Creation Research Society Quarterly 2019 55:196–211.

Unknotting the Nucleus: Regulation and Domain Modularity of Type II Topoisomerases

Joseph E. Deweese* and Salvador Cordova**

Abstract

Routine cellular processes such as transcription, replication, and cell division result in knots, tangles, and torsional stress in DNA. All living organisms produce proteins known as topoisomerases to alleviate these DNA topology challenges, which can lead to cellular dysfunction or death if unresolved. Type II topoisomerases manage DNA topology by generating a transient double-stranded DNA break in one segment of DNA and passing another segment of DNA through the break before resealing the broken DNA. Human type II topoisomerases are well-characterized anticancer drug targets, but there are severe off-target toxicities often associated with some of these drugs. Humans have two versions of topoisomerase II, and it is of clinical interest to selectively target one version of topoisomerase II in humans. Selective targeting requires a thorough understanding of the differences between the two versions, and the evidence presented here explores some of the key pieces of information regarding these differences including genomic, amino acid sequence, modification, and interaction data. We argue that the two versions of topoisomerase II differ in key regions that also are heavily modified via post-translational modifications, which may provide key insights into the regulation and separation of function between the two isoforms. Finally, we suggest that protein domains display modularity that may help us understand the design of these and other proteins by analogy to the idea of a dependency graph.

Introduction

Biochemistry is a fascinating discipline that involves studies at the interface of biology and chemistry through examining the structure and function of the four main classes of fundamental biological molecules: carbohydrates, lipids, nucleic acids, and proteins. In the cellular world, vast arrays of complex biomolecules interact in an elegant and highly sophisticated manner to carry out the various processes that separate living cells from inanimate matter. It is within this submicroscopic realm that protein machines called enzymes operate.

* Joseph E. Deweese (Corresponding author), Lipscomb University College of Pharmacy, Nashville, TN, joe.deweese@lipscomb.edu

** Salvador Cordova, FMS Foundation, Canandaigua, NY

Accepted for publication August 18, 2019

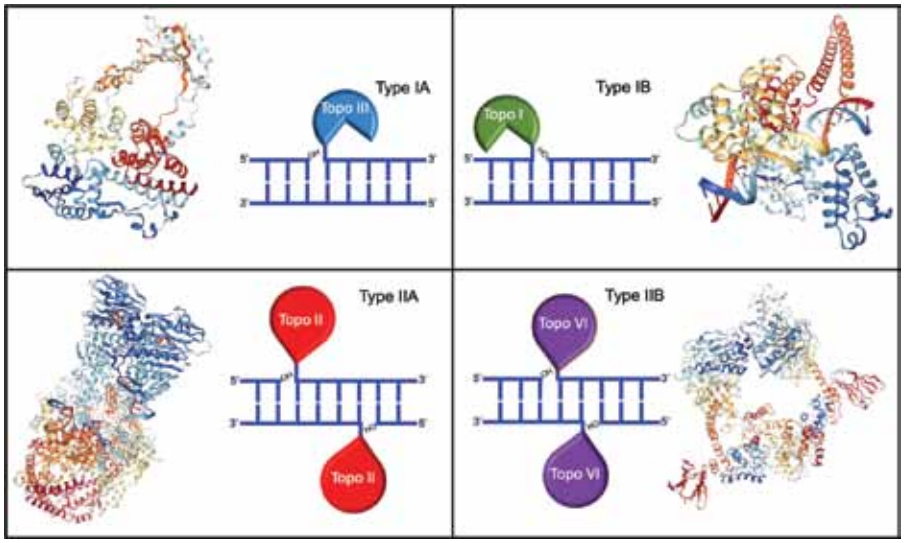


Figure 1. Families of Topoisomerases. Four of the five families of topoisomerases are represented with example structures at left or right. A diagram depicting the specific locations of DNA cleavage by each family and the location of the temporary covalent bond formed (either with the 5' or 3' end of the cleaved DNA) are shown using shapes linked to the DNA. Note: the diagrams do not show the full enzyme, but instead display a shape representation of a part of the enzyme. Structure images were derived from the Protein Databank (RCSB.org) with the structure file noted in parentheses as follows: Type IA: *E. coli* topoisomerase III (1D6M); Type IB: *H. sapiens* topoisomerase I (1A36); Type IIA: *S. cerevisiae* topoisomerase II (4GFH); and Type IIB: *M. mazei* topoisomerase VI (2Q2E).

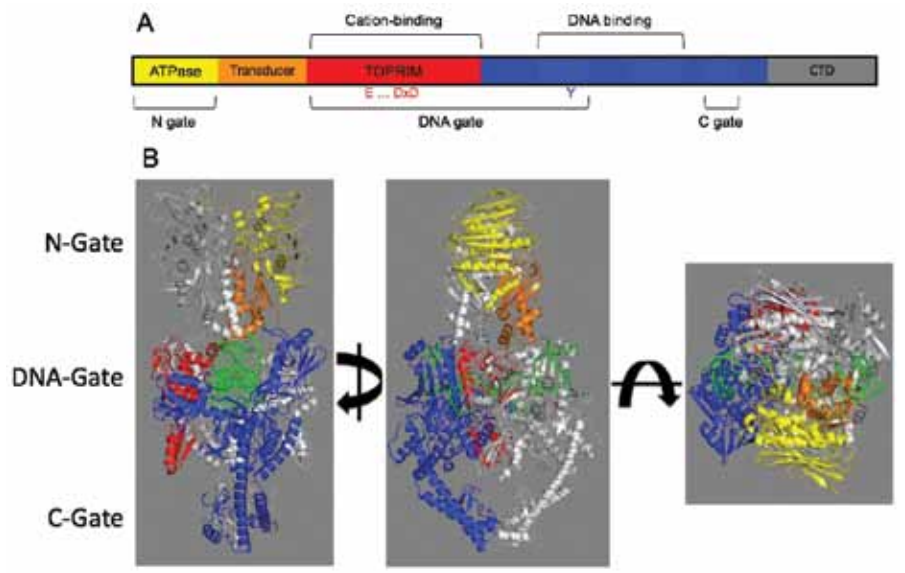


Figure 2. Domain and Biochemical Structure of Type IIA Topoisomerases. A) The domain map of a type IIA topoisomerase is depicted with some key regions identified along with key catalytic (Y) and metal-binding residues denoted (E.... DxD). B) Ribbon diagram based upon the crystal structure of *S. cerevisiae* topoisomerase II (RCSB 4FGH). One protomer is color-coded with the domain map. The other protomer is grey. Structure images were generated using Pymol.

Enzymes are responsible for catalyzing the very reactions that make it possible to harness energy from food, build new biomolecules, copy DNA, and many other routine cellular processes. Due to their critical roles, enzymes are also effective drug targets from a pharmaceutical perspective.

Cancer is one example of a disease where using specific drugs to target enzymes can be very effective. In fact, for several decades, drugs like doxorubicin and etoposide have been used to fight cancer (Murphy et al., 2017). Both of these drugs happen to impact the same enzyme found in the nucleus of cells: DNA topoisomerase II. There are a handful of other drugs that also impact topoisomerase II for the treatment of cancer, and there are additional drugs (e.g., levofloxacin) that target bacterial type II topoisomerases (DNA gyrase and topoisomerase IV) for the treatment of infections (Pommier et al., 2010).

While agents targeting topoisomerase II are effective, they are not without side effects (or adverse events, as they are called in the medical field). Anticancer drugs that target topoisomerases are often associated with some of the well-known side effects of chemotherapy like hair loss, gastrointestinal problems, and others. These effects are common to most drugs that impact enzymes/proteins involved in cell division and/or DNA function. Unfortunately, these effects are not necessarily the most concerning. For example, etoposide treatment is associated with a risk for developing acute myeloid leukemia (Pendleton et al., 2014). In other words, when a patient is treated with etoposide, there is a chance that the drug may actually cause a different cancer (called a secondary or treatment-related disease). Additionally, doxorubicin is a member of a class of drugs known as anthracyclines, which are associated with cardiotoxicity and have a lifetime dose limit maximum to reduce the risk of toxicity (Zhang et al., 2012; McGowan et al., 2017).

Both etoposide and doxorubicin have been extensively studied in the literature, but open questions still remain about the exact mechanism behind these adverse events. Therefore, it is critical to carefully examine the precise mechanisms of drugs to predict and possibly avoid some of these extreme side effects. Our work focuses on type II topoisomerases to explore their function in cells and to explore whether new agents can be developed that help avoid some of these toxicities.

In this manuscript, we will explore the structure and regulation of topoisomerases, and we will provide some reflection on possible novel targeting mechanisms, as well as implications for what these enzymes mean for a design perspective. Using a design-inspired viewpoint, we will examine enzymes as modular proteins and consider implications of a modular perspective on discovery of protein functions and the development of targeted therapeutics.

