The Characterisation of DNA Topoisomerase VI from *Methanosarcina mazei* Using Ensemble and Single-Molecule Methods

Shannon Jane McKie

John Innes Centre

Department of Biological Chemistry

June 2020

Statement

The work submitted within this thesis is entirely my own, except where due reference has been paid, and has not been submitted to this or any other university as part of any degree.

Dedication

I dedicate this thesis to my wonderful mother,

a truly transcendent woman whom I love dearly.

So, come up to the lab and see what's on the slab.

I see you shiver with anticipation.

But maybe the rain isn't really to blame,

So I'll remove the cause… but not the symptom!

Dr. Frank N. Furter (The Rocky Horror Picture Show)

Abstract

DNA topoisomerase VI (topo VI) is a type IIB DNA topoisomerase, found predominantly in archaea but with eukaryotic presence in various plant and algal species, and possibly plasmodia. In *Arabidopsis thaliana*, topo VI is an indispensable nuclear protein, essential for endoreduplication, despite the presence of other type IIA topoisomerases, capable of performing the same reactions. Whilst topo VI has been proposed to be a DNA decatenase since its discovery, robust evidence and a mechanism for its preferential decatenation activity was lacking. In this research, the activity of topo VI from the mesophilic archaeon *Methanosarcina mazei* (MmT6) was characterised using ensemble biochemistry, single-molecule magnetic tweezers and next-generation DNA sequencing (NGS). This demonstrated that MmT6 activity was significantly enhanced in the presence of catenated or braided DNA substrates, as opposed to supercoiled, through a strong preference for DNA crossings with angles close to 90°. In addition, MmT6 behaved as a true crossing sensor with dramatic increases in ATPase activity, DNA binding and the rate of strand-passage, with increasing DNA writhe. Taken together, these results strongly suggest that MmT6 has evolved an intrinsic preference for the unknotting and decatenation of interlinked chromosomes, simultaneously disfavouring the relaxation of supercoils, by sensing DNA crossings directly with geometries close to 90°. This provided an explanation for why topo VI homologues persist in higher eukaryotes during situations in which the genome undergoes rapid duplication, potentially as a dedicated DNA decatenase that cannot be substituted by other topoisomerases. It was also shown, using an NGS-based technique, that MmT6 cleaved DNA with the sequence preference 5'-[G/C][G/C]^[A/C/T], generating 2-base overhangs. The *in vitro* plasmid-based NGS technique developed to attain this result was extended to begin exploring the *in vitro* DNA cleavage preferences of other type IIA topoisomerases, namely *E. coli* DNA topoisomerase IV and *E. coli* DNA gyrase.

Acknowledgements

When I pause to consider all those who have supported my PhD journey, I feel immeasurable gratitude, as this was the most difficult, yet rewarding, challenge I have faced so far. Without this support I doubt this project would have been as fruitful as it has been, and so I would like to take this opportunity to sincerely thank those who came along for the ride.

To begin, and without hesitation, I would like to thank both of my supervisors: Prof. Anthony Maxwell and Dr. Keir Neuman. Many feel grateful to have had a single high-quality supervisor, whereas I was lucky enough to be supported by two of the most knowledgeable, attentive and empathetic men around. In many ways, they complemented each other perfectly, supplying me with different forms of support throughout the last 4 years, whenever I needed it. Prof. Anthony Maxwell, of the John Innes Centre, began my career by taking a chance on an undergraduate desperate for lab experience, recognising my aptitude for research and initiating a PhD collaboration with Dr. Keir Neuman. Even when I expressed fear over the prospect of living in the US for two years, he not only continued to help me with my self-belief, but also accompanied me during my first visit in the hope that it would calm my nerves, which it did immensely. Dr. Keir Neuman then welcomed me into his lab at the National Institutes of Health, Bethesda, where I was immediately made to feel like one of the group. Being someone who marvels at physics from the side-lines, but was often too scared to dive in, his patience in teaching me core biophysical concepts was remarkable and I'm a far more knowledgeable and technically skilled scientist as a result. I truly hope that both my supervisors will continue to be involved in both my career and life, as I have come to regard them both as extremely important people that have had an incredibly positive impact on my future.

With that, I would like to extend thanks to my supervisory team: Dr. Richard Bowater and Dr. Stephen Bornemann. They have watched me change from an anxious heap to a confident scientist, their supportive advice and scientific guidance aiding in this transformation.

Supervisors seldom support lab members alone and I feel extremely lucky to have been a part of two incredible groups of scientists. I will begin with the Maxwell group, of which I have been a part of for five years. Group members come and go but I would like to provide special thanks to Lesley Mitchenall and Dr. Thomas Germe, both of whom are close to my heart and have supported me both academically and emotionally. I would be two stone heavier and a whole lot sadder without our lunchtime walks together. I would also like to thank Dr. Natassja Bush for allowing me to work with her during my undergraduate degree, beginning my career in the topoisomerase field. For the Neuman group, special thanks are in order for the mentorship provided to me by Dr. Yeonee Seol. Without her help I wouldn't be the single-molecule scientist I am today, and I will always be very thankful for those times she helped a

Contents

List of Figures

List of Tables

List of Equations

Chapter 1

General Introduction

1.1   The Importance of DNA Topology

The DNA duplex is one of life's most fundamental molecules and therefore, maintaining its integrity is paramount. Not long after its structure was first elucidated in 1953 by James Watson, Francis Crick and Rosalind Franklin [1], potential topological issues associated with such a molecular structure recognised [2]. The consequences of topological perturbations in DNA are exemplified during replication in which the strands of the duplex are separated. This separation leads to the formation of positive supercoils ahead of the replication fork and precatenanes behind [3]. If the positive supercoils were not relaxed, binding of the replication machinery and progression of the replication fork would be impeded. Furthermore, during replication, the two daughter duplexes of closed-circular DNA remain interlinked as catenanes, and large linear chromosomes can become knotted and linked. During transcription, DNA strand separation also promotes positive supercoiling ahead and negative supercoiling behind the transcriptional complex. These structures must be resolved if the cell cycle is to proceed, allowing the cell to partition its genetic material and divide unhindered, as well as transcribe and synthesise polypeptides. During plasmid replication, negative supercoiling of the origin is required for initiation as local melting is necessary for protein binding [4]. And, in addition, the genome of *E. coli* is compacted, which is maintained in part by enzyme-dependent negative supercoiling [5]. The proteins responsible for performing these vital roles in DNA topology manipulation are called the DNA topoisomerases (topos) and they are crucial to the life of the cell. This indispensability, along with their prevalence across all organisms, makes DNA topoisomerases key anti-microbial and anti-cancer targets.

Chapter 1: Introduction

1.2   Describing DNA Topology

DNA topology, explored thoroughly by Bates and Maxwell [6], describes the structural consequences that occur as a result of the DNA double helix and includes supercoils, knots and catenanes. Due to the right-handed helical nature of B-form DNA, when the two strands are separated during transcription, positive supercoils are generated ahead of the fork and negative supercoils behind. In order to discuss DNA topology in this manner, a few terms warrant defining. The linking number (Lk) is a term which describes the number of times one strand of the duplex fully wraps around the helical axis, and for relaxed DNA there are 10.5 base pairs (bp) per each of these helical turns (Figure 1A). The following equation can be used to calculate the Lk for constrained, relaxed DNA, where N is the total number of bp in the molecule and h is the number of bp per helical turn (10.5).

$$Lk = \frac{N}{h}$$

Equation 1.1: Calculating the linking number (Lk) for DNA.

As an example, pBR322, a 4361 bp *E. coli* plasmid commonly used in topology assays, has a linking number of ~ +415. This value is positive as B-form DNA is a right-handed molecule. In other words, the two strands within the duplex do not simply lie parallel to each other, but rather, they wrap around each other in the right-handed direction. It is also an approximate value owing to the fact that a linking number is an integer, however the precise value obtained from using equation 1 is 415.3 as the DNA duplex is being treated as a linear segment. In reality, the plasmid is circular and the extra twisting needed to bring the two ends into register with each other will be distributed throughout the plasmid and count towards changing the topology away from the relaxed state. Therefore, the term more commonly employed when describing a fully relaxed plasmid is given the specific denotation, $Lk°$, which has a subtle difference to Lk in that it need not be an integer as it is not a true linking number, but rather a reference point for evaluating changes in topology. Furthermore, when discussing changes in topology, the change in Lk (ΔLk) is used rather than the new total Lk. For instance, if four full righthanded

turns were added to relaxed pBR322, the plasmid would be described as ΔLk = +4 rather than Lk = +419. Increasing the linking number by 4 in a small plasmid increases the torsional strain more significantly than it would in a large plasmid, so another useful topology term is the specific linking difference or superhelical density (σ), as it allows the ΔLk to be scaled with respect to the size of the DNA molecule.

$$\sigma = \frac{\Delta Lk}{Lk°}$$

Equation 1.2: Calculating the specific linking difference (σ).

By describing the extent of supercoiling in this manner, the topology of different sized plasmids can be compared.
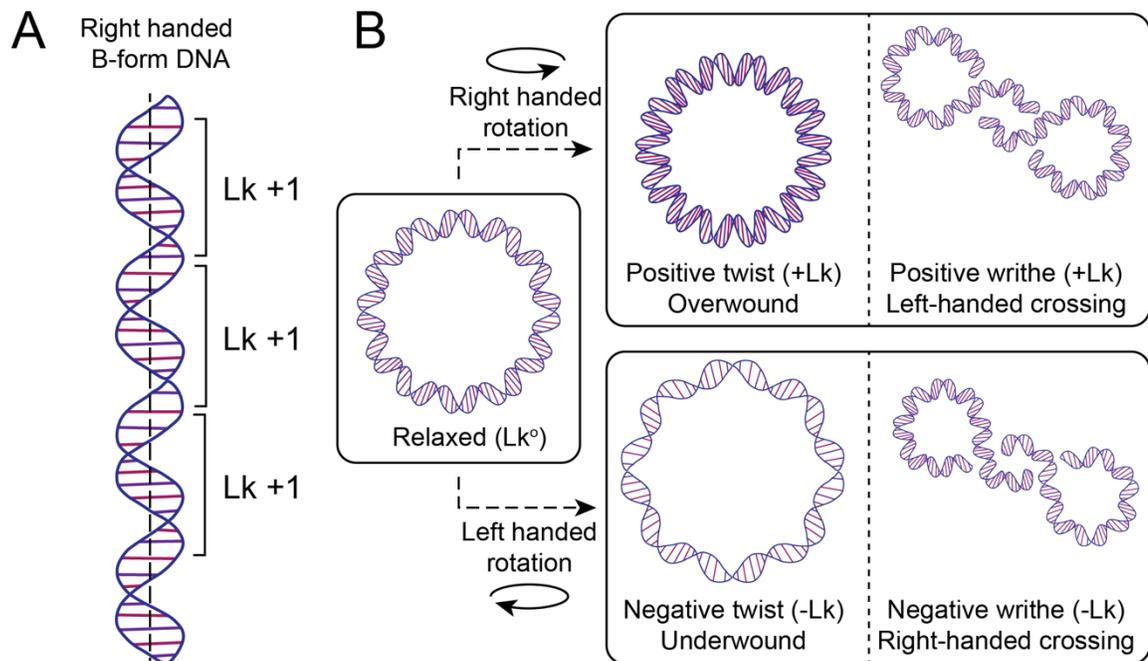
Figure 1.1: Pictorial representation of DNA topology. A: Representation of the B-form DNA duplex. The black vertical dashed line represents the helical axis and bracketed regions highlight each full helical twist, which is 10.5 bp (in relaxed DNA) and equals a linking number (Lk) of +1. B: Representation of twist and writhe in both positively- (overwound) and negatively- (underwound) supercoiled circular DNA.

Twist and writhe are important terms used to describe the physical manifestation of increasing or decreasing the Lk for a DNA duplex. If you imagine a strand of DNA constrained

Chapter 1: Introduction

at one end and rotate it 360º in a right-handed direction you will increase the Lk of the DNA segment by +1, generating positive twist in the DNA, meaning that one strand wraps fully around the other in less than 10.5 bp. This is analogous to what happens ahead of the replication fork. If you kept rotating the free end of the DNA in this direction, eventually twist would be converted to writhe and the helix would begin wrapping around itself, creating positive (left-handed) crossings. If you add twist into the helix in a left-handed direction, negative twist is introduced, decreasing the Lk and eventually converting to negative (right-handed) writhe (Figure 1B). This is analogous to what occurs behind the transcriptional fork and, *in vivo*, metabolically active DNA is generally negatively supercoiled as underwound DNA is easier to access [7]. Being the physical manifestation of the Lk, the twist and writhe within constrained DNA must add up to the Lk [8, 9].

$$Lk = Tw + Wr$$

Equation 1.3: The relationship between twist (Tw), writhe (Wr) and linking number (Lk).

As must the change in twist and writhe within that molecule, add up to the change in Lk.

$$\Delta Lk = \Delta Tw + \Delta Wr$$

Equation 1.4: The relationship among the change in twist (ΔTw), writhe (ΔWr) and linking number (ΔLk).

Alongside supercoiling, DNA duplexes can also become catenated or knotted. Catenanes are particularly prevalent within biological systems, being a natural product of circular DNA replication. A catenane is formed when two plasmids remain interlinked and this situation requires resolution before the cell can divide. Within the cell the DNA topoisomerases are responsible for mediating changes in twist and writhe by transiently cleaving one or both strands of the duplex and allowing either rotation of the DNA about its axis or passage of one or both strands through a transient break in one or both strands, thereby modifying the topology of DNA. They possess the ability to relax supercoils (or generate them, specifically in the case of DNA gyrase and reverse gyrase), unknot and decatenate DNA, all of which are vital functions and make topoisomerases fundamental to the physiology of the cell.
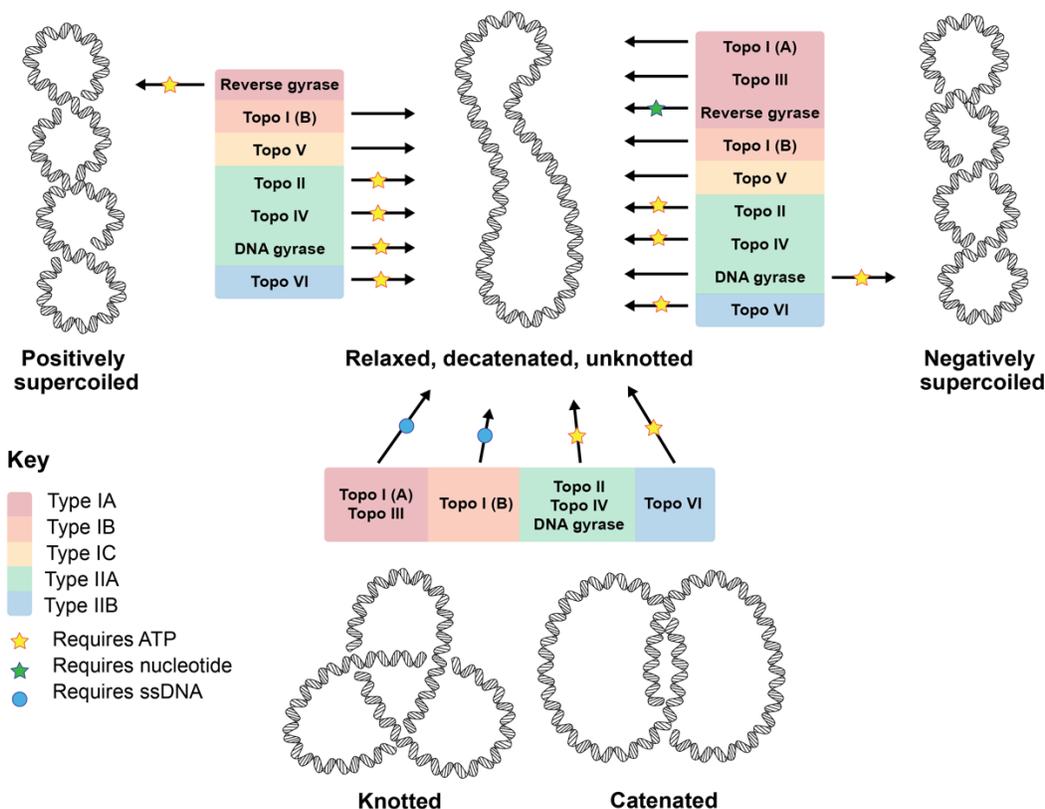
Figure 1.2: Topological activities of DNA topoisomerases: relaxation of positive and negative supercoils, decatenation, and unknotting. Colour is representative of the topo type: red is type IA, orange is type IB, yellow is type IC, green is type IIA and blue is type IIB. Yellow stars represent ATP-dependency, green stars represent nucleotide-dependency (e.g. reverse gyrase can relax negative supercoils using CTP, GTP and UTP) and the blue circles indicate that single-stranded or nicked DNA is required for activity.

1.3 DNA Topoisomerases

DNA topoisomerases (topos) are ubiquitous across all domains of life and indispensable to cellular survival due to their role in DNA topology maintenance. They are classified as type I or type II depending on whether they catalyse the formation/re-ligation of single- or double-stranded DNA breaks, respectively [10, 11]. In general all topoisomerases perform a similar task (i.e. interconverting the topological states of DNA), however, the precise ways in which this is achieved differs between enzyme groups and imparts certain activity preferences, for example: preferential decatenation rather than relaxation (Figure 1.2). These intriguing enzymes have been of keen scientific interest since the first was discovered in 1971 [12]. As

Chapter 1: Introduction

topos transiently disrupt the integrity of the duplex in order to maintain it, they must function in a highly coordinated and precise manner so as to avoid generating permanent DNA breaks. This mechanism makes topoisomerases vulnerable to poisoning, which is currently being exploited in the use of antimicrobial and anticancer therapeutics.

A key feature linking all topos, is the formation of a covalent DNA-protein intermediate in which a phosotyrosyl linkage is formed via a nucleophilic attack of the phosphate group in the DNA by the active site tyrosine of the topoisomerase. This can be the 5'-phosphate, in the case of the type IA and type II topos, or the 3'-phosphate, in the case of type IB and IC topos (Figure 1.3) [13]. Explored below are the numerous ways in which the topoisomerase family employ this DNA cleavage mechanism in the preservation of genome integrity.
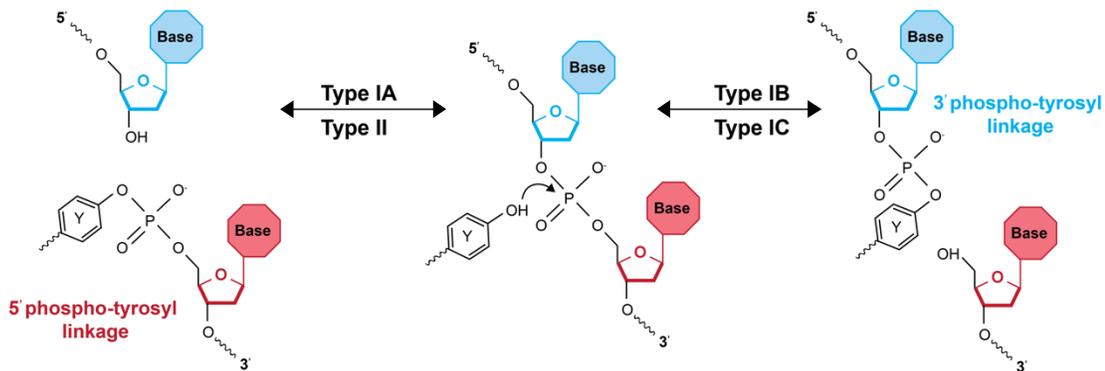
Figure 1.3: The reaction that takes place between the DNA backbone and the active-site tyrosine of the topoisomerase. The -OH group attacks the phosphate group, replacing either the 3' oxygen (in the case of type IB) or the 5' oxygen (in the case of type IA or type II), forming a covalent yet reversible bond. Image design based on [14].

1.3.1 Type I DNA topoisomerases

Type I topoisomerases are further categorised into A, B and C subgroups, depending on protein structure and mechanism. Whilst all type I topos function via the transient cleavage of a single DNA strand, and modify the DNA linking number in steps of one, the way in which this is utilised to modify DNA topology varies among the subcategories.

1.3.1.1 Type IA DNA topoisomerases

The type IA DNA topoisomerases (type IA topos), were the first to be discovered and characterised biochemically [12]. They function via the transient formation of a single-stranded DNA nick, forming a covalent $5^{'}$-phosphotyrosine linkage. Type IA topo activity is generally Mg^{2+}-dependent and ATP-independent, performing thermodynamically favourable reactions (aside from reverse gyrase, see 1.3.1.1.3).

1.3.1.1.1 DNA topoisomerase I

The first topo to be discovered was a type IA, originally called ω and now known as prokaryotic DNA topoisomerase I (topo IA) [12]. Isolated from *E. coli*, topo IA was shown to reduce the superhelical density of a negatively supercoiled plasmid in the absence of a nucleotide cofactor. It is hypothesised that the main cellular function of topo IA is in the prevention of hyper-negative supercoil accumulation during transcription, rather than global control of negative supercoiling [15]. It is likely that the activity of topo IA impacts the global supercoiling within the cell, however, this is seen as secondary to its primary function during transcription. Hyper-negative supercoiling has been demonstrated to result in stable R-loop formation (a DNA:RNA hybrid) and if unresolved, R-loops can go on to disrupt genome integrity and inhibit DNA metabolism [16]. Topo IA has been shown to prevent the excessive accumulation of R-loops. These nucleic acid structures are an ideal substrate for topo I due to the presence of single stranded-DNA [17], a feature required for optimal topo IA binding [18]. Furthermore, topo IA has been shown to directly interact with the $β^{'}$ domain of RNA polymerase via its C-terminal domain [19]. Together, these data strongly indicate that the main role of topo IA is topology maintenance, specifically during transcription, counteracting the activity of DNA gyrase (section 1.3.2.1.1) and preventing the formation of hyper-negative supercoils. Interestingly, when using positively supercoiled DNA containing a single-stranded loop, topo IA exhibited relaxation activity [20]. This implies that the topo IA preference for relaxing negatively supercoiled DNA arises mainly from the presence of single-stranded DNA (ssDNA) regions which are far more prevalent in negatively supercoiled DNA.

Chapter 1: Introduction

Topo IA is a monomeric 97-kDa protein with a toroidal conformation (Figure 1.4A and B). It has a positively charged central cavity large enough to accommodate a single DNA duplex (2.8 nm in diameter). The first topo I structure was solved using a truncated 67-kDa fragment, missing a portion of the C-terminal domain (CTD) and unable to relax negatively supercoiled DNA (however the ability to cleave ss-DNA remained intact) (Figure 1.4A) [21, 22]. Even though this structure led to invaluable insights on the mechanism of topo I, a full-length structure complexed with ss-DNA has also been solved [23] (Figure 1.4B). The presence of both the ss-DNA and the topo IA CTD, has facilitated even deeper insight into the mechanism of topo IA. This monomeric protein has 9 distinct domains, highlighted in Figure 1.4E. The 67-kDa structure (Figure 1.4A) contained domains 1 to 4, with domains 5 to 9 being included in the full-length structure (Figure 1.4B). Domains 1, 3 and 4 are believed to be the site in which the ssDNA substrate is bound and cleaved; the active site tyrosine (Y319) being present in domain 3. Domain 2 is a hinge region, and whilst the architecture of domains 1, 3 and 4 are in good agreement between the two structures, domain 2 shows the largest conformational change, indicating that the presence of the CTD and ssDNA substrate has a more profound effect on domain 2 architecture. Moreover, a flexible loop of charged residues in domain 2, disordered in the 67-kDa structure, is now visible in the full-length structure and hypothesised to play a role in strand passage as it is highly conserved amongst type IA topos. Domains 5 to 7 are 4-cys zinc ribbon domains and each coordinates a Zn^{2+} using 4 cysteines, a methionine and a hydrophobic residue (residue identity varies amongst domains). The ssDNA in this structure is bound to the CTD, with the 3'OH group interacting predominantly with domain 5. Domain 6 is the primary domain of the CTD which is in contact with the NTD, mostly with domains 2 and 4. Domains 8 and 9 are zinc ribbon-like domains and do not coordinate Zn^{2+} as the methionine has been substituted for a phenylalanine in domain 8, and a tyrosine in domain 9. Whilst these two domains play a role in activity, if they are removed the protein retains partial relaxation
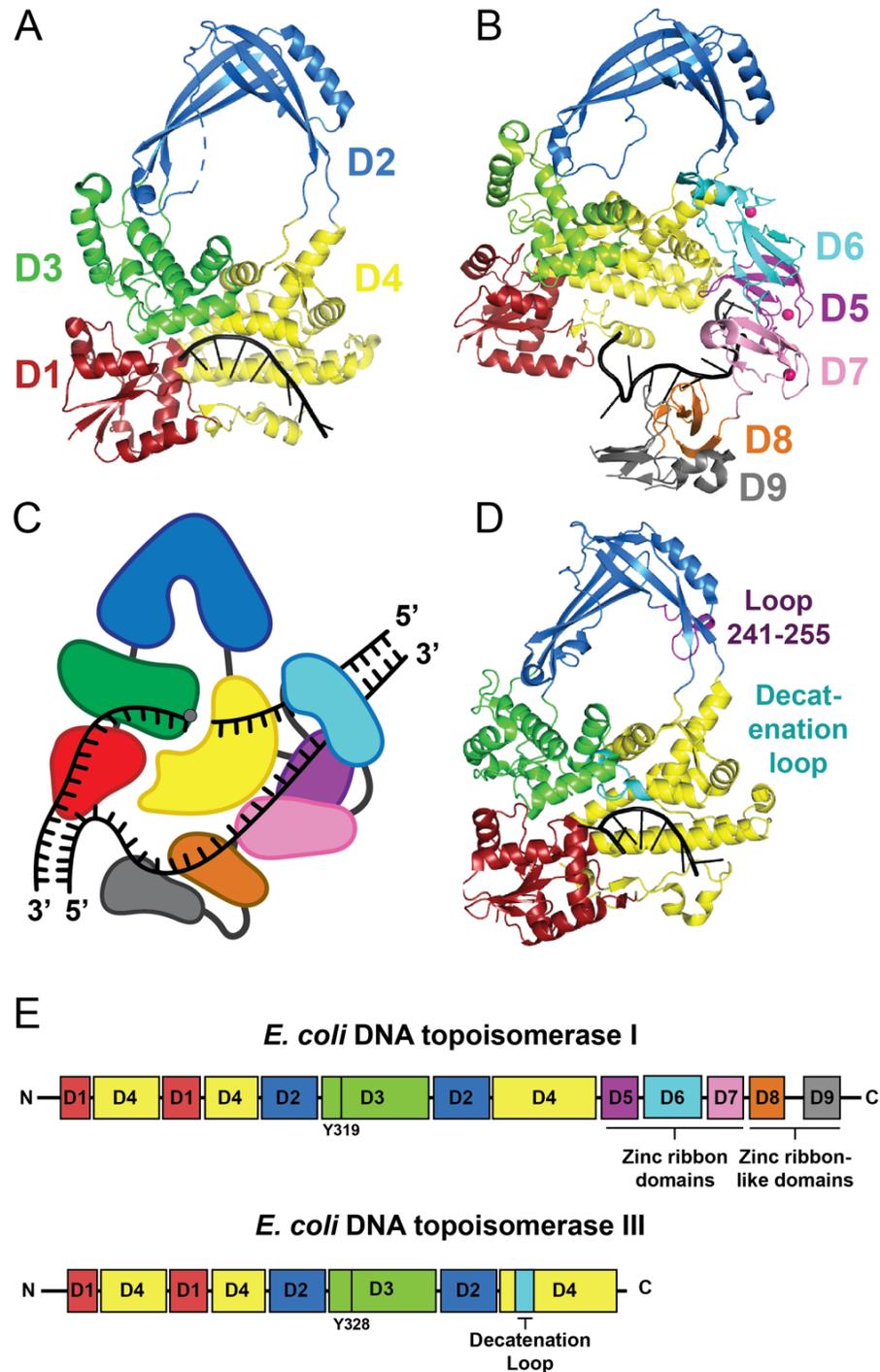
Figure 1.4: Protein structures and gene domain organisation of type IA DNA topoisomerases. A: 67-kDa structure of G-segment bound *E. coli* DNA topoisomerase I (topo IA). PDB: 3PX7 (Zhang et al., 2011, *unpublished*). **B:** Full-length (97-kDa) structure of T-segment bound *E. coli* topo IA. PDB: 4RUL [23]. **C:** Topo IA double chain model, demonstrating predicted DNA binding position for the G-segment (cleaved by the catalytic tyrosine in D3) [23]. **D:** Structure of *E. coli* DNA topoisomerase III (topo III). PDB: 2O54 [24]. **E:** Gene domain schematics for E. coli topo I and III. Black vertical bars indicate key residues. D stands for domains, which are colour coded throughout as follows; D1 is red, D2 is blue, D3 is green, D4 is yellow, D5 is purple, D6 is teal, D7 is pink, D8 is orange, and D9 is grey. DNA is in black.

activity [25]. However the activity is far more distributive, in comparison to the processive full-length topo IA, indicating that domains 8 and 9 are likely to be important for ssDNA binding. In light of this new structural information, a model for topo IA binding was proposed in which the gate strand (G-strand) is bound and cleaved by domains 1, 3 and 4, with the transport strand (T-strand) being bound by domains 5 to 9 (Figure 1.4C). It is then proposed that the way in which this enzyme alters topology of DNA is by the passing of the T-strand through the cleaved G-strand, in a process termed 'strand-passage'. The term strand-passage is applied to the reactions executed by all type IA and type II topoisomerases, the difference being that type IA topos pass a single strand or duplex (topo III, described below) through a break in another single strand, whilst type II topos form double-stranded breaks (DSBs) through which another duplex is passed.

1.3.1.1.2 DNA topoisomerase III

DNA topoisomerase III (topo III) is another integral member of the type IA subfamily and is highly conserved across not only prokaryotic lineages but also archaeal and eukaryotic. Whilst the focus here is mainly on prokaryotic topo III, it's worth noting that in eukaryotes such as Drosophila and mammals, two isoforms, topo IIIα and IIIβ, are found and appear to play distinct roles in cellular development [26]. As seen in Figure 1.4D, the full-length crystal structure of *E. coli* topo III closely resembles that of the 67-kDa structure of topo I (Figure 1.4A), with 4 distinct domains (Figure 1.4E), and also functions in an ATP-independent manner. Topo III lacks the zinc-binding domains of topo I, but has added sequence that appears to imbue the enzyme with highly efficient decatenation activity, allowing topo III to pass a duplex through a single-stranded break, providing the other strand contains a pre-existing nick. These regions are known as the decatenation loop and loop 241-255. Whilst the function of loop 241-255 remains unclear, the decatenation loop has been established to play an important role during decatenation [27]. If the loop is replaced with a homologous sequence from topo I, the

decatenation activity is almost completely abolished, thought to be a consequence of reduced protein interaction with the duplex strand.

1.3.1.1.3 Reverse gyrase

Reverse gyrase is a distinctive type IA topoisomerase, as unlike topo I and III, it utilises ATP to both relax negative supercoiling and input positive supercoils. It was first discovered in *Sulfolobus* [28] and has since been identified in many species of thermophilic and hyperthermophillic archaea and eubacteria, indicating that reverse gyrase is important for adaptation to high temperature environments [29]. This curious enzyme has two distinct regions with the NTD exhibiting sequence similarity with helicases from superfamilies I and II, whereas the CTD bears remarkable similarity to prokaryotic topo I (Figure 1.5A) [30]. In 2002, reverse gyrase from *Archaeoglobus fulgidus* became the first full length supercoiling enzyme to be crystallised, with and without the non-hydrolysable ATP analogue, ADPNP (Figure 1.5B) [31]. This structure revealed a 'padlock' shaped enzyme with two folds (H1 and H2) at the NTD that resembled the *E. coli* RecA protein's ATP-binding region. At the CTD, four domains (TOPO, TOPRIM, Zn1 and Zn2) are arranged in a toroidal structure and are highly homologous to domains 1-4 of prokaryotic topo IA. The helicase and topo domains of reverse gyrase interact via the latch domain, which has structural similarity to the *E. coli* transcription terminator, rho. The helicase domain of reverse gyrase has been shown to alternate between high and low DNA-affinity states in an ATP-dependent manner [32], which, in the presence of ATP, can transiently destabilise the DNA duplex [33]. The topo I-like domain has been shown to exhibit nucleotide-independent DNA relaxation when the helicase domain is deleted, a common feature amongst type IA topos [34]. It appears that the unique combination of these two protein regions, imparts reverse gyrase with the ability to positively supercoil DNA. Interestingly, in the absence of the latch domain, reverse gyrase is unable to supercoil due to abolishing the enzymes ability to unwind the DNA, thus functionally linking the destabilisation of the duplex to the strand passage reaction, whereby the protein increases the DNA linking number [33]. Also, reverse gyrase is able to relax negative writhe when ATP is replaced by CTP, GTP or UTP, but

as it is unable to hydrolyse these nucleotides, only ATP hydrolysis can support positive supercoiling [35].

Figure 1.5: *Archaeglobus fulgidus* **Reverse gyrase gene domain organisation and protein structure. A: Gene domain organisation of** *A. fulgidus* **reverse gyrase. B: Structure of** *A. fulgidus* **reverse gyrase. PDB: 4DDU [31]. Throughout domains are labelled as follows: Zn1 is grey, H1 is red, insert is yellow, H2 is green, latch is orange, TOPRIM is teal, Zn2 is purple, and TOPO is blue.**

1.3.1.2 Type IB DNA topoisomerases

The type IB topos, e.g. eukaryotic DNA topoisomerase I (topo IB), were first discovered in 1972 and found to be able to relax both positive and negative supercoils in closed circular DNA via a transient single-stranded cleavage event between the protein and the 3' phosphate of the DNA backbone [36]. Human topo IB was crystallised using an N-terminally truncated 70-kDa protein in complex with a 22-bp DNA duplex [37] (Figure 1.6), which led to a greater understanding of how this enzyme binds DNA and the proposal that it functions via the

controlled rotation mechanism, first described using Vaccinia topo I [38]. This mechanism is quite different to the strand-passage reaction employed by type IA and begins with topo IB creating a single-strand nick in the DNA. This then permits the rotation of DNA, the speed of which is controlled by friction within the enzyme cavity, before the re-ligation of the strand. The core domain, formed from the three subdomains (Figure 1.6) and the C-terminal domain (CTD) are strictly required for relaxation activity, and when expressed and purified individually, the reconstitution of these two domains is the minimal requirement for activity [39]. Both the N-terminal domain (NTD) (not crystallised in Figure 1.6) and the linker are poorly conserved among the type IB topos and less important for relaxation *in vitro* [40]. However, the NTD harbours 6 nuclear localisation signals (NLS) and is therefore necessary for activity *in vivo* [41]. Using single-molecule methods, it was shown that, on average, human topo IB removed 19 DNA turns per cleavage/re-ligation cycle [42]. The *in vivo* role of topo IB is believed to be in the removal of torsional strain during replication, transcription and chromosome condensation [40].
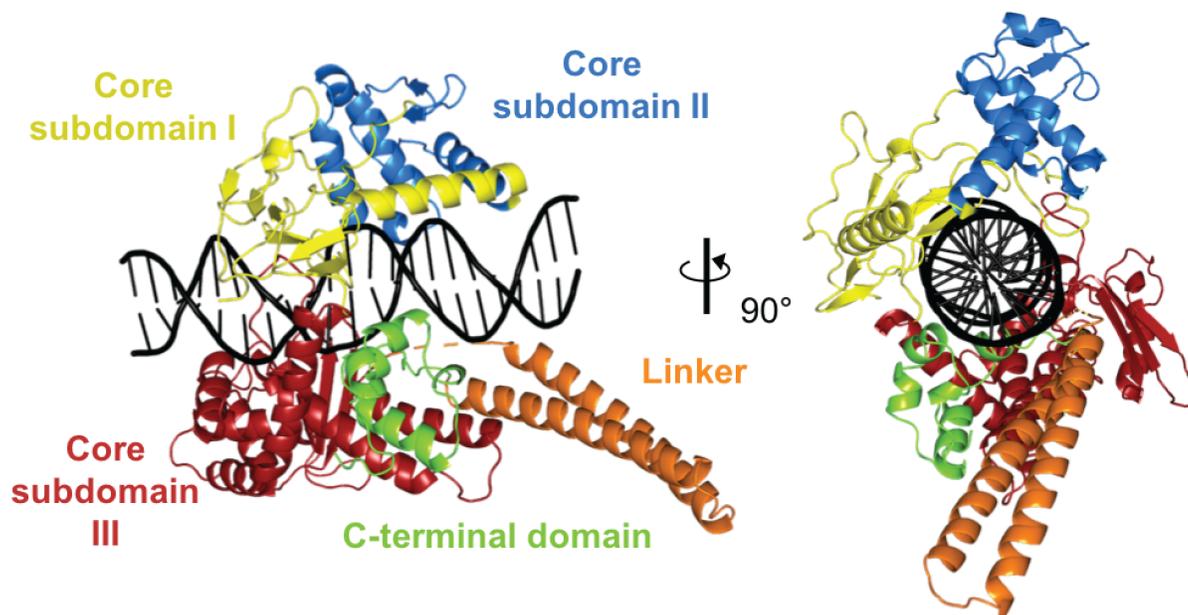
Figure 1.6: DNA-bound structure of human DNA topoisomerase IB. The core subdomain I is yellow, II is blue, III is red, C-terminal domain is green, linker is orange and DNA is black. PDB: 1A36 [37].

1.3.1.3 Type IC DNA topoisomerases

DNA topoisomerase V (topo V) was first isolated from the hyperthermophillic methanogen, *Methanopyrus kandleri* and described as the prokaryotic version of topo IB [43]. This was based on the findings that topo V activity was both ATP- and Mg^{2+}-independent, that it could efficiently relax both negative and positive supercoils, and that the covalent protein:DNA complex was formed with the 3' phosphate of the DNA backbone. In addition to these data,
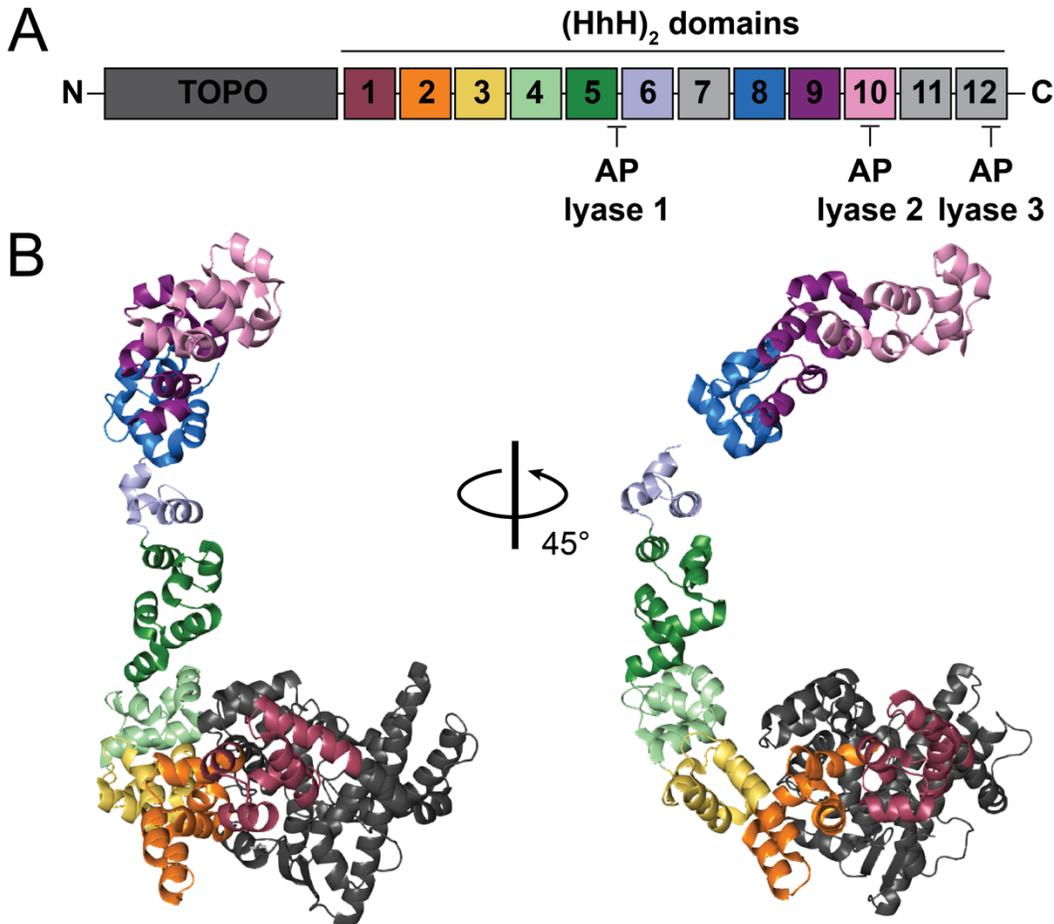
Figure 1.7: DNA topoisomerase V (topo V) gene domain organisation and protein structure. A: Gene domain organisation of *Methanopyrus kandleri* topo V. B: Structure of *M. kandleri* topo V. PDB: 5HM5 [31, 44]. Throughout domains are labelled as follows: TOPO is grey, (HhH)$_2$ domains are each coloured individually (colour scheme of gene domains matching that of structure).

a single molecule approach found that topo V relaxed DNA using controlled rotation, at a rate of 12 turns per cleavage/re-ligation cycle, on average [45]. For these reasons, topo V was initially regarded as type IB, however it has since been proposed that it constitute a new type

IC class due to structural and biochemical differences [44, 46]. The first topo V structure was a 61-kDa N-terminal fragment, from *M. kandleri*, and it revealed how divergent this protein was in relation to all other topoisomerases [46] (Figure 1.7). Not only did it lack any sequence similarities, the folds adopted by topo V are unique amongst topoisomerases and the active site itself is quite different, indicating an alternative cleavage/re-ligation mechanism. In addition to this, topo V also has DNA repair activity as an AP(apurinic or apyrimidinic)-lyase, involved in the repair of abasic DNA damage [47, 48]. This DNA repair activity functions independently of the topoisomerase activity as mutation of the active site tyrosine did not change topo V's ability to perform DNA repair [47]. With the resolution of the 97-kDa structure in 2016, topo V has proved to be a very curious enzyme with a total of four active sites contained within a single polypeptide; one topoisomerase site and three AP lyase sites [44] (Figure 1.7). The protein has a globular N-terminal region housing the topoisomerase active site, followed by 12 helix-hairpin-helix, motif 2 ((HhH)$_2$) domains. The first AP-lyase active site was determined to be between the fifth and sixth (HhH)$_2$ domains [48], the second in the vicinity of the tenth (HhH)$_2$ domain and the third at the extreme CTD within the twelfth (HhH)$_2$ domain [44]. Even though all topos are considered to be indirectly involved with DNA repair and maintenance, only topo V possesses the direct ability to repair abasic DNA damage. Whilst fascinating, this unique enzyme appears to only be present in a single species of archaeal halophilic hyperthermophiles, *M. kandleri*, living within hydrothermal vents of the deep ocean, and has therefore been postulated to have a viral origin [49].

1.3.2 Type II DNA topoisomerases

The type II topoisomerases (type II topos) are divided into A and B subgroups based on structural and evolutionary grounds [11]. All type II topos function via transient double-stranded DNA breaks (DSBs) and employ a strand-passage mechanism, through which DNA topology can be manipulated. They share key functional domains (explored below) and employ a divalent cation-dependent mechanism (preferably Mg^{2+}) in order to change Lk by two, generally

in pursuit of DNA relaxation and decatenation.

1.3.2.1 Type IIA topoisomerases

Type IIA topos are well-studied and include the prokaryotic enzymes, DNA gyrase (gyrase) and topoisomerase IV (topo IV), and eukaryotic topoisomerase II (topo II). All three enzymes have been shown to decatenate, unknot and relax positively and negatively supercoiled DNA. For topo IV and topo II, these activities are Mg^{2+}- and ATP-dependent, with the exception being the ATP-independent relaxation of negatively-supercoiled DNA by gyrase. In addition, gyrase can also introduce negative supercoils in the presence of ATP. Topo IV and gyrase are heterotetramers with an A_2B_2 subunit arrangement, whilst topo II is a homodimer, with the monomer seemingly arising from the fusion of the two gyrase subunits [50]. In general, strand passage for the type IIA topos begins with the binding of one DNA duplex, termed the gate segment (G-segment), to the enzyme at the DNA gate, which is then transiently cleaved forming a DSB with 5´ 4-base overhangs. Another duplex, termed the transport segment (T-segment), is captured by the ATP-operated clamp (N-gate) and passed through the break in the G-segment before it is resealed. The T-segment exits through the C-gate, resulting in a *ΔLk* of 2, and the ATP-operated clamp reopens allowing the enzyme to perform another consecutive round of strand-passage or disengage from the DNA (Figure 1.8). This reaction is also known as the two-gate model, owing to the alternate opening and closure of the N-gate and the DNA-gate [51]. The specifics of this reaction, such as the effect of DNA topology on enzyme activity, varies amongst the type IIA topos, with each type both intrinsically (e.g. structure) and extrinsically (e.g. protein-protein interactions and temporal/spatial regulation) adapted to preferentially perform different DNA topology manipulations. Together with type I topos, the type IIA topos are fundamental to cellular viability through the maintenance of DNA integrity. A key aspect that makes type IIA topos both fascinating and clinically-relevant, is this role in DNA-integrity protection, which is achieved through disruption of the DNA duplex. The formation and resealing of the DSB must be highly coordinated and rapid, so that type II topos

do not form extensive genotoxic damage. Under normal circumstances this activity is highly efficient, but also forms a juncture of vulnerability which can be exploited by drugs targeting either the activity of type II topos or stabilising the DSBs, preventing re-ligation.
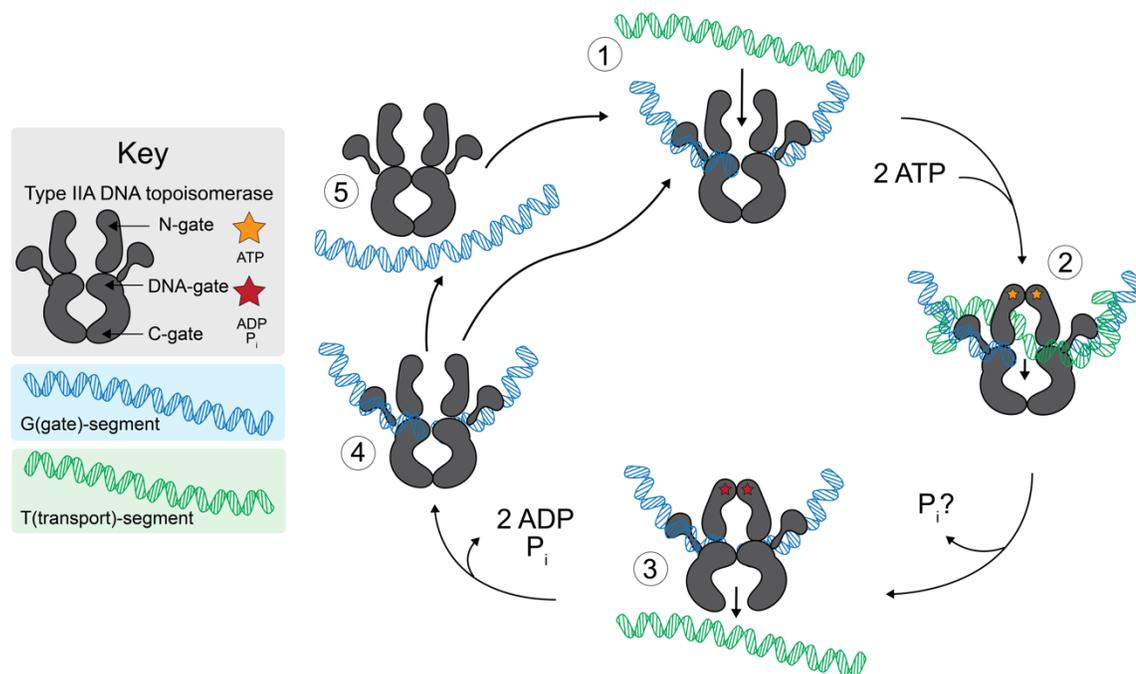
Figure 1.8: Generalised strand-passage mechanism for the type IIA DNA topoisomerases. (1) The G-segment is bound to the DNA-gate of the topo, and the T-segment is captured. (2) ATP binding drives dimerization of the N-gate and cleavage of the G-segment, through which the T-segment is passed. (3) The DNA-gate closes, G-segment re-ligated and the T-segment released through the C-gate. (4) ADP and phosphate dissociate and the N-gate reopens. (5) The enzyme may disengage with the G-segment, or, (1) remain bound and perform a consecutive strand-passage event.

Structurally, there are a number of key protein domains shared amongst the type IIA topos (as well as some also found in type IIB and type I topos) [13] (Figure 1.9). In the type IIA topos, all the following domains are present in pairs within the holoenzyme. Firstly, there are two domains fundamental to the Mg^{2+}-dependent double-stranded DNA cleavage activity, the WHD (winged-helix domain), also known as the 5Y-CAP domain, and the TOPRIM (topoisomerase/primase) domain, which are located in the prokaryotic A subunit (named GyrA in gyrase and ParC in topo IV). The WHD contains a helix-turn-helix fold, commonly found in other DNA-binding proteins, such as the *E. coli* catabolite activator protein (CAP) [52].

Chapter 1: Introduction

However, even though this domain is often associated with highly sequence-specific DNA-binding activity in alternative proteins, in topos it has seemingly evolved to supress DNA sequence preference, allowing topos to bind and act on DNA duplexes of any sequence [53]. This appears to be due to the two 'recognition' helices being much closer together in topos, preventing them from binding a major groove each and hence occupying a single major groove together, thus reducing the extent of the intermolecular interactions with the DNA and supressing sequence preference. Interestingly, with the advent of next-generation sequencing (NGS) techniques, mild sequence preferences are evident within different topo families, to differing degrees and in response to variable DNA sequences, indicating that the WHDs have partially retained the ability to bind and cleave particular DNA sequences more efficiently, however their activity on alternative sequences is not restricted by this [54]. This makes logical sense as DNA metabolism, particularly DNA replication, affects the entire genome and causes topological perturbations independent of DNA sequence, which topos must resolve. In addition to the DNA binding activity of the WHD, it also houses the catalytic tyrosine residue, involved in the attack on the scissile phosphate of the DNA backbone and forming the reversible covalent bond [55], and an isoleucine, which intercalates into the DNA and produces a ~150° bend in the G-segment [56], which is considered important for promoting DNA cleavage efficiency [57].

The second region important for DNA cleavage activity is the TOPRIM domain, formed from an α/β/α Rossmann-like fold [53]. This protein conformation is commonly found in proteins that act on DNA, including DnaG type primases, OLD family nucleases, RecR proteins and topos [58]. For type II and IA topos, this domain is fundamental to the transesterification reaction between the scissile phosphate and the catalytic tyrosine, as it contains a DxD motif which chelates Mg^{2+}, and a glutamate residue thought to act as a general acid/base during the reaction [58, 59]. Together, the DxD Mg^{2+} chelating motif of the TOPRIM domain and the active site tyrosine of the WHD, form a bipartite active site capable of forming transient breaks in the duplex [13]. In addition to its role in cleavage catalysis, the TOPRIM domain also houses conserved regions, namely the EGDS and PLRGK motifs, which form numerous interactions

with the G-segment and assist with DNA binding [53, 60].
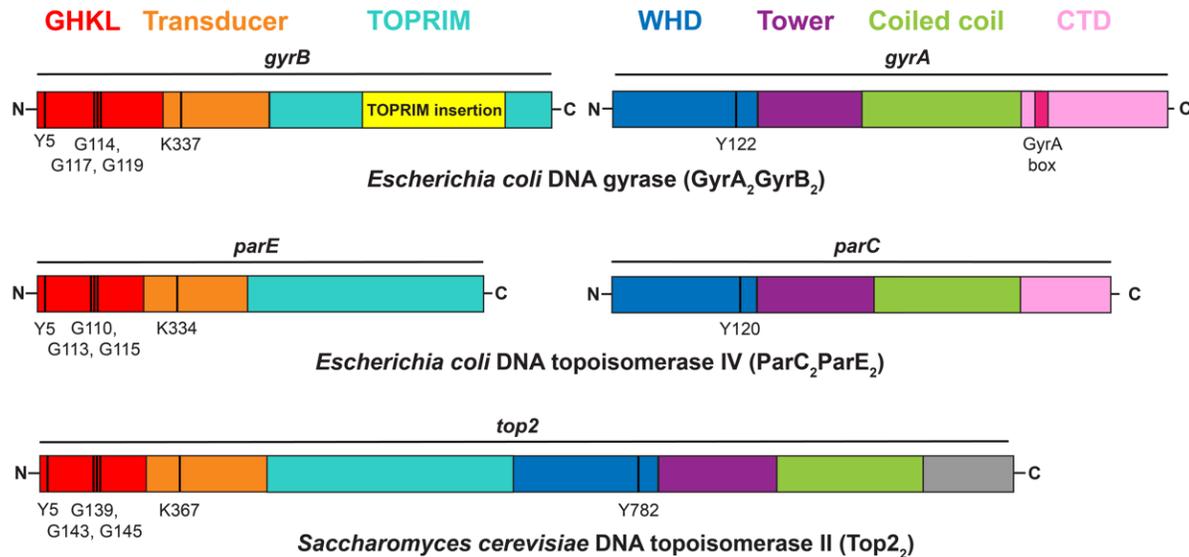
Figure 1.9: Type IIA DNA topoisomerase gene domain organisation of *E. coli* DNA gyrase, *E. coli* DNA topoisomerase IV (topo IV) and *Saccharomyces cerevisiae* DNA topoisomerase II (topo II). The domains are colour coded throughout, as follows: GHKL is red, transducer is orange, TOPRIM is teal, TOPRIM insertion is yellow, WHD is blue, tower is purple, coiled coil is green, CTD is pink (grey in the case of topo II), and the GyrA box is magenta. Vertical black bars represent key catalytic residues.

A third domain, only present in type IIA topos and thought to be important in DNA binding, is the tower. It contains a β-sheet that interacts with the G-segment as it leaves the WHD, anchoring the outer portion of the bent duplex, stabilising it and therefore promoting both DNA binding and cleavage efficiency [56]. It has been speculated, that as this domain is absent in type IA and IIB topos, it is likely they are unable to induce DNA bending [53]. However, for type IIB enzymes, there is mounting evidence in opposition to this, demonstrating that DNA bending is either induced or recognised to facilitate binding, albeit via a distinct mechanism [61, 62].

The fourth noteworthy domain is the GHKL (DNA Gyrase, Hsp90, bacterial CheA-family histidine kinases and MutL) ATPase domain. The ATP binding region of this domain is formed from an 8-stranded antiparallel β-sheet surrounded by several α-helices, and the binding of ATP stimulates B-subunit dimerisation, stabilising the N-gate in a closed conformation [13]. The

bound ATP is contacted by a lysine residue, which resides outside of the GHKL domain, in the transducer domain, and likely stabilises the ATP-hydrolysis transition state by forming a hydrogen bond to the γ-phosphate [63-65]. This interaction has been shown to facilitate significant rotation (11-18°) between the GHKL and transducer domains [64, 66, 67] and is thought to permit transduction of free-energy from ATP hydrolysis to the DNA-gate. The precise role(s) of ATP in type II topo activity is still unclear, however, along with evidence that it promotes DNA cleavage through transducer domain movement (although this clearly isn't necessary for gyrase-dependent DNA relaxation), it is hypothesised that ATP binding/hydrolysis modulate B-subunit dimerization, forming a stable interface which protects against the formation of genotoxic DSBs [68]. Gyrase may be able to circumvent this necessity due to its extensive DNA wrapping ability (see 1.3.2.1.1).

Even though there are a number of significant similarities within the structure and function of type IIA topos, there are also key differences that speak to their evolution into distinct enzyme types with varying preferential activities *in vivo*, explored in depth below.

1.3.2.1.1 DNA gyrase

DNA gyrase is a unique type IIA topoisomerase, found predominantly in bacteria [69], however it is also present in plants [70], apicomplexans [71] and archaea [72, 73]. In 1972, it was shown using sedimentation analysis, that the DNA within *E. coli* was negatively supercoiled [74], with gyrase, the enzyme responsible for this DNA topology, purified in 1976 [69]. Gyrase can relax supercoils (positive and negative), decatenate and unknot DNA, in an Mg^{2+}-dependent manner [11]. A unique feature of gyrase is that it can also input negative supercoils into closed-circular DNA molecules in the presence of ATP [69], and that the relaxation activity is ATP-independent [75] (unknotting and decatenation are ATP-dependent [76-78]). *E. coli* gyrase, the most extensively characterised, is a 374-kDa heterotetrameric enzyme formed from two A-subunits (GyrA) and 2 B-subunits (GyrB), 97-kDa and 90-kDa, respectively [75]. As gyrase is essential to bacterial viability, and absent in humans, it has had significant and ongoing clinical success

as a target for antibacterial agents [79, 80].

Structurally, DNA gyrase is formed of the same core protein domains as other type IIA topos (explored in 1.3.2.1, Figure 1.9), however, its unique ability to negatively supercoil DNA arises from its ability to wrap DNA via the CTDs of the GyrA subunit. This 35-kDa GyrA CTD is formed of six β-strands which form a β-pinwheel fold with two-thirds of the outer surface being basic, which indicates a role in DNA binding/bending [81-84] (Figure 10C). Also within the CTD is a structural motif, situated within a loop between β-strands 1 and 6, called the GyrA-box, which has been determined to be crucial to the supercoiling activity of gyrase [85, 86]. The GyrA-box is seven amino acids long (QRRGGKG), and if deleted or substituted with alanines, gyrase is rendered unable to negatively supercoil (whilst the decatenation and relaxation activities remain unaffected) [85]. Hence the presence of the GyrA-box in gyrase, and its absence in all other type IIA topos, suggests it is the major determinant of gyrase supercoiling activity. The way in which gyrase is thought to impart negative supercoiling begins with G-segment binding to the DNA-gate (formed from the WHD of GyrA and the TOPRIM domain of GyrB) which is then chirally wrapped around one of the GyrA CTDs, before being presented over the G-segment, at ~60°, as the T-segment, which forms a left-handed (positive) crossing. The passing of the T-segment in this orientation, through the G-segment, results in the conversion of the left-handed crossing to a right-handed (negative) crossing, thus adding negative writhe to the DNA molecule [87]. The length of DNA bound and wrapped by gyrase, and therefore protected from degradation, is between ~120-150 bp, as measured by nucleolytic digestion [88-90].

Attaining a full-length crystal structure of gyrase has continued to prove unsuccessful, however there are many examples of truncated, fused and domain structures, which have added greatly to our understanding of gyrase activity. The first crystal structure was of the GyrB N-terminal domains (NTDs) in complex with ADPNP, which revealed a dimer with an extensive interface between the two subunits, and forming a positively-charged 20 Å-wide cavity, large enough to accommodate a duplex [91] (Figure 10A). Follow-up mutagenesis studies showed that at least one of the arginines within the positively-charged central cavity is important for

21

DNA supercoiling and DNA-dependent ATPase stimulation [92]. The GyrB CTD has also been crystallised, from *E. coli* and *Mycobacterium tuberculosis,* either in isolation [93, 94] or as part of a GyrB-A fusion [95], revealing another dimeric structure with the TOPRIM domains forming the dimer interface and the transducer domains extending upwards [93]. The *E. coli* GyrB CTD structure, crystallised using a fusion of amino acids 388-804 of GyrB, directly to amino acids 1-525 of GyrA, allowed elucidation of a large 170 amino acid TOPRIM insertion, present in a number of bacterial phyla, including proteobacteria, acidobacteria, chlamydia and aquificales, but absent in others such as actinobacteria (e.g. *M. tuberculosis*) and firmicutes [95] (Figure 1.10B). The domain adopts a novel fold formed from a unique combination of β-sheets and α-helices, that protrudes from the TOPRIM domain, lying alongside the coiled-coil domain of the GyrA subunit. Its deletion leads to dramatic reduction in DNA binding, supercoiling and ATPase activity [95, 96].
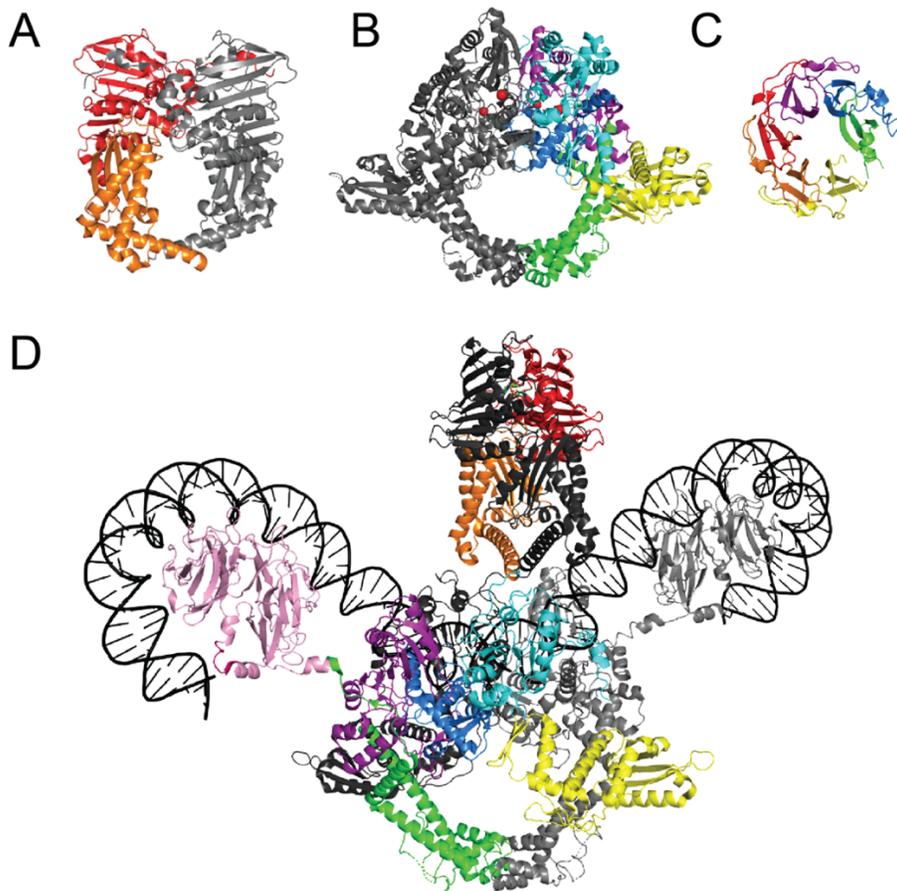
Figure 1.10: Structures of *E. coli* DNA gyrase. A. Homodimer of *E. coli* GyrB bound to ADPNP. PDB: 1EI1 [97]. B: GyrBA fusion of *E. coli* DNA binding/cleavage core. PDB: 3NUH [95]. C: *E. coli* gyrase GyrA CTD. Each β-strand is coloured differently to demonstrate the pinwheel structure.

22

PDB: 1ZI0 [83]. A-C are crystal structures. D: Cryo-EM structure of full length *E. coli* DNA gyrase bound to DNA and gepotidacin. PDB: 6KRW [98]. In figures A, B and D, domains are coloured as follows: GHKL is red, transducer is orange, TOPRIM is teal, TOPRIM insertion is yellow, WHD is blue, tower is purple, coiled coil is green, CTD is pink, and the GyrA box is magenta.

As technologies have advanced, cryogenic electron microscopy (cryo-EM) has become an excellent structural technique for use on large multimeric proteins that have been resistant to crystallisation. In 2019, the first full length structure of *E. coli* DNA gyrase in complex with DNA and the antibiotic gepotidacin, was solved using cryo-EM, with resolution of the DNA-binding domains reaching as high as 3.0 Å [98] (Figure 1.10D). This unprecedented structure gives an in-depth view of the overall architecture of DNA gyrase, revealing the spatial organisation of the domains, the position of the GyrA-box, and insight into DNA-cleavage site conformational changes, particularly in regard to the position of the TOPRIM insertion domain. This work has paved the way for future structural characterisation studies on not only gyrase, but other type IIA topos.

Gyrase is a vital cellular component, integral to bacterial DNA metabolism. Early on, it was found to be the target of nalidixic acid and oxolinic acid [99], as well as novobiocin and coumermycin [100], all of which are compounds previously used to inhibit DNA replication in *E. coli* [101, 102]. However, although it is clear that gyrase is fundamental to cellular viability, exactly which roles it plays *in vivo* has been more difficult to elucidate. A lot is known about the *in vitro* activity of gyrase, but whether this translates directly to its *in vivo* behaviour is not a straightforward question to answer. It is believed that the fundamental role of gyrase *in vivo* is the introduction of negative supercoiling. Indeed, if gyrase is inhibited, the genome becomes relaxed, indicating that gyrase plays a role in the homeostatic maintenance of the genome in a negatively supercoiled state [103-106]. Negative supercoiling has been shown to be important for the initiation of DNA replication and transcription as underwinding the DNA encourages melting of the origin and gene promoters [107, 108]. Using microarray analysis, inhibition of DNA gyrase (and topo IV) caused a variable effect on gene transcription [109]. The expression of certain genes were more sensitive to the reduced level of negative supercoiling, for instance,

nalidixic acid inhibition reduced the expression of catabolite sensitive genes, such as the lactose and galactose operons [110]. However, for some genes, a reduction in the level of negative supercoiling caused an increase in expression [109]. The expression of the genes encoding GyrA and GyrB are themselves sensitive to the superhelical density, with expression stimulated when the DNA is relaxed [106].

In addition to its role in maintaining a negatively supercoiled genome, gyrase is also considered vital during the elongation phase of replication and transcription, a role of which the effects are intertwined with homeostatic negative supercoiling. The twin supercoiled domain model, first described in 1987 [111], describes how DNA strand separation during transcription leads to positive supercoiling ahead of the fork and negative supercoiling behind (this is similar for replication, however precatenanes form behind). If the positive supercoiling is not relaxed, the built-up torsional strain will impede the progression of the protein machinery, and prevent DNA replication or transcription. Numerous lines of evidence support the role of gyrase in positive supercoil relaxation, including gyrase loss-of-function mutations resulting in a dramatic decline of both replication and transcription, as measured by the level of incorporated tritium (^{3}H)-labelled nucleotides [112]. Using NGS, the binding of both gyrase and topo IA were enriched in areas of high transcriptional activity, directly correlated to the binding of RNA polymerase, and also at the origin of replication, implicating gyrase and topo IA in these processes *in vivo* [113]. An *in vitro* DNA replication system also demonstrated a role for gyrase in the removal of positive supercoils ahead of the fork, whilst topo IV was found to favour removal of the precatenanes formed behind [114].

1.3.2.1.2 DNA topoisomerase IV

DNA topoisomerase IV (topo IV) was discovered in 1990 as a gyrase homologue, which was fundamental to chromosome segregation and cellular partitioning in *E. coli* [115]. It is composed of two ParC (homologous to GyrA) and two ParE (homologous to GyrB) subunits, with considerable structural similarity to gyrase, and was shown to relax positive and negative

supercoils and decatenate, in an ATP/Mg^{2+}-dependent manner [116]. Whilst topo IV shares numerous functional domains with gyrase (Figure 9), including the WHD, TOPRIM, GHKL, tower, coiled-coil and transducer domains, it is unable to supercoil DNA. As discussed previously, negative supercoiling by gyrase arises from its ability to wrap DNA (~130 bp) and present a T-segment with a positive chirality over the G-segment, prior to strand-passage [11, 87]. The topo IV ParC CTD, whilst similar to the GyrA CTD, has a more degenerate structure, with five β-strands rather than six, and the absence of the GyrA-box. Structurally, this suggested that topo IV lacks the ability to supercoil DNA, which was later confirmed biochemically [116]. However, as the outer surface of the ParC CTD is positively charged, this suggests a role in DNA binding and it is thought to mediate topo IV substrate specificity [117]. Indeed, when the activity of topo IV missing the ParC CTD is compared to wild-type, a clear reduction in rate is seen for the relaxation of negative supercoils (~10-fold), positive supercoils (~100-fold) and decatenation (~100-fold) [117]. Topo IV is known to preferentially act on positive supercoils (as opposed to negative) with a ~20-fold increase in rate [118], corroborated by single-molecule assays with the relaxation rate ~25-fold faster on positive supercoils [119]. Using a single-molecule magnetic tweezers DNA braid assay, the DNA crossing angle preference for topo IV was shown to be 85.5°, and while this crossing angle appears more frequently in positive writhe than negative (~3-fold) [119], it is not enough to explain the stark difference in relaxation rates [120]. It was then demonstrated that the CTD of ParC permits the highly processive relaxation of positive supercoils, whilst negative supercoil relaxation is perfectly distributive [120], hence providing an explanation for the preferential relaxation of positive writhe as a combination of both a crossing angle preference below 90° and DNA-binding activity of the ParC CTDs, the latter playing a larger role. The presence of the CTDs also dramatically increases the decatenation activity of topo IV, and together with preferential positive supercoil relaxation, suggests that the CTDs recognise DNA geometries more common in positive supercoils and catenanes than in negative supercoils, and may potentially be able to manipulate DNA geometry in order to achieve the preferential DNA crossing angles [117]. If the GyrA CTD is removed, the enzyme becomes unable to supercoil and has enhanced

decatenation activity, much like topo IV [121].

As DNA topo IV was identified in *E. coli* with partition defects, it was hypothesised early on that the main *in vivo* role for topo IV was decatenation and chromosome segregation [115]. This idea has been built upon since with numerous lines of evidence suggesting this is indeed the major role for topo IV within the cell. Using an *in vitro* replication system, topo IV was highly efficient at unlinking replicated daughter chromosomes, whilst gyrase failed to do so under all conditions tested [122]. So, even though gyrase does have limited ability to decatenate *in vivo* [123], it is only around 1% of the efficiency capable by topo IV [124]. Using NGS, topo IV binding/cleavage was also found to be specifically enriched at the dif site, where *E. coli* chromosomes are unlinked [125]. However, as mentioned previously, topo IV is also highly efficient at relaxing positive supercoils *in vitro* [117, 118], which has led to consideration that topo IV plays a role during replication and transcription, relaxing positive supercoils ahead of the protein machinery. It has been found that topo IV over-expression from a plasmid [126] or ParC/ParE gene locus duplications [127], can rescue cells with topo IA mutations. More recently, expression of ParC and ParE was shown to be amplified in *Bacillus subtilis* harbouring topo IA suppression mutations, with topo IV able to functionally replace all type IA topos in *B. subtilis* [128]. In addition, it was shown topo IV can, to an extent, support replication and transcription fork progression *in vivo* using cells encoding temperature-sensitive mutants of DNA gyrase [129]. Evidence such as this would indeed support the idea of topo IV providing functional redundancy, particularly for topo IA, and less so, gyrase, in certain organisms. However, it seems that the preferential decatenation activity of topo IV *in vivo,* arises from a combination of protein-protein recruitment via the *E. coil* SMC (structural maintenance of chromosome) complex, known as MukBEF [130-133], and temporal regulation [134]. These regulatory factors, alongside topo IVs intrinsic efficiency during decatenation, makes topo IV the major decatenation enzyme within bacteria.

1.3.2.1.3 DNA topoisomerase II

DNA topoisomerase II (topo II), structurally akin to its prokaryotic counterparts (albeit with the

A and B subunits fused into a single polypeptide chain), is the major type IIA topoisomerase found in eukaryotes and is vital to cellular vitality and survival. It is present in both lower eukaryotes, such as *Saccharomyces cerevisiae* (yeast) [135, 136] and *Drosophila melanogaster* [137], as well higher eukaryotes, such as *Xenopus laevis* [138] and humans [139, 140]. In higher eukaryotes, topo II arises as two isoforms, named topo IIα and topo IIβ, which have distinct cellular functions [141], but either can complement the loss of topo II in yeast [142, 143]. Although all topo II enzymes have the critical functional domains described in 1.3.3.1 (Figure 1.9), the structural variations, as well as variable regulation, has led to functionally non-redundant proteins, e.g. if you knock out topo IIα, topo IIβ cannot support DNA metabolism alone [144].

Yeast topo II, one of the more thoroughly characterised, is a homodimeric enzyme capable of relaxing positive and negative supercoils, decatenation and unknotting, using ATP and Mg^{2+} [145]. By employing yeast harbouring temperature sensitive mutations in topo II, it was found that absence of topo II was only lethal when the cells initiated mitosis, and this effect could be mitigated using compounds that block cell cycle progression [146]. This implied that the fundamental role of topo II in yeast is chromosome segregation and unknotting. Yeast topo II has also been demonstrated to play a key role in the transcription of long genes (>3kb), with its absence causing a stall in fork progression which cannot be rescued by topo I alone [147]. Topo II was found to be insensitive to commonly used type IIA prokaryotic inhibitors such as oxolinic acid, nalidixic acid and novobiocin [145], hence these compounds and their derivatives have had clinical success as antibiotics. Yeast topo II has also proved to be a useful screening tools in the quest to find human topo II inhibitors for use as chemotherapeutics and to explore drug resistance mutations [148-150].

Structurally, crystallised yeast topo II has led to significant advancements in the understanding of type IIA topos. One such example was the crystallisation of yeast topo II amino acids 408-1177, in complex with a 30 bp G-segment [56] (Figure 1.11A). It revealed how topo II bends the G-segment, producing a ~150° angle, the duplex ends extending towards the GHKL domain. It also revealed that the DNA bound to the DNA-gate is A-form, rather than B,

further demonstrating how topo II markedly deforms the G-segment upon binding. In 2012, another yeast topo II structure was presented in which a fully functional and minimally truncated topo II (amino acids 1-1177) was crystallised in complex with a 30 bp G-segment and ADPNP [151] (Figure 1.11B). It revealed that the GHKL domains sit atop the G-segment binding interface, with a 45° offset. For the first time the K-loop was revealed, which is a typically disordered (in previous structures) stretch of amino acids containing six consecutive lysine residues, housed within the transducer domain. It was shown to interact with the G-segment, upon which its structure became ordered enough to resolve. This then led to the mutagenesis of various lysines to either alanine or glutamate, to establish its role in topo II activity. It was found that the KKKK to AAAA mutation resulted in 75% reduction in relaxation activity and the KKKK to AEEA mutation resulted in a 90% reduction [151], the latter reduction likely due to the reversal of the electrostatic charge of the K-loop in the most highly conserved lysine positions. The analogous investigation for human topo IIα was carried out (however only three conserved lysines are present in human topo II), with both alanine and glutamate substitutions causing dramatic loss of activity, the latter again having a more severe effect. Decatenation activity was also dramatically reduced whilst cleavage levels remained similar to wild-type, therefore implicating the K-loop in strand passage, post G-segment cleavage, by preventing DNA-stimulated ATPase activity, further consolidated using ATPase assays [151].
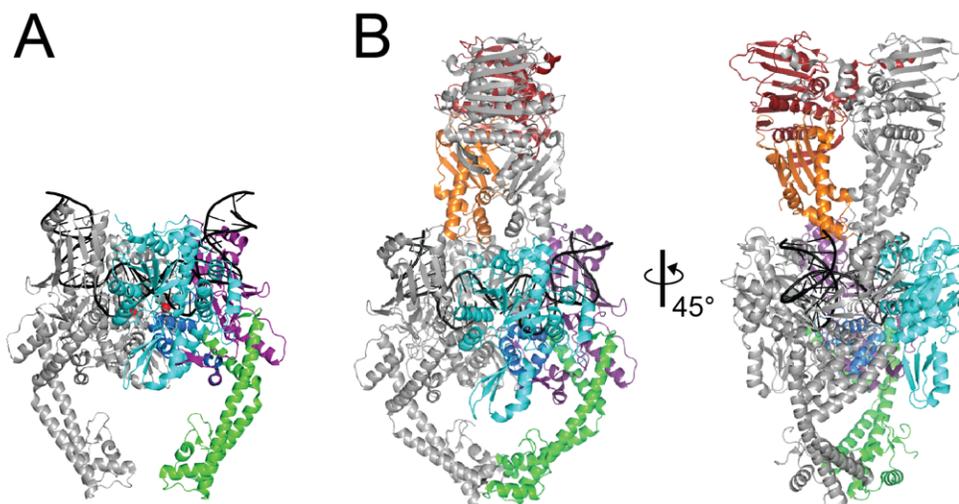
Figure 1.11: Structures of *S. cerevisiae* DNA topoisomerase II. A: DNA topo II (AA$_{408-1177}$) bound to G-segment DNA. PDB: 2RGR [56]. B: DNA topo II (AA$_{1-1177}$) bound to G-segment DNA and

ADPNP. PDB: 4GFH [151]. Domains are coloured as follows: GHKL is red, transducer is orange, TOPRIM is teal, WHD is blue, tower is purple, coiled coil is green, and DNA is black. The second Top2 subunit is grey.

A previously mentioned, higher eukaryotes encode two isoforms of topo II, namely topo IIα (180-kDa) and topo IIβ (170-kDa), and whilst they share similar *in vitro* biochemical activities to both each other and yeast topo II, topo IIα and β perform distinct cellular roles and are controlled by separate regulatory mechanisms [152]. Topo IIα is critical during cellular proliferation [153], with expression low during G_1, rising through S-phase and greatest during G_2/M-phase [154]. Using NGS, it was shown in human cells that topo IIα activity is higher in regions of compact genome, with activity correlating to gene length and transcription rate [155]. *In vitro*, topo IIα is more active in the relaxation of positive supercoils, than negative [156], which is dependent on the presence of the CTDs [157]. Using a single-molecule magnetic tweezers approach, wild-type topo IIα relaxed positive writhe twice as fast as negative but was roughly equal in processivity [158]. However, if the CTDs were removed, this chiral discrimination was reversed with ~1.4-fold increase in negative writhe relaxation over positive, whilst processivity again remained constant, indicating the CTDs control the sign of the chiral preference, not the processivity. This is in contrast to topo IV, as in the absence of the ParC CTDs, topo IV's chiral preference is removed through abolishing the enzymes processive nature during positive writhe relaxation [117]. This demonstrates the major evolutionary divergence between topo IIα and topo IV, with chiral discrimination mechanisms arising separately [158], which, interestingly, is absent for yeast topo II as it demonstrates no chiral preference in the relaxation of positive and negative writhe, with a preferred DNA crossing angle of 90° [120].

Topo IIβ expression patterns differ in comparison to topo IIα, with topo IIβ staying relatively constant throughout the cell cycle [154], but with variable levels of expression during fetal development and in different cell types [153]. Topo IIβ knockouts are lethal owing to disruption of neuronal differentiation [26]. Conditional knockouts implicated topo IIβ activity in retinal development [159] and ovulation [160], and the use of poisons showed it played a role

in spermatogenesis [161] and lymphocyte activation [162]. This has led to a picture of topo IIβ as an important factor in tissue-specific development and cellular differentiation, a concept elaborated upon using ChIP-seq [163]. It was found that topo IIβ interacts with the transcriptional regulator CTCF and the cohesion complex and acts on DNA located at the borders of chromosomal domains as well as conserved transcription factor binding sites. Topo IIβ activity cannot functionally replace topo IIα, unless topo IIβ CTDs are replaced with those from α, indicating the CTDs play a significant role in preferential *in vivo* activity [164]. Topo IIβ further diverges from topo IIα, with *in vitro* relaxation activity being highly distributive, differentially sensitive to inhibitory compounds, and more thermostable [140].

1.3.2.2 Type IIB DNA topoisomerases

The type IIB topoisomerases, despite also employing ATP/Mg^{2+}-dependent double-stranded DNA cleavage, are separated from type IIA based on structural and evolutionary grounds. The two enzymes within this class are DNA topoisomerase VI (topo VI) and DNA topoisomerase VIII (topo VIII). Arguably, they have the simplest protein architecture possible for a type II topo, with only two protein interfaces (rather than the three found in type IIA), meaning that only one interface is intact at any one time to prevent the formation of stable DSB. Therefore, whilst type IIB topos are also hypothesised to follow the general strand-passage reaction (explored in Figure 1.8 and 1.3.2.1.1), it is believed that activity must be much more highly coordinated in order to prevent the excessive formation of DSB. Topo VI, being the main focus of this research, is explored in depth in a subsequent section (see Section 1.4).

