA helix, as well as to take care of DNA catenanes (interlinked DNA molecules) that can occur at the termination of replication and knots in DNA that can arise during recombination events (Figure 1). These different forms of DNA: supercoiled, relaxed, catenated and knotted, represent topological isomers of double-stranded circular DNA, i.e. they have the same molecular mass, but differ in the path taken by the DNA strands.

DNA topology, literally how the two complementary single strands of DNA are intertwined, is an important topic in itself and is often the domain of mathematicians. However, for the purpose of discussing topoisomerases, a broad understanding is all that is necessary (see Box 1). In general terms, positive supercoiling (over-winding of DNA) is bad, with some exceptions, and moderate negative supercoiling (under-winding of DNA) is good; knots are generally bad and catenanes normally should be removed (Figure 1). Resolving these unwanted species is the job performed by topoisomerases, which, remarkably, carry out these reactions by transiently cleaving one or both strands of DNA (see below). DNA breakage is a danger to cell viability and the cleavage-religation process must be carefully controlled. Interruption of this process can lead to broken DNA and cell death, which is why these enzymes have been exploited as targets for both antibacterial chemotherapy and anti-cancer agents. As topoisomerases occur in all organisms (prokaryotes, eukaryotes and archaea) it is the differences between them that need to be established and exploited to develop successful clinical agents.

Mechanistic aspects of type II topoisomerases

DNA topoisomerases are divided into two types, I and II, depending on whether their reactions proceed via transient single- or double-stranded DNA breaks. They are further divided into subtypes: IA, IB, IC, IIA, IIB; these subtypes differ both

mechanistically and structurally (Figure 2). Type I enzymes are less relevant to this article, as they are not strictly molecular machines and do not use ATP. The exception is the type IA topoisomerase reverse gyrase that uses ATP to introduce positive supercoils into DNA. Despite this, the type I topoisomerases are still of considerable interest as drug targets (see below) and because the subtypes show fundamental mechanistic differences.

The type IB and C enzymes relax supercoiled DNA using a 'swivel' mechanism by making transient single-stranded breaks involving a covalent bond to the 3′-phosphate end of the DNA, and allowing the DNA to swivel within the enzyme in a controlled rotation. The type IA enzymes work by a strand-passage mechanism making transient covalent bonds to the 5′-phosphate end of DNA at the break site. The type IA enzymes also relax supercoiled DNA, with reverse gyrase being the only type IA enzyme capable of introducing supercoils. It consists of two domains, a helicase domain, which binds and hydrolyses ATP, and a topoisomerase domain, which cleaves the DNA. The helicase domain transiently unwinds the double-stranded DNA in an ATP-dependent reaction increasing the linking number (Lk) (see Box 1). The topoisomerase domain carries out strand passage by cleaving one strand of the DNA helix and passing the other strand through the cleaved strand before resealing. This strand-passage event results in positive supercoiling. Reverse gyrase has only been found in thermophilic archaea and eubacteria. These organisms are thought to have positively supercoiled genomes, which, in this case, is beneficial as positively supercoiled DNA is thought to be more resistant to the harmful effects of high temperature.

Type II topoisomerases use the free energy provided by ATP hydrolysis to carry out their reactions in which one segment of double-stranded DNA is cleaved and another segment is passed through the break. The two subtypes, type IIA and type IIB, are similar mechanistically but differ structurally (Figure 2). The type II topoisomerases are either homodimers (e.g. human topoisomerase (topo) IIα) or heterotetramers (e.g. prokaryotic topo IV). Although the quaternary structures may differ, all type IIA enzymes share similar domains (Figure 2). The biggest structural difference between the two subtypes is that the type IIB topoisomerases lack the bottom protein interface or 'exit gate', which has mechanistic implications as there is no post-strand-passage cavity (Figures 2 & 3). This post-strand-passage cavity has been suggested to have a 'proofreading' role, potentially allowing backtracking of the transported segment.