Topoisomerases: Regulators of DNA Topology

Topoisomerases are a family of enzymes that resolve topological challenges in DNA. These enzymes have been found in all three domains of living organisms and are even found in the genomes of some viruses (McClendon et al., 2006). Topoisomerases regulate DNA topology by generating transient single- or double-stranded DNA breaks. The first topoisomerase to be discovered was initially named the ω protein (Wang, 1971) and later became known as topoisomerase I (Wang et al., 1979). A few years later, DNA gyrase was found (Gellert et al., 1976) and later named topoisomerase II, though gyrase is still the common name (Wang et al., 1979). This naming turned out to be quite fortuitous since type I enzymes generate a single-stranded DNA break while type II enzymes form a double-stranded DNA break (Figure 1).

This has become the naming convention of the now five families of topoisomerases (Table 1). As seen in Table 1, type I and II enzymes are divided into families (IA, IB, etc.). For example, topoisomerase III is a type IA while topoisomerase IV is a type IIA. This classification system accounts for structural and biochemical characteristics. As depicted in Figure 1 and listed in Table 1, several different mechanisms (single- vs. double-strand break, 5' versus 3' linkage, spacing between breaks) are utilized by different families of topoisomerases. While the type I enzymes are a fascinating study, we will be focusing on the type IIA enzymes for the remainder of this review.

Type II Topoisomerases

Type II Topoisomerases are divided into two major families: IIA and IIB (Figure 1, lower portion). While Type IIB topoisomerases are present in some organisms, the Type IIA topoisomerases are apparently more abundant in nature, and the enzymes encoded in the human genome are Type IIA topoisomerases (as discussed below). Type IIA topoisomerases, known generically as topoisomerase II, are dimers or tetramers meaning these enzymes are made from two or four separate protein chains. In the case of Type IIA enzymes in eukaryotes, the enzymes are typically dimers with two copies of the same protein chain coming together to form the functional enzyme. In archaea and bacteria, the enzymes are generally tetramers with an A2:B2 structure meaning two copies of each of two different proteins. For example, DNA gyrase is a type II topoisomerase in bacteria, which is formed by two copies of the protein GyrA and two copies of GyrB.

The organization of the protein domains within topoisomerase II is similar across organisms. As seen in Figure 2A, common protein domains including the ATPase, TOPRIM, and other domains are shown. The organization of prokaryotic enzyme protein domains generally matches that of the eukaryotic proteins, except that the prokaryotic versions have separate protein chains for each half.

X-ray crystallography has enabled the acquisition of high-resolution structures of type II topoisomerases (Dong et al., 2007; Schmidt et al., 2010; Schmidt et al., 2012). This information has provided detail on how these protein domains relate to each other. As seen in Figure 2B, the ATPase domain is at the "top" while the TOPRIM and DNA binding domains are found in the middle. The protein then loops down to a lower domain (C-gate) before looping back up the sides. What is not present in this structure is the C-terminus, which is thought to exist on either side of the enzyme like a pair of "ears" or "wings." However, due to the relatively flexible nature of this domain, the crystal structure of the C-terminus has not yet been resolved.

Topoisomerase II performs a complex catalytic cycle that involves a series of steps and movements (Figure 3). First, topoisomerase II binds to a helix-helix cross-over, which may be formed by separate chromosomes or segments of the same DNA molecule (Zechiedrich et al., 1990; Roca et al., 1993). Second, topoisomerase II bends one of the two DNA segments and creates a temporary double-strand break (Deweese et al., 2009; Schmidt et al., 2010). The broken segment is called the gate segment. During the process of cutting the gate segment of DNA, topoisomerase II forms a covalent bond between an active site tyrosine amino acid and the 5' ends of the broken DNA segment (Liu et al., 1983; Sander et al., 1983; Zechiedrich et al., 1989; Mueller-Planitz et al., 2008; Schmidt et al., 2010). Due to the symmetry of topoisomerase II, this occurs with a four base-pair stagger, as depicted in Figure 1. As the DNA is bent and cleaved, topoisomerase II binds to ATP, which "closes" the top portion of the enzyme, called the N-terminal gate or

Table I. This table depicts families of topoisomerases and representative examples. Type I topoisomerases temporarily cut one strand while type II topoisomerases temporarily cut both strands of the double helix. Domains where specific enzymes are found are denoted: A, archea; B, bacteria; E, eukaryote. Linkage refers to the phosphotyrosine covalent connection between the topoisomerase active site and the cleaved DNA. Activity is classified by the ability to support relaxation (rel) or supercoiling (sup) and the direction of supercoiling is denoted: - for negative supercoils and + for positive supercoils. Question marks represent information that has not been determined or where results may differ between organisms. Representative species are denoted. It should also be noted that there are viral topoisomerases in the IA, IB, and IIA families that are not shown in the table.

Family	Enzyme	Rep. Species	Linkage	Mechanism	Cofactors	Activity
IA	topoisomerase I (B)	*E. coli*	5'	enzyme-bridged strand passage	Mg^{2+}	rel -
	topoisomerase III (α, β) (E)	*H. Sapiens*				
	topoisomerase III (A, B)	*E. coli*				
	reverse gyrase (A)	*Sulf. acidocaldarius*			Mg^{2+}, ATP	sup +
IB	topoisomerase IB (E, B)	*H. sapiens*	3'	controlled rotation/swiveling	none	rel -+
	topoisomerase IB mitochondrial (E)					
IC	topoisomerase V	*M. kandleri*	3'	controlled rotation/swiveling	none	rel -+
IIA	topoisomerase II (E)	*S. cerevisiae*	5'	double-strand passage, 4-bp overhang	Mg^{2+}, ATP	rel -+
	topoisomerase II (α, β) (E)	*H. sapiens*				
	topoisomerase IV (B)	*E. coli*				
	gyrase (A,B)	*E. coli*				sup -
IIB	topoisomerase VI (A, E)	*Sulf. Shibatae*	5'	doubls-strand passage, 2-bp overhang	Mg^{2+}, ATP	rel -+
	topoisomerase VIII (B, A)	*A. degensii*	?	?	Mg^{2+}, ATP(?)	rel -

N-gate (Lindsley et al., 1993; Roca et al., 1993; Classen et al., 2003; Wei et al., 2005; Bendsen et al., 2009; Schmidt et al., 2012). Third, with the N-gate closed, topoisomerase II pulls apart the gate segment, forming the DNA gate. The other segment of DNA, called the transport segment, transverses through the DNA gate (Osheroff, 1986). Fourth, the N-gate undergoes a twisting conformational change that appears to be accompanied by the hydrolysis of one ATP molecule (Roca et al., 1992; Lindsley et al., 1993; Harkins et al., 1998; Baird et al., 1999; Schmidt et al., 2012), which may prevent the transport segment from back-tracking to the N-gate (Schmidt et al., 2012). Fifth, the DNA gate can now close and allow the opening of the "bottom" of the enzyme, called the C-terminal gate or C-gate (Osheroff et al., 1987; Robinson et al., 1991; Bromberg et al., 2002). This occurs along with the hydrolysis of the second molecule of ATP (Roca et al., 1992; Lindsley et al., 1993; Harkins et al., 1998). Closing of the DNA gate also allows for the