1.3.2.2.1 DNA topoisomerase VIII

The most recent topoisomerase family discovered, in 2014, was the second member of the type IIB topoisomerases, named DNA topoisomerase VIII (topo VIII) [165]. Identified through database screening using the B subunit of *Sulfolobus shibatae* topo VI, topo VIII was classified as a highly divergent type IIB, in which the Bergerat fold (GHKL ATP-binding domain) and TOPRIM domain were contained within a single polypeptide chain, along with the H2TH and

transducer domains, which are characteristic of type IIB topos. Currently, 77 topo VIII enzymes have been identified in nine bacterial phyla, four in archaea, all within the euryarchaeota phylum, and one still unclassified [166]. Along with the dramatic sequence divergence from topo VI, being a homodimer, and being more common in bacteria than archaea, topo VIII is further distinct from topo VI due to its encoding being dependent on plasmids and integrated elements [166]. Of the topo VIII enzymes currently characterised biochemically, two exhibited Mg^{2+}-dependent relaxation of positive and negative supercoils (*Microscilla marina* and *Paenibacillus polymyxa*), however the *M. marina* topo VIII performed these reactions independent of ATP, a behaviour not typical of type IIB topos [165]. Topo VIII from *Ammonifex degensii,* in the presence of SDS and proteinase K, only demonstrated cleavage activity, however this was also ATP-independent. In all cases the activity wasn't strong, and moreover, many of the topo VIII enzymes found seem to be in a state of inactivation. So whether topo VIII plays a considerable role, if any, in the organisms in which it arises, is still unclear. In those cases where the topo VIII is encoded within a plasmid, other topos including topo VI and topo IA have also been detected, indicating that these enzymes may be important in the DNA metabolism of the plasmid itself [166].

1.4   DNA Topoisomerase VI

1.4.1 Overview

The principal enzyme, and until recently, the only enzyme, of the class type IIB is DNA topoisomerase VI (topo VI), the subject of this project. Topo VI was initially identified in the archaeal hyperthermophile, *Sulfolobus shibatae,* and reported as the first member of the type IIB topoisomerase family [167]. It has since been found throughout the archaea, a few bacterial species, and intriguingly, in eukaryotes such as plants and algae [168, 169]. Topo VI has an A_2B_2 heterotetrameric subunit arrangement with minimal homology to the type IIA topoisomerases, mainly found in the DNA-binding, magnesium-chelating and ATPase domains [64, 170]. The A subunit (Top6A) is homologous to the eukaryotic meiotic recombination factor, Spo11, which is a key protein involved in meiosis [171, 172]. Until recently, no B-subunit

Chapter 1: Introduction

(Top6B) homologue had been identified, until the focus was shifted from the gene sequence to the protein structure. Structural homologues of Top6B have been identified via bioinformatic analysis in various eukaryotes, including mouse and *A. thaliana*, and have been shown to interact with Spo11β [173, 174]. The link between topo VI and these eukaryotic meiosis proteins will be explored in further detail below. The domain organisation of topo VI differs from type IIA in that the WHD domain and TOPRIM domains are contained within the same polypeptide, Top6A. Furthermore, topo VI is thought to represent the simplest form a type II enzyme can have, with only two protein interfaces (instead of the three found in IIA), which open and close alternatively. This makes topo VI of keen interest in the dissection of the strand-passage reaction, particularly the role of ATP in the opening/closing of the N-gate [64].

Whilst topo VI is distinct from the type IIA topos, it also bears remarkable similarity in key domains which are responsible for topology manipulation, such as ATP, Mg^{2+} and DNA-binding regions. This, along with the structural simplicity of the enzyme, makes topo VI a marvellous candidate for in depth characterisation, which may provide not only insights into the workings of type IIB, but also IIA and eukaryotic meiotic factors.

1.4.2 Plant topoisomerase VI

Genomic analysis of *Arabidopsis thaliana* revealed three distantly related homologs of the Top6A/Spo11 subunit (AtSPO11-1, -2 and -3) and one of the Topo6B (AtTOP6B) subunit [175, 176]. These genes have since been bioinformatically identified within various plant species, including *Oryza sativa* (rice), *Medicago truncatula* and *Lycopersicon esculentum* (tomato). In *A. thaliana*, the AtSPO11-1 protein is vital for efficient meiotic recombination and the production of viable gametes, indicating its role is homologous to that of Spo11 from *Saccharomyces cerevisiae* [177]. AtTOP6B and AtSPO11-3 have high levels of expression and the proteins have also been demonstrated to interact. AtTOP6B has also been shown to interact with AtSPO11-2 (although AtSPO11-2 is not highly transcribed) but not AtSPO11-1 [175]. *A. thaliana* mutants, homozygous for null mutations in either AtTOP6B or AtSPO11-3, display identical phenotypes, with severely growth stunted plants failing to live much past 4-5 weeks

[178]. Furthermore, double mutants displayed a phenotype identical to that of either single mutant, demonstrating the proteins likely function in the same process, or even the same protein complex. This study suggested that the topo VI homologues play a key role in DNA replication/metabolism and cell growth. A process called endoreduplication, in which the genome is replicated multiple times in the absence of cellular division, is critical for plant growth to enlarge cells of the hypocotyl and leaves [179]. Endoreduplication is deficient in the AtSPO11-3/AtTOP6B mutants and can explain the dwarf aspect of the phenotype. Sugimoto-Shirasu *et al*. [169], found that mutations in either AtSPO11-3 or AtTOP6B did indeed halt endoreduplication, however the phenotype was not quite as severe as Hartung *et al*. had found [178]. The plants also produced fertile seeds, indicating that the topo VI homologues do not play such a central role in DNA metabolism (at least in the particular ecotype analysed) as previous work had suggested, and their primary role is likely during endoreduplication. However, *A. thaliana* possesses both topoisomerase II and DNA gyrase homologues, both of which are type IIA topoisomerases capable of the same reactions as topo VI, namely DNA decatenation, unknotting and relaxation. So why *A. thaliana* requires topo VI (AtSPO11-3 and AtTOP6B) exclusively during endoreduplication isn't clear. One possibility is that it functions to resolve endoreduplication specific DNA intermediates, however no such DNA species have yet been found. Topo VI may also be under distinct regulation and so expressed specifically during endoreduplication [180].

An important difference between archaeal and plant topo VI, is the identification of two other proteins that seemingly interact with plant topo VI to form the functional complex, which haven't been found in archaea, called BIN4 (brassinosteroid-insensitive4) and RHL1 (roothairless-1) [181, 182]. Mutants in either of these proteins results in plants with severely restricted growth due to the inability to complete the endoreduplication cycle. The BIN4 protein binds DNA and in the absence of a functional version, DNA damage response pathways are activated and endocycles are prematurely arrested leading to the prevention of cell expansion [182]. The RHL1 protein, likewise has DNA binding activity and when mutated, successive rounds of genome replication during endoredulplication is prevented after the first 2 endocycles

[181]. RHL1 shows sequence similarities to the topo IIα C-terminal domain, and whilst the function of this domain isn't fully understood, both it and RHL1 contain putative phosphorylation sites and the PEST sequence (thought to be involved in rapid protein degradation [183]. And, in addition, it is the C-terminal domain of topo IIα, which varies most from the IIβ isoform, indicating a potential role in differential regulation [184]. So, whilst it is plausible that RHL1 may be directly involved in the topo VI chemical reaction, its main function may possibly be regulation of the complex's activity [181].

Plant topo VI has also been shown to modulate the expression of numerous genes, such as those regulated by plant steroid hormones, brassinosteriods, implicating topo VI in plant growth and development via an alternative route to endoreduplication defects [185]. Another study implicated plant topo VI in cellular growth and development, through a role in chromatin organisation and transcriptional silencing [186]. Mutants in the *mid* gene, encoding the MIDGET (MID) protein, exhibited identical phenotypes as mutants in *rhl1*, *rhl2* and *top6B*, with MID demonstrating a physical interaction with RHL1, and that the absence of MID results in DNA damage check-point activation. Furthermore, *A. thaliana Top6B* mutants have been shown to have defects in epidermal development, producing dysfunctional root hairs and trichomes; and carbon partitioning, with cell walls becoming inundated with pectins and other proteins, and the matrix accumulating callose and starch [187]. The rice Top6A and B subunits were overexpressed in *A. thaliana*, with the resultant phenotype involving reduced abscisic acid (ABA) sensitivity, high salt tolerance and dehydration resistance [188]. A single base substitution in *A. thaliana topo6A,* caused defects in the reactive oxygen-species response, thought to be due to the failure of top6A binding in the activation of critical genes [189]. Together, these data indicate that topo VI likely plays a role in mediating the plant's response to endogenous and exogenous cues, integrating them through chromatin remodelling and transcriptional control, and is therefore an important determinant of cellular development.

1.4.3 Plasmodial topoisomerase VI

For plasmodia, the literature is presently in disagreement as to the presence of a functional topo VI within the genome. A review on *Plasmodium* genomics from 2003, claimed that plasmodia encoded both a Top6A and Top6B subunit based on homology to domains and subunit arrangement also found in archaeal topo VI proteins [190]. For instance, Top6A was identified due to the presence of both a WHD domain, harbouring the catalytic tyrosine residue, followed by a TOPRIM domain. It is true that Top6A has this architecture, but it is also present in Spo11. A subsequent paper from a different group, interested in the presence of Spo11 in the genomes of eukaryotes, stated that whilst plasmodia (apicomplexans) contained a Top6B homologue, it did not encode Top6A, only Spo11 [191]. The paper only speculated that Top6A may have been lost over time and as the Top6B is highly divergent from archaeal Top6B, with only N-terminal homology in the GHKL domain, that perhaps it functions independently of Top6A. The situation is further confounded when in 2016, a group claimed that topo VI from *Plasmodium falciparum* can complement the function of topo II in *Saccharomyces cerevisiae* when expressed transiently and that it is sensitive to the ATPase inhibitor radicicol [192].

If topo VI is present in plasmodia, it is hypothesised that it may play a potential role in schizogeny, asexual reproduction via numerous genome replications followed by cytokinesis, and could have promise as a novel antimalarial drug target. However, clearly substantially more research is required.

1.4.4   Archaeal topoisomerase VI

Most of the work previously carried out on archaeal topo VI comes from characterisation of the enzyme from the hyperthermophillic archaeon, *S. shibatae*. Other archaeal species analysed include the hyperthrmophillic *Methanococcus jannaschii* and the mesophilic *Methanosarcina mazei*. The *S. shibatae* topo VI (SsT6) is an A_2B_2 hetrotetramer with an A subunit of 47 kDa (Top6A) and a B subunit of 60 kDa (Top6B). It was shown to relax both positive and negative supercoils in the presence of ATP and Mg^{2+}, have no supercoiling activity, and to be more efficient in decatenation than in relaxation [167]. Double stranded DNA cleavage by SsT6 was

found in the presence of both nucleotide (ATP or ADPNP) and divalent cation (Mg^{2+} or Ca^{2+}), but in the absence of either nucleotide or cation, no cleavage was detected [193]. The optimal conditions to experimentally stabilise the cleavage complex of SsT6 in the absence of compounds was shown to be in the presence of 10 mM $CaCl_2$ and 1 mM ADPNP [194], with the enzyme becoming covalently bound to DNA at the 5' phosphate, as for type IIA topos [193, 194]. SsT6 has also been shown to generate a 2-base stagger during DNA cleavage, rather than the 4-base found in type IIA topos, and whilst no specific cleavage site sequence has yet been described, a potential preference for A and T nucleotides has been reported [194].
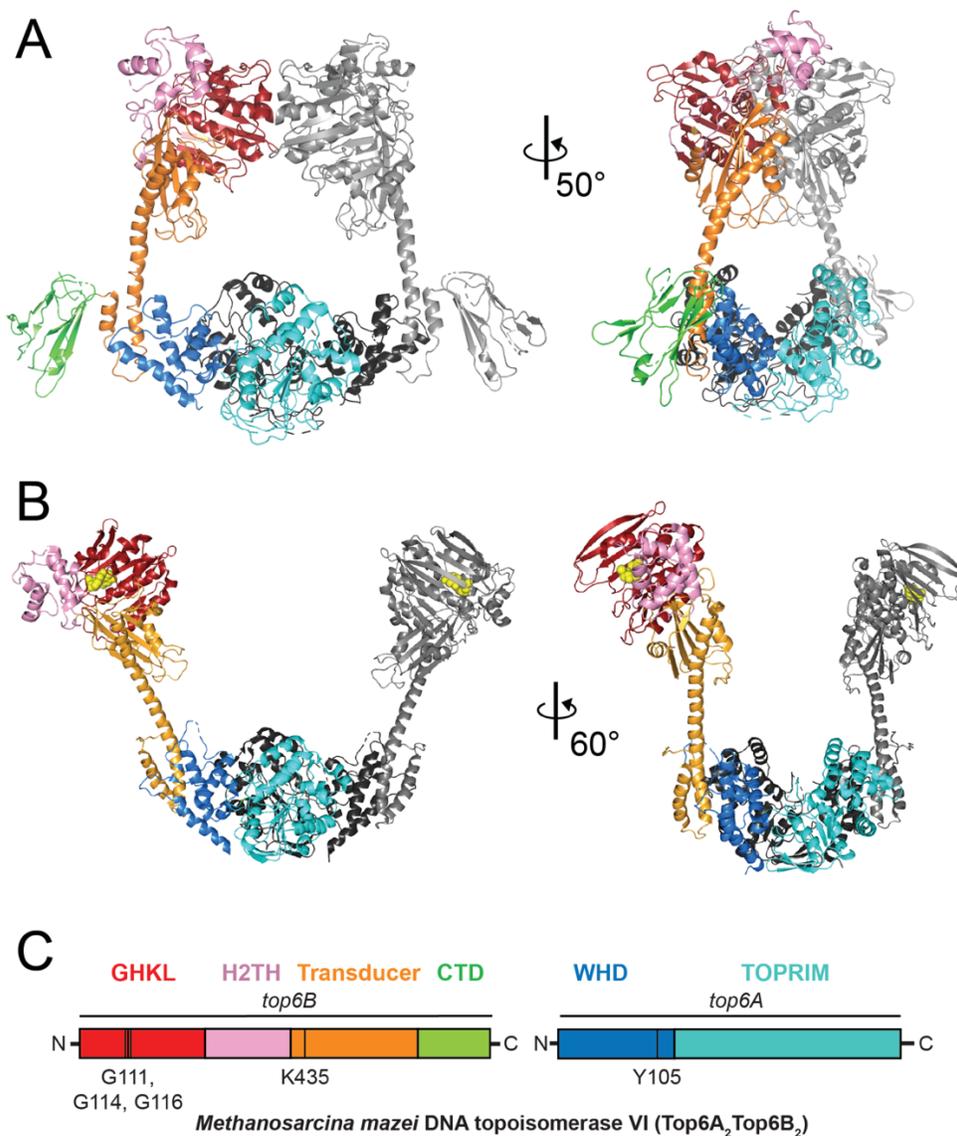
Figure 1.12: Protein structures and gene domain organisation of DNA topoisomerase VI (topo VI). A: *Methanosarcina mazei* **topo VI. PDB: 2Q2E [195]. B:** *Sulfolobus shibateae* **topo VI. PDB: 2ZBK**

[196]. C: Gene domain organisation of *M. mazei* topo VI. Domains are coloured throughout as follows: GHKL is red, H2TH is pink, Transducer is orange, Top6B CTD is green, WHD is blue and TOPRIM is teal. The other Top6B and Top6A subunits and light and dark grey, respectively. Vertical black bars in C represent key catalytic residues.

Structural work on archaeal topo VI has been carried out on the *S. shibatae*, *M. jannacshii* and *M. mazei* enzymes, revealing the domain architecture and quaternary structure of Top6A and Top6B, as well as the heterotetramer in a closed (Figure 1.12A) and open (Figure 1.12B) conformation. A truncated version of *M. jannaschii* Top6A, containing the DNA binding core, was crystallised and resolved at 2.0-Å resolution, shedding light on the DNA-cleavage mechanism of topo VI and revealed structural motifs shared between type IIA and IIB topos [170]. Namely the N-terminal winged-helix domain (WHD) that harbours the catalytic tyrosine residue and the C-terminal TOPRIM domain, previously referred to as the Rossman-like fold, which chelates divalent cations, specifically Mg^{2+}. In topo VI, this is achieved via conserved acidic residues Asp249, Asp251, and Glu197 (*M. jannaschii* numbering). Both oxygens of Asp251 are hydrogen-bonded to two water molecules, along with another two observed water molecules that complete the octahedral coordination geometry necessary for Mg^{2+} binding. Analysis of the structure of the Top6A homodimer revealed the formation of a positively electrostatic groove, 18 Å in diameter, seemingly capable of accommodating a DNA duplex. Furthermore, the catalytic tyrosine and the TOPRIM domain face the interior of the groove, plausibly in position to interact with a captured G-segment.

The Top6B subunit, responsible for the binding and hydrolysis of ATP was crystallised from *S. shibatae*, as both an un-liganded monomer and a nucleotide-bound dimer with three key domains described [64]. The N-terminal ATP-binding fold of Top6B was found to be a member of the GHKL family of ATPases (**G**yrase, **H**sp90, Histidine **K**inase, Mut**L**), a shared feature with the type IIA topos. Despite <15 sequence identity between the GHKL domains of Top6B and the B subunit (GyrB) of *E. coli* gyrase (type IIA), the overall structure of the GHKL core 112 residues was found to be very similar with a C_α r.m.s.d of only 1.17 Å. The second Top6B domain is called the helix-two-turns-helix (H2TH) motif and contains 4 α-helices. It is

found in all the type IIB enzymes, archaeal and eukaryotic, and is not present in any type I or type IIA topos. It is thought to play a role in G-segment bending and recognition, and coupling G-segment cleavage to strand-passage [62]. The third domain is known as the transducer domain, and remarkably again, shows high structural similarity with the equivalent domain of *E. coli* GyrB, giving a C_α r.m.s.d of only 1.54-Å despite <11% sequence identity. It has been proposed that for both topo VI and gyrase, this domain is key in transducing energy from the ATPase domain to the DNA binding/cleavage domain. This structural data demonstrated the increasing likelihood that type IIA and IIB topos are directly related in terms of evolution and despite being so structurally distinct, key mechanistic motifs have been highly conserved between the two families. It is known that the GHKL domains of type IIA topos and MutL proteins dimerise in response to ATP binding, and in the case of type IIA topos this dimerisation of the N-gate is key to the strand passage reaction, allowing the DNA gate to open without the BA dimers dissociating and producing genotoxic DNA breaks [197]. This dimerisation also occurs when nucleotide is bound to the Top6B subunit, and was demonstrated crystallographically using ADPNP [64]. This began to indicate that, although the structure and sequence of topo VI was quite distinct from type IIA topos, its mechanism likely follows a similar path to the strand-passage mechanism of type IIA topos. For a visual representation of topo VI gene domain organisation, see Figure 1.12C.

The enzyme this research project centres around is topo VI from the mesophilic archaeon *M. mazei* (MmT6). Multiple reasons make MmT6 an attractive candidate for topo VI characterisation, including the ability to express it in *E. coli* and an optimum activity temperature of ~37°C. The optimum temperature for activity of *S. shibatae* topo VI is ~80°C, making assays troublesome to perform. Its suitability was further consolidated through the full-length crystallisation of the MmT6 heteroteramer to 4-Å resolution (Figure 12A) [195]. This was done in the absence of nucleotide, with an AB dimer occupying each asymmetric unit of the crystal and the heterotetramer recapitulated through a crystallographic two-fold rotation axis. The heterotetramer appears to form a rough square with overall dimensions of ~170 Å x 130 Å x 70 Å and an internal, electrostatically positive cavity of adequate size to accommodate 2 DNA

38

duplexes. The Top6A/Top6B dimer interface is formed from the tight association of a helical hair pin that extends from the Top6B transducer domain, with two α-helices found in the N-terminal region of Top6A.

The Top6A subunits form a dimer, creating a 20-Å wide, positively-charged channel, in which the active site tyrosine of the WHD domain and the Mg^{2+}-binding centre of the TOPRIM domain of each Top6A are in position to interact with a G-segment. It is likely that this Top6A dimer, in association with the Top6B dimer, cleaves the G-segment to produce a DSB in which the 5' end of each strand is covalently bound to the tyrosine (Y105) of each Top6A. This structural data is in agreement with that obtained from the crystallisation of the truncated DNA binding core of Top6A, explored above [170]. In the MmT6 structure, the less well conserved, N-terminal 70 amino acids of Top6A can now be seen, including an α-helix (amino acids 17-37) that, along with the first α-helix of the WHD (amino acids 71-94), forms the binding interface with Top6B. These two α-helices are linked by a region of disordered amino acids (44-67), conserved amongst topo VI homologues as well as Spo11, which contains two arginine residues (Arg44 and Arg67) and one asparagine residue (Asn48), and appears to jut out over the DNA-binding groove and may interact with the G-segment (Figure 1.13B).

As previously discussed, the Top6B subunit of topo VI has an N-terminal GHKL domain responsible for the binding and hydrolysis of ATP. This is followed by the H2TH domain, and the transducer domain. The transducer moves relative to the GHKL domain upon ATP binding/hydrolysis and transmits this movement to the DNA binding core of Top6A. The movement of the transducer domain is mediated by a conserved lysine residue (Lys435 in MmT6 and Lys427 in SsT6) that reaches into the GHKL domain and interacts with the γ-phosphate of the bound nucleotide via a hydrogen bond (Figure 1.13A). Once the nucleotide is hydrolysed and the phosphate released, the lysine residue turns away from the active site, returning the transducer domain back to its relaxed state [67]. Following the transducer domain is a C-terminal domain (CTD), which is absent in the SsT6 Top6B [64]. It adopts an immunoglobulin-type fold, placing it in a functionally diverse protein family, with key roles in the binding of macromolecules, typically in the form of carbohydrates and proteins. The complete

deletion of the MmT6 CTD modestly reduces activity *in vitro* but does not change the enzyme's preference for different DNA substrates. Interestingly, the few archaeal species in which the Top6B CTD is present, are those which also possess DNA gyrase. It has been hypothesised, therefore, that the CTD may be involved in regulating topo VI spatially, localising it to certain cellular compartments [195].
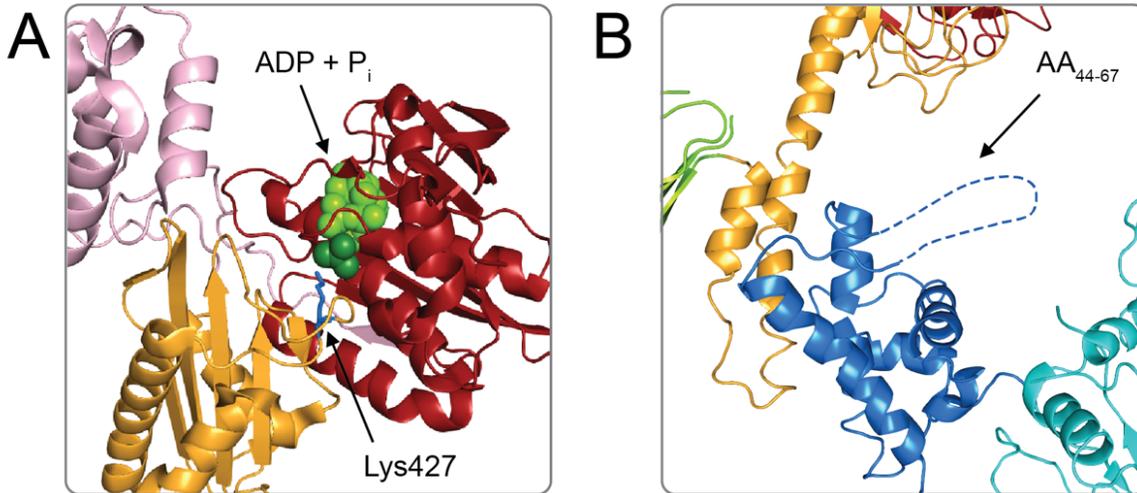
Figure 1.13: Close ups of DNA topoisomerase VI structures. A: *S. shibatae* topo VI close up of the GHKL (red) domain bound to ADP and phosphate. The 'switch lysine' present in the transducer domain, Lys427 (435 in *M. mazei*) can be seen reaching into the ATP-binding pocket. PDB: 2ZBK [196]. B: Close up of the interface between Top6A and Top6B, from *M. mazei.* Two α-helices from the transducer domain of Top6B, interact with two α-helices from the WHD of Top6A, forming a highly-stable non-covalent interface. The blue dotted line represents the conserved and disordered amino acids 44-67, believed to form a loop that extends over and interacts with the bound G-segment. PDB: 2Q2E [195].

Along with solving the structure of the MmT6 holoenzyme, the work done by Corbett *et al.*, 2007 [195] also delved further into the ATP-mediated conformational changes of the SsT6 enzyme using small-angle X-ray scattering (SAXS) analyses, complementing and adding to previous crystallographic analyses of the Top6B subunit in the apo-state and bound to AMP-PNP, ADP·AlF$_4^-$, ADP·P$_i$, and ADP [64, 67]. This work gave great insight into the ATP-mediated DNA cleavage reaction of topo VI, making it possible to model the strand passage mechanism by which the enzyme likely functions. The 2-gate model of strand passage, used to describe the working of all type II topos, involves the formation of two protein interfaces that

alternately open and close in the presence of ATP/Mg^{2+} to allow the capture and passage of a T-segment through a cleaved G-segment before its re-ligation [198]. The MmT6 holoenzyme structure substantiates the 2-gate model, by demonstrating the existence of two protein interfaces that can alternately open and close to allow the capture of two DNA segments. In the absence of ATP, the Top6A subunits are dimerised, whilst the Top6B subunits remain open, allowing the capture of the G- and T-segments. Upon ATP binding the Top6B subunits dimerise, forming a strong protein interface. Due to the Lys345 residue mentioned above, ATP binding is coupled to transducer domain movement, which transmits this conformational change to the WHD of Top6A. This is thought to cause the catalytic tyrosine of the WHD to become correctly positioned in the vicinity of the Mg^{2+} binding centre of the TOPRIM domain, forming the bipartite active site capable of G-segment cleavage. The subsequent hydrolysis of an ATP to ADP + P$_i$ then to ADP (P$_i$ release), explored using the SsT6 B subunit structures, showed an 11° outward rotation of the transducer domain [67]. If the Top6B/Top6A dimer interface is present, this transducer domain movement should theoretically be able to separate the Top6A subunits, allowing the passage of the T-segment. Hydrolysis and release of the second ATP, should then reset the enzyme, allowing it to undergo another strand passage cycle. A simplified schematic of the topo VI mechanism can be seen in Figure 1.14.

Recently, a study was published detailing extensive biochemical characterisation of the DNA sensing regions of the *M. mazei* Top6B subunit [62]. In this study, site-directed mutagenesis revealed three conserved, positively charged regions within Top6B that are of importance in sensing DNA bends and crossings by MmT6 and tightly coupling strand scission by the Top6A subunit to ATPase activity in the Top6B subunit. The three motifs explored were the H2TH domain, the KGRR$_{(186-189)}$ loop, and the WKxY loop, with key residues of each motif substituted to either alanine or glutamate residues. Using a combination of biochemical and biophysical techniques, their active role in sensing the T-segment and coupling this to ATPase activity was revealed. Compared to wild-type topo VI, mutations in the KGRR and WKxY loops abolished strand-passage, whilst it was significantly impaired by mutations in the H2TH domain. The ATPase activity in the presence of supercoiled DNA was increased for KGRR

mutants, indicating futile cycling (ATP hydrolysis in the absence of strand-passage); roughly equal to wildtype for H2TH mutants; and completely abolished for WKxY mutants. Binding to supercoiled DNA is also moderately affected for H2TH and KGRR mutants, with no binding detected for the WkxY mutant. This research led to the conclusions that the WkxY loop was integral to strong G-segment binding; the KGRR loop (of the GHKL domain) is involved in sensing the T-segment and coupling its presence to ATPase activity; and the H2TH promotes G-segment bending in the presence of ATP, and aids in coupling ATP hydrolysis to DNA cleavage and strand-passage [62].
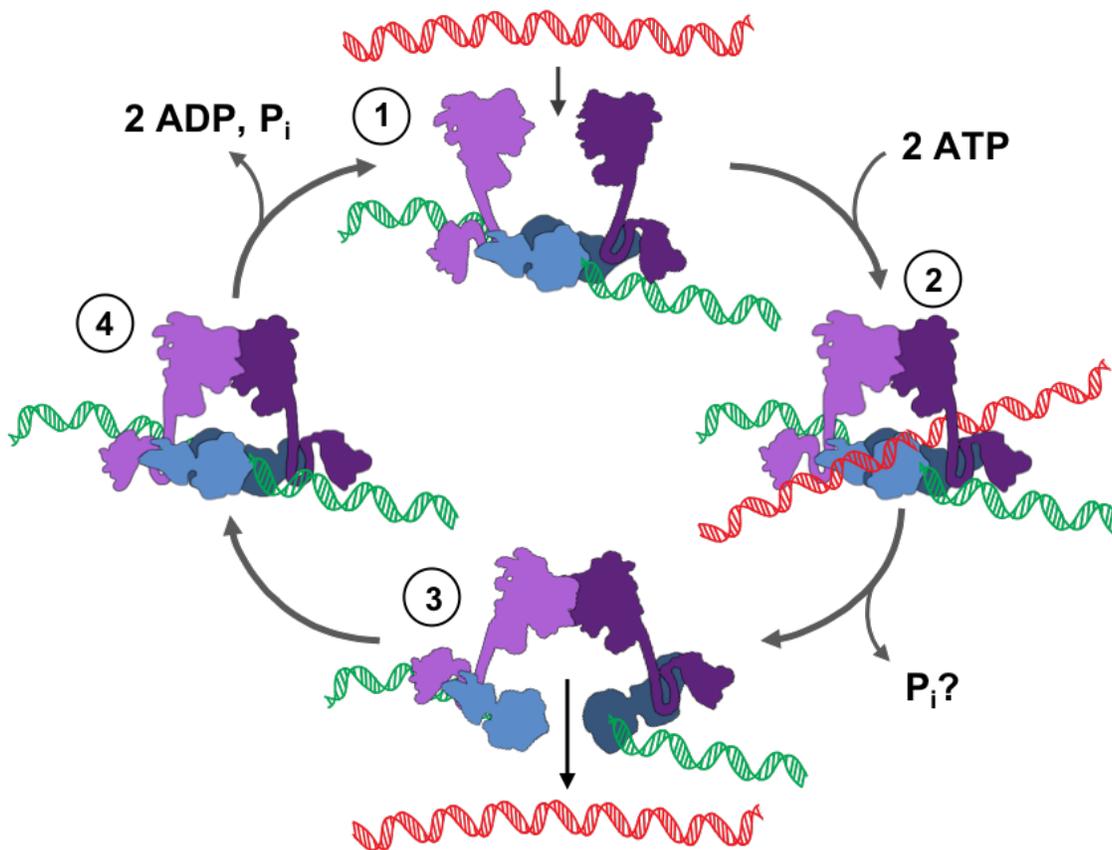
Figure 1.14: Simplified DNA topoisomerase VI reaction mechanism. (1) The heterotetramer binds a G-segment (green) to the Top6A subunits (light and dark blue), and captures a T-segment (red), within the enzymes central cavity. (2) An ATP molecule binds to each Top6B subunit (light and dark purple) causing Top6B dimerisation. (3) ATP is hydrolysed and an inorganic phosphate (P_i) is released which leads to conformational changes causing G-segment cleavage and duplex opening, so as to allow the T-segment to pass through and exit the enzyme. (4) The G-segment is re-ligated and the dissociation of ADP and P_i triggers the opening of the Top6B subunits so that then enzyme can perform a consecutive round of strand passage or disengage from the DNA [62, 195].

1.4.5 Spo11

Spo11, a 45-kDa protein, was first identified in *Saccharomyces cerevisiae* using assays designed to characterise the genomes of yeast deficient in sporulation and therefore meiosis [199, 200]. Spo11 initiates meiotic recombination in prophase I via the formation of double-stranded DNA breaks (DSBs), becoming covalently bound to the 5' phosphate via an invariant tyrosine residue, not only found in all topo VI enzymes but also throughout the topoisomerase family [172]. Mutation of this tyrosine residue to a phenylalanine in Spo11, removes its ability to form DSBs [171]. Once Spo11 generates a DSB it is endonucleolytically removed from the DNA, likely via the Mre11-Rad50-Xrs2 complex, yielding a Spo11-oligonucleotide complex [201]. The resultant double-strand break is processed before undergoing recombination (or, less frequently, repair), thus making Spo11 an important component in the generation of genetic diversity [202]. Spo11 orthologs appear in numerous eukaryotic organisms as a meiosis-specific factor, including fungi, invertebrates, mammals and plants [203, 204].

When the *topo6A* gene was sequenced, the homology to *Spo11* was recognised early on, and the tyrosine responsible for the topo-like activity of the Spo11 protein was determined [171]. Spo11 also generates a 2-base stagger upon cleavage [205], as does topo VI [194], thus strengthening the evolutionary relationship between them. However, Spo11 has been notoriously difficult to express and purify for *in vitro* analysis, hence making characterisation of Top6A particularly important for Spo11 research, which has already led to key insights on the potential catalysis mechanism of Spo11-induced DSBs. For instance, it was found that when Top6A was expressed alone it was extremely insoluble [193] and only when expressed in the presence of Top6B, or purified from the parental organism [167], could considerable amounts of active protein be collected [195]. This indicated that Top6B plays a key role in the stabilisation of Top6A and led to the notion that there may be Top6B homologs also present, which stabilise Spo11 and may also play a role in meiosis. Bioinformatics did not reveal any homologs based on DNA sequence alone, however when the focus was shifted to structural homology, candidates were identified [173, 174]. In two papers, published back to back, Top6B-like

homologs were explored in *A. thaliana* [174]*, S. cerevisiae* (Rec102), *Shizosaccharomyces pombe* (Rec6), *Drosophila melanogaster* (MEI-P22) and mouse [173]. *A. thaliana* encodes 3 Spo11 homologs (AtSpo11-1, AtSpo11-2 and AtSpo11-3), with AtSpo11-3 being the Top6A subunit homolog, necessary for endoreduplication (see 1.4.2). AtSpo11-1 and AtSpo11-2 are believed to play key, non-redundant roles in DSB-formation during meiosis, with mutations in either producing similar phenotypes and leading to the postulation that they function as a heterodimer [177, 206, 207]. In 2016, it was demonstrated that a structural homolog of archaeal Top6B (AtTop6B-like) was present in *A. thaliana,* with its absence causing plant sterility through defects in female gamete development [174]. Using a yeast triple-hybrid assay, AtTop6B-like was important for the formation of the AtSpo11-1/AtSpo11-2 heterodimer. Structurally, AtTop6B shared key conserved domains in the GHKL and transducer domains. In mice, another structural homolog was identified (MmTop6B-like), which was shown to interact with Spo11-β, but not Spo11-α, via a 38-amino acid insertion at the NTD of Spo11-β, which is absent in Spo11-α [173]. In tissues, MmTop6B-like had similar expression profiles to Spo11-β, expressed only in the testes or ovaries during meiosis. Homozygous truncation mutations in MmTop6B-like rendered adult mice deficient in spermatogenesis, with MmTop6B-like being demonstrated to be directly involved with DSB formation during meiosis. Later in 2016, another paper emerged detailing similar findings for a rice top6B-like protein also necessary during meiotic recombination [208, 209]. Very recently, in 2020, a further study was published in which the *Saccharomyces cerevisiae* Spo11 complex was purified from insect cells, after decades of trial and error, achieved only through the identification and dual expression of Top6B-like homologues, Rec102 and Rec104, as well as a third protein, Ski8 [210]. This has allowed in vitro biochemical characterisation of Spo11 to finally begin. The work done on Top6B-like homologs is yet another example of how the characterisation of archaeal topo VI has facilitated breakthroughs in the study of meiosis.

1.4.6   DNA topoisomerase VI inhibition

When it comes to inhibition of type II topos, two key mechanisms have been described. One

44

class of inhibitors are known as cleavage-complex stabilisers, as they bind between the double stranded break and prevent its re-ligation. This leads to stalling of replication machinery and the accumulation of lethal DSBs. This mode of action is used by the prokaryotic type IIA inhibitor class, the quinolones [99, 211], and the eukaryotic anti-cancer drug, etoposide, which targets topo II [212]. The other potent class of topoisomerase drugs are the aminocoumarins. These molecules competitively bind to the ATP domain, preventing the binding of ATP, therefore hindering the ATP-dependent function of the enzyme [213]. The aminocoumarins, whilst very potent *in vitro*, are not used clinically due to their high levels of toxicity [214]. For topo VI, a few inhibitors have been identified including the Hsp90 inhibitor radicicol, which is an ATPase inhibitor [215]. A screen was also performed using topo II inhibitors, and whilst the mechanism of action isn't clear, some topo II drugs did inhibit SsT6. In 2013, a novel microtitre based assay used for the high-throughput screening of drugs against type II topos, was described [216]. Here novel inhibitors against MmT6 were identified. These included m-amsacrine, suramin, hexylresorcinol, 9-aminoacridine, purpurin and quinacrine. M-amsacrine, 9-aminoacridine and quinacrine were shown to prevent DNA cleavage by MmT6, suramin and purpurin blocked DNA binding and hexylresorcinol's mode of action remains unknown. This study also looked at the effects of growing plants in the presence of hexylresorcinol. Plants exhibited symptoms akin to those of the topo VI knockouts, including endoreduplication defects, but it cannot be said with absolute certainty that this is due to the inhibition of plant topo VI as hexylresorcinol may have a number of cellular targets. However, the evidence of hexylresorcinol's *in vitro* inhibition of MmT6 provides indirect evidence that inhibition of plant topo VI may well be the main cause of the phenotype. This demonstrated the potential that MmT6 research has in advancing knowledge of plant, and potentially plasmodial, biology, which are organisms that are important in agriculture and human disease.

1.5 Single-Molecule Methods

1.5.1 Overview

In the past, exploring the function and mechanism of topoisomerases was dominated by

ensemble or "bulk" experiments that have led to a great body of data and understanding on the enzymes' activity. However, the averaging effect that comes from working with many topo molecules in solution has the tendency to obscure the finer mechanistic details. As technologies and scientific understanding increases, the ways in which one can explore at the molecular level are expanding. One key set of techniques being used progressively more for the characterisation of topos, are single-molecule assays. As the name suggest, these assays involve using exceedingly low concentrations of protein to attain data on the activity of a single enzyme reacting with a single substrate. Mapping conformational changes in the *B. subtilus* DNA gyrase in solution has been achieved through single-molecule fluorescence (förster) resonance energy transfer (smFRET) [217]. FRET analysis has shown that the GyrB subunits dimerise in the presence of ATP and reopen post ATP hydrolysis. Also using FRET it was shown that the DNA gate is predominately in the closed conformation during the reaction, providing evidence that DNA gate opening is a rare and brief stage due to the danger of DSBs [218]. Fluorescence has also been used *in vivo*, studying the localisation of DNA gyrase, topo IV and topo I in *Bacillus subtilis* [219]. Using fluorescently-labelled subunits, it was shown that the topo IV subunits are uniformly distributed in the nucleoid, however DNA gyrase and DNA topoisomerase I dynamically associated with DNA replication forks. This indicated an important role for gyrase and topo I in topology maintenance during replication.

In terms of single-molecule *in vitro* techniques that have shown to be well-suited to the study of topos, there are two which allow the precise manipulation of DNA topology, with the resolution to detect single strand-passage events. These are the magnetic tweezers technique and the optical trap. During the course of this research the magnetic tweezers technique has been employed in the characterisation of MmT6, and therefore described in depth below (see 1.5.2 and Chapter 4). Briefly, the optical trap, in this context, involves tethering a DNA duplex to a micron-sized bead which can be manipulated with nano-scale precision using highly focussed light [220]. This set up can vary, for instance, one extreme end of the DNA duplex can be tethered to a slide and the other to a bead which can be rotationally manipulated using light polarisation. Another set up, used in topo research, involves tethering the duplex at both

extreme ends to micron sized beads, one of which is held still in the optical trap whilst the other is rotated by an attached micro-pipette. This assay was implemented in characterisation of topo IV, demonstrating its chiral preference for positive supercoils [119].

1.5.2 Magnetic tweezers

By researching in the lab of Dr. Keir Neuman, the opportunity to explore MmT6 activity using single molecule assays, namely the magnetic tweezers technique, was facilitated. This method provides an elegant way in which the activity of a single topoisomerase on a single DNA duplex can be tracked in real time and with high accuracy. For in depth reviews see cited articles [221-223]. The set-up is relatively simple and consists of a sample cell containing DNA that is tethered to the slide at one extreme end, and to a paramagnetic micron-sized bead at the other, all via noncovalent interactions. Two permanent magnets fixed 1 mm from each other are placed above the sample cell and produce a magnetic field gradient that exerts an upward force on the tethered magnetic bead. The sample cell is placed on top of an inverted oil immersion microscope, which relays the image of the slide to a camera. The magnetic beads produce diffraction rings and the appearance of these rings can be used to measure the distance of the bead from the surface of the sample cell. This allows the extension of the DNA to be tracked, with the measurement of extremely small changes in extension (1-20 nm, depending on magnetic force) being possible. As well as being able to apply a relatively constant force over an extensive range (potentially reaching in excess of 200 pN) to the DNA, this set up can also manipulate the DNA topology. By rotating the magnets above the sample cell, the DNA tethered bead will also rotate causing the introduction of twist and writhe into the DNA. As DNA becomes more twisted, the energy stored will eventually be converted to writhe, forming a plectoneme. The production of this plectoneme is accompanied by a decrease in the DNA extension and therefore can be measured (see Chapter 4). Topos, such as MmT6, can relax plectonemes and as they do so, the DNA extension increases and this can be accurately measured. So, the magnetic tweezers form a powerful tool that can be used to study the activity of topos at the single molecule level with the ability to capture single strand passage events.

The magnetic tweezers technique has been used to show that topo IV is more active in relaxing positive supercoils over negative supercoils, due in part to a preferential DNA crossover angle of 85.5° [120]. Acute crossover angles are more often found in DNA that is positively supercoiled, whilst the angle is obtuse for negatively supercoiled DNA. TopoIV has also shown to be highly processive when relaxing positive supercoils, as a result of its CTDs, while being perfectly distributive on negatively supercoiled DNA. This work also demonstrated that for yeast topo II, the angle preference was 90°, hence providing an explanation for its non-preferential activity in the presence of positive and negative writhe [120]. Also, human topo IIα, as described in section 1.3.2.1.3, was demonstrated to be equally processive on positive and negative writhe, however the rate increase in the presence of positive is dictated by the CTDs, with the removal of the CTDs changing the chiral sign preference rather than the processivity [158]. A more comprehensive swivel mechanism for topo IB was established, showing that the enzyme nicks a single strand allowing the DNA to rotate in a controlled manner arising from friction generated within the enzyme cavity, with a chance of re-ligation with each full rotation [42]. More recently, the magnetic tweezers were used to measure the gate-opening dynamics of the type IA topos, *E. coli* topo I and III [224]. Gate opening is an event so transient that it had been impossible to measure using other biochemical or crystallographic techniques. The magnetic tweezers approach applies a constant force to assay these enzymes at the single-molecule level, allowing the elucidation of this previously uncaptured event. Single-molecule data, such as those mentioned, complements that obtained from ensemble experiments, together providing a deeper understanding into topo mechanism and function. This project has aimed to achieve this for *M. mazei* topo VI, elucidating previously unseen activities with implications for the mechanism and *in vivo* roles of this little-known enzyme.

1.6 Next-Generation Sequencing Techniques

The use of next-generation sequencing techniques in the field of topos has been gaining momentum over the last 10 years, as a way of studying topo-dependent genome cleavage patterns *in vivo* [54]. For a comprehensive literature review on the topic, see Chapter 5, Section

5.1, in which the current published data utilising NGS techniques will be discussed and used to inform on the generation of the novel technique utilised in the data collection for that chapter.

1.7 Project Aims

Despite its surprisingly ubiquitous nature in all domains of life, less is known functionally about topo VI and its homologues in comparison to type IIA. This project has sought to integrate the use of ensemble assays with the single-molecule technique, the magnetic tweezers, to perform an in-depth analysis of the activity of wild-type *M. mazei* topo VI. This is with the purpose to gain a deeper understanding of the mechanism, particularly how DNA topology modulates activity, and inform speculations into topo VI's role *in vivo*. As *M. mazei* topo VI expression and purification is accessible, and a full-length crystal structure is already obtained [195], the research reported here further develops *M. mazei* topo VI as a model topo VI enzyme, highly amenable to various forms of *in vitro* analysis. Further knowledge gained may be translated into finding inhibitors of the eukaryotic topo VI enzymes, which may have potential as herbicidal or anti-malarial agents; and a deeper understanding into the mechanism and function of Spo11.

Chapter 2

Materials and Methods

2.1 Ensemble Assay Methods

This section details the materials and methods used in the acquisition of data presented and discussed in Chapter 3.

Table 2.1: Names and constituents of frequently used buffers and solutions for data collected using ensemble techniques.

Buffers/solutions	Details
MmT6 dilution buffer (DB)	20 mM HEPES pH 7.5, 10% (v/v) glycerol, 100 mM potassium glutamate and 2 mM β-mercaptoethanol. Buffer stored at -20°C.
MmT6 minimal buffer (MB)	20 mM bis-tris propane pH 7, 100 mM potassium glutamate and 1 mM DTT. Buffer stored as 5x concentrate at -20°C.
MmT6 relaxation buffer (RB)	20 mM bis-tris propane pH 7, 100 mM potassium glutamate, 10 mM $MgCl_2$, 1 mM DTT and 1 mM ATP. Buffer stored as 5x concentrate at -20°C.
MmT6 binding buffer (BB)	20 mM bis-tris propane pH 7, 100 mM potassium glutamate, 10 mM $MgCl_2$ and 1 mM DTT. Buffer stored as 5x concentrate at -20°C.
STEB (2x)	100 mM Tris·HCL pH 8.0, 40% (v/v) sucrose, 100 mM EDTA, 0.5 mg/ml Bromophenol Blue
SAB (5x)	125 mM Tris (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 10% (v/v) β-mercaptoethanol and 0.002% (v/v) bromophenol blue.
TAE	40 mM Tris-Acetate pH 8.0, 1 mM EDTA

Purification buffer A	20 mM HEPES pH 7.5, 10% (v/v) glycerol, 800 mM NaCl, 20 mM Imidazole, 2 mM β-mercaptoethanol and EDTA free protease inhibitors.
Purification buffer B	20 mM HEPES pH 7.5, 10% (v/v) glycerol, 150 mM NaCl, 20 mM Imidazole, 2 mM β-mercaptoethanol, and EDTA free protease inhibitors.
Purification buffer B1	20 mM HEPES pH 7.5, 10% (v/v) glycerol, 150 mM NaCl, 500 mM Imidazole, 2 mM β-mercaptoethanol, and EDTA free protease inhibitors.
Purification buffer B2	20 mM HEPES pH 7.5, 10% (v/v) glycerol, 800 mM NaCl, 20 mM Imidazole, 2 mM β-mercaptoethanol.
Purification buffer C	20 mM HEPES pH 7.5, 10% (v/v) glycerol, 300 mM NaCl, 2 mM β-mercaptoethanol and EDTA free protease inhibitors.
Auto induction media (AIM)	10 g/L tryptone, 5 g/L yeast extract, 3.3 g/L ammonium sulphate, 6.8 g/L potassium dihydrogen phosphate, 0.5 g/L glucose, 2 g/L lactose, and 0.15 g/L magnesium sulphate

2.1.1 *M. mazei* DNA topoisomerase VI expression and purification

Both subunits of MmT6 (Top6A and Top6B) were expressed from a polycistronic duel-expression vector (Corbett et al., 2007), transformed into Rosetta 2 + pLysS *Escherichia coli* expression cells. They were grown for 24 hr at 37 °C in autoinduction growth media (AIM) (Formedium) in the presence of kanamycin (50 mg/ml) and Chloramphenicol (35 mg/ml). The culture was then centrifuged using the Sorvall RC6+ centrifuge in the FS9 rotor for 8 min, at 8000 rpm and 4°C. Supernatant was discarded and the pellet resuspended in buffer A (refer to Table 2.1 for all buffers mentioned, unless otherwise stated).

Chapter 2: Materials and Methods

Cells were lysed under high pressure using the Avestin high-pressure homogeniser. Samples were then spun at 18,500 rpm for 1 hr at 4°C with the Sorvall RC 6+ centrifuge and SS34 rotor. The lysate was then passed over a HisTrap™ FF Ni^{2+} column (5 ml/min, GE Life Sciences) and washed with purification buffer B. The protein was stepped of in purification buffer B1 and then loaded on to a HiTrap™ SP Sepharose HP column (5 ml/min, GE Life Sciences) followed in tandem by a HiTrap™ Q Sepharose HP column (5 ml/min, GE Life Sciences). The SP Sepharose column was removed before the protein was stepped off from the Q Sepharose with purification buffer B2. Protein concentration was assessed using solution absorbance at 280 nm, determined using the Lamba 25 spectrophotometer, the molecular weight and the theoretical extinction coefficient for MmT6 (for the heterotetramer: 220 kDa and 187,260 M^{-1}cm^{-1}, respectively). Samples were also analysed after each column procedure via the addition of 5x SAB and boiling for 5 min, before being loaded onto a RunBlue^{TM} 12% (w/v) precast SDS acrylamide gel from Expedeon, at 180V for ~40 min. The His-tag was cleaved using tobacco etch virus protease (TEV) (New England BioLabs) at a 1:50 concentration ratio (TEV:MmT6). The cleaved protein was then passed over a HisTrap™ FF Ni^{2+} column (5 ml/min, GE Life Sciences) and washed with purification buffer B, with the cleaved protein collected in the flow-through. The un-cleaved protein, his-tag and TEV were then stepped off in buffer B1. Fractions of MmT6 were pooled and concentrated using Amicon Ultra 15 mL centrifugal filter units (10 kDa cut-off, MilliporeSigma) before being passed down a Superose 6 10/300 (GE Life Sciences) column in buffer C, at 0.2-0.5 mLmin^{-1}. MmT6-containg fractions were subsequently concentrated as before and protein concentration determined using the method described above. All proteins stored at -80°C.

2.1.2 SEC-MALS and iSCAT

Size Exclusions Chromatography coupled to Multi-Angle Light Scattering (SEC-MALS) and Inferometric Scattering Technology (iSCAT) were techniques used to explore the holoenzyme structure of MmT6. Both were conducted with the help of the NHLBI Biophysical platform (Bethesda, USA). Provision of MmT6 in dilution buffer was all that was required, with each

technique being carried out by the platform. For SEC-MALS, two samples of MmT6 were provided 0.45 µM and 2.27 µM, and for iSCAT, one sample of 20 nM MmT6. The data from both assays was personally plotted and analysed using Igor Pro 7 (Wavemetrics).

2.1.3 DNA relaxation assays

Generally, the final MmT6 concentrations used ranged from 0.4 to 114 nM (20-40 nM being optimal) and were assayed in the presence of 2 nM pBR322* (see Chapter 3 for more details on pBR322*), either negatively supercoiled or positively supercoiled (Inspiralis), relaxation buffer, dilution buffer, with a 30 µl final volume. Samples were incubated at 21 or 37°C for 30 min before the reaction was stopped with the addition of 20 µL 2 x STEB and 30 µL chloroform: isoamyl alcohol (24:1 v/v). Samples were vortexed thoroughly, centrifuged at 13000 rpm using an Eppendorf table-top centrifuge for 3 min before being loaded onto a 1% (w/v) agarose gel. Generally, gels were run for 10-15 hrs at 2 Vcm^{-1}, before being stained with 0.5 µg/mL ethidium bromide (EtBr) and imaged under UV illumination using the Syngene Gel Doc system. DNA topology (negatively supercoiled, positively supercoiled, relaxed and linear), divalent metal ion identity and concentration (MgCl$_2$, MnCl$_2$, ZnCl$_2$, CaCl$_2$, CoCl$_2$ and NiCl$_2$ titrated at 0.5-100 mM), pH (6-9), temperature (20-90°C), use of potassium glutamate versus acetate, and incubation time were varied depending on the particular assay being performed. For details on specific relaxation assay conditions, see Figure legends associated with data presented in Chapters 3-5.

2.1.4 DNA decatenation assays

These assays are highly similar to DNA relaxation assays, aside from the DNA substrate utilised. Generally, the final MmT6 concentrations used ranged from 0.1 to 40 nM (0.5-5 nM being optimal depending on the catenated substrate in use) and were assayed in the presence of 250 ng of kinetoplast DNA (kDNA) or 2 nM of singly-catenated plasmids (Inspiralis), relaxation buffer, dilution buffer, with a 30 µL final volume. Samples were incubated at 21 or 37°C for 30 min before the reaction was stopped with the addition of 20 µL 2x STEB and 30 µL

chloroform: isoamyl alcohol (24:1). Samples were vortexed thoroughly, centrifuged at 13000 rpm using an Eppendorf table-top centrifuge for 3 min before being loaded onto a 1% (w/v) agarose gel. Generally, gels were run for 10-15 hrs at 2 Vcm^{-1}, before being stained with 0.5 µg/mL ethidium bromide (EtBr) and imaged under UV illumination using the Syngene Gel Doc system.

2.1.5 DNA cleavage assays

Generally, the final MmT6 concentrations ranged from 5 nM up to 160 nM (~160 nM being optimal) and were assayed in the presence of 3-4 nM pBR322* (Inspiralis), minimal buffer, dilution buffer, 1 mM ADPNP, 10 mM MgCl$_2$ and the final volume made up to 30 µl. Samples were incubated at 37°C for 30 min before the addition of 3 µL proteinase K (New England BioLabs) and 3 µl 2% (w/v) sodium dodecyl sulphate (SDS). Samples were incubated for 1 hr at 37°C before the reaction was stopped with the addition of 20 µL 2 x STEB and 30 µL chloroform: isoamyl alcohol (24:1). Samples were vortexed thoroughly, centrifuged at 13000 rpm using an Eppendorf table-top centrifuge for 5 min before being loaded onto a 1% (w/v) agarose gel. Generally, gels were run for 8-12 hrs at 2 Vcm^{-1}, before being stained with 0.5 µg/mL ethidium bromide (EtBr) and imaged under UV illumination using the Syngene Gel Doc system. DNA topology (negatively supercoiled, positively supercoiled, relaxed and linear), divalent metal ion concentration (MgCl$_2$, MnCl$_2$, ZnCl$_2$, CaCl$_2$, CoCl$_2$ and NiCl$_2$ titrated at 0.5-100 mM), pH (6-9), temperature (30-50°C), use of potassium glutamate versus acetate, and incubation time were varied depending on the particular assay being performed. For details on specific DNA cleavage assay conditions, see figure legends associated with data presented in Chapters 3-5.

For samples in which ethylenediaminetetraacetic acid (EDTA) was used to stimulate DNA cleavage reversal (Chapter 3, Section 3.2.5.6), the addition of EDTA occurred between MmT6 incubation and proteinase K/SDS treatment, and unless otherwise stated, was incubated at room temperature for 5 min. The concentration of EDTA was titrated 8-100 mM, and the incubation time varied 1-30 min.

2.1.6 DNA binding assays

To asses MmT6 DNA binding in the absence of cleavage, three separate assays were employed, all of which involved the use of a nitrocellulose membrane capture technique.

The first assay began with incubating 16 nM MmT6 with 10 nM pBR322*, which contained negatively-supercoiled, relaxed and linear pBR322* (Inspiralis) within the same sample, in MmT6 in binding buffer (BB, see Table 2.1) in 500 µL total volume, for 30 min at 37°C. The 0.45 µm nitrocellulose Centrex MF filters (Whatman) were prewashed with 500 µL of BB by centrifuging at 2000 rpm for 5 min using the Allegra X-I4R centrifuge (Beckman Coulter). The prewash flow-through was discarded and the binding assay applied to the column and centrifuged in the same manner as just described. The eluent (unbound DNA) was collected, before another 500 µL of BB was added to wash the filter. Once centrifuged, the wash was collected and the protein-bound DNA was then displaced from the filter using 500 µL 0.2% (w/v) SDS solution and centrifugation. The unbound DNA and MmT6-bound DNA fractions were applied to 30-kDa spin columns, prewashed with 10 mM Tris·HCl pH 7.5, and spun for 10 min at 14,000 rcf using an Eppendorf table top centrifuge. The eluent was discarded and the sample washed with 500 µL of 10 mM Tris·HCl pH 7.5, centrifuged again for 10 min at 14,000 rcf. This was repeated twice to fully exchange the buffers before the spin filter was inverted in a new collection tube and concentrated samples were eluted by centrifuging at 1000 rcf for 2 min. The DNA concentration of the samples was analysed using the Nanodrop. Using a 1% (w/v) native agarose gel, 100 ng of DNA from both the unbound and MmT6-bound DNA samples was run for 12 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination. Data used in construction of the graph (Chapter 3, Figure 3.22) was extracted from the gels using ImageJ and analysed in Excel.

The second assay used to explore MmT6 binding, followed a very similar protocol to the first, with 16 nM MmT6 being incubated with different topologies of 10 nM pBR322* separately, namely negatively supercoiled, positively supercoiled, relaxed and linear (obtained

Chapter 2: Materials and Methods

from Inspiralis), but without the requirement for gel electrophoresis. MmT6-bound DNA was analysed as a percentage of the total DNA added, using Microsoft Excel.

The third assay, first described in Litwin *et al*. [225], was used to attain a measure of relative MmT6 binding affinity for pBR322* topoisomers of varying superhelical density. To begin, extended distributions of negatively- and positively-supercoiled topoisomers were generated by incubating 11 nM negatively-supercoiled pBR322* with 100 nM *Archeoglobus fugidus* (gifted by Yeonee Seol, NHLBI, NIH, Bethesda, USA) in 50 mM Tris·HCl (pH 8.0), 10 mM NaCl, 10 mM MgCl$_2$, and 1 mM ATP, (30 μL reaction volume) at 95°C for 10 and 20 s, respectively. The reaction was stopped via the addition of 125 mM EDTA, 3 μL 2% (w/v) SDS and 2 μL proteinase K (New England BioLabs), and incubated for 1 hr at 37°C. The pBR322* was then purified using the Qiagen PCR clean up kit and ran on a 1% (w/v) agarose gel for 13 hr at ~2 Vcm^{-1}, with and without 1.5 μg/mL chloroquine, to validate the topology. Agarose gels were stained with 0.5 μg/mL ethidium bromide and imaged under UV illumination. The gel containing chloroquine (Chapter 3, Figure 3.23C) was soaked in 1x TAE for ~1 hr, prior to ethidium bromide staining.

Once the pBR322* topoisomer distributions were attained, MmT6 (8 nM) was incubated with 5 nM pBR322*, either positively or negatively supercoiled, in BB for 30 min at 37°C. The MmT6-bound DNA was separated from the unbound DNA using 0.45μm nitrocellulose Centrex MF filters (Whatman), and cleaned using 30 kDa Amicon spin filters, exactly as described above. The MmT6-bound and unbound DNA samples were run on a 1% (w/v) agarose gel for 13 hr at ~2 Vcm^{-1}, stained with 0.5 μg/mL ethidium bromide and imaged under UV illumination. The concentration of each topoisomer was measured as a function of band intensity, using ImageJ and Igor Pro 7 (Wavemetrics). In ImageJ, the gel image was corrected for the background using a rolling ball radius of 100, and the pixel intensity profile for each lane was extracted and imported into IgorPro7, where each peak was fitted to a Gaussian to attain the area. The area of each peak was used as a measure of topoisomer concentration, with the ratio of unbound to MmT6-bound DNA providing a measure of the relative dissociation

constant (K_d) for that particular topoisomer. See Chapter 3, Section 3.2.6, and [225] for further detail.

2.1.7 ATPase assays

2.1.7.1 PK/LDH assay

Constituents and their final concentrations can be found in table 2.2. Typical sample volumes were made up to 100 μL and placed in a flat-bottom 96-well microtitre plate. *E. coli* DNA gyrase (200 nM) (gifted by Lesley Mitchenall, JIC, UK) in the presence of 21 nM relaxed pBR322* was used as a control. The concentration of MmT6 used was 750 nM and the concentration of supercoiled pBR322* varied from 21 nM to 181 nM. The solution absorbance at 340 nm was measured using the CLARIOstar plate reader (BMG LabTech) every minute up to 60-90 min at 37°C. ATP/ADP were omitted from the solutions and added to the relevant samples after 10 min. Novobiocin, 1 μM, was used to inhibit the gyrase control and was added after 50 min. The output data were exported to and analysed using Microsoft Excel.

Table 2.2: Constituents of ATPase assay and respective final concentrations. NADH, PEP, PK/LDH (rabbit muscle), ATP, ADP and MgCl₂ purchased from Sigma Aldrich.

Constituent	Final Concentration
Nicotinamide adenine dinucleotide (NADH)	400 μM
Phosphoenolpyruvic acid (PEP)	800 μM
Pyruvate Kinase/Lactate Dehydrogenase (PK)	0.9-1.5/1.35-2.1 units
Adenosine triphosphate (ATP)	1 mM
Adenosine diphosphate (ADP)	1 mM
5X Minimal Buffer	1X
Magnesium Chloride (MgCl₂)	10 mM

Chapter 2: Materials and Methods

2.1.7.2 Radioactive ATPase assay

2.1.7.2.1 Purification of [γ-^{32}]-ATP

The following procedure was developed and modified in [226]. An ATP purification column was prepared using P-2 gel (Biorad). In 10 mL of P-2 wash buffer (10 mM Tris-HCl pH 8, 50 mM NaCl, 1 mM EDTA), 1.67 g of P-2 was added and left to hydrate overnight at room temperature. The supernatant was removed from the hydrated gel, which was then washed 4 times with degassed P-2 wash buffer. The gel was then poured into a 4 mL Econo-Column® (Biorad) and allowed to settle before P-2 buffer was slowly pumped through the column at 0.2 mL/min. A 100 μL ATP solution was made using 97 μL of 100 mM ATP, 2 μL of 33 mM [γ-^{32}]-ATP and 1 μL of a 1% (w/v) bromophenol blue solution dissolved in DMSO. Any residual P-2 wash buffer was carefully removed without disturbing the column matrix and the [γ-^{32}]-ATP solution added. Once the ATP solution had fully entered the column, the column was filled to the top with P-2 buffer. The ATP solution was left to migrate down the column, and using a Geiger counter pointed at the base of the column, the arrival of the radioactive [γ-^{32}]-ATP was monitored. Once the [γ-^{32}]ATP was close to eluting from the column, single drops were collected as separate fractions until the radioactive signal intensity dropped again. Using TLC plates, 0.5 μL drops were spotted from each [γ-^{32}]-ATP fraction, 1 cm from the bottom of the plate. The plate was then positioned upright with the ATP spots at the bottom in TLC running buffer (0.5 M lithium chloride in 1 M glacial acetic acid). The running buffer was allowed to move at least two thirds up the TLC plate before being removed and fully dried using heat. The dry TLC plates were then exposed to a phosphor screen for an hour before being imaged using the Typhoon FLA 7000 plate reader. The fractions shown to contain the most [γ-^{32}]-ATP and the least released phosphate were pooled and the concentration attained using absorbance at 260 nm (ATP ε = 15400 M^{-1}cm^{-1}) (Figure 2.1).
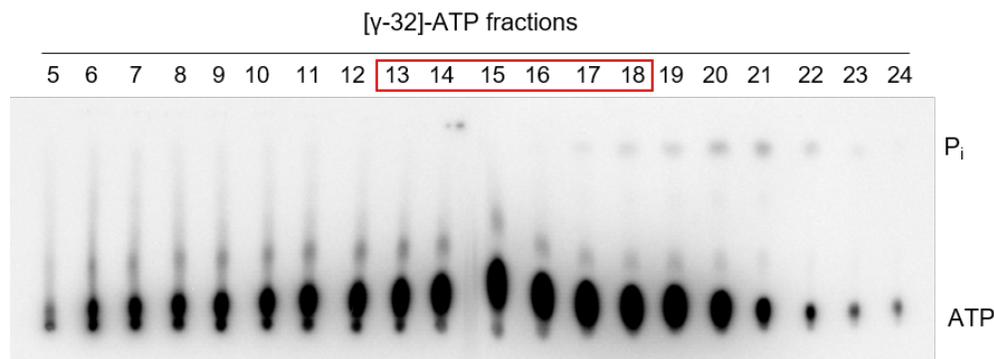
Figure 2.1: TLC plate imaged using the Typhoon FLA 7000 plate reader. Each ATP fraction was tested for its concentration of [γ-^{32}]-ATP (bottom signal) and released phosphate (top signal). Those fractions highlighted in red were pooled and used in the MmT6 radioactive ATP hydrolysis assay.

2.1.7.2.2 [γ-^{32}]-ATP hydrolysis activity of MmT6

This assay was, in effect, a relaxation time course, but replacing ATP with [γ-^{32}]-ATP. The reaction constituents were as follows; 0.5 µM MmT6 (A_2B_2), 430 nM DNA (relaxed, linear and supercoiled pBR322* were explored), 450 µM [γ-^{32}]-ATP, 10 mM MgCl$_2$ and 1x minimal buffer (20 mM bis-tris propane (pH 7), 100 mM potassium glutamate and 1 mM DTT), in a 55 µL reaction volume total. At each time point (1, 2, 4, 6 ,8, 10, 15, 20, and 30 minutes) a 5 µL aliquot was taken from the reaction and quenched using 5 µl of 2% (w/v) SDS and 100 mM EDTA. A 1 µL aliquot of each time point was then spotted onto a TLC plate and treated as above (2.1.7.2.1 Purification of [γ-^{32}]-ATP). The intensity of the spots were calculated using ImageJ and the amount of ATP hydrolysed at each time point calculated by looking at the proportion of [γ-^{32}]-P$_i$ released compared to unhydrolysed [γ-^{32}]-ATP (see Chapter 3 for more details).

2.2 Magnetic Tweezers Methods

This section details the materials and methods used in the acquisition of data presented and discussed in Chapter 4, collected using the magnetic tweezers (MT).

Chapter 2: Materials and Methods

Table 2.3: Names and constituents of frequently used buffers and solutions, for data collected using the single-molecule magnetic tweezers.

Buffers/solutions	Details
Wash buffer (WB)	1x PBS, 60 µM BSA and 0.02 (v/v) Tween-20
Bead wash buffer (BWB)	10 mM Tris-HCl pH 7.5 and 2M NaCl
KOH-ethanol solution	10% (w/v) KOH pellets dissolved in 95% (v/v) Ethanol
MmT6 MT activity buffer	20 mM Bis-Tris propane pH7, 100 mM potassium glutamate, 10 mM $MgCl_2$, 1 mM DTT, 0.01% (v/v) tween-20, 30 µM BSA
MmT6 MT storage buffer	20 mM HEPES pH 7, 40% (v/v) glycerol, 2 mM DTT, 15 µM BSA, 100 mM potassium glutamate

2.2.1 5-kb coilable DNA substrate synthesis

A 5-kb DNA segment from pET28b was amplified via PCR, using 5 ng/µL of pET28b, 0.5 µM of 5kbF and 5kbR primers (Table 2.4), 200 µM dNTPs, 1x Phusion buffer and 1 unit/50 µL of Phusion DNA polymerase (ThermoFisher). The biotin and digoxigenin DNA handles were also PCR-amplified using BioF/BioDigR and DigF/BioDigR respectively (Table 2.4) and 5 ng/µL pBlueScript plasmid as the template. Thermocycler conditions are shown in Table 2.5. All DNA products were run on a 1% (w/v) agarose gel to determine size followed by gel extraction using the Qiagen gel extraction kit. The 5-kb fragment, biotin-labelled and digoxigenin-labelled handles were all digested separately with BsaI-HF (New England Biolabs) using 5 µL of enzyme to 1 µg of DNA and incubating at 37°C for 6 hr, followed by the addition of another 3 uL of BsaI-HF and a further incubation at 37°C for 6 hr. The BsaI-HF was deactivated by incubating the samples at 65°C for 20 min, which were then purified using the Roche PCR clean up kit. The biotin and digoxigenin handles were ligated to the 5-kb fragment by incubating

2 nM of each DNA constituent with 1 unit T4 DNA ligase (New England Biolabs) per 50 µL

reaction, for 3 hr at 25°C followed by 15 hr at 16°C. A sample of the ligation reaction was then

run on a 1% (w/v) agarose gel to check the ligation efficiency. If the sample contained various

products (e.g. 5-kb fragment with only 1 ligated handle) then the DNA sample was gel extracted

and cleaned using the Qiagen gel extraction kit. See Figure 2.2 for illustration of the protocol

as well as gel-based analysis of the DNA fragments synthesised.
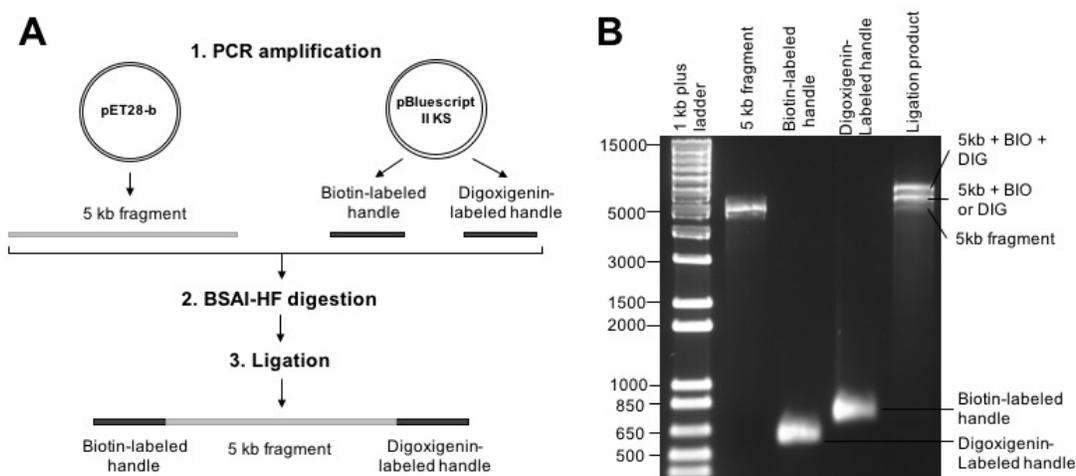
Figure 2.2: Synthesising the magnetic tweezers 5-kb coilable DNA substrate. A: The stages involved in generating the 5-kb coilable DNA substrate. 1. The 5-kb fragment is amplified via the polymerase chain reaction (PCR) from the pET28-b plasmid, whilst the biotin- and digoxigenin-labelled DNA handles are PCR amplified from pBluescript II KS. 2. All fragments are then digested using the restriction enzyme BsaI-HF. 3. The handles are then ligated to the 5-kb fragment and gel purified. B: DNA fragments used in the generation of the 5-kb coilable DNA substrate, before and after ligation, run on a 1% (w/v) agarose gel, alongside a 1-kb plus DNA ladder. The first band of the last lane (ligation product) is the desired substrate and was gel purified to separate it from DNA with either one or no handle.

2.2.2 5-kb braiding DNA substrate synthesis

A 5-kb DNA segment from pET28b was amplified via PCR, using 5 ng/µL of pET28b, 0.5 µM

of 5kbBF and 5kbBR primers (Table 2.4), 200 µM dNTPs, 1x Phusion buffer and 1 unit/50 µL

of Phusion DNA polymerase. Thermocycler conditions are shown in Table 2.5. The PCR

product was gel extracted using the Qiagen gel extraction kit and any nicks re-ligated by

incubating 2 nM DNA with 1 unit T4 DNA ligase (New England Biolabs) per 50 µL reaction, for

Chapter 2: Materials and Methods

3 hrs at 25°C followed by 15 hr at 16°C. In the single crossing assays the DNA duplex used was 3-kb, rather than 5-kb, and was kindly gifted in its purified form by Yeonee Seol (NHLBI, NIH, Bethesda, US).

2.2.3 12-kb force calibration DNA substrate synthesis

A 12-kb DNA segment was amplified via PCR, using 5ng/μL of λ DNA as the template, 0.5 μM of 12kbF and 12kbR primers (Table 2.4), 200 μM dNTPs, 1x Phusion buffer and 1 unit/50 μL of Phusion DNA polymerase. Thermocycler conditions are shown in Table 2.5. As there were no contaminating PCR products, gel extraction wasn't required.

Table 2.4: Primers used in the production of the 5-kb coilable DNA substrate used in the magnetic tweezer assays.

Primer name	Sequence
5-kb forward (5kbF)	GCTGGGTCTCGCAACGAAACCCGACAGGACTATAAAG
5-kb reverse (5kbR)	GCTGGGTCTCGACCACTCTGATGCCG CATAGTTAAG
Biotin handle forward (BioF)	GCTGGGTCTCGGTTGTTCCCTTTAGTGAGGGTTAATTG
Digoxigenin handle forward (DigF)	GCTGGGTCTCGTGGTTTCCCTTTAGTGAGGGTTAATTG
Biotin/digoxigenin handle reverse (BioDigR)	TATAGTCCTGTCGGGTTTCG
12-kb biotinylated forward (12kbF)	Bio-GCGGGTCTCGCAACAATCTGCGATGATGACTATG
12-kb digoxygenin reverse (12kbR)	Dig-GCGGGTCTCGACACCTTTGCGATGTACTTGACAC

5-kb braiding forward (5kbBF)	Bio-GCTGGGTCTCGCAACGAAACCCGACAGGACTATAAAG
5-kb braiding reverse (5kbBR)	Dig-GCTGGGTCTCGACCACTCTGATGCCGCATAGTTAAG

Table 2.5: Thermocycler programs for the PCR amplification of the 5 kb coilable, biotin/digoxigenin DNA handles, 12 kb force calibration and 5 kb braiding substrates.

	DNA substrate							
	5 kb coilable DNA		Biotin or digoxigenin DNA handles		12 kb force calibration DNA		5 kb braidable DNA	
PCR stage	Temp	Time	Temp	Time	Temp	Time	Temp	Time
Initial denaturation	98°C	2 min	94°C	2 min	98°C	30 s	98°C	2 min
Denaturation	98°C	30 s	94°C	20 s	98°C	10 s	98°C	30 s
Annealing	66°C	15 s	60°C	30 s	64°C	15 s	66°C	15 s
Extension	72°C	2 min 30 s	68°C	2 min	72°C	6 min	72°C	2 min 30 s
Final extension	72°C	5 min	68°C	5 min	72°C	10 min	72°C	5 min
Hold	4°C	hold	4°C	hold	4°C	hold	4°C	hold

2.2.4 Cover glass cleaning

The cover glass slides were placed vertically in a Teflon slide holder and submerged fully in a KOH-ethanol solution (10% (w/v)) in a 1 L glass beaker and sonicated for 30 minutes. The slides were then transferred to another 1 L beaker containing deionised water and sonicated for a further 30 min. The slides were then sonicated in deionised water twice more to fully remove the KOH-ethanol solution. The Teflon slide holder was removed from the beaker and

rinsed with deionised water, followed by 95% (v/v) ethanol. The slides were then fully dried by microwaving for 3 min at full power and stored in an airtight container at room temperature.

2.2.5   Sample cell assembly

The flow cell was composed of two different cover glass slides; the base was #1½ (24 mm x 60 mm) and the top was #1 (24 mm x 60 mm). The top cover glass had 2 mm inlet holes at either side, made using a bead blaster with aluminium particles. The top and bottom cover glass slides were adhered to each other using double-sided tape with a central channel cut out using a CO_2 laser engraver/cutter. Nylon chimneys, one larger serving as a buffer reservoir, and one smaller serving as an outlet channel, were placed over the 2 mm inlet holes at either side of the top cover glass. An airtight seal was produced using vacuum grease. Epoxy was then used around the base of the chimneys to securely fasten them to the slide and left to cure for an hour. Sample cells were then stored in an airtight container at room temperature. See Figure 2.3 for an illustration of the sample cell construction and Figure 2.4 for a photograph of a complete sample cell.
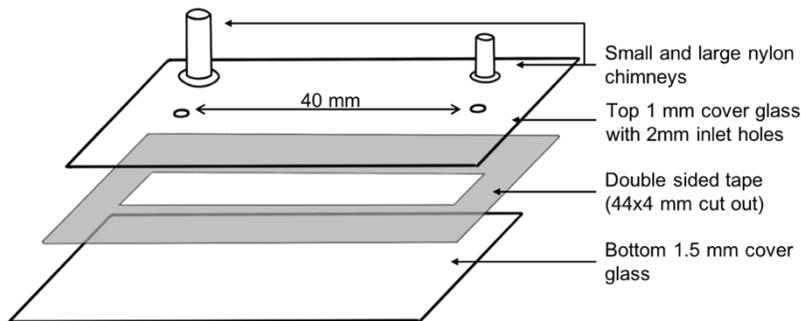
Figure 2.3: Magnetic tweezers sample cell assembly, showing the arrangement of the different parts, which come together to form the complete sample cell. The bottom (1.5 mm) cover glass is adhered to the top (1 mm) cover glass via a piece of double sided tape containing a 44x4 mm cut out. The nylon chimneys (large and small) are sealed over the top of the inlet holes cut into the top cover glass, using vacuum grease and epoxy.
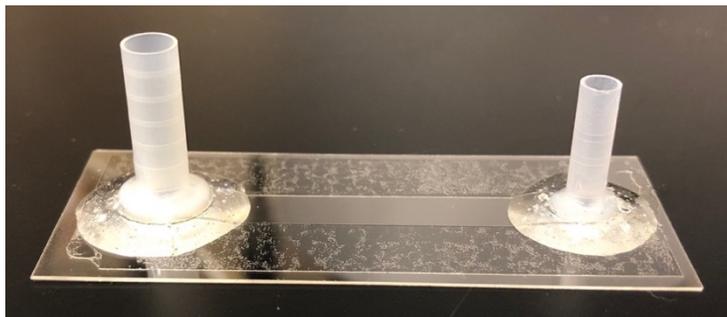
Figure 2.4: Photograph of an assembled sample cell.

2.2.6 Stuck bead generation

The following allows for the generation of 8-10 stuck bead sample cells. In a 1.5 μL Eppendorf tube, 200 μL of deionised water was mixed with 2 μL of 2 μm polystyrene beads (4% (w/v)) and mixed by quickly vortexing. The bead solution was then sonicated for 5 min. A 200 μL aliquot of acetone was added to the bead solution, vortexed quickly and then 40 μL introduced immediately to the sample chamber. The sample cell was then incubated for 5-10 min at room temperature before the sample cell was washed with 200 μL of deionised water (being careful not to fully remove the water) followed by 300 μL of 100% (v/v) ethanol (HPLC grade). Using a vacuum line, the ethanol was completely removed from the sample chamber and the sample cells baked on a hot plate at 125°C for 15 s. The stuck beads were then checked under a microscope.

2.2.7 DNA tether formation for supercoil relaxation assays

A 0.2 mg/ml anti-digoxigenin solution was made in 1% (w/v) PBS. With n = number of sample chambers, in a 1.5 μL Eppendorf tube, 45 x n μL of 1% (w/v) PBS, 3 x n μL of anti-digoxigenin solution and n μL of 0.5 nM coilable DNA substrate were added and incubated for 30 min at room temperature. A 40 μL aliquot was introduced to the sample chamber and left to incubate at 4°C overnight. The sample chamber was then washed with 200 μL of wash buffer (see Table 2.3).

2.2.8 Magnetic bead attachment to DNA for supercoil relaxation assays

Chapter 2: Materials and Methods

In an Eppendorf tube, 10 µL of a 1% (w/v) magnetic particle suspension (1 µm diameter, Dynabeads MyOne Streptavidin T1; 35601, Invitrogen) was added to 200 µL of bead wash buffer (BWB) and vortexed. The tube was then placed on a magnet to collect the beads whilst the supernatant was pipetted off. This was repeated, before another 200 µL of BWB was added and the mixture sonicated for 5 min at room temperature. A 40 µL aliquot of the magnetic bead solution was then introduced to the sample chamber containing the tethered DNA molecules and left to incubate at room temperature for 10 min. The unbound beads were washed out by gently flowing through 1 mL of wash buffer (WB).