DNA topology

If a linear double-stranded DNA molecule is closed into a circle with both strands sealed, then the two strands are 'linked' together a number of times, corresponding to the number of double-helical turns in the original linear molecule. This number, the linking number (Lk), must be an integer and its value will depend on the number of base pairs in the DNA (N) and the helical repeat (h), generally taken to be ~10.5 in B-form DNA:

$$Lk = N/h$$

Strictly, the linking number of this most 'relaxed' circular molecule will be the closest integer to N/h – the ends may need to be twisted slightly to join up. The exact non-integer value of N/h is called Lk°, representing the imaginary undistorted circular molecule. If the ends of the original linear molecule had been rotated a number of turns either in the same direction as the helical twist of the double helix, or in the opposite direction, before joining, this changes the linking number, and the circular DNA is said to be 'supercoiled', and has a linking difference, ΔLk, where:

$$\Delta Lk = Lk - Lk^{\circ}$$

ΔLk can be positive (corresponding to positive supercoiling) or negative (corresponding to negative supercoiling). The distinction between Lk° and the true integer value of Lk for the relaxed molecule is largely insignificant for molecules of plasmid size or above.

Lk can be broken down into two parameters: Tw, the twist of the DNA, i.e. how many times the two strands coil around the helix axis (the number of double-helical turns), and Wr, writhe, the coiling of the helix axis in space:

$$Lk = Tw + Wr$$

Tw and Wr can be interconverted and do not have to be integers. The term DNA supercoiling, the higher order coiling of the DNA, is really describing the writhe, but the overall distortion of the molecule is really represented by the linking difference (ΔLk), and the density of supercoiling (σ), or specific linking difference, is defined as:

$$\sigma = \Delta Lk / Lk^{\circ}$$

For a fuller description of supercoiling, knots and catenanes see Bates, A.D. and Maxwell, A. (2005) DNA Topology, Oxford University Press, Oxford.

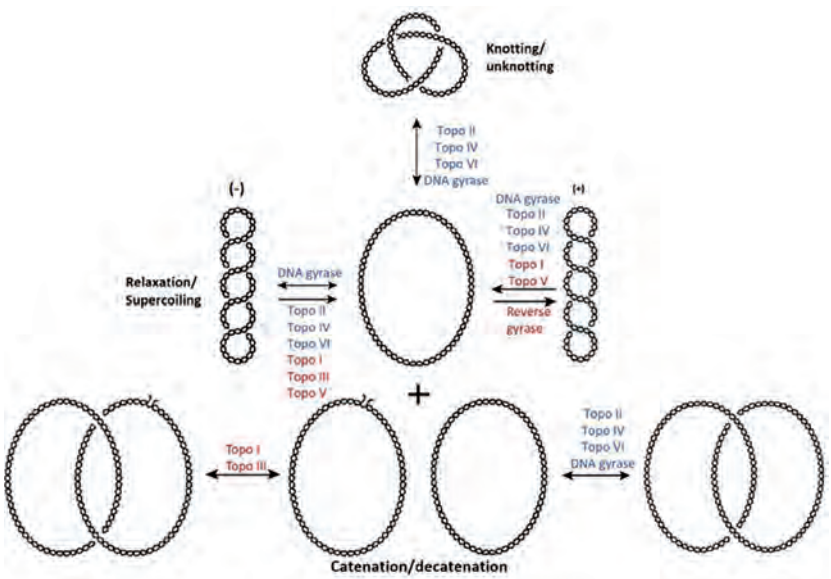


Figure 1. Various DNA topological reactions catalysed by DNA topoisomerases. Type I topoisomerases are indicated in red and the type II topoisomerases are indicated in blue. Arrows represent the direction of the reaction. For illustrative purposes, in the catenation/decatenation reactions, the non-nicked substrates are drawn as relaxed; however, it is more likely that these are supercoiled in this reaction. Figure has been adapted from Bush et al. (2015).