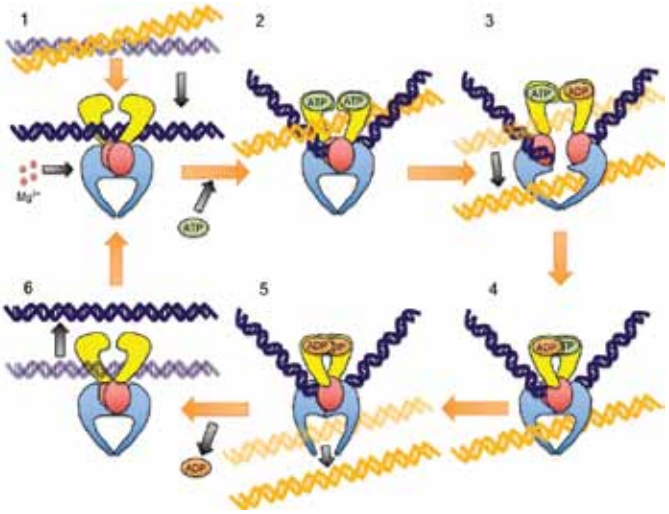
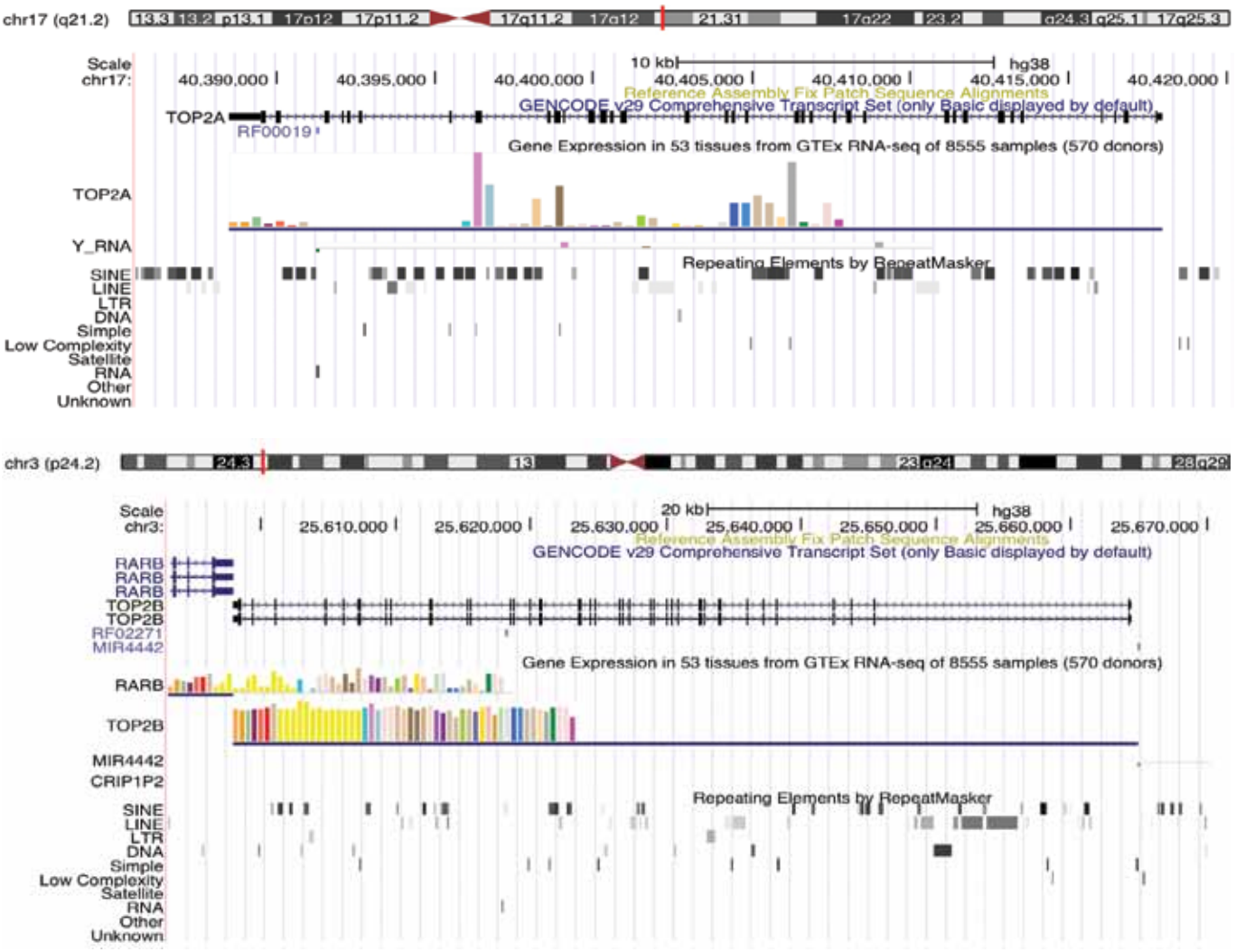


Figure 3. Catalytic Cycle of Type IIA Topoisomerases. 1) The dimeric enzyme binds a helix-helix crossover (one segment in gold, the other in purple). 2) The gate segment of DNA is bent and cleaved. The transport segment is captured by the N-terminal clamp in the presence of ATP. 3) The gate segment ends are separated and the transport segment can cross to the lower portion of the enzyme. This is accompanied by the hydrolysis of one ATP molecule to ADP. 4) The gate segment is ligated and the N-terminal portion of the enzyme rotates. 5) The C-gate opens and releases the transport segment. The second molecule of ATP is hydrolyzed. 6) The C-gate closes and the N-gate opens allowing release of the gate segment of DNA. The enzyme is now reset for another round of catalysis.

Figure 4. Gene Browser Features of the human TOP2A Gene Locus. Data generated online at http://genome.ucsc.edu/ (Kent, et al., 2002).

Figure 5. Gene Browser Features of the human TOP2B Gene Locus. Data generated online at http://genome.ucsc.edu/ (Kent, et al., 2002).



gate segment to be ligated (i.e., original bonds reformed and no longer bound to the enzyme). Once the transport segment is released, topoisomerase II can now close the C-gate, release the ADP, and open the N-gate so the gate segment can be released.

This fascinating process involves several highly coordinated and orchestrated events. For example, the ability to cut DNA requires the enzyme to be in a conformation where the DNA gate is closed. This is because each active site requires amino acids from separate halves of the enzyme to be complete. As discussed earlier, topoisomerase II in humans and other eukaryotes is a homodimer—meaning two copies of the same protein. The TOPRIM domain, which coordinates two Mg^{2+} ions from each protein chain interacts with the active site tyrosine from the opposite protein chain to form a functional catalytic site. While the DNA gate is open, or if the protein chains are separated, DNA cleavage and ligation cannot take place. This design feature acts as a fail-safe mechanism that prevents release of the DNA ends from the protein before they are re-attached.

Exploring Two Isoforms of Topoisomerase II

Type II topoisomerases appear to be required for life in living organisms. Even "simple" life forms such as yeast (e.g., *Saccharomyces cerevisiae*) or bacteria (*e.g., Escherichia coli*) encode at least one type II topoisomerase with many organisms encoding two separate enzymes. Humans encode two separate type IIA topoisomerases: topoisomerase IIα and topoisomerase IIβ. These two enzymes are encoded by separate genes on separate chromosomes and controlled by distinct promoters (Nitiss, 2009; Pommier et al., 2016). While topoisomerase IIβ is widely expressed across a range of tissues (Consortium, 2013), topoisomerase IIα expression appears to follow progression of the cell cycle with an increase in S-phase and mitosis (Nitiss, 2009). Based upon cellular and biochemical evidence, topoisomerase IIα is more involved in DNA replication and chromosome segregation, which fits with the observed expression profile (Nitiss, 2009). Topoisomerase IIβ is more involved in regulation of topology during transcription and for chromatin remodeling (Nitiss, 2009). Recent evidence also indicates a role for topoisomerase IIβ in domain looping (Uuskula-Reimand et al., 2016).

In mammals, loss of topoisomerase IIβ at the cellular level can be tolerated (Errington et al., 1999). However, the situation appears to be different at the organismal level. Mice lacking topoisomerase IIβ die of neural developmental failure (Yang et al., 2000). More recent experiments suggest that this is because of a critical role for topoisomerase IIβ in regulation of chromatin topology during the process of chromatin remodeling where regions of DNA are "turned on or off" through movement of regulatory components in the DNA. For example, topoisomerase II-mediated DNA strand breaks have been shown in specific cases to be an integral part of regulating gene expression (Bunch et al., 2015; Madabhushi et al., 2015). In contrast, topoisomerase IIα is absolutely required at the cellular and organismal levels (Nitiss, 2009). Loss of topoisomerase IIα leads to problems with chromosome segregation and cell division (Grue et al., 1998; Akimitsu et al., 2003; Carpenter et al., 2004). During replication, newly synthesized chromosomes, referred to as sister chromatids, become intertwined or catenated. The interlinking of these sister chromatids must be resolved prior to mitosis in order for chromosome segregation to take place properly. Topoisomerase IIα is required to carry out this critical function in cells, and topoisomerase IIβ is unable to fully complement for the loss of the other isoform. Similarly, topoisomerase IIα is unable to fully complement for the loss of topoisomerase IIβ.

What is the reason behind this difference? Why could one isoform not complement for the other? Is it possible that by examining the differences between the proteins we might be able to identify components that are critical for isoform-specific roles in cells? These questions have health relevance as well. Human type II topoisomerases are also cancer drug targets, but the clinically used cancer drugs cannot differentiate between either of the isoforms. As mentioned earlier, some of these drugs have severe risks (i.e., secondary leukemia for etoposide and cardiotoxicity for the anthracyclines) associated with their use, and there is evidence that these adverse events may involve topoisomerase IIβ. Interestingly, this is the version of the enzyme that is present in all tissues, while growing tissues generally express topoisomerase IIα (Drake et al., 1989; Capranico et al., 1992). Therefore, it is of clinical and pharmaceutical interest to be able to develop agents that can selectively target topoisomerase IIα (and avoid topoisomerase IIβ).

The Tale of the Tails

Topoisomerase IIα and IIβ share approximately 65% amino acid identity, though they differ in length by over 90 amino acids (β is longer). Most of the shared amino acid sequence is found in the core domains: ATPase, TOPRIM, DNA binding/cleavage, and C-gate. In these regions, identity rises to >80% (similarity above 90% in some regions). Both the extreme N-terminus and the C-terminus are distinct between the isoforms. For example, the C-terminal domain is very different with only 30% amino acid identity. In evolutionary terms, this portion is considered "variable" compared to the rest of the protein, but it makes up around 25–30% of the protein at roughly 400 amino acids.

Evolutionists believe topoisomerases I and II evolved independently (convergent evolution) multiple times (Forterre et al., 2007; Forterre et al., 2009). In

the specific case of topoisomerase IIα and IIβ, the evolutionary assumption is that these two isoforms originated from a common ancestral topoisomerase II via duplication followed by subsequent divergence (Lang et al., 1998; Sng et al., 1999). The timing of this event is thought to be around the time of the origin of vertebrates since these two isoforms appear to be common to vertebrates, but not typically found in other life forms (Lang et al., 1998). However, the evolutionary mechanism for the divergence of the C-termini is unknown.