2.2.9 Slide preparation for DNA braiding

In contrast to slide preparation for supercoiling, better results are achieved for braiding when the DNA is incubated with the magnetic beads and then introduced into a slide coated in anti-digoxigenin. First, 10 µL of a 1% (w/v) magnetic particle suspension (1 µm diameter, Dynabeads MyOne Streptavidin T1; 35601, Invitrogen) was added to 200 µL of bead wash buffer (BWB) and vortexed. The tube was then placed on a magnet to collect the beads whilst the supernatant was pipetted off. This was then repeated, before 20 µL of wash buffer (WB) was added and the mixture sonicated for 5 min at room temperature. The starting concentration of the bead suspension was 16.7 pM, which was diluted to 8.35 pM after washing. The braided DNA substrate was added in 5-fold excess to the final bead concentration. The DNA and magnetic bead mixture were left to incubate overnight at 4°C with rotation. The slide surface was prepared by incubating it with 0.015 mg/ml anti-digoxigenin in 1x PBS and was left overnight at 4°C. The following day, the slide was washed with 200 µL of WB and 5-10 µL of the magnetic bead-DNA mixture was gently introduced into the sample cell and left to incubate at room temperature for 1 hr. The unbound beads were then washed out by gently flowing through 1 mL of wash buffer (WB).

2.2.10 The magnetic tweezers instrument

The instrument, built by Andrew Dittmore and later modified by myself and Yeonee Seol (NHLBI, NIH, Bethesda, USA), is comprised of five main elements; the stages on which the sample cell is mounted, two permanent magnets fixed 1 mm apart, a red LED, an oil immersion objective and a camera. See Chapter 4 for both a photograph of the instrument itself and a schematic. The sample cell is mounted on top of two stages, one responsible for fine movements (minimum movement is 1 nm) and the other for coarse (minimum movement is 10 nm) in the x and y directions. Directly above and parallel to the sample cell are two permanent magnets fixed 1 mm from each other. They are powered by a motor which both rotates the magnets and moves them in the z coordinate. The distance of the magnets from the sample cell determines the force on the DNA. Below the sample cell is an oil immersion objective, capable of visualising the stuck and tethered beads, 1-2 μm in diameter. The image is captured by a camera which relays it to the computer and allows the user to see the contents of the sample cell. Specialised programs for operating the instrument are written using Lab view 2018 software (National Instruments). For the instrument used here, the program coding was done by Andrew Dittmore and Yeonee Seol (NHLBI, NIH, Bethesda, USA).

2.2.11 Camera resolution determination

To increase the field of view, a new camera was purchased and installed (acA1920-155um Basler ace from Basler). As this camera was more sophisticated than the previous, it was necessary to attain the resolution in nmpixel^{-1}. To do this a tethered DNA molecule was calibrated at high force and then while the program was tracking its position, the sample cell was moved 2 μm in either the x or y direction. The data recorded was then opened using Igor Pro 7 (Wavemetrics). Depending on whether the data was for movement in the x or y direction, a graph was made with x or y position against time. An average of magnetic bead position before and after the movement was calculated and the difference between them calculated in pixels. Because the sample cell was moved 2 μm, the pixel number was divided by 2000 to get the nmpixel^{-1} of the new camera. In this case, the resolution of the camera was 98.66 nm pixel^{-1} in x and 99.5 nm pixel^{-1} in y.

Chapter 2: Materials and Methods

2.2.12 Force extension calibration

The force on the tethered DNA was dependent on how close the magnets were to the sample cell. The closer to the sample cell, the higher the force on the DNA. To obtain an accurate measurement of the force on the DNA, as topos show force dependence in their activity, a force-extension calibration was performed. This was done using 12-kb tethered DNA molecules.

Tethered 12-kb DNA molecules were calibrated at high force with a stuck bead as a reference. Once their position was being tracked correctly, the force was dropped rapidly from high (~10 pN) to extremely low (~0 pN), achieved by moving the magnets from as close to the sample cell as possible to 8 cm away, at which point the effective force on the DNA applied by the magnets is ~0. The difference in extension measured will later be used as the DNA contour length (z) in the calculation of the force. The force was then increased to maximum. The software (Lab View 2018) was then programmed to move the magnets away from the sample cell by 0.1 mm every 20 s, until it was 6 cm from its starting position. At each position the beads displacement in the x coordinate was recorded. To generate an accurate force-extension curve, numerous data sets are required, in this case, 5 different DNA molecules were each assayed thrice. Fully derived in Chapter 4, the following equation was used to calculate the magnetic force on the DNA in z, as a function of magnet distance from the surface of the sample cell,

$$F_z = \frac{k_B T z}{\langle x^2 \rangle}$$

Equation 2.1: Calculating the magnetic force on the DNA-tethered bead.

Where F_z is the force on the DNA in z, $k_B T$ is the Boltzmann constant multiplied by room temperature expressed in kelvin (4.06 pN nm), z is the DNA contour length and x is the average displacement of the bead in the x coordinate. The assay described above provides both the DNA contour length (z) and the average displacement of the bead in the x coordinate (x), which are used to calculate the force (F_z). All data analysis was performed using Igor Pro 7 (Wavemetrics).

2.2.13 MmT6 -80°C storage for use in magnetic tweezers

Using two different strategies to probe MmT6 activity raised some questions as to how MmT6 should be stored for use in the magnetic tweezers in a manner which minimises protein waste and maintains MmT6 stability. In comparison to ensemble methods that use MmT6 in the concentration range 1-1000 nM, the magnetic tweezers require a far lower range from 0.025 to 2 nM. It was therefore necessary to store the protein at a much lower stock concentration in the -80 °C, of 300 nM, which was kept stable by the addition of 15 μM BSA and 40% (v/v) glycerol. When in use, the protein is diluted using the standard dilution buffer from Table 2.1. It is important for the magnetic tweezers assay that glycerol concentration is kept to a minimum to prevent non-specific interactions, or "sticking", of the magnetic bead to the surface.

2.2.14 MmT6-dependent supercoil relaxation assays

DNA tethers were first determined to be coilable by performing a calibration in which the DNA extension was measured as a function of magnet rotation. Once a coilable DNA tether was found and calibrated in MmT6 MT activity buffer (Table 2.3), MmT6 at the desired concentration (0.25-2 nM) was added to the sample cell with 1 mM ATP at 21°C, under 0.4 pN. The DNA tether was then supercoiled via the rotation of the magnets with 16-18 rotations producing a plectoneme on which MmT6 acted, relaxing each supercoil and increasing the DNA extension in distinct steps of ~110 nm. The data output was DNA extension change over time, and was stored as a delimited text file. The text file was opened and analysed in Igor Pro 7 (Wavemetrics) using a T-test based procedure, coded by Yeonee Seol (NHLBI, NIH, Bethesda, USA) [227]. All MT data analysed in this work has been done using this procedure in Igor Pro 7 (Wavemetrics), unless otherwise stated. The analysis procedure used a sliding window to determine time points at which the DNA extension changed and provided information on the size of the extension change and the rate of change, enabling determination of the average plectoneme relaxation rate, recorded as strand-passage events/min. This was plotted as a function of MmT6 concentration and fitted to the equation,

$$V_0 = \frac{V_{max}[E]}{K_{d,app} + [E]}$$

Equation 2.2: Michaelis-Menten-like equation.

To determine how the level of supercoiling affected MmT6 activity, 750 nM MmT6 was incubated with 1 mM ATP at 21°C and under 0.4 pN, and left to relax a plectoneme of 20 supercoils. The dwell time between each strand-passage event was extracted and plotted as a function of the supercoil number (also known as the 'DNA crossing number') during the dwell time in question, for both positive and negative supercoiling. For instance, the average dwell time when there were 2 supercoils was ~80 s, whilst the average dwell time when there were 20 supercoils was ~10 s. These data were fitted to an inverse relationship, where the dwell time was found to be inversely proportional to the number of supercoils present.

To determine how force affected MmT6 relaxation activity, force was varied whilst MmT6 concentration was kept constant (750 pM). For positive supercoiling, the force regime was 0.2, 0.5, 1, 1.5 and 2 pN, and for negative supercoiling it was 0.2, 0.4, 0.6 and 0.8 pN. The DNA tethers were calibrated to determine the buckling point at which DNA writhe begins to form, and the DNA extension change per magnet rotation, for each of the forces assayed. This allows the same number of DNA supercoils to be input regardless of the force and facilitates data analysis. Once calibration was complete, the plectoneme was incubated with 750 pM MmT6, 1 mM ATP at 21°C and the average rate of relaxation (strand-passage events/min) at each force on both positive and negative supercoils was extracted and plot as a function of force.

2.2.15 MmT6-dependent braid unlinking assays

DNA tethers were first determined to be braidable by performing a calibration in which the DNA extension was measured as a function of magnet rotation. Once a braidable DNA tether was found and calibrated in MmT6 MT activity buffer (Table 2.3), MmT6 at the desired concentration (0.025-0.9 nM) was added to the sample cell with 1 mM ATP at 21°C, under 0.5 pN. The average rate of braid unlinking was extracted and plotted as a function of MmT6 concentration. The data was also fitted to the Michaelis-Menten-like equation (equation 2.2). As MmT6

displayed a moderately processive behaviour during braid unlinking, further data analysis was performed. A 'burst' is defined as MmT6-dependent unlinking activity in which two or more strand-passage events occur in quick succession. The processivity rate (average rate of unlinking per burst, measured as strand-passage events/min), processivity step size (measured as number of strand-passage events/burst) and the average dwell time (measured in seconds) between bursts, were all extracted and plotted as a function of MmT6 concentration (see Figure 2.5 for labelled MT trace of MmT6-dependent braid unlinking).
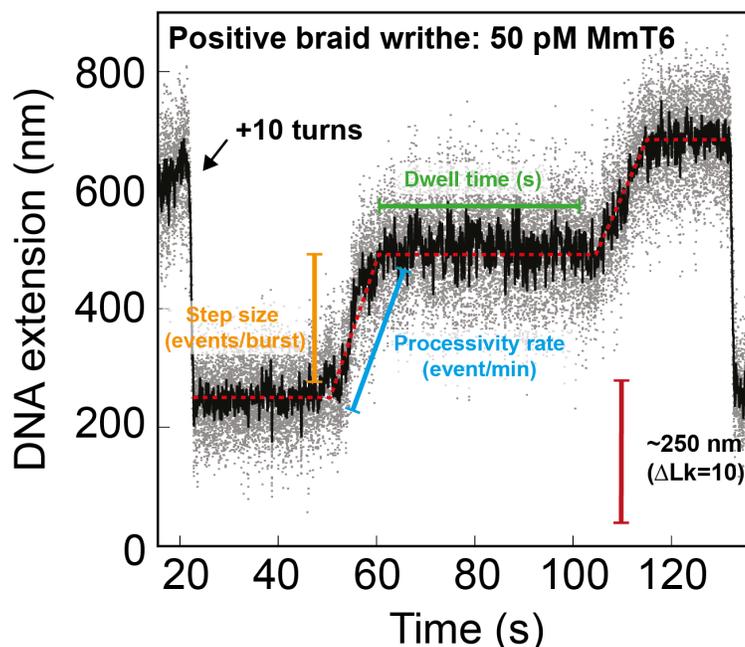
Figure 2.5: Labelled MT trace of MmT6-depedent braid unlinking. The average braid unlinking rate was calculated as the change in DNA extension over time, measured from the lowest DNA extension to the highest and expressed in strand-passage events/min (events/min). For this example of a positive braid being unlinked by 50 pM MmT6 under 0.5 pN at 21°C, the average rate was 6 events/min. However the data can be analysed to a greater depth by extracting the average processivity rate (blue line), the average step size (yellow line) and the average dwell time (green line). In this example, for the portions of the trace labelled, the processivity rate was 75 events/min, the step size was 5 events and the dwell time was 45 s. The light grey points are the raw data with the black representing the smoothed data, achieved using the Savitzky–Golay filter (smoothing window of 199 data points) in Igor Pro 7 (Wavemetrics). The red dashed line indicates periods of inactivity (horizontal portions) and processive bursts of multiple strand passage events (diagonal portions). The red line is a scale bar, highlighting how the unlinking activity is occurring in steps larger than a ΔLk of 2.

Chapter 2: Materials and Methods

To determine how force affected MmT6 unlinking activity, force was varied whilst MmT6 concentration was kept constant (500 pM). For positive supercoiling, the force regime was 0.5, 1, 2, 3 and 4 pN, and for negative supercoiling it was 0.5, 1, 2 and 3 pN. The DNA braids were calibrated to determine the DNA extension change per magnet rotation, for each of the forces assayed, which facilitates data analysis. Once calibration was complete, the braid was incubated with 500 pM MmT6, 1 mM ATP at 21°C and the average rate of relaxation (strand-passage events/min), the processivity rate (strand-passage events/min), the step size (strand-passage events/burst) at each force on both positive and negative supercoils was extracted and plot as a function of force.

2.2.16 MmT6-dependent single crossing assays

First described in [120], the single crossing assay was employed to explore the DNA crossing geometry preference of MmT6 in detail. To begin the DNA extension as a function of magnet position must be precisely measured for the input of the first crossing. This is achieved by measuring the DNA extension every tenth of a magnet rotation, and fitting the resulting curve to a geometric function, shown below.

$$L = \sqrt{L_0^2 - 4e^2 sin^2(n\pi)} - r + \sqrt{r^2 - e^2}, |n| < 0.5$$

Equation 2.3: Geometric function used to determine the DNA spacing and extension, where *L* is the measured DNA extension (dependent on the force applied), *L₀* is the maximum DNA extension, *e* is half the separation distance between the DNA molecules, *n* is the number of magnet rotations, and r is the radius of the magnetic bead to which the tethers are bound [120].

This equation allowed the extension of the DNA and the spacing between the two DNA tethers to be calculated, both of which were important pieces of information for execution of DNA crossing-angle simulations. Once calibrated, the single crossing was incubated with 900 pM MmT6 and 1 mM ATP at 21°C. A high concentration of MmT6 was used as the reaction is complete in a single cycle, so as to ensure T-segment engagement, rather than G-segment binding, was rate limiting. Force was varied as single braids were used to attain multiple data sets. Changing the force under which the assay was conducted changed the crossing angle

geometry and in effect presented MmT6 with a different DNA crossing geometry on which to act. The assay involved the input of a single crossing in one direction, allowing MmT6 to relax the crossing, before the input of a single crossing with the opposite chirality and so on. This was done until a single data set contained 50-100 unlinking events for both positive and negative DNA crossings. The data were then plotted as histograms and fit to single exponentials to attain the characteristic (or average) relaxation times for the unlinking of positive (τ_L) and negative (τ_R) DNA crossings, and the ratio of these times was calculated (τ_L/τ_R).

The assumption was that the rate-limiting step in the MmT6-dependent unlinking of the single crossing was sensitive to DNA crossing geometry. In order to relate the unlinking time ratio (τ_L/τ_R) to a value for the preferred MmT6 DNA crossing angle, simulations were performed. Here, two different types of simulation were used, Monte Carlo (MC) and Brownian Dynamics (BD). The BD simulations were conducted by Parth Desai (NHLBI, NIH, Bethesda, USA). These simulations use alternative methods to attain the distribution of crossing angles experienced by the single crossing and their execution relies on the data extracted from the calibration, namely the DNA spacing and extension, as well as the force under which the assay was conducted. Using the assumption that the ratio of unlinking times (τ_L/τ_R) is equal to the ratio of angle probabilities between negative ($P_R(\alpha_0)$) and positive ($P_L(\alpha_0)$) DNA crossings ($\tau_L/\tau_R = P_R(\alpha_0)/ P_L(\alpha_0)$), the experimental data was related to a precise value for preferred DNA crossing geometry. In other words, if a positive crossing was unlinked 3-fold faster than negative, the preferred angle was defined as the angle that was 3-fold more probable in positive crossings over negative. This assay was performed on 14 different DNA crossing geometries and the experimental data was related to the both the MC- and BD-derived crossing angle distributions to attain an averaged value for the MmT6 preferred DNA crossing angle. For each data set, this value was plotted against the average angle of the positive crossing distribution and the force, demonstrating the insensitivity of this value to both these parameters.

2.3 Utilising Next Generation Sequencing to Explore Type II DNA Topoisomerase cleavage specificity *in vitro* Using Plasmid DNA

Chapter 2: Materials and Methods

This section details the materials and methods used in the acquisition of data presented and discussed in Chapter 5.

Table 2.6: Names and constituents of frequently used buffers and solutions, for data collected using the NGS techniques.

Buffers/solutions	Details
MmT6 minimal buffer	20 mM bis-tris propane pH 7, 100 mM potassium glutamate and 1 mM DTT. Buffer stored as 5x concentrate at -20°C.
EcT4 minimal buffer	40 mM Tris-HCl pH 7.5, 100 mM potassium acetate and 1 mM DTT. Buffer stored as 5x concentrate at -20°C.
EcGyrase minimal buffer	35 mM Tris-HCl, 24 mM potassium chloride, 2 mM DTT, 6.5% glycerol (w/v) and 0.1 mg/mL BSA

2.3.1 Gel-based type II topo DNA cleavage pattern

The first approach implemented in the analysis of type II topo cleavage specificity was an agarose gel-based assay, in which topo cleavage on plasmid DNA, coupled with plasmid linearization and agarose gel electrophoresis, allowed resolution of topo-specific DNA fragmentation patterns. The type II topos used were *M. mazei* topo VI (MmT6), *E. coli* topo IV (EcT4) (gifted by Yeonee Seol) and *E. coli* gyrase (EcGyrase) (Inspiralis). Buffer conditions for each topo can be found in Table 2.6. For MmT6, 500 nM MmT6 was incubated with MmT6 minimal buffer, 11 nM pBR322* (linear, positively and negatively supercoiled), 2 mM ADPNP and 10 mM $MgCl_2$ for 30 min and 37°C. For EcT4, 150 nM EcT4 was incubated with EcT4 minimal buffer, 11 nM pBR322* (linear, positively and negatively supercoiled), 2 mM ADPNP and 10 mM $CaCl_2$ for 30 min and 37°C. For gyrase, 64 nM gyrase was incubated with gyrase

minimal buffer, 11 nM pBR322* (linear, positively and negatively supercoiled, and relaxed) and 20 mM CaCl$_2$ for 30 min and 37°C. For MmT6 and EcT4 cleavage reactions, this was followed by the addition of 2 μL EcoRI and 2 μL of 10x cut smart buffer (New England BioLabs, #R3101S), and then a further 30 min incubation at 37°C to linearise pBR322* in samples containing relaxed, negatively or positively supercoiled DNA (linear pBR322*, obtained from Inspiralis, has already been cut with EcoRI prior to use here). A 3 μL aliquot of 2% (v/v) SDS and 3 μL of proteinase K (New England BioLabs, #P8107S) is then added to all samples which are incubated for 1 hr at 37°C. The samples were then treated using 20 μL STEB and 30 μL chloroform:IAA (24:1) and centrifuged at max speed for 5 min using a table-top Eppendorf centrifuge. The gyrase protocol differs as EcoRI is not active in EcGyrase minimal buffer. Once the EcGyrase cleavage reaction was complete, 3 μL of 2% (v/v) SDS and 3 μL of proteinase K were added and the reactions are incubated for 1 hr at 37°C. This is followed by DNA extraction using the QIAquick PCR purification kit to the recommended protocol (Qiagen, #28106). The DNA concentration was measured using the Nandrop™ and then treated with EcoRI (as recommended), and incubated at 37°C for 30 min. This was followed by the addition of 5 μL STEB. For all topo reactions, 200-400 ng of DNA was loaded onto a 1% (w/v) TAE agarose gel alongside 5 μL of a 1-kb plus DNA ladder (Invitrogen™, #10787018) and run for 13 hr at 20 V. The gels were stained using SyBr green I (ThermoFisher Scientific, #S7567) (2 μL in 200 uL 1x TAE) for 1 hr, and destained in 1x TAE for 30 min before imaging using UV illumination.

To extract the sizes of the fragments produced by the different topos, the image of the agarose gel was analysed using the following software. To begin the image was opened in ImageJ, the background subtracted using a rolling ball radius of 100, and then saved as an 8-bit jpeg file. This file was then opened in GelAnalyzer, in which each lane was presented as a graph of band intensity against pixel number (Figure 2.6A). Each peak was assigned an Rf value (migration distance), so using the 1-kb ladder, a calibration curve was plotted in Igor Pro 7 (Wavemetrics) for Rf value against DNA fragment size in base pairs (bp) (Figure 2.6B). This could then be used in conjunction with Rf values extracted for each fragment arising as a result

of topo-dependent cleavage, to calculate their size, which was later compared to the data from next generation sequencing (NGS).
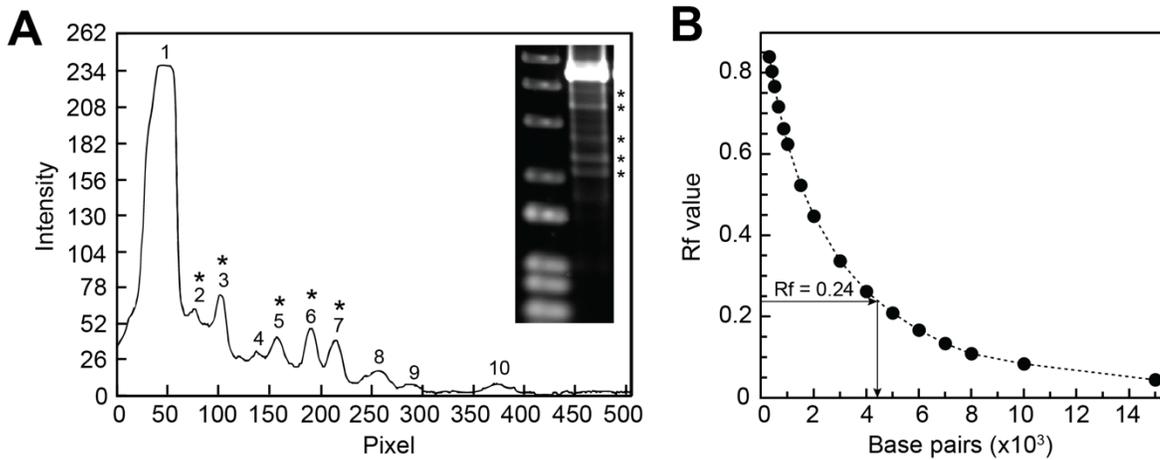
Figure 2.6: Agarose gel analysis of topo-dependent DNA fragmentation patterns. A: Example of MmT6-dependent cleavage on linear pBR322*, with a 1-kb plus ladder to the left. The plot represents a pixel intensity value for the MmT6 lane plotted against the pixel number, with each peak representing a band on the gel (those peaks starred in order from left to right are highlighted in the gel image from top to bottom).

2.3.2 DNA preparation for Illumina library preparation and sequencing

Reactions were set up in which MmT6, EcT4 and EcGyrase were incubated with pBR322* and allowed to cleave the DNA. For MmT6, in 300 µL total, 500 nM MmT6 was incubated with MmT6 minimal buffer, 10 mM MgCl$_2$, 2 mM ADPNP and 10 nM pBR322* (negatively supercoiled), for 30 min at 37°C. For EcT4, in 300 µL total, 240 nM EcT4 was incubated with EcT4 minimal buffer, 20 mM CaCl$_2$, 2 mM ADPNP and 10 nM pBR322* (relaxed, negatively and positively supercoiled For EcGyrase, 64 nM EcGyrase was incubated with EcGyrase minimal buffer, 20 mM CaCl$_2$ and 11 nM pBR322* (relaxed, negatively and positively supercoiled), for 30 min at 37°C. The DNA only control involved incubating 11 nM negatively supercoiled pBR322* in 1x MmT6 minimal buffer at 37°C for 30 min. For the MluCI restriction enzyme control, 11 nM of negatively supercoiled pBR322* was incubated with MluCI (New England BioLabs, #R0538) to the recommended protocol. All samples were sonicated, 10 s on (100 watts) followed by 20 s rest, for a total of 10 min on time (30 min in total), to sheer the

DNA into 100-500 bp fragments. The protein bound DNA was separated from unbound DNA using 0.45µm nitrocellulose Centrex MF filter (Whatman) pre-washed with the minimal buffer of the relevant protein. This stage was skipped for DNA-only and MluCI controls as no protein is covalently linked to the protein in these samples. The samples were each loaded onto the nitrocellulose column and centrifuged at 2000 rpm for 5 min using the Allegra X-I4R centrifuge from Beckman Coulter. The unbound DNA flow-through was discarded and the membrane washed using the addition of 300 µL of the relevant minimal buffer, and centrifuged again as previously described. The wash flow-through was discarded and the protein-bound DNA was eluted from the membrane using 200 µL of 0.8% (v/v) SDS, centrifuging the sample once more. The QIAquick PCR purification kit (Qiagen, #28106) was used to clean each sample, to the recommended protocol, and the presence of DNA (still attached to protein) was assessed using the NanoDrop™. The samples were then taken to the NHLBI sequencing core, in which the first round of Illumina sequencing adapter ligation was performed. As the protein is still attached to the DNA, only unbound DNA ends will be ligated to adapters with a specific bar code, allowing the DNA ends to be separated once sequenced. For DNA-only and MluCI controls, both ends of the DNA fragments are adapter ligated at this stage as no DNA end is protected by protein attachment. Once the first round of adapter ligation was complete, the topo-related samples were treated using 3 µL of 2% (w/v) SDS and 3 µL of proteinase K (New England BioLabs, #P8107S), and incubated for 1 hr at 37°C. The samples were cleaned again using the QIAquick PCR purification kit (Qiagen, #28106) to the recommended protocol, before the addition of 3 µL calf intestinal alkaline phosphatase (CIP) and 10x cutsmart buffer (New England BioLabs, #M0290), incubated for 2 hr at 37°C, before the addition of a further 3 µL of CIP and incubation overnight at 37°C. The Samples were then cleaned again using the QIAquick PCR purification kit (Qiagen, #28106) to the recommended protocol. The DNA was then quantified using the NanoDrop™, before the samples were taken back the NHLBI sequencing core for the second round of adapter ligation and sequencing using the Illumina MiSeq platform. Fragments were paired end sequenced to 75 bp (PE75).

Chapter 2: Materials and Methods

Whist the end DNA library generation stages were outsourced to the NHLBI sequencing core, in general the protocol is as follows. The first stage in DNA library generation is to repair the DNA ends so that they are all blunt. The end repair mixture used contained T4 DNA polymerase, *E. coli* DNA Polymerase I Klenow fragment and T4 polynucleotide kinase, which resected 3' overhangs and filled in 5' overhangs. The DNA fragments are then adenylated so that 5' and 3' adapters can be ligated. Adapters are known as Y-shaped and contain unique sequences which allow the DNA to be barcoded. This facilitates the parallel sequencing of multiple pooled samples, or in this case, allows the two ends of the DNA to be distinguished. The DNA is then solid-phase PCR amplified to generate clusters of identical fragments and sequenced using fluorescent reversible terminator bases (Illumina).

2.3.3 Sequencing data analysis

Data obtained from the NHLBI sequencing core was in the file format fastq.gz, and contained sequence data for thousands to millions of DNA fragments. Analysis was performed using the NIH high performance computing system, BioWulf, and the coding pipeline created by Vijender Chaitankar, a member of the NHLBI bioinformatics core. In descriptive terms, the first stage was to unzip the sequence data files, remove primer contamination and run an alignment code that assigned each sequence to a position on the reference genome, which in this more unusual case, was the *E. coli* plasmid, pBR322*. This was then visualised using Integrated Genome Viewer (IGV), which allowed the visualisation of where these reads mapped to the designated reference material. The code used for data analysis is shown below (written by Vijender Chaitankar).

```
sinteractive —mem=32g —cpus-per-task=8
module load trimmomatic bwa samtools;
java -jar $TRIMMOJAR SE -threads 8 ./SM4853-4_1_S7_L001_R1_001.fastq
./trim/SM4853-4_1_S7_L001_R1_001.fastq
```

```
ILLUMINACLIP:/usr/local/apps/trimmomatic/Trimmomatic-
0.36/adapters/TruSeq3-SE.fa:2:30:10 SLIDINGWINDOW:2:15 CROP:14;
bwa index /home/kimr3/MiSEQ/plasmid_reference/pBR322sequence.fa;
bwa aln -t 8 -n 0 /home/kimr3/MiSEQ/plasmid_reference/pBR322sequence.fa
./trim/SM4853-4_1_S7_L001_R1_001.fastq > ./align/SM4853-
4_1_S7_L001_R1_001.sai;
bwa samse /home/kimr3/MiSEQ/plasmid_reference/pBR322sequence.fa
./align/SM4853-4_1_S7_L001_R1_001.sai ./trim/SM4853-
4_1_S7_L001_R1_001.fastq | samtools view
-q 25 -F 3844 -h - | grep "XT:A:U\|@*" | samtools sort -@ 8 -o
./align/SM4853-4_1_S7_L001_R1_001.bam - ;
samtools index ./align/SM4853-4_1_S7_L001_R1_001.bam;
```

Commands of note are in bold. The code involves the use of pre-scripted modules called trimmomatic and BWA samtools, the former being used for trimming Illumina sequencing data, removing adapter sequence and low-quality reads, and the latter being used to align the filtered sequence to the reference genome. The command SE instructs the computer to align the data as single end, which generated alignment files in which only one of the paired ends is included. During sequencing, fragments were sequenced from one end before being flipped *in situ* and then sequenced from the other. This not only sequenced both ends of the fragment (important here as all DNA ends could have potentially been associated with a topo) but also linked the two sequences as having originated from the same fragment. In order to separate topo-cleaved ends from background noise, the data was aligned as a single end. The sequencing was also performed to a depth of 75 bp, which is a level of detail unnecessary here, so the sequence was cropped to 14 bp prior to alignment (see command CROP:14). Through bioinformatic analysis, by Vijender Chaitankar (NHLBI, NIH, Bethesda, US) it was determined that 14 bp was short enough that noise was reduced, whilst still allowing each sequence to be uniquely aligned to pBR322*. The final command of note, -q 25, allowed the optimisation of the quality of reads which are mapped to the reference.

Chapter 2: Materials and Methods

In brief, as this is described in detail in Chapter 5, the motivation behind designing this assay was to reveal if there was any sequence specificity for cleavage by MmT6 specifically, and by extension, other type II topos, and the length of the base-stagger generated during cleavage. The answer to the first question should be readily attainable by looking at sequence density bias around the plasmid. If there was a preference for a sequence, more reads would cluster to a particular region. However, if there were no density bias and all regions of the plasmid were represented equally, then this would provide evidence for MmT6 having little to no sequence specificity. The second question, the size of the overhang generated, is a harder question to answer and relied on the presence of specificity. As the enzyme is known to cleave double-stranded DNA, forming two covalent 5' phosphotyrosyl linkages, one to each strand of the duplex, this meant that during the end repair stage, these overhangs were filled in on both sides of the cut site. This should result in a two-fold enrichment of the sequence of the overhang and it is this footprint which allowed cleavage overhang size determination. Without specificity, potentially the whole plasmid would be enriched equally and so the stagger would be hard to glean from the data. Please refer to Chapter 5 for a comprehensive description of the protocol rational.

As it currently stands, the way in which the preferred sequences were analysed once aligned involved the use of visual inspection. For certain data sets (see Chapter 5) the peaks were obvious and displayed a characteristic shape, allowing for easy identification. The sequence at each peak was extracted and in the case of MmT6 data, it was analysed using WebLogo online software (http://weblogo.threeplusone.com/) [228], in order to assess any base identity preferences around the cleavage sites. In the case of EcT4 and EcGyrase data analysis, an online wordcount software was used (http://www.bioinformatics.nl/cgi-bin/emboss/wordcount#forms::wordcount), as these data sets did not resolve the precise cleavage site. However, regions of consecutive cytosine or guanidine residues were found to be enriched within the sequences using the aforementioned online tool. When this technique reaches maturity, code will be generated allowing for unbiased peak detection, as well as a more complex and thorough analysis of sequence preference.

Chapter 3

Characterisation of *M. mazei* Topo VI using Ensemble Biochemistry

3.1 Introduction

The ensemble characterisation of DNA topoisomerase (topo) activity *in vitro* can take multiple forms, with perhaps the most prevalent technique being agarose-gel electrophoresis. Due to varying plasmid topologies that alter the rate at which the plasmid migrates within the gel, the topology manipulation performed by topos can be assayed in a substantial amount of detail. For the characterisation of MmT6, this technique was indispensable.

As seen in Figure 3.1, supercoiled (SC) plasmid DNA migrates the fastest through the gel due its more compact structure. Under these conditions, nicked DNA, also known as open circular (OC), travels the slowest as the nick prevents any twist or writhe from being stored in the plasmid, which therefore means it has the most open structure. If the plasmid is linearised (L) it will run true to size, should a DNA size marker be run simultaneously. In the case of pBR322*, the plasmid used here, the linearised form runs between the open circular and supercoiled species. If the DNA is partially relaxed, generally this means that the plasmid will run as a set of bands known as topoisomers, running progressively faster than the open circular. This is due to varying degrees of supercoiling still being present in the DNA, with increasing writhe leading to further migration through the gel. When DNA is relaxed by topos, rather than yielding a single fully-relaxed product, they reach an equilibrium in which different degrees of supercoiling remain. For type IA topos and type IIB topos, this equilibrium is thermodynamically expected, whereas type IIA topos are able to further relax the DNA, producing even narrower topoisomer distributions that are beyond the thermodynamic equilibrium [229-231]. This is known as topology simplification, and while the cause of this

phenomenon has been debated for some time, it remains unresolved. As MmT6 is a type IIB topo, it has been shown not to perform topology simplification, despite utilising ATP, which demonstrated that the free energy from ATP hydrolysis was not the determinant of topology simplification by type IIA topos [231]. Using atomic force microscopy (AFM), it has also been suggested that MmT6 bends the G-segment to a similar degree as type IIA [61], hence eliminating G-segment bending as the lone contributing factor of topology simplification [229]. The subject will not be addressed further in this work, however it is worth noting that MmT6 has proven to be a highly useful probe for exploring the possible mechanisms of topology simplification. A potential explanation explored in [231] was that geometric selection of the DNA crossing bound may be contributing. This remains a plausible hypothesis as Chapter 4 will detail that the MmT6 DNA crossing angle preference is ~88°, which is extremely close to 90°, like yeast topo II [120], which incidentally also doesn't significantly simplify topology despite being a type IIA topo [229]. The issue is confounded, however, as *E. coli* topo IV, known to simplify topology to a high degree [61, 229], has a DNA crossing angle preference of ~85.5° [120].
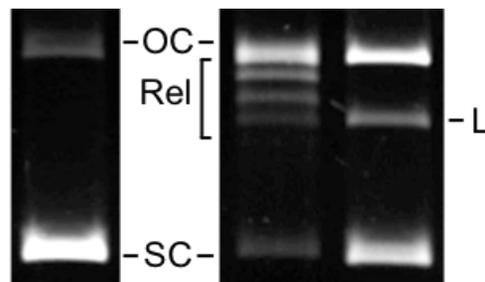
Figure 3.1: The effects of DNA topology on the rate of migration through an agarose gel in absence of DNA intercalators. This illustrates how supercoiled (SC), open circular/nicked (OC), relaxed (rel), and linear (L) pBR322* migrate through a 1% (w/v) native agarose gel, run for 10 hr at ~2 Vcm^{-1}, stained with 0.5 μg/mL ethidium bromide and imaged under UV illumination. For further details on the assay from which the gel was taken see Figure 3.6.

Although gel electrophoresis was not the only technique to be utilised in this chapter, it comprised a major component of this work and is often a key technique used in ensemble topo characterisation. Other techniques employed here in Chapter 3 were the PK/LDH (Section 3.2.7.1) and radioactive ATP (Section 3.2.7.2) assays exploring MmT6 ATP hydrolysis activity,

as well as a membrane capture technique used to explore MmT6-dependent DNA binding (Section 3.2.6), all of which are described in full within the respective sections.

3.2   Results and Discussion

3.2.1   MmT6 expression and purification

This project has centred around the in-depth characterisation of DNA topoisomerase VI from *M. mazei* (MmT6), using sensitive techniques of both an ensemble and single molecule nature. It has therefore been imperative that the enzyme be produced in considerable quantities and to a high degree of purity. The presence of contamination from other DNA-cleavage enzymes or ATPases were unacceptable, hence the optimisation and technical skill acquisition of MmT6 protein production was paramount when this research began. Fortunately, James Berger and colleagues (Johns Hopkins University, Baltimore, MD, USA) had previously isolated and cloned MmT6, developing a robust purification protocol designed to produce protein of crystallographic standard [62, 195], as well as facilitating the work done here by gifting the MmT6 expression plasmid.

When topo VI was first discovered and characterised in 1994, it was done so by purification from the hyperthermophillic archaeon, *Sulfolobus shibatae* [167]. This seminal work revealed a heterotetrameric protein, formed of two Top6A and two Top6B subunits, of 47 and 60 kDa, respectively, with ATP/Mg^{2+}-dependent DNA relaxation and decatenation activities. Subsequently, evidence of over-expression and purification from *E. coli* was attained, demonstrating that active *S. shibatae* topo VI (SsT6) could be reconstituted from separately expressed/purified Top6A and Top6B subunits [193]. The Top6B subunit could be robustly expressed and purified, remaining soluble throughout, however Top6A required purification from inclusion bodies after denaturation and renaturation. This significantly affected the amount of active protein attained and led researchers to conclude that both subunits required dual expression from the same plasmid. Following work detailed the use of a pET3b-based expression plasmid, which allowed both subunits to be expressed simultaneously, thus allowing

the expression and purification of considerable amounts of active SsT6 in *E. coli* [194]. This approach has continued, with MmT6 being expressed from another dual expression vector derived from pST39 [232], which also expressed both Top6A and Top6B together. It is hypothesised that Top6A, in particular, requires the presence of Top6B in order to maintain structure and stability. It was recognised early on that Top6A is highly homologous to the eukaryotic meiotic factor Spo11 [171], the *in vitro* expression of which has also plagued Spo11 research for some time. Very recently, the *Saccharomyces cerevisiae* Spo11 complex has finally been purified from insect cells, after decades of trial and error, achieved only through the identification and dual expression of Top6B-like homologues, Rec102 and Rec104 [210].

As previously mentioned, the MmT6 expression plasmid used in this research, was gifted by James Berger (Johns Hopkins University, Baltimore, MD, USA), with the expression and purification protocols implemented largely unchanged from previously published work [62, 195]. This begins with the chemical transformation of the MmT6 plasmid into competent Rosetta™ 2 (Merck) *E. coli* expression cells, followed by expression in auto induction growth media (AIM) at 37°C for 24 hr. When samples of the expression culture were analysed using denaturing polyacrylamide gel electrophoresis (PAGE), MmT6 expression was confirmed with Top6B (69-kDa) and Topo6A (42-kDa) clearly visible (Figure 3.2).
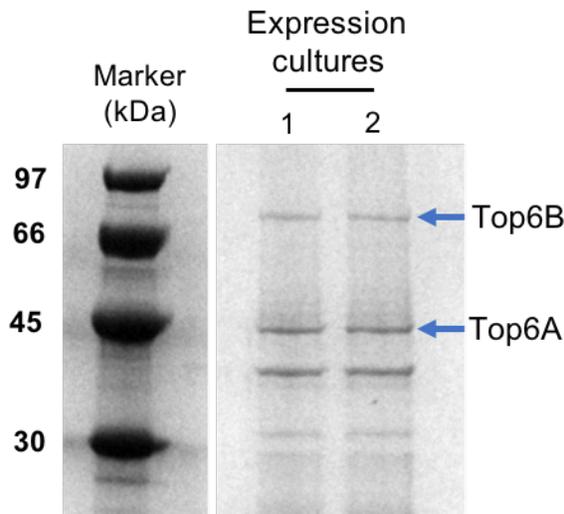
Figure 3.2: MmT6 expression. Rosetta 2 *E. coli* expression cells were transformed with the polycistronic dual expression vector carrying both the Top6A (45 kDa) and Top6B (65 kDa) subunit genes, and incubated at 37°C for 24 hrs. A 15 μL aliquot of the culture was run on a

precast 10% (w/v) polyacrylamide gel (Sigma) under denaturing conditions and stained with *Instant*Blue™ (Sigma). Blue arrows indicate the position of the Top6A and Top6B proteins.
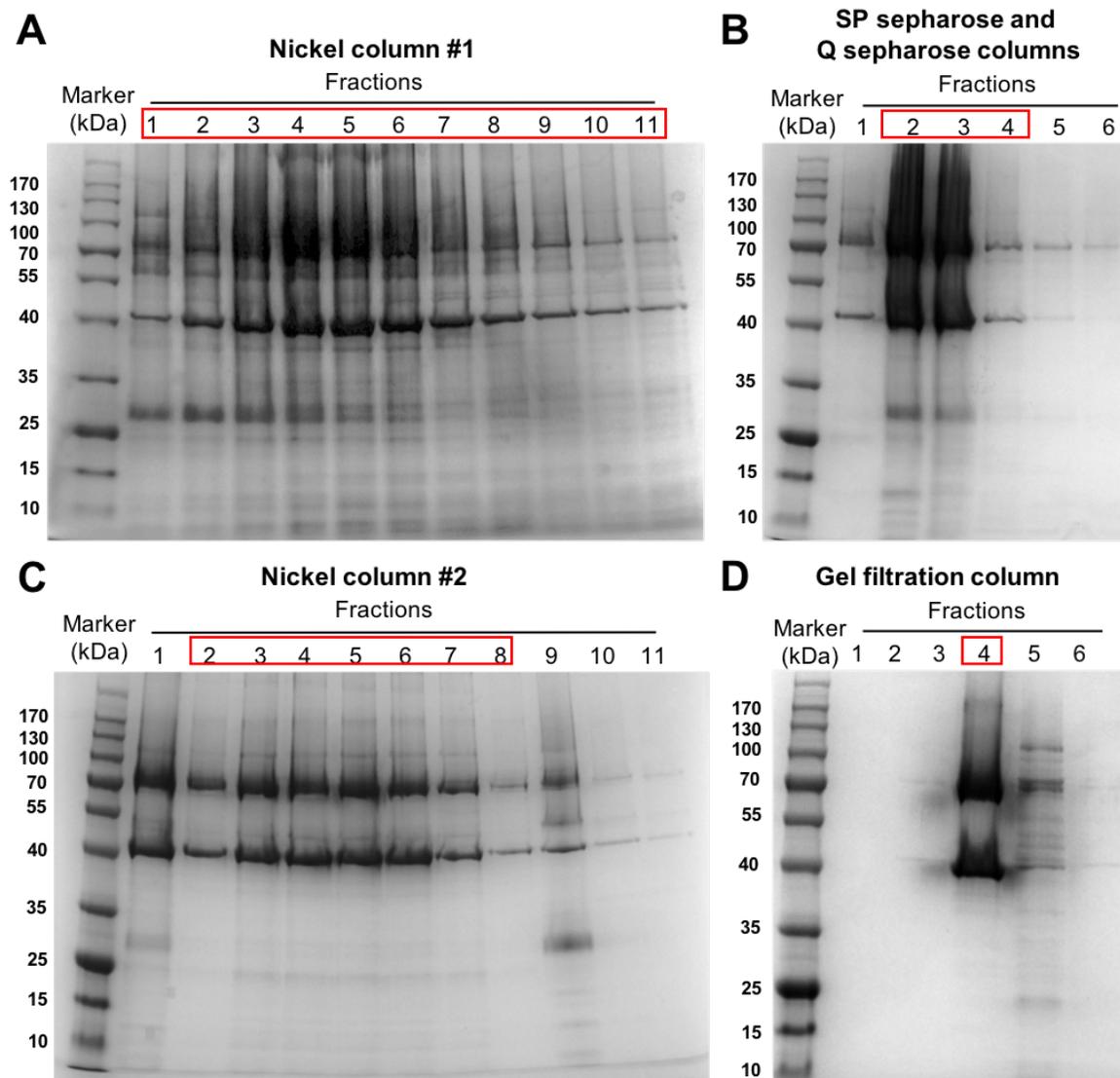
Figure 3.3: MmT6 purification. A: Nickel column number 1. As MmT6 was his-tagged, a nickel column (HisTrap™ FF Ni^{2+} column, GE Life Sciences) was used, with the MmT6 eluted using an imidazole gradient into 5 mL fractions. B: The fractions from A were pooled and passed over an SP Sepharose column (HiTrap™ SP Sepharose HP column, GE Life Sciences) followed directly by a Q Sepharose column (HiTrap™ Q Sepharose HP column, GE Life Sciences). Proteolysed Top6B remained bound to the SP sepharose, while folded Top6A and Top6B were bound to the Q Sepharose. The SP Sepharose column was removed and the Top6A and Top6B were eluted from the Q Sepharose using an NaCl gradient into 5 mL fractions. C: Post his-tag removal (via tobacco etch virus (TEV) protease incubation), the MmT6 was run down a second nickel column (HisTrap™ FF Ni^{2+} column, GE Life Sciences) to remove the his-tagged TEV and cleaved MmT6 his-tag. MmT6 was eluted from the column in 5 mL fractions and concentrated down to 0.5 mL. D: The concentrated MmT6 was put down a Superose 6 10/300 (GE Life Sciences) gel filtration

column and eluted in 1 mL fractions. Fractions from each column were run on 10% (w/v) denaturing polyacrylamide gels (Sigma) and stained using _Instant_Blue™ (Sigma). Pooled fractions for all figure panels, which went on to subsequent purification steps, are highlighted in red (in D, the highlighted fraction represents the final product used for activity testing).

Following expression, MmT6 expression cultures were lysed, and the soluble MmT6 was subject to purification. The Top6B protein was his-tagged, which allowed Top6B/Top6A dimers to be purified using a nickel column (Figure 3.3A). This proved effective in separating the MmT6 protein from the bulk cellular extract. The MmT6 was eluted from the nickel column using an imidazole gradient, with fractions containing MmT6 being subject to further purification using both SP and Q Sepharose columns (Figure 3.3B). The purpose of this stage of the protocol was to remove any mis-folded or proteolysed Top6B. The Q Sepharose column was attached directly after the SP Sepharose, with proteolysed Top6B remaining bound to the SP, whilst fully folded Topo6A/Top6B dimers were bound to the Q Sepharose. By removing the SP Sepharose column, and then applying an NaCl gradient, the Topo6A/Top6B dimers were eluted (Figure 3.3B). Fractions containing MmT6 were then pooled and treated with tobacco etch virus (TEV) protease to remove the his-tag, prior to purification using a second nickel column. However, this time MmT6 without the his-tag was immediately eluted from the column whilst the cleaved his-tags and the his-tagged TEV protease remained bound (Figure 3.3C). To achieve increased purity, the MmT6 flow-through fractions were concentrated before being applied to a gel filtration column (Figure3.3D). Using size exclusion chromatography allowed the separation of the MmT6 from smaller contaminating proteins (see lane 4 versus lane 5 in Figure 3.3D).

Post gel filtration, the MmT6 was concentrated once more before relaxation activity was assessed (Figure 3.4), to determine whether active MmT6 protein had been purified in the absence of contaminating DNA cleavage enzymes. Without the gel filtration column, often activity testing would reveal topo-independent DNA linearisation, suggesting nuclease contamination, which, for numerous aspects of this research, was unacceptable. In general, MmT6 achieved optimal relaxation of 2 nM negatively supercoiled pBR322* between 10-40 nM

MmT6, in the presence of 10 mM MgCl$_2$ and 1 mM ATP, incubated for 30 min at 37°C. Once a robust protocol for the expression of MmT6 was established, the work detailed over the remainder of this thesis was facilitated.
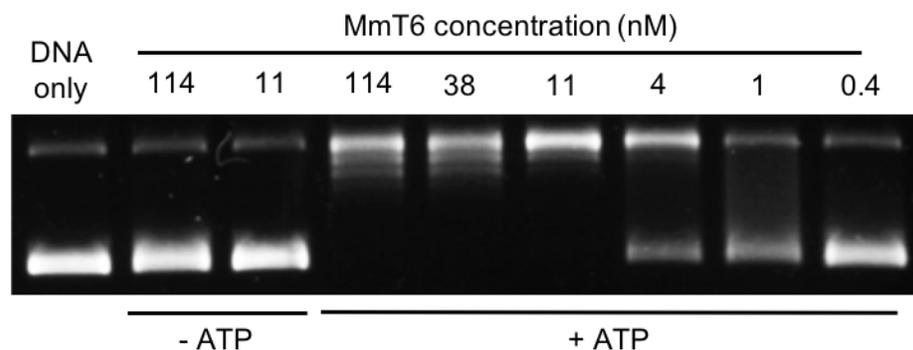
Figure 3.4: MmT6 activity testing. Post gel filtration, MmT6 was used in a DNA relaxation assay to assess protein activity. MmT6 (0.4-114 nM) was incubated with 2 nM negatively supercoiled pBR322* with or without 1 mM ATP, at 37°C for 30 min, and then run on a 1% (w/v) native agarose gel for 15 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination.

3.2.2 SEC-MALS and iSCAT

This section details preliminary assays exploring the overall complex formation of MmT6 as a heterotetramer formed of two Top6A and two Top6B subunits, ~220 kDa in size, using Size Exclusion Chromatography coupled to Multi-Angle Light Scattering (SEC-MALS) and Interferometric Scattering Technology (iSCAT). Both techniques allow the size of molecules to be determined using light scattering [233, 234]. Here, they were employed to determine whether MmT6 was composed of Top6A or Top6B monomers, Top6A/Top6B dimers or heterotetramers formed of two Top6A/Top6B dimers. SsT6, when first discovered, was demonstrated to be a heterotetramer in solution [167], with its crystal structure also demonstrating this [196]. MmT6, whilst the crystal structure was of the Top6A/Top6B dimer [195], also appeared to a heterotetramer in solution, with the major molecule size determined using SEC-MALS shown to be ~200 kDa, regardless of whether 2.27 or 0.45 µM MmT6 was used (Figure 3.5A). As SEC-MALS required relatively large amounts of protein, iSCAT was also employed which allowed molecular size determination at 20 nM MmT6. This similarly revealed that the majority of MmT6 molecules in solution were ~210 kDa (see peak 3, Figure

3.5B), and therefore heterotetramers. The iSCAT technique revealed further detail showing molecules that were either ~60-70 kDa, determined to be Top6B monomers (see peak 1, Figure 3.5B) or ~100-110 kDa, determined to be Top6B/Top6A heterodimers (see peak 2, Figure 3.5B). These are likely present in very small quantities due to low-level complex degradation, rather than representing a single data point for the equilibrium between the dimer and heterotetramer. However, iSCAT measurements at different MmT6 concentrations would be necessary to determine whether MmT6 begins to dissociate into dimers as concentration decreases. It is highly likely that the Top6B detected is only present due to having not formed a dimer with Top6A during expression, rather than dissociating from Top6A in solution as no Top6A was detected. However, size limitations associated with iSCAT may prevent visualisation of free Top6A, therefore future work is required to fully understand MmT6 complex formation.
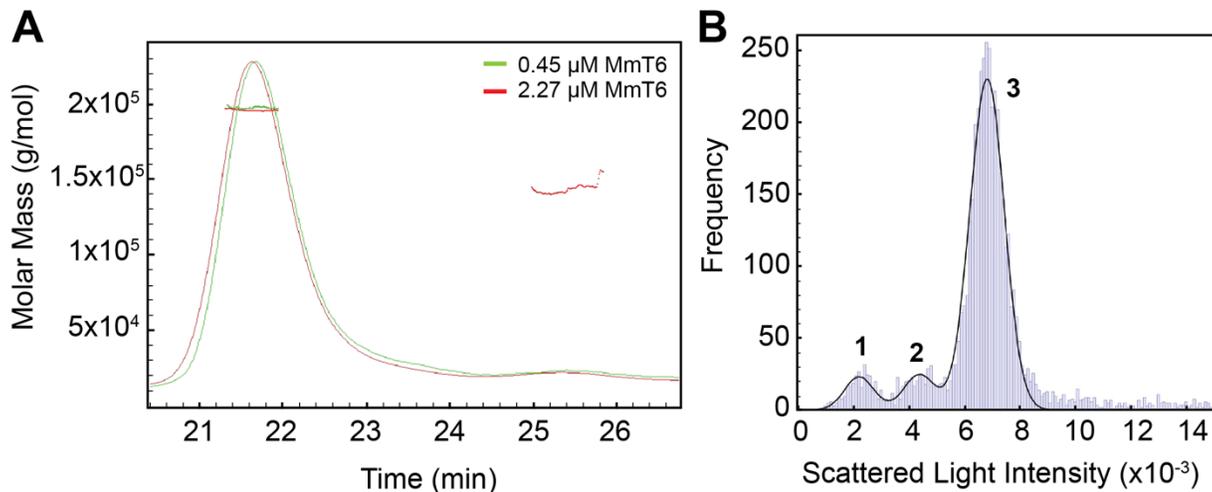
Figure 3.5: MmT6 complex size determination using Size Exclusion Chromatography coupled to Multi-Angle Light Scattering (SEC-MALS) and Interferometric Scattering Technology (iSCAT). A: SEC-MALS data from 0.45 µM (green line) and 2.27 µM (red line) MmT6. Only one major peak was confidently measured to have a size of ~200 kDa. B: iSCAT data from 20 nM MmT6. Peak 1 represents molecules sized ~60-70 kDa, determined to be Top6B monomers. Peak 2 represent molecules sized ~100-110 kDa, determined to be Top6B/Top6A heterodimers. Peak 3 represent molecules sized ~200-220 kDa, determined to be $Top6B_2Top6A_2$ heterotetramers.

3.2.3 MmT6 supercoil relaxation

The following section will detail experimental work conducted to explore the DNA relaxation activity of MmT6 using ensemble agarose gel-based methods. As described above, the relaxation activity of MmT6 can be probed in this way due to the topological status of a plasmid altering the way it migrates through a gel in response to an electrical field. Supercoiled plasmid DNA is more compact and will travel faster than the more openly-structured relaxed DNA, with the linearised plasmid running true to size, its gel position dependent on the size of the plasmid used. In this work the plasmid pBR322* (4325 bp) was heavily utilised, which is of *E. coli* origin and derived from one of the first widely-used cloning vectors, pBR322 (4361bp) [235]. The pBR322* plasmid has minor sequence changes, namely a single G to T transversion making pBR322* a high copy number plasmid [236], and a deletion which disrupted the tetracycline resistance gene.

Negatively supercoiled pBR322*, was used in preliminary assays exploring MmT6 relaxation activity conditions and established a starting point, based on previous work [62, 167, 171, 193-195], for the optimisation and characterisation thereof. The first topo VI to be characterised, *S. shibatae* topo VI (SsT6), revealed an enzyme capable of DNA relaxation and decatenation activity in the presence of ATP and Mg^{2+} [167], also shown later to be true for MmT6 [62, 195]. At the onset of this project, it was appropriate to confirm these findings with the MmT6 purified above, which indeed demonstrated the strict necessity for the both ATP and Mg^{2+} during DNA relaxation (Figure 3.6). As will be seen later, DNA cleavage activity by MmT6 was promoted by the use the use of ADPNP, a non-hydrolysable analogue of ATP, which stabilised the MmT6-dependent DNA cleavage complex and generated linearised pBR322* (Figure 3.6) (see section 3.2.5 for detail on MmT6-dependent DNA cleavage). With basic conditions for activity established and in agreement with prior work, the activity of MmT6 was systematically explored, beginning with DNA relaxation and followed by decatenation (Section 3.2.4), DNA cleavage (Section 3.2.5), DNA binding (Section 3.2.6) and ATPase activity (Section 3.2.7).
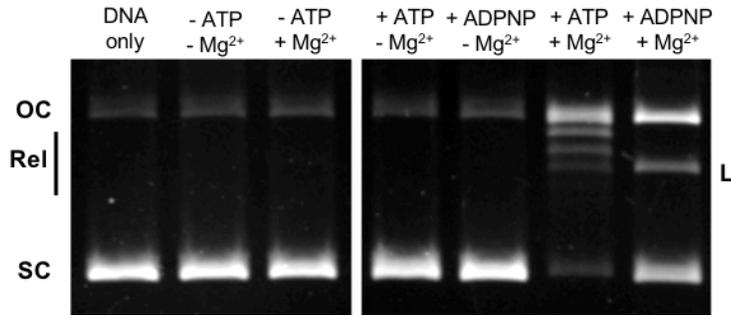
Figure 3.6: MmT6 activity conditions. MmT6 (20 nM) was incubated with 2 nM negatively supercoiled pBR322* at 37°C for 30 min. The conditions varied were the presence or absence (+ or - respectively) of 1 mM ATP, 1 mM ADPNP or 10 mM MgCl$_2$. The DNA only sample is incubated in the absence of MmT6. Samples containing ADPNP were treated with SDS and proteinase K, prior to electrophoresis. Samples were run on a 1% (w/v) native agarose gel for 10 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination. OC: open circular or nicked, Rel: relaxed, SC: supercoiled, and L: linear.

3.2.3.1 DNA topology

There are numerous variables affecting *in vitro* MmT6 activity, which can be modified to further explore its enzymatic properties and mechanism. The first described here is DNA topology. Depending on the direction of helical rotation, constrained DNA can become either positively- (left-handed) or negatively- (right-handed) supercoiled, known as overwound and underwound, respectively [6]. Many topos display a preference for DNA supercoil chirality, for instance *E. coli* topo IV and human topo IIα preferentially relax positively-supercoiled DNA [119, 120, 156], whilst *E. coli* topo I prefers negatively-supercoiled [12, 20]. Some topos, like yeast topo II and human topo IIβ display no preference, relaxing both chiralities at highly comparable rates [120, 145, 237]. To ascertain whether MmT6 displayed a chiral preference for relaxation, it was titrated into an assay containing 2 nM of either positively- or negatively-supercoiled pBR322*, 1 mM ATP, 10 mM MgCl$_2$, and incubated at 21 or 37°C for 30 min. Both temperatures were explored owing to the fact that both the ATPase (section 3.2.7) and single-molecule magnetic tweezers (Chapter 4) assays had to be conducted at 21°C. *M. mazei* itself, is a mesophilic archaeon that lives in freshwater and marine environments, with an optimal temperature of 30-40°C [238, 239]. And so, whilst 21°C is below the optimum, it is not unlikely, based on its habitat, that *M. mazei* would experience, and persist in, reduced temperatures. This means

that exploring activity at a lower temperature does not render the finding insignificant to *in vivo* extrapolations, and also offers a solution to slowing the reaction so that more subtle effects can be explored. The chirality-dependence on MmT6 relaxation rate is a prime example, as at 21°C a more significant difference in rate can be seen, an effect which is reduced considerably at 37°C (Figure 3.7). Based on the results in Figure 3.7, at 21°C, in the presence of negative supercoils, ~2-fold more enzyme is required to achieve a similar level of relaxation as positive supercoils, indicating a mild preference for the relaxation of positive writhe. In Chapter 4, the magnetic tweezers data corroborates this result, however, the *in vivo* significance of this 2 to 3-fold rate enhancement on positive writhe isn't clear, particularly as at 37°C, the effect is barely noticeable. *E. coli* topo IV and human topo IIα display ~25-fold and 10-fold rate enhancements, respectively, in the presence of positive writhe over negative at optimal temperatures [119, 120, 156], implying this chirality-dependent behaviour is more likely to effect *in vivo* roles. Nevertheless, both the data discussed here and in Chapter 4, suggest MmT6 has an intrinsic mild preference for the relaxation of positive DNA writhe. When SsT6, the first topo VI characterised, was tested in the presence of both positive and negative supercoils, no appreciable difference could be seen between the two when incubated with SsT6 at optimal temperature [167]. This is in agreement with data here, which shows that at 37°, MmT6 relaxation activity on either chiralities is very similar, as determined from the gels.
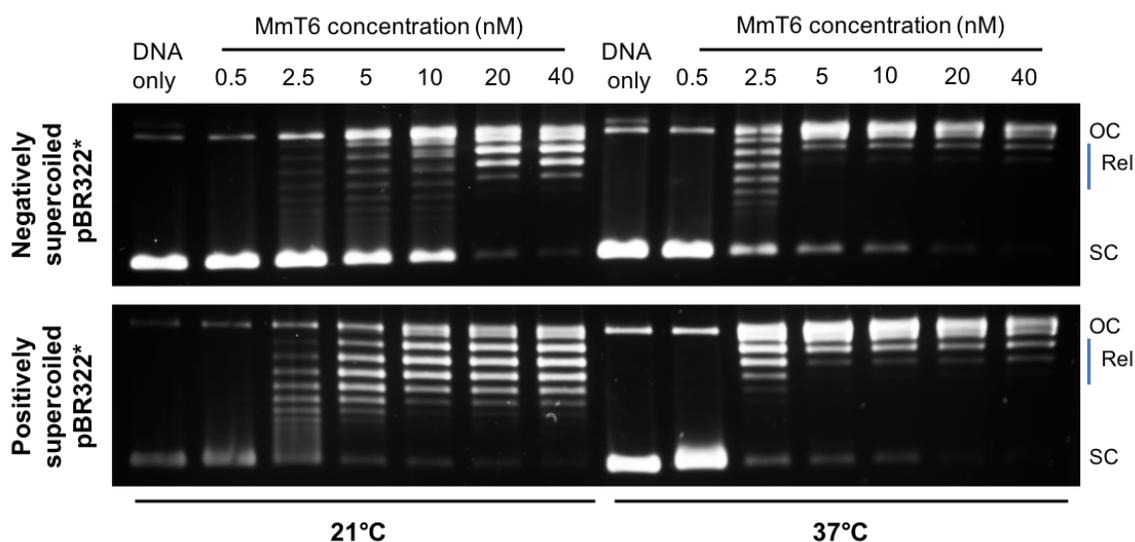
Chapter 3: Ensemble Results

Figure 3.7: MmT6 titration, exploring relaxation activity of negatively or positively-supercoiled pBR322* at 21°C and 37°C. MmT6 concentration was varied from 0.5 to 40 nM, and was incubated with 2 nM pBR322*, 1 mM ATP, 10 mM MgCl$_2$, for 30 min, at either 21°C or 37°C. OC: open circular or nicked pBR322*, Rel: relaxed topoisomers of pBR322*, and SC: supercoiled pBR322*. The DNA-only sample was incubated in the absence of MmT6. Samples were run on a 1% (w/v) native agarose gel for 12 hr at ~2 Vcm^{-1}, stained with 0.5 μg/mL ethidium bromide and imaged under UV illumination.

This result is further supported by exploring the ensemble relaxation behaviour of MmT6 over time (Figure 3.8). Once again, the difference in rate can be seen more clearly when assayed at 21°C, showing the relaxation of positively-supercoiled pBR322* to reach completion 2-3-fold faster than negatively-supercoiled, in the presence of 20 nM MmT6. Interestingly, increasing the temperature from 21 to 37°C had a more significant effect on rate for the relaxation of negative supercoils, increasing rate ~2.5-fold (full relaxation achieved by 6 min for 37°C, and by 15 min for 21°C), compared to the ~0.5-fold increase for positive (full relaxation achieved by 4 min at 37°C, and by 6 min at 21°C). As explored thoroughly in Chapter 4, the DNA crossing geometry preference for MmT6 is slightly below 90°, an angle more common in positive supercoils than negative. Therefore, the angle fluctuation frequency, which would increase as a function of temperature, would likely have a more profound effect on negative supercoil relaxation, with negative supercoils sampling the MmT6 preferential angle more frequently. As discussed above, and in Chapter 4, it is not clear whether this subtle difference in relaxation rates has *in vivo* implications for MmT6 activity, however as *M. mazei* encodes a DNA gyrase, having MmT6 preferentially relax positive writhe over negative may reduce any interference with the DNA gyrase activity. It may also allow MmT6 to assist in the relaxation of positive writhe generated as a result of DNA replication and transcription. However, as will be shown, MmT6 is likely a preferential DNA decatenase, therefore differential effects in relaxation activity may be a by-product of its decatenation mechanism.

Time courses, such as those in Figure 3.8 can also inform on the processivity of the enzyme in question. In previous work, MmT6 was determined to be highly distributive, using a plasmid-chase assay [62]. Visual inspection of the gel in Figure 3.8, indicated that MmT6 under

the conditions used here, appeared to function via a mixture of processive and distributive behaviours, with negative supercoil relaxation appearing more processive, while the activity on positive was more distributive. The argument for processivity originates from the fact that relaxed topoisomers can be seen at very early time points (1-2 mins) for both chiralities, but more significantly in the case of negative supercoil relaxation. This form of analysis, however, is not definitive and further exploration would be extremely useful. In Chapter 4, the processivity of MmT6 will be explored again but it is important to note that magnetic tweezers DNA substrates are not identical to those found within cells, which means *in vitro* ensemble characterisation using supercoiled plasmid DNA may be more informative when it comes to addressing MmT6 processivity *in vivo*.
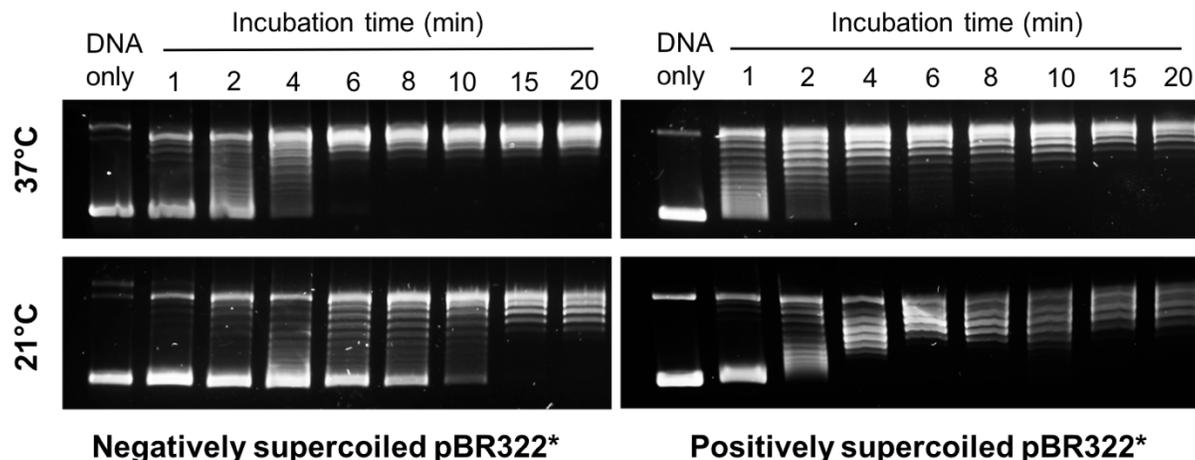
Figure 3.8: Time course of MmT6-dependent relaxation of negatively- or positively-supercoiled pBR322* at 21°C and 37°C. Assay was conducted using 20 nM MmT6, 1 mM ATP, 10 mM MgCl$_2$ and 2 nM pBR322*. The DNA-only sample was incubated in the absence of MmT6. Samples were run on a 1% (w/v) native agarose gel for 10 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination.

3.2.3.2 Temperature

As discussed in Section 3.2.3.1, varying the temperature at which the relaxation reaction was incubated, assisted in revealing subtle enzymatic effects. Here, the temperature was varied more dramatically and with the purpose of exploring the thermostability of MmT6, through relaxation activity. As the optimum growth temperature of *M. mazei* is 30-40°C [119, 120, 156], with MmT6 activity determined optimal at 37°C [62, 195], the results shown in Figure 3.9

Chapter 3: Ensemble Results

substantiate this, finding that the MmT6 purified in this work also relaxed negatively supercoiled pBR322* optimally at 40°C. Relaxation activity, whilst diminished, was detected at 20-30°, with a far steeper decrease in activity as the temperature exceeded 40-50°C. This suggested MmT6 activity is not particularly stable at high temperatures which may be a function of the enzyme itself becoming structurally disrupted, or instead may be due to the interaction with the DNA becoming less efficient. Previous work [62], as well as work done in this chapter and in Chapter 4, indicated that MmT6 preferentially binds DNA crossings, rather than binding the G-segment before subsequently capturing a T-segment. The movement of the DNA would increase with temperature, which may mean DNA crossings are less stable in position through time and may affect the MmT6 relaxation rate through prevention of DNA binding. However, SsT6, despite being from a hyperthermophile and determined optimally active at 80°C, was reduced to being only 40% active in DNA relaxation after a 40 min preincubation at 80°C, even though this temperature was determined optimal in the absence of preincubation [167]. This suggested SsT6 was not particularly thermostable, which may be true for all topo VI enzymes. A likely cause of this instability is the overall holoenzyme structure. Topo VI enzymes are heterotetramers with all four subunits held together via non-covalent interactions, and as described in Section 3.2.1, the stable expression and purification of topo VI relies on dual expression of both the Top6A and Top6B subunits. This is hypothesised to be due to one or both subunits relying on the other for stability. In addition to this, as described in section 3.2.2, the enzyme appears to be a heterotetramer in solution, meaning the Top6A/Top6B dimer must find and interact with another dimer, to form stable and active enzymes prior to interaction with DNA. It may be, therefore, that with both these structural implications, an increase in temperature rapidly disrupts the heterotetramer arrangement, first separating the two Top6A/Top6B dimers, before disrupting the interaction between the Top6A and TopB subunits within each dimer. Future work using biophysical techniques would be necessary to further explore the effects of temperature on MmT6 structure and activity, for example, preincubation of MmT6 at various temperatures followed by SEC-MALS.
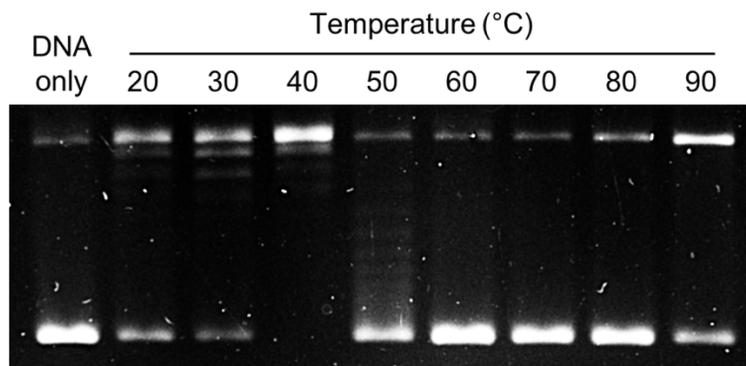
Figure 3.9: Effect of temperature on MmT6-dependent relaxation activity. MmT6 (37 nM) was incubated with 2 nM negatively supercoiled pBR322*, 1 mM ATP and 10 mM MgCl$_2$, for 30 min at 20 to 90°C. The DNA only sample was incubated in the absence of MmT6 and at 90°C. Samples were run on a 1% (w/v) native agarose gel for 10 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination.

3.2.3.3 Divalent metal cation identity and concentration

All type II and type IA enzymes require Mg^{2+} during topology modulation, as it is integral to topo binding and cleavage activity, likely due to a coordination or stabilisation effect between the active-site tyrosine and the scissile phosphate, or their transition states [59]. In addition, type IB activity is stimulated in the presence of Mg^{2+}, however it does not directly interact with the protein [240, 241]. The TOPRIM domain of type II and type IA topos harbours a DxD motif which coordinates Mg^{2+} [58]. This motif can, however, bind other divalent metal cofactors that may or may not permit topo activity. A common example is the use of Ca^{2+} in DNA-cleavage stimulation by type II topos such as DNA gyrase [242, 243] and *Drosophila melanogaster* topo II [244]. Although the main drive behind exploring divalent metal ion identity related effects on MmT6 activity began with the goal to generate more substantial DNA cleavage (Section 3.2.5.2), its effects on relaxation activity were also systematically explored.

The divalent metal ions used, all derived from the chloride salt, were magnesium (Mg^{2+}), manganese (Mn^{2+}), zinc (Zn^{2+}), calcium (Ca^{2+}), cobalt (Co^{2+}) and nickel (Ni^{2+}), titrated into a MmT6 relaxation assay at 0.5 to 100 mM (Figure 3.10). Zinc, calcium and nickel did not support any relaxation, however calcium has been shown to support human topo II [245] and *Mycobacterium tuberculosis* DNA gyrase [246] relaxation, as well as cleavage by SsT6 [194]. Mg^{2+}, as expected, was the optimal divalent metal ion cofactor during relaxation at a

concentration of 5-25 mM. However, at higher concentrations of Mg^{2+} (50 mM and above), relaxation activity was abolished. Possible explanations for this involve disruption of DNA binding through effects of increased ionic strength, however with reference to Figure 3.17 exploring DNA cleavage as a function of metal ion cofactor identity, 50 mM Mg^{2+} supported robust levels of DNA cleavage. Therefore, the lack of relaxation activity at 50 mM cannot be due to interference with G-segment binding. Rather, it suggested that 50 mM Mg^{2+} disrupts relaxation activity through prevention of T-segment capture, rather than G-segment binding and cleavage. At 100 mM $MgCl_2$, both relaxation and cleavage were inhibited, which may indicate that this concentration abolished DNA binding altogether.
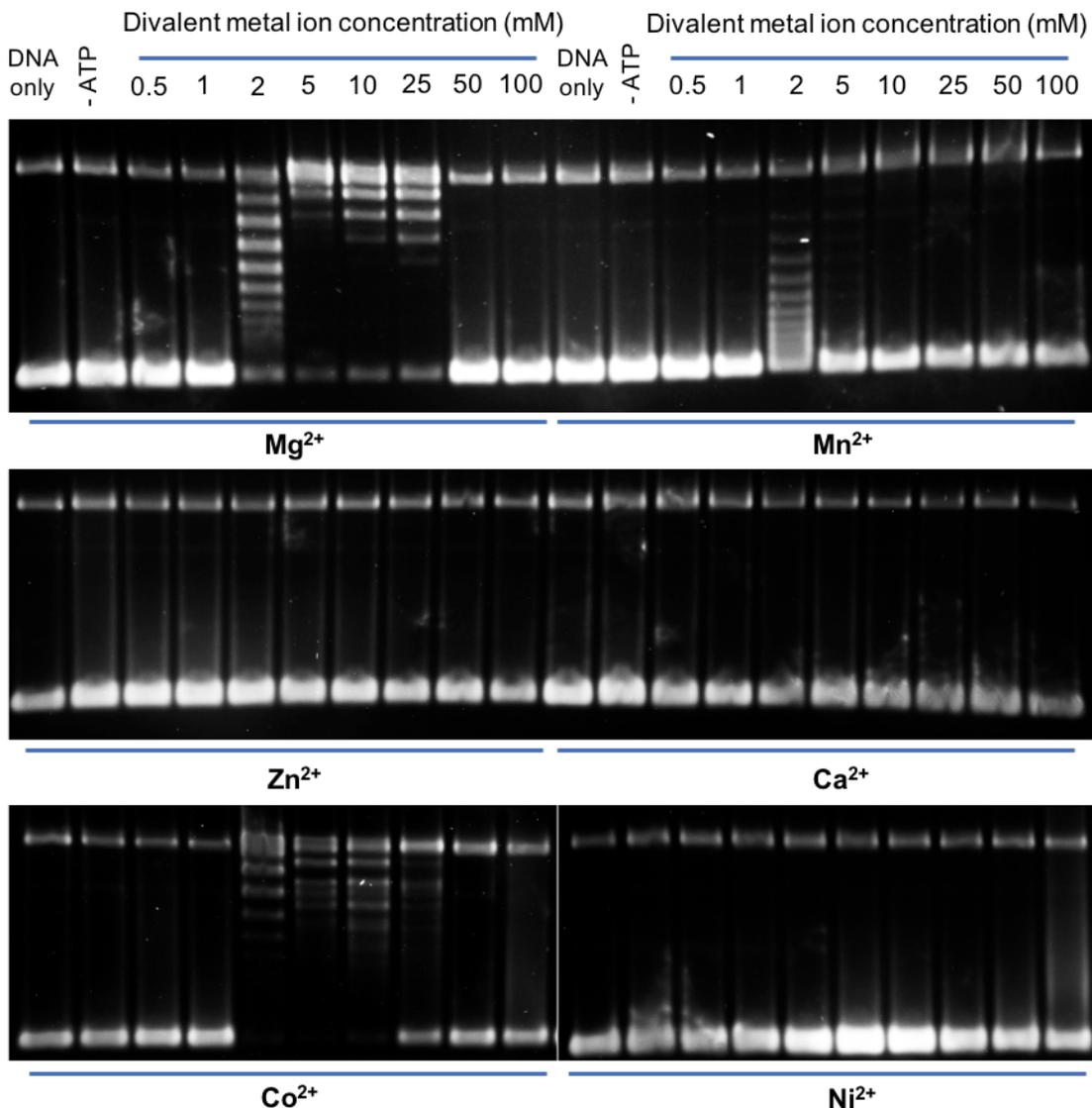
Figure 3.10: Exploring the effects of varying divalent metal ion identity and concentration on the relaxation activity of MmT6. MmT6 (20 nM) was incubated with 1 mM ATP, at 37°C for 30 min in

the presence of different divalent metal ions. Divalent metal ions explored were magnesium chloride (MgCl$_2$ or Mg^{2+}), manganese chloride (MnCl$_2$ or Mn^{2+}), zinc chloride (ZnCl$_2$ or Zn^{2+}), calcium chloride (CaCl$_2$ or Ca^{2+}), cobalt chloride (CoCl$_2$ or Co^{2+}), and nickel chloride (NiCl$_2$ or Ni^{2+}), with each varied in concentration from 0.5 to 100 mM. The DNA-only sample was incubated in the absence of MmT6 and in the presence of 100 mM of the respective metal cation salt. All samples were run on a 1% (w/v) native agarose gel for 15 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination.

Mn^{2+} supported a modest amount of relaxation at 2 mM, in line with results found for human topo II [245, 247] *M. tuberculosis* DNA gyrase [246] and Vaccinia topo I [248], but not for *Drosophila melanogaster* topo II [249]. Interestingly, Co^{2+} supported the most relaxation after Mg^{2+}, with 2 mM being optimal. This was also found to be the case for human topo IIα [250] and Vaccinia topo I [251], but not Yeast type I topo [136]. This is surprising considering the *in vivo* genotoxic effects of Co^{2+}, hypothesised to be caused by either interference with DNA repair proteins [252], oxidative damage [253] or indeed poisoning of topo II activity promoting formation of DNA cleavage complexes [250]. The cleavage promoting effect of Co^{2+} can also be seen for MmT6 in Figure 3.10, as increasing Co^{2+} concentration results in the progressive appearance of a linear band between 5-25 mM. As no DNA linearisation is seen above 25 mM, this suggested the DNA damage is dependent on MmT6 activity, rather than protein-independent oxidative damage.

3.2.3.4 Potassium glutamate versus potassium acetate

M. mazei is an acetate-metabolising archaeon, found to convert acetate to methane as a source of energy once the extracellular concentration surpassed 0.4 mM [254, 255]. In assays conducted here, potassium glutamate was selected as the monovalent salt at 100 mM, based on previous work [62, 195]. However, as MmT6 is from an organism that preferentially metabolises acetate, the effect of using potassium acetate, rather than glutamate, on relaxation activity was explored (Figure 3.11). At the concentration commonly employed in MmT6 reactions (100 mM) both glutamate and acetate supported robust relaxation of 2 nM negatively supercoiled pBR322* by 37 nM MmT6 (Figure 3.11). Glutamate is known to have an effect on

the quality of gel electrophoresis resolution which can be seen here with topoisomer bands in the presence of acetate being far more defined. This indicated that replacing glutamate with acetate (at 100 mM) could provide a technical solution to producing gels with more distinct topoisomer distributions. However, increasing acetate concentration had a more severe inhibitory effect on relaxation activity, fully inhibiting MmT6 by 300 mM, whilst glutamate permitted some level of relaxation until 500 mM. As both were the potassium salt, this difference in activity was likely due to differential effects of the acetate and glutamate ions. Any future work should encompass the use of concentrations below 100 mM.
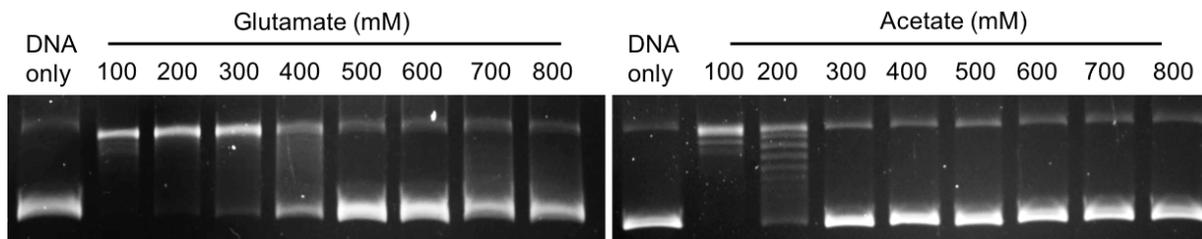
Figure 3.11: Exploring the effects of potassium glutamate and potassium acetate on MmT6-dependent relaxation activity. MmT6 (37 nM) was incubated with 2 nM negatively supercoiled pBR322*, 1 mM ATP and 10 mM MgCl$_2$, for 30 min at 37°C. The buffer conditions were varied, using 100-800 mM potassium glutamate or potassium acetate. DNA only samples were incubated in the absence of MmT6. All samples were run on a 1% (w/v) native agarose gel for 10 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination

3.2.4 MmT6 decatenation activity

Topo VI enzymes are hypothesised to be preferential DNA decatenases, based on the organisms in which these enzymes are found and the processes in which topo VI is integral, such as in *Arabidopsis*, where topo VI was shown to be vital during endoreduplication [169], as well as preliminary *in vitro* characterisation of archaeal SsT6 [167]. This section details the results of using two different DNA substrates in the characterisation of MmT6 decatenase activity, namely kDNA (Section 3.2.4.1) and singly-catenated plasmids (Section 3.2.4.2).