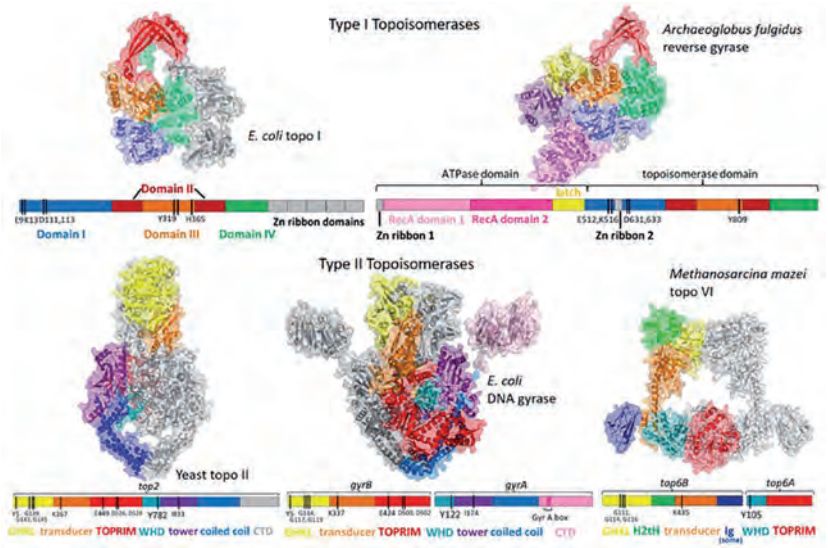


Figure 2. Domain diagrams and structures of type I and type II DNA topoisomerases. Type I topoisomerases: *E. coli* topo I (left-hand side; PDB: 4RUL) and *Archaeoglobus fulgidus* reverse gyrase (right-hand side; PDB: 1GKU) coloured by their various domains indicated in the domain diagrams below. Blue is Domain I, red – Domain II, orange – Domain III, green – Domain IV, grey shows the Zn-binding domains. The pink (light and dark) on the *A. fulgidus* reverse gyrase indicates the two RecA domains and the yellow shows the latch domain. Type II topoisomerases: yeast (*Saccharomyces cerevisiae*) topo II (left-hand side) – truncated structure (lacking the CTD) showing amino acids 1-1177 (PDB: 4GFH), *E. coli* DNA gyrase (middle) – assembled from a number of truncated structures (PDB: 1EI1, 3NUH and 3L6V) and *Methanosarcina mazei* topo VI (right-hand side, PDB: 2Q2E) all with their respective domain diagrams below. Yellow shows the GHKL domain which binds ATP, orange indicates the transducer domain, red the TOPRIM domain which binds the Mg^{2+} ion cofactor, teal is the winged-helix domain (WHD), which contains the catalytic tyrosine, purple is the tower domain and blue shows the coiled coil domain. The grey on the yeast topo II domain diagram (missing in the structure) shows the C-terminal domain (CTD), which is structurally and functionally different to the *E. coli* DNA gyrase CTD (light pink). The *M. mazei* structure has two additional domains in the helix-2-turn-helix (H2tH in green) and the Ig domain (in dark blue).

All type II topoisomerases use strand passage to relax DNA or resolve knots and catenanes, but only DNA gyrase can introduce supercoils into DNA. Strand passage occurs by a two-gate mechanism in which one segment of DNA (the T or transported segment) is passed through a transient break in another (the G or gate segment) (Figure 3). Supercoiling via gyrase works by a slightly different mechanism in which DNA is wrapped around the C-terminal domains of GyrA, presenting the T segment to the G segment such that strand passage, driven by the binding and hydrolysis of ATP, occurs in the correct direction (Figure 3).