Along with collaborators, we found that when comparing the two human isoforms and assuming a common ancestor for them, their sequences appear to be under purifying selection for functional constraint (Deweese et al., 2019). So, it appears that these regions in the two isoforms are different and are under selective pressure to remain that way. From a design perspective, we would argue that this is evidence that these proteins were designed separately (using common modules) to fulfill different functions.

The C-terminus is the least well-understood portion of the enzyme. As noted earlier, no structural data is available for this region in the eukaryotic enzymes, and early research in this area suggested that cells could survive the loss of large segments of this region. However, a more thorough examination is needed.

There are several lines of evidence for complex regulatory roles for the C-terminus in these enzymes. First of all, the evidence from previous studies, mentioned above, regarding the inability of either isoform to fully complement loss of the other even though the catalytic portion of the enzyme is essentially the same. Second, there are a large number of post-translational modifications in this region including phosphorylation, sumoylation, and others (see discussion below). Third, there is evidence for isoform-specific protein-protein interactions with the C-terminus, discussed below (Deweese et al., 2019). Fourth, exchanging this region between the two proteins leads to a change in biochemical and cellular function (Linka et al., 2007). Thus, given our understanding of the similarities and differences between topoisomerase IIα and IIβ, we suggest that the C-terminus is a prime candidate for a complex regulatory role. In other words, why does topoisomerase IIβ not complement for the loss of topoisomerase IIα? There must be something that one version has that the other is missing, and most of this difference (with some exception at the extreme N-terminus) is found in the C-terminus. Note that the N-terminus also likely plays an important role, but it is only 20–40 amino acids. Therefore, our current focus is on the larger C-terminal domain.

In the following sections, we will explore several pieces of data related to the C-terminus that help provide some insight into the differences between the two isoforms. It has been hypothesized that by examining the C-terminus, we may be able to develop more selective anticancer agents that may reduce adverse events (Murphy et al., 2017).

Surveying Differences between Topoisomerase IIα and IIβ

Genomic Differences

In terms of genomic context, the genes topoisomerase IIα (*TOP2A*) and IIβ (*TOP2B*) are found on different chromosomes. Using the UCSC Genome Browser (http://genome.ucsc.edu/), we can gain a better look at the genetic context and components of these genes (Kent et al., 2002). *TOP2A* is on the long arm of Chromosome 17 at chromosome band 17q21.2 (Figure 4), while *TOP2B* is on the short arm of chromosome 3 at chromosome band 3p24.2 (Figure 5). Their exon structures are very similar, but they vary in the intronic regions with a higher density of SINES (Short Interspersed Nuclear Elements) found in the intronic regions of *TOP2A*. According to the original mapping reports, several exons in the N- and C-terminus differ in their start/end positions (Lang et al., 1998; Sng et al., 1999). In addition, several LINEs (Long Interspersed Nuclear Elements) are found in *TOP2B* that are not found in the other isoform. Interestingly, the genes for both isoforms contain at least one non-coding RNA gene within an intron. The non-protein coding genes are different between the two genes and found inside different introns.

Based upon the gene expression data, clearly *TOP2B* is much more broadly expressed across a range of tissues, while *TOP2A* is much more selectively expressed. Interestingly, 5' of both isoforms are retinoic acid receptor genes. Retinoic acid receptor alpha (*RARA*) is 5' of *TOP2A* while *RARB* is 5' of *TOP2B* (Figure 5).

In reviewing the sequence alignments, it is clear that some regions of the protein are largely shared, as discussed above. However, the extreme N-terminus and the C-terminus are very different between these two isoforms (Figure 6). Several examples of regions of the sequence alignments are shown including snapshots within the largely shared core catalytic regions and in the N- and C-terminus where these two proteins differ so significantly.

Post-Translational Modifications

To examine the post-translational modification (PTM) data, we used the online databases Phosphosite Plus (Hornbeck et al., 2015) and UniProt. Phosphosite Plus enables researchers to examine locations and types of modifications along with the literature and proteomic evidence for those modifications. As seen in Figure 7, both isoforms are known to have many putative sites for PTMs—primarily phosphorylation and ubiquitination. However, much of the evidence in

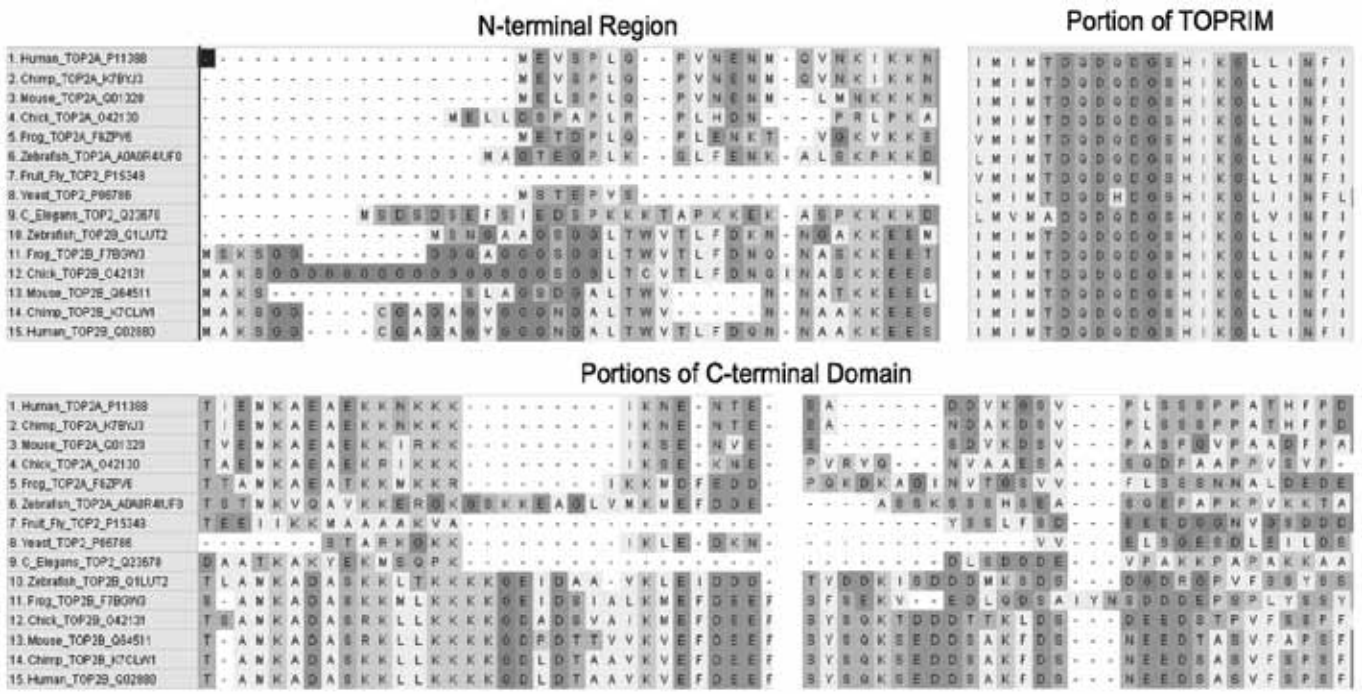


Figure 6. Segments from Sequence Alignments of Type IIA Topoisomerases. Sequences from 15 type IIA topoisomerases are shown. Sequences were retrieved from Uniprot. Selected example regions are shown including the extreme N-terminus, a portion of the TOPRIM domain (metal-ion binding region), and two portions of the C-terminal domain are shown. Alignments were generated using MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets (Kumar, Stecher, and Tamura 2015).

the database for these modifications is from high-throughput proteomics studies. Thus, there have been relatively few research papers published on roles and significance of specific modifications. For example, there is clear evidence for regulatory roles in promoting activity of topoisomerase IIα for phosphorylation at Ser-1213 (Wells et al., 1995; Ishida et al., 2001). However, most sites do not have such detailed characterization.

It should be noted that both isoforms have similar numbers of putative modifications, but interestingly, around half of those modifications (91/191 for Top2A) are found in the C-terminal domain (Deweese et al., 2019). Also, when examining both isoforms, a number of the modification sites are not shared between isoforms—especially in the C-terminus (Deweese et al., 2019). Thus, there are unique modification profiles for both enzymes that may help explain the separate functional regulation. Currently, we are continuing to examine the database evidence for these modifications to characterize the differing profiles between the isoforms.

Protein-Protein Interactions

Protein-protein interactions are also a major factor in the regulation of function in cells. The cellular environment is a very busy, crowded place. The ability of proteins to interact with specificity and cause changes in activity in other proteins is still an amazing thing to consider. These interactions require proteins to have specific sequences, to fold in complex three-dimensional shapes, and to be able to interact in the chemical and pH environment where they operate.

Previously, we showed using The BioGrid database (Oughtred et al., 2019) that topoisomerase IIα and IIβ collectively interact with ~143 proteins and about 34 of these interact with both isoforms (Deweese et al., 2019). More recently, this number has changed to ~153 proteins with only 32 of these showing evidence for interaction with both isoforms. Figures 8 and 9 depict interaction networks generated by The BioGrid database. It should be noted again that this data is primarily from high-throughput studies rather than functional studies. So, little is known about the significance (or lack thereof) for most of these interactions. However, these networks provide a starting point for us in our ongoing analysis of the interactions and modifications of the C-terminal domain of topoisomerase II.