3.2.4.1 Decatenation using kDNA

The use of kinetoplast DNA (kDNA) in exploring the *in vitro* decatenation activity of topos, has provided a means to attaining a qualitative assessment of decatenation activity. This substrate originates from the mitochondria of trypanosomes and is formed of an extended network of interlinked DNA plasmids of two sizes, called maxicircles (20-40 kb) and minicircles (0.5-1kb) [256]. Its use can be a qualitative measure of decatenation activity, with plasmids released from the network as a result of topo-dependent activity, which can then be visualised using gel electrophoresis. When SsT6 was first characterised, the use of kDNA revealed that 3-fold less SsT6 was required to fully decatenate kDNA, than was required for the relaxation of the same amount of negatively supercoiled plasmid [167], resulting in the hypothesis that SsT6 was a preferential DNA decatenase. However, with the inability to quantify or control how many plasmids are interlinked and by how much, comparing topo-dependent decatenation activity to supercoil relaxation using a gel-based approach is limited. Regardless of this, MmT6 decatenation activity was explored using kDNA at both 37 and 21°C (Figure 3.12). If Figure 3.12 is compared to Figure 3.7, which explored the effect of MmT6 concentration on supercoil relaxation, low-level decatenation activity can be seen at both temperatures using 0.5 nM MmT6, whilst supercoil relaxation isn't detected until 2.5 nM MmT6 is used. This suggested that ~5-fold less MmT6 is required for decatenation and is in agreement with previous work done on SsT6[167]. The caveat, as previously alluded to, is that quantifying how many strand passage events are required to release plasmids from the kDNA network, compared to how many are required to begin resolving more relaxed topoisomers of supercoiled DNA, was not possible. It may be that 5-times fewer stand-passage events are required to fully decatenate kDNA than fully relax supercoiled DNA. Nevertheless, this result suggested that MmT6, like SsT6, is a preferential decatenase.
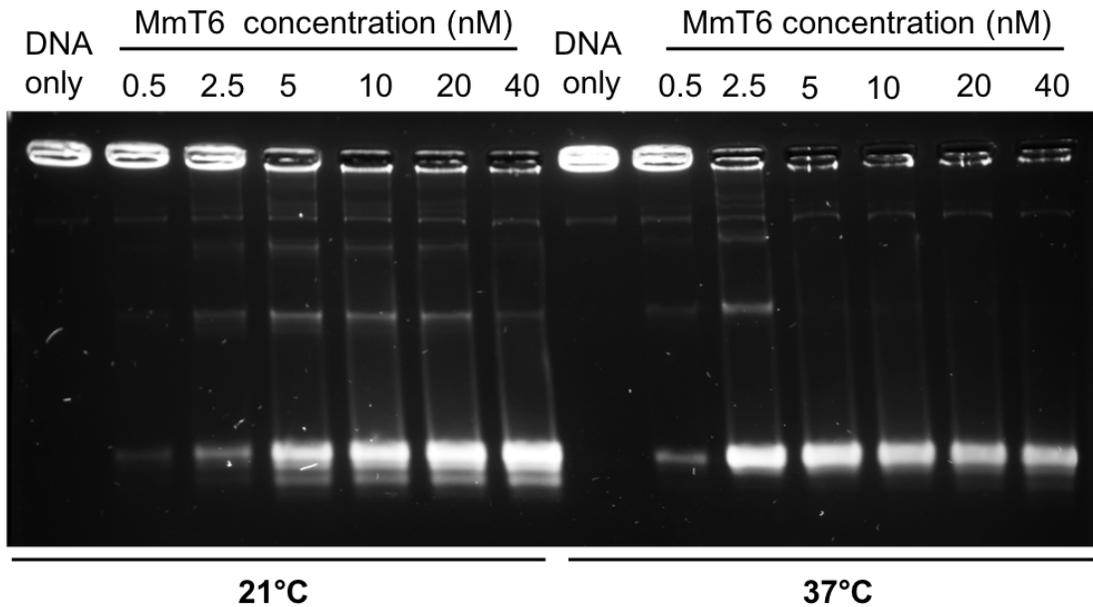
Figure 3.12: MmT6 decatenation activity using kDNA. MmT6 (0.5-40 nM) was incubated with 250 ng of kDNA in the presence of 1 mM ATP and 10 mM MgCl$_2$, at 37°C or 21°C for 30 min. As plasmids are released from the network by MmT6, they can migrate into the gel, allowing the decatenation reaction to be visualised. All samples were run on a 1% (w/v) native agarose gel for 15 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination.

3.2.4.2 Decatenation using singly-catenated plasmid DNA

As just discussed, the use of kDNA in the quantitative analysis of topo-dependent decatenation activity is not possible due to the inability to quantify the number of strand-passage events required to fully decatenate the substrate. This, therefore, left the topo field with no reliable way to explore the ensemble *in vitro* decatenation activity in a calculable sense. The use of single-molecule DNA braids, implemented in Chapter 4, has been used as a proxy for exploring decatenation activity [257], however, as DNA braids and catenated plasmids are not identical in nature, this approach is also limited in characterising bona fide decatenation activity. Recently, Nidda Waraich (JIC, Norwich, UK; Waraich et al. BioTechniques, in press, 2020) established a protocol, which generated a catenated substrate of two different-sized plasmids, catenated by a single link. This substrate allowed the quantitative analysis of MmT6 decatenation activity, and moreover, as the plasmids are also negatively supercoiled, both the decatenation and relaxation activity can be monitored within the same reaction (Figure 3.13). As seen in Figure 3.13, decatenation activity was detected at MmT6 concentrations as low as

100 pM, with the substrate fully decatenated using 500pM MmT6, even though the relaxation of the plasmids wasn't detected until 5 nM MmT6. This result indicates that MmT6 could be as much as ~10-fold more active during decatenation than relaxation. However, it must be considered that only a single strand-passage event was required to fully separate the two plasmids, whilst full relaxation required considerably more. This being said, the fact that MmT6 displayed low-level decatenation activity at a concentration 50-fold lower than the first sign of relaxation (0.1 versus 5 nM), suggested that there is a preference for decatenation by MmT6. If this was purely a function of how many strand-passage events are required, you would expect to see a mixture of decatenation and relaxation occurring simultaneously at lower MmT6 concentrations. As relaxation is entirely absent below 5 nM MmT6, this suggested that a unique feature of catenanes accelerated MmT6 activity. As explored thoroughly in Chapter 4, this is likely to be due to a preferential DNA crossing geometry close to 90° for MmT6, which appears far more often in catenanes than supercoils [119].

As the protocol for the generation of this substrate has only recently been refined, significant quantities were not available. However, once more of the substrate is produced, exploring MmT6 cleavage, binding and ATPase activity in its presence would be invaluable to further consolidating the hypothesis of MmT6 and other topo VI enzymes as preferential DNA decatenases.
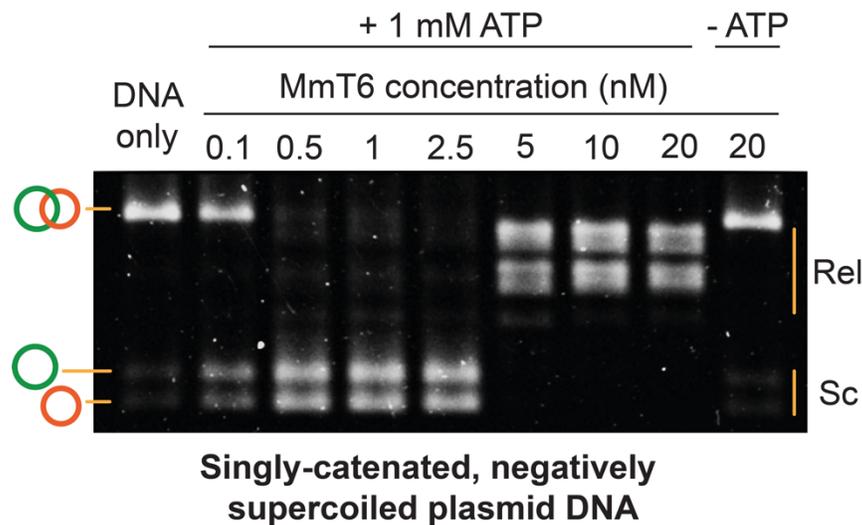
Figure 3.13: MmT6 decatenation activity using singly-catenated, negatively supercoiled plasmid DNA. A singly catenated, negatively-supercoiled plasmid (2 nM) was decatenated by 0.1-20 nM

MmT6 in the presence of 1 mM ATP and 10 mM MgCl$_2$, incubated at 37°C for 30 min. The catenated plasmids vary in size and when decatenated, can be seen as two bands that migrate further than the catenated substrate (labelled with the orange and green rings, left-hand side). As they are also negatively supercoiled; the relaxation of the plasmids can be seen at MmT6 concentrations ~10 fold higher than when full decatenation is seen. OC: open circular, Rel: relaxed and Sc: supercoiled. All samples were run on a 1% (w/v) native agarose gel for 15 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination.

3.2.5 MmT6 DNA cleavage

Exploring DNA cleavage activity of topos using gel-based methods can inform on the specifics of this stage of the reaction, in the absence of canonical activity. Topo-dependent cleavage can be explored for a number of purposes, such as characterising inhibitory compounds [258-260], determining cleavage kinetics [241, 261] and studying DNA cleavage sequence preferences [54]. Here, MmT6 DNA cleavage activity was explored with the purpose of both gaining deeper mechanistic insight into the MmT6 reaction, as well as optimising DNA cleavage in order to facilitate work presented in Chapter 5, exploring MmT6 DNA sequence preferences.

As DNA cleavage is a highly transient stage of the topo reaction mechanism, exploring it requires disruption of the canonical topo reaction. In many cases, this is achieved through the use of cleavage-complex stabilising compounds, which prevent re-ligation by binding at the sites of topo-dependent DNA breakage [259]. Other methods include the substitution of Mg^{2+} for Ca^{2+}, which promotes formation of the DNA cleavage complexes in *E. coli* DNA gyrase [242] and *D. melanogaster* topo II [244], or the use of ADPNP, which enhanced cleavage stabilisation by SsT6 [194]. For MmT6, with reference to Figure 3.6, the starting conditions for stimulating DNA cleavage were incubation of ~10-fold more MmT6 than used in relaxation assays, with 3-4 nM negatively supercoiled pBR322*, in the presence of 10 mM MgCl$_2$ and 1 mM ADPNP, at 37°C for 30 min.

In addition to the aforementioned modifications to *in vitro* protocols, resolution of topo-cleaved DNA on a gel required further processing as a result of the formation of the covalent topo-DNA complex. Type II and type IA topos form 5$^{'}$ phosphotyrosyl linkages with the DNA back bone, which are reversible during the canonical mechanism, but become stabilised during

DNA cleavage assays. If the topo is not removed the DNA will fail to migrate correctly through the gel and make resolution of the DNA species challenging. Therefore, all assays detailed in this section, post initial incubation, were followed by treatment with sodium dodecylsulphate (SDS) and proteinase K (PK), which successfully removed the MmT6 and allowed electrophoretic separation of the linear, nicked and supercoiled DNA species (see Chapter 2 for DNA cleavage assay methods).

3.2.5.1 DNA topology

As with the optimisation of DNA relaxation, MmT6-dependent DNA cleavage was explored as function of DNA topology. MmT6 (120 nM) was incubated with either 4 nM negatively-supercoiled, relaxed or linearised pBR322*, in the presence of 10 mM $MgCl_2$ and 1 mM ADPNP, at 37°C for 30 min. As expected, using negatively supercoiled pBR322* stimulated the most cleavage, followed by relaxed (Figure 3.14). The amount of cleavage on linear is harder to assess, however, very little change in the DNA substrate occurs when incubated with MmT6 (Figure 3.14). This preliminary foray into cleavage activity also highlighted another aspect of MmT6 cleavage activity, in that a significant proportion of the DNA was nicked post MmT6 incubation. This could have been due to contamination with a nickase, however as will be discussed in section 3.2.5.6, treatment with ethylenediaminetetraacetic acid (EDTA) caused nick reversal, which is a hallmark of topo-dependent cleavage activity (nickase and nuclease activity is not reversible).
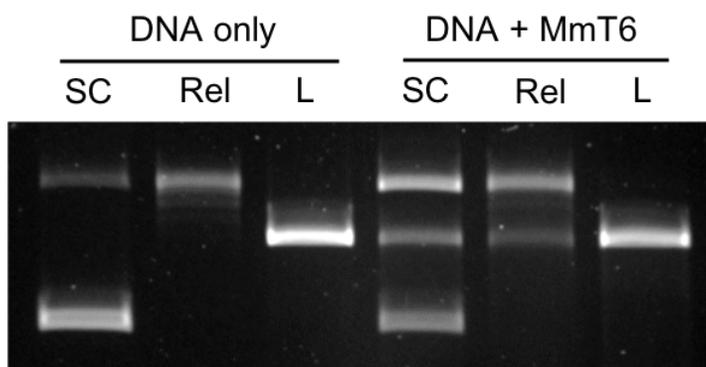
Figure 3.14: Effect of DNA topology on DNA cleavage activity of MmT6, using negatively-supercoiled, relaxed or linear pBR322*. MmT6 (120 nM) was incubated with 4 nM of either

negatively supercoiled (SC), relaxed (Rel) or linear (L) pBR322*, and 1 mM ADPNP and 10 mM MgCl$_2$, for 30 min at 37°C. Samples were run on a 1% (w/v) native agarose gel for 10 hr at ~2 Vcm^{-1}, stained with 0.5 μg/mL ethidium bromide and imaged under UV illumination.

Having established that supercoiled DNA stimulated more MmT6-dependent cleavage, further assays were conducted to explore if chirality also had an effect on cleavage, particularly with reference to section 3.2.3.1 which demonstrated the ~2-fold relaxation rate enhancement in the presence of positively supercoiled pBR322*. A simple explanation for why positive supercoils were relaxed more quickly would have been that MmT6 either bound or cleaved positive writhe more readily, however Figures 3.15-3.17 demonstrated this was not the case. To begin, an MmT6 titration was performed with either negatively- or positively-supercoiled pBR322*, which indicated that the more MmT6 used, the more DNA cleavage is attained (Figure 3.15). This assay also showed that negatively-supercoiled DNA was both nicked and linearised more than positively-supercoiled DNA, meaning the relaxation rate enhancement in the presence of positive writhe was not likely to be due to higher levels of DNA binding or cleavage. Just to note, the reason the outer lanes on the agarose gel in Figure 3.15 are less intense was due to uneven UV-transillumination, rather than sample loading errors.
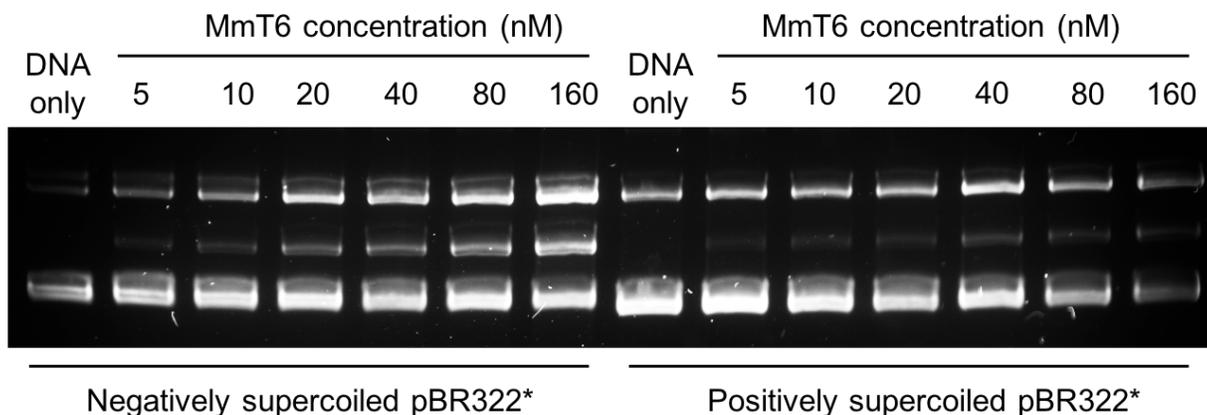
Figure 3.15: MmT6 titration to explore DNA cleavage activity using negatively and positively supercoiled pBR322*. MmT6 concentration was varied from 5 to 160 nM and incubated with 4 nM pBR322*, 1 mM ADPNP, 10 mM MgCl$_2$, at 37°C for 30 min. All samples were run on a 1% (w/v) native agarose gel for 10 hr at ~2 Vcm^{-1}, stained with 0.5 μg/mL ethidium bromide and imaged under UV illumination.

To explore the chirality-dependent cleavage reaction further a time course was performed, using 160 nM MmT6, 4 nM pBR322*, 10 mM $MgCl_2$ and 1 mM ADPNP, incubated at 37°C with samples taken at time points specified in Figure 3.16. Using ImageJ, the data attained using electrophoresis was quantified, with the percentage of DNA either remaining supercoiled, or becoming nicked or linearized, plotted as a function of incubation time. One interpretation of the MmT6 nicking behaviour would be that the single-stranded DNA break is an intermediate on the way to full DNA linearisation. This was found to be the case for *E. coli* DNA gyrase cleavage in the presence of either Ca^{2+} or inhibitors, in which DNA nicking increased at early time points only to decrease as DNA linearization increased, indicating the nicks were an intermediate [262]. As seen in Figure 3.16, this was not the case for MmT6 cleavage on either negatively- or positively-supercoiled pBR322*, with both the nicking and double-stranded DNA cleavage activity occurring simultaneously, peaking early on and then remaining constant. This implied that DNA nicking was an intrinsic feature of *in vitro* MmT6 DNA cleavage activity when incubated in the presence of ADPNP, and may suggest that for MmT6, nicking one DNA strand is not particularly well coupled to the cleavage of the second, unlike what has been found for *D. melanogaster* topo II [263] human topo IIα [264] and *E. coli* gyrase [262]. For both DNA supercoil chiralities, the amount of nicked DNA was ~50% and represented the most abundant DNA species produced by MmT6. However, the amount of linear DNA was ~40% for negative writhe and only ~28% for positive writhe, which further demonstrated that higher levels of cleavage were supported by MmT6 in the presence of negatively-supercoiled DNA. One aspect to consider, is that detection of DNA cleavage only confirms the presence of a G-segment, whilst relaxation activity requires both a G- and T-segment. Therefore, as MmT6 DNA cleavage activity was found to be enhanced in the presence of negative writhe but relaxation occurred more rapidly in the presence of positive writhe, this suggested that G-segment only binding/cleavage may be elevated when the DNA is underwound, whilst overwound DNA promotes G- and T-segment binding or T-segment strand-passage.
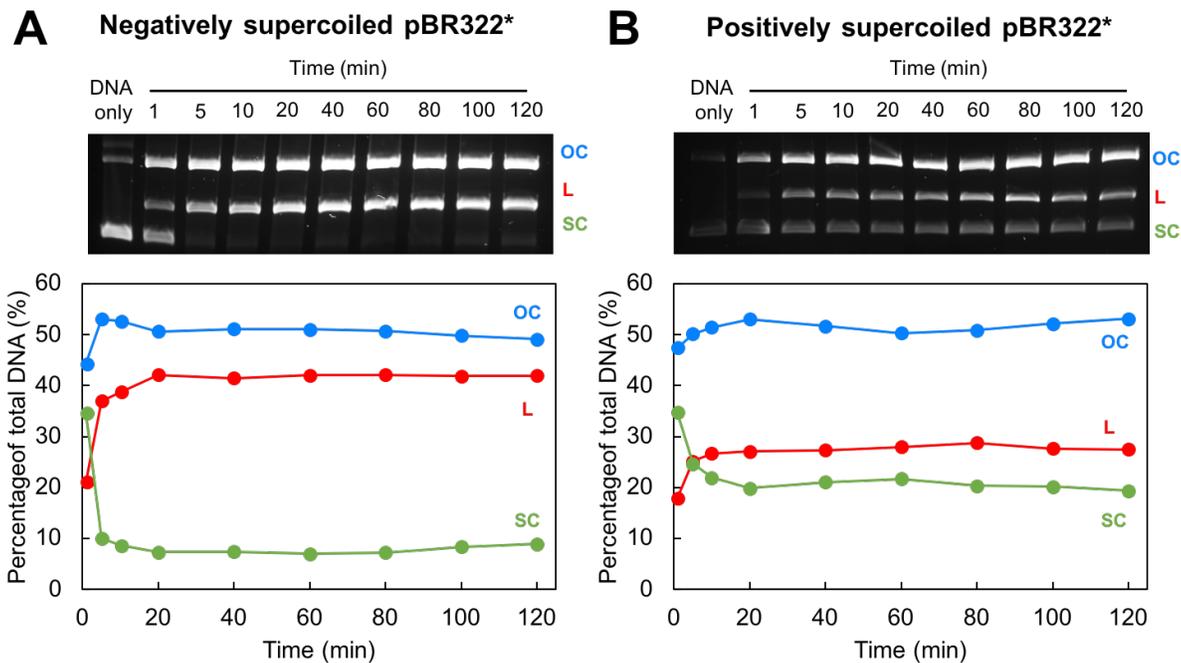
Figure 3.16: Time course of MmT6-dependent DNA cleavage activity on negatively (A) or positively-supercoiled (B) pBR322*. Assays were conducted using 160 nM MmT6, 1 mM ADPNP, 10 mM MgCl$_2$ and 4 nM pBR322*, incubated at 37°C for 30 min. The DNA-only samples were incubated in the absence of MmT6. Samples were run on a 1% (w/v) native agarose gel for 10 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination. Data used in construction of the graphs was extracted from the gels using ImageJ and analysed in Excel. OC: open circular or nicked (blue), L: linear (red), and SC: supercoiled (green).

3.2.5.2 Divalent metal cation identity and concentration

As was conducted for relaxation (Figure 3.10), the effect of divalent cation identity and concentration was explored for MmT6-dependent DNA cleavage, using 0.5 to 100 mM of Mg^{2+}, Mn^{2+}, Zn^{2+}, Ca^{2+}, Co^{2+} and Ni^{2+}, all in the form of the chloride salt. As with all cleavage reactions reported here for MmT6, 1 mM ADPNP was used in place of 1 mM ATP. SsT6, from the hyperthermophile *S. shibatae*, was shown to cleave DNA in the presence of 1 mM ADPNP and 10 mM of either Mg^{2+} or Ca^{2+} [194]. Interestingly, in the presence of ADPNP and Mg^{2+}, a large amount of DNA nicking was found, similar to that found here for MmT6 (Figures 3.15-3.17). However when the divalent cation was Ca^{2+}, slightly more linearisation was detected and significantly less nicked DNA [194]. This was not the case for MmT6 (Figure 3.17) as Ca^{2+} stimulated significantly less DNA cleavage than Mg^{2+}, and was also surpassed in efficacy by

Mn^{2+} and Co^{2+} (Figure 3.17 and 3.18). The most MmT6-depenedent DNA cleavage was detected in the presence of 25 mM $MgCl_2$, indicating that a 5-fold increase in Mg^{2+} concentration above what was determined optimal for relaxation (5 mM, see Figure 3.10), maximised the amount of DNA linearisation achieved. As mentioned in section 3.2.3.3, DNA relaxation was inhibited by 50 mM Mg^{2+}, whereas DNA cleavage was still robust up until 100 mM. Both these
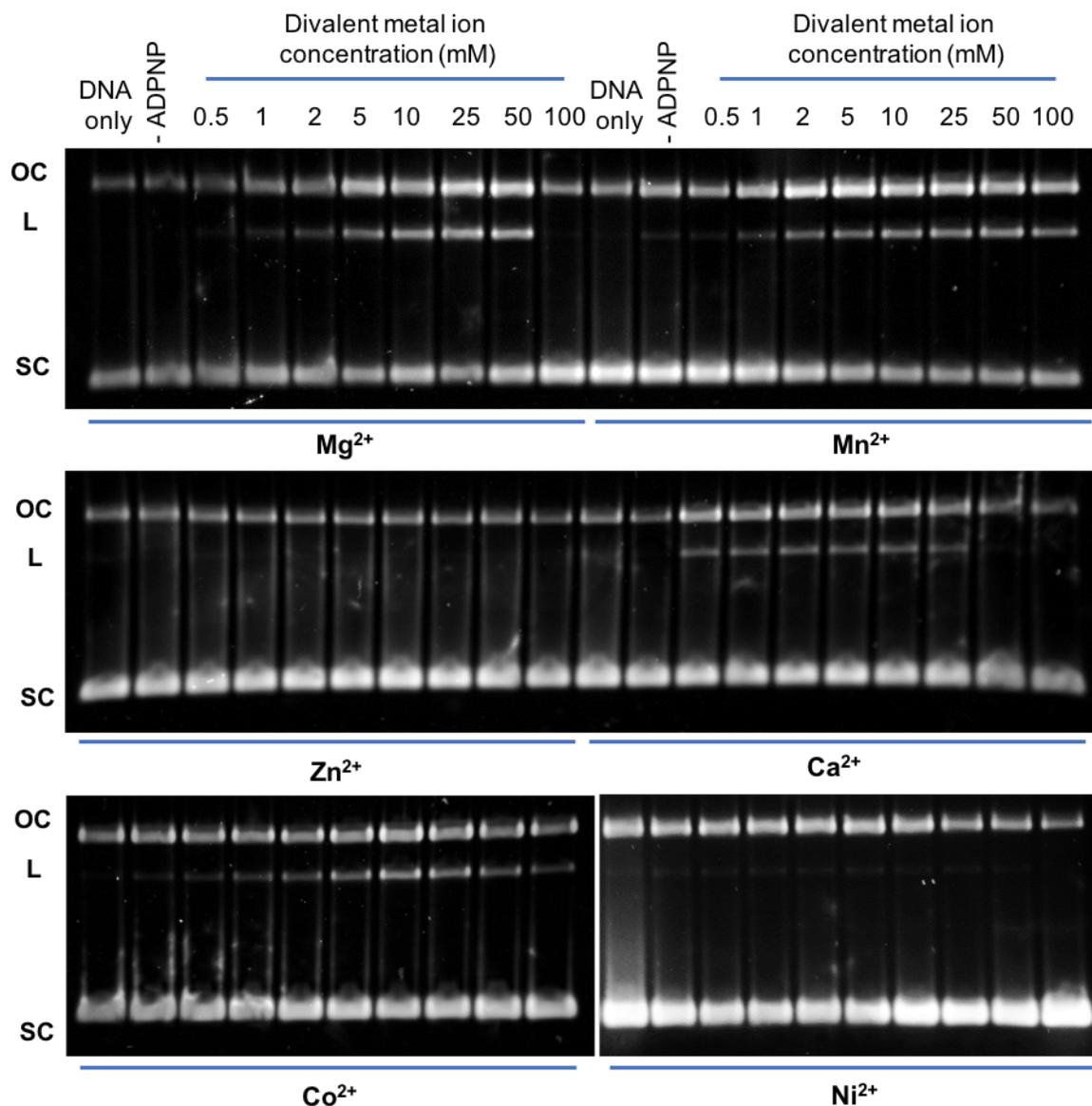
Figure 3.17: Exploring the effects of varying divalent metal ion identity and concentration on the DNA cleavage activity of MmT6. MmT6 (160 nM) was incubated with 1 mM ADPNP and 4 nM negatively-supercoiled pBR322*, at 37°C for 30 min in the presence of different divalent metal ions. Divalent metal ions explored were magnesium chloride (MgCl₂ or Mg²⁺), manganese chloride (MnCl₂ or Mn²⁺), zinc chloride (ZnCl₂ or Zn²⁺), calcium chloride (CaCl₂ or Ca²⁺), cobalt chloride (CoCl₂ or Co²⁺), and nickel chloride (NiCl₂ or Ni²⁺), with each varied in concentration from 0.5 to 100 mM. The DNA-only sample was incubated in the absence of MmT6 and in the presence

of 100 mM of the respective metal cation salt. The sample without ADPNP was run in the presence of MmT6 and 100 mM metal ion. All samples were run on a 1% (w/v) native agarose gel for 15 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination. OC: open circular or nicked, L: linear, and SC: supercoiled.

findings indicated that a reaction stage post DNA cleavage was more sensitive to Mg^{2+} and was rate limiting, potentially T-segment capture or strand-passage. Mn^{2+} also stimulated an appreciable level of DNA linearisation by MmT6, and while the maximum (15 ± 3%) was ~10% less that that stimulated by 25 mM Mg^{2+} (26 ± 4%), Mn^{2+}-stimulated cleavage remained at a constant maximum level between 25 and 100 mM, rather than being inhibited at 100 mM, which occurred in the presence of Mg^{2+} (Figure 3.17). The cause of this inhibition at 100 mM Mg^{2+} could be due the disruption of DNA binding by MmT6, or potentially prevention of ADPNP binding to Top6B, as Mg^{2+} is important for nucleotide binding to the GHKL domain [59], whilst Mn^{2+} (as well as Co^{2+}, Ca^{2+}, Zn^{2+} and Ni^{2+}) is not.

No cleavage was detected in the presence of Zn^{2+}, and only a very small amount in the presence of Ni^{2+}, which is unlikely to be MmT6-dependent as it is also present in the DNA-only sample and in the absence of ADPNP (Figure 3.17 and 3.18). Neither of these metal cofactors were able to stimulate relaxation either (Figure 3.10), indicating that either DNA cleavage is inhibited, or DNA binding altogether. Ni^{2+} is also a well-known cause of oxidative DNA damage, along with Fe^{2+} (iron) and Cu^{2+} (copper) [265], both of which weren't explored here for this reason. As the work here was mainly concerned with exploring conditions in order to maximise the levels of MmT6-dependent cleavage for research reported in Chapter 5, further exploration into how these various metal ions affected MmT6 activity was not a priority. However, future work could encompass exploring DNA binding in the presence of each cation to ascertain further detail on which reaction stage is being inhibited.
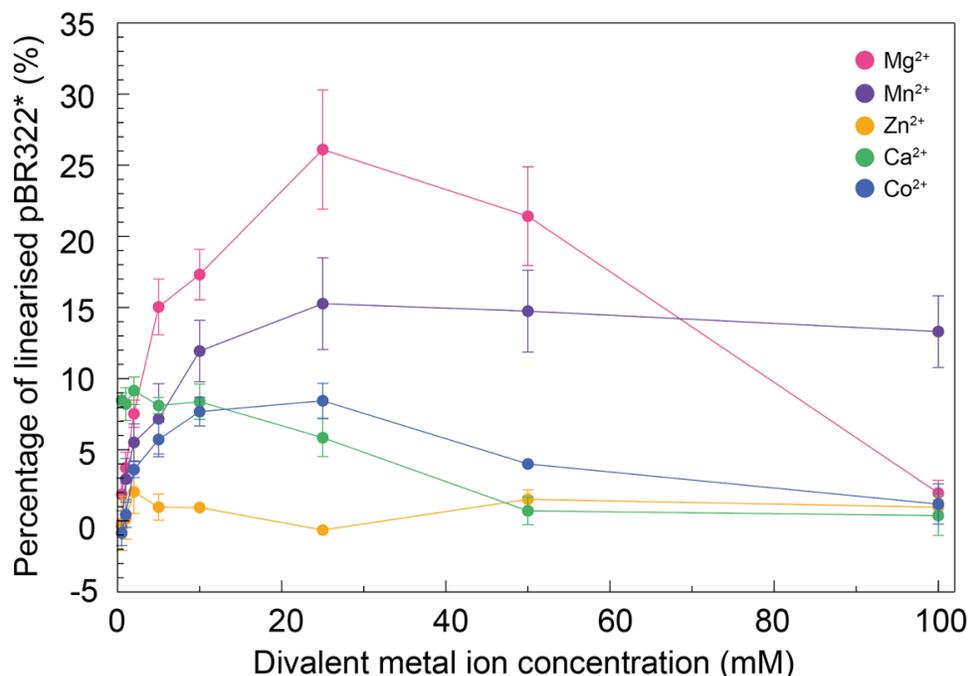
Figure 3.18: MmT6-depednent pBR322* linearization in the presence of different divalent metal ions. Using Image J and Excel, data was extracted from 3 separate assays, examples of these are seen in Figure 3.17, and the percentage of linearised DNA ± standard error (SE) was plot as a function of divalent metal ion concentration. The divalent metal ions explored were Mg^{2+} (pink), Mn^{2+} (purple), Zn^{2+} (yellow), Ca^{2+} (green), an Co^{2+} (blue).

3.2.5.3 Buffer pH

The next variable explored, for both DNA relaxation and cleavage by MmT6 was buffer pH, using Bis-Tris propane at pH 6-9 (Figure 3.19). Being able to assay such a wide range of pHs using a single buffer is facilitated in this case by the fact that Bis-Tris propane has two pK_a values, 6.8 and 9.0. The relaxation portion of this assay has been presented alongside the cleavage assay, rather than being discussed above in section 3.2.3, as pH appeared to affect relaxation and cleavage in different ways, and side by side comparison facilitated discussion more easily. MmT6-dependent DNA relaxation and cleavage is optimal at ~pH 7, however lower pH (pH 6) inhibited relaxation more significantly than higher pH (pH 8-9), whilst DNA cleavage remained robust at pH 6 (Figure 3.19). This suggested that a reaction stage post DNA cleavage was sensitive to acidic pH and was rate-limiting during the relaxation reaction, such as T-segment capture or strand-passage. DNA cleavage is inhibited progressively more between pH 8-9, which is true for relaxation, hence inhibition of DNA binding or cleavage likely

explains the relaxation inhibition under these conditions. D. *melanogaster* topo II demonstrated enhanced DNA cleavage (mainly nicking) at pH 5 [263], and human topo II was found to be a potent DNA damaging agent *in vivo* and *in vitro* under acidic conditions [266], corroborating the data found here, which also found DNA cleavage by MmT6 to remain robust at lower pH. Future work should explore the effect of even lower pH on MmT6 as the range used here was narrow.
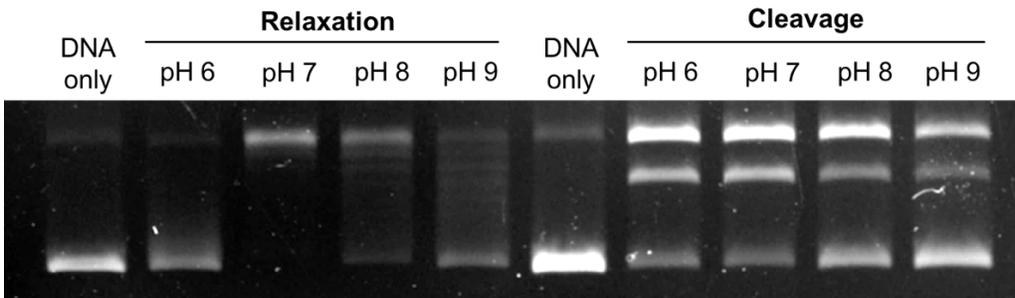
Figure 3.19: Exploring the effect of buffer pH on MmT6-dependent relaxation and DNA cleavage activity. For relaxation, MmT6 (12 nM) was incubated with 2 nM negatively supercoiled pBR322*, 1 mM ATP and10 mM MgCl$_2$, for 30 min at 37°C. For cleavage, MmT6 (120 nM) was incubated with 3 nM negatively supercoiled pBR322*, 1 mM ADPNP and10 mM MgCl$_2$, for 30 min at 37°C. For both relaxation and cleavage, the pH of the buffer (Bis-Tris propane) was varied from 6 to 9. The DNA-only sample was incubated in the absence of MmT6 and at pH 7. Samples were run on a 1% (w/v) native agarose gel for 10 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination

3.2.5.4 Temperature

Similar to data presented in Figure 3.9, exploring the effect of temperature on MmT6-dependent relaxation, the effect of temperature on DNA cleavage was determined, albeit across a narrower range. Over 30-50°C, MmT6 DNA cleavage activity did not change (Figure 3.20), even though relaxation was severely compromised at 50°C (Figure 3.9), which again suggested that a stage post DNA cleavage is sensitive to elevated temperature, the likely cause being disruption of T-segment capture or strand-passage (as was found for Mg^{2+} concentration and acidic pH). T-segment capture is an attractive candidate to explain the differences between relaxation and DNA cleavage activity at 50°C, as it has been mentioned in this chapter and will be explored further in both sections below and in Chapter 4, as well as in published work [62], MmT6 preferentially acts on preformed DNA crossings in which the T- and G-segments are bound simultaneously. At higher temperatures, DNA will move more rapidly and to a greater

degree, so may be impeding T-segment binding, whilst maintaining similar levels of G-segment engagement, therefore affecting DNA relaxation but not necessarily DNA cleavage. Future work should take advantage of the effects that temperature, Mg^{2+} and pH have on MmT6 relaxation and cleavage activity to further parse out the specifics of the reaction.
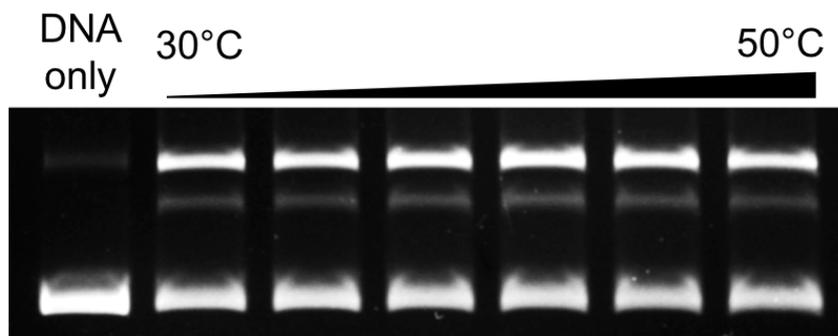
Figure 3.20: Effect of temperature on MmT6-dependent DNA cleavage activity. MmT6 (80 nM) was incubated with 4 nM negatively supercoiled pBR322*, 1 mM ADPNP and 10 mM MgCl₂, for 30 min at 30 to 50°C. The DNA only sample was incubated in the absence of MmT6 and at 50°C. Samples were run on a 1% (w/v) native agarose gel for 10 hr at ~2 Vcm⁻¹, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination.

3.2.5.5 Potassium glutamate versus potassium acetate

As already discussed in Section 3.2.3.4, *M. mazei* is an acetate-metabolising archaeon, found to convert acetate to methane as a source of energy once the extracellular concentration surpassed 0.4 mM [254, 255]. So, along with exploring the effect of replacing glutamate with acetate in DNA relaxation, the effect on DNA cleavage was also explored (Figure 3.21). The results are underwhelming, showing that cleavage activity in the presence of both acetate and glutamate reduced with increasing concentration, the presence of either salt causing the DNA to migrate less efficiently which reduced resolution of the DNA species at higher concentrations. These results also suggested that the loss of relaxation activity was a function of disrupted DNA binding or cleavage.
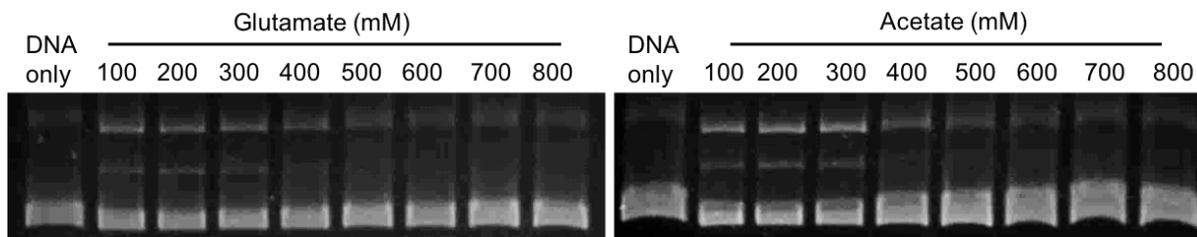
Figure 3.21: Exploring the effects of potassium glutamate and potassium acetate on MmT6-dependent DNA cleavage activity. MmT6 (120 nM) was incubated with 3 nM negatively supercoiled pBR322*, 1 mM ADPNP and 10 mM MgCl$_2$, for 30 min at 37°C. The buffer conditions were varied, using 100-800 mM potassium glutamate or potassium acetate. DNA only samples were incubated in the absence of MmT6 and using 800 mM potassium glutamate or potassium acetate. All samples were run on a 1% (w/v) native agarose gel for 10 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination

3.2.5.6 EDTA-reversal of DNA cleavage complexes

Topos, of course, are not the only enzymes which utilise DNA breakage during their reaction mechanism. There are numerous others such as endonucleases; including those important for apoptosis [267] and restriction endonucleases which protect bacteria and archaea from invading viral DNA [268], exonucleases; including those critical for proof-reading and polymerase error correction during DNA replication [269], DNA nickases; such as those involved in the CRISPR/Cas9 genome editing system [270], and the meiotic Spo11 protein; which generates double-stranded DNA breaks during meiotic recombination [203]. However, even though all of these generate DNA cleavage sites on DNA, Spo11 in particular using a similar mechanism to topos [204], the cleavage is not reversible. Endonucleases, exonucleases and nickases do not form covalent DNA/protein intermediates and once the DNA is cleaved, they dissociate. Even though Spo11 does form a covalent intermediate, it requires endonucleolytic excision [203], as the DNA break cannot be reversed by the enzyme. Topos, on the other hand, can efficiently re-ligate the DNA backbone post strand-passage, and it is this behaviour which distinguishes topo-dependent DNA cleavage from other forms.

As previously described, ADPNP was utilised here as its addition in place of ATP promoted the formation of MmT6/DNA cleavage complexes. In other words, ADPNP, as it is nonhydrolyzable, caused MmT6 to freeze during the reaction cycle, post DNA cleavage,

remaining covalently bound to the 5' scissile phosphate. This allowed the usually extremely transient DNA cleavage reaction by MmT6 to be explored, along with indicating that ATP hydrolysis is not required for DNA cleavage or those events preceding it, but is necessary for those which follow. However, these MmT6/DNA-cleavage complexes are still reversible, which was found to occur in the presence of ethylenediaminetetraacetic acid (EDTA) (Figure 3.22). These assays were conducted as a means to determine whether the elevated levels of DNA nicking in the presence of MmT6 was the consequence of a contamination, and began with the addition 40 mM EDTA to an MmT6 cleavage reaction, which was then incubated for 5 min at room temperature, before the addition of SDS and PK. As seen in Figure 3.22A, the sample treated with EDTA had far less DNA cleavage, with all of the double-stranded and most of the single-stranded cleavage absent as compared to a sample in which EDTA was omitted. This strongly suggested that, not only the DNA linearisation, but also the nicking activity was caused by MmT6 activity. It is thought that EDTA reverses the cleavage complex and promotes DNA re-ligation through its ability to chelate Mg^{2+}, removing it from the topo active site, and this indicated that the presence of the metal cofactor is an important determinant in cleavage state maintenance [262]. This reaction was explored further, finding that 8-16 mM EDTA was optimal for cleavage reversal, with increasing EDTA concentration having less of a reversal effect (Figure 3.22B). A potential explanation for this is that higher concentrations of EDTA strip the metal from the MmT6 active site too quickly for resealing to occur, suggesting the reaction is multistep and not as straightforward as just removing the Mg^{2+} from the active site. Another potential explanation involves the binding of ADPNP. This is coordinated by a Mg^{2+} ion and the binding of nucleotide to Top6B is thought to cause Top6B subunit dimerisation, forming the second protein interface and thus preventing the stabilisation of double-stranded breaks once the DNA gate is opened to allow strand-passage [64]. At high EDTA concentrations, it is possible that the Mg^{2+} is stripped from the GHKL domain, causing ADPNP to dissociate and the Top6B subunits to separate. This could potentially lead to DNA gate separation with MmT6 heterotetramers becoming dimers and thus preventing any re-ligation from occurring.
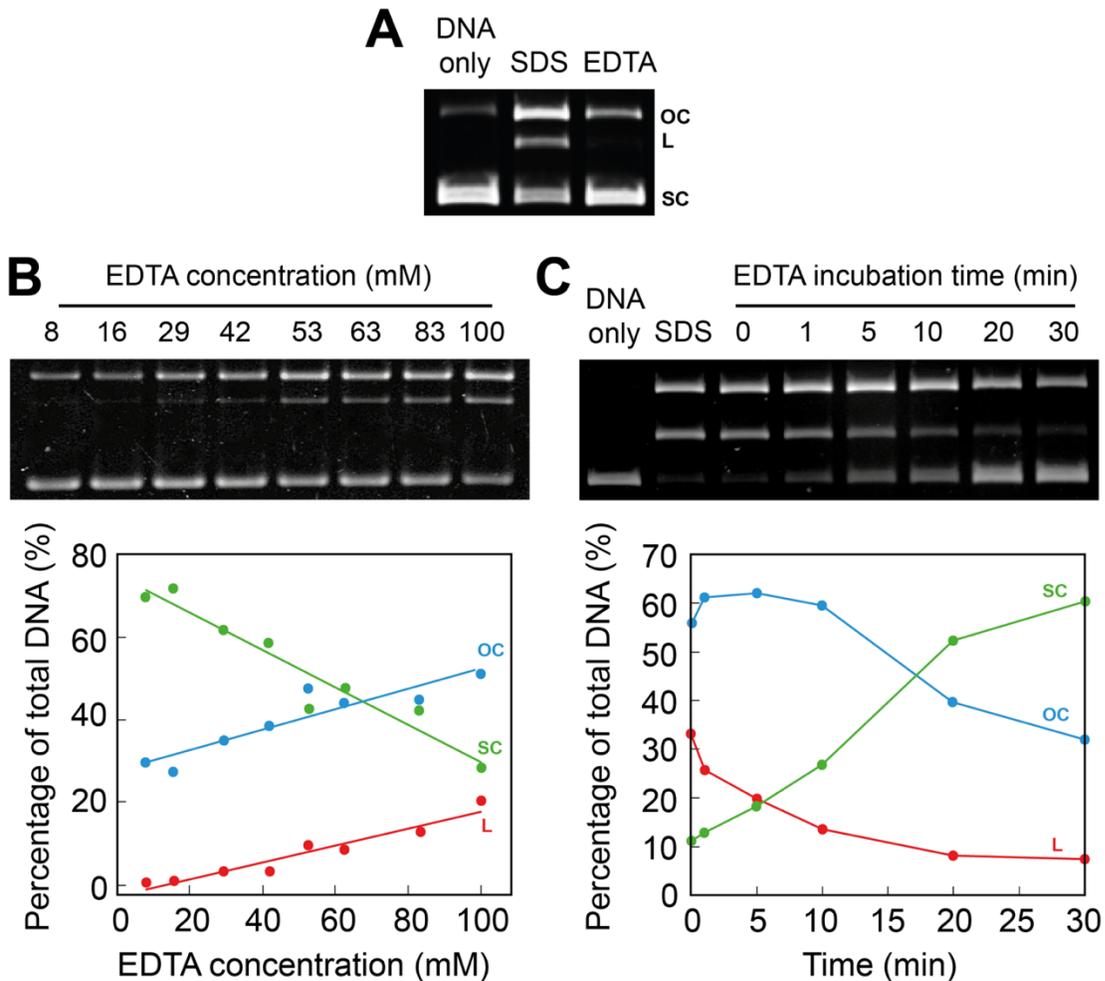
Figure 3.22: EDTA-dependent reversal of MmT6 cleavage complexes. A: Preliminary assay in which 136 nM MmT6 was incubated with 4 nM negatively supercoiled pBR322*, 1 mM ADPNP and 30 mM MgCl$_2$, for 30 min at 37°C. For the 'SDS' sample, 3 μL of 2% SDS were added to the reaction mixture, whilst for the 'EDTA' sample, 40 mM EDTA was added and incubated at room temperature for 5 min before the addition of the SDS. This was followed by the addition of proteinase K (see Chapter 2 for method details). The DNA-only sample was incubated in the absence of MmT6. **B:** Effect of EDTA concentration on MmT6 cleavage complex reversal. A cleavage assay was performed in which 80 nM MmT6 was incubated with 4 nM negatively supercoiled pBR322*, 10 mM MgCl$_2$ and 1 mM ADPNP, were incubated for 30 min at 37°C. This was followed by the addition of EDTA (8-100 mM) and a 5-min incubation at room temperature. **C:** The effect of incubation time in the presence of EDTA on MmT6 cleavage complex reversal. MmT6 (80 nM) was incubated with 4 nM negatively supercoiled pBR322*, 1 mM ADPNP and 10 mM MgCl$_2$, for 30 min at 37°C. This was followed by the addition of 30 mM EDTA, and further incubation at room temperature (0-30 min). For all samples presented in this figure, 3 μL 2% SDS and 3 μL proteinase K (New England Biolabs) were added post EDTA incubation (see Chapter 2 for cleavage assay protocol). All samples were run on a 1% (w/v) native agarose gel for 8-12 hr at ~2 Vcm^{-1}, stained with 0.5 μg/mL ethidium bromide and imaged under UV illumination. Data used in construction of the graphs was extracted from the gels using ImageJ and analysed in Excel. OC: open circular or nicked (bue), L: linear (red), and SC: supercoiled (green).

As well as being sensitive to EDTA concentration, the extent of MmT6 DNA cleavage reversal was also dependent on the incubation time, with the most reversal occurring after a 30-min incubation in the presence of 30 mM EDTA (Figure 3.22C). In general, both the nicks and double-stranded breaks are resealed simultaneously, however, early on there appeared to be a low-level increase in the amount of nicked DNA, which may indicate that the double-stranded DNA breaks are sealed using a nicked intermediate. This was also found for the reversal of DNA gyrase cleavage complexes, formed in the presence of different cleavage state stabilising drugs [262]. However, in the case of MmT6, more work on this topic is required to gain a deeper understanding of cleavage reversal (re-ligation) kinetics.

3.2.6 MmT6 DNA binding

Exploring the DNA binding behaviour of topos in the absence of catalytic activity, can inform on the optimal DNA substrate for binding, which can reveal further information on the mechanism and role of topos. For MmT6, work had been published prior to the research reported here, using fluorescently labelled DNA duplexes [62]. This work demonstrated that MmT6 had a preference for binding DNA duplexes longer than 40 bp, indicating a more extensive G-segment interaction than originally thought, and that the binding of supercoiled DNA had a ~60-fold preference over sheared salmon sperm DNA, indicating a preference for binding to DNA crossings and/or bends [62]. Along with other supporting data, these results led to the hypothesis that MmT6 preferentially binds and distributively relaxes single DNA crossings. The work reported here began prior to publication of the aforementioned study [62], however it supports and elaborates upon the DNA binding behaviour of MmT6 as a true DNA crossing sensor, using alternative and complementary methods.

All assays employed here involved the use of a filter-binding assay that captured MmT6/DNA complexes using nitrocellulose membranes, an established technique that has been utilised in various ways to research DNA binding and cleavage complex formation by various topos [156, 225, 258, 271-277]. For type II topos, catalytic activity in most cases (aside

from DNA gyrase) is dependent on the presence of ATP, whilst DNA binding only requires Mg^{2+}, therefore allowing these two activities to be assayed separately using the presence or absence of ATP, respectively. By way of controls, nitrocellulose, in the form of 0.45 µm nitrocellulose Centrex MF filters (Whatman), was found to bind MmT6-bound DNA complexes whilst unbound DNA was eluted from the column, allowing the enrichment of MmT6/DNA complexes for further analysis. An initial assay involved incubating 16 nM MmT6 with different 10 nM pBR322* topologies simultaneously and electrophoretically comparing the MmT6-bound DNA to that which was unbound (Figure 3.23). The pBR322* mixture contained negatively-supercoiled, relaxed and linear plasmid, with MmT6 visually displaying a binding preference for supercoiled and nicked DNA, with linear pBR322* and the most relaxed topoisomers being bound less strongly. This is in agreement with the previous study which found MmT6 bound supercoiled DNA more strongly than linear [62], and tentatively indicated that MmT6 DNA binding increased with increasing superhelical density of topoisomers, a concept that will be elaborated upon below. Through the course of this work, particularly while exploring DNA cleavage by MmT6, a tendency to induce nicking in the presence of 1 mM ADPNP was found (Section 3.2.5). The result in Figure 3.23 indicated that MmT6 also displayed a preference for binding nicked DNA, as well as inducing it, the significance of which is unclear. Topo IV and topo II have been shown to bind and stabilise 4-base overhangs [56, 229], demonstrating a preference for binding to a transition-state species, a behaviour MmT6 may also share. It is perhaps not surprising that a type II topo would tightly bind to exposed DNA overhangs or nicks as their canonical activity relies on the transient cleavage and re-ligation of the duplex, mitigating the risk of releasing the DNA break as much as possible.
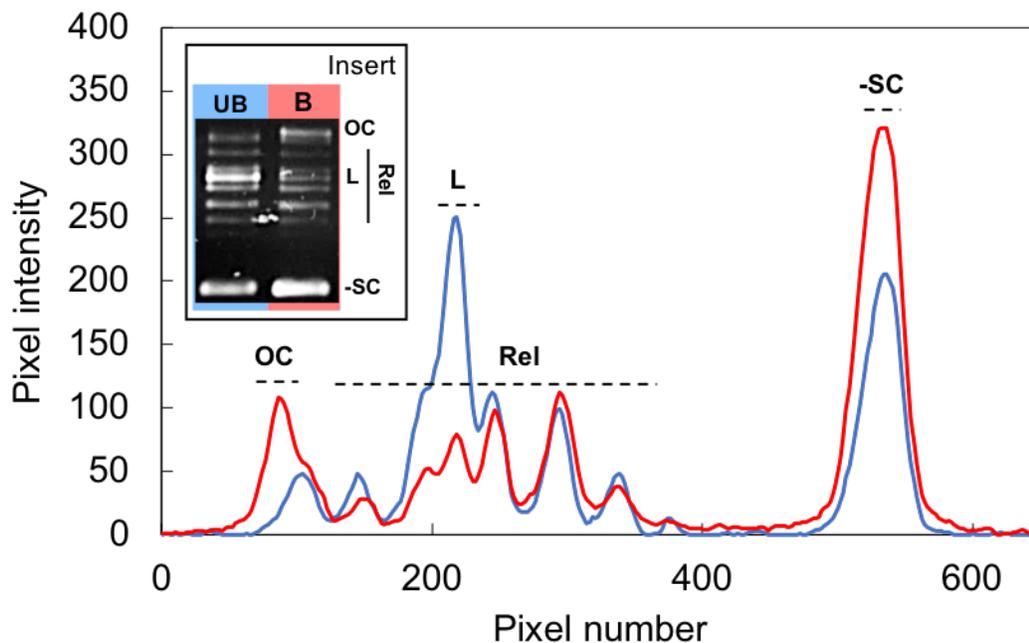
Figure 3.23: MmT6 DNA-binding activity using a mixture of linear, relaxed and negatively supercoiled pBR322*. The insert shows the gel analysed (using ImageJ) to form the main figure. MmT6 (16 nM) was incubated with a 10 nM mixture of pBR322* (linear, relaxed and negatively supercoiled DNA) and 10 mM $MgCl_2$ for 30 min at 37°C. The unbound DNA (UB, blue) was separated from the MmT6-bound DNA (B, red) using 0.45µm nitrocellulose Centrex MF filters (Whatman) (see Chapter 2 for DNA binding assay methods). DNA samples were run on a 1% (w/v) native agarose gel for 12 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination. Data used in construction of the graph was extracted from the gels using ImageJ and analysed in Excel. OC: open circular or nicked, L: linear, Rel: relaxed and SC: supercoiled (green).

In order to attain a clearer picture of the binding affinity of MmT6 to different pBR322* topologies, another nitrocellulose-capture technique was employed in which 16 nM MmT6 was incubated with 10 nM negatively-supercoiled, positively-supercoiled, relaxed and linear pBR322* separately. The MmT6-bound DNA was captured, proteolyzed and purified prior to concentration determination. In Figure 3.24, the percentage of the total DNA in solution which was bound by MmT6 is shown. As expected, linear pBR322* was bound the least, with negatively-supercoiled pBR322* being bound preferentially. However, unexpectedly positively-supercoiled and relaxed pBR322* were bound with equal affinity, despite MmT6 displaying a faster rate of relaxation on positively-supercoiled DNA over negatively-supercoiled (Section 3.2.3, Figure 3.8). A simple explanation for the increased rate during positive supercoil

relaxation would have been an increased MmT6 binding affinity for positive writhe, however this does not seem to be the case. The binding result was supported by exploring DNA cleavage by MmT6 on both positive and negative writhe, with MmT6 inducing more DSBs on negatively-supercoiled DNA than positively-supercoiled (Section 3.2.5, Figure 3.16), a behaviour which can be described by differences in binding affinity. However, a conundrum arose, as why would MmT6 display a higher rate of relaxation on positive writhe, whilst supporting more binding and cleavage on negative writhe? Relaxation assays, by nature, indicate that during MmT6's interaction with the DNA substrate, both the G-segment and T-segment were present. For binding and cleavage activity, the presence of the T-segment is necessarily not compulsory and it may be that MmT6 binds and cleaves a G-segment alone more frequently when the DNA is underwound. This behaviour would initially suggest activity should be higher on negatively supercoiled DNA, but does not necessitate this outcome as for relaxation to occur both the G- and T-segments must be present. This suggested, instead, that the acceleration of a reaction stage post DNA binding and cleavage, was responsible for the increased rate during positive supercoil relaxation, and as explored in Chapter 4, was probably a consequence of the DNA crossing geometry facilitating strand passage. A similar behaviour has been documented for human topo IIα, demonstrating that whilst topo IIα relaxed positive writhe 10-fold more rapidly than negative, it maintained higher levels of binding and cleavage on negatively-supercoiled DNA [156]. The preference for positive supercoils seemed to rely on the elements within the topo IIα C-terminal domains, however CTD removal did not change cleavage behaviour, indicating that the CTDs promote activity on positive writhe through DNA geometry preference for strand passage, rather than increasing cleavage [157]. Human topo IIβ also maintains lower levels of cleavage on positively-supercoiled DNA whilst relaxing positive and negative writhe at comparable rates [157, 237], furthering suggesting that cleavage rates, as well as DNA binding, are not enough to inform on overall topo activity when considered alone.
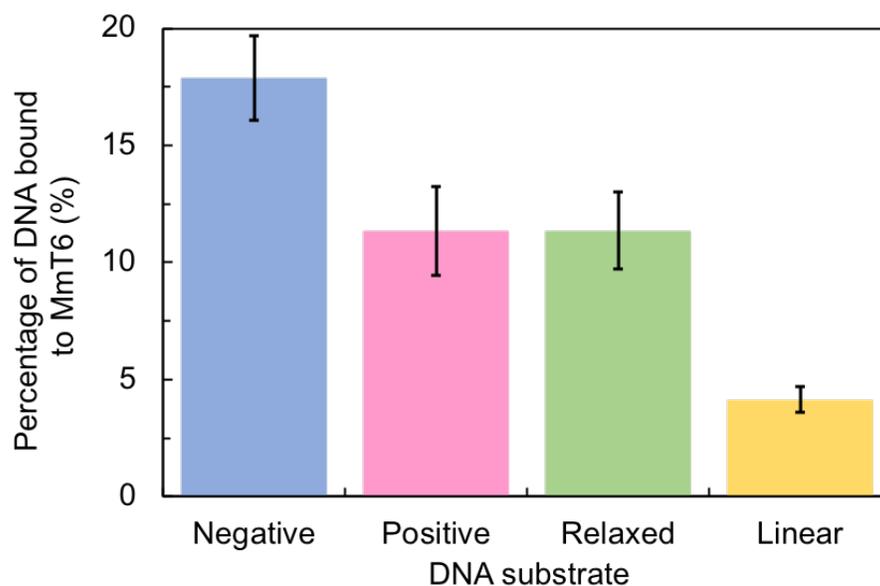
Figure 3.24: DNA binding activity of MmT6 on negatively supercoiled, positively supercoiled, relaxed or linear pBR322*. MmT6 (16nM) was incubated with a 10 nM mixture of pBR322* (linear, relaxed and negatively supercoiled DNA) and 10 mM MgCl₂ for 30 min at 37°C. The unbound DNA was separated from the MmT6-bound DNA using 0.45µm nitrocellulose Centrex MF filters (Whatman). The MmT6-bound DNA concentration was measured using the Nanodrop™, enabling calculation of the percentage of pBR322* that was bound to MmT6, ± SE (see Chapter 2 for DNA binding assay methods). For each DNA substrate the assay was repeated three times.

Whilst measuring the percentage of DNA bound to MmT6 for the different pBR322* topologies indicated a preference for binding more tightly to supercoiled DNA, this form of analysis can be taken deeper, employing a topoisomer-binding assay first described in [225]. This assay allowed the calculation of the relative dissociation constant (K_d) for MmT6 on different topoisomers of pBR322*. The K_d is used to describe how readily a molecular complex will reversibly separate into its constituents, and is therefore a measure of binding affinity. The K_d is calculated using equation 3.1, where E is the concentration of unbound MmT6, D_x is the concentration of unbound 'x' topoisomer and ED_x is the concentration of the 'x' topoisomer bound to MmT6,

$$K_{d(x)} = \frac{[E][D_x]}{[ED_x]}$$

Equation 3.1: Calculating the K_d for topoisomer 'x'. _E_ is the free enzyme concentration, _D_ is the concentration of unbound DNA and _ED_ is the concentration of MmT6-bound DNA.

Chapter 3: Ensemble Results

It is then possible to determine the ratio of K_d values for two different topoisomers equilibrated in solution, and therefore provide a measure of the relative affinity of MmT6 to one over the other, by dividing the K_d equations for both the 'x' and the 'y' topoisomer

$$\frac{K_{d(x)}}{K_{d(y)}} = \frac{[E][D_x]}{[ED_x]} * \frac{[ED_y]}{[E][D_y]}$$

Equation 3.2: Calculating the K_d for topoisomer 'x' relative to topoisomer 'y'. _E_ is the free enzyme concentration, _D_ is the concentration of unbound DNA and _ED_ is the concentration of MmT6-bound DNA.

At equilibrium, the relative concentration of free enzyme is equal, regardless of the topoisomer in question, so the variable is dropped from the equation 3.2, leading to equation 3.3.

$$\frac{K_{d(x)}}{K_{d(y)}} = \frac{[D_x][ED_y]}{[D_y][ED_x]}$$

Equation 3.3: Simplified version of equation 3.2, calculating the K_d for topoisomer 'x' relative to topoisomer 'y'. _D_ is the concentration of unbound DNA and _ED_ is the concentration of MmT6-bound DNA.

This demonstrates that the relative K_d value for each topoisomer can be measured using the ratio of concentrations of unbound DNA topoisomers ($[D_x]/[D_y]$) and MmT6-bound DNA topoisomers ($[ED_y]/[ED_x]$), producing a value that described the affinity of MmT6 for topoisomers of varying superhelical density. The lower the relative K_d, the higher the affinity of MmT6 for that particular topoisomer. The relative concentrations of each topoisomer in the unbound and MmT6-bound fractions was determined using densitometric analysis of agarose gel images.

Commercially available pBR322* (obtained from Inspiralis) can be purchased as relaxed, negatively supercoiled and positively supercoiled, however none of these options provided extended distributions of different topoisomers on which to measure MmT6 binding affinity. Therefore, _Archeoglobus fugidus_ reverse gyrase was employed to convert highly negatively supercoiled pBR322* into extended distributions of both positive and negative topoisomers (Figure 3.25). Using a time course it was determined that incubating 11 nM

negatively supercoiled pBR322* with 100 nM reverse gyrase at 95°C generated an extend distribution of negatively and positively supercoiled topoisomers after 10 s and 20 s, respectively (Figure 3.25A). This was then coupled to a DNA purification protocol (see Chapter 2), which generated pure DNA samples of the desired topology and purity (Figure 3.25B). To confirm the DNA topology, the samples were run in the presence of 1.5 μg/mL chloroquine (CQ). The DNA intercalation of CQ has the effect of making negative topoisomers run more relaxed and positive topoisomers to become more positive, which can be seen in Figure 3.25C, confirming the expected DNA topology of both topoisomer distributions.
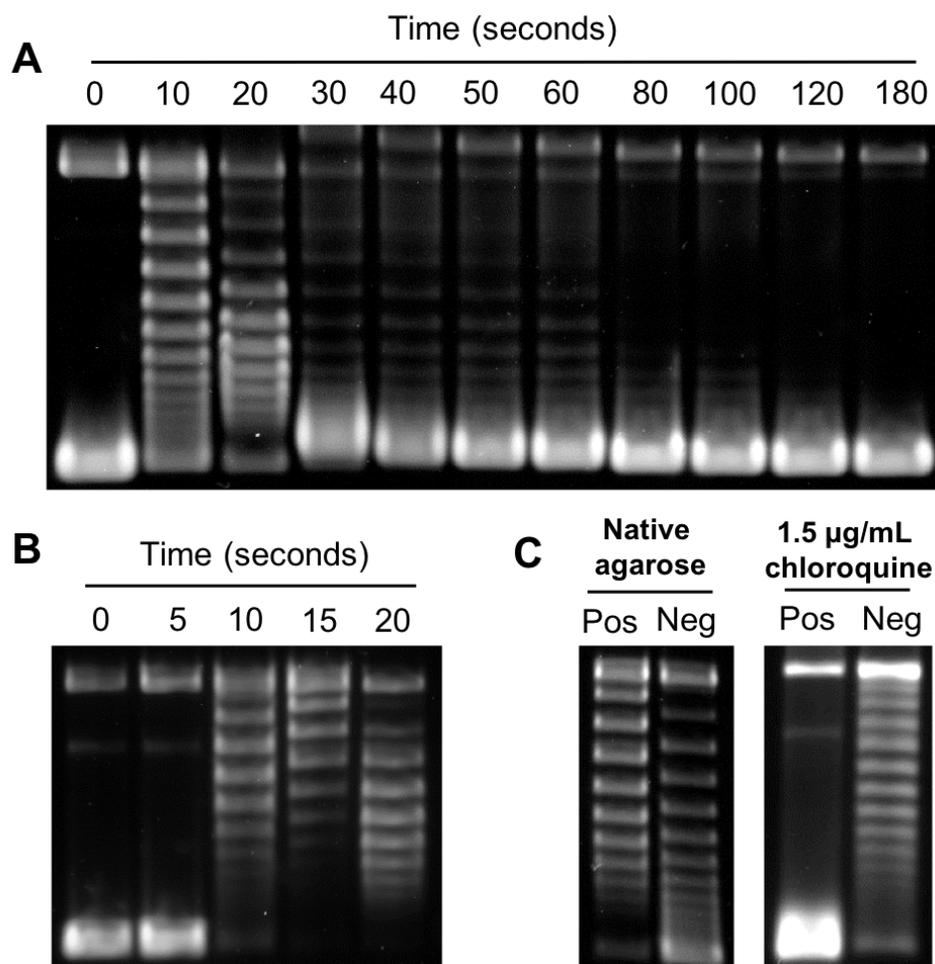
Figure 3.25: Utilising *Archeoglobus fugidus* reverse gyrase to generate extended positive and negative topoisomer distributions of pBR322*. A: Time course of *A. fulgidus* reverse gyrase activity on negatively supercoiled pBR322*. Reverse gyrase (100 nM) was incubated with 11 nM negatively supercoiled pBR322* and 1 mM ATP, at 95°C, to determine the appropriate incubation time for the generation of negative and positive topoisomer distributions (10 and 20 s, respectively). B: Time course of *A. fulgidus* reverse gyrase activity on negatively-supercoiled pBR322*, as in A, but followed by DNA purification prior to gel electrophoresis. C: The topology

of purified pBR322* from B was confirmed using agarose gel electrophoresis in the presence of 1.5 µg/mL chloroquine. All samples were run on a 1% (w/v) agarose gel for 13 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination. The gel containing chloroquine (see Figure 3.25C) was soaked in 1x TAE for ~1 hr, prior to ethidium bromide staining. Pos: positively supercoiled pBR322*, and Neg: negatively-supercoiled pBR322*.

A limitation of producing DNA topology distributions in this manner was that there was no control over the amount of each topoisomer produced. For instance, in the negatively supercoiled sample there may be more of the topoisomers with lower superhelical density than high, with the reverse being true for the positively supercoiled sample. This may hinder direct comparison between MmT6 binding on positive and negative writhe, however, it does not affect drawing conclusions within the data sets for each chirality.

Once pBR322* samples of the desired topology were attained, 5 nM pBR322* was incubated with 8 nM MmT6 and 10 mM MgCl$_2$ before separation of the MmT6-bound DNA from the unbound DNA using nitrocellulose membrane capture. It is worth noting here, that for all topo-binding assays of this nature, it is advisable to keep the molar concentration of topo below that of the DNA in order to prevent multiple topo molecules binding to a single plasmid, a situation that the membrane-capture technique cannot distinguish. In all cases presented here, the MmT6 concentration is moderately above the DNA, and this is due to MmT6 DNA binding being too weak to bind sufficient amounts of DNA for accurate quantification post membrane capture, particularly for binding to linear DNA. Data presented in Figure 3.24, in which the MmT6:DNA ratio was 1.6:1 demonstrated that whilst MmT6 was in excess only ~17% of the negatively supercoiled DNA was bound, with less than 5% linear being bound. This suggested that, while some plasmids may be bound by more than one MmT6 molecule, it likely wasn't happening frequently enough to dramatically affect the results, particularly as no evidence of cooperativity in MmT6 binding activity has been documented.

Once the MmT6-bound and -unbound DNA samples were purified they were run simultaneously on a 1% native agarose gel for 13 hr at ~2 Vcm^{-1} and stained with 0.5 µg/mL ethidium bromide. As previously described, the binding affinity of MmT6 to particular

topoisomers was calculated by measuring the band intensity on images of agarose gels (see Figure 3.26A for an example of a gel). The ratio of unbound DNA to MmT6-bound DNA concentrations allowed determination of the relative K_d and provided a measure of MmT6 affinity for each topoisomer. As seen in Figure 3.26B, the K_d decreased for increasing superhelical density, with more supercoiled topoisomers of both positive and negative chirality being bound more tightly by MmT6. This further indicated that MmT6 recognised and bound DNA crossings preferentially, as highly supercoiled topoisomers have more opportunities for MmT6 binding, than less supercoiled ones, and will therefore be bound more frequently. An interesting comparison is topo IV DNA binding behaviour assayed using the same technique [225]. Whilst this work didn't incorporate the use of positively-supercoiled DNA as was done here, topo IV was shown to exhibit increased binding with increased negative supercoiling, also demonstrating that topo IV had a preference for supercoiled DNA. However, this was hypothesised to be due to its recognition or induction of DNA bending, which occurs far more frequently in DNA as the superhelical density is increased. Indeed, using the single-molecule magnetic tweezers, topo IV exhibited only a single binding position on a plectoneme of multiple DNA crossings, which generated the hypothesis that it was bound to the distal end of the plectoneme [120], while MmT6 activity indicated multiple plectoneme binding sites (see Chapter 4 for MmT6-related magnetic tweezers assays). This difference between MmT6 and topo IV may be evident in the relationship between the binding affinity and superhelical density. For topo IV this relationship was linear whilst for MmT6 binding, the data is not well described using a linear equation. However, even though the data is not presented with fits, negative topoisomer binding data (Figure 3.26B, blue circles) can be well described using an inverse relationship, similar to that found using the magnetic tweezers (Chapter 4), exploring the relaxation rate as a function of superhelical density. This provided further evidence that MmT6 preferentially binds DNA crossings. The data for positive supercoil binding has demonstrated once again that MmT6 binding to positive supercoils is not equivalent in nature to negative supercoil binding. However, in both the positive and negative data sets, a large portion of the topoisomers could not be resolved leading to a loss of information about binding to more highly

supercoiled molecules. This precluded the ability to accurately fit the data, an issue which was more acute for positive topoisomer binding. However, the data indicated that MmT6 binding to negative topoisomers with lower superhelical density is tighter than binding the equivalent positive topoisomers. Although, as previously described, the inability to control the concentration of each topoisomer between the different chiralities may prevent accurate comparison between the data sets. Nevertheless, this result was supported by the data presented in Figure 3.24, in which MmT6 demonstrated more robust binding to negatively supercoiled DNA. In future work, optimisation of the use of chloroquine and netropsin in resolving more negative and positive topoisomers, respectively, could inform further on the MmT6 DNA binding behaviour as a function of superhelical density and chirality. This would also help circumvent issues associated with assigning linking numbers to topoisomers that are very highly supercoiled. Here the most highly-supercoiled band was described as having 25 supercoils (based on a specific linking difference of ~0.06 [225]), however in reality it was a collection of different topoisomers which couldn't be resolved by conventional agarose-gel electrophoresis equipment. An alternative would be to employ capillary gel electrophoreisis, which has been demonstrated to accurately resolve highly supercoiled topoisomers [278]. Another possibility would be to explore binding in the presence of single topoisomers.
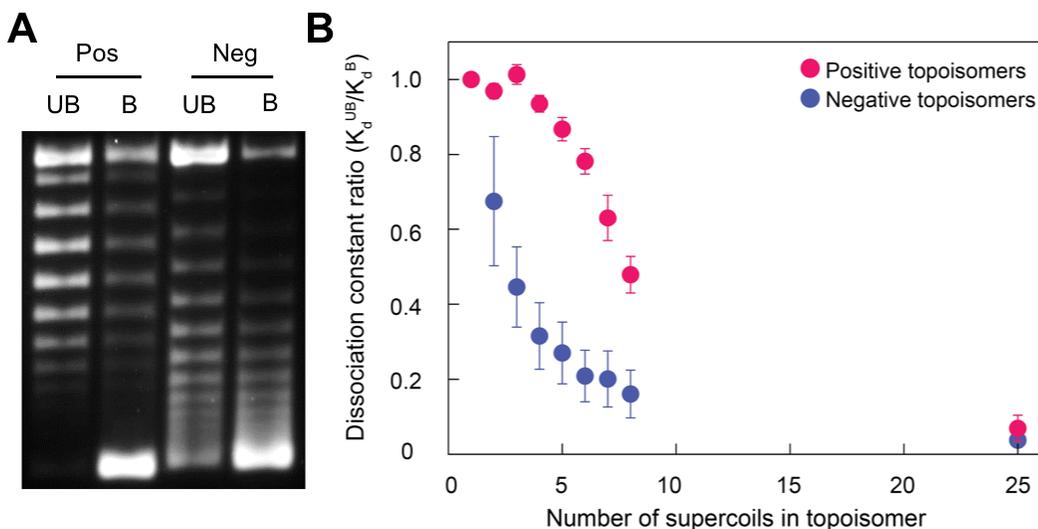
Figure 3.26: MmT6 DNA binding affinity as a function of the number of supercoils in pBR322*. A: MmT6 (8nM) was incubated with 5 nM pBR322*, either positively (pos) or negatively (neg) supercoiled, and 10 mM MgCl$_2$, for 30 min at 37°C. The MmT6-bound DNA (B) was separated from the unbound DNA (UB) using 0.45µm nitrocellulose Centrex MF filters (Whatman). The bound and

unbound DNA samples were run on a 1% (w/v) agarose gel for 13 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination. **B:** Calculation of the dissociation constant ratio (K_d^{UB}/K_d^B) for MmT6-dependent binding to positively (pink) and negatively (blue) supercoiled topoisomers, plot as a function of the number of supercoils within each topoisomer. Data was extracted from the gel in Figure 3.25A using ImageJ and analysed in IgorPro7 (Wavemetrics). Method was first described in [225]. Data is derived from the average of three separate assays.

3.2.7 MmT6 ATPase activity

All type II topos hydrolyse ATP during DNA topology manipulation, with the only exception being the ATP-independent relaxation of negative supercoils by DNA gyrase [11]. It has therefore been of great interest to characterise how these enzymes hydrolyse ATP, under which circumstances it is stimulated, and, whilst it remains a critical point of discussion, why ATP is necessary in those cases where the reaction is thermodynamically favourable [279, 280]. For instance, MmT6 relaxes DNA to thermodynamic equilibrium just like the ATP-independent activity of topo IA [231], yet ATP is strictly required for MmT6 activity. A number of structural and biochemical studies have been performed using topo VI enzymes (including MmT6) in which the aim was to dissect ATP turnover and explore the DNA-dependence on ATPase activity of both wild type and mutant topo VI enzymes [62, 64, 67, 195].

The first crystal structures of the Top6B subunit were achieved using a truncated form of *S. shibatae* topo VI (SsT6), both with and without ADPNP [64]. These structures revealed that in the presence of nucleotide, the Top6B subunits dimerised, much like that which has been demonstrated for type IIA topos and MutL, all of which harbour the ATP-binding GHKL domain [91, 281-284]. Along with the fact that ADPNP is sufficient to cause DNA cleavage, these data indicate that ATP hydrolysis is not important for DNA binding, N-gate dimerisation or G-segment cleavage in topo VI. Another study delved further still by determining the structure of truncated SsT6 Top6B in complex with various ATP analogues, representing various stages of the Top6B ATP-turnover cycle [67]. This work demonstrated that the major structural changes that occur within Top6B happen upon ATP biding and P$_i$ release. Nucleotide binding alone drove Top6B subunit dimerisation, which appeared to be tightly coupled to DNA

cleavage and suggested a major role of ATP for topo VI was in formation of a stable Top6B subunit interface, which caused conformational changes in Top6A and led to DNA cleavage. Tightly coupling DNA cleavage to the dimerisation of the Top6B subunits would protect against the formation of permanent double strand breaks [68], and as topo VI enzymes have only two gates, rather than three like type IIA, this coupling may be even more important. The ATP-bound conformation, known as the restrained state (achieved using ADPNP and ADP·AIF^{4-}), sequesters the bound nucleotide from solvent and positions critical amino acids for ATP hydrolysis [67]. Top6B in complex with ADP and P_i retained the restrained conformation, indicating hydrolysis does not directly cause structural changes within the enzyme [67]. Other experiments had already found that ADP/P_i release was rate limiting for topo II and tightly coupled to strand passage, suggesting that a conformational change was required for said release [285]. The Top6B structural work corroborated this by demonstrating that Top6B only returned to a relaxed state in the absence of the scissile P_i. This study then speculated that the T-segment itself, is responsible for causing conformational changes which lead to its passage, as well as facilitating P_i release by re-exposing the ATP binding site to solvent. Indeed, it was recently demonstrated that MmT6 ATP hydrolysis was dependent on regions of the Top6B subunit which are extremely likely to interact with the T-segment [62].

With this previous data in mind, the work reported in the remainder of this section details further exploration of the ATP-hydrolysis activity of MmT6 as a function of DNA topology, using two alternative methods: the pyruvate kinase/lactate dehydrogenase-coupled (PK/LDH) assay (section 3.2.7.1) and the [γ-^{32}]-ATP assay (3.2.7.2).

3.2.7.1 PK/LDH ATPase assay

The PK/LDH assay is a spectroscopic technique which couples the hydrolysis of ATP to the oxidation of reduced nicotinamide adenine dinucleotide (NADH), the concentration of which can be tracked as a function of absorbance at 340 nm (A_{340}) (NADH absorbs strongly at 340nm, whilst NAD$^+$ does not). This is achieved by conducting the topo-dependent ATPase reaction in the presence of pyruvate kinase (PK), lactate dehydrogenase (LDH) and their respective

substrates; phosphoenolpyruvate (PEP) and NADH. As seen in Figure 3.27, ATP hydrolysis, such as by a type II topo, generates ADP which is a substrate for PK. PK uses the ADP to convert PEP into pyruvate while also resynthesising the ATP. The pyruvate generated by PK is the substrate for LDH, which converts it to lactate by oxidising NADH, therefore the ATP hydrolysis by the enzyme in question is coupled to the oxidation of NADH and a reduction in A_{340} (Figure 3.27). The Beer-Lambert law (Equation 3.4) can then be used to convert the change in A_{340} to a concentration value with the loss of NADH equimolar to the hydrolysis of ATP, where A is the absorbance, ε is the molar extinction coefficient, c is the molar concentration and ℓ is the path length.

$$A = \varepsilon c \ell$$

Equation 3.4: The Beer-Lambert law. A is the absorbance, ε is the molar extinction coefficient, c is the molar concentration and ℓ is the path length.