Although we have an understanding of the mechanism of strand passage, the role of ATP hydrolysis is less clear. With gyrase it seems obvious that the free energy is utilized to drive negative supercoiling. However, this is not the case for the other type II topoisomerases, which catalyze reactions that, in general, do not require energy. One proposal is that binding and hydrolysis of ATP is used to stabilize or destabilize protein interfaces within type II topoisomerases allowing efficient strand passage while reducing the risk of accumulating potentially toxic double-stranded breaks.

Type II topos as drug targets

A major reason for the continued interest in DNA topoisomerases over the last 40 years or so has been their utility as drug targets both for antibacterial and anti-cancer chemotherapy. The quinolone class of antibiotics, of which ciprofloxacin is the best known, are the only clinically successful bacterial topoisomerase inhibitors. They have a dual-targeting mechanism of action whereby they can inhibit both type II bacterial topoisomerases (gyrase and topo IV). Their mode of action is well-studied and involves stabilizing a cleavage complex with bound DNA. This can be bacteriostatic (i.e. prevents replication, most likely a result of stalling of the replication fork due to the presence of a gyrase-DNA complex), or bactericidal (i.e. kills the bacteria) due to induction and processing of double-stranded breaks in DNA. As with most antibiotics there is significant resistance, and it is imperative that new agents to replace quinolones are found.

Aside from quinolones, there are several other small molecules and toxins that inhibit gyrase and, in some cases, also topo IV. These include novobiocin and simocyclinone, and the proteinaceous toxins microcin B17 and CcdB (Figure 4). Novobiocin is an aminocoumarin compound produced by *Streptomyces* species. These are extremely potent inhibitors of gyrase (and topo IV) that act by competitively inhibiting the ATPase reaction. Despite having little similarity to ATP, aminocoumarins bind at the ATPase active site (Figure 4). In contrast, simocyclinones, which also contain an aminocoumarin moiety, do not inhibit the ATPase reaction, but instead bind to two pockets in GyrA preventing the enzyme's interaction with DNA (Figure 4). There are issues with both the aminocoumarins and simocyclinones in terms of toxicity and solubility that currently preclude their use in clinical medicine.

Recently, a new novel class of gyrase-specific compounds with a distinct mode of action was discovered: the thiophenes. These compounds bind allosterically in a pocket away from the quinolone-binding site, resulting in a conformation that promotes cleavage complexes. These compounds have not been optimized for clinical use, and it is yet to be determined if this will be feasible.

Although bacterial type I topoisomerases (e.g. topo I) are potential targets for antibacterial chemotherapy, there are currently no clinical agents that target these compounds; this remains an active area of research.

Human topoisomerases (topo I and II) have been successfully exploited in anti-cancer chemotherapy; examples include camptothecin, etoposide and amsacrine. Like quinolones, these drugs stabilize cleavage complexes with DNA. In addition to these topo II poisons, there are other compounds that act via different modes of action, including inhibiting ATP hydrolysis.

Topo VI, a type IIB topoisomerase, is not currently a drug target, as archaea are not generally regarded as pathogens. However, its occurrence in plants and Plasmodium parasites (see below) raises the possibility of targeting topo VI for the development of herbicides and antimalarial drugs. A few inhibitors of archaeal topo VI have been reported.

Finding new topoisomerases in unexpected places

DNA topoisomerases have been found in all organisms, but there is considerable variation in the number and types of topoisomerases. For example, *Mycobacterium tuberculosis* has two topoisomerases, topo I (type IA) and gyrase (type IIA). In contrast, *Escherichia coli* has four topoisomerases with an additional enzyme in each class, topo III (type IA) and topo IV (type IIA). DNA gyrase in *M. tuberculosis* possesses enhanced decatenation ability, an activity normally attributed to topo IV. In recent years, gyrase and topo VI have been found in eukaryotes including *Arabidopsis thaliana* (a model plant species) and *Plasmodium falciparum* (a causative agent of malaria).