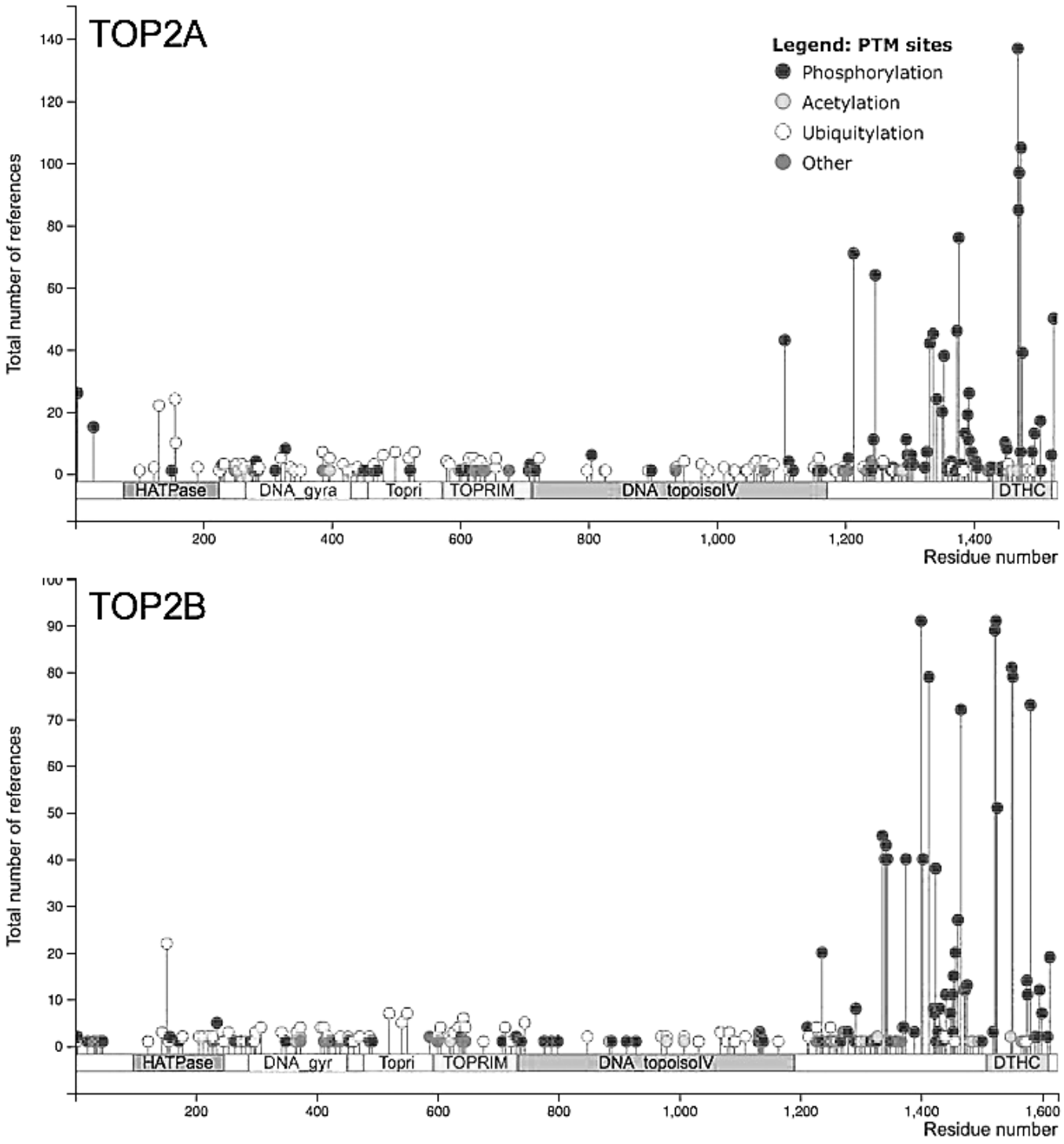


Figure 7. Post-Translational Modifications of Top2A and Top2B from Phosphosite Plus. Modification maps were generated by Phosphosite Plus (phosphosite.org) for *H. sapiens* Top2A (https://www.phosphosite.org/proteinAction.action?id=2303) and Top2B (https://www.phosphosite.org/proteinAction.action?id=5866). All sites with at least one reference are shown mapped onto the domain structures of the enzymes.

Importantly, it is clear that both isoforms interact with dozens of proteins—with less than half of those proteins shared between the isoforms.

Summary

As discussed above, topoisomerase IIα and IIβ in humans represent two versions of an enzyme that have distinct roles in cells. Even though these enzymes share a catalytic mechanism, they are regulated independently in cells and cannot compensate for the loss of the other isoform. We hypothesize that much of this regulatory function is embedded within the C-terminus of the protein with some regulation also occurring at the N-terminus and possibly at regions within the protein. This model for isoform specific functional regulation of topoisomerase II is not likely unique to this family of proteins, but instead it represents a broader model for protein regulation.

This brings up another important concept: modular protein domains. Several of the domains within topoisomerase II are shared between the isoforms but the regulatory domains appear to be distinct. Could it be that the domains of proteins serve as modules in the original design and thus we see shared protein domains across diverse organisms and protein types? It is also possible that the regulatory domains have taxon specific differences that enable these proteins to operate in the specific cellular context of a given species or genus, for example. This may help explain why some proteins share large portions of information across diverse organisms but also have some regions that are highly specific for a given subset of organisms.

Protein Domains as Modules

Ewert (2018) has argued that protein families can be analyzed statistically using the traditional tree model or using a dependency graph model. In his analysis, he examined the presence or absence of protein families across a wide range of organisms. His work found that the pattern of the presence and absence of protein families among organisms fit a dependency graph model better than a ancestry tree model (Ewert, 2018). This result was true for each of the protein databases examined. While this does not prove the protein families are sorted in a modular fashion, it is supportive of the possibility. In other words, imagine a designer providing specific families of proteins based upon the needs of specific organisms.

We would like to suggest for it to be considered that this may not be limited to the level of protein families. Many proteins have two or more domains that represent smaller sections of the whole protein but may have a discrete function (e.g., the GHKL ATPase domain of topoisomerase II). What if protein domains are sorted in a modular fashion? How could we know if that was the case? How could we test it?

Interestingly, there is a concept in evolutionary biology called "promiscuous domains" (Basu et al., 2008). These protein domains "show a tendency to occur in diverse domain architectures" (Basu et al., 2008). According to the cited study, these domains seem to appear and come and go along the supposed evolutionary spectrum in a manner not consistent with simple common descent. We suggest that this pattern may fit a "modular" concept better than a tree-like common descent concept, but there is much work that remains to be done in this area.

Further, it is becoming apparent that proteins may also have organism- and species-specific differences that may be finely tuned for a given organism. In other words, there may be high levels of constraint on protein sequences that are organism-specific, which may preclude the ability of these sequences to evolve. This is an interesting area of investigation, and we hope that other scientists will carefully consider that subtle protein differences among organisms may have implications at the organism level.

Discussion: Implications and Open Questions

Topoisomerases are fascinating molecular machines responsible for relieving "topological entanglements"—knots and tangles—in the DNA. These enzymes operate by transiently breaking one or both strands of DNA. As drug targets, these enzymes have been exploited in the treatment of cancer. The type II topoisomerases, human topoisomerase IIα and IIβ, are very similar throughout their catalytic regions but differ widely at the N- and C-termini. These regions likely serve regulatory roles for these proteins, and we hypothesize that selective targeting of the C-terminus may enable the generation of selective topoisomerase II drugs that may reduce adverse events in patients.

As mentioned above, topoisomerases are essential to all known forms of life. These enzymes are needed in the simplest of organisms to facilitate the maintenance of DNA. Thus, even the smallest known genomes require topoisomerase activity to facilitate maintenance. The Neo-Darwinian paradigm requires multiple independent yet highly convergent origins for these proteins both among and within the individual topoisomerase families. In other words, the Neo-Darwinian paradigm would require type IIA topoisomerases to originate more than once independently—yet converging on common modules and domains that are widespread (e.g., TOPRIM, GHKL ATPase, etc.).