This technique has been widely utilised in the investigation of topo ATPase activity [63, 286-294], including for MmT6 [62], in both the mechanistic characterisation of topo ATPase activity and the exploration of drug-based topo inhibition. Benefits of using this approach include the continuous nature of the measurement and the ATP re-synthesis, meaning that changes to the ATP:ADP ratio needn't affect enzyme activity as ADP is continually re-phosphorylated. However, for the work reported in this section, the PK/LDH assay did not prove efficient for MmT6 ATPase activity in terms of substrate usage for reasons which will be explored below.
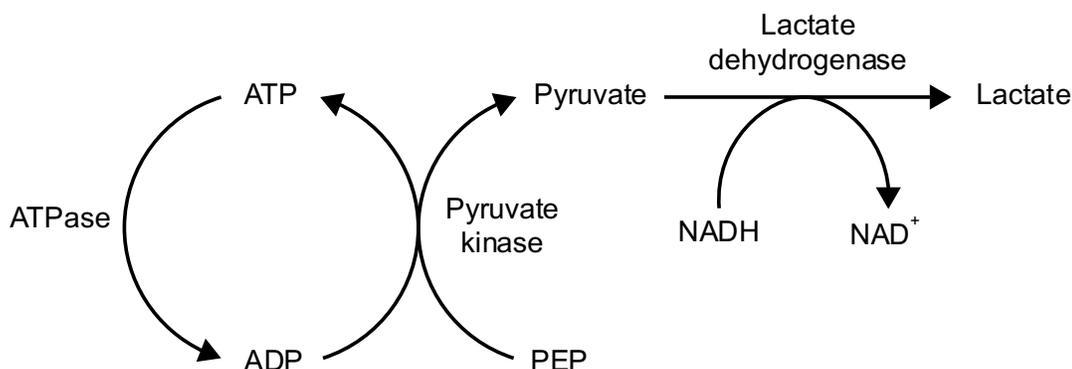
Chapter 3: Ensemble Results

Figure 3.27: Biochemistry of the PK/LDH ATPase assay. ATPase activity (here from the activity of MmT6) converted ATP to ADP + P$_i$. The ADP created was used a substrate for pyruvate kinase (PK), which converted phosphoenolpyruvate (PEP) to pyruvate and regenerated the ATP. Pyruvate was then converted to lactate by lactate dehydrogenase (LDH), which involves the conversion of NADH to NAD$^+$. Therefore, the ATP hydrolysis reaction by the ATPase was linked to the loss of NADH in solution, followed by measuring the decrease in absorbance at 340 nm (A$_{340}$).

To begin, a control was conducted using 200 nM *E. coli* DNA gyrase in the presence of 21 nM relaxed pBR322* (Figure 3.28). For negative supercoiling, the specific activity of gyrase ATP hydrolysis is known to be ~1 s^{-1} [295, 296]. Knowing this meant that the PK/LDH system could be tested using an enzyme for which a large body of data already existed. With reference to Figure 3.28, four samples were prepared; without NADH (green line) to attain a measure of the back ground A$_{340}$; with ADP added after 10 min (blue line) to make sure the PK and LDH enzymes functioned as expected; gyrase and relaxed pBR322* with ATP added after 10 min (yellow line); and the last contained gyrase and relaxed pBR322* with ATP added after 10 min and novobiocin (novo) added after 45 min (red line). Novo is a potent ATPase inhibitor of gyrase, which functions via competitively binding the GHKL domain, therefore preventing ATP binding [100, 297]. The addition of 2 mM ADP in the control sample (Figure 3.28, blue line) demonstrated the system functioned correctly as the A$_{340}$ dropped dramatically to background levels, congruent with the loss of NADH, which is used at 400 µM. The samples containing gyrase and relaxed pBR322* (red and yellow lines) were allowed to equilibrate for 10 min prior to the addition of 2 mM ATP. Once ATP was added the A$_{340}$ declined linearly as ATP hydrolysis by gyrase caused NADH oxidation. To confirm that this loss of NADH was dependent on gyrase ATP hydrolysis activity, 10 µM novo was added to one sample (red line) at 45 min, and caused the A$_{340}$ decline to cease. This confirmed that the change in the A$_{340}$ was indeed caused by gyrase ATP hydrolysis activity (Figure 3.28). The gradient of the slope from the gyrase, pBR322* and ATP containing sample (yellow line) allowed the calculation of gyrase ATPase specific activity, which for this assay was 0.2 s^{-1}. Even though this was a fifth of what was expected (~1 s^{-1}), there are a number of explanations for this. The gyrase concentration was

difficult to accurately assess and as this value was used in the calculation of the specific activity, this may be a source of significant error. The activity of gyrase also changes between purification batches and it is possible the preparation used here was not as active as those used in previous research. However, as seen in Figure 3.28B, DNA gyrase behaved as expected, with samples from both the gyrase-containing samples being run a 1% (w/v) agarose gel to assess topology. The sample which contained gyrase and ATP successfully supercoiled the relaxed pBR322*, whereas the sample containing novo is more relaxed, consistent with the ATP-independent relaxation of negative supercoils by gyrase. In the latter sample, the DNA began relaxed and prior to novo addition, gyrase would have significantly supercoiled it. However, during the remaining 65 min, novo would have prevented ATP binding, so gyrase instead switched to relaxation, which does not require ATP [11].
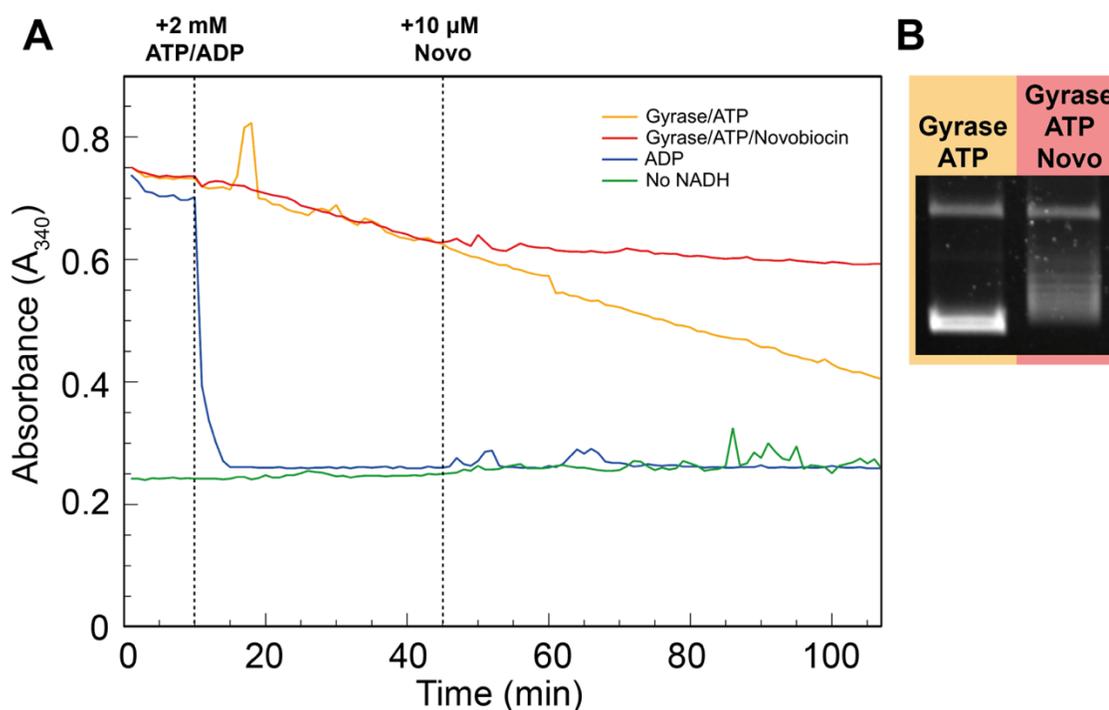
Figure 3.28: Using the PK/LDH ATPase assay to explore ATP hydrolysis activity of _E. coli_ DNA gyrase. A: A_{340}, in the presence of 200 nM gyrase, 2 mM ATP and 21 nM relaxed pBR322*, decreased over time due to the ATP hydrolysis activity of gyrase, leading to a reduction in NADH concentration. An aliquot of 2 mM ATP was added after 10 min to allow the system to equilibrate. At 45 min 10 μM novobiocin (novo) was added to one of the two samples containing 200 nM gyrase, 2 mM ATP and 21 nM relaxed pBR322* (red line), and not to the other (yellow line). As controls, a sample without NADH (green line), to show the baseline absorbance, and a sample in which 2 mM ADP, instead of 2 mM ATP, was added at 10 min to establish the PK and LDH enzymes were functioning as expected (blue line). The gradient of the yellow line was used to calculate

the specific activity of gyrase. B: The samples containing 200 nM gyrase, 2 mM ATP and 21 nM relaxed pBR322*, with and without 2 µM novo (red and yellow, respectively), were ran on a 1% (w/v) agarose gel for 10 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination.

Once the gyrase control confirmed that the PK/LDH system functioned as expected, attention was turned to using MmT6 and negatively supercoiled pBR322* in order to explore MmT6 ATP hydrolysis activity. Initial assays, using similar concentrations of enzyme (200 nM) and DNA substrate (21 nM), as used for the gyrase, failed to yield any results with no decrease in A_{340} detected. A low-level ATPase signal for MmT6 was finally achieved using 750 nM MmT6 in the presence of 93 nM negatively supercoiled pBR322*, a DNA:enzyme ratio of 1:8 (Figure 3.29A). The slope of the line is shallow and also short lived, indicating that MmT6 not only hydrolyses ATP more slowly than type IIA topos, but also appeared to either be tightly coupled to the presence of DNA writhe or the enzyme did not remain active in solution for a considerable length of time. The former explanation is supported by data which will be presented in the following section (Section 3.2.7.2). Whilst it was unclear whether the slope was caused by MmT6 activity, an ATP hydrolysis rate was calculated regardless and found to be 0.029 s^{-1} or 1.7 min^{-1}. Recently-published data exploring MmT6 ATPase activity utilised the PK/LDH assay, used a DNA:Enzyme ratio of 1:4 (500 nM MmT6 and 136 nM supercoiled plasmid), reporting the specific activity of MmT6 ATP hydrolysis to be ~2.3 min^{-1} [62]. This added some validity to the specific activity value calculated from the data in Figure 3.29A. A more robust MmT6 ATPase signal was also measured here using this same ratio, incubating 181 nM negatively supercoiled pBR322* with 750 nM MmT6 (Figure 3.29B). This produced a more significant decrease in A_{340}, which also demonstrated a relatively shallow gradient and only proceeded for 30 min before the A_{340} remained relatively constant (only showing the background level of reduction in A_{340} caused by the oxidation of NADH in solution over time). The specific activity calculated using this data was 0.032 s^{-1} or 1.92 min^{-1}, which is not only highly similar to the value calculated from the data in Figure 3.29A, but also to the value reported in the previously-published work [62]. Together, this suggested the activity was indeed MmT6-dependent, and

by taking a sample every 5 min post the addition of ATP, the change in DNA topology could be tracked during the course of the reaction (Figure 3.29C). This preliminary gel-based data further indicated that MmT6 ATP hydrolysis activity is highly coupled to the presence of it substrate, and potentially indicated that once the DNA was fully relaxed, background MmT6-dependent ATP hydrolysis did not occur. This is in line with data previously reported, in which the basal MmT6 ATP hydrolysis activity in the absence of DNA was ~0 s^{-1}. This is not generally the case with type IIA topos, which have a far more appreciable background ATP hydrolysis activity when the optimal DNA substrate has been converted to product or is not present at all. Gyrase is a good example, maintaining around 5% ATPase activity in the absence of DNA, and even though linear DNA is neither the optimal substrate nor able to store topological alterations attempted by gyrase activity, it can fully stimulate the DNA-dependent gyrase ATPase activity [295]. During supercoiling, it is thought that as gyrase can wrap DNA and present the T-segment by doing so, linear DNA facilitates this behaviour, regardless of whether topology changes can be stored or not. MmT6 ATP hydrolysis activity was around 30-fold lower than that published for gyrase, and raised important questions relating to the *in vivo* roles of MmT6. Previous MmT6 work highlighted this, suggesting there may be an unknown cofactor or condition that could accelerate MmT6 activity *in vivo*, as the current value for ATPase activity would render MmT6 a highly inefficient DNA relaxase that would be unable to support DNA replication [62]. An alternative explanation will be provided in Chapter 4, as data suggested MmT6 was highly adapted to the preferential removal of catenanes, which in turn disfavoured supercoil relaxation. As neither the singly catenated DNA substrate used in Section 3.2.4.2, or positively supercoiled pBR322*, were available at the time of performing these ATPase assays, their affects have not been explored, so future work should encompass this.
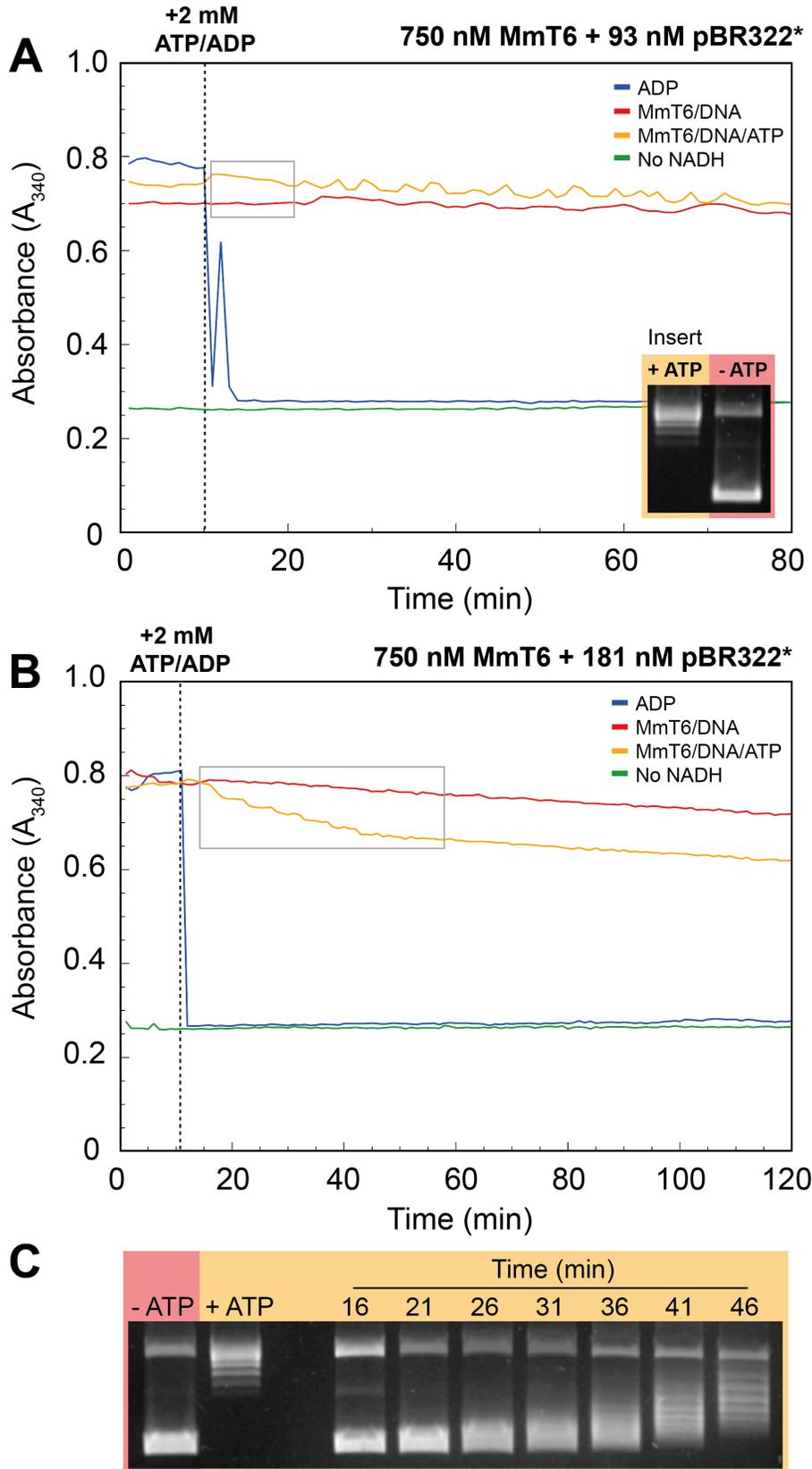
Figure 3.29: Using the PK/LDH ATPase assay to explore ATP hydrolysis activity of MmT6. A: The A$_{340}$, in the presence of 750 nM mMT6, 2 mM ATP and 93 nM negatively supercoiled pBR322*, decreased over time due to the ATP hydrolysis activity of MmT6, leading to a reduction in NADH concentration. The 2 mM ATP was added to one sample containing 750 nM MmT6 and 93 nM

negatively supercoiled pBR322* after 10 min to allow to system to equilibrate (yellow line), but not to the other (red). As controls, a sample without NADH (green line), to show the baseline absorbance, and a sample in which 2 mM ADP, instead of 2 mM ATP, was added at 10 min to establish the PK and LDH enzymes were functioning as expected (blue line). The gradient of the portion of yellow line within the grey box was used to calculate the specific activity of MmT6. The inset shows the DNA after 80 min for the samples containing MmT6 and DNA (with and without ATP), which were run on a 1% (w/v) agarose gel for 10 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination. B: Same as in A, however, 750 nM MmT6 was incubated with 181 nM of negatively supercoiled pBR322*. C: DNA samples from Figure 3.29B, were taken during the course of the reaction (every 5-6 min) once the ATP had been added to the sample containing 750 nM MmT6 and 181 nM negatively supercoiled pBR322*, as well as after 120 min for both the MmT6 samples, with and without ATP. These samples were run on a 1% (w/v) native agarose gel for 10 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination.

Using the PK/LDH assay to explore MmT6 ATP hydrolysis was technically challenging owing to tight coupling of activity to substrate concentration. This meant that in order to see an appreciable decrease in A_{340}, excessive amounts of supercoiled pBR322* were necessary in order to achieve the desired concentration in a 100 µL sample volume, which is quite typical when using 96-well plates for this form of assay. It was, of course, not impossible as both the two data sets reported here and the previously published MmT6 characterisation data were in agreement as to the final value for the specific activity of MmT6 ATPase activity, but also the DNA:enzyme ratio necessary to reliably detect the activity. However, an alternative approach was employed, described in the following section, which allowed more efficient substrate usage and higher sensitivity.

3.2.7.2 Radioactive ATPase assay

Due to the technical drawbacks just described (Section 3.2.7.1) for using the PK/LDH assay in the exploration of MmT6 ATP hydrolysis activity, a radioactive ATPase assay was employed. This assay involved the use of [γ-^{32}]-ATP and measuring the amount of [γ-^{32}]-P$_i$ released as a result of MmT6-dependent activity by analysing the samples with thin-layer chromatography (TLC). In effect, this assay was akin to a relaxation assay, with the difference being the final

product analysed, which here was the unhydrolysed [γ-^{32}]-ATP and released [γ-^{32}]-P$_i$, achieved using TLC plates and a phosphorscreen, rather than looking at DNA topology using an agarose gel. Using radioactivity meant that a smaller reaction volume could be utilised, and ATPase activity could be detected with higher sensitivity; both of these factors facilitating a more comprehensive exploration of how DNA topology modulated MmT6 ATPase activity.

When exploring ATPase activity of topos it is incredibly useful to have a competitive inhibitor that specifically binds the protein of interest and prevents further ATP hydrolysis, such as the use of novo with *E. coli* gyrase (Figure 3.28). Using novo allowed confirmation that the ATPase activity being measured was of gyrase origin. Highly effective inhibitors of MmT6 ATPase activity are lacking, which is a further limitation currently effecting topo VI research. However, there is one compound named radicicol which was shown to inhibit *S. shibatae* topo VI (SsT6) via a competitive mechanism similar to that of GyrB and novo [196, 215, 298]. For SsT6, full inhibition of activity was achieved with 200 µM radicicol, but when tested against MmT6, 600 µM and above was needed (Figure 3.30). This demonstrated that radicicol does not inhibit MmT6 to the same degree as SsT6, with neither cases coming close to the potency of novo on DNA gyrase. Nevertheless, 2 mM radicicol was employed here with the aim of trying to establish basal MmT6 activity, as contaminating ATPases, even in minute quantities, could affect the results due to the techniques sensitivity.
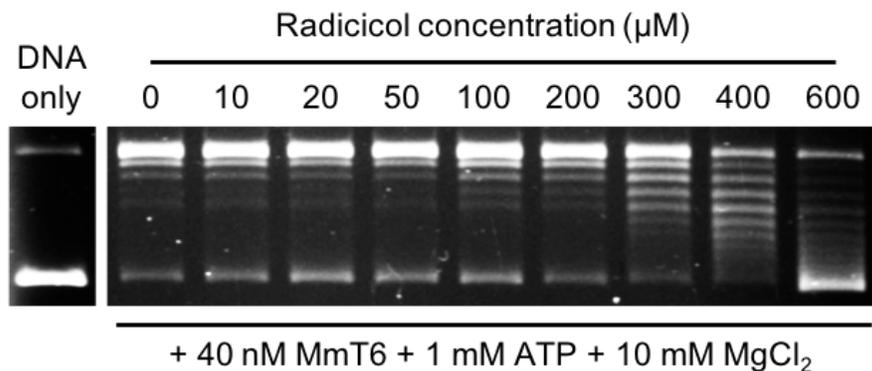
Figure 3.30: The effect of radicicol on MmT6 relaxation activity. MmT6 (40 nM) was incubated with 2 nM negatively supercoiled pBR322*, 1 mM ATP, 10 mM MgCl₂, and radicicol (1 to 600 µM), for 30 min at 37°C. The samples were run on a 1% (w/v) native agarose gel for 10 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination.

These assays explored MmT6 ATP hydrolysis activity when 1 μM MmT6 was incubated with 430 nM of either negatively supercoiled, relaxed or linear pBR322*, or in the absence of DNA (with or without 2 mM radicicol). Samples were taken at time points 1, 2, 4, 6, 8, 10, 15, 20 and 30 min and quenched using 5 μl of 2% SDS and 100 mM EDTA, for samples containing DNA. For those samples incubated in the absence of DNA, the time points were 10, 20, 30, 40, 50 and 60 min and 2 μM MmT6 was used. A 1 μL aliquot of each time point was then spotted onto a TLC plate and placed in a solution of 0.5 M lithium chloride and 1 M glacial acetic acid. The TLC plate was then fully dried and exposed to a phosphorscreen for 1 hr prior to visualisation using the Typhoon FLA 7000 plate reader. Results for the TLC plates can be seen in Figure 3.31, in which the $[\gamma\text{-}^{32}]$-ATP and released $[\gamma\text{-}^{32}]$-P$_i$ can be seen as a lower and upper signal, respectively. Much like gel electrophoresis, the released $[\gamma\text{-}^{32}]$-P$_i$ can be separated from the $[\gamma\text{-}^{32}]$-ATP as it is smaller and will therefore travel at a faster rate up the TLC plate. The ATP-alone control demonstrated that there was no background released $[\gamma\text{-}^{32}]$-P$_i$ in the ATP preparation. A very small amount of released $[\gamma\text{-}^{32}]$-P$_i$ was detected when MmT6 was incubated in the absence of DNA, which was mostly absent when 2 mM radicicol was added. Progressive accumulation of released $[\gamma\text{-}^{32}]$-P$_i$ can be seen when MmT6 is in the presence of all DNA substrates, however the spots are clearly more intense for the sample containing supercoiled pBR322* (Figure 3.31).

The data visualised by the TLC plates was converted to change in ATP concentration over time by analysing the TLC plate images using GelAnalyzer. The image was converted to a plot of pixel intensity and the area of the two peaks were calculated. The ratio in peak area allowed the calculation of how much ATP had been hydrolysed in each sample (Figure 3.32), which was then plotted against time to establish the rate of MmT6-dependent ATP hydrolysis.

Chapter 3: Ensemble Results
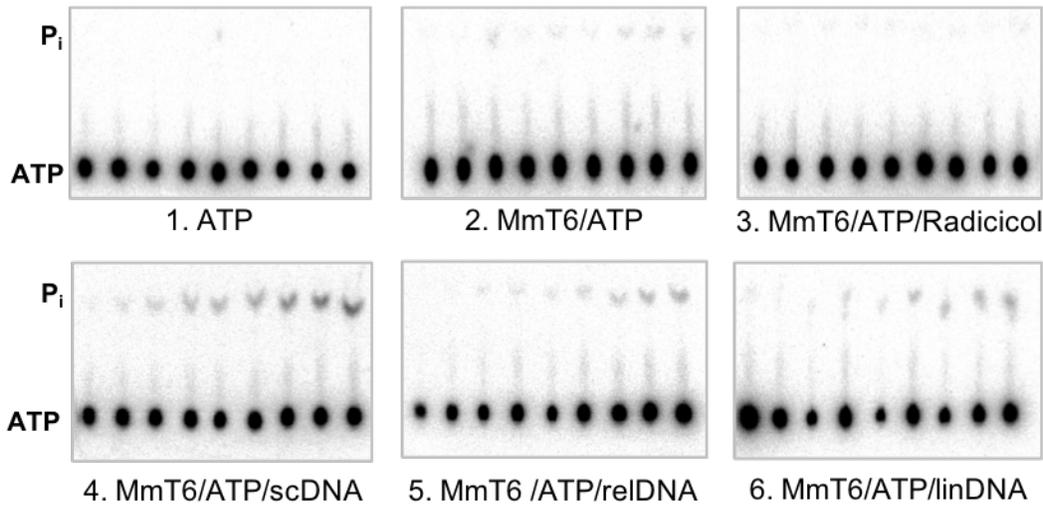
Figure 3.31: Examples of thin layer chromatography (TLC) plates showing MmT6-dependent P$_i$ release. The substrates common amongst all 6 TLC plates are 450 µM radioactive ATP ([γ-^{32}]-ATP) and 10 mM MgCl$_2$. TLC plate 1 was [γ-^{32}]-ATP only, plate 2 was 1 µM MmT6 and 450 µM [γ-^{32}]-ATP, plate 3 was 1 µM MmT6, 450 µM [γ-^{32}]-ATP and 2 mM radicicol, plate 4 was 1 µM MmT6, 450 µM [γ-^{32}]-ATP and 430 nM negatively supercoiled pBR322*, plate 5 was 1 µM MmT6, 450 µM [γ-^{32}]-ATP and 430 nM relaxed pBR322*, and plate 6 was 1 µM MmT6, 450 µM [γ-^{32}]-ATP and 430 nM linear pBR322*. All TLC plates imaged using the Typhoon FLA 7000 plate reader.

$$\frac{V1}{V1+V2} \times 450 = P_i \text{ released } (\mu M)$$
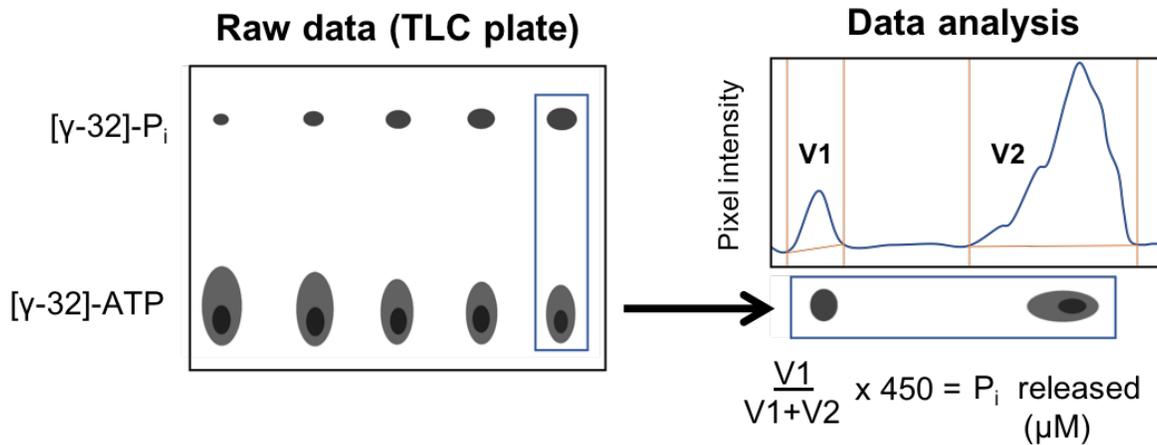
Figure 3.32: Illustration of how the raw data obtained from TLC plates was analysed for ATP hydrolysis rate of MmT6. The TLC plate shows 2 dark spots per sample, the lower being unhydrolysed [γ-^{32}]-ATP and the upper being released [γ-^{32}]-phosphate. Using GeneAnalyzer software, the pixel intensity is measured for each sample, which produced peaks and the volume of these were calculated. Using the equation shown in the illustration, the concentration of [γ-^{32}]-phosphate released by MmT6-dependent activity was calculated from these peak volumes (V1 and V2) and used to determine the rate of ATP hydrolysis.

All data derived from these assays can be seen in Figure 3.33. MmT6 ATPase activity was maximised in the presence of supercoiled pBR322* as compared to relaxed or linear, with the specific activities calculated to be 0.093 ± 0.01 s^{-1}, 0.002 ± 0.001 s^{-1} and 0.002 ± 0.002 s^{-1}, respectively (Figure 3.33A and D). This meant that the presence of DNA writhe enhanced MmT6 DNA-dependent ATPase activity by 4.7-fold. Previously published work on MmT6 also documented a ~5-fold increase in MmT6 ATPase activity when in the presence of supercoiled plasmid as opposed to sheared salmon sperm DNA [62], further validating the work described here. An intriguing aspect of the results found in the presence of supercoiled pBR322* was that after 15 min, the ATP hydrolysis rate declined to a rate similar to that achieved for linear and relaxed pBR322* (Figure 3.33A). To explore this, a DNA relaxation time course was performed in tandem using the identical DNA:enzyme ratio which was utilised in the radioactive ATPase assay (Figure 3.33B). This demonstrated that at 15 min all of the negatively supercoiled pBR322* had been fully converted to a relaxed product, providing further evidence that the ATPase activity of MmT6 was tightly coupled to the presence of a supercoiled substrate. In addition to this control over ATPase activity in the presence of DNA, the data obtained with MmT6 alone demonstrated that the basal MmT6 ATPase activity was incredibly low, at 0.008 ± 0.001 s^{-1}, which is 0.5 min^{-1} (Figure 3.33C and D). This is also in line with the previous work, which found the basal rate ≈ 0 s^{-1}, with the heightened sensitive of the radioactivity assay over the PK/LDH assay potentially allowing a more detailed assessment. This may also be why the values calculated for MmT6 specific ATPase activity were ~3-fold higher when measured using the radioactive ATPase assay, as opposed to the PK/LDH assay. In any case, all values calculated here and in prior published work using MmT6 [62] are relatively similar to each other and display highly equivalent behaviours, both in the DNA:enzyme ratio necessary to reliably collect data and the activity maximisation of MmT6 in the presence of DNA writhe. This adds a great deal of confidence that the work performed here is reflective of the true MmT6 behaviour and also established the use of either the radioactive ATP or PK/LDH assays as a means of exploring this.

Chapter 3: Ensemble Results
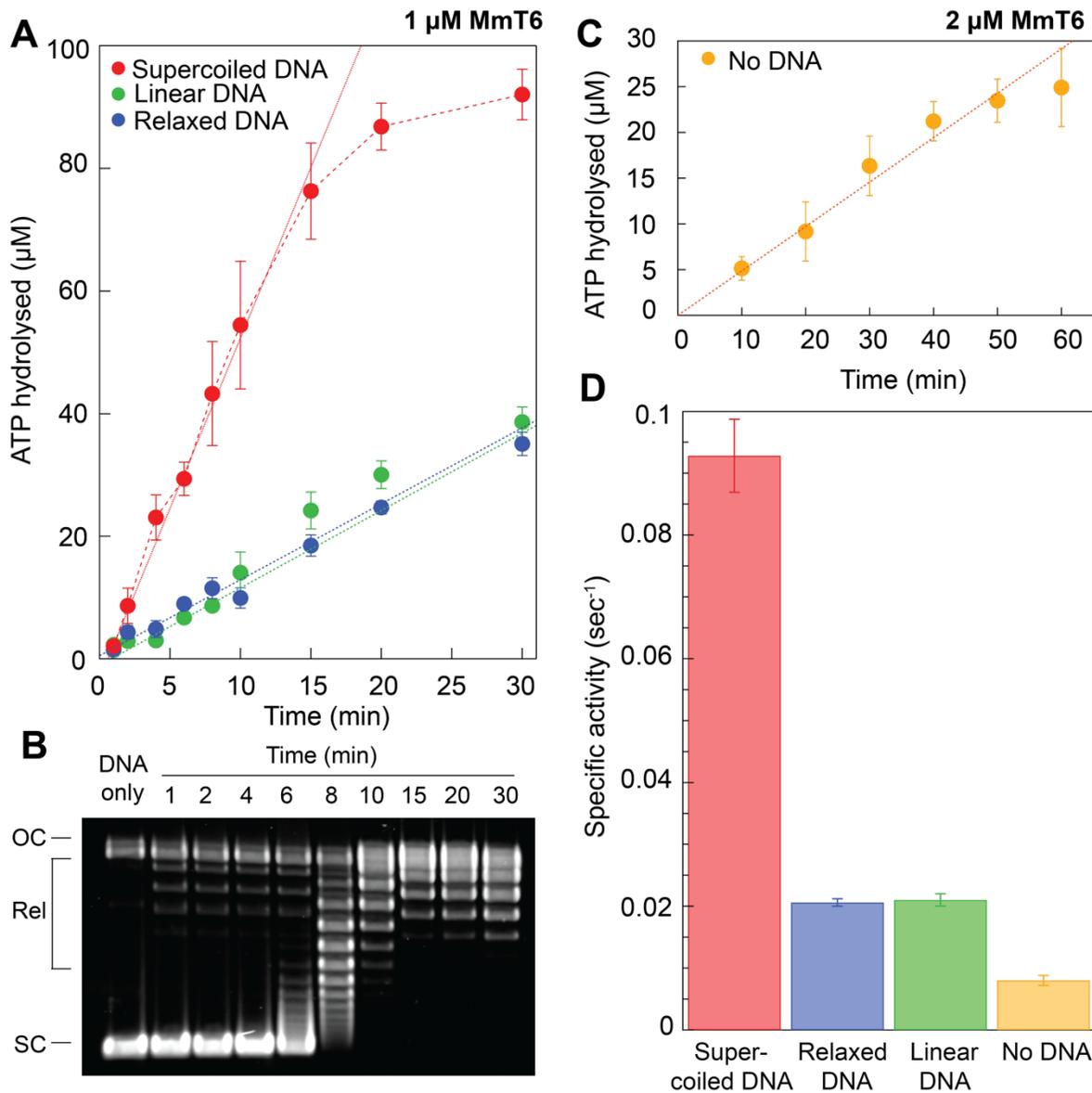
Figure 3.33: The effect of DNA topology on the ATP hydrolysis activity of MmT6. A: The radioactive ATPase assay was employed to measure the rate of ATP hydrolysis of MmT6 (1 μM) ± SE in the presence of 430 nM pBR322*, either negatively supercoiled (red), relaxed (blue) or linear (green), 450 μM [γ-32]-ATP and 10 mM MgCl$_2$, at 21°C. The data was fit well by a linear equation, in which the gradient was used to determine rates of ATP hydrolysis, and therefore the specific activity of MmT6 (sec^{-1}) in the presence of different pBR322* topologies. For the sample containing negatively-supercoiled DNA, the linear fit included data points from 1 to 15 min, whilst for relaxed and linear all data points were used in the fit. Data is the average across three separate assays. B: Agarose gel based MmT6-dependent relaxation time course performed to accompany the supercoiled data in Figure 3.33. The same protein:DNA ratio was employed (2:1), in which 39 nM MmT6 was incubated with 17 nM negatively supercoiled pBR322*, 450 μM ATP and 10 mM MgCl$_2$, at 21°C, with samples taken at the same time points as Figure 3.33 (1, 2, 4, 6, 8, 10, 15, 20 and 30 min). The samples were run on a 1% (w/v) native agarose gel for 15 hr at ~2 Vcm^{-1}, stained with 0.5 μg/mL ethidium bromide and imaged under UV illumination. OC: open circular or nicked,

Rel: relaxed, and Sc: supercoiled. C: Basal rate of MmT6 hydrolysis ± SE, in the absence of DNA, was measured using the radioactive ATPase assay, by incubating 2 µM MmT6 with 450 µM [γ-32]-ATP at 21°C, taking a sample at 10, 20, 30, 40, 50 and 60 min (yellow). The data were fit to a linear equation, which was used to determine the rate of ATP hydrolysis and therefore the specific activity of MmT6 (sec^{-1}) in the absence of DNA. D: The specific activity of MmT6 ATP hydrolysis ± SE, in the presence of different DNA topologies, supercoiled (red), relaxed (blue), or linear (green) or no DNA (yellow).

3.3 Conclusions

Throughout this chapter, the utilisation of various biochemical techniques in the characterisation of MmT6 activity has led to gaining a deeper insight into how MmT6 functions and under which conditions, providing the ideal backdrop for single-molecule characterisation, detailed in Chapter 4. The data described here reveal a topo that has a mild rate enhancement for the relaxation of positive writhe, increased activity in the presence of catenated DNA molecules, and a significant preference for the binding and relaxation of DNA writhe. Taken together, the data reveals a picture of MmT6 as a DNA crossing sensor, binding and resolving crossings in an ATP-dependent fashion, with loss of writhe leading to loss of ATP hydrolysis activity. DNA topology has a striking effect on activity, with ATPase and DNA binding enhanced 5-fold and 4-fold, respectively, in the presence of negatively supercoiled pBR322* as opposed to linearised pBR322*. This data suggested that topological features of supercoiled DNA were recognised and acted upon by MmT6, likely the intramolecular DNA crossings formed due to the increased superhelical density, and that these crossings stimulate MmT6 activity, which is minimal in their absence. MmT6 ATPase activity is tightly coupled to presence of DNA writhe, with MmT6 basal activity in the absence of DNA practically non-existent. The explanation for this behaviour may be due to the enzyme only operating with two protein interfaces which must alternately open and close to both allow the passage of the T-segment, as well as protect from forming stabilised double-stranded DNA breaks once the G-segment has been cleaved. Type IIA enzymes do not appear to have quite such tightly coupled ATPase activity to the presence of both the G- and T-segment within the enzyme cavity, potentially due to the presence of the third protein interface, known as the C-gate. This may afford a higher level of protection against generating double-stranded DNA breaks.

MmT6 ensemble characterisation using biochemical approaches has allowed deeper insight into the in vitro behaviour of topo VI enzymes as well as facilitate further considerations as to its *in vivo* roles. With a preference for *in vitro* decatenation, and an intrinsic inefficiency during supercoil relaxation, it is likely that MmT6 is a preferential DNA decatenase *in vivo* that functions via the direct detection of DNA crossings. This is likely modulated by a preference for a particular DNA crossing geometry which is found more frequently in DNA catenanes than DNA supercoils. This hypothesis will be expanded upon in the chapters to come (particularly Chapter 4), with a complete picture presented in Chapter 6.

3.4 Future Work

During the discussion of the data presented in this chapter, numerous areas which require further study have been highlighted. This section will bring these suggestions together, as well as speculate on further biochemical characterisation of MmT6 which is currently outstanding. With regards to ATPase characterisation, exploring how catenated DNA and positively supercoiled DNA modulate this activity would inform further on the preferential activity of MmT6, which was not completed during the work reported here. If MmT6 preferential relaxes positive writhe over negative, and preferentially decatenates rather than relaxes, an even greater stimulation of ATPase activity should be present than was seen in the presence of negative writhe. Measurements of MmT6 binding on catenated DNA may also be illuminating.

On particularly intriguing and unresolved element of *in vitro* MmT6 activity explored here is that DNA cleavage and binding are higher for negative writhe, despite a ~2-fold preference for the relaxation of positive writhe. Further exploration into why this occurs may advance understanding on the true DNA binding and cleavage behaviour of MmT6. The explanation offered here is that during binding and cleavage, only the G-segment needs to be present, however for relaxation to occur, both the G- and the T-segment are necessary. The techniques used to explore binding and cleavage do not provide a way of distinguishing whether one or both DNA segments were present and therefore it may be that binding and cleavage are facilitated to a higher degree on negative writhe through an enhancement of MmT6 G-segment

only interactions. Perhaps the underwound DNA interacts with elements of Top6A more readily, and/or can be bent more easily, G-segment bending being a behaviour already known for MmT6 [61]. It may also more closely resemble a transition-state.

Another facet of MmT6 activity touched upon here is that increasing magnesium concentration alters relaxation and cleavage activity in different ways. Relaxation is inhibited at 50 mM Mg^{2+} whereas cleavage at this concentration remains robust. A potential explanation for this is reminiscent of the data discussed in the previous paragraph. For cleavage only the G-segment need be present and so perhaps 50 mM Mg^{2+} does not prevent MmT6 binding to the G-segment, but does prevent the binding and passage of the T-segment, possibly through electrostatic effects mediated by the Mg^{2+} or interference with nucleotide binding and Top6B dimerisation. The way in which Mg^{2+} concentration effects MmT6 activity is worthy of further analysis, with deeper characterisation of how Top6A coordinates and utilises Mg^{2+} ions via the TOPRIM domain not only informing on topo VI enzymes, but potentially also type IIA topos and Spo11. Also, further exploration of the effects of Co^{2+} on MmT6 activity may be enlightnening.

Chapter 4

Characterisation of *M. mazei* Topo VI Using Single-Molecule Magnetic Tweezers

4.1   Introduction

The utilisation of magnetism in biological research began in 1950 with *in vivo* experiments exploring physical qualities of the cytoplasm, in which cells phagocytosed magnetic particles that were then manipulated using external magnetic fields [299]. Later, single DNA duplexes were tethered to a slide at one extreme end, and to a magnetic bead at the other, in which DNA elasticity [300] and supercoiling [301] were explored by manipulating the DNA molecule using external permanent magnets. This formed the initial groundwork for the development of the magnetic tweezers (MT) as a broadly applicable technique in the single-molecule study of biological polymers. Whilst the work reported here relies on manipulation of DNA, single-molecule MT have also been used in research on microtubules [302], collagen [303], and protein folding/unfolding [304, 305] (see [306] for a comprehensive review). However, DNA remains the most commonly explored biopolymer using MT, not only allowing the characterisation of physical DNA properties [300, 301, 307-309] but also the activity of proteins that interact with DNA such as helicases [227, 310, 311], polymerases [312, 313] and topoisomerases (topos) [119, 120, 158, 222, 224, 314]. As the magnetic tweezers are able to achieve forces of 10^{-3}-10^2 piconewtons (pN), have a temporal resolution of 10^{-1}-10^{-2} and, most importantly, can apply torque [315], they have been particularly effective in topo research (see Chapter 1, section 1.5 for more information) as DNA topology can be precisely controlled and manipulated, with the activity of topos being tracked at high temporal-spatial resolution [222].

There are variations on the magnetic tweezers set up (see [222, 315]), with the basic configuration employed in the research reported here (Figure 4.1). A sample cell, housing the

experimental substrates, is mounted above an inverted oil immersion microscope objective, atop motorised stages. Above the sample cell, two permanent magnets are fixed 1 mm apart
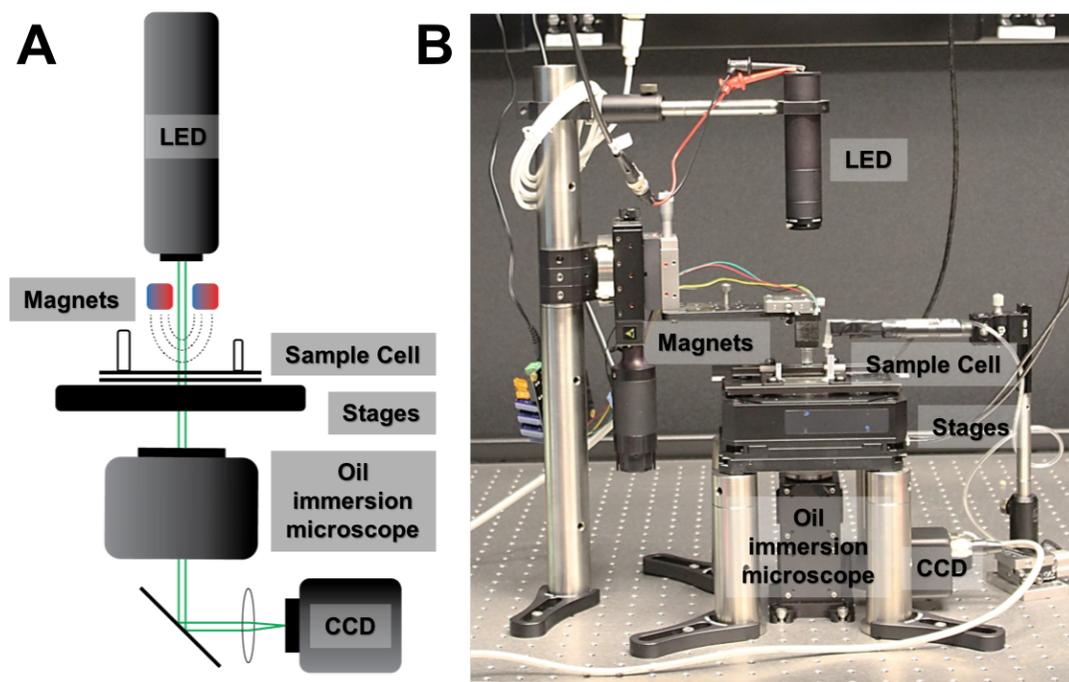
Figure 4.1: Magnetic tweezers (MT) instrument set up. A: Simplified schematic highlighting the major components of the MT instrument. B: Photograph of the MT instrument used in the research reported here.

which generate the magnetic field used to manipulate the magnetic beads within the sample cell. An LED light source is focused on the sample cell and relays an image to a CCD camera. See Figure 4.1A for a schematic and 4.1B for a photograph of the actual instrument used. Within the sample cell and via non-covalent interactions, DNA is tethered at one end to the slide surface via the interaction between digoxigenin and antidigoxigenin, and at the other end to a micron-sized, streptavidin-coated paramagnetic bead via interactions with DNA-incorporated biotin. Either a single torsionally-constrained duplex, or two unconstrained duplexes can be attached to the bead, depending on whether DNA supercoiling or braiding is being explored, respectively (Figure 4.2). The height of the bead from the surface of the sample cell is obtained using the diffraction rings, which are produced as a consequence of light scattered by the bead interfering at the camera. The appearance of these concentric rings around the bead being tracked change depending on the position of the bead from the sample cell surface. Rotation of the magnets above the sample cell causes the entrained magnetic

bead within to rotate, thereby also rotating the attached DNA and changing its linking number. For a single, intact, torsionally-constrained duplex, this rotation causes twist to accumulate before the DNA buckles to generate writhe, forming a supercoiled structure known as a plectoneme (Figure 4.2A). The formation of the plectoneme is accompanied by a reduction in DNA extension, bringing the bead closer to the surface of the sample cell. Therefore, the activity of topos, which generally relax supercoils, can be tracked as a function of DNA extension. With each topo-dependent removal of a supercoil, the DNA extension increases and this can be measured in real time and with high resolution, affording the ability to see single strand-passage events (Figure 4.2A). For DNA braiding, the situation is similar, however as the DNA is unconstrained (e.g. via a DNA nick, or a single interaction with the sample cell surface), twist cannot be stored in the duplexes as the magnets are rotated. Instead, the DNA molecules wrap each other, generating writhe without changing the twist. The generation of a DNA braid, also results in reduction of DNA extension, and topo-dependent unlinking of the braid causes the DNA extension to increase, allowing the activity to be tracked in the same manner as supercoil relaxation (Figure 4.2B).
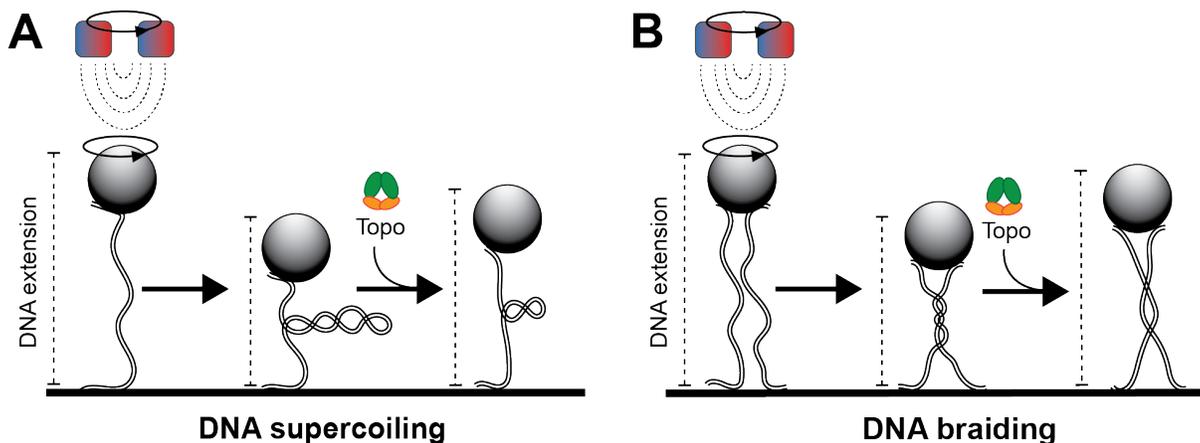
Figure 4.2: Illustration of DNA topology manipulation from within the sample cell. A: *DNA supercoiling*. Rotation of the fixed magnets above the sample cell causes the bead-bound DNA duplex within to rotate, producing a supercoiled structure called a plectoneme, which decreases the DNA extension. DNA topoisomerases (topos) can relax the plectoneme, removing supercoils and increasing the DNA extension. B: *DNA braiding*. Rotation of the fixed magnets above the sample cell, causes the two bead-bound DNA duplexes within to wrap each other, producing a DNA braid, which decreases the DNA extension. Topos can unlink the braid, increasing the DNA extension.

Here, using topo VI from *M. mazei* as a model (MmT6), the single-molecule magnetic tweezers assay has been employed, in combination with ensemble biochemistry (described in Chapter 3), to gain a deeper insight into the topo VI mechanism, particularly how the topology of the DNA substrate modulates activity. Previous research, whilst instructive, has left us with the conundrum that the activity of topo VI seems far below that which is necessary to support the cell in terms of transcription- and replication-induced supercoil relaxation [62]. The data presented here forms a compelling case for the preferential decatenation activity of MmT6 as an intrinsic quality of the enzyme, arising from a strict DNA crossing angle preference of 88°, which in turn significantly disfavours the removal of supercoils. MmT6 also behaves as a true DNA crossing sensor, with a dramatic increase in the rate of strand passage with increasing DNA writhe. Moreover, these data extend to providing an explanation for why topo VI remains a vital component of some eukaryotic systems, such as during endoreduplication in plants [169, 178], despite the presence of other highly-efficient type IIA topos. Topo VI may be a dedicated decatenase, the activity of which will not be dominated by the necessity to relax transcription- and replication-induced supercoils (discussed in Chapter 6).

4.2   Results and Discussion

4.2.1 Magnetic force calibration

The MT assay is a technique that facilitates the manipulation of DNA topology under precise force, allowing for in-depth characterisation of how DNA tension and topology affects the catalytic activity of MmT6. As described above, in the set up used here, a DNA duplex was tethered to a sample slide at one extreme end, and to a 1 µm magnetic bead at the other. Two permanent magnets were fixed parallel to the sample cell, above the tethered DNA, and exerted an upwards magnetic force on the bead. The magnitude of this force is controlled by the distance between the sample cell (housing the DNA tethered magnetic beads) and the permanent magnets. As the magnets moved further from the slide, the force on the DNA duplex decreased. To calibrate the force applied to the DNA as a function of magnet position, the DNA

Chapter 4: Single-Molecule Results

extension and the variance of the bead position in the x coordinate were experimentally determined. They are related to the applied magnetic force (F_z) by the equation:

$$F_z = \frac{k_B T z}{\langle x^2 \rangle}$$

Equation 4.1: Calculating the magnetic force on the DNA-tethered bead.

Where F_z is the force applied to the bead by the magnets, k_B is the Boltzmann constant, T is the temperature in Kelvin, z is the DNA extension and x is the displacement of the bead in the x coordinate. To derive this equation, some physical concepts require consideration.

To begin, the tethered bead is considered as an inverted simple pendulum. For a simple pendulum the downward force acting on the suspended object is gravity. For the DNA tethered bead, the equivalent force is upwards and supplied by the magnetic field. In an ideal system, a simple pendulum when displaced from equilibrium, oscillates back and forth infinitely in a sinusoidal manner. In reality, friction and air resistance will cause a pendulum to eventually stop swinging and return to its equilibrium position. For the magnetic tweezers system, the bead is in water and at the microscopic scale this could be compared to a macroscopic pendulum submerged in treacle. If you were to displace this treacle-submerged pendulum it would move back to equilibrium but no longer oscillate around it, and this is described as a critically damped system. Whilst the bead is, in reality, being constantly displaced from equilibrium due to the Brownian motion of water particles, the force from the magnetic field constantly, and without oscillation, brings the bead back to equilibrium. Another assumption for a simple pendulum is that the tether is mass-less in relation to the suspended mass and can therefore be disregarded. The magnetic tweezers set up achieves better than this as mass is not a considered variable in the calculations. A third simple pendulum assumption is that the tether cannot bend nor stretch and whilst DNA can undergo these deformations, in these assays, the position of the bead is averaged over such a long period of time so that an accurate measurement can be obtained. The final assumption addressed here is that a simple pendulum swings in a perfect plane. Again, whilst the magnetic bead does move in the x, y and z coordinates, the calculations focus only on its movements in the x direction and this is recorded

over a considerable period such that an accurate average can be determined. It's important to note here that even for a simple pendulum in a perfect system, approximations and simple assumptions are made during calculations.
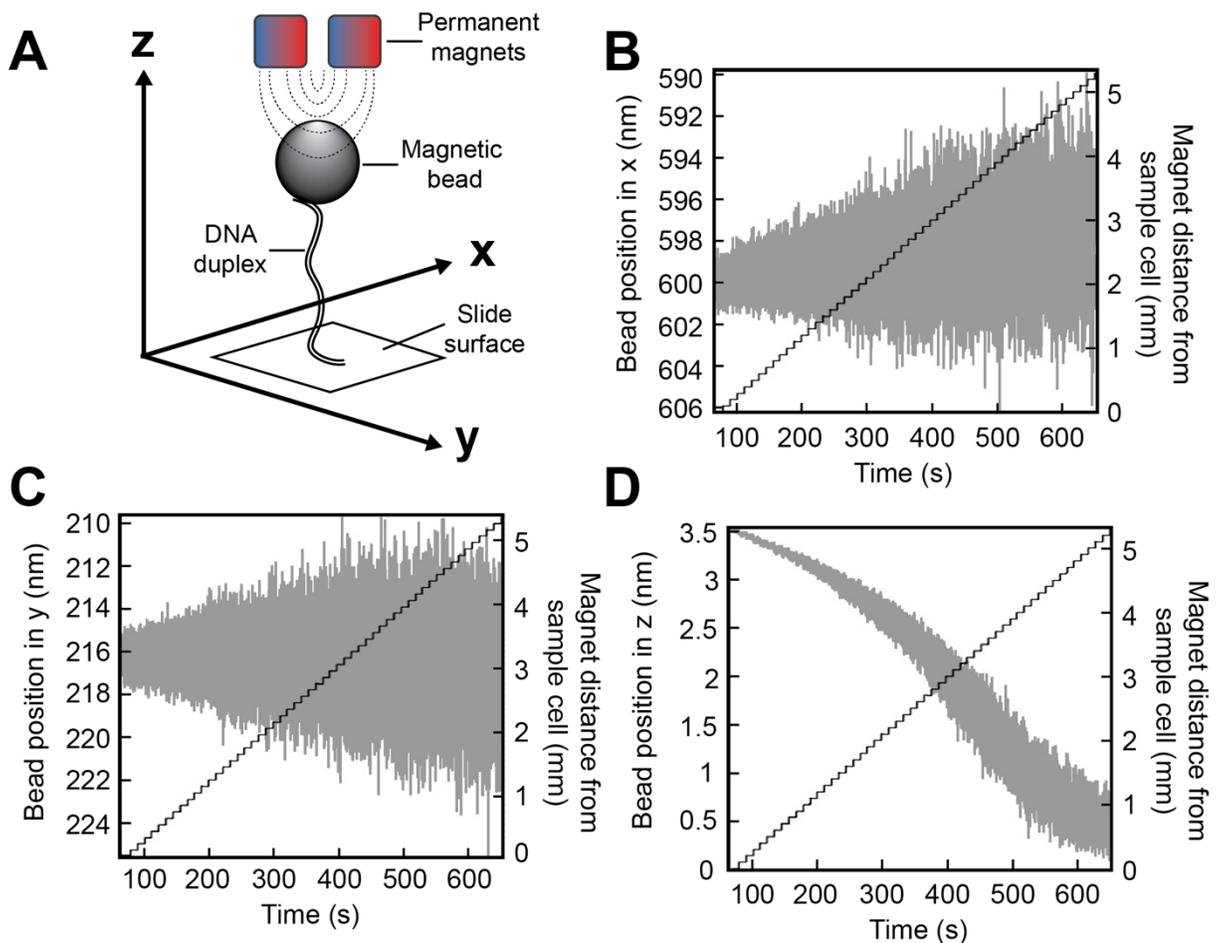
Figure 4.3: Effect of magnet proximity on bead displacement. A: Illustration of coordinate annotation, from the perspective of the tethered DNA duplex. The x and y coordinates are parallel to the sample cell surface, while z is perpendicular. B-D: The displacement of the bead (left y axis, grey line) over time (x axis) as a function of magnet position from the sample cell (right y axis, black line), for the x, y and z coordinates, labelled B, C and D, respectively.

The displacement of the bead from equilibrium arises from constant bombardment by water molecules in all directions, i.e. Brownian motion, and this displacement in the x coordinate is related to the force in z. Shown in Figure 4.3 and 4.4, as the magnets were moved further from the sample cell, the magnetic force acting on the bead decreased, which allowed the bead displacement due to Brownian motion to increase, as the magnets exerted less upward force to counteract this. This demonstrated how the beads displacement is related to

the magnetic force. In the z coordinate, not only does the bead move more drastically from equilibrium, the DNA extension also decreases, as the upward force on the bead diminishes (Figure 4.3D). In the x coordinate the bead is entrained by the magnetic field lines and so maintains alignment. Whilst it is possible to use the variance of the bead position in the y coordinate to calculate the force in z, being perpendicular to the magnetic field lines results in the bead being able to freely rotate, and hence the radius of the bead must be considered. For the purposes of this research, using the bead displacement in the x coordinate is sufficient. However, in systems using very short pieces of DNA, using the bead displacement in y can increase accuracy of measurements as the need to include the radius of the bead increases the extension value. Here, a 12 kb DNA duplex was utilised to experimentally determine the magnetic force. This is because the longer the DNA is, the less stiff it becomes and therefore the speed of the oscillation is reduced. This is important as the DNA mustn't oscillate near the frame rate of the camera if accurate measurements are to be gathered.
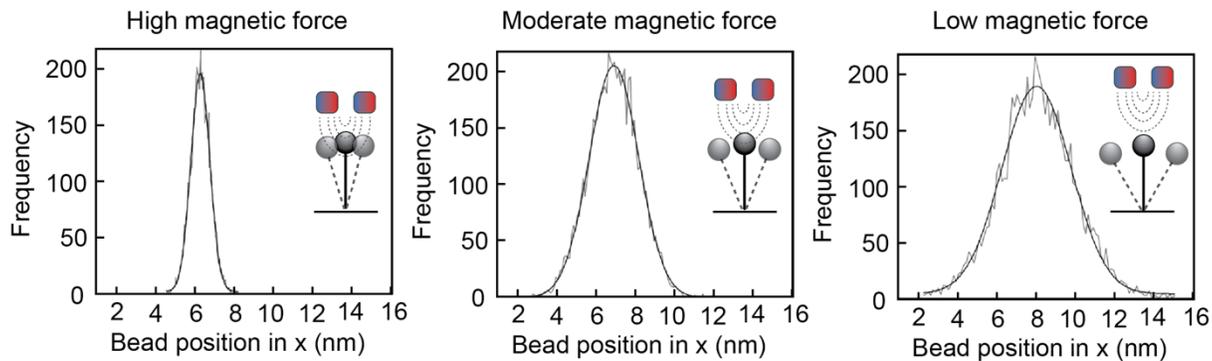
Figure 4.4: The distribution of the beads position in x as a function of magnet position from the sample cell. When the magnets are close to the sample cell, they exert a high magnetic force on the bead, reducing the extent of its displacement from equilibrium. As the magnets are moved further from the sample cell, the magnetic force deceases and the bead displacement increases.

To begin deriving an equation to calculate the magnetic force on the bead, the physics of the simple pendulum is used. With reference to figure 4.5, some terms need defining: F_x is the restoring force which occurs in the opposite direction to the displacement force, F_z is the upwards force on the magnetic bead from the magnetic field gradient, T is a tensile force
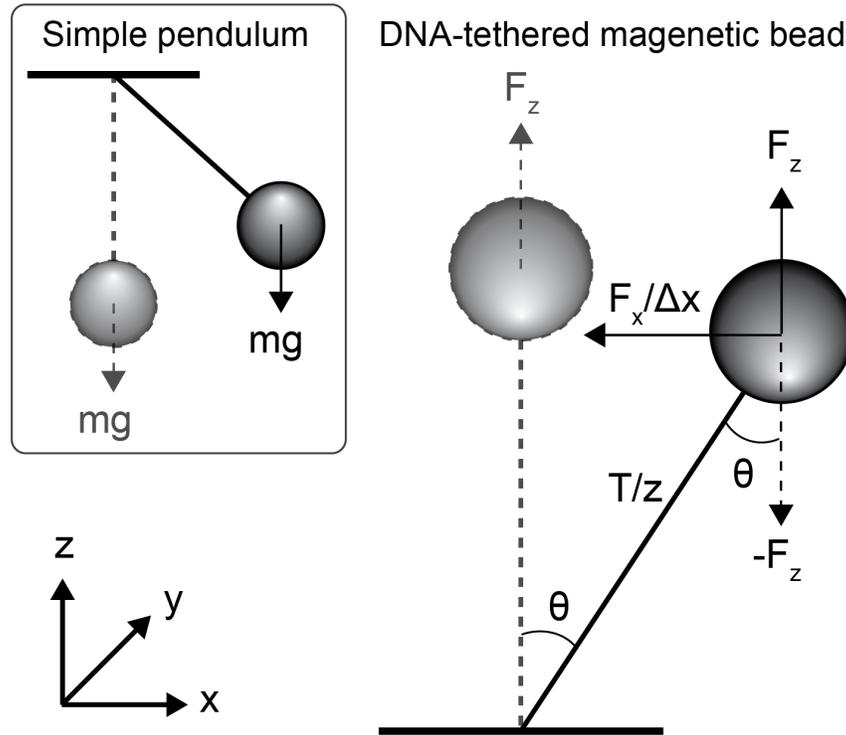
Figure 4.5: The simple pendulum and the DNA-tethered magnetic bead. The simple pendulum is acted upon by the downwards force of gravity multiplied by the mass of the tethered object. The DNA tethered bead displaced from equilibrium is acted upon by the upward magnetic force (F_z). F_x is the restoring force which occurs in the opposite direction to the displacement from equilibrium, Δx is the distance of displacement from equilibrium in the x direction, T is a tensile force provided by the DNA tether, z is the DNA contour length and θ is the angle of displacement from the equilibrium.

provided by the DNA tether, *Δx* is the distance of displacement from equilibrium in the x direction, *z* is the DNA contour length and *θ* is the angle of displacement from equilibrium. Using trigonometry, we know,

$$T = \frac{-F_z}{\cos \theta}$$

Equation 4.2: Trigonometry relates the tensile force (*T*) to the magnetic force in z (*F_z*) and the angle of displacement (*$\cos\theta$*).

$$F_x = T \sin \theta$$

Equation 4.3: Trigonometry relates the restoring force (*F_x*) to the tensile force (*T*) and the angle of displacement (*$\sin\theta$*).

Or in terms of distance,

Chapter 4: Single-Molecule Results

$$\Delta x = z \sin \theta$$

Equation 4.4: Trigonometry relates the distance of displacement (Δx) to the DNA contour length (z) and the angle of displacement ($sin\theta$). Equation 4.3 written in terms of distance, rather than force.

The small angle approximation is then used which states that if the angle of displacement is small then $cos\theta \approx 1$, which then follows that $F_z \approx T$, so,

$$F_x \approx \frac{-F_z}{z} \Delta x$$

Equation 4.5: The restoring force (F_x), is roughly equal to the magnetic force in z (F_z), divided by the DNA contour length (z) and multiplied by the beads displacement in x (Δx).

Equation 4.5 highly resembles Hooke's law, which describes the force needed to displace a resting spring (F_x) in terms of its stiffness, otherwise known as the spring constant k (measured in N/m), multiplied by the distance of displacement (Δx),

$$F_x = -k \, \Delta x$$

Equation 4.6: Hooke's law.

So, now the movement of the bead in the x coordinate follows the physics of a spring, another equation can be considered which describes the potential energy (PE) stored in a spring when it is displaced from equilibrium, given by,

$$PE = \frac{1}{2} k \Delta x^2$$

Equation 4.7: Elastic potential energy equation.

Where k is the spring constant and Δx is the distance of displacement. This equation can then be combined with the equipartition theorem equation. The equipartition theorem states that within a closed system the thermal energy is on average distributed equally amongst each component, or degree of freedom, and this is equal to $\frac{1}{2}k_B T$, where k_B is the Boltzmann constant and T is the temperature in kelvin. The movement of the bead in x (Δx) can be considered a single degree of freedom, which means the equipartition equation can be set to equal equation 4.7. This then leads to,

$$k = \frac{k_B T}{x^2}$$

Equation 4.8: Equipartition theorem combined with the elastic potential energy equation.

If equation 4.8 is combined with equation 4.5 which finds k ($F_x/\Delta x$) equal to F_z divided by the DNA contour length (z), you can then solve for F_z:

$$F_z = \frac{k_B T z}{\langle x^2 \rangle}$$

Equation 4.1: Calculating the magnetic force on the DNA-tethered bead.

As is well established, the Boltzmann constant multiplied by room temperature expressed in kelvin is 4.06 pN nm. The average displacement of the bead in x (Δx), along with the DNA contour length (z) were experimentally determined and used to calculate the force on the bead as a function of magnet position.
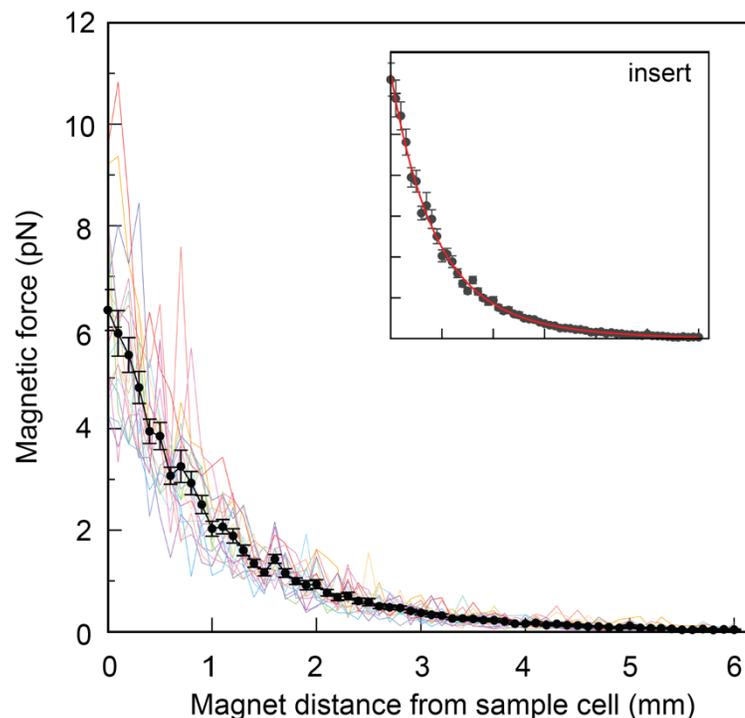
Figure 4.6: Magnetic force on the bead as a function of magnet position from the sample cell. All 15 data sets are rainbow coloured with the average force at each magnet position plotted ± SE in black (black circles and black line). The insert shows the double exponential line of fit in red, with the equation F = +3.4674 exp(-0.64175Zmag)+2.6918exp(-6.3144Zmag).

Data was experimentally obtained and analysed in IgoPro7 (wavemetrics) using a procedure coded by Dr Keir Neuman (NHLBI, NIH, Bethesda, US). For all 15 data sets, the extension of the DNA at each magnet position and the variance of the bead position in the x

coordinate were extracted and used to calculate the force at each position (Figure 4.6). The data sets were then averaged and fit to a double exponential with the equation,

$$F = +3.47 \exp(-0.64 Zmag) + 2.69 \exp(-6.31 Zmag)$$

Equation 4.9: Double exponential equation fit to magnetic tweezers force calibration data. *F* is the force on the magnetic bead and *Zmag* is the position of the magnets above the sample cell.

This equation was then inputted into the software that runs the MT instrument so that the experimentalist can either get an accurate idea of the force on the DNA at a given magnet position, or input a particular force and have the magnets move to the appropriate position automatically. This became an invaluable tool in subsequent data collection.

4.2.2 MmT6 supercoil relaxation

To begin exploring the single-molecule behaviour of MmT6, the MT supercoiling assay was employed, described in detail elsewhere [221, 222]. As opposed to agarose gel analysis of DNA topology, the magnetic tweezers imparts not only control over the precise level and chirality of DNA supercoiling, but also allows the real-time detection of supercoil relaxation by MmT6 with the ability to capture single strand-passage events. Rotating the magnetic bead anticlockwise, leads to the formation of positive writhe (left-handed crossings), whilst negative writhe (right-handed crossings) is generated by rotating the bead clockwise (Figure 4.7). Generally, data collection began with the calibration of the DNA tether. The DNA extension is plotted as a function of DNA turns, forming a plot known as a hat curve, owing to its shape (Figure 4.7). At low force (0.2 pN), as the magnets were rotated either clockwise or anticlockwise, the DNA extension decreased symmetrically as writhe was generated (Figure 4.7, light grey). However, at high force the hat curve becomes asymmetric, as the clockwise rotation promoted DNA melting over the formation of negative writhe. Generally, this didn't affect MmT6-dependent assays, as a lower force (0.4 pN) was found to be optimal. It is a consideration, however, in section 4.2.3, exploring the effect of force on MmT6 activity. Calibration of the chosen DNA tether is important as it established whether the bead was

indeed tethered by a single, coilable DNA molecule. Sometimes the bead was found to be tethered to two or more duplexes, generating an uncharacteristic hat curve, or the DNA was nicked, meaning twist and writhe cannot be stored in the molecule and so DNA extension was insensitive to magnet rotation, remaining fully extended.
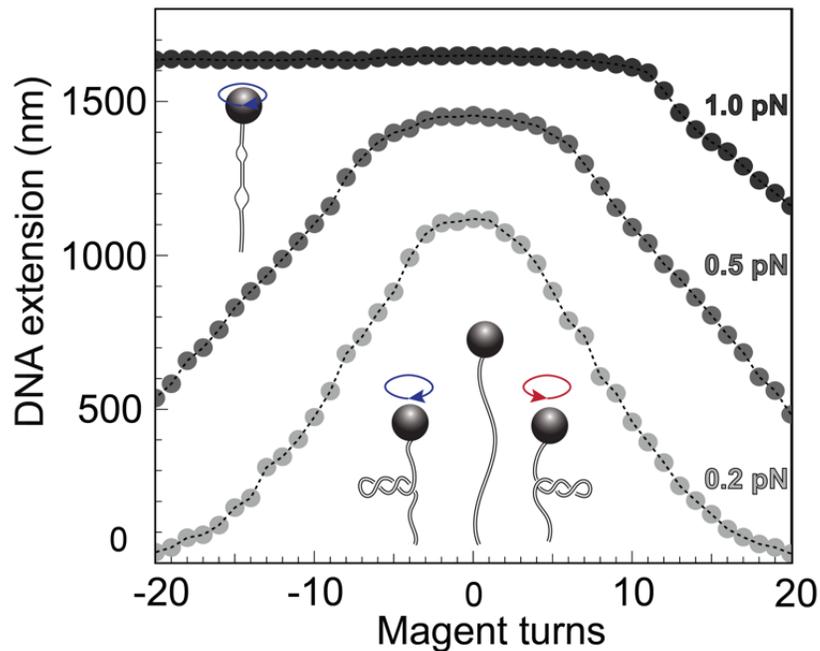
Figure 4.7: Coilable 5-kb DNA tether hat curve. DNA extension is plotted as a function of magnet turns at three forces, 0.2 pN (low force, light grey), 0.5 pN (low-moderate force, medium grey) and 1.0 pN (high force, dark grey). At high force the DNA extension is insensitive to clockwise magnet rotation, as DNA is melted rather than negatively supercoiled. At low-to-moderate force, DNA extension decreased as the magnets were rotated clockwise/anticlockwise, due to the formation of DNA writhe.

As previously described, rotation of the DNA leads to a reduction in its extension due to plectoneme formation, which can be tracked using the diffraction pattern of the attached magnetic bead. Upon addition of MmT6 and ATP, the DNA extension increased over time as supercoils were removed by MmT6, hence the relaxation reaction was tracked via DNA extension change over time (Figure 4.8). Based on the calibration (Figure 4.7), the removal of a supercoil at 0.4 pN equated to an extension change of ~110 nm, with MmT6-dependent relaxation occurring as an abrupt increase in of ~110 nm. This is equal to a change in Lk of 2,

Chapter 4: Single-Molecule Results

and was expected of MmT6 as it is a type II topo. Figure 4.8 demonstrates three examples of MmT6-dependent plectoneme relaxation, in which 16 positive turns were added to the DNA molecule before being relaxed in steps of ~110 nm (ΔLk=2). Once the DNA extension exceeds the threshold set, the magnets automatically recoil the DNA.
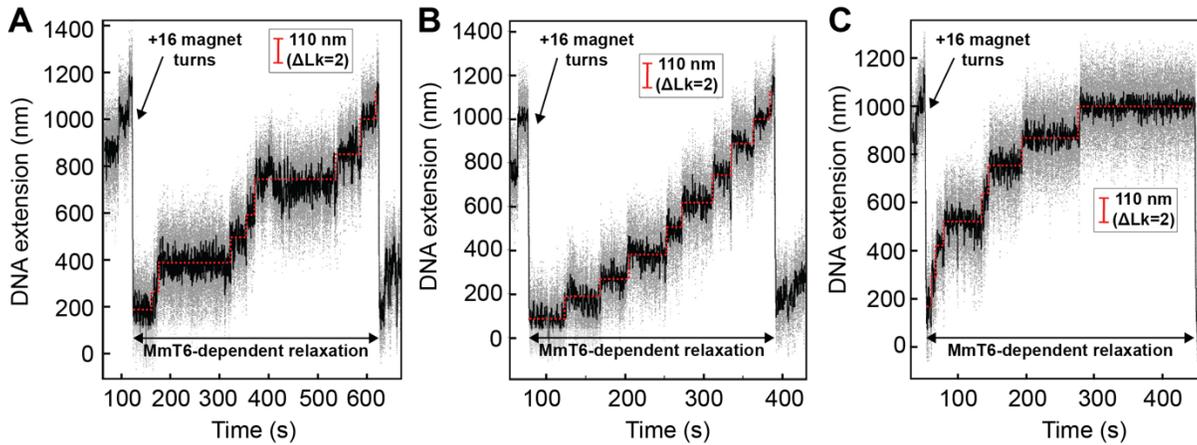
Figure 4.8: MmT6-dependent relaxation of a plectoneme. A, B and C: All three traces are examples of MmT6-dependent plectoneme relaxation of positive writhe using 500 pM MmT6, 1 mM ATP, under 0.4 pN at 21°C. The single-headed arrows indicate the point at which the magnets are rotated, adding 16 positive turns to the DNA. The double headed arrows indicate the period of time in which MmT6 relaxed the positive writhe. This is evident as abrupt DNA extension increases of ~110 nm, which are highlighted using a dashed red line, and represent single strand-passage events by MmT6. The light grey points are the raw data (frame rate = 200 Hz) with the black representing the smoothed data, achieved using the Savitzky–Golay filter in Igor Pro 7 (smoothing window = 199 data points).

Once established that MmT6 was amenable to exploration using the MT set up, the chirality-dependent relaxation activity of MmT6 was measured as a function of MmT6 concentration (250-2000 pM) (Figure 4.9). The observed DNA extension changes revealed that MmT6 relaxed negative supercoils 2-3 times slower than positive supercoils. The measured average relaxation rates for supercoils of either chirality increase as a function of MmT6 concentration with the data fit to a Michaelis-Menten-like rate equation, resulting in an apparent K_d ($K_{d,app}$) 3-fold lower for the relaxation of positive DNA writhe.
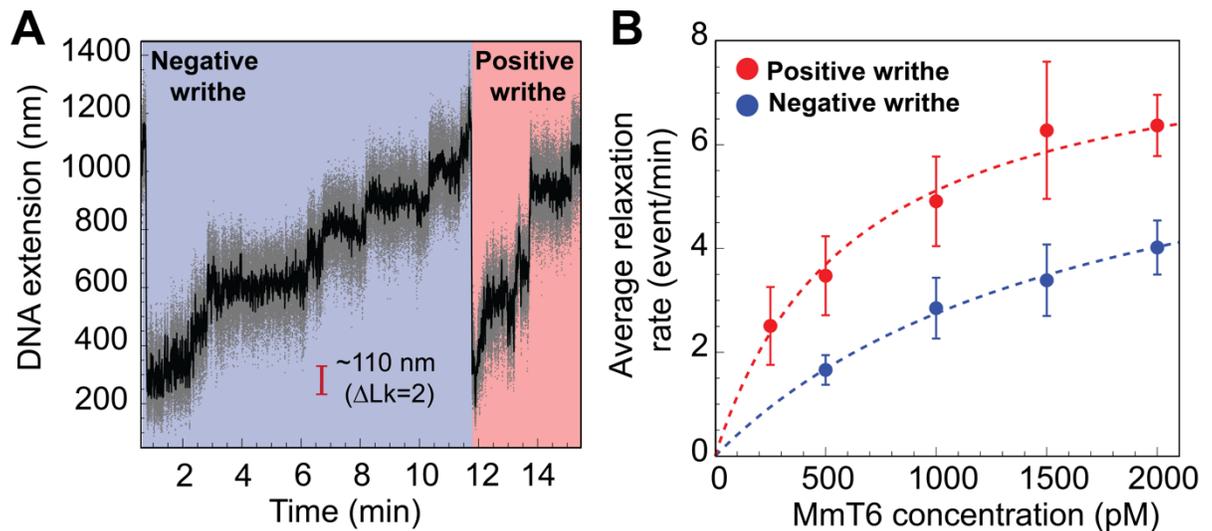
Figure 4.9: Magnetic tweezers supercoil relaxation data. A: Example trace of MmT6-dependent supercoil relaxation. Data collected under 0.4 pN, at 21°C, using 500 pM MmT6 and 1 mM ATP. Each strand passage event is evident as an abrupt DNA extension increase of ~110 nm, corresponding to a ΔLk of 2. The light grey points are the raw data with the black representing the smoothed data, achieved using the Savitzky–Golay filter in Igor Pro 7. Relaxation of negative writhe is highlighted in blue, and positive in red. Positive writhe is relaxed faster, resulting in shorter events (<1 min) being compressed when plotted on the same axis as negative supercoil relaxation. B: Average relaxation rate of MmT6 (± SEM) on positive and negative supercoils as a function of MmT6 concentration, collected under 0.4 pN, at 21°C, and with 1 mM ATP. Data were fitted (dashed lines) to a rate function ($V_0 = \frac{V_{max}[E]}{K_{d,app}+[E]}$). Raw data was analysed in Igor Pro 7 using a T-test-based method, first described in [227].