The genes for eukaryotic gyrases are nuclear-encoded but the proteins are organelle-targeted in both *A. thaliana* and *P. falciparum*. The enzyme is targeted to both the chloroplast and mitochondria in *A. thaliana*, and to the apicoplast, a relict plastid, in *P. falciparum*. The requirement for gyrase in eukaryotes may at first seem unnecessary given the ability of these organisms to negatively supercoil DNA through the action of DNA wrapping around histones. However, in accordance with the bacterial origins of these organelles, they appear to have retained prokaryotic aspects of organellar DNA replication and organization.

In *A. thaliana* there are two genes for GyrB; it is currently not clear what the roles of the two B subunits are. *A. thaliana* gyrase is sensitive to quinolone drugs, like its prokaryotic counterpart,

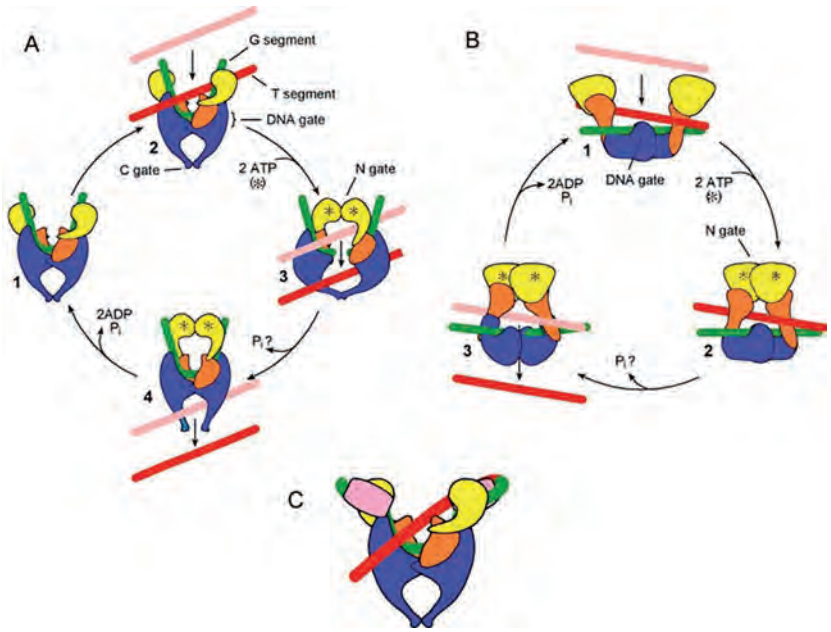


Figure 3. Mechanism of type II topoisomerases. The DNA T-segment (coloured pink to red) is transported through the enzyme-stabilized double-stranded break in the DNA G-segment (green). (A) is the core strand-passage mechanism of type IIA topoisomerases. (B) shows the type IIB topoisomerase mechanism, note the lack of the post-strand-passage cavity. (C) shows the wrapping of the DNA around DNA gyrase C-terminal domains (light pink). For all panels the domains are approximately coloured by domain: yellow is the GHKL domain, orange shows the TOPRIM domain and blue indicates the cleavage-religation domain; adapted from Bates et al. 2011.

and it has been shown that plants can be killed by these compounds, raising the possibility of developing herbicides based on quinolones.

Relatively little is known about *P. falciparum* gyrase. This is possibly due to the difficulties of heterologous protein expression attributable to the AT-rich genome of the malarial parasite. *P. falciparum* gyrase B has successfully been characterized, but the A subunit has proved more difficult, with only truncated fragments producing soluble protein. The initial indication of the presence of gyrase in *P. falciparum* came from the sensitivity of the apicoplast to ciprofloxacin. Given the drastic health impacts of malaria, developing this enzyme as a drug target is viewed with importance.