In fact, evolutionists in the topoisomerase field have acknowledged that topoisomerases present a difficulty for evolutionary explanations:

> The phylogenetic distribution of DNA topoisomerases is thus *quite puzzling and clearly does not agree with the classical universal tree of*

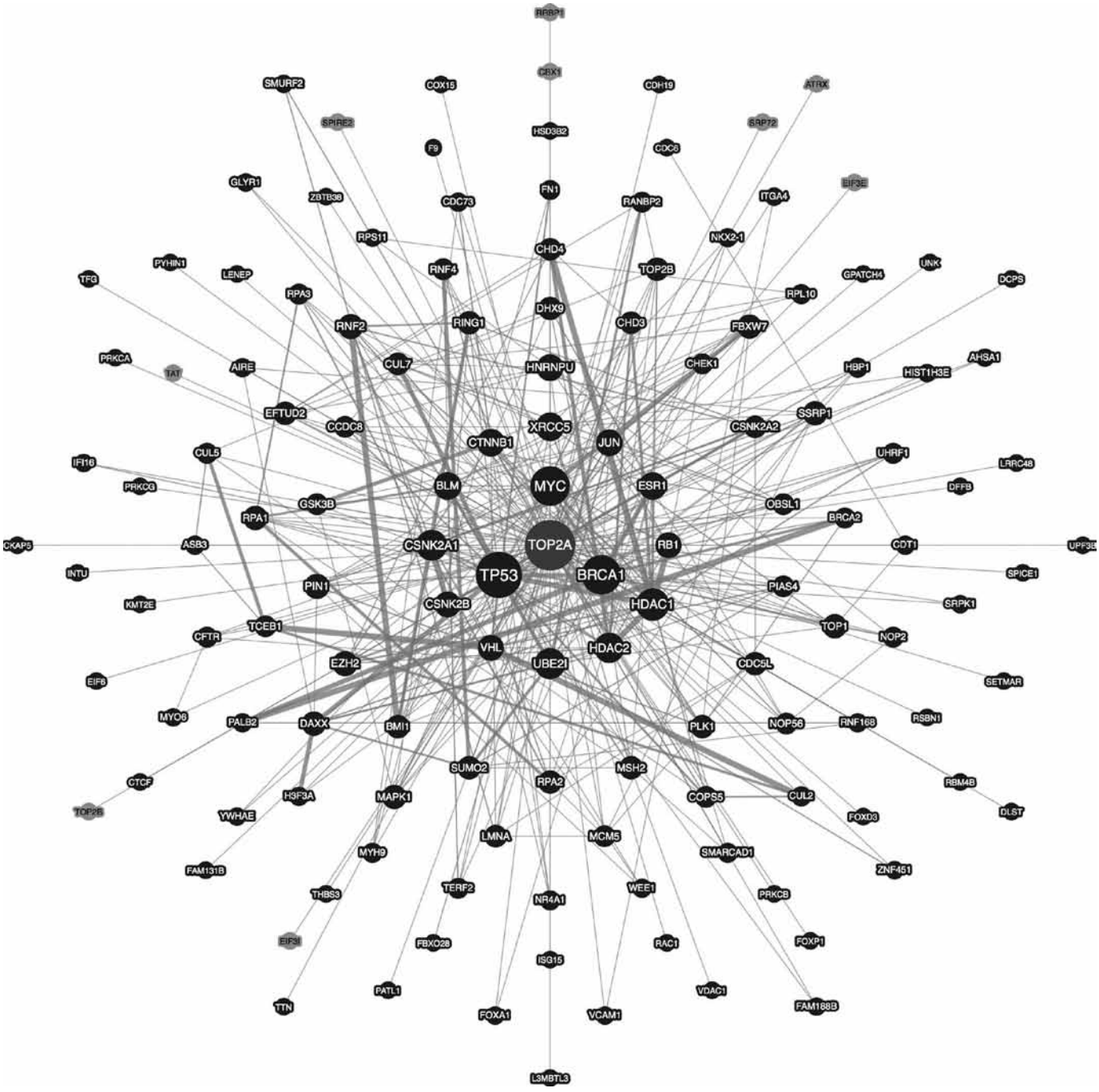


Figure 8. Protein-Protein Interaction Network for Top2A. The BioGrid (BioGrid.org) database was used to generate an interaction map for Top2A (https://thebiogrid.org/113006/summary/homo-sapiens/top2a.html). Partners are shown with those having the most evidence being depicted as closer to Top2A in the center.

life: neither with any of the alternative models such as the bacteria-first model nor with the ring of life model. Another puzzling problem is the phylogenetic position of viral DNA topoisomerases. Whereas in some cases their placement can be explained by HGT [Horizontal Gene Transfer] from a cellular host (e.g. mimivirus Topo IA and IB), in other cases (e.g. Poxvirus Topo IB, T4 Topo IIA) the viral DNA topoisom-

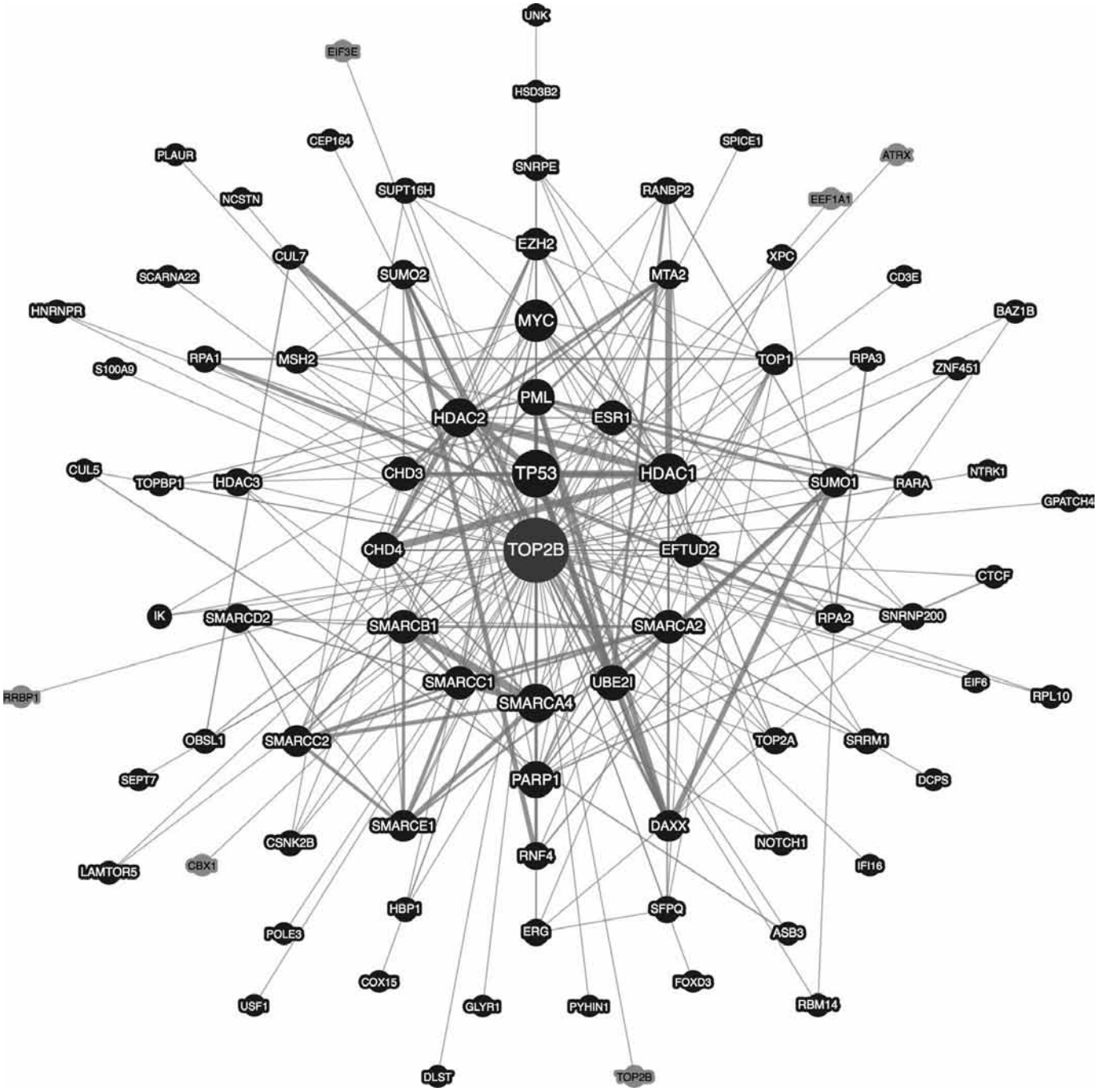


Figure 9. Protein-Protein Interaction Network for Top2B. The BioGrid (BioGrid.org) database was used to generate an interaction map for Top2B (https://thebiogrid.org/113008/summary/homo-sapiens/top2b.html). Partners are shown with those having the most evidence being depicted as closer to Top2B in the center.

erases form well-defined subfamilies that are only distantly related to their cellular counterparts. (emp. added) (Forterre et al., 2007)

Ironically, some evolutionists see the large number of diverse topoisomerases as counterintuitive. For example, in a paper on the phylogenomics and origin of topoisomerases, the authors muse: "An intelligent designer would have probably invented only one ubiquitous Topo I and one ubiquitous Topo II to

facilitate the task of future biochemists. The reality turned out to be quite different, and more interesting."(Forterre et al., 2009) Aside from the obvious note that these authors have no real clue what "an intelligent designer would have" probably done, their presumption assumes a very reductionist perspective on living systems.