In line with the single-molecule relaxation assay results, the preferential relaxation of positive supercoils by MmT6 was also supported by an agarose gel based approach (Figure 4.10), showing the relaxation of supercoiled pBR322* to occur at 4 min for positively supercoiled and 8 min for negatively supercoiled. An aspect of topo VI activity, previously described in [62] and built upon here, is that the rate of reaction, whilst reliable, is far slower than type IIA topos in supercoil relaxation. Even at high concentrations above those considered standard in single-molecule measurements (e.g. 2 nM MmT6), the maximum rate measured for positive and negative supercoil relaxation was only 6.4 ± 0.6 and ~4 ± 0.5 strand passage events min[-1], respectively. Wendorff and Berger found the maximal rate of ATP hydrolysis by MmT6, using the PK/LDH system, to be ~3 ATP min[-1] during relaxation of negatively supercoiled plasmid [62]. Here the ATPase rate was determined to be ~5 ± 0.9 ATP min[-1],

using a radioactive ATPase assay (see Chapter 3). If it is assumed that a single strand-passage event requires the hydrolysis of two ATP molecules, both values for the MmT6 ATPase rate correspond well with a rate of ~1.5 - 3.5 strand passage events min^{-1} on negative supercoils, between 0.25- 2 nM MmT6. As discussed [62], this rate is not only a fraction of the relaxation rate of type IIA topos, but it also seems unlikely that MmT6 could support cellular DNA metabolism as a DNA relaxase in *M. mazei* (replication rate for M. mazei calculalted in Wendorff *et al.,* 2020 [62]).
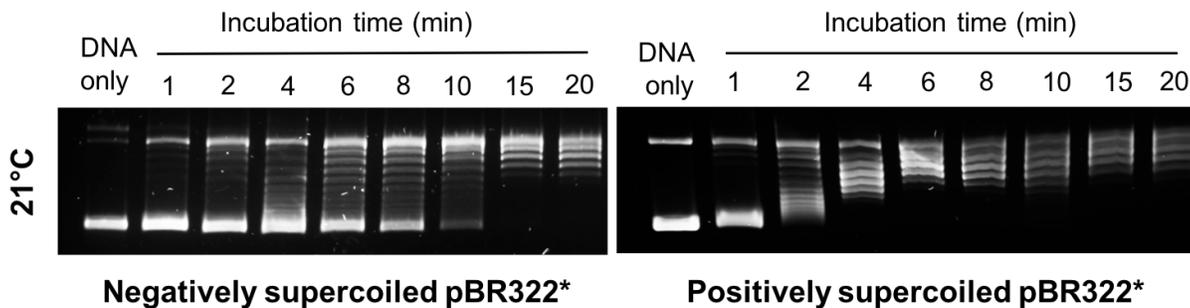
Figure 4.10: Agarose gel-based supercoil relaxation time course. Negatively or positively supercoiled pBR322* was incubated at 21°C, with 20 nM topo VI, the reaction was halted at consecutive time points using 50 mM EDTA. Samples were run on a 1% (w/v) native agarose gel for 15 hr at ~2 Vcm^{-1}, stained with 0.5 μg/mL ethidium bromide and imaged under UV illumination.

The processivity of type II topo activity, i.e. the number of T-segments passed during a single G-segment binding event, can vary depending on the enzyme being assayed and the DNA substrate present. For MmT6, each strand-passage event is evident in the trace, as an abrupt DNA extension change of ~110 nm (at 0.4 pN), which is expected for the resolution of a single DNA crossing (Figures 4.8 and 4.9A). This provides preliminary evidence of distributivity, as many seconds to minutes elapsed between strand passage events, indicating that MmT6 binds and resolves a single crossing, before disengaging the G-segment and rebinding the next crossing. With the magnetic tweezers affording this level of detail, this analysis can be taken further, extracting each dwell time between strand-passage events and plotting them as a function of the number of DNA crossings within the plectoneme (Figure 4.11). In the average relaxation rate analysis, the automatic recoiling of the DNA was initiated 2-3 DNA crossings from full relaxation as it was evident early on that MmT6 struggled to fully relax

plectonemes. Using 750 pM MmT6, a concentration determined to be effective for both positive and negative plectoneme relaxation, MmT6 was left to fully relax the DNA (Figure 4.11A and B). This produced a characteristic trace shape with the dwell time duration dramatically increasing as the DNA approached a fully relaxed state. Plotting these dwell times as a function of the level of DNA supercoiling clearly demonstrated that the rate of reaction decreased with decreasing DNA crossing number, which fit well to an inverse relationship, where the time taken to relax the crossing was inversely proportional to the number of crossings present (Figure 4.11C).
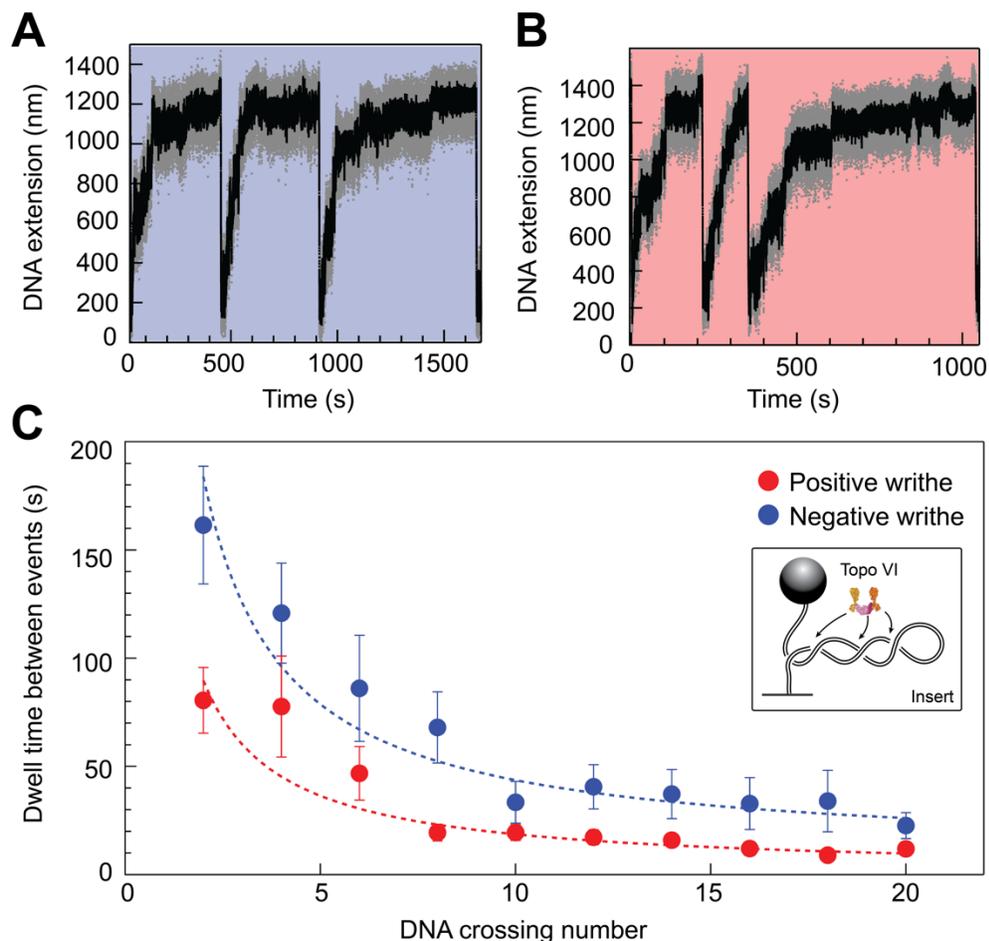
Figure 4.11: Topo VI relaxation rate is dependent on the level of DNA supercoiling. A: An example trace of 750 pM topo VI fully relaxing 20 negative supercoils in a 5 kb DNA duplex under 0.4 pN at 21°C. DNA extension is plotted against time, with the dramatic extension decrease corresponding to manual rotation of the magnets causing rapid DNA supercoiling. The slower DNA extension increase corresponds to topo VI-dependent relaxation activity. B: Same as in A, aside from the DNA being positively supercoiled. C: Dwell time (± SEM) between topo VI-dependent strand-passage events, plotted against the level of DNA supercoiling. Data were fitted to an inverse function, where the time taken for topo VI to perform a strand-passage reaction is

inversely proportional to the number of DNA crossings present in the substrate. Inset illustrates MmT6 binding behaviour, based on the data, as a highly distributive crossing sensor, with each DNA crossing within the plectoneme constituting a binding position for topo VI. Data analysed in Igor Pro 7, using a *T*-test based method, first described in [227].

This further supports MmT6's distributive nature in supercoil relaxation, but also cements the idea, described earlier [62], that this enzyme recognises and binds DNA crossings as a primary substrate, making each crossing along the plectoneme a potential site of activity. This is the converse of topo IV plectoneme binding behaviour, which was found to bind at a single site, likely at the end loop [120]. This DNA crossing binding behaviour of MmT6 is further supported by the MmT6 ATPase and DNA binding assays described in Chapter 3. The ATPase rate was not only dramatically stimulated in the presence of DNA writhe but also ATPase activity was tightly coupled to the presence of DNA writhe, with ATP hydrolysis decreasing once the DNA reached a relaxed state. The DNA binding of MmT6 in the absence of ATP, assayed using a nitrocellulose membrane capture technique, also indicateed that with increasing levels of supercoiling there is increased MmT6 binding.

4.2.3 Effect of Force on MmT6 Supercoil Relaxation

Alongside shedding light on the behaviour of MmT6, the enzyme titration also provided a useful guide for determining the most appropriate enzyme concentration to use in further supercoil relaxation assays. Based on both the plectoneme relaxation time and the unambiguity of the trace (i.e. being able to clearly distinguish each strand-passage event), the MmT6 concentration of 750 pM was chosen to assay the effect of force on the supercoil relaxation activity. Whilst a constant magnetic force on the DNA could cause tensile-related effects on topo activity, being able to control this variable is an important asset of the MT. It can facilitate the deceleration of reaction steps, and allow the elucidation of previously uncapturable events. This proved invaluable in the study of *E. coli* topo I and III kinetics, in which the gate opening dynamics could be measured for the first time [224].
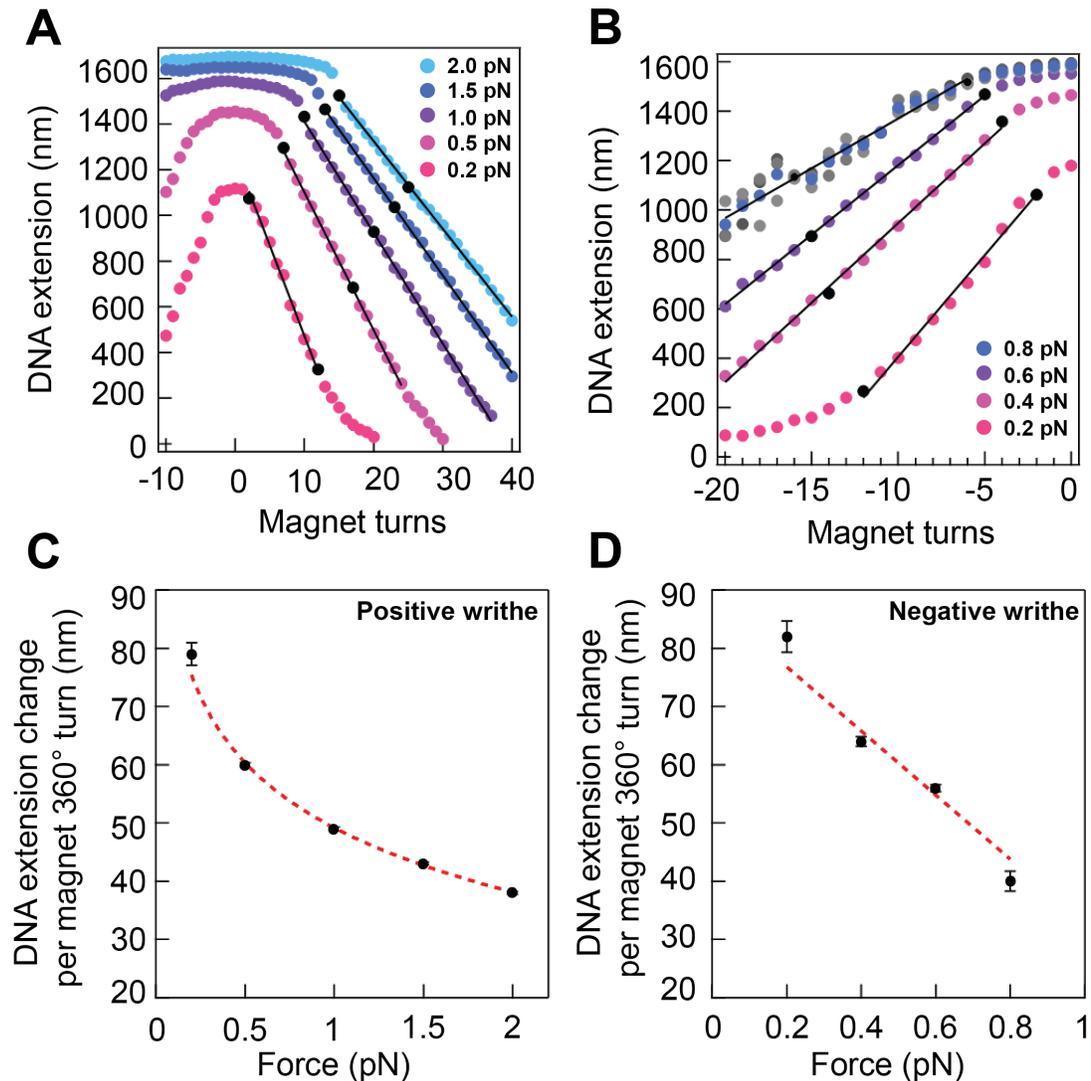
Figure 4.12: DNA extension as a function of magnet turns for different forces acting on a DNA plectoneme. A: Positive supercoiling at 0.2 (magenta), 0.5 (pink), 1.0 (purple), 1.5 (dark blue) and 2.0 pN (light blue). The first black circle is the buckling point in which writhe is formed and the second is 10 magnet rotations later, producing 10 supercoils. The black line is a linear fit (y=bx+a), the gradient of which is the DNA extension change per magnet turn, plotted in C. B: Negative supercoiling at 0.2 (magenta), 0.4 (pink), 0.6 (purple), and 0.8 pN (dark blue). The black circles denote the same as in A, with the DNA extension change per magnet turn being plotted in D. C: DNA extension change ± SE per magnet rotation as a function of force, for positively supercoiled DNA. The data was heuristically fitted to a power trendline (y=ax^b) (red dashed line). D: DNA extension change ± SE per magnet rotation as a function of force, for negatively supercoiled DNA. The data was heuristically fit to a linear trendline (y=bx+a) (red dashed line).

At low force the DNA extension as a function of magnet turns is symmetrical whether positive or negative turns are added. However, as described above, high force promotes DNA melting rather than negative writhe formation for negative supercoiling (Figure 4.7) [316]. This

means that force-dependent effects on MmT6 positive supercoil relaxation can be assayed to a higher degree of force compared to negative. As force is increased, there are two important aspects which change and therefore require consideration when conducting assays and comparing data. These are the magnet turns at which the DNA buckles, converting twist to writhe, and the extension change per magnet turn. This information was determined by collecting data on the DNA extension as a function of magnet turns for each force used (Figure 4.12). The force regime for positive supercoils was 0.2, 0.5, 1, 1.5 and 2 pN (Figure 4.12A and C), whilst for negative supercoils it was 0.2, 0.4, 0.6 and 0.8 pN (Figure 4.12 B and D). For each force the number of turns at which the DNA buckled to form a plectoneme, and the DNA extension change per turn, was different. For example, clockwise magnet rotation at 0.2 pN caused the DNA to buckle at turn 3 and each subsequent turn corresponded to a change in DNA extension of ~79 nm (Figure 4.12A and C). However, at 1.5 pN, the DNA buckled at ~13 turns, and the DNA extension change per turn was only ~43 nm (Figure 4.12A and C). This is information is vital to both conducting the experiments and analysing the data. Regardless of force, the same number of DNA crossings in the plectoneme (i.e. writhe) must be input, so knowledge of the buckling point facilitated this. Data analysis required knowledge of the extension change per linking number expected as the DNA is relaxed.

Once all the necessary calibration data was obtained (Figure 4.12), the average MmT6 relaxation rate in the presence of a plectoneme with a writhe of 10 at each force was analysed. The outcome can be seen in Figure 4.13. The average relaxation rate on positive supercoils exhibited a biphasic character, with the rate increasing from very low (0.2 pN) to moderate force (0.5-1 pN), and then decreasing once more as the force increased further (Figure 4.13A). This is partially reflected in the negative supercoil data, however as the force cannot be varied to the same extent without causing DNA melting, the data extracted is limited (Figure 4.13B). Nevertheless, this behaviour suggested that either a single rate-limiting step was being changed in two ways, or that increasing the force changed two different steps. The true meaning of this data is still unclear however a possible model is that at low force the rate is decreased due to increased supercoil movement, which may hinder MmT6 binding efficiency.

160

As the force increased, supercoil diffusion decreased and this may allow more efficient MmT6 binding. Indeed, it has been previously demonstrated that increasing the force reduced the diffusion rate of the plectoneme [317]. The subsequent rate decrease at high force may be a function of increasing twist. As force is increased, more twist is required before the DNA buckles to form writhe. For supercoiled circular DNA, twist and writhe are partitioned into ~25% twist and ~75% writhe [6]. For the MT setup, the DNA tension provided by the magnetic force results in twist remaining constant once the DNA buckles to form writhe, with each rotation post-buckling entirely accommodated by increasing writhe [158, 223]. As the force increases, more twist is required before buckling occurs which results in plectoneme formation with a higher twist density. Increased twist could hinder a variety of stages during the MmT6 reaction such as DNA binding and cleavage, therefore reducing the rate. When the force-dependent supercoil relaxation activities of *E. coli* DNA gyrase (gyrase) were assayed, a similar biphasic result was reported for the relaxation of positive writhe, in which gyrase activity increased when force was increased from 0.2 to 1 pN [318]. This was predicted to be due to torque-dependent acceleration of the DNA-wrapping, which is subsequently inhibited as torque increased further. As MmT6 cannot wrap DNA, this can't explain the data reported here.
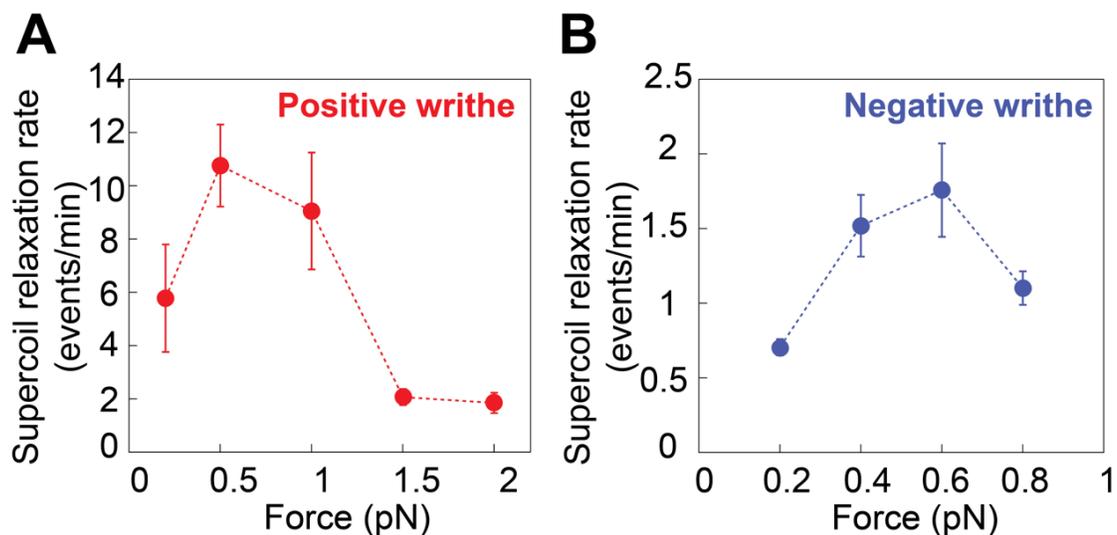
Figure 4.13: The effect of force on MmT6 supercoil relaxation activity. A: The average rate of positive supercoil relaxation ± SE by 750 pM MmT6 as a function of force, using 1 mM ATP at 21°C. B: The average rate of negative supercoil relaxation ± SE by 750 pM MmT6 as a function of force, using 1 mM ATP at 21 °C.

4.2.4 MmT6 Unlinking Activity Using DNA Braids

As demonstrated in section 4.2.2, the distributive nature of MmT6 relaxation activity, as assayed using MT, was evident independently of the supercoil chirality, however positive crossings were still relaxed 2-3-fold faster than negative. Interestingly, a few type IIA topoisomerases including *E. coli* topo IV and human topo IIα also demonstrated preferential relaxation of positive writhe [120, 158]. The detailed basis for chiral preference likely varies amongst type II topos, however one commonality was that the enzymes preferred a particular DNA crossing geometry for either binding or T-segment capture. In order to explore how the geometry of the DNA crossings and the presence of twist, play a role in chirality sensing by MmT6, the MT-based DNA braiding assay was utilised. In this assay, rather than tethering a single torsionally constrained duplex, two unconstrained DNA duplexes were attached to a single magnetic bead, which, upon rotation of the magnets, wrapped the duplexes around one another (see Figure 4.2B), creating writhe without changing the twist [257]. In contrast to the supercoiled substrate, DNA writhe is created directly in the braiding system (rather than via the conversion of twist to writhe), therefore the magnets are rotated clockwise for positive crossings and anticlockwise for negative. This braided DNA substrate is far more akin to a catenated substrate than supercoiled, and allows the exploration of how writhe affects enzymatic behaviour in the absence of changing twist. It is worth noting, that 20 magnet rotations (at low force) generate a plectoneme which requires 10 relaxation events to become fully relaxed; whilst for braiding 10 magnet rotations produces a braid which is fully unlinked via 10 strand-passage events. This arises as a consequence of supercoiled DNA being intramolecular writhe and braided DNA being intermolecular writhe.

As with supercoiled DNA, the braided substrate required calibration in terms of DNA extension as a function of magnet turns (hat curve), before the addition of MmT6. The hat curve of a bead tethered by two DNA duplexes has a highly characteristic shape, and must be confirmed prior to data collection (Figure 4.14). A braid hat curve should be triphasic, beginning with a dramatic decrease in DNA extension as the first turn is executed, which is followed by a

shallower extension decrease as the braid is formed. At ~20 turns (for 0.5 pN), the braid buckled and wrapped around itself forming a four-duplex supercoiled braid structure with each magnet turn producing a larger DNA extension decrease than for braiding. In general, assays are performed using the single crossing and braid substrates only, as activity on the supercoiled braid structure is not quantifiable, due to difficulty in characterisation as a result of its topological complexity.
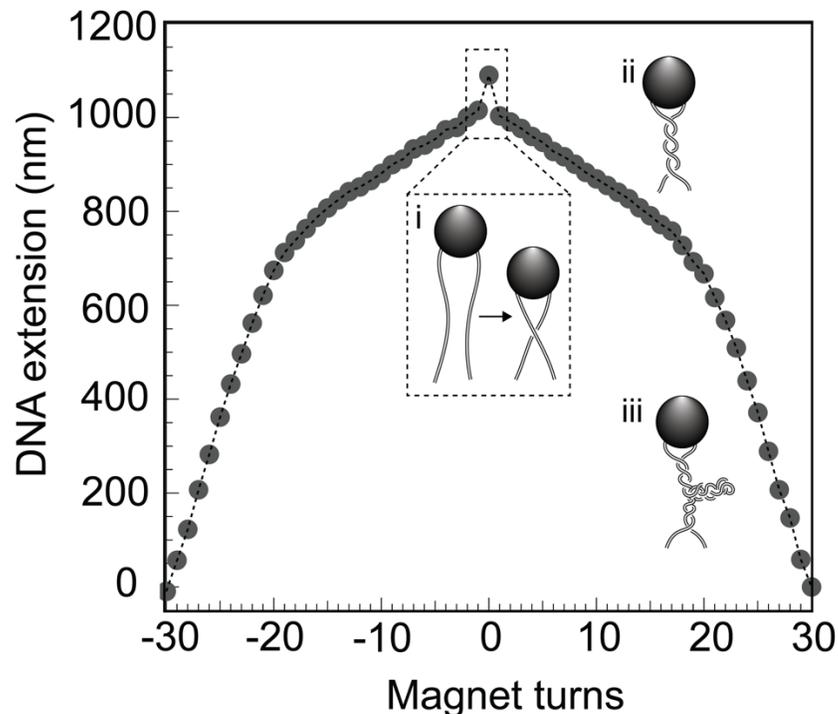
Figure 4.14: Braided DNA tether hat curve. DNA extension was plotted as a function of magnet turns. The braid curve is symmetrical, with the first full positive or negative turn resulting in a dramatic extension decrease (see i, in dashed box). Subsequent turns cause the DNA extension to decrease with a shallower gradient as the DNA braid is formed (see ii). Eventually the braid buckles upon itself and produces a supercoiled braid structure (see iii), with each magnet turn resulting in a larger DNA extension decrease than for braiding. Data collected under 0.5pN, at 21°C, using 5 kb DNA tethers.

In the case of MmT6, the braided DNA substrate had a surprising effect on MmT6 activity. In Figure 4.15A, an example trace of 50 pM MmT6 activity (0.5 pN), shows how a positive braid is relaxed in two rapid bursts which corresponded to 5 strand-passage events at rates of 0.6 to 1.25 strand-passage events sec^{-1}. In Figure 4.15B, using 100 pM MmT6 on a

negative braid, the situation was similar with rapid bursts of 2 or more consecutive strand-passage events also occurring. It was evident that the average rate of braid unlinking had increased ~5-fold above that of supercoil relaxation (e.g. 500 pM MmT6 gave an average rate of 3.5 ± 0.8 strand-passage events/min on positively supercoiled DNA, and 18.9 ± 2.3 strand-passage events/min on positive braids) (Figure 4.16A). Moreover, MmT6 was capable of sustaining robust activity at concentrations ~5-fold lower than was achievable in supercoil relaxation, with a mild processive behaviour (e.g. passing a consecutive T-segment whilst remaining bound to the initial G-segment). Rates for MmT6 processive activity (dwell times between events not included in the average) (Figure 4.16B) are comparable to the rates of other type IIA topos, for instance *E. coli* DNA gyrase, measured using a rotor bead tracking technique, showed 1 strand passage event sec^{-1} [319], and *E. coli* topo IV has been documented to have a rate as high as 2.5 strand passage events sec^{-1} using a positively supercoiled magnetic tweezers assay [119].

The average unlinking rates by topo VI at concentrations ranging from 50 to 900 pM, were well described by a Michaelis-Menten-like function, providing V_{max} values of 21.4 ± 0.5 and 10.9 ± 1.3 strand passage events min^{-1}, and $K_{d,app}$ to be 66.5 ± 7.3 and 164.2 ± 72.4 pM, for the unlinking of positive and negative braids respectively (Figure 4.16A). The chiral preference remains, with the V_{max} being 2-fold higher, and the $K_{d,app}$ being 2.5-fold higher, for positive braid-unlinking over negative. This further suggests the chiral selection originates from a DNA-crossing geometry preference for MmT6 which occurs at a stage between binding and strand-passage. Based on DNA binding experiments (Chapter 3, Section 3.2.6), the chiral selection may be occurring after initial G-segment binding as MmT6 exhibits tighter binding to negative writhe than it does positive, suggesting that perhaps the cleavage of the G-segment and/or passage of the T-segment through the break, is sensitive to the crossing angle. However, as this magnetic tweezers data doesn't delineate between binding, cleavage and strand passage, exactly when the chiral selection takes place isn't clear. Moreover, the binding assays don't discriminate between DNA crossing binding or G-segment alone, so it is plausible that MmT6 binds non-productively to negative writhe more than it does on positive.
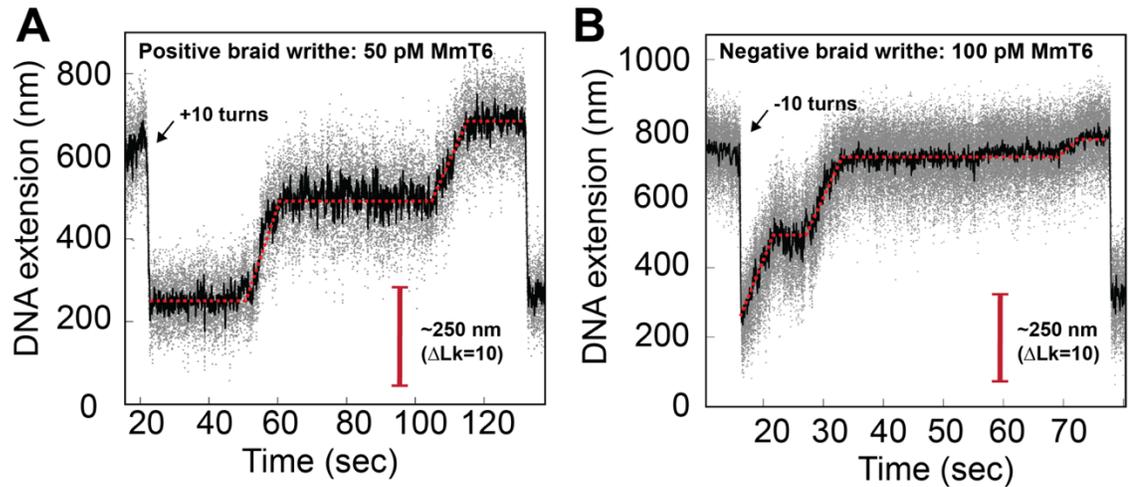
Figure 4.15: MmT6-dependent braid unlinking examples. A: Positive braid writhe unlinked by 50 pM MmT6, with 1 mM ATP, under 0.5 pN at 21°C. The black arrow indicates where 10 clockwise (+) magnet rotations were initiated. The light grey points are the raw data with the black representing the smoothed data, achieved using the Savitzky–Golay filter in Igor Pro 7. The red dashed line indicates periods of inactivity (horizontal portions) and processive bursts of multiple strand passage events (diagonal portions). The solid red line is a scale bar, highlighting how the unlinking activity is occurring in steps larger than a ΔLk of 2. B: Negative braid writhe unlinked by 100 pM MmT6, with 1 mM ATP, under 0.5 pN at 21°C. All annotations are the same as in A, aside from the black arrow which denotes the addition of 10 anticlockwise (-) rotations. Kinetic parameters measured can be seen described in Figure 2.5.

The MmT6 V_{max} for unlinking was around 3-fold higher than for supercoil relaxation and largely due to bursts of two or more unlinking events in rapid succession. This rapid unlinking could be due to either simultaneous binding and unlinking by multiple enzymes or a processive unlinking by a single enzyme, with these scenarios distinguishable based on the enzyme concentration dependence. Extracting the rates of the processive bursts in the absence of the dwell times, as well as the size of these bursts (termed step size and measured in number of strand-passage events), the values remain fairly constant across large changes in MmT6 concentration suggesting that this burst activity is likely due to processive unlinking by a single enzyme (Figure 4.16B and C). On the other hand, the duration of the dwell times between events decreases with increasing MmT6 concentration, further indicating that a single enzyme drives a burst of multiple unlinking events (Figure 4.16D). This striking outcome strongly suggests that a unique geometrical property manifested in DNA braids, but absent/rare in

Chapter 4: Single-Molecule Results

supercoils, allows stable MmT6 binding to the DNA substrate during multiple strand-passage cycles.
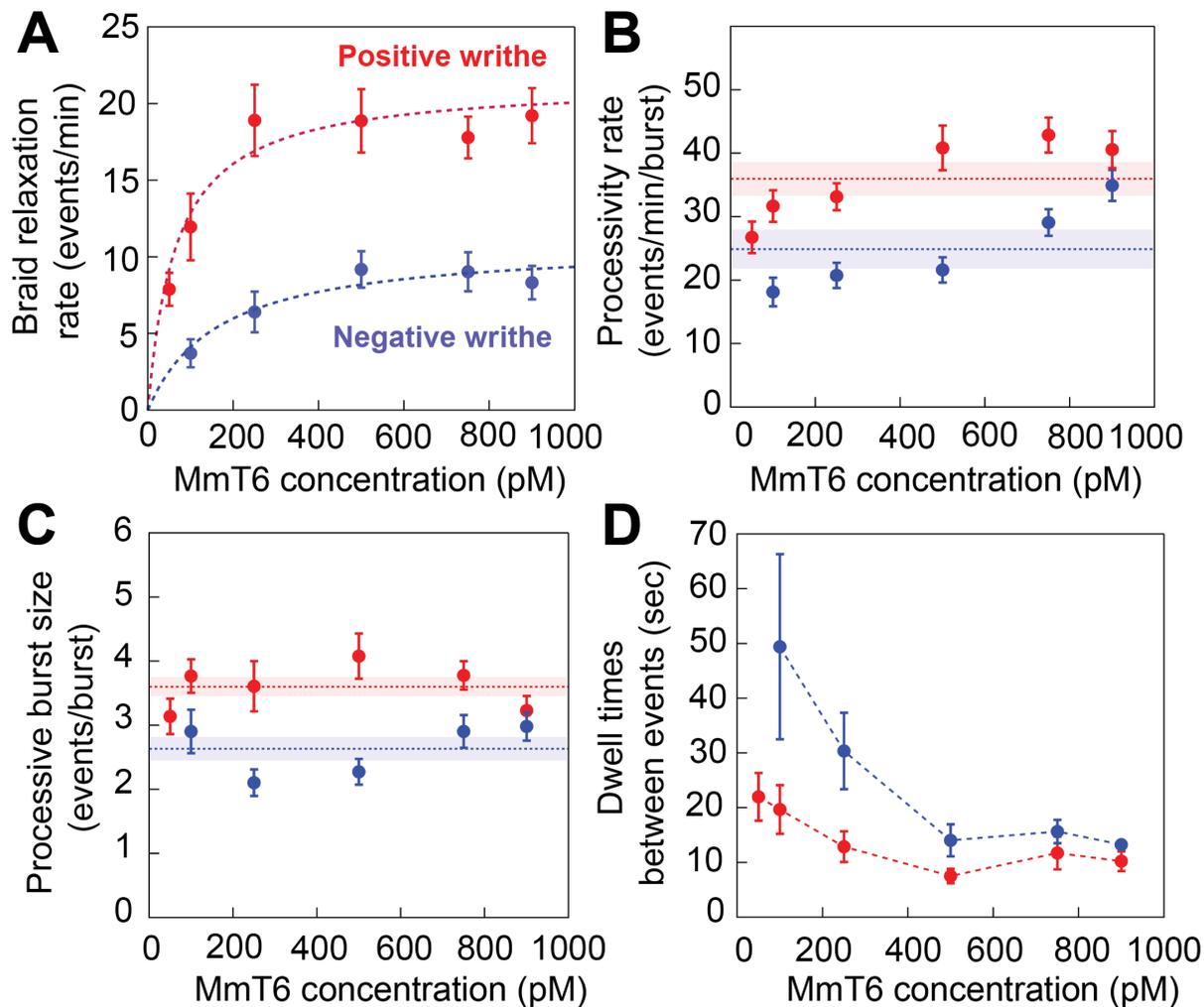
Figure 4.16: MmT6-dependent braid unlinking. A: The average rate of topo VI braid unlinking activity (± SEM), both positive and negative chirality, measured as the number of strand-passage events/min and plotted as a function of topo VI concentration. Data were fit (dashed lines) to a Michaelis-Menten-like function ($V_0 = \frac{V_{max}[E]}{K_{d,app}+[E]}$). **B:** The processive burst rate of topo VI (± SEM), measured as the number of strand-passage events/min within each burst, and plotted as a function of topo VI concentration. A burst is defined as rapid topo VI activity corresponding to the passage of two or more consecutive T-segments, in which individual strand-passage events cannot be discerned. **C:** The processive burst size of topo VI (± SEM), measured as the number of strand-passage events/burst, and plotted as a function of topo VI concentration. **D:** The dwell times between processive bursts of topo VI activity, plotted as a function of topo VI concentration. A dwell time is defined as a period of time in which the DNA extension remains constant, reflecting lack of topo VI-dependent braid unlinking activity. In Figures 4A-D, data was

collected under 0.5 pN force, at 21°C, using 1 mM ATP, with topo VI activity on positive DNA braids being highlighted in red, and in blue for negative DNA braids.

The DNA crossing angle of the G- and T-segment is considered an important factor in the DNA substrate preference of type II topos. Recent work from the Neuman lab (NHLBI, NIH, Bethesda, USA) (data not shown) suggested all angles within a braid are wider than those found in supercoils, dependent on DNA tension, spacing and temperature, something that was also documented in Charvin *et al.* [223]. More work is required to fully understand DNA braid geometry, however it is likely that this change in DNA-crossing angle stimulated MmT6 activity as MmT6 transitioned from being perfectly distributive on supercoils to mildly processive on braids. A potential explanation is that this is a function of DNA-crossing angle, with the diffusion of the T-segment into the MmT6 cavity at the appropriate juxtaposition to the G-segment, occurring rapidly for braids, and rarely for supercoils. MmT6 behaviour may also be affected by twist, since braided DNA is unable to store any perturbations in twist, only writhe. In supercoiled molecules the twist is changed dramatically from the relaxed state, depending on the force. Therefore, MmT6 processivity on braided DNA could be due to more stable binding, or faster cleavage/re-ligation, in the absence of changing twist. In Section 4.2.6, the effect of DNA geometry on MmT6 activity will assayed directly.
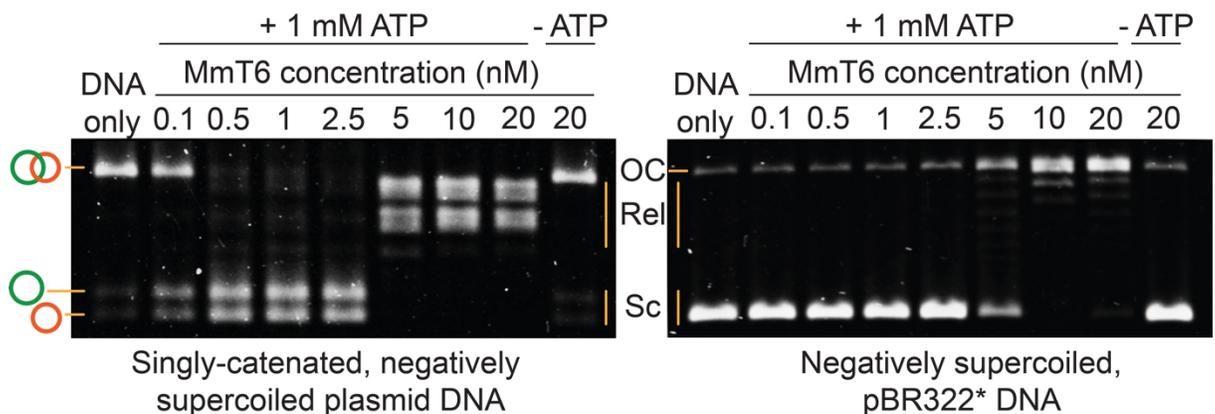
Figure 4.17: Agarose gel based assay of DNA decatenation and relaxation by MmT6. On the left, a singly catenated, negatively supercoiled plasmid substrate is decatenated by 0.1-20 nM MmT6. The catenated plasmids vary in size and when decatenated, can be seen as two bands that migrate further than the catenated substrate. As they are also negatively supercoiled, the

Chapter 4: Single-Molecule Results

relaxation of the plasmids can be seen at MmT6 concentrations ~10 fold higher than when full decatenation is seen. This is further corroborated by a relaxation assay performed using negatively supercoiled pBR322* (right-hand gel), where relaxation activity is not detected until ~10 fold the MmT6 concentration necessary for decatenation. OC: open circular, Rel: relaxed and Sc: supercoiled. Both reactions were incubated for 30 min at 37°C.

As mentioned above, the braided DNA substrate is often likened to a catenated substrate and topo-dependent unlinking activity is considered a proxy for exploring decatenation activity. This parallel between braided and catenated DNA arises due to the braid also forming intermolecular writhe with DNA crossing geometries more similar to catenanes than supercoils [223]. Therefore, the data presented here may indicate that MmT6 is more efficient at decatenation than supercoil relaxation, *in vivo*, in line with its predicted role. However, decatenation has been a more challenging activity to quantify using traditional gel-based topology assays, due to the reliance on kinetoplast DNA (kDNA). This substrate originates from the mitochondria of trypanosomes and is formed of an extended network of interlinked DNA plasmids of two sizes called maxicircles (20-40 kb) and minicircles (0.5-1kb) [256]. As discussed in Chapter 3, its use can be a qualitative measure of decatenation activity, with plasmids released from the network as a result of topo-dependent activity, which can be visualised using gel electrophoresis. However, with the inability to quantify or control how many plasmids are interlinked and to what degree, comparing a topo-dependent decatenation activity to supercoil relaxation using a gel-based approach is limited. Recently, a DNA substrate was developed by Nidda Waraich (JIC, Norwich, UK; Waraich et al. BioTechniques, in press, 2020. See Section 3.2.4.2), in which two plasmids of different sizes are singly-linked. This facilitated a far more quantitative approach to studying decatenation activity and allowed the comparison of MmT6-dpendent decatenation activity to supercoil relaxation. As seen in Figure 4.17, MmT6 activity during decatenation was detectable at MmT6 concentrations 10-fold lower than for supercoil relaxation (0.1-0.5 and 5 nM MmT6, respectively). This further supports the MT-based data, which also demonstrated robust activity in braid unlinking at concentration 5-10-fold lower than plectoneme relaxation (25-50 and 250-500 pM, respectively). Again, this strongly

suggested that there was a unique feature of both braids and catenated DNA molecules that enhanced MmT6 activity.

4.2.5 Effect of Force on MmT6 Unlinking Activity

As was achieved for supercoil relaxation activity (section 4.2.3), varying the tensile force on the braid was conducted to illuminate further nuances of MmT6 behaviour. In the case of plectonemes, the force is perpendicular to the plectoneme axis, which makes parsing the effects of force on the supercoils and MmT6 activity more theoretically challenging as it's unclear whether the plectoneme itself experiences any of this force directly, or if the effects of force arise from the accompanying changes in twist. In the case of a DNA braid, the magnetic force runs parallel to the DNA, meaning force has a direct effect on the DNA crossings, and likely the activity of MmT6. The calibration for this form of analysis was simpler for DNA braids, than it was for supercoils. Due to the absence of changing the twist, DNA braids remain chirally symmetrical at high force, unlike the DNA melting which occurs during negative supercoil formation. This results in the ability to assay the effect of force on MmT6-dependent negative braid unlinking to the same degree as positive, in theory. In reality, as explored below, MmT6 was more sensitive to force during negative braid unlinking and did not display activity to the same force maximum as for positive braid unlinking. In supercoils, knowledge of both the buckling point (where twist is converted to writhe), and the DNA extension change per magnet rotation, as a function of force, were necessary for accurate assay execution and data analysis. However, for braids, only the latter piece of information is relevant as writhe is input immediately upon magnet rotation, independent of the force.
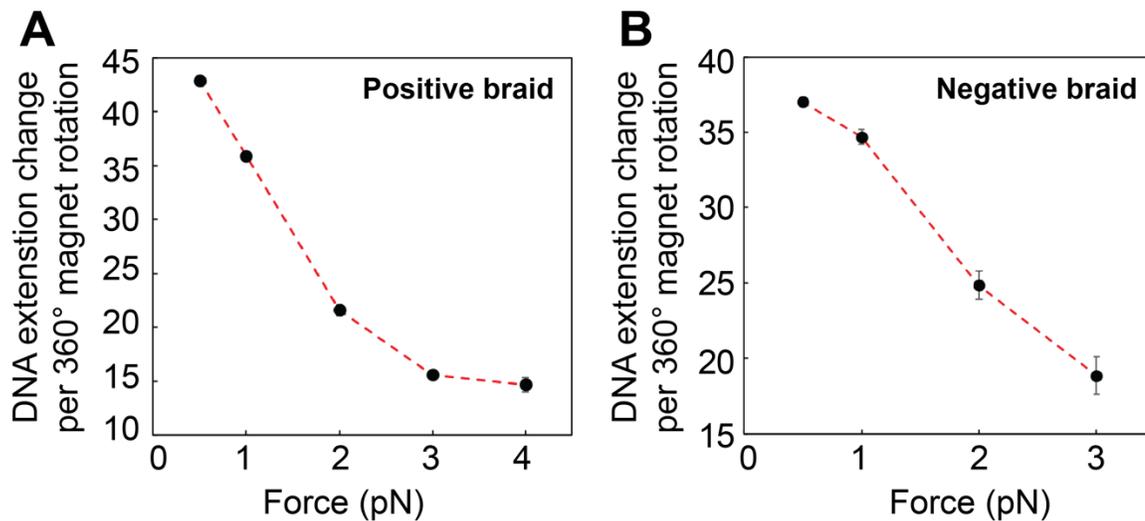
Figure 4.18: DNA extension change per magnet rotation, for different forces acting on a DNA braid. A: The DNA extension change per magnet rotation for a positive DNA braid under 0.5, 1 ,2, 3 and 4 pN of force ± SE. Data points black circles, joined by a red dashed line. B: The DNA extension change per magnet rotation for a negative DNA braid under 0.5, 1 ,2 and 3 pN of force ± SE. Data points black circles, joined by a red dashed line. Please note that for most data points in A and B, the error bars are too small to be visualised.

The force regime for positive braid unlinking was 0.5, 1, 2, 3 and 4 pN, and for negative braid unlinking it was 0.5, 1, 2 and 3 pN. Calibration data for the DNA extension change per magnet rotation can be seen in figure 4.17. The values for positive (4.18A) and negative (4.18B) braids are similar, but as mentioned above, should be symmetrical. The reason this is not the case for this data is entirely due to different braids being utilised for positive verses negative braid unlinking. Depending on the duplex spacing of the DNA braid, the extension change per magnet rotation can be subtly different. Whilst this shouldn't affect the results dramatically it is worth noting as DNA spacing modulates the angle distributions of the DNA crossings, and as explored in future sections, MmT6 is sensitive to DNA crossing angles.

Varying the force in the magnetic tweezers braiding system (Figure 4.19) revealed a similar effect on MmT6 activity to that of supercoil relaxation. For positive-braid writhe, increasing the force from 0.5 to 1 pN resulted in a moderate rate increase, whereas for negative-braid writhe the rate at these two forces was the same (4.19A). This is reminiscent of the initial rate increase in supercoil relaxation when force was increased from 0.2 to 0.5 pN (Figure 4.13), with the potential explanation for this being the same, e.g. a low-level increase

in force reduces the diffusion rate of DNA crossings, enhancing MmT6 DNA-crossing binding efficiency. This is then followed by a significant drop in unlinking rate for both positive and negative braids as force is increased further. Negative-braid unlinking was more sensitive to force increases, with no activity detected at 4 pN. MmT6 was still able to unlink positive braids at 4 pN, however activity was severely compromised. This decrease in activity was determined to be largely a consequence of an increase in dwell times between unlinking events as force increased (Figure 4.19B). Both processivity and step size were largely unaffected, however there was a mild decrease in both for the unlinking of positive braids with high force reducing both to the level measured for negative braid unlinking (Figures 4.19 C and D). This force-dependent rate reduction could be the result of two different factors. It may be due to a force-dependent step in the catalytic cycle that becomes rate limiting, for example, transferring a DNA strand against DNA tension.
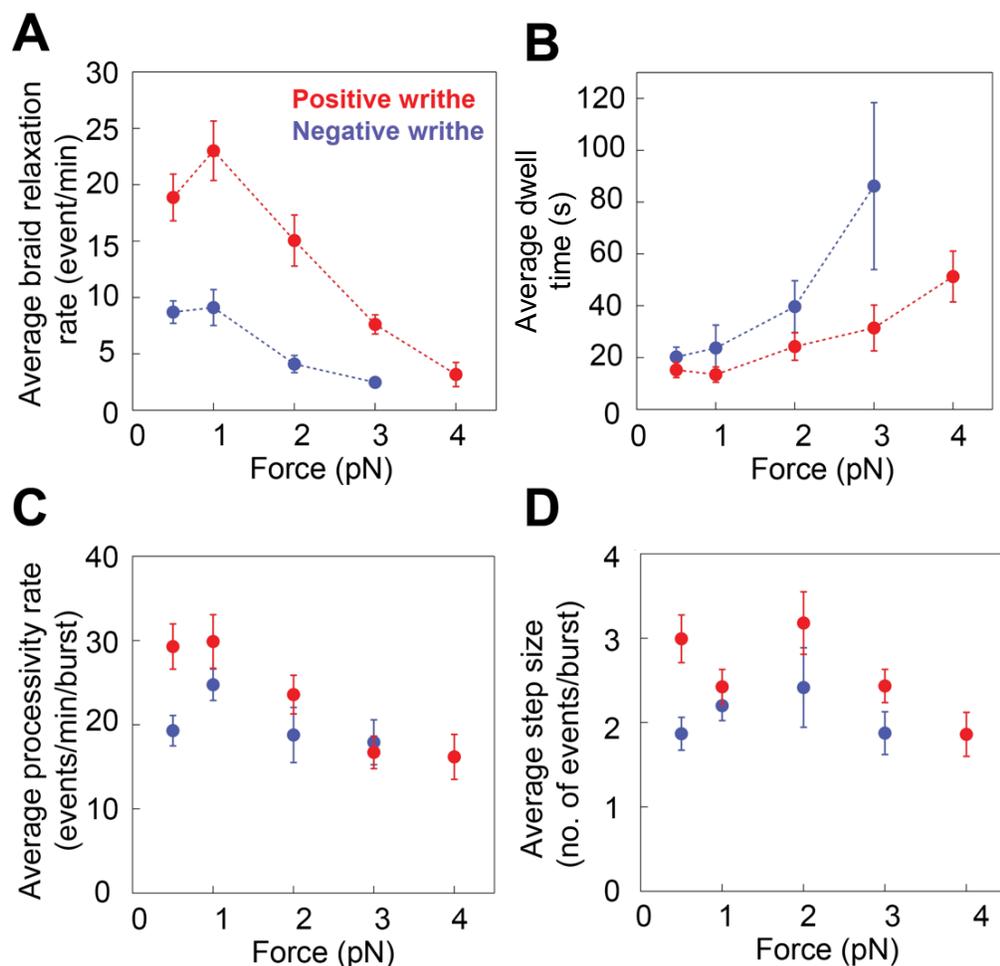
Figure 4.19: The effect of force on MmT6 unlinking activity. A: The average rate of positive (red) and negative (blue) braid unlinking (strand-passage events/min) (± SEM), plotted against force.

B: The average dwell times (s) between unlinking events for positive (red) and negative (blue) braids (± SEM), plotted against force. C: The average processivity rate (strand-passage events/min/burst) for positive (red) and negative (blue) braid unlinking (± SEM), plotted against force. D: The average step size for positive (red) and negative (blue) braid unlinking bursts (± SEM), plotted against force. Assays in Figures A-D conducted using 500 pM topo VI and 1 mM ATP, at 21°C.

However, it is more likely to be due to the increased force causing a change in the DNA crossing angle distribution disfavouring MmT6 activity, possibly affecting both binding time and strand-passage efficiency. A similar form of analysis was completed for *E. coli* DNA gyrase (gyrase), in which unlinking of positive and negative writhe was measured for different forces. Like MmT6, gyrase had a distinct preference for unlinking positive braid writhe, however, unlike MmT6, it was able to efficiently unlink positive braids at high force (5 pN), and unable to unlink negative braids under all force conditions assayed (0.2 – 5pN). These distinctions between MmT6 and gyrase are likely due to the way in which the enzymes engage with DNA, as wrapping of the DNA around the gyrase GyrA C-terminal domains, heavily favours interaction with positive DNA crossings, disfavouring negative. Whilst MmT6 demonstrates a mild chiral preference for positive crossings, it does not wrap DNA and therefore does not disfavour negative braid unlinking as strongly.

4.2.6 MmT6 DNA Crossing Geometry Preference

Based on the data reported here, for both supercoil relaxation and braid unlinking, MmT6 displayed a 2 to 3-fold rate enhancement during positive DNA crossing resolution, compared to negative. In combination with exploration into how force affects MmT6 activity, these results suggest that MmT6 may have a preference for a particular DNA crossing angle, that is; the angle in which the G and T-segment juxtapose, which is more common in positive crossing than negative. Due to the rate increase in the presence of positive writhe, this indicated that the angle preference was below 90° as acute angles are more common in positive DNA crossings [119, 120]. In order to explore MmT6's preferential DNA geometry in more detail, unlinking rates of a single crossing by MmT6 were measured, in which the crossing geometry

can be well-defined and therefore its effect on activity determined. A single crossing, in this context , describes one full 360° rotation of the magnetic bead, which creates an intertwined DNA substrate that is fully relaxed by type II topos in a single strand-passage event. This method was first described in [120] and applied to *E. coli* topo IV. Shown in Figure 4.14, when the braid goes from fully relaxed to a single crossing, achieved by one 360° magnet rotation, there is a distinct drop in DNA extension. This allowed straightforward measurement of the unlinking rate of a single crossing of either chirality. These measurements were conducted at high MmT6 concentration as they were complete in a single catalytic event, and this ensured T-segment binding, rather than G-segment binding, was rate limiting. This, when combined with Monte-Carlo simulation-derived crossing angle distributions, allows determination of the MmT6 crossing angle preference. This is facilitated due to single positive and negative crossings being identical in every respect, aside from the crossing angle distributions. An enzyme with a preference for a crossing angle below 90° will have increased activity on positive crossings, as is true for MmT6 and topo IV [119, 120], whereas if the enzyme binds preferentially to perfectly symmetric crossing angles (90°), there would be no difference between the rate of relaxation on positive and negative crossings, as seen for yeast topo II [120].
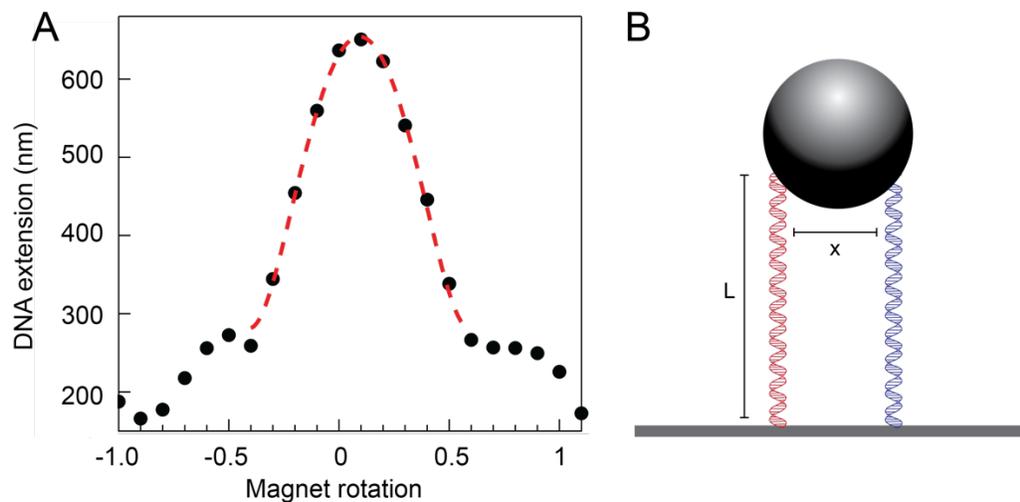
Figure 4.20: Braid tether calibration and geometric fit. A: The DNA extension of a 3 kb double tether, plot as a function of magnet rotations and fit to a geometric function [120]: $L = \sqrt{L_0^2 - 4e^2 sin^2(n\pi)} - r + \sqrt{r^2 - e^2}, |n| < 0.5,$ **where L is the measured DNA extension (dependent**

Chapter 4: Single-Molecule Results

on the force applied), L_0 is the maximum DNA extension (dependent on the force), e is half the separation distance between the DNA molecules, n is the number of magnet rotations, and r is the radius of the magnetic bead to which the tethers are bound. The first term in the expression describes the extension of a twisted swing, whilst the second two terms are a correction for the spherical bead, see [120] for a full description. The output from the fit in terms of DNA spacing are found in Table I, Appendix II. B: Geometric parameters of the DNA tethers, attained through the calibration and subsequent geometric fit (seen in Figure 7A), namely the length of the DNA tethers (L) and the separation distance between them (x, or $2e$). Along with the force and number of turns, this information is used in Monte Carlo (MC) and Brownian Dynamics (DB) simulations.

To begin, the braid being assayed requires precise calibration of the first crossing, meaning DNA extension is measured for every 10th of a magnet rotation. This data was then fitted to a geometric function, which allows the calculation of both the length and spacing of the DNA attached to the bead (see Figure 4.20). This information, along with the magnetic force applied to the bead, as DNA tension affects the DNA crossing angle distribution, is integral to executing DNA crossing angle simulations as a means of obtaining the angle distributions of the DNA crossing geometry being assayed. Once calibration is complete, data collection in the presence of MmT6 was undertaken, adding one full turn in one direction, allowing MmT6 to unlink the DNA crossing, before adding a full turn in the opposite direction and so on (Figure 4.21A and B). The times taken to relax either positive or negative crossings were plotted as histograms and fit with single exponentials to attain the characteristic unlinking times for the crossing geometry being assayed (Figure 4.21C). In agreement with results from DNA braids and supercoils, the characteristic unlinking time from the fitting on positive crossings (τ_L) for the braid assayed in Figure 5 was 8.4 ± 1.4 s, which was 4.7-fold faster than that of negative crossings (τ_R) (43.8 ± 6.9 s), suggesting that a positive crossing geometry facilitates topo VI activity through the higher probability of forming the preferred crossing angle (Figure 4.21C). For all data collected and analysed in this section see Appendix II Figures 1-3 and Table 1. To relate this experimentally derived data to a more precise value for the DNA crossing angle preference of MmT6, both Monte Carlo (MC) (Figure 4.22A) and Brownian Dynamics (BD) simulations were performed to determine the distribution of crossing angles formed for a given
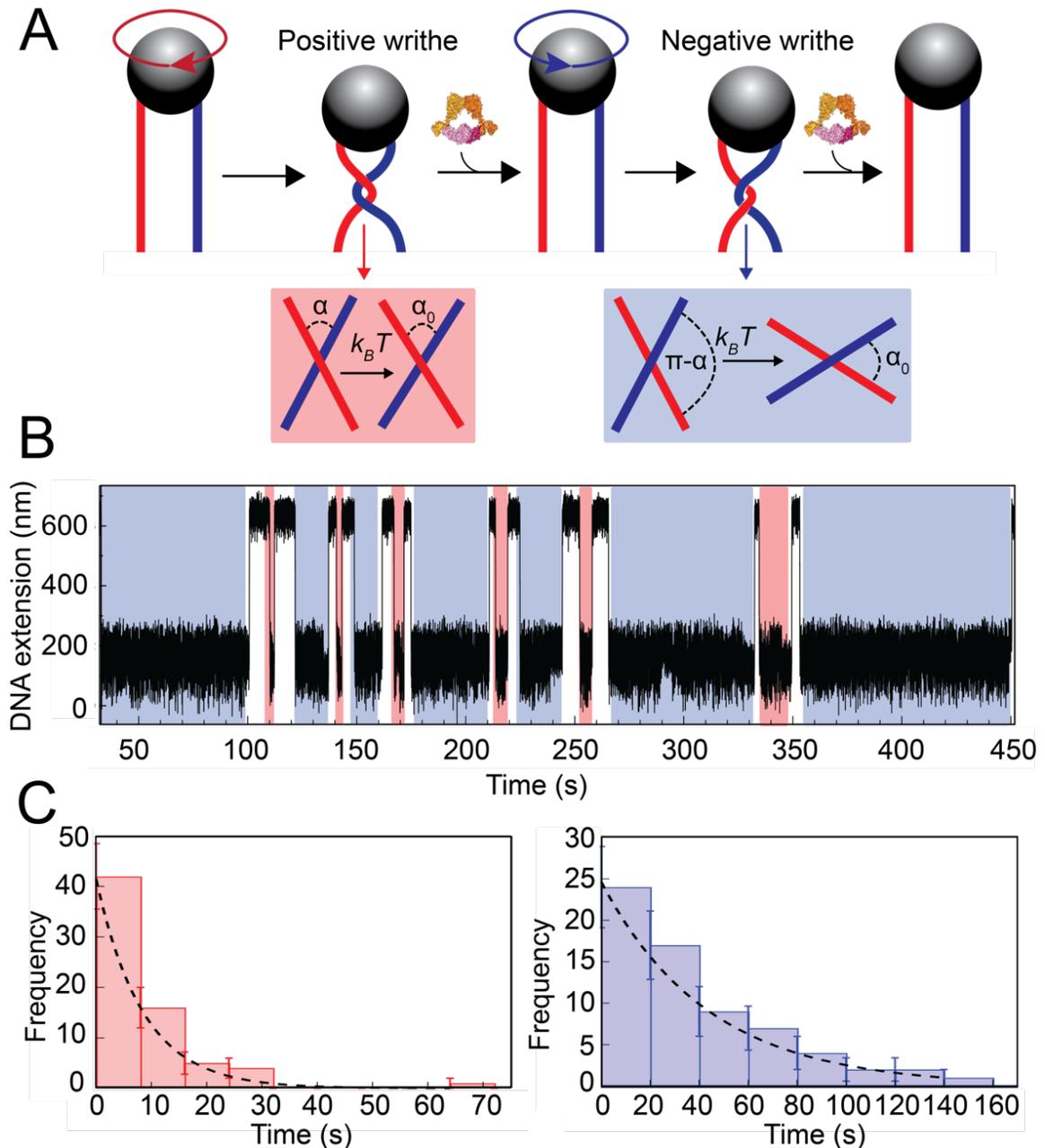
Figure 4.21: Experimental data collection and analysis of MmT6 single-crossing unlinking activity. A: Single crossing assay schematic with DNA crossing geometry for positive (left-handed, red) and negative (right-handed, blue) DNA writhe. One 360° clockwise magnet rotation imparts a positive crossing, which is unlinked by MmT6, followed by the generation and subsequent MmT6-dependent unlinking of a negative DNA crossing, formed by a 360° anticlockwise magnet rotation. The crossing angle is defined as the clockwise angle between the top and bottom DNA strands, and for positive crossings this is an acute angle (α), and obtuse for negative crossings (π-α). For MmT6, achieving the preferred angle (α_0<90) requires a smaller thermal fluctuation of positive crossings, therefore there is a higher probability of α_0 being achieved than for negative crossings. B: Single crossing unlinking data, collected for a braid formed from 3-kb DNA tethers, spaced 624 nm apart, at a force of 1 pN, using 900 pM MmT6 and 1 mM ATP. Positive crossings (red) were relaxed more rapidly than negative crossings (blue). C:

Chapter 4: Single-Molecule Results

Distributions of the MmT6-dependent unlinking times for negative (blue bars) and positive (red bars) crossings, of the data shown in B. The data (N=100) were fitted with single exponentials, P(t) = Aexp(-t/τ), giving characteristic unlinking times of τ_R= 43.8 ± 6.9 for negative crossings, and τ_L= 8.4 ± 1.4 for positive crossings, with a ratio of τ_L/τ_R = 0.19 ± 0.04.

DNA geometry and force. The BD simulations were executed by Parth Desai (NHLBI, NIH, Bethesda, USA). The crossing angle distribution were obtained by simulating the thermal fluctuations of two DNA molecules, the movements of which were dependent on DNA length, the tension applied to the DNA, temperature, and the spacing between the DNA duplexes. So, whilst the average crossing angle was not necessarily the preferred crossing angle (α_0), this fluctuation resulted in α_0 being formed at a certain frequency, which was related to the measured unlinking rate of a single crossing (assuming that achieving α_0 is rate limiting). The MC and BD simulations can predict how probable α_0 is under the given DNA crossing geometry for positive and negative crossings. If the assumptions that the time taken to unlink the crossing is dependent on a single rate limiting step, which is likely as the unlinking times are exponentially distributed, and that this step is dependent on the DNA crossing angle, then τ_L/τ_R is equal to the ratio of negative and positive crossing angle probabilities (Figure 4.22B). In other words, for the crossing geometry in Figures 4.21 and 4.22, the positive crossing is relaxed 4.7 times faster than the negative, and so using this ratio, the angle that appears 4.7 times more often in positive crossings than negative can be determined. This was repeated for fourteen different crossing geometries, with all resulting in extremely similar values for crossing angle preference, independent of the average or most probable crossing angle for that particular braid geometry (Figure 4.22C) or the force under which the data was collected (Figure 4.22D0.

Together these data indicated that the preferred crossing angle for MmT6 was 87.8° ± 0.4°, when using MC simulations, and 87.4° ± 0.4°, when using BD simulations. Supported by MC simulations done in Stone *et al.* [119], this angle preference can account for the consistent ~2-fold difference in rate between the positive and negative supercoil relaxation and braid unlinking assays, with 87.4-87.8° appearing in positive crossings ~2-fold more frequently. This

value is not dissimilar to that found for topo IV, attained using the same assay, of 85.5° [120].

However, one distinct difference is that topo VI was far more strict about the crossing angles

on which it would act. If the imposed crossing angle (α_L) is defined as the most common angle
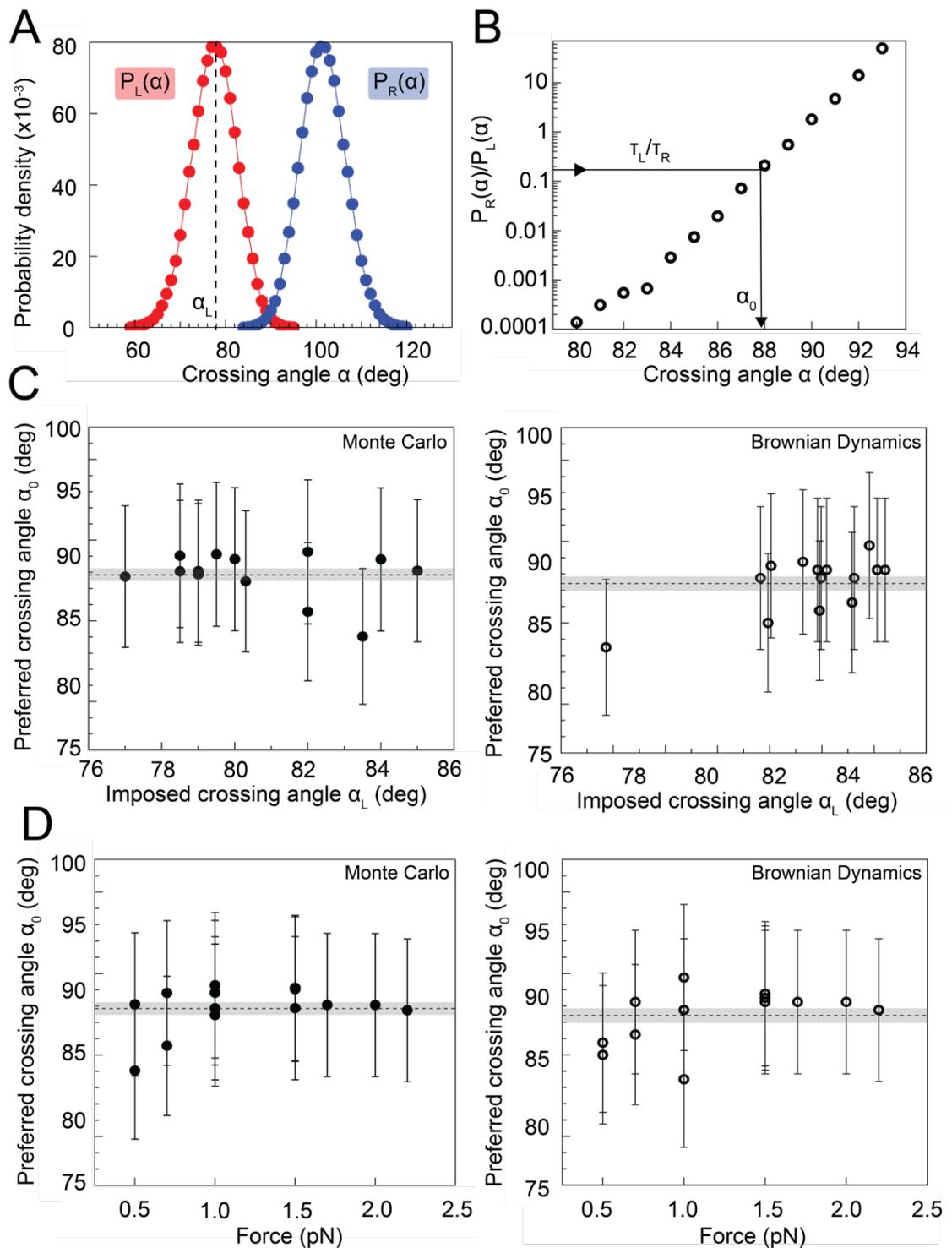
Figure 4.22: Preferred DNA crossing angle measurements for topo VI. A: Crossing angle probability distributions for single positive ($P_L(\alpha)$, red) and negative ($P_R(\alpha)$, blue) crossings, from Monte Carlo (MC) simulations [223] for the tether geometry and force seen in Figure 5. The

Chapter 4: Single-Molecule Results

positive angle probability distributions were attained directly, whereas the negative distributions were derived from the relationship $P_R(\alpha) = P_L(180°-\alpha)$. For Brownian dynamics (BD) simulations, the negative crossing angle distributions, like the positive, were measured directly. The imposed crossing angle (α_L, black-dotted line) is the most probable angle for the positive crossing angle distribution, in this case ≈79°. B: Ratio of negative to positive probability from Figure 4.22A plotted on a semilogarithmic axis. Using the relationship $\tau_L/\tau_R = P_R(\alpha_0)/ P_L(\alpha_0)$, the preferred angle ($\alpha_0$) can be determined, illustrated by the black arrow. For this tether geometry and force, $\tau_L/\tau_R =$ 0.19 ± 0.04, which gave α_0 of 87.9° ± 4.4° when using MC simulations, and 83.5° ± 4.2° when using BD simulations. The error associated with α_0, as is the same for Figures 6C and D, is the combination of the statistical and systematic error. C: Preferred crossing angles (α_0) from fourteen different DNA tether geometries, plotted against the most probable positive crossing angles (α_L), or imposed crossing angles, as measured by MC simulations (filled circles, left-hand plot) and BD simulations (open circles, right-hand plot). D: Preferred crossing angles (α_0) from fourteen different DNA tether geometries, as measured by MC simulations (filled circles, left-hand plot) and BD simulations (open circles, right-hand plot), plotted against the force on the DNA tether. The combined average preferred crossing angle for topo VI, ascertained using MC simulations, was 87.8° ± 0.4° (± SEM), and using BD simulations, was 87.4° ± 0.4° (± SEM), both values represented by the dotted line and error shading in plots C and D.

to appear in the positive distribution (see Figure 4.22A), then the most acute α_L that MmT6 could unlink was 77°, anything lower than this and no activity was detected. However, topo IV seemed far more versatile and was able to unlink crossings with α_L as low as 50° [18]. This may suggests that MmT6 either cannot remain bound to a disfavoured crossing for long enough to achieve the preferred angle, or that it cannot exert any control over the crossing angle, whilst topo IV may be able to. MmT6 is the most structurally simple type II topo and so it is possible that the lack of large CTDs and/or a third protein interface, renders it a far more passive crossing sensor. Taken together, the data obtained from the single crossing assays not only provides an explanation for the chiral selectivity of MmT6, but also strengthens the idea that topo VI is a preferential decatenase. Using MC simulations, it has been shown that angle distributions within catenanes are distributed around 90°, whilst supercoils are more acute at 60° [19]. A strict crossing angle preference of 88°, would indeed predispose topo VI to being far more active in decatenation than supercoil relaxation. Moreover, simulating the crossing angle distributions using two distinct simulation techniques and attaining strikingly similar

values for MmT6 α_0, not only adds confidence to the accuracy of this value but also supports the use of either simulation technique in DNA crossing angle distributions.

4.3   Conclusions

In this study, it has been demonstrated that not only is MmT6 a perfectly distributive and extremely slow supercoil relaxase, confirming work done in [62], but that in the presence of braided and catenated DNA substrates, its rate increased 5 to 10-fold, matching rates determined for other type IIA topos on their respective optimal substrates [19, 35]. Along with the observation that MmT6 has an extremely strict preference for DNA crossing geometries close to 90°, which appear far more frequently in catenanes than they do in supercoils [119], these data strongly indicates that MmT6 is a preferential decatenase, which in turn intrinsically disfavours supercoil relaxation (Figure 4.23). This leads to providing a potential explanation for the presence of topo VI in higher eukaryotes, during situations in which the genome is rapidly replicated, such as endoreduplication in plants, explored in depth in Chapter 6.

The chiral selectivity for positive writhe, explained by the crossing angle preference of 87.4-87.8°, was an unforeseen result. Whilst being consistently significant, on both plectonemic and braided DNA substrates, the 2 to 3-fold increase in rate on positive writhe is not particularly stark, unlike the difference seen for topo IV, which relaxed positive writhe 21-fold faster than negative [120]. So whether this rate difference for MmT6 has *in vivo* implications isn't clear, however as *M. mazei* encodes a DNA gyrase which maintains the genome in a negatively-supercoiled state, perhaps the slight angle preference for positive DNA crossings, prevents MmT6 from interfering with the activity of DNA gyrase. It may also indicate, as supercoils ahead of replication forks are positive, that MmT6 can help support the relief of replication and transcription induced torsional strain in combination with other topos.

Another surprising outcome of this work was the switch from being highly distributive in the presence of supercoils, to mild processivity on braids. There are two main differences between the magnetic tweezers supercoiling and braiding substrates: (1) braids do not accumulate twist upon magnet rotation, whilst supercoils do, and (2) the DNA-crossing angle

geometries within a braid are wider than those within the plectoneme. So it is likely that one of,

or a combination of the two, is promoting processivity by MmT6. The more likely candidate is

the DNA-crossing angle, promoting binding and strand passage on braids, as opposed to twist-

related effects. This is due to results obtained by varying the upwards force on the DNA by the

magnets (Figures 4.13 and 4.19). For braids, the main effect on the MmT6 reaction caused by

increasing force, is an increase in dwell time between relaxation events (Figure 4.19B),

suggesting that the time it takes for MmT6 to bind increased as angle distributions narrowed.

There was also a moderate decrease in processivity rate and step size on positive braids, which

again could be a function of the preferred angle becoming less probable as force increased

(Figure 4.19C and D). Therefore, MmT6 may have struggled to remain bound to the G-segment

long enough to capture and pass a consecutive T-segment at the correct juxtaposition.

However, this does not rule out the possibility that twist played a role, as using high force during

positive supercoil relaxation (Figure 4.13A), also caused a dramatic decrease in rate. The main

change in a supercoil with increasing force is the amount of twist required to buckle the DNA

and produce writhe, hence increased twist may also be inhibitory during supercoil relaxation.

The data trend obtained from the force titration experiments on both positive and negative

supercoil relaxation (Figure 4.13A and B) also indicated that increased force affected either

one rate limiting step in two ways, or two different steps. This is clearer for positive supercoil

relaxation, as high force promotes melting of negatively supercoiled DNA, thereby precluding

measurements at high forces (Figure 4.13). From low to medium force (0.2-1 pN) the rate

increases modestly before rapidly declining at high force (>1 pN). The initial rate increase may

be a function of slowing the diffusion of the crossings within the plectoneme so that they are

more easily captured by topo VI, as indeed it has been demonstrated that increasing the force

can have a dramatic effect on the diffusion of the plectoneme itself [317]. The rate decrease at

high force may be a function of increasing twist perturbations away from the relaxed state, that

could affect a variety of stages during the MmT6 reaction such as DNA binding and cleavage.

Taken together, the data presented in this chapter provides both supporting evidence

for the previously documented inefficiency in supercoil relaxation [62], but also extends to

providing an explanation for this as a function of a strict crossing angle preference close to 90°.

It was demonstrated that MmT6 was more active and processive during braid unlinking, and together with ensemble measurements of decatenation activity, it seems likely that MmT6 is preferential decatenase, which intrinsically disfavours the removal of supercoils.
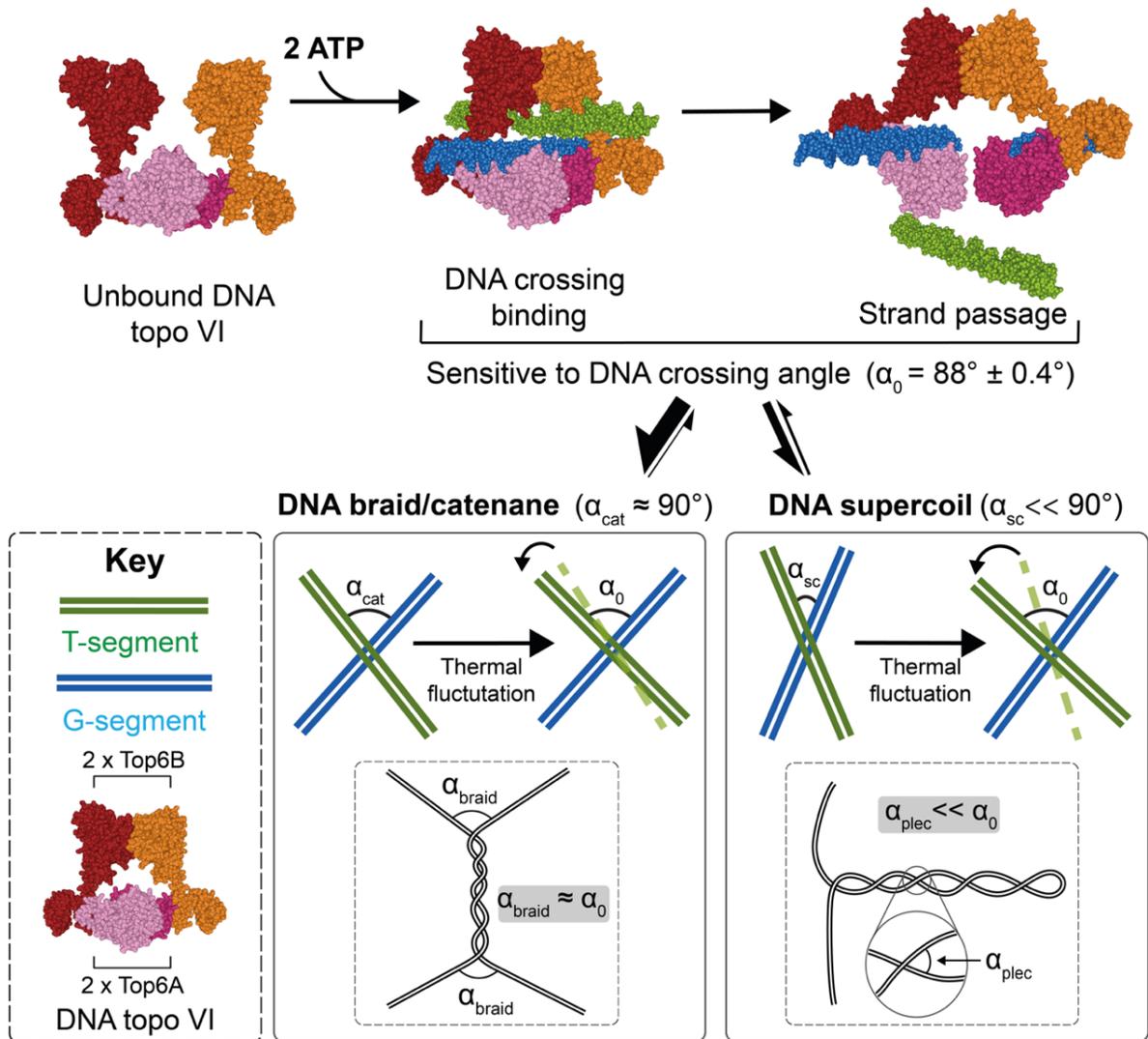
Figure 4.23: Summary model for chirality-dependent MmT6 activity. MmT6 DNA binding and strand passage has a chiral selectivity, which results in the passage of the T-segment occurring more rapidly when the DNA crossing angle ≈ 90°. This angle geometry is far more probable in braided and catenated DNA substrates than in supercoils, hence MmT6 preferentially decatenates whilst disfavouring the relaxation of supercoils. Larger thermal fluctuations must take place for supercoils to reach the preferred MmT6 crossing angle, and are therefore less efficiently relaxed. Chiral selection of positive supercoils over negative supercoils can also be explained by an angle preference of slightly below 90°, as this angle appears more often in

positive supercoils than negative. MmT6 and DNA structures modelled in PyMol, using PDB: 2Q2E (MmT6)[195] and PDB: 1ZEW (B-form DNA) [320].