Topo VI (type IIB) is a DNA-relaxing enzyme originally found in archaea. In plants it is required for endoreduplication, a polyploidization process responsible for the enlargement of plant cells, which determines plant size. Topo VI is related to type IIA topoisomerases in terms of structure/function, but differs in the strand-passage mechanism (Figures 2 & 3). It is a heterotetramer composed of two subunits, A and B. Subunit A is a homologue of the Spo11

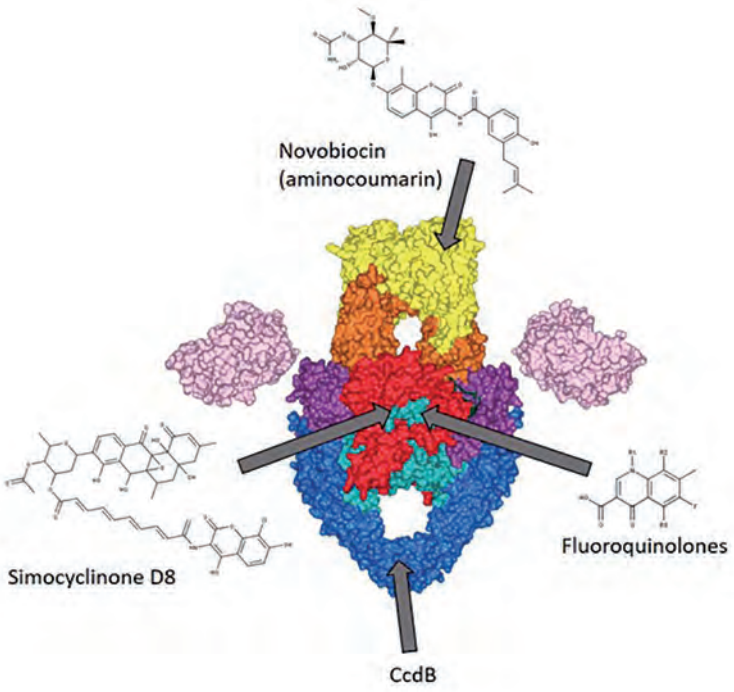


Figure 4. Model of full-length DNA gyrase from *Mycobacterium tuberculosis* based on homology modelling to other known structures, indicating the binding sites of inhibitors; adapted from Nagaraja et al. 2017.

protein, required for the double-strand DNA breaks that initiate recombination in meiosis; subunit B has conserved ATP binding and hydrolysis domains. *Arabidopsis* topo VI function appears to require the activity of two accessory proteins, RHL1 and BIN4.