From a design perspective, the diversity of topoisomerase forms makes sense in light of seemingly ever-increasing evidence for specialization of these enzymes. As discussed above, specialization is evident with topoisomerase IIα and IIβ in the fact that each enzyme performs some distinct nuclear functions and cannot fully complement for the loss of the other. This specialization has been acknowledged by others in the topoisomerase field. For example, leading topoisomerase researchers speculate why there are so many topoisomerases:

> "We are now beginning to answer that question, and the picture that has emerged is that each of these enzymes has a set of specific functions and that their specialization allows for precise coordination, especially in complicated DNA transactions that are required for replication, transcription and chromosome segregation. Our emerging understanding of chromosome nuclear territories, DNA repair, chromatin-remodeling complexes and super-enhancers, in addition to replication and transcription factories and chromatin looping, underscores the importance of understanding the six vertebrate DNA (and RNA) topoisomerases and integrating that knowledge into our study of genome function." (Pommier et al., 2016)

Indeed, the topoisomerase field is just beginning to grasp the depth of complexity of these amazing molecular machines. Having six topoisomerases in vertebrates may not seem logical to some reductionists. Using a design perspective, we see living organisms as engineered systems designed to live and function in diverse environments and contexts. As our understanding of the complexities of the genome increases, the need for more complex regulatory mechanisms becomes apparent. These mechanisms clearly extend to the control of fundamental enzymes involved in DNA metabolism including topoisomerases and a number of other critical enzymes.

It is hard to imagine any one of the topoisomerases developing via an unguided process once—let alone multiple times independently—because they are essential to life, and thus present chicken-egg paradoxes. Evolving topoisomerases incrementally via natural selection would face severe challenges. For example, a partially functional topoisomerase that cuts but cannot ligate (reconnect) the DNA ends would result in permanent strand breaks. If the protein chains of topoisomerase II cannot form homodimers (eukaryotes) or heterotetramers (bacteria), then the enzyme would fail to perform its function since the active site of topoisomerase II requires amino acids from two separate protein chains to function in precise spatial proximity.

Aside from these fundamental issues, there are many other questions that could be raised about the supposed evolution of topoisomerases. What about the communication taking place within the topoisomerase II domains during the stages of the catalytic cycle? What about the complex strand passage mechanism and the coupling of ATP hydrolysis? How could that system develop incrementally? A partially functional topoisomerase II does not solve the major challenges. An objection may be raised by specialists in the field that an enzyme called Spo11 exists, which is able to generate strand breaks (Romanienko et al., 2000). In recent years, this protein has been shown to function in meiotic recombination and has similarity in structure to the subunit A of topoisomerase VI. It actually forms a heterotetramer with a second protein, Top6BL (Robert et al., 2016; Vrielynck et al., 2016). This highly specialized role for Spo11/Top6BL highlights the increasing specificity of function among these enzymes. Thus, Spo11 does not answer any of the above questions since it also serves a critical, highly-specialized cellular role.

These questions bring up a final thought: how do topoisomerases get to where they are needed and "know" when and where to act? This is a complex question and one that will likely be solved in the next few years. We hypothesize that localization is a function of the C-terminus of the protein and likely involves protein-protein interactions, post-translational modifications, and chromatin-topoisomerase interactions.

Conclusion

From a design perspective, these enzymes have been present since the beginning and are uniquely suited to increasingly specialized roles. The more we learn of the localization, regulation, and specialization of topoisomerases, the more complex and nuanced the picture becomes. Thus, we argue that these amazing "molecular scissors" are precisely what Paul spoke of in Romans 1:20—the "things that have been made" that enable us to see His everlasting power and divinity.

Acknowledgements

The authors would like to thank several research students in the Deweese lab who have assisted in various stages of this project. Their names have been omitted from the submission. We would like to thank Dr. Jeff Tomkins for helpful discussions regarding the UCSC Gene Browser and the structure of the TOP2A and TOP2B genes. We thank Drs. Jean Lightner and Kevin Anderson for helpful feedback on an early version of the

manuscript. Also, we would like to thank the Creation Research Society, the FMS Foundation, and Lipscomb University College of Pharmacy and Health Sciences for funding.

References

Akimitsu, N., N. Adachi, H. Hirai, M.S. Hossain, H. Hamamoto, M. Kobayashi, Y. Aratani, H. Koyama, and K. Sekimizu. 2003. Enforced cytokinesis without complete nuclear division in embryonic cells depleting the activity of DNA topoisomerase IIalpha. *Genes to Cells* 8(4):393–402.

Baird, C.L., T.T. Harkins, S.K. Morris, and J.E. Lindsley. 1999. Topoisomerase II drives DNA transport by hydrolyzing one ATP. *Proc. Natl. Acad. Sci. U.S.A.* 96(24):13685–13690.

Basu, M.K., L. Carmel, I.B. Rogozin, and E.V. Koonin. 2008. Evolution of protein domain promiscuity in eukaryotes. *Genome Research* 18(3):449–461.

Bendsen, S., V.H. Oestergaard, C. Skouboe, M. Brinch, B.R. Knudsen, and A.H. Andersen. 2009. The QTK loop is essential for the communication between the N-terminal atpase domain and the central cleavage—ligation region in human topoisomerase IIalpha. *Biochemistry* 48(27):6508–6515.

Bromberg, K.D., C. Hendricks, A.B. Burgin, and N. Osheroff. 2002. Human topoisomerase IIα possesses an intrinsic nucleic acid specificity for DNA ligation. Use of 5' covalently activated oligonucleotide substrates to study enzyme mechanism. *Journal of Biological Chemistry* 277(34):31201–31206.

Bunch, H., B.P. Lawney, Y.-F. Lin, A. Asaithamby, A. Murshid, Y.E. Wang, B.P.C. Chen, and S.K. Calderwood. 2015. Transcriptional elongation requires DNA break-induced signalling. *Nature Communications* 6:10191.

Capranico, G., S. Tinelli, C.A. Austin, M.L. Fisher, and F. Zunino. 1992. Different patterns of gene expression of topoisomerase II isoforms in differentiated tissues during murine development. *Biochimica et Biophysica Acta* 1132(1):43–48.

Carpenter, A.J., and A.C. Porter. 2004. Construction, characterization, and complementation of a conditional-lethal DNA topoisomerase IIalpha mutant human cell line. *Molecular Biology of the Cell* 15(12):5700–5711.

Classen, S., S. Olland, and J.M. Berger. 2003. Structure of the topoisomerase II ATPase region and its mechanism of inhibition by the chemotherapeutic agent ICRF-187. *Proc. Natl. Acad. Sci. USA* 100(19):10629–10634.

Consortium, G.T. 2013. The Genotype-Tissue Expression (GTEx) project. *Nature Genetics* 45(6):580–585.

Deweese, J.E., K.G. Hoang, R.A. Menzie, C.A. Fief, C.A. Ayes, J.M. Keck, J.T. Wilson, S. Cordova, and C.W. Nelson. 2019. The variable C-terminal domain of human type II topoisomerases as a functionally relevant therapeutic target. *The FASEB Journal* 33(1_supplement):793.794–793.794.

Deweese, J.E., and N. Osheroff. 2009. Coordinating the two protomer active sites of human topoisomerase II: nicks as topoisomerase II poisons. *Biochemistry* 48(7):1439–1441.

Dong, K.C., and J.M. Berger. 2007. Structural basis for gate-DNA recognition and bending by type IIA topoisomerases. *Nature* 450(7173):1201–1205.

Drake, F.H., G.A. Hofmann, H.F. Bartus, M.R. Mattern, S.T. Crooke, and C.K. Mirabelli. 1989. Biochemical and pharmacological properties of p170 and p180 forms of topoisomerase II. *Biochemistry* 28(20):8154–8160.

Errington, F., E. Willmore, M.J. Tilby, L. Li, G. Li, W. Li, B.C. Baguley, and C.A. Austin. 1999. Murine transgenic cells lacking DNA topoisomerase IIbeta are resistant to acridines and mitoxantrone: analysis of cytotoxicity and cleavable complex formation. *Molecular Pharmacology* 56(6):1309–1316.

Ewert, W. 2018. The dependency graph of life. *BIO-Complexity* 3:1–27.

Forterre, P., and D. Gadelle. 2009. Phylogenomics of DNA topoisomerases: their origin and putative roles in the emergence of modern organisms. *Nucleic Acids Research* 37(3):679–692.

Forterre, P., S. Gribaldo, D. Gadelle, and M.C. Serre. 2007. Origin and evolution of DNA topoisomerases. *Biochimie* 89(4):427–446.

Gellert, M., K. Mizuuchi, M.H. O'Dea, and H.A. Nash. 1976. DNA gyrase: an enzyme that introduces superhelical turns into DNA. *Proc. Natl. Acad. Sci. USA* 73:3872–3876.

Grue, P., A. Grasser, M. Sehested, P.B. Jensen, A. Uhse, T. Straub, W. Ness, and F. Boege. 1998. Essential mitotic functions of DNA topoisomerase IIα are not adopted by topoisomerase IIβ in human H69 cells. *Journal of Biological Chemistry* 273(50):33660–33666.