4.4   Future work

The MT approach has gained popularity in the characterisation of topo activity at the single-molecule level, due to its exquisite sensitivity, highly quantifiable measurements and full control over DNA topology. The work reported here not only expands on this, but represents the first foray into the single-molecule behaviour of topo VI, using MmT6 as a model enzyme. A lot of insight has been gained from this work and the groundwork has been laid for future mechanistic studies on this curious enzyme. This could include exploration into the effects of salt concentration, the effect of twist density on activity, or mutational studies in which MmT6 is systematically explored for protein regions important for various aspects of activity (similar to that done using ensemble techniques [62]). The MT set up was used to explore DNA binding by topo IV, finding it both bends and over twists DNA upon binding in the absence of ATP [321], and this is another example of an assay which could be applied to topo VI. As mentioned in Chapter 1 and section 4.2.3, the MT set up facilitated the first measurement of gate opening dynamics for *E. coli* topo I and III [224]. This has not been achieved for a type II topo, and MmT6 would make a marvellous candidate for the exploring gate-opening, due to it having only two protein interfaces (N-gate and DNA-gate), rather than the three present in type IIA topos (N-gate, DNA-gate and C-gate).

Chapter 5

Utilising Next Generation Sequencing to Explore Type II DNA Topoisomerase Sequence preference *in vitro* Using Plasmid DNA

5.1   Introduction

The work included in this section (Section 5.1) has been adapted from previously published work by McKie *et al.,* 2020 [54] and appears with full permission.

As they are vital to cellular survival and constitute effective antibiotic and anti-cancer targets, understanding the DNA binding and cleavage profiles of DNA topoisomerases (topos) is of great interest to the field. Whereas the fundamental biochemistry of the topo reactions has been established, how their activity is regulated across the genome and through time *in vivo* has only recently begun to be addressed in a systematic manner with the utilisation of next-generation sequencing (NGS) technologies. NGS platforms have been adapted to generate genome-wide maps and sequence context of binding and cleavage of topos. Continuous refinements of these techniques have resulted in the acquisition of data with unprecedented depth and resolution, which has shed new light on *in vivo* topo behaviour. The binding and activity of topos over the genome *in vivo* was difficult to study until the advent of NGS. Over and above traditional chromatin immunoprecipitation (ChIP)-seq approaches that probe protein binding, the unique formation of covalent protein–DNA linkages associated with DNA cleavage by topos affords the ability to probe cleavage and, by extension, activity over the genome. NGS platforms have facilitated genome-wide studies mapping the behaviour of topos *in vivo*, how the behaviour varies among species and how inhibitors affect cleavage. Many NGS approaches achieve nucleotide resolution of topo binding and cleavage sites, imparting an

extent of information not previously attainable. Discussed below are the current examples in which NGS has been utilised to perform *in vivo* topo binding and cleavage activity.

5.1.1 Next-generation sequencing

The first widely used sequencing technique was developed in 1977 by Frederick Sanger and colleagues, which employed the use of chain terminating dideoxynucleotides [322]. It was later automated by Applied Biosystems with the development of the AB370 in 1987, which accelerated sequencing time and increased accuracy by using capillary electrophoresis [323]. Whilst this was a revolutionary technique for the time, and facilitated the completion of the first human genome in 2004 [324], this 'first generation' sequencing method had its limitations. Sanger sequencing is relatively low-throughput, time-intensive, and expensive. Efforts to ameliorate sequencing resulted in the development of affordable, high-throughput technologies. This led to a number of innovations in the DNA sequencing field and the birth of next-generation sequencing (NGS) platforms [325], which allow cell-free library preparation and are able to perform millions of parallel sequencing reactions, detecting the output directly without electrophoresis. This meant that genomes could be fully sequenced in an extraordinarily short timeframe. The main drawback, particularly for the first NGS platforms, was the relatively short read length. However, this was circumvented through the development of new algorithmic tools for aligning short reads.

In 2005 the first NGS platform was released by Roche, the 454 genome sequencer [326]. The detection method used, called pyrosequencing, relied on the activity of pyrophosphate, which would emit light when a base was incorporated, and allowed the sequencing of ~200,000 reads, ~110 bp long per run, generating 20 Mb of data. More recent advances made reads of 700 bp possible, generating 0.7 GB of data in less than 24 h with the 454 GS FLX Titanium system. In 2006, Sequencing by Oligo Ligation Detection (SOLiD) by Agencourt, was bought by Applied Biosystems [323]. This NGS platform uses octamer oligos, each with a unique fluorescent label, that interrogate the first two bases in the sequence adjacent to the sequencing primer. Once the label colour is detected, it is released by cleavage

between bases 5 and 6 of the octamer, and another probe will hybridize to the next two free bases and so on. The process is then repeated using a sequencing primer one nucleotide shorter than the first. In this manner, the first-round sequences bases in positions 4, 5, 9, 10, 14, 15, etc., whilst the second round sequences 3, 4, 8, 9, 13, 14, etc., until sequencing primer position 0 is reached [327]. Due to each base being interrogated twice, early SOLiD systems benefitted from high accuracy, with more recent versions able to sequence 85 bp fragments with 99.99% accuracy [323]. Also, in 2006, the Illumina platform was released, which facilitates detection using fluorescently-labelled reversible-terminator nucleotides. Each nucleotide is uniquely coloured allowing its incorporation to be detected, before the subsequent removal of the 3' terminator group along with the fluorescent probe. Numerous Illumina platforms have since been manufactured, employing this technique, such as MiSeq, HiSeq and NextSeq, with the main difference between these systems being the depth of parallel sequencing power and data output. The most recently released Illumina technology, Novaseq, can perform up to 40 billion sequencing reactions in parallel, of up to 250 bp in length with an output maximum of 6000 Gb. This sequencing scale has opened the door to a wide range of innovative approaches that leverage these platforms to develop unique methodologies to probe field-specific questions. The research cited here relies on using one of the three aforementioned platforms. In most instances, the choice of platform is somewhat arbitrary, largely dictated by the best available sequencing depth and length. A comprehensive overview of sequencing platforms is provided in [323].

5.1.2 Non-specific DSB mapping

Genome metabolism is highly coordinated, dynamic and staggeringly complex, with numerous proteins interacting with both the duplex and other proteins, to maintain DNA integrity while facilitating DNA processing. DNA within each cell is subject to an assault from both endogenous and exogenous factors, resulting in various forms of damage that require either repair or tolerance [328]. Arguably one of the more toxic forms of damage is the induction of DNA breaks, particularly DSBs, which can lead to carcinogenic translocations or cell death [329]. One early

use of NGS was to explore the DSB landscape of the genome, to better understand when and where the majority of DSBs occur. One such study developed a generalized technique called BLESS (direct in situ Breaks Labelling, Enrichment on Streptavidin and next-generation Sequencing) [330]. By labelling DSBs in situ, the level of false positives caused by genome shearing during extraction was minimized, allowing the resolution of the "breakome" in both human and mouse cells. The cells were fixed and lysed to attain purified nuclei, which were then subject to biotinylated adapter ligation. The genomic DNA could then be fragmented and enriched for endogenously labelled DSBs using streptavidin. Exposing cells to aphidicolin, an inhibitor of DNA polymerase, it was found that replication stress is a major driver of DSB generation, particularly within genes and satellite regions. BLESS was also able to accurately identify telomere ends as well as Sce-1 endonuclease sites.

A subsequent study developed an alternative technique, designed to overcome noise and background complications associated with BLESS, as well as to gain information on DSB end structure [331]. Termed END-seq, this protocol involves embedding mouse thymocytes or lymphocytes within agarose before soaking the agarose in proteinase K solution, followed by an RNase solution. The agarose embedded cells are then end repaired and adapter ligated, using a specifically indexed and biotinylated adapter. The DNA is then extracted, sheared, and immunoprecipitated using streptavidin, before ligation of the second adapter. The DNA library was single-end sequenced using Illumina Hiseq2500 or Nextseq500. This technique proved to be highly-sensitive, with nucleotide resolution at break sites, and led to a number of new insights in the field of T- and B-cell research. It also has potential as a tissue specific CRISPR-Cas9 off target cleavage detection protocol. A limitation, however, is evident in the end repair stage in which 3'-overhangs are resected and 5'-overhangs are filled-in to blunt the DNA for adapter ligation. This means that any information concerning the length of the overhang is lost.

5.1.3 Spo11 DSB mapping

One of the first studies to implement the use of NGS in topoisomerase-specific genome-wide DSB mapping, explored the generation of meiosis specific DNA cleavage by the

topoisomerase-related protein, Spo11, in *Saccharomyces cerevisiae* [205]. Spo11 acts as a homodimer and is highly homologous to the topoisomerase VIA subunit [203]. It is responsible for DSB generation during meiosis, which ultimately leads to chromosomal recombination, making Spo11 an important factor in driving evolution through the generation of genome diversity. Spo11 homodimers form covalent 5' phosphotyrosine linkages on both strands of a DNA duplex, generating a DSB with a 2-base overhang, a process highly reminiscent of the type II topoisomerases. The main difference being that Spo11 DSBs are irreversible and are followed by endonucleolytic cleavage adjacent to the bound Spo11. This releases the Spo11 monomer covalently attached to an oligonucleotide (oligo) and generates 3' overhangs, which facilitate strand invasion and recombination. The released Spo11-bound oligos are separated into two distinct subpopulations of equal quantity, longer (21–37 nt) and shorter (<12 nt). Pan and colleagues used these oligos to develop an NGS-based approach, termed Spo11-oligo-seq, designed to overcome limitations encountered by previous studies employing DNA microarray analysis, to attain a comprehensive genome-wide Spo11-mediated DSB map with nucleotide resolution at the cleavage sites [205].

Nuclear extracts were obtained from synchronous meiotic yeast cells and the Spo11-bound oligos captured via immunoprecipitation and then released by proteolytic digestion (Figure 5.1A). The oligos were then 3′-GTP tailed using terminal deoxynucleotidyl transferase (TdT) and ligated to a double-stranded DNA adapter using T4 DNA ligase before the complementary strand was synthesized using Klenow polymerase. The DNA was purified using denaturing polyacrylamide gel electrophoresis (PAGE), before a second round of 3′-GTP end tailing, adapter ligation and strand synthesis. The original oligo cannot be further modified as the 3′ adapter has an inverted dT and the 5′ DNA end is blocked by the phosphotyrosine residue. Sequencing adapters were then added via PCR before the oligos were sequenced using the 454 platform (Roche). The fragments could then be aligned and bioinformatically processed, generating an in-depth representation of Spo11-DSB activity genome wide. This revealed a complex picture in which the DSB landscape in *S. cerevisiae* is influenced by numerous factors in a hierarchical manner, and that the concept of DSB 'hotspots' [332] is

misleading. Spo11 demonstrates an opportunistic cutting activity where most sites can constitute a cleavage site but with variable probability that is modulated by numerous other factors.

5.1.4 DNA topoisomerase II DSB mapping

In 2014, human topoisomerase II (topo II), both the α- and β-isoforms collectively, became the first true topoisomerases explored at the genome-wide scale [333]. Human colon cancer cells (HCT116) were treated with etoposide (a topo II cleavage-complex stabilizing agent) and the genomic DNA extracted. Here, two different methods were implemented to capture both SSBs and DSBs, named SSB-seq and DSB-seq, respectively. In order to label SSBs, nick-translation was used, a technique first described in 1977, which enables labelling of DNA at a nick site [334]. In this case, digoxigenin-labelled dUTP was used so that the nicked DNA fragments, once sonicated, could be immunoprecipitated using anti-digoxigenin. For DSB-seq, the DNA was 3'-end tailed using TdT in the presence of biotinylated dNTPs so that after sonication, the DNA could be enriched using streptavidin-coated beads. S1-nuclease was used to cleave the 3'-end tailing and then both the SSB- and DSB-seq DNA populations were adapter ligated and sequenced using Illumina.

This study highlighted the tendency of low-dose etoposide to produce SSBs, as opposed to DSBs, indicating that etoposide predominantly acts on a single monomer of topo II, suggesting in turn that only one cleavage site is poisoned, leaving the other to be freely re-ligated. This is in agreement with a previous study that found only 3% of etoposide-stabilized cleavages are DSBs [335]. It is worth noting, that whilst topo I (type IB) can form single-stranded nicks, these nicks do not have the free 3'-OH group, so are not substrates for nick-translation and won't be labelled, making SSB-seq topo II-specific. Whole-genome analysis revealed that topo II SSBs and DSBs were enriched at transcription start sites and more prevalent in highly expressed genes, further demonstrating the cellular necessity of topo II in regulating topology during transcription.

In 2016, came a comprehensive study looking specifically at murine liver topo IIβ activity, using ChIP-seq, ChIP-exo and Hi-C [163]. For an in-depth ChIP-seq protocol see [336], however, in short, DNA protein interactions were stabilized using a crosslinking agent and topo IIβ-bound DNA was enriched using anti-topo IIβ antibody. The crosslinking was reversed, and the DNA underwent library preparation in which the ends were repaired, A-tailed, and adapters were ligated before sequencing using the Illumina HiSeq2500 platform.

Topo IIβ and topo IIα, whilst being structurally and catalytically alike, are functionally distinct, performing different cellular functions. Topo IIβ knockouts are lethal owing to disruption of neuronal differentiation [26]. Conditional knockouts implicated topo IIβ activity in retinal development [159] and ovulation [160], and the use of poisons showed it played a role in spermatogenesis [161] and lymphocyte activation [162]. This has led to a picture of topo IIβ being an important factor in tissue-specific development and cellular differentiation, a concept elaborated upon by Uusküla-Reimand and colleagues with the help of ChIP-seq [163]. They found topo IIβ interacts with the transcriptional regulator CTCF and the cohesin complex and acts on DNA located at the borders of chromosomal domains as well as conserved transcription factor binding sites.

Topo IIα activity genome-wide was also explored in depth in 2018 [155], using human leukaemia K562 cells treated with etoposide, mitoxantrone, genistein or p-benzoquinone, all of which are known topo II inhibitors. DNA was extracted and sonicated before bead-based immunocapture using rabbit anti-topo IIα polyclonal antibody. The DNA was released from the topo IIα using calf alkaline phosphatase (CIP), an enzyme capable of phosphodiesterase activity at 3% of its canonical monophosphatase activity. CIP is therefore able to partially remove the phosphotyrosine residue, allowing for end repair and adapter ligation. The generated libraries were then sequenced using the illumina HiSeq2500 platform. In agreement with past studies, they found that the majority of topo IIα cleavage sites were SSBs rather than DSBs, the latter only making up 1.3–3.4% of cleavage events. Reads clustered in regions of compact genome, reminiscent of the cleavage profile of yeast Spo11, and with a preference for regions with high transcriptional activity. They also identified chromosomes that consistently

contained more topo IIα cleavage sites than others. In particular chromosome 11 that includes the KMT2A gene, which houses a drug-induced leukemia translocation break point. The KMT2A gene, along with others, forms the KMT2A recombinome, and a proportion of these genes demonstrated significantly higher levels of topo IIα cleavage. They also concluded that topo IIα cleavage sites were enriched in genes necessary for DNA metabolism and transcriptional regulation.

As protocols are refined, an increasingly more precise DSB landscape for topo II is emerging. In 2019, DSB maps were generated with cleavage-site nucleotide resolution for human and yeast topo II [337]. The protocol, termed cleavage complex (CC)-seq (Figure 5.1B), begins with genomic DNA extraction from human or yeast cells treated with etoposide, followed by a phenol-chloroform extraction to remove bulk proteins, leaving protein-bound DNA in the aqueous fraction. The DNA is then sonicated and enriched using a silica membrane, which will only bind DNA that is protein bound. The library preparation then occurs in two phases, with the P7 adapter being ligated to the protein-free DNA end. The topo II is then cleaved using TDP2, an enzyme evolved to rescue stalled topo II cleavage complexes in vivo [338], before the P5 adapter is ligated and the library is sequenced. In this study the NGS platforms used were the Illumina MiSeq and NextSeq 500.

As this study was conducted on both human and yeast cell lines, species-specific conclusions about topo II activity could be made. For both human and yeast, topo II activity was increased in gene dense regions of the genome. However, only a correlation with transcription rate and gene length was observed in human cells. For yeast, topo II activity was concentrated in intergenic regions, which are numerous and evenly spaced. For human cells, topo II activity was found in the gene bodies just downstream of the transcription start site and was maximized for genes that are longer and more actively transcribed. This strengthens the hypothesis that topo II in human cells plays an integral role in topology maintenance during transcription. This gene-local activity pattern may not be necessary in yeast as the genome is both far smaller and less complex. There is also the potential alternative that in yeast cells topo I may be more active in relieving transcriptional stress on the duplex than topo II, whilst in

human cells this role is more evenly shared. It is worth noting that for topo IIα specifically [155], activity in the absence of cleavage-complex stabilizing drugs was found to be located at the distal ends of genes, with the addition of drugs causing a proximal shift.
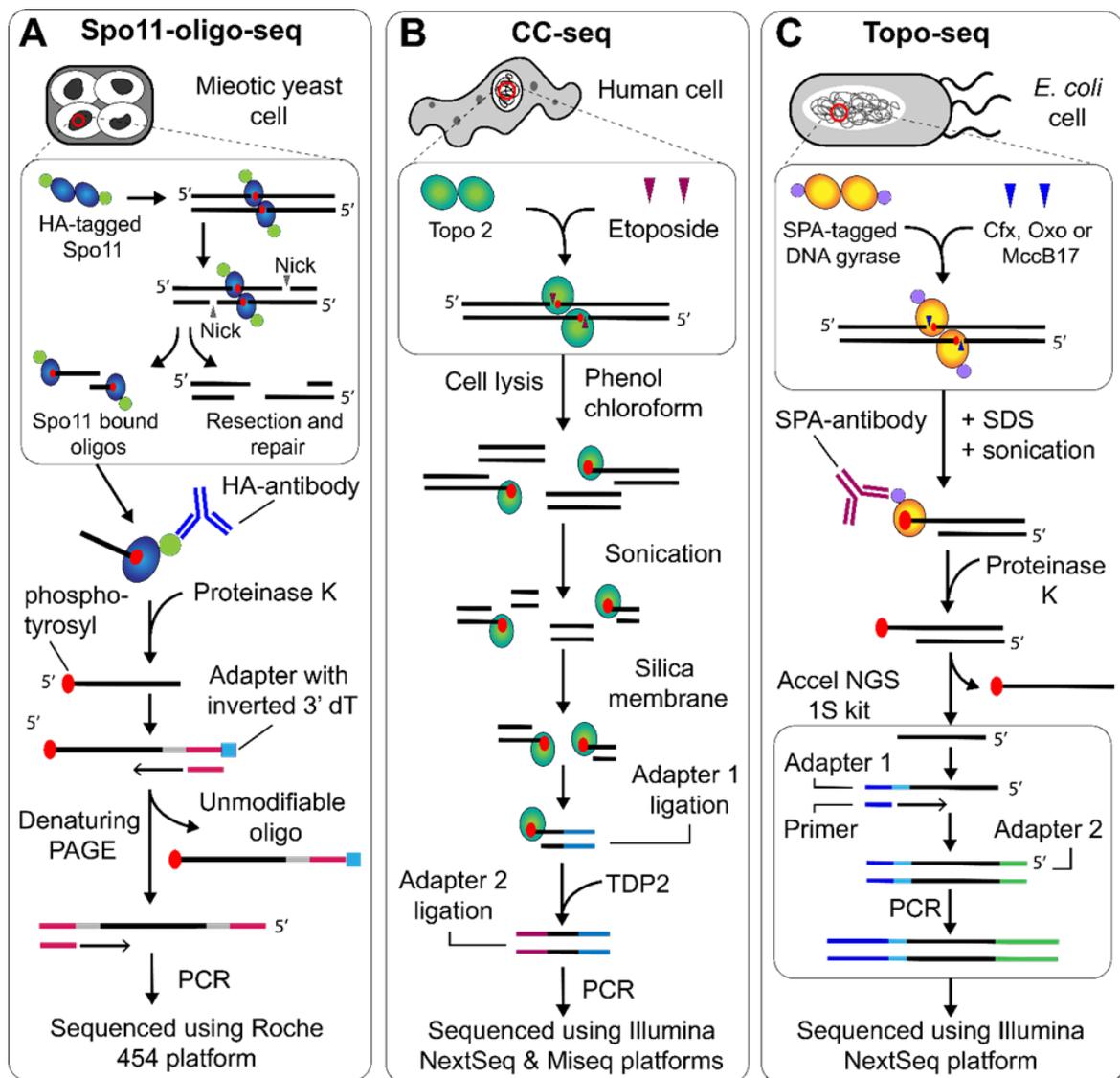
Figure 5.1: NGS protocols to map topo cleavage sites genome wide. (A) Spo11-oligo-seq began with DNA extraction from meiotic *S. cerevisiae* cells. HA-tagged Spo11-bound DNA oligos are enriched with the HA-antibody, before the protein is removed using proteinase K. The 3' end is extended using TdT and an adapter is ligated which contains an inverted dT residue. The complementary strand is synthesized, followed by separation using denaturing PAGE, 3' end tailing and adapter ligation on the newly synthesized strand, which can then be sequenced using the Roche 454 platform [205]. (B) CC-seq, used to map topo II cleavage, began with DNA extraction from etoposide treated human cells. The bulk protein is removed using phenol-chloroform extraction, whilst the protein bound to DNA remains in the aqueous phase. The DNA can then be sonicated and enriched using a silica membrane. The first ligation attached adapters to the sonicated end, which was then followed by TDP2-dependent removal of the topo II, before

a second adapter ligation followed by sequencing using the Illumina NextSeq and Miseq platforms, [337]. (C) Topo-seq, used to map gyrase cleavage, began with DNA extraction from E. coli cells treated with ciprofloxacin (cfx), oxolinic acid (oxo) or microcin B-17 (MccB17), followed by sonication. Gyrase cleavage complexes were then enriched using the SPA-antibody before proteinase K treatment and use in the Accel NGS 1S kit to generate fragments that can be sequenced using the Illumina Nextseq platform [339].

An analysis of etoposide-induced topo II cleavage complexes (topoIIcc) in mouse and human cell lines used the aforementioned END-seq protocol, which was adapted to explore both processed and unprocessed topoIIcc [340]. Mouse embryonic fibroblasts (MEFs) depleted of the cohesin subunit, SMC3, exhibited a dramatic decrease in topoIIcc, whereas Smc3+/+ control cells exhibited topoIIcc enrichment at cohesin sites. Not only was this decrease in topoIIcc proportional to coehsin levels in MEFs, similar results were obtained with a HCT-116 human cell line depleted of RAD21, another member of the cohesin complex. Moreover, depleting MEFs for the WAPL protein, responsible for cohesin removal from DNA, also increased the levels of etoposide-induced DNA damage. Interestingly, whilst topo II activity as well as binding (also measured here using ChIP-seq) was dependent on cohesin occupancy; the converse relationship was not apparent, with RAD21 and CTCF levels remaining unchanged in topoIIβ−/− B cells.

Whilst the activity and binding of topo II was found to be dependent on cohesin, the processing of these sites was sensitive to transcriptional activity. Etoposide-induced topoIIcc occurring within transcription start sites (TSS) or gene bodies were more likely to produce chromosomal translocations than those occurring in intergenic regions, with higher levels of translocations in regions of high transcription and decreased levels of translocations on addition of the transcriptional inhibitor 5,6-dichloro-1-beta-d-ribofuranosylbenzimidazole (DRB). In eukaryotic cells, the conversion of a topoIIcc to a protein-free DSB is mediated by TDP2. The processing of topoIIcc in the presence of DRB was decreased, with more lesions remaining protein-bound, whilst transcriptional activation by INFβ increased levels of protein-free DSBs. Hence, the processing and repair of topoIIcc is stimulated by transcriptional activity, and this in turn leads to an increase in translocation rate. This effect seemed to be dependent on

proteasome activity, rather than TDP2, with proteasome inhibition by epoxomicin increasing the number of topoIIcc that are reversible (protein-bound) and decreasing the amount of protein-free DSBs 5.5-fold.

5.1.5 DNA topoisomerase IV DSB mapping

In 2018, NGS was used to explore DNA topoisomerase IV (topo IV) binding and cleavage sites genome wide in *E. coli* [125, 341]. To explore binding in the absence of cleavage, a traditional ChIP-seq method was employed in which topo IV was allowed to bind DNA and then crosslinked, followed by antibody-based fragment capture. To explore topo IV cleavage complexes, a novel approach was developed, called NorfIIP, relying on norfloxacin, a quinolone that traps cleavage complexes produced by prokaryotic type II topoisomerases.

Four highly enriched topo IV binding sites were found, one of which corresponded to the dif site, a previously confirmed hotspot of topo IV cleavage. The dif (deletion induced filamentation) site is located at the replication terminus and is important for chromosome dimer resolution [342]. To identify additional binding sites, the raw sequencing data were filtered for sites showing the highest Pearson correlation with the sequencing pattern at the dif site. This approach led to the identification of 19 additional topo IV binding sites throughout the chromosome. These sites were found frequently within intergenic regions and spanned 200 bp with no significant consensus sequence. Topo IV binding enrichment was also detected in GC-rich regions. By combining this binding data with specific topo IV cleavage sites, enhanced with norfloxacin, it was found that only certain binding sites correspond to topo IV cleavage sites, indicating that not all DNA binding results in topo IV cleavage of the DNA. The dif site remained strong for both binding and cleavage by topo IV, as well as a site at 2.56 Mb. However, the other strong binding sites were not present in the norfloxacin-dependent cleavage data. A characteristic topo IV-dependent peak profile was found in which two 170 bp peaks were present, separated by a 130 bp reduction in sequencing coverage. This was hypothesized to be due to the presence of topo IV residues remaining covalently bound to the 5′ ends of the cut site and preventing adapter ligation and sequencing. This characteristic signal was used to

algorithmically detect genome-wide topo IV cleavage sites. Using this approach 88 sites, common across three different experiments, were identified. As expected, topo IV cleavage at the dif site was the most enriched. Interestingly, most NorflIP sites were not observed in the ChIP-seq binding experiments, indicating that not only does binding not strongly correlate with cleavage, but that cleavage does not strongly correlate with binding. No strong consensus sequence preference for topo IV cleavage was detected, beyond a slight bias for GC dinucleotides and an increased spacing between GATC motifs around the cleavage sites.

5.1.6 DNA gyrase DSB mapping

In 2018, an NGS approach was developed that allowed the precise identification of gyrase cleavage sites genome-wide in *E. coli*, facilitating in depth analysis of drug driven cleavage in vivo [339]. The method, named Topo-seq (Figure 5.1C), involved trapping gyrase cleavage complexes using either oxolinic acid (OXO), ciprofloxacin (CFX), or microcin B 17 (MccB17) and immunoprecipitating the resulting fragments. They then underwent a single-strand library preparation protocol that only ligates adapters to the 3′ free strand from the cut site. This method was developed to avoid issues caused by the 5′ covalent linkage to the active site tyrosine of gyrase. Proteinase K treatment leaves the phosphotyrosyl residue at the 5′ cut site, which drastically reduces ligation efficiency of the sequencing adapters. In previous studies, the use of CIP or TDP2 offered an enzymatic solution. However, the efficiency of these enzymes is not always optimal, leading to the introduction of potential artefacts. In topo-seq, this effect is mitigated entirely by solely focusing on the unbound 3′ single strand. By sequencing only these strands from either side of the cut site, the expected characteristic signal of gyrase cleavage involves a bimodal peak with a sharp 4-base sequence loss in the centre, indicative of gyrase cleavage. This unique signal allowed gyrase cleavage sites to be accurately called and subsequently analysed.

Using this approach, Sutormin and colleagues identified 4635 gyrase cleavage sites (GCSs) when cells were treated with CFX, 5478 when treated with OXO and 732 when treated with MccB17. An in-depth comparison among these inhibitors and their cleavage sites revealed

not only inhibitor-specific cleavage patterns, but also inhibitor concentration-dependent effects. Having identified a wealth of cleavage sites, robust analysis of the gyrase sequence preference was facilitated, finding the gyrase binding motif to be both extensive and degenerate. This motif is ~130 bp in length, symmetrical on either side of the cleavage site and present in all cleavage sites, regardless of the inhibitor used. It exhibits a periodic GC content fluctuation, reminiscent of the binding pattern seen for eukaryotic nucleosomes. The only inhibitor-specific differences identified were associated with the precise cleavage site. The tendency for CFX and OXO to intercalate and drive cleavage at guanine nucleotides was confirmed. However, the sequence preference for MccB17 exhibited a variable pattern and was unlike oxo and cfx. This likely reflects the fact that OXO and CFX have a different mode of action to MccB17.

Gyrase showed a strong cleavage preference for highly transcribed loci, producing the most distinct enrichment of GCSs around the rRNA operons, known to have high rates of transcription. Furthermore, the use of rifampicin, a transcriptional inhibitor, not only decreased the number of cfx-dependent cleavage sites by half, but also produced significant relocation of GCSs away from rRNA operons and decreased gyrase avoidance of poorly transcribed regions. Together, these data implicate *E. coli* gyrase directly in topology manipulation during transcription.

5.1.7 DNA topoisomerase IA DSB mapping

The first type IA topoisomerase to be explored using NGS was done in concert with gyrase and RNA polymerase (RNAP) from *Mycobacterium tuberculosis* [113]. Whilst *M. tuberculosis* is a well-known and devastating human pathogen, with resistant strains beginning to dominate in some areas of the world [343], it is also a curious bacterial species in regard to topoisomerase encoding. Whilst the majority of bacterial species encode 4 topoisomerases: two type IA topoisomerases (topo I and topo III) and two type IIA topoisomerases (DNA gyrase and topo IV), *M. tuberculosis* has only one type IA (topo I) and one type IIA (DNA gyrase) [344]. The rationale behind this study was to obtain direct evidence of topo I and DNA gyrase interacting at sites of transcription, to further support the twin-supercoiled domain model of transcription

and the role that topoisomerases play in removing accumulated torsional stress in the DNA caused by the transcriptional machinery. Hypothesized by Liu and Wang in 1987, the twin-supercoiled domain model describes how the separation of the DNA strands during transcription and replication would lead to positive supercoiling ahead of the protein machinery and negative supercoiling behind [111]. Indeed, numerous in vitro studies support this hypothesis, demonstrating that transcription alters DNA topology [345]; inhibition of DNA gyrase leads to positive supercoil accumulation [105] and mutation of topo I causes negative supercoil accumulation [346]. However, direct evidence of the association of bacterial topoisomerases and the transcriptional machinery was not obtained until a ChIP-seq-based protocol was implemented genome wide using *M. tuberculosis* [113]. ChIP-seq involves the crosslinking of protein to DNA, the DNA of which is enriched for fragments bound to a protein of interest, usually using a protein-specific antibody. So, the data set produced by the subsequent sequencing of this DNA library generates a binding profile for the protein, rather than a direct read out of cleavage activity.

In this study, the binding of *M. tuberculosis* gyrase and topo I were compared to the binding of RNAP. In agreement with the twin-supercoiled domain model, both topo I and gyrase were found to be specifically enriched in regions of transcriptional activity, with a high correlation with RNAP binding. As expected, RNAP was enriched at promoter sequences with topo I enrichment proximally-located and gyrase enrichment both overlapping with RNAP and downstream. In addition, both gyrase and topo I were highly enriched at the origin of replication, directly implicating them in replication, whilst only gyrase displayed a binding peak at the termination site. This is likely due to *M. tuberculosis* gyrase having a dual activity of both negative supercoiling and efficient decatenation [347], the latter being a role usually fulfilled by topo IV in other bacteria.

In 2019, another study from the same group, explored topo IA activity in more depth, contrasting binding and cleavage in *Mycobacterium smegmatis*, a close relative of *M. tuberculosis* [348]. In this study, four libraries were produced, three to probe topo IA cleavage, and another for topo IA binding. The first of the topo IA cleavage protocols involved the use of

a poisonous topo IA mutant, with a substitution in the topoisomerase-primase (TOPRIM) domain (D108A), previously shown to cause accumulation of cleavage complexes [349]. Once transformed into competent *M. smegmatis* cells, an increase in cleavage complexes and stimulation of the SOS response were detected, demonstrating that the mutant protein was being expressed and interacting with the genome. The second technique designed to probe topo IA cleavage implemented the use of a newly discovered topo IA inhibitor, imipramine [350]. Imipramine is a tricyclic antidepressant that has been shown to act specifically on topo IA from mycobacteria and prevent re-ligation, thereby increasing DNA cleavage complexes. The third technique, designed to probe binding, followed a traditional ChIP-seq type method. They also managed to capture wild-type topo IA cleavage sites, in the absence of inhibitors. In all four cases, topo IA-specific DNA fragments were immunoprecipitated and sequenced using the Illumina Genome Analyzer IIx.

Being able to contrast binding and cleavage, as with topo IV [125, 341], revealed a similar disconnect between binding and catalytic activity. Whilst binding was shown genome-wide and somewhat evenly distributed, for topo IA, the cleavage profiles were far more region-specific. Multiple regions displayed topo IA binding enrichment where cleavage was completely absent. In vitro assays demonstrated that the enriched sequence was cleavable. However, topo IA inhibition by Nucleoid Associated Proteins (NAPs) measured in vitro suggests that the inhibition of topo IA cleavage in vivo could arise from the presence of NAPs. A unique feature amongst type IA topoisomerases is the mycobacterial topo IA sequence preference for both binding and cleavage, which was identified in previous in vitro studies [351-353]. Analysis of the genome-wide cleavage sites of *M. smegmatis* topo IA revealed a similar consensus sequence as previously established for topo IA in vitro, indicating that topo IA binds and cleaves DNA in a sequence-dependent manner in vivo.

As with the previous study looking at topo IA and gyrase co-localization in *M. tuberculosis*, topo IA in *M. smegmatis* exhibited increased occupancy and cleavage associated with RNAP and highly transcribed genes. However, in *M. smegmatis* both cleavage and binding were detected at the termination site, a result not previously seen, implicating topo IA in

chromosome segregation and suggesting that, like mycobacterial gyrase, topo IA also has dual activities in both relaxation and decatenation. In most bacterial species the dedicated type IA topo involved in decatenation is topo III, an enzyme absent in mycobacterial species. *In vitro* assays performed in this study did indeed show that *M. smegmatis* topo IA had significant catenation and decatenation activity, and therefore may also be involved in chromosome segregation along with gyrase, in vivo.

5.1.8 DNA topisomerase IB DSB mapping

Currently only one study has mapped genome-wide activity of human topo IB, a type IB topoisomerase [354]. Using HCT116 human colon cancer cells, topo IB binding was assayed using a ChIP-seq method, and cleavage by a method termed Top1-seq. Top1-seq involved the use of camptothecin, a topo IB inhibitor, to generate covalent cleavage complexes, which were enriched using immunoprecipitation and then sequenced.

Topo IB binding enrichment was highly correlated with transcription, with 67% of DNA-bound topo IB interacting with transcribed genes both upstream of the TSS and downstream of the transcription termination site. By quantifying the transcription level of genes using RNA-seq, a method developed to allow sequencing of the transcriptome [54], increased topo IB occupancy was shown to correlate with increased levels of transcription, providing further evidence for its role in transcriptional topology regulation in eukaryotic cells, much like topo IA in prokaryotes. Comparison of the binding landscape of topo IB to RNAPII revealed their localization was highly correlated, in a manner that suggested a direct protein-protein interaction, rather than topo IB being solely bound to regions of torsional stress within the genome. To explore this, cellular extracts were subjected to a pull-down assay, which confirmed the direct interaction between topo IB and RNAPII in the absence of DNA, indicating that topo IB is in integral member of the transcriptional machinery and is recruited to sites of RNAPII engagement.

As mentioned, this study used a technique called Top1-seq, which was designed to capture catalytically-engaged topo IB, using a brief camptothecin treatment. Whilst binding of

topo IB was high at promoters, particularly in concert with RNAPII, catalytically engaged topo IB was diminished at TSS but enriched within gene bodies, again with a positive correlation with transcription level. Cleavage was also biased towards the template strand, indicating that topo IB activity is coupled to the movement of the transcriptional machinery. Using complementary in vitro assays, RNAPII was found to enhance topo IB relaxation activity and increase processivity under physiological ionic conditions, whilst none of the other general transcription factors seemed to have an effect on topo IB relaxation activity. The interaction between RNAPII and topo IB was mapped to the CTD of RNAPII and the NTD of topo IB. An interaction that increased the time topo IB spends in the cleaved state, enhancing the processive removal of supercoils.

5.1.9 Remarks on the current literature

Since the advent of NGS technologies in 2005, numerous techniques have been developed to explore genome-wide DNA binding and cleavage landscapes, which has both answered and posed new questions concerning *in vivo* topo behaviour. Encouraging evidence in support of the widely accepted twin-domain model, has established a greater understanding of topoisomerase activity during transcription and to a lesser extent replication. Topo II and topo I in eukaryotes exhibit higher levels of activity in the vicinity of highly transcribed genes [333, 354], as has also been shown for mycobacterial topo I and gyrase [348]. Topo IV, on the other hand, demonstrates a clear preference for the dif site in *E. coli* [125, 341], supporting the idea that decatenation is its primary cellular function [122].

Using NGS as a means to explore topo cleavage across the genome, encouraged innovation in protocol development to circumvent issues which arise as a result of the covalent linkage between the protein and the DNA backbone. Proteinase K will remove the majority of the protein but will not cleave the phosphotyrosyl bond, leaving the adduct to interfere with sequencing adapter ligation. Two enzymatic solutions have been used to remove the adduct, namely CIP and TDP2, the first of which is very inefficient and the second far less commercially accessible. Using a single-strand DNA library preparation kit, as Sutormin and colleagues did

when exploring DNA gyrase activity [339], offers another elegant solution by only sequencing the free 3' DNA strand.

One particularly intriguing result has come from contrasting topoisomerase binding with cleavage activity. For some proteins, e.g., topo IV and topo IA, the correlation between them is so low that most binding sites do not constitute cleavage sites and vice versa. The nature of the techniques used to explore binding verses cleavage (such as using cleavage complex stabilizing agents which bias cleavage preference) and the platforms used to sequence the fragments, may contribute to this discrepancy. However, *in vitro* biochemical data also suggest that the relationship between cleavage and binding is not so clear. For example, topo IIα whilst being more catalytically active on positively supercoiled DNA, maintains higher levels of cleavage on negatively supercoiled DNA [156]. This raises questions as to the determinants of topo binding and cleavage sites; how do the factors known to influence binding and activity including DNA topology and sequence play off against protein-mediated recruitment or inhibition in trans, and what other factors may influence topo binding and cleavage in vivo. It is clear the answer is not straightforward, with topo IB being directly recruited as part of the transcriptional machinery [354] and mycobacterial topo IA [348] and prokaryotic DNA gyrase [339] exhibiting more substantial DNA sequence preferences. As with all cellular processes, the mechanisms governing DNA topology maintenance are often not only protein-specific but also species-specific. A difficulty with collecting such vast amounts of data is the interpretation thereof, which has been challenging in some cases due to the complexity of the cellular environment. One possible approach would be the use of in vitro NGS techniques to map topo binding and cleavage in which DNA topology and sequence can be systematically controlled in addition to the use of specific proteins and precise buffer conditions, including inhibitor addition. Such an approach would complement genome-wide studies and could provide insights that could aid in the interpretation of genome-wide results. NGS is undoubtedly a very powerful technique, both in isolation and as part of an experimental repertoire, which will hopefully be further implemented in the field of DNA topoisomerases.

The work presented in the remainder of this chapter demonstrates a novel *in vitro* technique, which can be used to further explore topo cleavage activity, imparting a far higher level of control over the assay conditions. This technique, named SHAN-seq (Simplified High Accuracy eNd-sequencing) could potentially allow the systematic evaluation of how different topos bind and cleave DNA, how inhibitors and other proteins affect this behaviour, and reveal any DNA topology-dependent effects. The latter, in particular, is a factor of which one has very little to no control over *in vivo*. Data such as this would be highly complementary to *in vivo* studies and could help inform future conclusions drawn from the vast quantities of data often produced by NGS techniques. SHAN-seq is accessible, relatively simple, and fast, allowing the in-depth characterisation of DNA cleavage activity *in vitro* and under specifically controlled conditions.

5.2 SHAN-seq protocol and optimisation

SHAN-seq is a novel *in vitro* technique that employs the use of accessible equipment and materials, paired with NGS, to thoroughly explore type II topo-dependent cleavage activity. The initial motivation behind the development of this technique was to find an innovative way to explore potential *M. mazei* topo VI (MmT6) sequence preferences and elucidate the base-stagger generated during cleavage. For topo VI, this has only be explored once before using radioactivity, for *S. shibatae* topo VI (SsT6) [194]. This work found that SsT6 generated ATP-dependent DSBs in which the overhang formed was 2 bases. This was a new finding for a type II topo, as all type IIA topos explored thus far generate 4-base staggers [355]. This strengthened the division between the type IIA and B topos, and the 2-base stagger was also confirmed to be produced during Spo11-dependent cleavage [205, 356], reinforcing the relationship between Spo11 and the Top6A subunit. However, this has not since been shown for any other topo VI enzyme, and therefore attaining evidence of this using MmT6 would solidify the conclusion that this is a hallmark of topo VI activity. For SsT6, little sequence preference could be determined, however cleavage hotspots were clearly visible when SsT6

cleaved pBR322 and a preference for an adenine (A) or thymine (T) within the overhang was detected. This work clearly demonstrated that, whilst robust sequence preference could not be determined, there was some unknown factor predisposing SsT6 to higher levels of activity in precise regions. The hope for SHAN-seq was that it could reveal far more detail on the cleavage activity of MmT6, potentially finding sequence preference and confirming the base stagger. In reality the scope of this work increased to include preliminary exploration into the cleavage activity of *E. coli* DNA topoisomerase IV (EcT4) and *E. coli* DNA gyrase (EcGyrase) as well as MmT6, and also demonstrated its efficacy in exploring DNA cleavage by other enzymes such as restriction enzymes.

SHAN-seq, as previously stated, involves the use of accessible materials and equipment and follows a relatively simple protocol. The assay begins with a typical topo cleavage assay, in which the enzyme of choice is allowed to cleave the DNA, usually under conditions that favour cleavage-complex stabilisation. For example, the conditions used here for MmT6 were 2 mM ADPNP and 10 mM $MgCl_2$, for EcT4 were 2 mM ADPNP and 20 mM $CaCl_2$, and for EcGyrase were 20 mM $CaCl_2$ in the absence of nucleotide, all of which are a deviation from optimal activity conditions and increase the number of cleavage-complexes produced. Therefore, all data reported here involves cleavage in the absence of inhibitors known to stabilise the cleavage complex, and therefore may represent more 'natural' cleavage behaviour. Of course, if drug-dependent cleavage effects were of interest, these conditions could be easily amended to satisfy the research question at hand. This is also true for the DNA substrate used, in which the experimenter has the freedom to control the topology of the DNA used as well as full rein over the DNA sequence. For instance, a suspected cleavage site from other experiments such as *in vivo* NGS assays, could be engineered into the DNA substrate used during SHAN-seq. Here pBR322* was used, to explore the effect of DNA topology on topo-dependent cleavage.

Once the cleavage reaction is complete, the DNA, with the topo still covalently attached, must be sheared to produce 100-500 bp DNA fragments, a prerequisite for most NGS platforms. DNA fragments any longer than 500 bp will begin to impact both the adherence of

DNA to the sequencing chip and the *in situ* PCR reaction. The method of DNA shearing used here was sonication. Optimisation of sonication conditions are necessary to ensure the DNA is fully fragmented, but not over-processed so as to disrupt the covalent DNA-protein interaction. Here, sonication conditions were established which resulted in the vast majority of the DNA substrate being sheared to the appropriate length (Figure 5.2A).
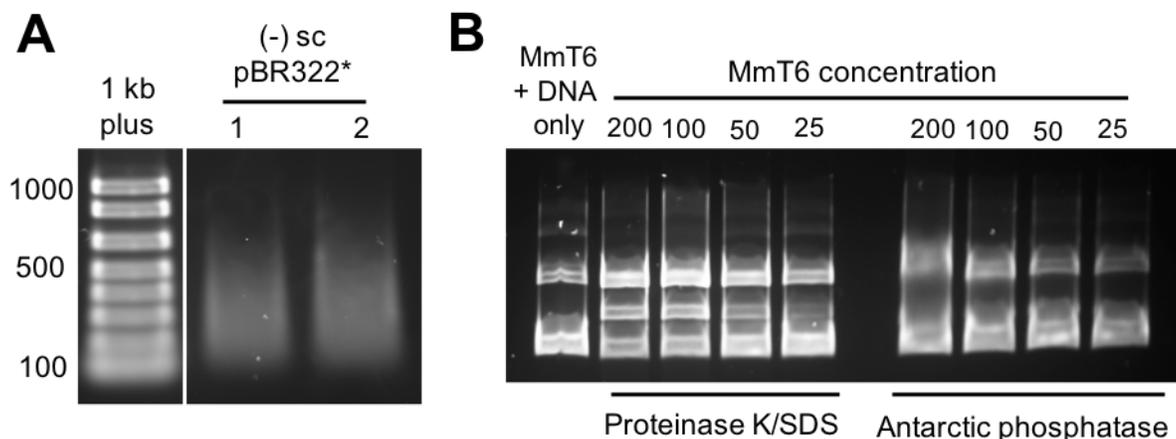
Figure 5.2: Sonication and alkaline phosphatase optimisation. A: Two samples of sonicated 10 nM negatively supercoiled pBR322* and run on a 1% (w/v) agarose gel alongside a 1 kb plus DNA ladder, at 80V for 2 hr. B: Topo VI titration of cleavage activity (MmT6 incubated in the presence of 1 mM ADPNP and 10 mM MgCl$_2$ for 30 min at 37°C) treated with either proteinase K and SDS or Antarctic alkaline phosphatase, and run on a 1% (w/v) agarose gel at 80V for 2 hr.

After sonication, the sample is applied to a commercially available nitrocellulose column (0.45μm nitrocellulose Centrex MF filter, Whatman). As nitrocellulose is negatively charged, DNA is unable to bind, whilst proteins generally bind readily. Therefore, any DNA that is protein-bound will remain embedded within the nitrocellulose membrane, whilst unbound DNA is eluted. This provides an enrichment technique independent of the need for specific protein antibodies, and allows the enrichment within the sample for protein-bound DNA, thereby decreasing the level of noise associated with the sequencing. This technique is potentially cruder than antibody-based enrichment, with an increased likelihood that some nonspecific DNA-protein interactions and DNA associated with contaminating proteins, will also be captured. This is mitigated however by the use of nitrocellulose wash steps and highly pure

protein, respectively, and in general, the results within this chapter indicate that this enrichment technique is effective in capturing topo-cleaved DNA fragments. The protein-bound DNA was eluted using 0.8% sodium dodecyl sulphate (SDS) solution and purified, with the resultant DNA product remaining covalently attached to protein.

The purified DNA was then subjected to the first round of sequencing adapter ligation, which was performed by the NHLBI sequencing core at the National Institutes of Health (Bethesda, USA). This stage is crucial to attaining nucleotide resolution of the topo cleavage site. As some DNA ends in the sample were covalently bound to topo, they were protected from being processed during adapter ligation. The sequencing adapters are important as they are a means to adhering the fragments to the sequencing chip surface, allowing primer binding for both PCR amplification of the library and sequencing, as well as uniquely identifying the DNA samples using DNA indexes. Usually indexing is used to discriminate between two or more assay samples which have been pooled together and sequenced in parallel. In this case, the indexed adapters are being used to discriminate between sonication-dependent and topo-dependent DNA cleavage sites. This is achieved through performing the adapter ligation phase in two stages. The first stage, as the DNA remains covalently bound to protein, results in only protein-free DNA ends being repaired and ligated. The adapters used in Illumina DNA library preparation are called Y-shaped, owing to the fact that each primer contains sequence which allows it to form a duplex with its complementary partner, but also sequence which uniquely maps back to that specific primer. This results in the ability to also discriminate between the top and bottom DNA strands, with the 5$'$ end ligated to a different index to the 3$'$ side. Once the DNA is sequenced, this will produce a data set which can be separated into four parts, each specific to a particular strand of a particular end of the fragment being sequenced. However, before the DNA is ligated to the adapters, it must be blunted using an end repair protocol. The DNA is treated with an enzyme cocktail (T4 DNA polymerase, *E. coli* DNA Polymerase I Klenow fragment and T4 polynucleotide kinase), which resects 3$'$ overhangs and fills in 5$'$ overhangs. This is crucial to resolving the size of the base stagger as a DNA duplex cleaved by topo will produce a cut site in which both sides have a 5$'$ overhang. This means the DNA within the

overhang will be filled in on both sides of the cut site, doubling its presence within fragments, relative to the adjacent nucleotides. It is this sequence enrichment which leads to precise identification of topo cleavage sites, with the number of bases enriched against the background informing on the size of the overhang (Figure 5.3). However, during this first end repair stage, the topo bound DNA ends are not processed, only the ends produced by sonication. Following end repair, blunt ends are phosphorylated, followed by adenylation, which then allows the subsequent ligation of the adapters.

Once the sonicated DNA ends were adapter-ligated, the DNA library was treated with proteinase K and SDS to remove the majority of the protein. Whilst most of the protein is removed in this way, proteinase K cannot hydrolyse the phospho-tyrosyl linkage formed between the 5' phosphate of the DNA and the active site tyrosine of the topo. This adduct still requires removal however, as its presence dramatically reduces adapter ligation efficiency. As mentioned in the introduction, the two enzymatic solutions available for the removal of the phospho-tyrosyl are TDP2 and alkaline phosphatase (specifically calf-intestinal alkaline phosphatase, CIP). TDP2 has evolved to process topo-dependent cleavage complexes *in vivo* and can do so efficiently [338], whilst CIP is primarily a monophosphatase but reportedly retains diphosphatase activity at 3% of its monophosphatase activity [155]. Both these solutions have drawbacks, with TDP2 being far less commercially available and CIP being very inefficient. Regardless of this, CIP was used in an extended protocol (e.g. overnight incubation) to remove the adduct. Initially, Antarctic alkaline phosphatase was used, and during optimisation was found to be ineffective at removing the protein from the DNA in the absence of a preliminary proteinase K/SDS treatment (Figure 5.2B). CIP was then procured for use instead due to its higher activity. Once incubation with CIP was complete, the DNA was returned to the NHLBI sequencing core for the second round of end repair and adapter ligation.
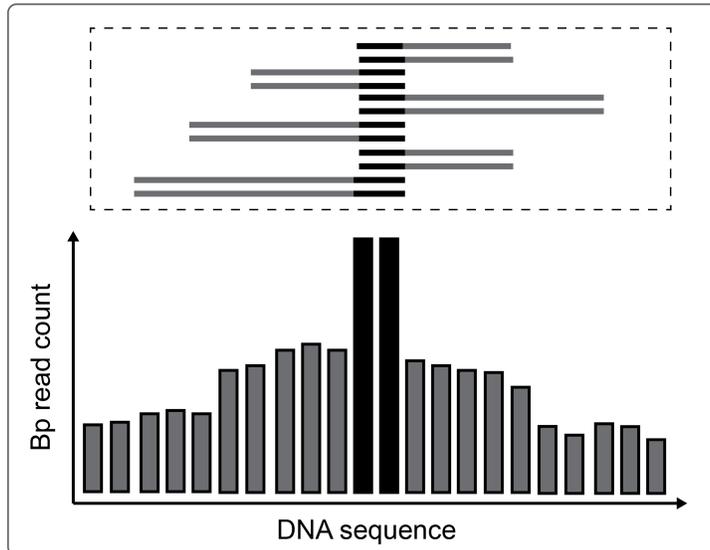
Figure 5.3: Expectation of sequencing data output. Sequencing data is represented as a histogram of how many times a particular nucleotide was sequenced within the DNA library. Due to the 5' overhang type II topos generate on both sides of the cleavage site, this sequence will be enriched (black bars) relative to the nucleotides on either side (grey bars).

Once the DNA ends were no longer protected by topo, they were end repaired and ligated to unique adapters forming the complete DNA library. As described in section 5.1.1, the Illumina platform uses fluorescently-labelled reversible-terminator nucleotides during sequencing. DNA fragments are adhered to the flow cell surface at one extreme end and undergo bridge PCR in which a cluster of each fragment is generated, in order to increase the intensity of the fluorescent signal. Each nucleotide is uniquely coloured allowing its incorporation to be detected and recorded, before the 3'-terminator group and the fluorescent probe are removed and another nucleotide is introduced. Potentially, this allows millions of fragments to be sequenced in parallel. See figure 5.4 for an illustration of the SHAN-seq protocol.
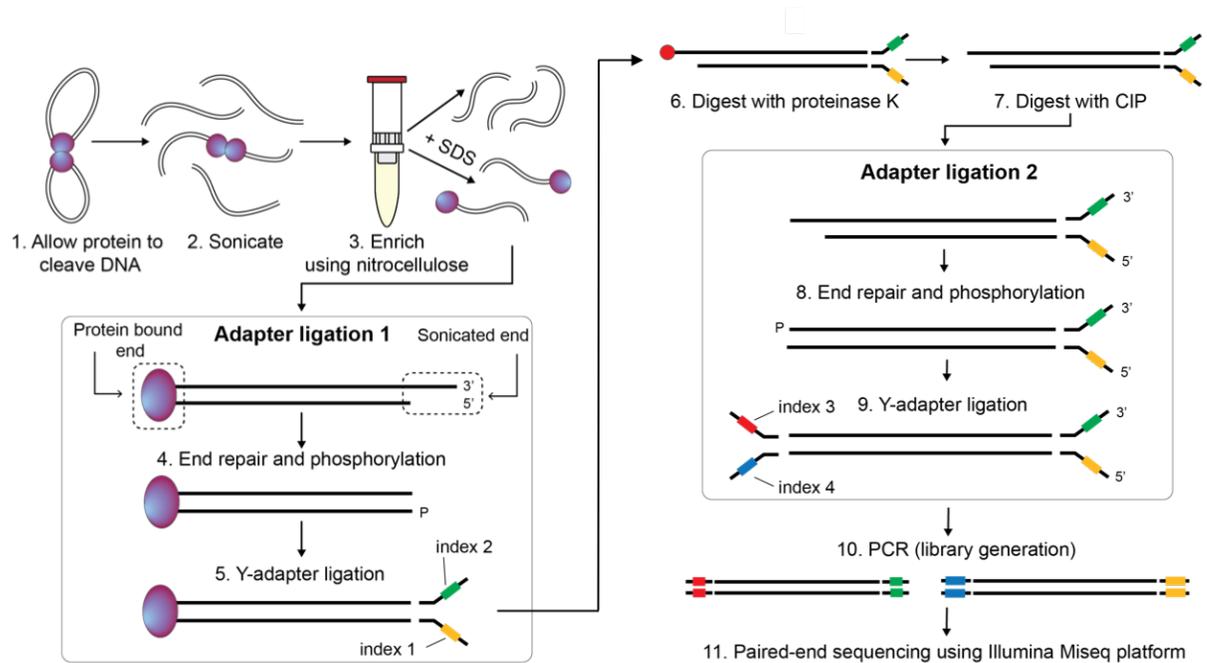
Figure 5.4: SHAN-seq protocol. 1. Topo cleaves the plasmid DNA. 2. DNA with covalently-bound topo is sonicated to shear the plasmid into 100-500 bp fragments. 3. The sample is enriched for topo-bound DNA using a nitrocellulose column. The topo-DNA complexes are eluted from the column using SDS. 4. DNA undergoes the first round of adapter ligation on the exposed DNA ends. This begins with end repair to blunt the DNA, followed by phosphorylation. 5. The DNA is then adenylated (not shown) so that the Y-shaped sequencing adapters can be ligated. The 5' and 3' ends are uniquely barcoded using index 1 (yellow) and 2 (green), respectively. 6. The topo is then removed from the DNA using proteinase K and SDS. This leaves a phospho-tyrosyl adduct attached to the 5' strand. 7. The adduct is removed using calf intestinal alkaline phosphatase (CIP). 8. The DNA is then subject to a second round of end repair and phosphorylation. 9. Following a second adenylation step (not shown) a different Y-shaped adapter is ligated in which the 5' and 3' ends are barcoded with index 3 (red) and 4 (blue) respectively. 10. The DNA library is then produced via PCR to generate 2 pools of fragments, those barcoded with index 1 and 4, and those with index 2 and 3. 11. The DNA is then paired-end sequenced using the Illumina Miseq platform.

5.3   Results and Discussion

5.3.1 Agarose gel-based cleavage patterns

If topos cleave DNA in a preferential manner, this should be observable using *in vitro* ensemble methods, as a reproducible DNA fragmentation pattern, as visualised via electrophoresis. This has been documented in the past for *S. shibatae* topo VI [194], *E. coli* gyrase [89, 357-360], bacteriophage T4 topo II [361], *Drosophila melanogaster* topo II [362, 363], and mammalian

Chapter 5: NGS Results

topo II [364], to name a few. Often no strong consensus sequence is found among the preferential cleavage sites, however there is clearly one or more determinants driving the increased cleavage at certain sites for most topoisomerases. This is also strongly influenced by intercalating drugs, such as oxolinic acid, amsacrine and doxorubicin, which bias cleavage to certain DNA sequences, as seen for gyrase and topo II [339, 365, 366]. In the data presented here, topos are stimulated to form cleavage complexes in the absence of specific drugs, with the goal to attain a representation of the enzymes more 'natural' cleavage activity. This was achieved by using ADPNP, the non-hydrolysable ATP analogue, for *M. mazei* topo VI (MmT6) and *E. coli* topo IV (EcT4), and replacing $MgCl_2$ with 20 mM $CaCl_2$ for EcT4 and *E. coli* DNA gyrase (EcGyrase). Each topo was incubated with pBR322*, which was either linear, negatively supercoiled, positively supercoiled or relaxed. This was performed at relatively high topo concentrations to capture as many cleavage complexes as possible (MmT6 - 500 nM, EcT4 - 150 nM and EcGyrase - 64 nM). The pBR322* was then linearised using an EcoRI digestion, with the linear pBR322* being the exception due to previous EcoRI-digestion prior to the topo-dependent cleavage assay. As there is only a single EcoRI site in pBR322*, EcoRI will linearise any uncleaved pBR322* and produce two or more fragments on those which have been cleaved by topo. The DNA samples were run on a 1% (w/v) agarose gel, loading 200-400 ng of DNA in order to maximise the likelihood of clearly resolving the cleavage pattern (Figure 5.5). All three topos fragmented the DNA in a reproducible pattern, unique to each topo and in some cases preliminary evidence of topology-dependent cleavage effects can be seen.

DNA cleavage by MmT6 clearly demonstrates its preference for particular DNA sites with bright DNA fragment bands appearing in all three samples of the same size, regardless of the DNA topology (Figure 5.5A). Interestingly however, there is a visual difference dependent on the DNA topology. For negatively and positively supercoiled DNA the appearance of the DNA fragments is far more 'smeared' than it is for the linear sample. This could indicate that sequence preference is more important for stimulating cleavage when the DNA is linear. As discussed in chapters 3 and 4, MmT6 is preferentially attracted to DNA crossings, cleaving writhed DNA with a far higher activity than linear, an effect also documented elsewhere [62].

Therefore the 'smeared' appearance of the supercoiled DNA samples could indicate that as writhe increases, the necessity for cleavage at specific sites decreases. It may also indicate, more simply, that the level of cleavage is increased in general with increasing writhe, however the preferential sites do not visually appear brighter than in the linear sample.
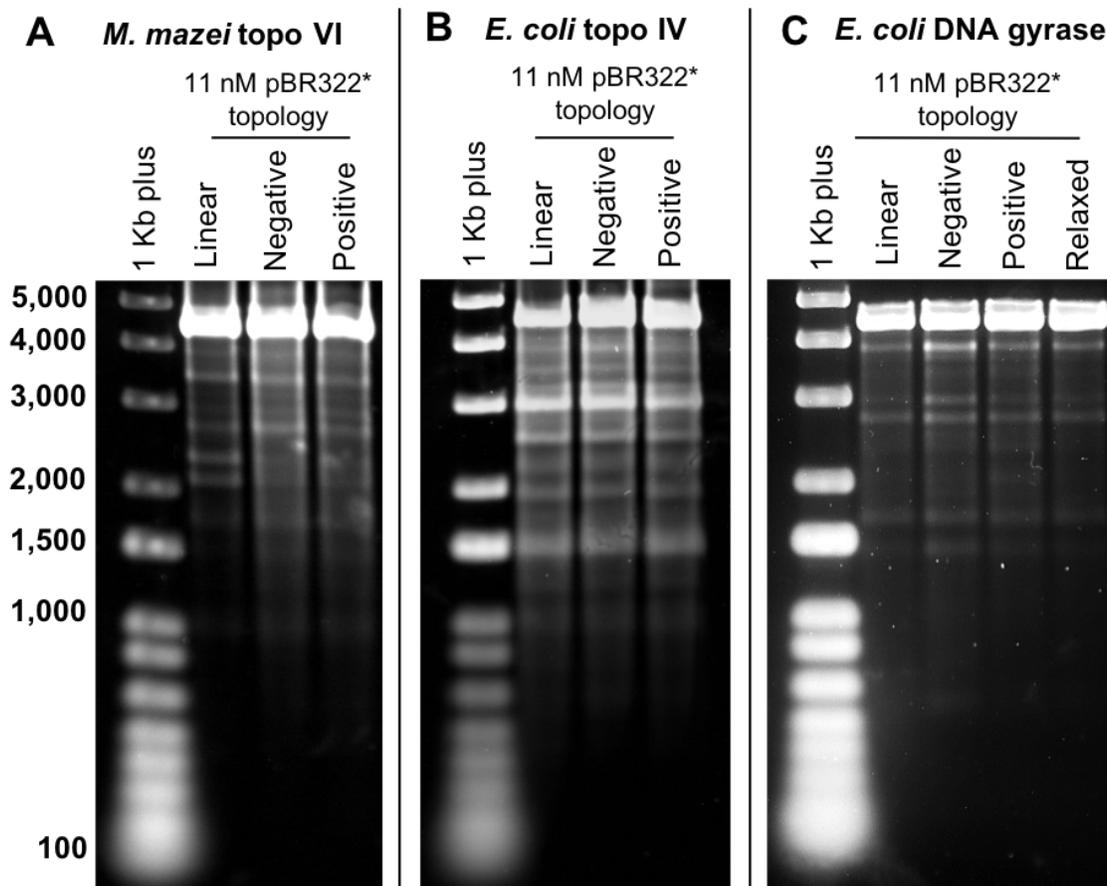
Figure 5.5: Agarose gel-based cleavage patterns produced by type II topos. A: Cleavage activity of 500 nM *M. mazei* topo VI in the presence of 11 nM of linear, negatively supercoiled and positively supercoiled pBR322*. B: Cleavage activity of 150 nM *E. coli* topo IV in the presence of 11 nM of linear, negatively supercoiled and positively supercoiled pBR322*. C: Cleavage activity of 64 nM *E. coli* gyrase in the presence of 11 nM of linear, negatively supercoiled, positively supercoiled and relaxed pBR322*. All assays run on a 1% (w/v) agarose gel.

DNA cleavage by EcT4, also forms a reproducible DNA fragmentation pattern, which is insensitive to DNA topology (Figure 5.5B). Out of the three topos characterised in this work, EcT4 demonstrates the most cleavage and generates the most complex cleavage pattern. This activity also appears to be independent of DNA topology with each lane being visually identical.

Chapter 5: NGS Results

One particularly strong site stands out at ~3000 bp, with another at ~1500 bp, the latter presumably being the partner fragment to the former (full length pBR322* is ~4.4 kb).

DNA cleavage by EcGyrase demonstrated an unexpected result, with one cleavage site being significantly more represented in the sample which contained negatively supercoiled pBR322*. This indicates there is a topology-sensitive EcGyrase cleavage site that is cleaved more efficiently when the DNA is underwound. This is intriguing as it indicates that gyrase cleaves its product (negatively supercoiled) more readily than its substrate (positively supercoiled or relaxed). In addition, the previously recorded 'strong gyrase site' (SGS) in pBR322, at position 990, is apparently absent in the data presented here. Much of the work done previously using EcGyrase, focused on cleavage site preferences using quinolone-derived inhibitors [275, 360]. These inhibitors function by intercalating between the topo-dependent DNA breaks and preventing re-ligation, thereby increasing the number of stalled cleavage complexes and killing cells through a disruption in DNA replication and transcription [367, 368]. Multiple avenues of evidence suggest that using drugs to explore DNA cleavage by topos strongly biases the cleavage specificity [339, 360, 369, 370]. For instance, quinolone-derived drugs (oxolinic acid and ciprofloxacin) demonstrate a clear preference for cleavage at G residues, while Microcin B-17, another gyrase inhibitor, promotes DNA cleavage with an alternative and far more degenerate sequence preference [339]. This may explain why the SGS of pBR322* has not been cleaved in this work, as no quinolone inhibitor was used. This provides further evidence of how dramatically certain drugs can bias cleavage, with their use producing data which may poorly represent the 'natural' cleavage by the enzyme. As most NGS studies require the use of inhibitors to stimulate topo-specific cleavage *in vivo* (discussed in the 5.1), this calls into question how much this data can truly inform on the normal activity of topos, with the inhibitors used potentially disrupting this behaviour.

Whilst it is not possible to determine the precise cleavage sites using this method alone, this data provided both encouragement that the three topos tested cleave pBR322* preferentially and that this should be evident in the sequencing, but also provided an excellent reference point for comparing to the NGS data.

5.3.2 NGS controls

To test whether SHAN-seq was capable of determining cleavage sites and overhangs using plasmid DNA *in vitro*, a restriction enzyme control was conducted. Like topos, restriction enzymes also generate DSBs, and in some cases form 5' overhangs. Restriction enzymes are distinct however, in that cleavage is irreversible, the protein does not form a covalent interaction with the DNA back bone, and in many cases the sequence specificity is extremely strict. The enzyme chosen here was MluCI, which cleaves pBR322* seven times at the sequence AATT, generating 4-base 5' overhangs. Linear pBR322* was incubated with MluCI and then sonicated. As the enzyme does not remain covalently bound to the DNA, the sample cannot be enriched using the nitrocellulose columns, nor can the two extreme ends of the DNA be uniquely indexed, meaning the sample is end-repaired and adapter ligated in one stage prior to sequencing. Along with the MluCI positive control, a DNA-only negative control was also performed in which 10 nM linear pBR322* was sonicated and sequenced. This was important as it is necessary to know how the plasmid fragments due to sonication in the absence of protein (Figure 5.6A). Generally speaking, sonication will fragment the DNA randomly but AT-rich areas are known to be more sensitive to shearing [371].

The NGS results clearly demonstrated all seven of the MluCI restriction sites (Figure 5.6B), appearing precisely where they were predicted to. At three positions, the overhang is not present due to the fragment size being below the lower sequencing size limit. For instance, the site at bp 59 would generate a 59 bp fragment, and the two sites at bp 1320 and 1334 (Figure 5.6C) would produce a 10 bp fragment between them, both of which are below 100-500 bp size recommendation and unlikely to be sequenced. However, this small caveat (e.g. the overhang enrichment may not be resolvable if cleavage sites are too numerous or close together) also demonstrated that despite this, the peaks were found precisely where expected. In addition, four of the seven sites did demonstrate a 4-base enrichment at the AATT sequence (Figure 5.6D), validating that this method can determine both sequence preference and overhang size. It can also inform on the type of overhang generated, 5' or 3'. This is due to the

end-repair stage of adapter ligation. As previously mentioned, the enzyme mixture used to blunt the DNA resects 3ʹ overhangs and fills in 5ʹ overhangs. Whilst this has the effect of doubling the sequence within 5ʹ overhangs, this sequence is removed from 3ʹ overhangs and would be discernible in the NGS data as a loss of sequence corresponding to the size of the overhang, rather than the enrichment seen here (Figure 5.6D).

In addition to validating that SHAN-seq performs as intended, this control also demonstrated that it has value in the characterisation of other types of DSB-inducing enzymes, such as restriction enzymes, making it a broadly applicable and versatile technique in the field of DNA-protein interactions.
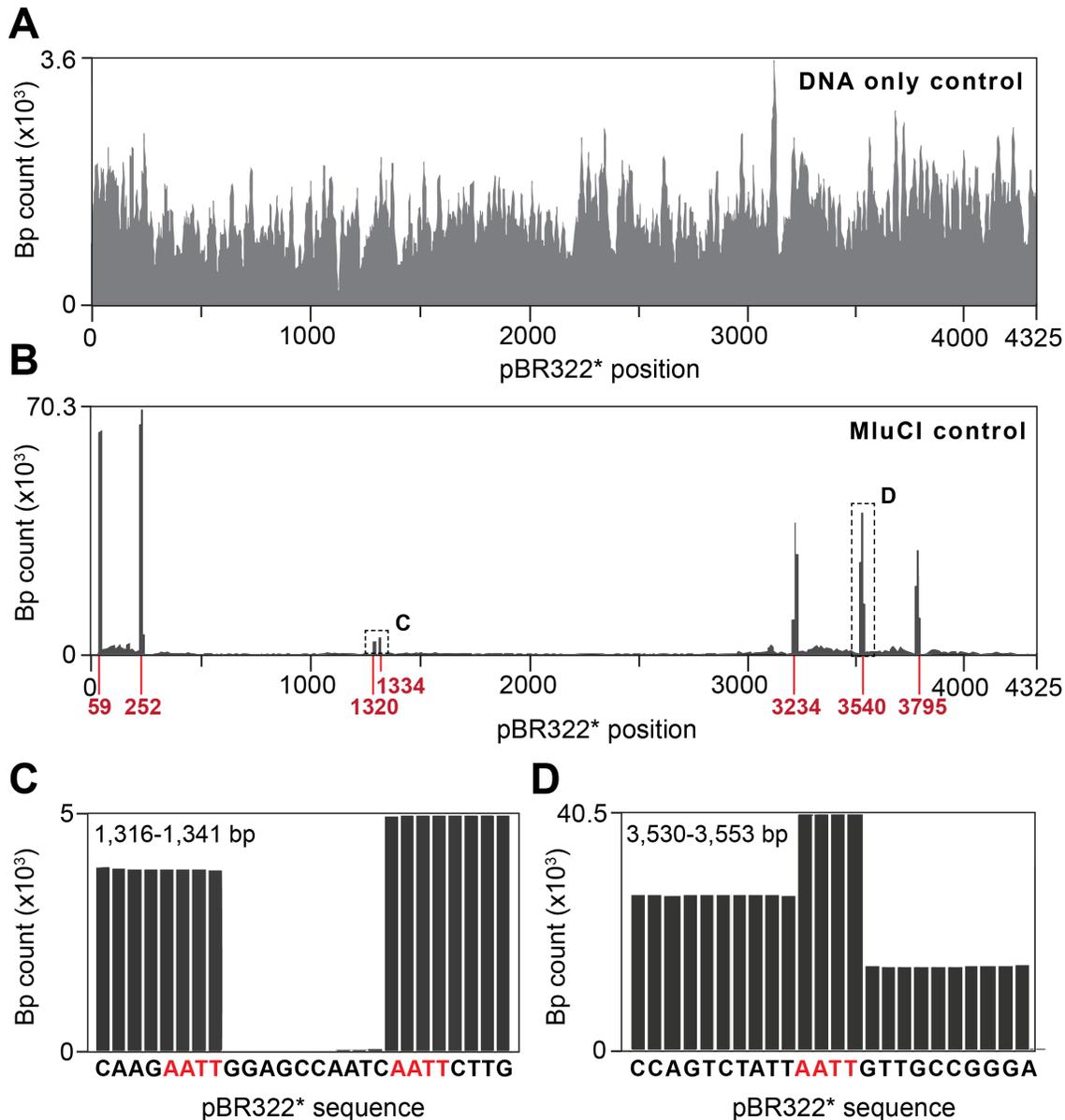
Figure 5.6: SHAN-seq NGS controls. For all plots, data is displayed as nucleotide (bp) read count at each position along the pBR322* sequence. A: Negatively supercoiled pBR322* sonicated and sequenced using Illumina Miseq PE-75. B: Linear pBR322* cleaved by MluCI, sonicated and sequenced using Illumina Miseq PE-75. The red bars and numbers indicate the predicted MluCI restriction sites. C: Close up view of the peaks between 1,316 and 1,341 bp. D: Close up view of the peak at 3,530-3553 bp. The sequence in red in parts C and D is the MluCI cleavage specificity (AATT).

5.3.3 DNA Topoisomerase VI NGS Results

DNA topo VI from *M. mazei* (MmT6), was both the focus of this research and the incentive behind the development of this technique. As previously remarked, exploration into the sequence preference and cleavage overhang size for topo VI has only been conducted once before, using the *S. shibatae* enzyme [194]. It was found that topo VI produced DSBs in the presence of ATP, generating 2-base overhangs with no strong sequence preference beyond an increased probability of an A or T being present in the overhang. Expanding on this preliminary foray into topo VI cleavage behaviour was an important aspect of the work reported here.

MmT6 (500 nM) was incubated with 10 nM negatively supercoiled pBR322* in the presence of ADPNP and 10 mM MgCl$_2$, conditions which were established as optimal for cleavage in Chapter 3. As described in Section 5.2, sequencing adapter ligation of DNA cleaved by MmT6 occurred in two phases with the intention being the unique indexation of the topo cleavage sites, allowing the data to be separated from DNA breaks produced by sonication. This proved highly effective, as not only was the resolution of MmT6-dependent cleavage sites facilitated, but this also highlighted some current drawbacks, discussed below, associated with the DNA preparation phase that will be improved upon in future work (see section 5.5).

The NGS data for MmT6 cleavage showed clear peaks across pBR322* (Figure 5.7), in comparison to the DNA-only control (Figure 5.6A). As seen in figure 5.7A, there were two peaks in the same position for all DNA strands sequenced, at positions ~2.3 and ~3.4 kb. This was not predicted as the 5' and 3' sonicated strands (yellow and green respectively, see Figure

5.7) should have a very different appearance to the 5´ and 3´ MmT6-cleaved strands (red and blue respectively). There are, however, a number of explanations to account for this result.
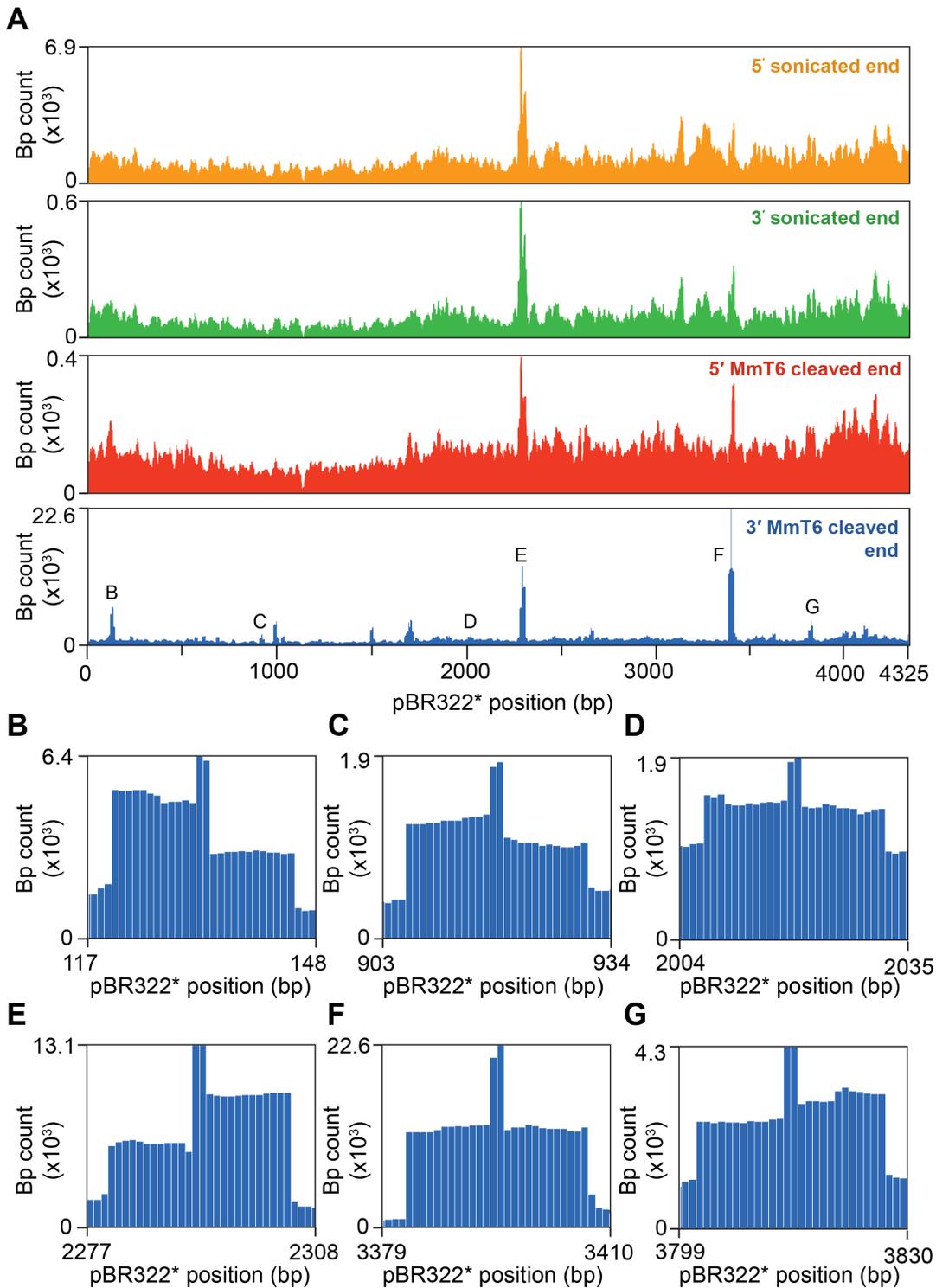
Figure 5.7: NGS data for *M. mazei* topo VI (MmT6) cleavage activity. All data are plotted as nucleotide read count (read depth) for each position along pBR322*. A: NGS data for MmT6 cleavage, plotted in regard to the strand sequenced, 5´ sonicated end (yellow), the 3´ sonicated strand (green), the 5´ MmT6 cleaved (red) end and the 3´ MmT6 cleaved end (blue). On the blue

plot, the peaks enlarged in parts B-G are labelled with their respective letter. B-G: Close up images of the MmT6 peaks, as labelled in A, showing the characteristic peak shape indicative of MmT6 cleavage activity, which appears insensitive to read depth.

If it is assumed that peaks associated with MmT6-cleaved ends are true MmT6 cleavage sites, the peaks in the same positions in the sonicated-end data may have been adapter-ligated during the first round, despite originating from MmT6 cleavage. Reasons for the DNA end no longer being protected by covalently bound MmT6 include MmT6-induced G-segment bending causing structural vulnerabilities within the DNA, therefore sensitising it to sonication near the cleavage site. Another possibility is the conversion of MmT6-induced nicks into double-stranded breaks (DSBs) during sonication. As discussed in Chapter 3, ADPNP stimulates MmT6 to cause significant levels of nicks, along with DSBs, and it has been observed that DNA is more likely to shear within a 10 bp radius of pre-existing nicks [372].