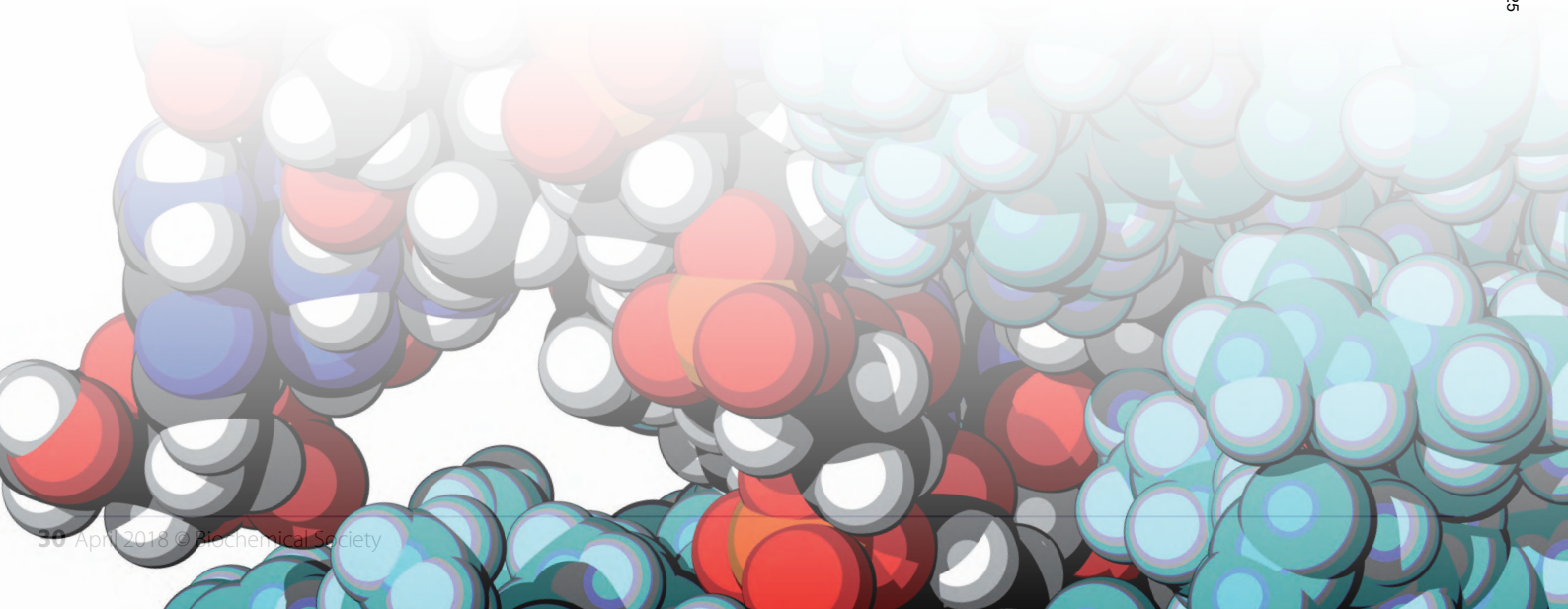
Genome sequencing of *P. falciparum* revealed the presence of a topo VI homologue. The lifecycle of *P. falciparum* involves a stage termed 'schizogony' that resembles endoreduplication, suggesting that topo VI may be involved in this process. Indeed, topo VI could also serve as a new target in malarial therapy.

A new addition to type IIB topoisomerases is topo VIII, a homodimer, with the fusion of the B and A subunits of type IIB topoisomerases into a single polypeptide. Three members of this new subfamily have been found: one from a mesophilic bacterium, one from a thermophilic bacterium and one from a bacterial plasmid. Although there is limited similarity between topo VIII and topo VI, the modular organization is the same. The architecture of topo VIII can be compared with that of topo VI based on the structural similarities of the individual domains. The role of topo VIII remains to be established.

Future directions

DNA topoisomerases are of considerable interest in terms of mechanistic enzymology and there are important aspects of their reactions that we still do not understand, such as the utilization of ATP by the type II enzymes. Their key roles in many DNA-associated processes has made them of considerable interest to several fields, including chromosome biology, DNA replication, transcription, etc. Their large size has made structural analysis challenging but recent success using X-ray crystallography and the application of cryo-electron microscopy will shed new light on structure/mechanism issues. They will remain subjects of intense study due to their importance as drug targets, and the finding of new topoisomerases will further illuminate their various roles in DNA metabolism. We confidently expect to continue to find new 'twists' to the topoisomerase story. ■

Work in A.M.'s lab is funded by the Biotechnology and Biosciences Research Council (UK) Institute Strategic Programme Grant BB/P012523/1. We thank Andy Bates for helpful comments.