Harkins, T.T., T.J. Lewis, and J.E. Lindsley. 1998. Pre-steady-state analysis of ATP hydrolysis by *Saccharomyces cerevisiae* DNA topoisomerase II. 2. Kinetic mechanism for the sequential hydrolysis of two ATP. *Biochemistry* 37:7299–7312.

Hornbeck, P.V., B. Zhang, B. Murray, J.M. Kornhauser, V. Latham, and E. Skrzypek. 2015. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Research* 43(Database issue):D512–520.

Ishida, R., R. Takashima, T. Koujin, M. Shibata, N. Nozaki, M. Seto, H. Mori, T. Haraguchi, and Y. Hiraoka. 2001. Mitotic specific phosphorylation of serine-1212 in human DNA topoisomerase IIalpha. *Cell Structure and Function* 26(4):215–226.

Kent, W.J., C.W. Sugnet, T.S. Furey, K.M. Roskin, T.H. Pringle, A.M. Zahler, and D. Haussler. 2002. The Human Genome Browser at UCSC. *Genome Research* 12(6):996–1006.

Lang, A.J., S.E. Mirski, H.J. Cummings, Q. Yu, J.H. Gerlach, and S.P. Cole. 1998. Structural organization of the human TOP2A and TOP2B genes. *Gene* 221(2):255–266.

Lindsley, J.E., and J.C. Wang. 1993. On the coupling between ATP usage and DNA transport by yeast DNA topoisom-

erase II. *Journal of Biological Chemistry* 268(11):8096–8104.

Linka, R.M., A.C. Porter, A. Volkov, C. Mielke, F. Boege, and M.O. Christensen. 2007. C-terminal regions of topoisomerase IIα and IIβ determine isoform-specific functioning of the enzymes in vivo. *Nucleic Acids Research* 35(11):3810–3822.

Liu, L.F., T.C. Rowe, L. Yang, K.M. Tewey, and G.L. Chen. 1983. Cleavage of DNA by mammalian DNA topoisomerase II. *Journal of Biological Chemistry* 258:15365–15370.

Madabhushi, R., F. Gao, A.R. Pfenning, L. Pan, S. Yamakawa, J. Seo, R. Rueda, T.X. Phan, H. Yamakawa, P.C. Pao, R.T. Stott, E. Gjoneska, A. Nott, S. Cho, M. Kellis, and L.H. Tsai. 2015. Activity-Induced DNA Breaks Govern the Expression of Neuronal Early-Response Genes. *Cell* 161(7):1592–1605.

McClendon, A.K., J.S. Dickey, and N. Osheroff. 2006. Ability of viral topoisomerase II to discern the handedness of supercoiled DNA: bimodal recognition of DNA geometry by type II enzymes. *Biochemistry* 45(38):11674–11680.

McGowan, J.V., R. Chung, A. Maulik, I. Piotrowska, J.M. Walker, and D.M. Yellon. 2017. Anthracycline Chemotherapy and Cardiotoxicity. *Cardiovascular Drugs and Therapy* 31(1):63–75.

Mueller-Planitz, F., and D. Herschlag. 2008. Coupling between ATP binding and DNA cleavage by DNA topoisomerase II: A unifying kinetic and structural mechanism. *Journal of Biological Chemistry* 283(25):17463–17476.

Nitiss, J.L. 2009. DNA topoisomerase II and its growing repertoire of biological functions. *Nature Reviews: Cancer* 9(5):327–337.

Osheroff, N. 1986. Eukaryotic topoisomerase II. Characterization of enzyme turnover. *Journal of Biological Chemistry* 261:9944–9950.

Osheroff, N., and E.L. Zechiedrich. 1987. Calcium-promoted DNA cleavage by eukaryotic topoisomerase II: trapping the covalent enzyme-DNA complex in an active form. *Biochemistry* 26:4303–4309.

Oughtred, R., C. Stark, B.J. Breitkreutz, J. Rust, L. Boucher, C. Chang, N. Kolas, L. O'Donnell, G. Leung, R. McAdam, F. Zhang, S. Dolma, A. Willems, J. Coulombe-Huntington, A. Chatr-Aryamontri, K. Dolinski, and M. Tyers. 2019. The BioGRID interaction database: 2019 update. *Nucleic Acids Research* 47(D1):D529-D541.

Pendleton, M., R.H. Lindsey, Jr., C.A. Felix, D. Grimwade, and N. Osheroff. 2014. Topoisomerase II and leukemia. *Annals of the New York Academy of Sciences* 1310:98–110.

Pommier, Y., E. Leo, H. Zhang, and C. Marchand. 2010. DNA topoisomerases and their poisoning by anticancer and antibacterial drugs. *Chemistry and Biology* 17(5):421–433.

Pommier, Y., Y. Sun, S.N. Huang, and J.L. Nitiss. 2016. Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. *Nature Reviews: Molecular Cell Biology* 17(11):703–721.

Robert, T., A. Nore, C. Brun, C. Maffre, B. Crimi, H.M. Bourbon, and B. de Massy. 2016. The TopoVIB-Like protein family is required for meiotic DNA double-strand break formation. *Science* 351(6276):943–949.

Robinson, M.J., and N. Osheroff. 1991. Effects of antineoplastic drugs on the post-strand-passage DNA cleavage/religation equilibrium of topoisomerase II. *Biochemistry* 30(7):1807–1813.

Roca, J., J.M. Berger, and J.C. Wang. 1993. On the simultaneous binding of eukaryotic DNA topoisomerase II to a pair of double-stranded DNA helices. *Journal of Biological Chemistry* 268:14250–14255.

Roca, J., and J.C. Wang. 1992. The capture of a DNA double helix by an ATP-dependent protein clamp: a key step in DNA transport by type II DNA topoisomerases. *Cell* 71(5):833–840.

Romanienko, P.J., and R.D. Camerini-Otero. 2000. The mouse Spo11 gene is required for meiotic chromosome synapsis. *Molecular Cell* 6(5):975–987.

Sander, M., and T. Hsieh. 1983. Double strand DNA cleavage by type II DNA topoisomerase from *Drosophila melanogaster*. *Journal of Biological Chemistry* 258(13):8421–8428.

Schmidt, B.H., A.B. Burgin, J.E. Deweese, N. Osheroff, and J.M. Berger. 2010. A novel and unified two-metal mechanism for DNA cleavage by type II and IA topoisomerases. *Nature* 465(7298):641–644.

Schmidt, B.H., N. Osheroff, and J.M. Berger. 2012. Structure of a topoisomerase II-DNA-nucleotide complex reveals a new control mechanism for ATPase activity. *Nature Structural & Molecular Biology* 19(11):1147–1154.

Sng, J.H., V.J. Heaton, M. Bell, P. Maini, C.A. Austin, and L.M. Fisher. 1999. Molecular cloning and characterization of the human topoisomerase IIa and IIb genes: evidence for isoform evolution through gene duplication. *Biochimica et Biophysica Acta* 1444:395–406.

Uuskula-Reimand, L., H. Hou, P. Samavarchi-Tehrani, M.V. Rudan, M. Liang, A. Medina-Rivera, H. Mohammed, D. Schmidt, P. Schwalie, E.J. Young, J. Reimand, S. Hadjur, A.C. Gingras, and M.D. Wilson. 2016. Topoisomerase II beta interacts with cohesin and CTCF at topological domain borders. *Genome Biology* 17(1):182.

Vrielynck, N., A. Chambon, D. Vezon, L. Pereira, L. Chelysheva, A. De Muyt, C. Mezard, C. Mayer, and M. Grelon. 2016. A DNA topoisomerase VI-like complex initiates meiotic recombination. *Science* 351(6276):939–943.

Wang, J.C. 1971. Interaction between DNA and an Escherichia coli protein omega. *Journal of Molecular Biology* 55(3):523–533.

Wei, H., A.J. Ruthenburg, S.K. Bechis, and G.L. Verdine. 2005. Nucleotide-dependent domain movement in the ATPase domain of a human type IIA DNA topoisomerase. *Journal of Biological Chemistry* 280(44):37041–37047.

Wells, N.J., and I.D. Hickson. 1995. Human topoisomerase II alpha is phosphorylated in a cell-cycle phase-dependent manner by a proline-directed kinase. *European*

Journal of Biochemistry 231(2):491–497.

Yang, X., W. Li, E.D. Prescott, S.J. Burden, and J.C. Wang. 2000. DNA topoisomerase IIβ and neural development. *Science* 287(5450):131–134.

Zechiedrich, E.L., K. Christiansen, A.H. Andersen, O. Westergaard, and N. Osheroff. 1989. Double-stranded DNA cleavage/religation reaction of eukaryotic topoisomerase II: evidence for a nicked DNA intermediate. *Biochemistry* 28:6229–6236.

Zechiedrich, E.L., and N. Osheroff. 1990. Eukaryotic topoisomerases recognize nucleic acid topology by preferentially interacting with DNA crossovers. *EMBO Journal* 9:4555–4562.

Zhang, S., X. Liu, T. Bawa-Khalfe, L.S. Lu, Y.L. Lyu, L.F. Liu, and E.T. Yeh. 2012. Identification of the molecular basis of doxorubicin-induced cardiotoxicity. *Nature Medicine* 18(11):1639–1642.