Further reading

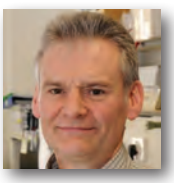
- Bates, A.D. and Maxwell, A. (2005) DNA Topology, Oxford University Press, Oxford
- Bush, N.G., Evans-Roberts, K. and Maxwell, A. (2015) DNA topoisomerases. EcoSal Plus, **6**
- Collin, F., Karkare, S. and Maxwell, A. (2011) Exploiting bacterial DNA gyrase as a drug target: current state and perspectives. Appl. Microbiol. Biotechnol., **92**, 479–497
- Bates, A.D., Berger, J.M. and Maxwell, A. (2011) The ancestral role of ATP hydrolysis in type II topoisomerases: prevention of DNA double-strand breaks. Nucleic Acids Res., **39**, 6327–6339
- Nagaraja, V., Godbole, A.A., Henderson, S.R. and Maxwell, A. (2017) DNA topoisomerase I and DNA gyrase as targets for TB therapy. Drug Discov. Today, **22**, 510–518
- Chan, P.F., Germe, T., Bax, B.D., et al. (2017) Thiophene antibacterials that allosterically stabilize DNA-cleavage complexes with DNA gyrase. Proc. Natl Acad. Sci. USA, **114**, E4492–E4500
- Tse-Dinh, Y.C. (2009) Bacterial topoisomerase I as a target for discovery of antibacterial compounds. Nucleic Acids Res., **3**, 731–737
- Evans-Roberts, K.M., Mitchenall, L.A., Wall, M.K., Leroux, J., Mylne, J.S. and Maxwell, A. (2016) DNA gyrase is the target for the quinolone drug ciprofloxacin in Arabidopsis thaliana. J. Biol. Chem., **291**, 3136–3144
- Nadal, M. (2007) Reverse gyrase: an insight into the role of DNA-topoisomerases. Biochimie, **89**, 447–455
- Gadelle, D., Krupovic, M., Raymann, K., Mayer, C. and Forterre, P. (2014) DNA topoisomerase VIII: a novel subfamily of type IIB topoisomerases encoded by free or integrated plasmids in Archaea and Bacteria. Nucleic Acids Res., **42**, 8578–8591

Natassja Bush grew up in South Africa where she undertook her initial BSc and BSc(Hons) degree in Zoology at the University of KwaZulu-Natal studying how seasonal variations and periods of acclimation affect the basal metabolic rate of the African Rock Kestrel. She moved to the UK in 2008 and completed her MSc in Biotechnology at the University of East Anglia, where she developed an interest in topoisomerases. She has worked within the field for 8 years and has just recently completed her PhD investigating the role of DNA gyrase in illegitimate recombination in the Biological Chemistry Department at the John Innes Centre; she is currently a postdoc in the Maxwell lab. Email: tash.bush@jic.ac.uk.

Monica Agarwal gained her PhD in Molecular Biology from Heriot-Watt University in Edinburgh, studying the subcellular dynamics of the Arabidopsis MPK4 signalling pathway. She currently works as a postdoctoral researcher in Tony Maxwell's group in the Department of Biological Chemistry at the John Innes Centre. Her scientific interests focus on understanding and exploiting the role of type II DNA topoisomerases in DNA replication and recombination, and on the development of new targets for new drugs. Email: monica.agarwal@jic.ac.uk.

Sara Henderson is currently working towards her PhD under the supervision of Professor Tony Maxwell in the Department of Biological Chemistry at the John Innes Centre. Her specific scientific interests focus on developing the structural understanding of DNA gyrase from Mycobacterium tuberculosis, *and exploiting it as a target for new antituberculotic agents. E-mail: sara.henderson@jic.ac.uk.*

Nidda Waraich has a BSc degree in Biochemistry from the University of Liverpool. She is currently a PhD student working in the Maxwell lab in the Department of Biological Chemistry at the John Innes Centre (Norwich, UK). Her research efforts are focused upon studying the eukaryotic DNA gyrases from A. thaliana *and* P. falciparum. *Email: nidda.waraich@jic.ac.uk.*



Tony Maxwell is a Project Leader in Biological Chemistry at the John Innes Centre (Norwich, UK) and Honorary Professor in Biological Sciences at the University of East Anglia. He has worked for more than 30 years on DNA topoisomerases and has particular interests in the mechanistic aspects of these enzymes and their exploitation as targets for antimicrobial chemotherapy. He is a co-founder of the spinout company Inspiralis Ltd, and is currently a Wellcome Trust Investigator. Email: tony.maxwell@jic.ac.uk.