Despite the presence of two peaks (positions ~2.3 and ~3.4 kb) in all data sets, the data associated with the MmT6-cleaved strands has revealed both new information on MmT6 cleavage activity, but also highlighted a potential weakness in the SHAN-seq protocol. The 5$^{'}$ and 3$^{'}$ MmT6-cleaved NGS data are quite distinct form one another, despite sharing the same origin. Firstly, the depth of sequencing is 56-fold higher for the 3$^{'}$ strand than the 5$^{'}$ strand (Figure 5.7A, blue and red plots, respectively). The most likely cause for this lies in the use of CIP to remove the phospho-tyrosyl adduct. Whilst reportedly CIP has diphosphatase activity at 3% of the canonical monophosphatase activity [155], it appears likely that the adduct is not being removed efficiently and is therefore interfering with adapter ligation/indexing the 5$^{'}$ strand of the MmT6 cleavage sites.

Sequencing the 3$^{'}$ strand from the MmT6 cut site, which is unmodified during cleavage, revealed visually obvious peaks with a highly characteristic shape (Figure 5.7A, blue plot). For close-ups of these peaks see Figure 5.7B-G. Across the pBR322* sequence, 15 'strong' and 14 'weak' MmT6 cleavage sites were identified using visual inspection based on a peak shape and sequencing depth. The peaks were 26 nucleotides in width with nucleotides 13 and 14 enriched relative to the twelve on either side. The read depth of nucleotides 13 and 14 was

roughly the sum of the read depth of the left twelve nucleotides and of those on the right. This originates from each sequence being cropped to 14 bp in length prior to alignment to the reference genome. As expected from MmT6 cleavage, the sequence in the overhang will be enriched relative to the sequence directly surrounding it, and this will inform on the size of the overhang, as discussed in Sections 5.2 and 5.3.2. In the case of MmT6 the width of this enrichment is two nucleotides, therefore indicating that MmT6 generates DSB with 2-base overhangs. This is exactly in line with expectations for MmT6, as both *S. shibatae* topo VI and Spo11 have also been confirmed to produce DSBs with 2 bp overhangs [194, 205]. This result helps to cement the conclusion that, in general, topo VI enzymes cleave DNA, generating 2-base overhangs, and this forms a critical distinction from the type IIA topos.

To attain further confidence that the peaks seen within the NGS data are indeed MmT6 cleavage sites, the NGS data set was compared to the gel-based data from section 5.3.1. The size of the fragments seen on the gel were compared to the 'theoretical' fragment sizes which would be produced if MmT6 were to cleave linear pBR322* at the site of each peak. For instance, in the NGS data there was a peak at nucleotide 3394, that if cleaved would form two fragments, 3394 and 931 bp in size. In Figure 5.8, The NGS data is plotted alongside the gel-based data. Stronger cleavage sites, identified on the NGS data as 'a' and 'b', are coloured dark grey for the NGS and gel-based data plot. The NGS strong sites were determined visually, with the two sites (positions ~2.3 and ~3.4 kb) chosen based on a read depth of 10,000 or
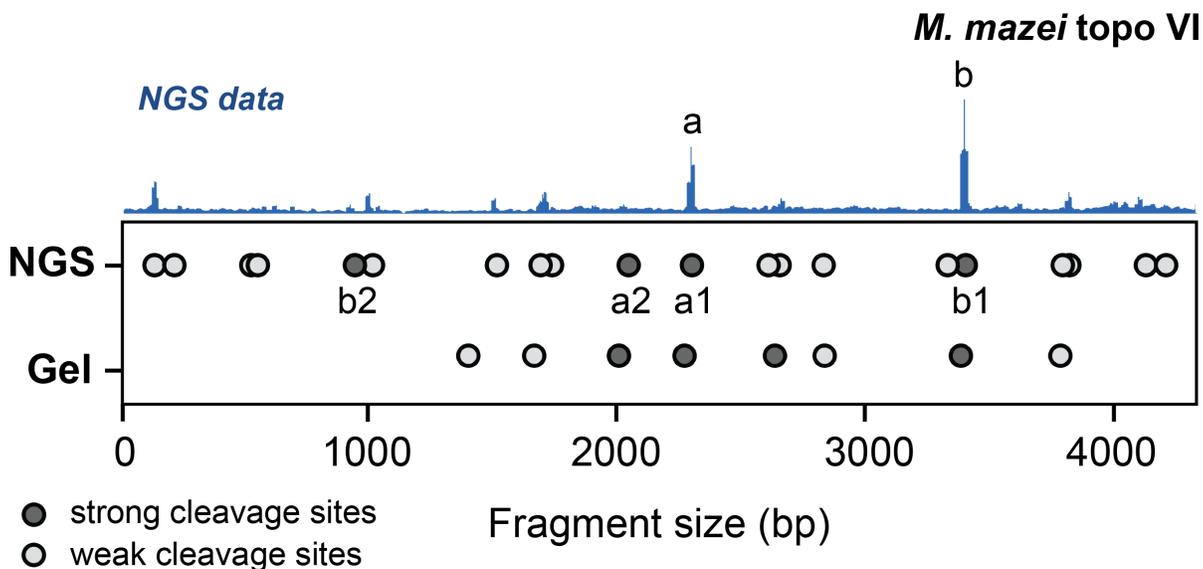
Figure 5.8: NGS data versus gel-based data for MmT6-dependent fragmentation of pBR322*. The NGS data is on top in blue. The strong cleavage sites are labelled as 'a' and 'b'. Below, the fragment sizes predicted by the sequencing were compared to those extracted from the gel-based assay. Strong cleavage sites (large peaks in NGS data/bright cleavage bands in gel-based data) are dark grey, while other weaker cleavage sites are light grey. For the NGS fragment sizes, the strong sites are labelled with 1 or 2 to demonstrate they are partners with both being formed should MmT6 cleave linearised pBR322* at the position indicated on the NGS data.

above. As cleavage at these sites would produce two fragments, the data points are labelled as 1 or 2, to demonstrate they are partners. All sites derived from the NGS data are plotted as partner fragments. This is not possible for the gel-based data, and instead the fragments which are clearly resolvable are plotted, with brighter bands being categorised as 'strong' cleavage sites. By comparing these two data sets, further confidence in the validity of the NGS data is obtained as many strong cleavage sites within the NGS data match well to the fragment sizes extracted from analysis of the gel in section 5.3.1. In some cases, strong cleavage sites seen in the NGS data correspond well with brighter bands seen on the gel (see peaks 'a' and 'b' in Figure 5.8). Also, a bright band on the gel corresponds to a fragment ~2.8 kb in size, with the sequencing indicating that was due to the combined effect of cleavage at two weaker sites which both form a 2.8-kb fragment.
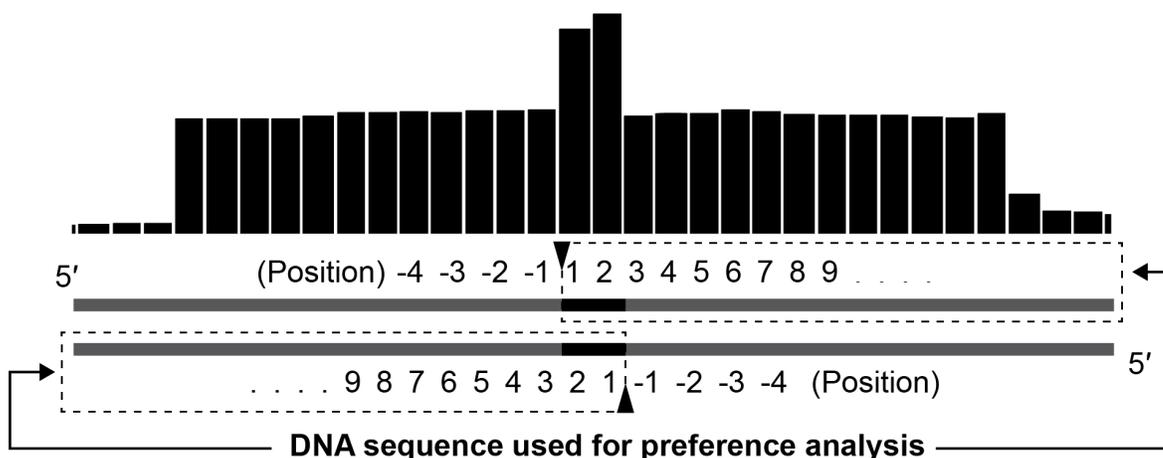
Figure 5.9: Illustration of which pBR322* sequences were extracted from each peak on the MmT6 NGS data to use in sequence preference analysis. The nucleotide numbering is used in Figure 5.10 which analyses sequence identity of positions 1-35. The black arrows indicate the scissile phosphate, immediately upstream of nucleotide 1.

Chapter 5: NGS Results

With further confidence that the NGS data of the 3$'$ MmT6-cleavaed strand was reflective of true MmT6 cleavage activity, analysis of the sequence at each of these sites was permitted. To see a full list of the sequences used in the analysis, see Table II in the Appendix III. As the 2-base enrichment is present, it is possible to extract the sequence at the precise site of cleavage and this was done for the 5$'$ sequence immediately downstream of the cleavage site on the top and bottom strands (see Figure 5.9). Using the browser-based software WebLogo [228], the sequence present at the strong and weak cleavage sites were analysed for nucleotide indentity similarities at each position, both separately and together. No strict consensus was found, but there was a clear preference for a C in position 4 and for a G/C in position 3 (which corresponded to a increased preference for a C/G in position -1 and for G in position -2 on the opposite strand, see Figure 5.9). This is particularly clear for the strong sites (Figure 5.10A), with this preference present but less significant for the weak sites (Figure 5.10B), and therefore also diltued when strong and weak sites are analysed toegther (Figure 5.10C). This may explain why sites analysed in Figure 5.10B are cleaved less frequently than those in Figure 5.10A. Another difference is that an A/T identity is common in position 8 for strong sites, however this is not the case for the weak sites and could be another factor causing less cleavage at these sites. Overall, the sequence preference is weak, yet the cleavage sites seem precise. This indicates that there may be a yet unknown feature that unites these sites which has yet to be analysed. For instance, there is evidence that type II topos, including topo VI, can bend the G-segment [61], and so it is plausible that MmT6 cleavage site sequences may represent DNA that is more easily bent. As MmT6 has been shown to bind the G- and T-segment as a DNA crossing and that the T-segment interacts with Top6B [62], this form of analysis overlooks the possibilty of T-segment binding preference. Future work could include simulations of supercoiled pBR322* to gain perspective on probable T segment DNA regions that interact with a given G-segment.
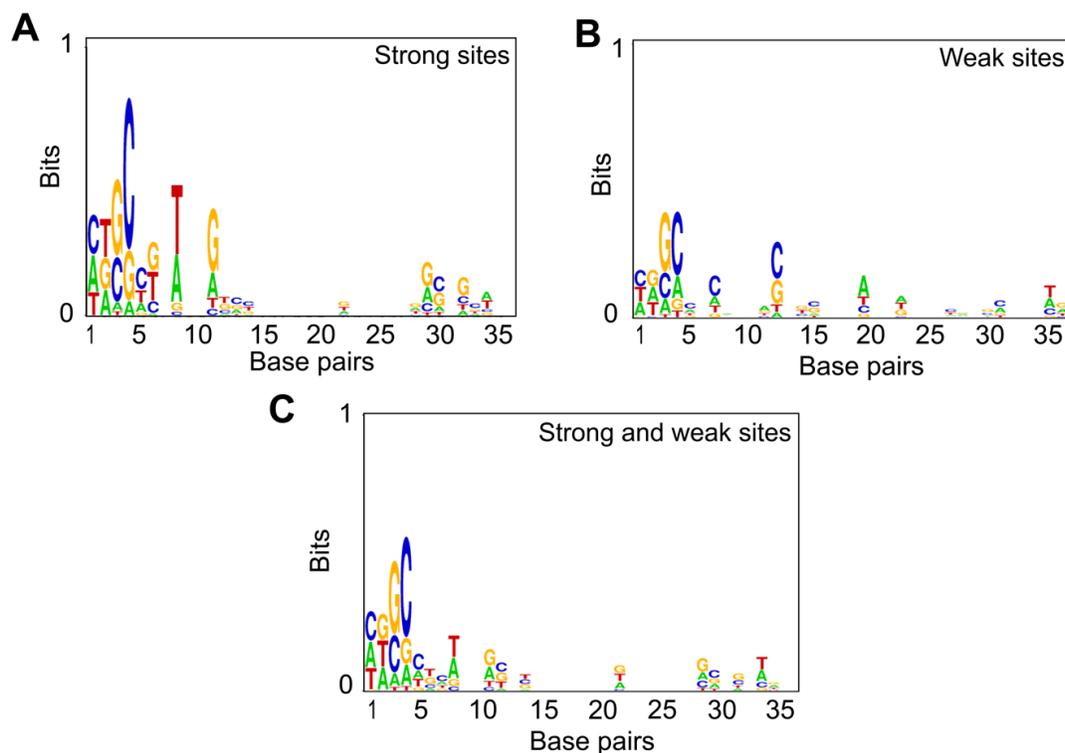
Figure 5.10: Sequence preference analysis of MmT6 cleavage sites using WebLogo. A: Sequence preference analysis for strong MmT6 cleavage sites. B: Sequence preference analysis for weak MmT6 cleavage sites. C: Sequence preference analysis for strong and weak MmT6 cleavage sites. The cleavage occurs just upstream of nucleotide 1, see Figure 5.9.

When DNA cleavage characterisation was conducted using *S. shibatae* topo VI (SsT6), the enzyme was found to generate a 2-base stagger, like MmT6, but no sequence preference was found aside from an increased probability of A or T nucleotides present in the overhang [194]. Using MmT6 data, this result was somewhat supported by the finding that 66% of nucleotides, positioned 5' of the cut site, were an A or a T. This work also found SsT6 had a strong cleavage site around the SspI restriction site (position 4168 on EcoRI-linerised pBR322*). Even though this does not correspond to either of the strong cleavage sites identified here, cleavage activity is detected in this region for MmT6, both upstream and downstream of the SspI restriction site. However, it is not strong cleavage as measured using the NGS and gel-based data. There are a number of explanations for why the MmT6 data generated in this current work, does not match up well with work done on the SsT6. MmT6 and SsT6 are different enzymes encoded for by different archaeal organisms, which live in extremely different

Chapter 5: NGS Results

conditions. *M. mazei* is a mesophilic archaeon, living in freshwater and marine environments, with an optimal temperature and pH of 30-40°C and 6.8-7.2, respectively [238, 239]. *S. shibatae* is a hyperthermophillic organism isolated from acidic high temperature environments with an optimal temperature and pH of 70-75°C and 2-3, respectively [373]. These wildly different environments will no doubt have had consequences for the adaptation of the intracellular proteins in each of these organisms. This is also evident biochemically, as SsT6 was found to perform robust DNA cleavage in the presence of $CaCl_2$, whilst MmT6 activity in the same conditions is barely detectable (see Chapter 3). Another explanation for the differences between MmT6 and SsT6 DNA cleavage activity could be due to the different DNA topology they likely interact with. *M. mazei* encodes a DNA gyrase and the genome is negatively supercoiled, whilst *S. shibatae* encodes a reverse gyrase and likely has a relaxed or positively supercoiled genome [374]. This data warns against taking information attained from the characterisation of a single species and applying it more broadly.

Sequence preference analysis for *S. cerevisiae* Spo11 found cleavage is promoted 3' of C nucleotides and disfavoured 3' of G [205, 375] with the dinucleotides cleaved most often being CT, CA, TA and CC. A similar form of analysis was conducted for the MmT6 cleavage sites (Figures 5.11 and 5.12). The nucleotides immediately upstream and downstream of the cut site were labelled -1 and 1, respectively. As evident in the previous analysis (Figure 5,10), the nucleotide in position -2 was found to exhibit nucleotide identity preference and so was also analysed (Figure 5.11A). The most common nucleotide in position -2 for MmT6 cleavage was G (66%) followed by C (17%), T (8%), and A (3%) (Figure 5.11B). In position -1 the most frequently observed nucleotide was C (59%), followed by G (31%), T (7%), and A (3%) (Figure 5.11C). This means that cleavage is also promoted 3' of C nucleotides, as had been found for Spo11 [205]. The difference however, is that MmT6 cleavage is not disfavoured 3' of G nucleotides, with G being the second most frequently observed in position -1, unlike Spo11 which disfavours G [205, 375]. The identity of the dinucleotides at position -2 and -1 can be seen in Figure 5.12A with the 3 most common sites being GC, GG and CC. Interestingly, for MmT6, G nucleotides are completely absent from position 1, a cleavage behaviour also seen

for Spo11 [205] (Figure 5.11D). And in addition, the identity of the dinucleotides which were cleaved is very similar to Spo11, with those more likely cleaved by MmT6 being CC, CA, CT
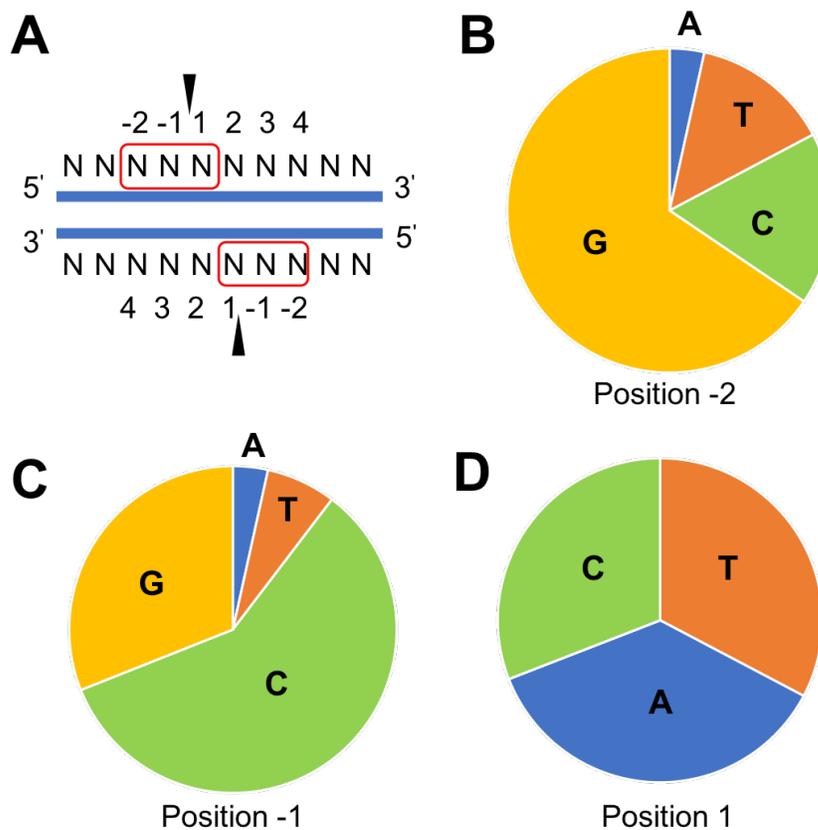
Figure 5.11: Sequence analysis of the precise MmT6 cleavage site. A: Position numbering of the nucleotides 2 bp upstream and 1 bp downstream of the cleavage site. The scissile bound is denoted as the two black triangles. B: Nucleotide identity as a percentage (%) for position -2. C: Nucleotide identity as a percentage (%) for position -1. D: Nucleotide identity as a percentage (%) for position 1. In all charts G is yellow, C is green, T is orange and A is blue.

and GA (Spo11 cleavage sites: CT, CA, TA and CC [205]) (Figure 5.12B). The sequence preference derived for *S. cerevisiae* Spo11 was 5$^{'}$-C^[A/C/T] (the '^' symbol denoting the scissile phosphate). For MmT6 the sequence preference was found to be 5$^{'}$-GC^[A/C/T], or perhaps more reflectively denoted as 5$^{'}$-[G/C][G/C]^[A/C/T]. The similarities between these preference sequences could indicate DNA binding is similar between Top6A and Spo11. However, the difference regarding G-nucleotide occupancy speaks to their evolutionary distance, being two distinct enzymes performing different cellular activities. Recent molecular characterisation of the *S. cerevisiae* Spo11 complex, also explored the change in Spo11

cleavage behaviour *in vivo* as a consequence of the F260A point mutation [210], an amino acid predicted to be within the DNA binding channel [376]. This mutation disrupted the wildtype dinucleotide preference for cleavage, reducing the frequency of C in position -1, whilst also reducing the suppression of G in this position. This preliminary result offers intriguing insight into the potential amino acids which modulate DNA sequence preferences for both Spo11 and Topo VI.
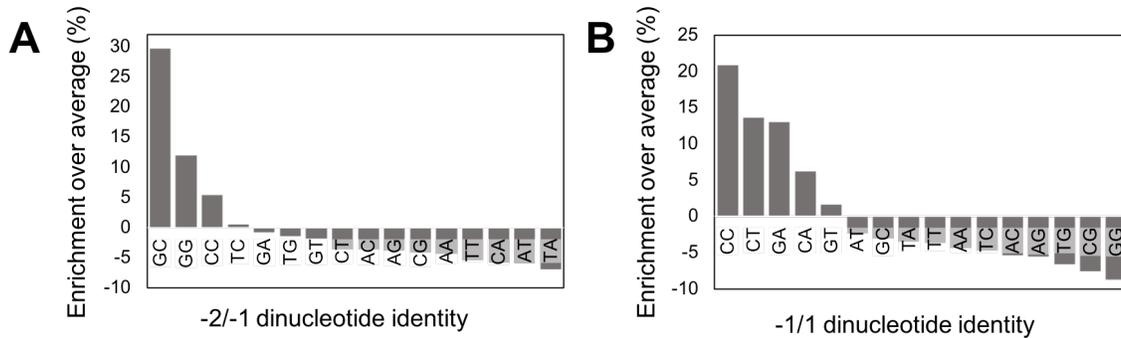
Figure 5.12: MmT6 cleavage site dinucleotide identity preference. A: Dinucleotide identity as a percentage (%) for positions -2 and -1. B: Dinucleotide identity as a percentage (%) for positions -1 and 1. Nucleotide position numbering as seen in Figure 5.11A. All data points have been corrected for the average dinucleotide enrichment of pBR322*.

5.3.4 DNA Topoisomerase IV NGS Results

Topo IV from *E. coli* (EcT4) was the second topo to be explored using SHAN-seq. It is a type IIA topo, which produces DSBs with 4-base overhangs. As evident from the gel in section 5.3.1, in the presence of 2 mM ADPNP and 20 mM $CaCl_2$ (instead of the optimal ATP and $MgCl_2$), many EcT4-dependent cleavage complexes are formed on pBR322* in a manner seemingly insensitive to DNA topology. In order to gain insight into the sequence preference of topo IV at these sites, SHAN-seq was used.

In total, 4 strong sites (read depth ≥ 700 bp count) and 11 weak sites (read depth < 700 bp count) were identified visually based on a characteristic peak shape seen in the EcT4 data and absent in the DNA only control (Figure 5.13). All peaks were present in each sample, insensitive to the DNA topology (Figure 5.13A), which is in agreement with the gel-based data (see Figure 5.5). The characteristic peak shape for EcT4 data was formed from sequence

enrichment spanning ~60-70 bp, with a bimodal peak separated by a central reduction in read depth, spanning ~15-20 bp (Figure 5.13B-D). Whilst this allowed the identification of EcT4 activity, the shape formed precluded the ability to precisely map the cleavage site and therefore the nucleotide composition at the scissile phosphate or the size of the overhang. One explanation for this outcome was the use of sonication breaking the DNA around EcT4 binding sites, potentially due to DNA deformation upon binding producing structural vulnerabilities. This would therefore lead to enrichment around the cleavage site, but sequence loss at the precise location. Another explanation is CIP not efficiently removing the phospho-tyrosyl adduct, which goes on to disrupt adapter ligation of sequences from the cleavage site. Interestingly, a similar peak shape indicative of EcT4 cleavage activity was found using an *in vivo* NGS-technique called NorFliP [125, 341]. In this study, EcT4 activity was studied genome-wide using the cleavage-complex stabiliser norfloxacin, forming two peaks ~170 bp wide with a central sequencing loss of 130 bp. This was attributed to the $5^{'}$ adduct interfering with adapter ligation as in this study only proteinase K was used, with no solution offered in the paper for the adducts' removal. Regardless, EcT4 was shown to generate hundreds of cleavage sites across the genome, consistent with data presented here indicating that of the three topos assayed, T4 maintains the highest levels of cleavage.

To gain further confidence that the peaks seen within the NGS data correspond to bona fide EcT4 cleavage sites, the NGS data was compared to the gel-based data. As described for MmT6 analysis (section 5.3.3), pBR322* fragment sizes for cleavage at each peak on the NGS data were compared to the fragments seen on the gel with stronger sites labelled on the NGS data and represented as dark grey circles on the graph (Figure 5.14). For the NGS data, strong sites were identified based on a read depth count of ≥700, whilst for the gel they were identified based on the band brightness. As for MmT6, the correlation between the two data sets is very good, with strongly enriched peaks corresponding to bright bands seen on the gel, and an overall agreement between all data points. So, despite not having resolved the cleavage site or seen the predicted 4-base stagger enrichment, the peaks within the NGS data are likely valid EcT4 cleavage sites.
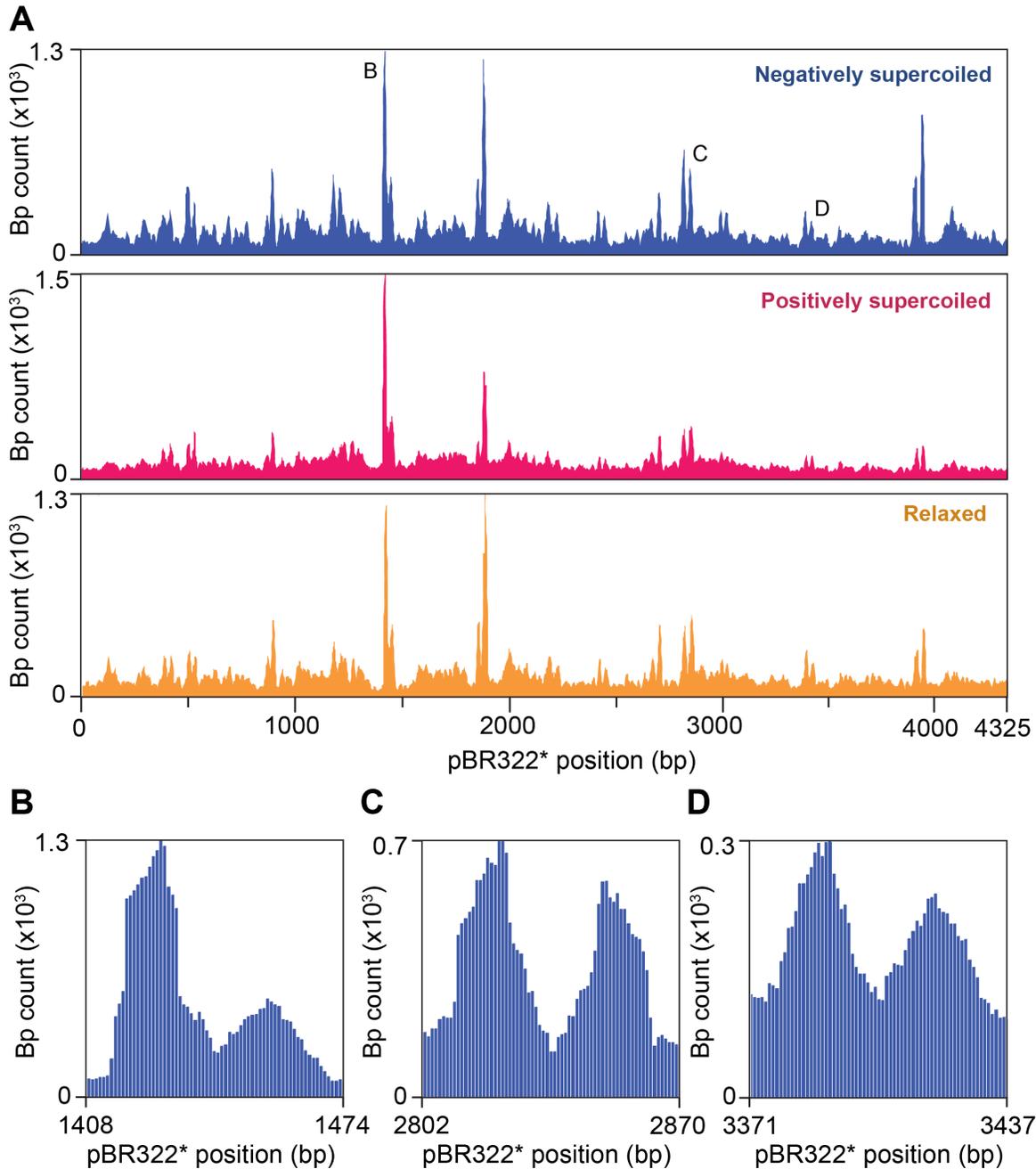
Figure 5.13: NGS data for *E. coli* topo IV (EcT4) cleavage activity. All data are plotted as nucleotide read count (read depth) for each position along pBR322*. A: NGS data for EcT4 cleavage, plot in regard to the topology of the DNA substrate used, negatively supercoiled (blue), positively supercoiled (red) and relaxed (yellow). On the negatively supercoiled (blue) plot, the peaks enlarged in parts B-D are labelled with their respective letter. B-D: Close up images of the EcT4 peaks, as labelled in A, showing the characteristic peak shape indicative of EcT4 cleavage activity, which appears insensitive to read depth.

E. coli Topo IV
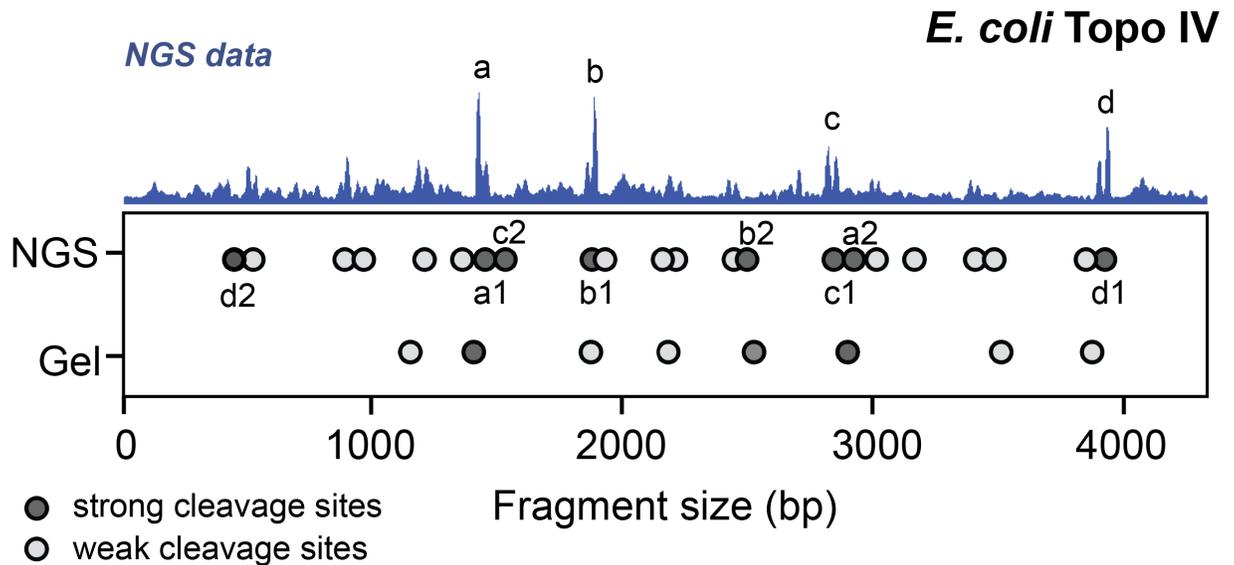
Figure 5.14: NGS data versus gel-based data for EcT4-dependent fragmentation of pBR322*. The NGS data is on top in blue. The strong cleavage sites are labelled as 'a', 'b', 'c' and 'd'. Below, the fragment sizes predicted by the sequencing were compared to those extracted from the gel-based assay. Strong cleavage sites (large peaks in NGS data/bright cleavage bands in gel-based data) are dark grey, while other weaker cleavage sites are light grey. For the NGS fragment sizes, the strong sites are labelled with 1 or 2 to demonstrate they are partners with both being formed should EcT4 cleave linearised pBR322* at the position indicated on the NGS data.

The type of sequence preference analysis performed on the MmT6 data was facilitated due to the resolution of the precise cleavage site. For EcT4, the cleavage site position can be roughly predicted (e.g. centrally located between the two peaks) but not precisely. This led to sequence composition analysis of the DNA within each peak, as compared to the whole pBR322* sequence. To do this, a 56 bp sequence, spanning the entire EcT4 cleavage site was extracted and analysed for the presence of certain dinucleotides (Figure 5.15). See Table III in Appendix III for information on each sequence extracted. A clear preference for C/G rich regions was found with the dinucleotides CG, GG, GC, CC and GT all shown to be enriched in the EcT4 cleavage sites relative to the background. This is in agreement with the EcT4 *in vivo* NGS study, which also found GC dinucleotides enriched near the cleavage sites [125, 341]. This may reflect T4 DNA binding preference and as these GC rich regions lie in the vicinity of the cleavage site, they may be involved in G-segment bending or interactions with the C-terminal domain (CTD) of the ParC subunit.

Figure 5.15: Dinucleotide identity preference for EcT4 cleavage sites. The enrichment (%) is calculated from the frequency that a particular dinucleotide appears in the EcT4 cleavage site sequences minus its frequency across pBR322* as a whole.

Using non-NGS techniques and promoting cleavage by EcT4 in the presence of quinolone-derived inhibitors found a preference for cleavage after a purine (G or A) with the following nucleotide most commonly being A or T, with a C in this position disfavoured. Whilst the data presented here cannot speak to this, they also found G-rich tracts to be common upstream of cleavage [275], a feature also observed here. Using *Staphylococcus pneumoniae* topo IV (SpT4), it was found that cleavage site-specificity in the presence of various quinolone drugs could be compressed into the consensus sequence, G(G/C)(A/T)A^GNNCT(T/A)N(C/A) [370]. Without cleavage site resolution, it is difficult to attain any agreement among this sequence and the EcT4 derived data presented here, however as seen for MmT6 and SsT6, there is often a fair amount of disparity between enzymes from different organisms. In addition, most, if not all, work published on the subject of topo-dependent cleavage has been achieved through the use of complex-stabilising drugs, which, as previously described, bias cleavage to certain sites which may not reflect strong cleavage sites in the absence of drugs.

5.3.5 DNA Gyrase NGS Results

The third topo explored using SHAN-seq was *E. coli* DNA gyrase (EcGyrase). It has been the subject of a number of studies looking at DNA binding and cleavage activity in the context of the DNA sequence [339, 360, 377, 378]. The studies often use quinolone-derived inhibitors to stimulate cleavage by EcGyrase. Here, only $CaCl_2$ was used to stimulate cleavage with the hope of attaining data on the more 'natural' cleavage site preference of EcGyrase.
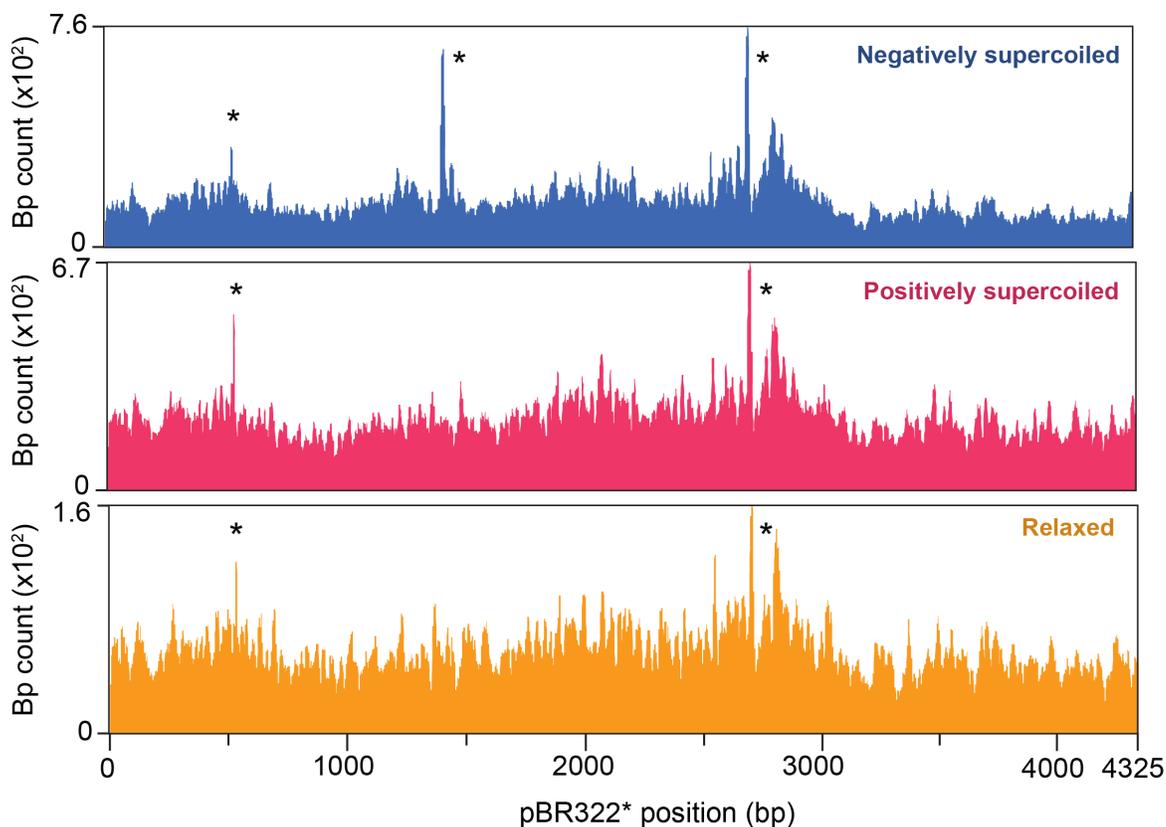
Figure 5.16: NGS data for *E. coli* DNA gyrase (EcGyrase) cleavage activity. All data are plotted as nucleotide read count (read depth) for each position along pBR322*. Data is displayed depending on the DNA topology of pBR322* used in the sample, negatively supercoiled (blue), positively supercoiled (red) and relaxed (yellow). EcGyrase dependent cleavage sites are labelled using asterisks.

Whilst the assay was limited by the availability of highly pure gyrase, a preliminary look into EcGyrase cleavage activity *in vitro* as a function of DNA topology was facilitated using SHAN-seq. Only three sites could be tentatively selected as EcGyrase-dependent, based on the read depth relative to the background (see Figure 5.16, peaks highlighted using asterisks). Interestingly, the second of the three sites was only enriched in the sample that contained

negatively-supercoiled pBR322*, suggesting this site was topology-sensitive. When this site, and the other two were compared to the agarose gel-based data from Section 5.3.1, the agreement was excellent, including the topology-sensitive site (Figure 5.17). In the gel (see Figure 5.5) there were two fragments (~1.5 and 3 kb) which were far more enriched in the sample using negatively-supercoiled DNA, which the NGS data suggested is due to topology-sensitive cleavage around base 1500 of pBR322*. This preliminary data suggests that DNA topology can affect DNA cleavage through a modulation of EcGyrase sequence preference.
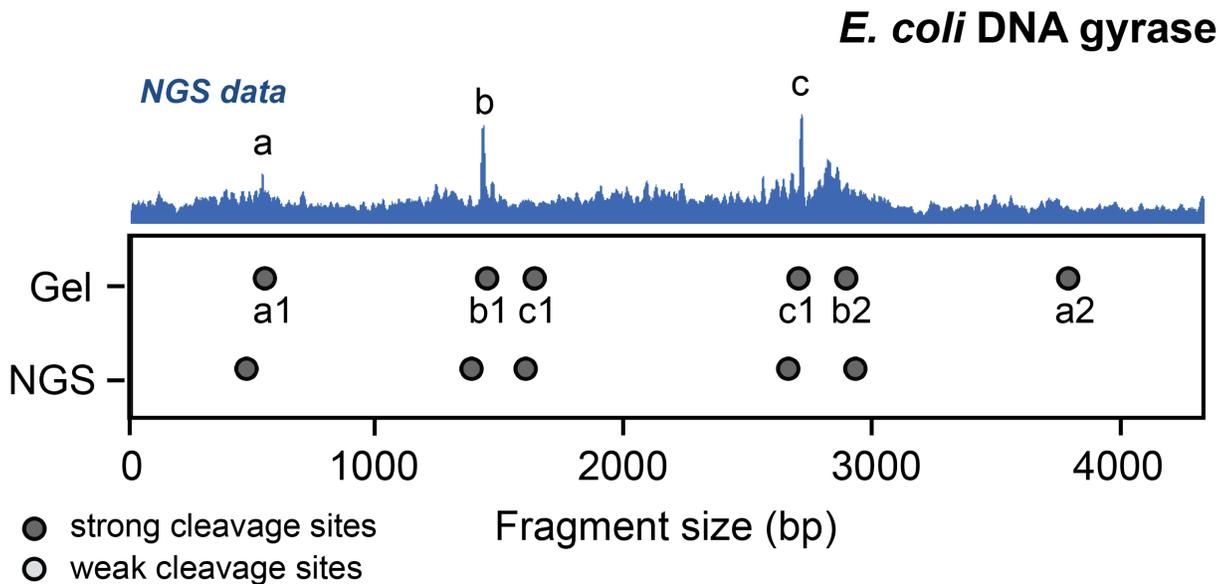
Figure 5.17: NGS data versus gel-based data for EcGyrase-dependent fragmentation of negatively-supercoiled pBR322*. The NGS data is on top in blue. The strong cleavage sites are labelled as 'a', 'b' and 'c'. Below, the fragment sizes predicted by the sequencing were compared to those extracted from the gel-based assay. Strong cleavage sites (large peaks in NGS data/bright cleavage bands in gel-based data) are dark grey, while other weaker cleavage sites are light grey. For the NGS fragment sizes, the strong sites are labelled with 1 or 2 to demonstrate they are partners with both being formed should EcGyrase cleave linearised pBR322* at the position indicated on the NGS data.

Unfortunately the data was unable to inform on the precise cleavage site at each of the three peaks, which prevents analysis of the cleavage site preference. However, sequences, 140 bp in length, were extracted with 70 bp either side of the peak centre and analysed for dinucleotide identity (see Table IV, Appendix III for sequence information). This revealed an

enrichment of dinucleotides made up of G and C, or C and T (Figure 5.18). In a previous study, also using an NGS technique called topo-seq, EcGyrase cleavage activity was analysed *in vivo* in the presence of oxolinic acid, ciprofloxacin and microcin B17. This work showed that there was a GC-content fluctuation that occurred either side of the precise cleavage site, which was hypothesized the be involved in binding to the GyrA CTDs and appeared independent of the drug used. The data collected using SHAN-seq only covers three sequences, and without nucleotide resolution of the cleavage site, the sequences cannot be analysed confidently for fluctuations in GC content. However, there are numerous tracts of three of more G/C nucleotides in the sequences with the GC content being 66, 64.5 and 59.6%, all of which are higher than the overall GC content of pBR322* (53%). This may indicate a preference for wrapping DNA with periodic repeats of G or C nucleotides. The preference for EcGyrase to wrap DNA with fluctuations in GC content was also found using radioactivity assays and it was determined that the G and C nucleotides were exposed as the DNA wrapped the GyrA CTDs [377], a concept elaborated upon in [339].

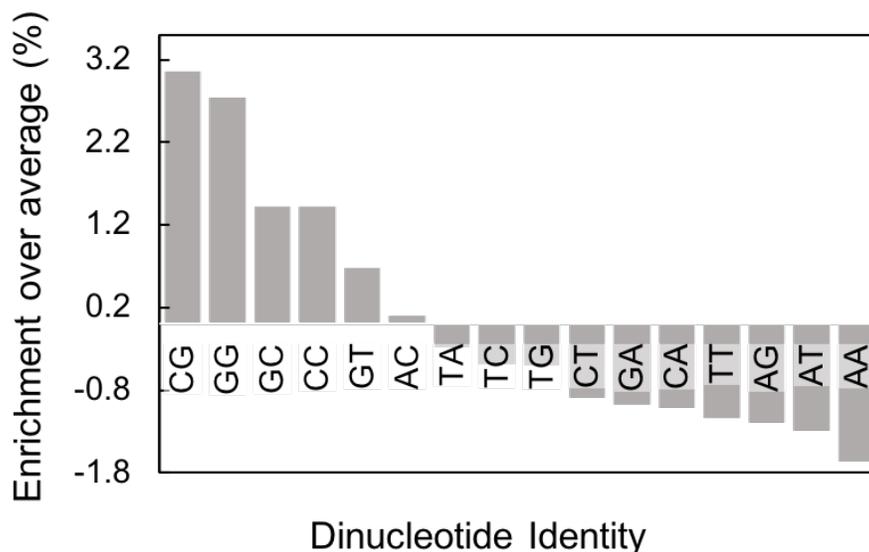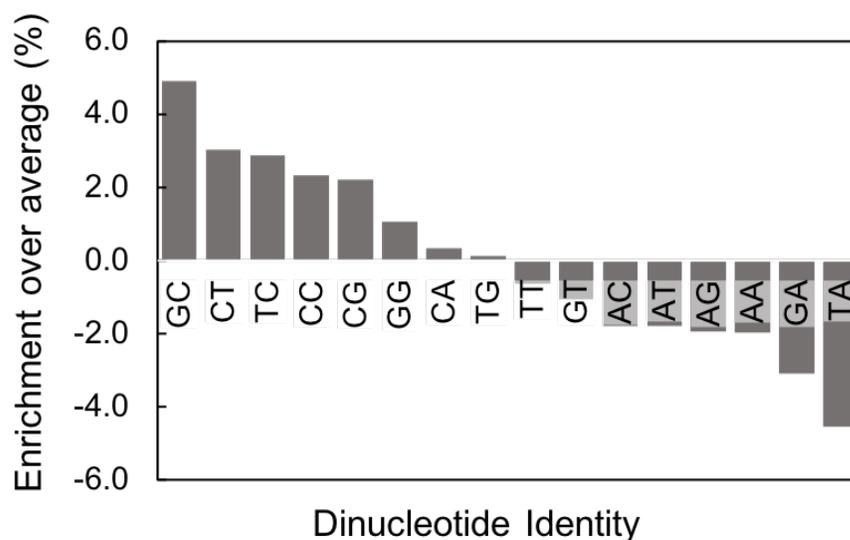**Figure 5.18: Dinucleotide identity preference for EcGyrase cleavage sites. The enrichment (%) is calculated from the frequency a particular dinucleotide appears in the EcGyrase cleavage site sequences minus its frequency across pBR322* as a whole.**

The pBR322 plasmid has been used in topology assays for some time and in the past, a strong gyrase cleavage site was determined at position 990, when EcGyrase cleaved

pBR322 in the presence of oxolinic acid [358-360]. This site is not represented by the work reported here, either in the NGS or gel-based assay. This strongly suggests that some cleavage sites are inhibitor-specific, and whilst oxolinic acid induced a high level of cleavage by EcGyrase at this site, in its absence it wasn't cleaved at all (see Figures 5.16 and 5.17). Therefore, any extrapolation, based on this observation, that the site at 990 may represent a cleavage site for gyrase under normal circumstances would be false. Indeed further work is required, particularly establishing gyrase cleavage site-specificity in the absence of cleavage-complex stabilising drugs, which is a topic far more unexplored. SHAN-seq, when optimised, offers an elegant solution to studying the drug-dependent and -independent EcGyrase cleavage activities.

5.4   Conclusions

The use of SHAN-seq has provided new data on the cleavage activity of MmT6, EcT4 and EcGyrase, all in the absence of inhibitors and providing insights previously unseen. This technique gives the experimenter complete control over the reaction conditions, such as DNA topology, DNA sequence, inhibitors and other proteins, all of which can be systematically assayed. The use of nitrocellulose filters in the enrichment of topo-bound DNA also allows proteins without specific antibodies to be explored in a highly accessible manner. Despite stages of SHAN-seq requiring optimisation (explored in section 5.5 below), its ability to elucidate precise cleavage sites, inform on the size of the overhang and whether it's on the 5$'$ or 3$'$ strand, has been confirmed through the use of restriction enzyme MluCI and MmT6.

MmT6 was shown to cleave pBR322* with the preference 5$'$-[G/C][G/C]^[A/C/T], generating 2-base 5$'$ overhangs. EcT4 demonstrated a unique cleavage site peak shape which allowed the identification of 15 potential cleavage sites. Even though nucleotide resolution of the cleavage site was absent, the sequences demonstrated an elevated GC content. It also showed that EcT4 activity is independent of DNA topology. EcGyrase was shown to generate the least amount of cleavage, with one particular site (~1.5 kb) clearly demonstrating topology-

sensitive activity, with its cleavage only occurring on negatively supercoiled pBR322* (not positively supercoiled or relaxed).

Using code written and executed by Rachel Kim (NHLBI, NIH, Bethesda, USA), the NGS data was circularised, creating an informative figure for use in comparison between different samples (Figure 5.19). As seen in Figure 5.19A, the cleavage profiles of MmT6, EcT4 and EcGyrase were compared. This revealed that the topology sensitive site for EcGyrase was also enriched in the EcT4 data, however for EcT4 this site was not sensitive to topology. The data can also be compared among different topologies (Figure 5.19B), highlighting topology-sensitive sites. This method of data presentation could be extended to comparing cleavage profiles using different inhibitors, or any variable the experimenter decides to explore. In conclusion, SHAN-seq is an accessible technique, which has proven successful in the exploration of topo-dependent cleavage activity in vitro using plasmid DNA.
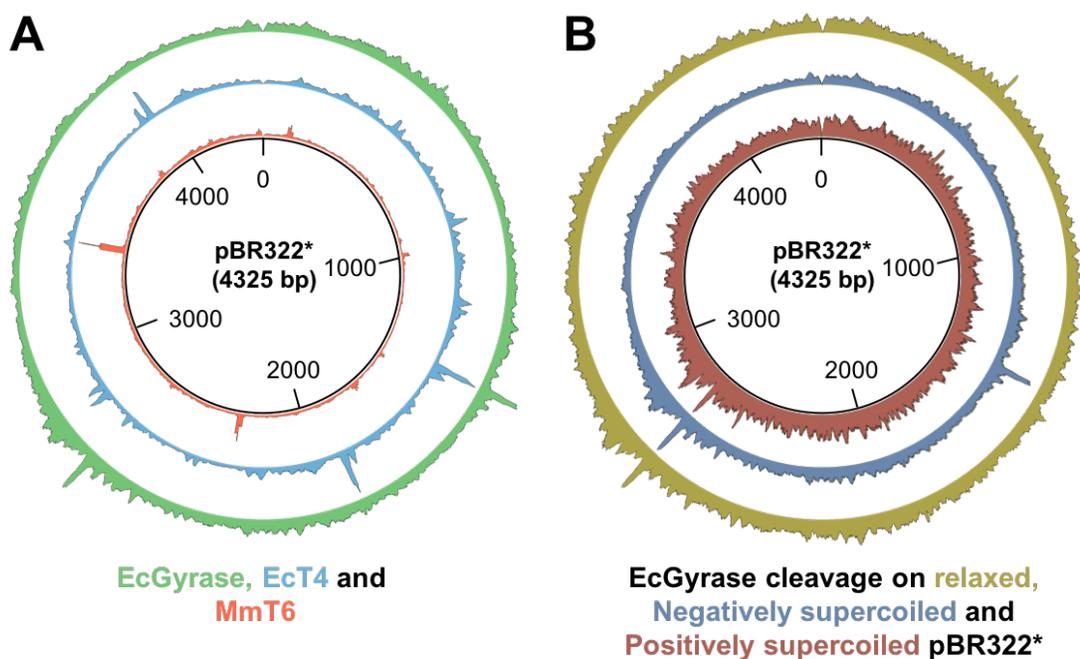
Figure 5.19: Circular plots of NGS data. A: Comparison of NGS data generated from the cleavage of negatively supercoiled pBR322* by EcGyrase (green), EcT4 (blue) and MmT6 (red). B: Comparison of NGS data generated from the cleavage of relaxed (yellow), negatively supercoiled (blue) and positively supercoiled (red) pBR322* by EcGyrase. Code written and executed by Rachel Kim.

5.5   Future Work

Chapter 5: NGS Results

Whilst SHAN-seq shows great promise in the characterisation of topo-dependent cleavage activity, during the course of this chapter, certain pitfalls have been highlighted. One major issue, which has plagued all topo NGS studies, is the removal of the phospho-tyrosyl adduct. Without its removal, adapter ligation of the 5' strand from topo cleavage sites will not occur and therefore won't be sequenced. There have been a number of solutions to this, namely enzymatic digestion. CIP was the choice here, being readily accessible through commercial means and evidence that its diphosphatase activity is sufficient to remove the adduct [155]. However, evidently CIP did not function efficiently, exemplified in the MmT6 NGS data from the MmT6-dependent 5' strand, which was barely sequenced whilst the 3' MmT6-dependent strand was sequenced to a significant depth. The adapter ligation stage in which the 3' and 5' ends are uniquely indexed offers an alternative approach, by allowing the 3' NGS data to be separated from the 5' data. In the case of MmT6, this worked surprisingly well and despite the failure of the 5' MmT6-clevaed end sequencing, the opposite strand revealed the information independently. This was similar to an NGS technique employed for gyrase cleavage site characterisation in which a single stranded DNA library generation kit was used, in which only the 3' strand was sequenced [339]. This work was also able to reveal the nucleotide overhangs associated with gyrase cleavage activity, demonstrating further the expendable nature of the 5' strand. In some cases, the enzyme TDP2 was employed, which is the protein responsible for rescuing stalled topo II *in vivo* by cleaving the phospho-tyrosyl [338]. However, caveats to the use of TDP2 include commercial inaccessibility and where it is available, the quantity and purity required arelacking. If large amounts of highly pure TDP2 were made, this could form a more robust approach to removing the adduct *in vitro*, and TDP2 itself has been used with success in in vivo topo NGS studies [337].

Another potential issue is the use of sonication. Based on the data presented here, particularly the EcT4 data, there is evidence that DNA is breaking in the vicinity of the cleavage site and this is being sequenced instead of the precise cleavage site. One potential explanation for this is that during sonication the DNA is sheering around the bound protein. However that would suggest the DNA becomes protein-free and therefore would be indexed in the first stage

of adapter ligation. This wasn't the case as the data is derived from the EcT4-cleavaed DNA end. That then leads to possibility that the DNA is shearing away from the cut site but remaining bound to protein. Topos tend to have extended DNA binding surfaces involved in docking/bending of the G-segment or sensing the T-segment, therefore it is possible that DNA will remain bound to the topo, despite sonication-dependent cleavage. Whilst sonication is the gold-standard in generating random DNA fragments, it may too harsh on the cleavage complexes. Enzymatic fragmentation options exist, which will be explored as SHAN-seq is further optimised.

An alternative use for this method, which has not yet been discussed, is in the exploration of topo DNA-binding activity. Rather than allowing the enzyme to cleave the DNA, binding can be explored by omitting nucleotide, in the case of MmT6. The DNA could then be sonicated and then enzyme removed prior to sequencing. This would allow the comparison of DNA binding to DNA cleavage activity, which, as discussed in section 5.1, is often not well correlated, and has raised some important questions about topoisomerase activity *in vivo*. The caveat with exploring binding is that it would be difficult to obtain data on the exact binding position of the topo, however it can still inform on whether certain regions of the DNA substrate are bound more frequently than others.

The continuation of work done on SHAN-seq is currently the role of Rachel Kim (NHLBI, NIH, Bethesda, US), with the goal to establish an optimised SHAN-seq protocol, before exploring the activity of different topos on various DNA topologies and with inhibitors. This is in the hope that SHAN-seq can be translated into a technique of use to the wider topo community.

Chapter 6

Discussion and Conclusions

6.1   Introduction

When this research project began, it was with the aim to further characterise *Methanosarcina mazei* DNA topoisomerase VI (MmT6) *in vitro*, to advance our understanding of type IIB topos and develop MmT6 as a tractable model system for expanding knowledge of other topo VI homologues. Both the presence of topo VI in eukaryotes [180] and the homology to Spo11 [171], makes research on this little-known archaeal enzyme important in a broader context, findings of which will hopefully be translated into furthering eukaryotic research in the field of DNA metabolism.

The research conducted here was fortunate enough to be a transatlantic collaboration between the UK-based lab of Prof. Anthony Maxwell (JIC, Norwich, UK), and the US-based lab of Dr. Keir Neuman (NIH, Bethesda, USA). The former lab specialises in the utilisation of ensemble biochemical techniques, and the latter in the use of the biophysical single-molecule technique, the magnetic tweezers. This allowed MmT6 characterisation to benefit from the application of multiple, complementary techniques, which facilitated acquisition of novel results and revealed previously unseen enzymatic behaviours that have led to generating a well-supported hypothesis of MmT6 activity. The remainder of this final chapter is concerned with bringing all these findings together and discussing implications of the data with an expanded perspective.

6.2   Characterisation of *M. mazei* topo VI using ensemble and single-molecule methods

The use of both ensemble and single molecule methods in the characterisation of MmT6 produced highly-complementary data, and revealed that MmT6 had preference for binding DNA writhe with a crossing geometry of 87.4-87.7°, thus predisposing MmT6 to decatenation and

braid unlinking, rather than supercoil or plectoneme relaxation. An overview of the results amassed during this project can be seen in Figure 6.1, highlighting findings which correspond to each step of the hypothesised MmT6 reaction cycle.
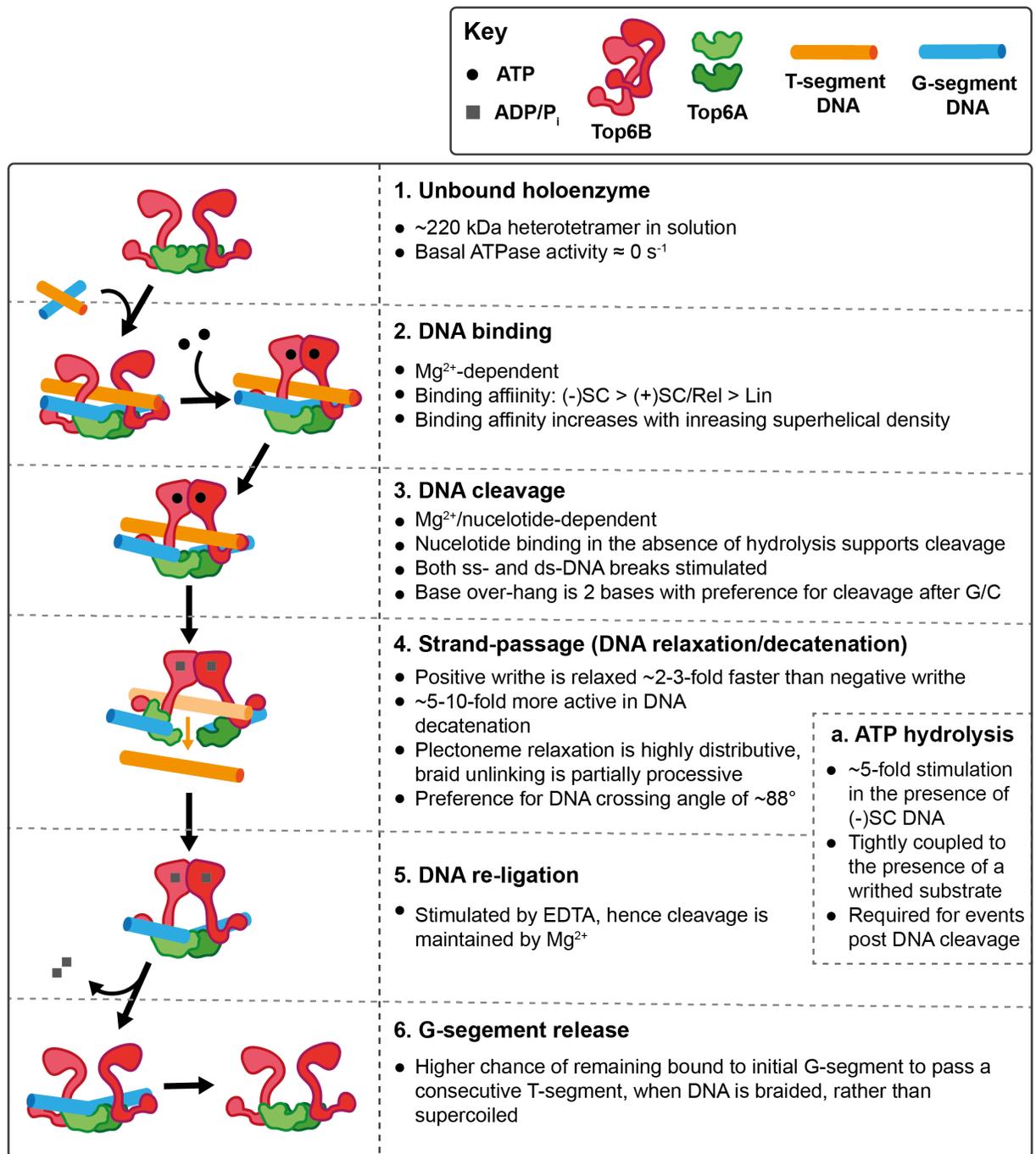
Figure 6.1: Results attained over the course of this research project and the MmT6 reaction cycle steps on which they inform. The MmT6 reaction cycles has been illustrated with reference to data attained in [62, 64, 67, 195]. An unbound MmT6 holoenzyme (1) interacts with a DNA crossing formed of the G- (blue) and T-segment (orange), with the Top6B subunits dimerising in response to ATP binding (2). The G-segment is cleaved (3), the T-segment is passed through the cleaved G-segment (4), before the G-segment is re-ligated (5). ATP hydrolysis occurs during strand

Chapter 6: Discussion and Conclusions

passage and re-ligation (a), with the release of ADP/P$_i$ causing the Top6B subunits to separate. The enzyme either remains bound to the G-segment, ready to capture a consecutive T-segment, or dissociates from the G-segment (6). The results discussed through this thesis are summarised beside each numbered stage of the MmT6 reaction cycle. Abbreviations are Ethylenediaminetetraacetic acid (EDTA), negatively supercoiled DNA ((-)SC), positively supercoiled DNA ((+)SC), relaxed (Rel) and linear (Lin), and magnesium ions (Mg^{2+}).

MmT6 was shown to be a heterotetramer in solution, meaning the holoenzyme does not assemble on the DNA from Top6A/Top6B dimers or Top6A and Top6B monomers. It also demonstrated that basal ATP hydrolysis activity was almost non-existent in the absence of a DNA substrate, indicating that DNA is a key activator of MmT6 ATPase activity. In the presence of DNA, MmT6 bound in a Mg^{2+}-dependent fashion, with increased affinity for writhed DNA substrates. It was demonstrated, using a membrane capture technique, that the relative MmT6 dissociation constant decreased as DNA superhelical density increased. This DNA writhe-dependent MmT6 stimulation was also seen when exploring ATP hydrolysis activity using a radioactive ATPase assay, and strand passage measured using the magnetic tweezers. The MmT6 ATP hydrolysis rate was ~5-fold higher in the presence of negatively supercoiled pBR322*, compared with either linear or relaxed pBR322*, and was tightly coupled to the presence of a writhed DNA substrate. Utilisation of the magnetic tweezers plectonemic DNA substrate allowed analysis of MmT6 strand-passage rate as a function of the number of supercoils. This demonstrated that MmT6 rate decreased with lowering supercoil density, with the data fitted to an inverse relationship in which strand-passage rate was inversely proportional to the number of supercoils. These data, along with research published during the course of this project [62], indicate that MmT6 directly senses and resolves DNA crossings, preferentially binding both the G- and T-segment simultaneously. This is unlike the hypothesised behaviour of type IIA topos, which are believed to bind the G-segment before capturing the T-segment [11, 51].

Using either the magnetic tweezers plectoneme or supercoiled pBR322*, in single-molecule and ensemble assays respectively, MmT6 was a largely-distributive and slow supercoil relaxase, demonstrating an ATP hydrolysis rate ~30-fold lower than type IIA topos.

236

As previously discussed [62], this would likely render MmT6 unable to keep up with the demands of DNA replication and transcription within *M. mazei* cells, and led to the reasoning that *in vitro* assays may be excluding a condition or molecule which enhances MmT6 rate *in vivo*. However, the use of the magnetic tweezers braided DNA substrate enhanced MmT6 average rate of reaction ~5-6-fold as compared to plectoneme relaxation at the same MmT6 concentration. In addition, MmT6 behaviour went from being highly distributive in the presence of supercoils to partially processive during braid unlinking. This suggested a factor more common in braids than plectonemes was responsible for enhancing MmT6 activity and allowing the enzyme to remain bound to the initial G-segment for long enough to capture and pass a consecutive T-segment. In addition, when the rate was extracted for the processive bursts only, the rate of MmT6 was no longer dramatically lower than that of other type IIA topos. Using a positive braid, the MmT6 average processive rate was measured to be 0.5-0.7 strand-passage events s^{-1}, which is far more comparable to that found for *E. coli* DNA gyrase at 1 strand passage event s^{-1} [319], and *E. coli* topo IV at 2.5 strand passage events s^{-1} [119]. In addition, when the processive rate for braid unlinking is compared to the average rate of plectoneme relaxation, MmT6 activity was enhanced 7-23-fold, depending on the MmT6 concentration used. DNA braids have been used as a proxy for studying decatenation activity, despite structural differences [257], therefore it was necessary to analyse MmT6 activity through ensemble techniques using catenated plasmids. This data supported the DNA braid unlinking results, demonstrating 5-10-fold less MmT6 was required for efficient decatenation activity, than for supercoil relaxation, under identical conditions.

Using both agarose gel-based ensemble techniques and the single-molecule magnetic tweezers, at 21°C, MmT6 displayed a weak yet significant 2-3-fold rate enhancement in the presence of positive writhe over negative. This was demonstrated not to be a function of enhanced binding or cleavage on positive writhe, with both these measurements demonstrated to be higher for negative writhe. Along with the data discussed in the previous paragraph, this suggested MmT6 had a DNA crossing geometry preference. This was explored in depth using the magnetic tweezers single crossing assay, first described in [120]. A preference for positive

Chapter 6: Discussion and Conclusions

writhe indicated that MmT6 had a preference for acting on juxtaposed G- and T-segments with angles lower than 90°, as these angles are more common in positive, with negative writhe angles more commonly obtuse. The preferred MmT6 DNA crossing angle was determined to be 87.4-87.7°, and even though this value is very close to 90°, it is enough to affect the chiral-dependent rate of relaxation. Monte Carlo simulations of plectonemes with a single supercoil at 0.4 pN showed that the preferred MmT6 angle appears ~2.8-fold more often in a single positive supercoil, compared to a negative supercoil (Figure 6.2A). In addition, when catenane angles were simulated in previously published work [119], the angle distributions were far broader and centred at 90°. Therefore, not only did the crossing angle preference explain the difference in rate on positive and negative writhe, but extended to providing a mechanism for the preferential decatenation activity of MmT6. Furthermore, this preferred angle is far less common in more supercoiled substrates, distributions of which peak around 60° [119], presenting an explanation for the highly inefficient MmT6 DNA relaxase activity.
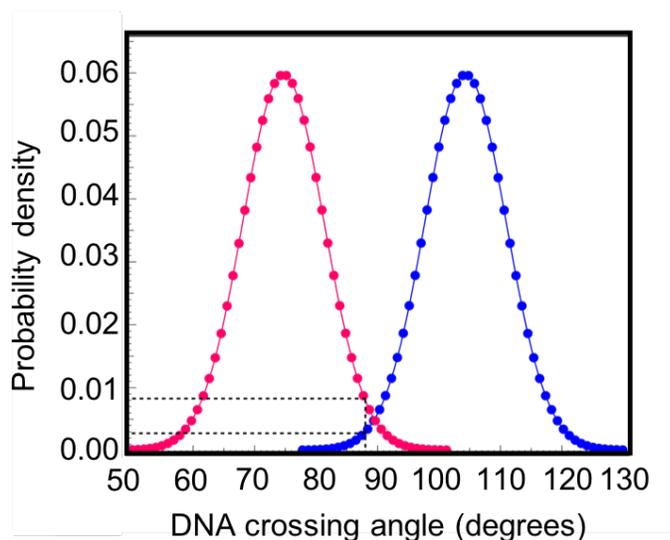
Figure 6.2: Supercoil DNA crossing angle distributions measured using Monte Carlo (MC-) simulations. DNA crossing angle distributions for a single positive (pink circles) and a single negative (blue circles) supercoil, measured using an MC-simulation under 0.4 pN, using a 5-kb DNA duplex. Black dotted lines indicate the probability of the MmT6 preferred angle (~87.55°) for a positive or negative crossing.

In addition to the use of the ensemble and single-molecule techniques, mentioned above, a Next Generation Sequencing (NGS) technique was developed to explore MmT6 DNA

238

cleavage sequence preference and measure the base-overhang generated during double-stranded cleavage *in vitro*. The technique, named SHAN-seq, was able to not only reveal both these pieces of information about MmT6, but also allowed a preliminary glance into the sequence preferences of other type II topos, *E. coli* DNA gyrase and topoisomerase IV. It is hoped, that once SHAN-seq is refined, that it will be a useful and accessible technique to the field of DNA topoisomerase research. For MmT6, it revealed that MmT6 cleaved the DNA generating 2-base overhangs with a preference for cleavage after G/C dinucleotides. Further work is necessary to understand this sequence preference, however, the work described in Chapter 5, demonstrated the validity of this technique as well as showcased its promise in the collection of *in vitro* topo-dependent cleavage data.

6.3   *M. mazei* topo VI findings in a broader context

6.3.1 Understanding the activity of topo VI from the archaeal perspective

Archaea constitute the third domain of the tree of life, alongside the bacteria and eukarya. Archaea earnt their distinction from bacteria, previously having been grouped together, based on their unique ribosomal proteins and RNA [379]. In terms of topoisomerase encoding, archaea are quite diverse. Almost all contain topo VI, aside from the order Thermoplasmatales, which instead encodes a DNA gyrase, whilst some, such as the Methanosarcina order encode both gyrase and topo VI [72]. In addition to type II topos, all archaea, aside from the Thuamarchaea, encode one or two copies of a type IA topo, which, in terms of sequence similarity, resembles bacterial and eukaryotic topo III, as opposed to bacterial topo IA [380]. This has led to the belief that topo VI must be integral to the removal of positive supercoils during transcription and replication, particularly in archaea lacking a gyrase, as topo III is known to have preferential decatenation activity [381] and topo IA is known to relax negative writhe, not positive (unless a ss-DNA region is present). However, the work reported here solidifies the observation found in [62], that *M. mazei* topo VI is an extremely inefficient DNA relaxase, and provides evidence of its preferential decatenation activity. If this behaviour is true of all archaeal topo VI enzymes, then those organisms which lack DNA gyrase, would struggle to

relieve torsional stress ahead of replication forks at a pace required by the cell. The current literature confounds the issue as the first archaeal type IA topo explored *in vitro*, from the hyperthermophile *Desulfurococcus amylolyticus*, also termed topo III, did exhibit robust relaxation activity of both positive and negative writhe at 95°C. However, it has also been established in *Sulfolobus solfataricus*, that of the three type IA topos it encodes, two are reverse gyrases [382] (another type IA topo important in positive supercoiling of hyperthermophillic genomes [383]), and the third is topo III, exhibiting preferential decatenation activity [384]. It may be that in archaea which lack DNA gyrase or a type I topo capable of positive supercoil relaxation, that topo VI has evolved more efficient DNA relaxation activity, in a similar sense to how *Mycobacterium tuberculosis* DNA gyrase has evolved to be efficient in both DNA supercoiling and decatenation in the absence of topo IV [347]. *M. mazei* encodes both topo VI and DNA gyrase, as well as two uncharacterised type IA topos, meaning that topo VI may not be required during positive supercoil relaxation and hence has evolved to preferentially decatenate. It is also important, as was done in Wendorff *et al.* [62], to consider the doubling time of the archaea in question, as the slow relaxation by topo VI may be adequate to support the growth of that particular organism. This did not seem to be the case for *M. mazei* and the involvement of unknown accessory factors that may enhance rate or processivity of topo VI has been postulated [62].

6.3.2 Implications for eukaryotic systems that express topo VI homologues

Whilst research on the eukaryotic topo VI and Spo11 proteins has the potential to create a large impact in eukaryotic cellular understanding, the proteins have been notoriously hard to express and purify for investigation *in vitro*. Therefore, utilising the far more accessible archaeal enzymes, has already shed light on the mechanism of the eukaryotic counterparts and will hopefully lead to advancements in their expression and *in vitro* characterisation, and, in addition, provide a means of screening compounds that can then be used against eukaryotic enzymes in herbicidal and anti-cancer treatments. For example, the Hsp90 inhibitor radicicol

was found to inhibit *S. shibatae* topo VI using *in vitro* techniques [215], which later led to the discovery that it also inhibited human type IIA topos and therefore had promise as an anti-cancer treatment [385]. Furthermore, using a high-throughput screen, the compound hexylresorcinol was found to inhibit *M. mazei* topo VI *in vitro*, which was subsequently confirmed to also effect the growth of *A. thaliana* seedlings, potentially through the inhibition of plant topo VI [386]. So whilst archaea are not always considered fundamental to furthering our knowledge on eukaryotic systems, archaeal topo VI has given important insight into the field of eukaryotic DNA metabolism, with its *in vitro* characterisation also advancing understanding on the mechanism of Spo11 [171, 172]. The work reported here hopes to maintain this trend, providing not only intriguing results into archaeal DNA metabolism, but leading to a deeper understanding of eukaryotic biology, in particular the role of topo VI in plants.

Genomic analysis of *Arabidopsis thaliana* revealed three distantly related homologs of the Topo6A/Spo11 subunit (AtSPO11-1, -2 and -3) and one of the Topo6B (AtTOP6B) subunit [175, 176, 178]. AtTOP6B and AtSPO11-3, the plant topo VI subunits, have high levels of expression and the proteins have also been demonstrated to interact [38]. *A. thaliana* mutants, homozygous for null mutations in either AtTOP6B or AtSPO11-3, display identical phenotypes, with severely growth-stunted plants failing to live much past 4-5 weeks [178]. Furthermore, double mutants displayed a phenotype identical to that of either single mutant, demonstrating the proteins likely function in the same process, or even the same protein complex. These studies suggested that the topo VI homologues play a key role in DNA replication and cell growth. A process called endoreduplication, in which the genome is replicated multiple times in the absence of cellular division, is critical for plant growth to enlarge cells of the hypocotyl and leaves [179]. Endoreduplication is deficient in the AtSPO11-3/AtTOP6B mutants and can explain the dwarf aspect of the phenotype [30, 31]. However, *A. thaliana* possesses four type I topoisomerases [180], along with both topoisomerase II and DNA gyrase, both of which are type IIA topos capable of the same reactions as topo VI, namely DNA decatenation, unknotting and relaxation. So why *A. thaliana* requires topo VI exclusively during endoreduplication is not clear. The research done here, whilst conducted on an archaeal topo VI, hopes to help provide

an explanation for this, as topo VI has an intrinsic preference for decatenation, arising as a function of a strict DNA crossing angle preference. This means topo VI not only has a significant rate increase in the presence of catenanes and braids, but strongly disfavours the relaxation of supercoils. During the endocycle, the genome is replicated to variable degrees, however in wild-type *A. thaliana* trichomes this can be as high as 32C [179], with mutants in topo VI unable to replicate the genome past 8C [169, 178]. With increasing DNA replication, comes elevated levels of transcription, and therefore extensive pressure on the protein machinery involved in these processes. With both replication and transcription involving strand separation, significant levels of DNA supercoiling will occur which is efficiently dealt with by the type I and type IIA topos. It is possible, therefore, as these proteins are extremely efficient at relaxation, that those topos which are known to be able to decatenate *in vitro*, become subsumed by the necessity to relieve torsional stress generated ahead of replication and transcription forks *in vivo*. With this reaction so heavily disfavoured by topo VI, it can decatenate and unknot the genome unhindered, allowing the cell to continue through the endocycle. The research described here clearly does not rule out the possibility of protein-protein recruitment, temporal regulation, or other activity-modulating factors which lead to the indispensability of topo VI in endoreduplication, but does provide a rationale that is independent of the requirement of these yet unknown components. In *A. thaliana*, topo VI accessory proteins have been identified, known as RHL1 and BIN4, which are believed to interact with topo VI directly [181, 182], however the mechanism through which they interact is still not well understood. They are worth mentioning as it is still possible that they enhance or direct topo VI activity in some way, however there have been no archaeal homologues found of either RHL1 or BIN4.

6.4   Conclusions

Throughout the course of this research project, DNA topoisomerase VI from *M. mazei*, has been further characterised using both ensemble and single-molecule methods. This has led to the generation of the hypothesis that MmT6 is DNA crossing sensor with an angle preference

of ~87.4-87.7°, which predisposes it to decatenation whilst disfavouring the removal of supercoils. In the presence of a more geometrically-preferred substrate (DNA braid or catenated plasmids), activity is enhanced 5-23-fold (dependent on which metrics you compare), with the data collected here forming robust supporting evidence for the preferential decatenation activity of MmT6. However, further work is still required. Nonetheless, MmT6 has proven to be a highly tractable model system for archaeal type IIB characterisation, amenable to the use of both ensemble and single-molecule techniques; these findings hopefully aiding both eukaryotic topo VI and Spo11 research.

List of abbreviations

A_{340}: absorbance at 340 nm

ADP: adenosine diphosphate

ADPNP: 5′-adenylyl β,γ-imidodiphosphate

AIM: autoinduction media

ATP: adenosine triphosphate

B: bound DNA

BB: MmT6 binding buffer

Bp: base pairs

BSA: bovine serum albumin

BWB: bead wash buffer

CaCl$_2$/Ca^{2+}: calcium chloride/calcium ion

CCD: charge coupled device

CIP: calf intestinal alkaline phosphatase

CTD: C-terminal domain

CoCl$_2$/Co^{2+}: cobalt chloride/cobalt ion

DB: MmT6 dilution buffer

DSB: double-stranded DNA break

DTT: dithiothreitol

EcGyrase: *E. coli* DNA gyrase

EcT4: *E. coli* DNA topoisomerase IV

EDTA: ethylenediaminetetraacetic acid

EtBr: ethidium bromide

G-segment: gate segment

Gyrase: DNA gyrase

h or hr: hour(s)

HCl: hydrochloric acid

244

iSCAT: inferometric scattering technology

kDNA: kinetoplast DNA

KOH: potassium hydroxide

LED: light emitting diode

Lk: linking number

MB: MmT6 minimal buffer

MgCl$_2$/Mg^{2+}: magnesium chloride/magnesium ion

min: minute(s)

MmT6: *Methanosarcina mazei* DNA topoisomerase VI

MnCl$_2$/Mn^{2+}: manganese chloride/manganese ion

MT: magnetic tweezers

NaCl: sodium chloride

NADH/NAD$^+$: nicotinamide adenine dinucleotide reduced/oxidised

Neg: negatively supercoiled

NGS: next-generation sequencing

NiCl$_2$: nickel chloride/NiCl$_2$

NTD: N-terminal domain

OC: open circular/nicked plasmid DNA

PAGE: polyacrylamide gel electrophoresis

PBS: phosphate buffer solution

PCR: polymerase chain reaction

PEP: phosphoenolpyruvic acid

PK: proteinase K

PK/LDH: pyruvate kinase/lactate dehydrogenase

Pos: positively supercoiled

RB: MmT6 relaxation buffer

Rel: relaxed plasmid DNA

s/sec: second(s)

SC: supercoiled plasmid DNA

SDS: sodium dodecylsulphate

SEC-MALS: size exclusion chromatography coupled to multi-angle light scattering

SHAN-seq: simplified high accuracy end sequencing

SSB: single-stranded DNA break

SsT6: *Sulfolobus shibatae* DNA topisomerase VI

TDP2: tyrosyl-DNA phosphodiesterase 2

TLC: thin layer chromatography

Topos: DNA topoisomerases

Topo IA: DNA topoisomerase IA

Topo IB: DNA topoisomerase IB

Topo II: DNA topoisomerase II

Topo III: DNA topoisomerase III

Topo IV: DNA topoisomerase IV

Topo V: DNA topoisomerase V

Topo VI: DNA topoisomerase VI

Topo VIII: DNA topoisomerase VIII

T-segment: transport segment

UB: unbound

WB: wash buffer

w/v: weight/volume

ZnCl$_2$/Zn^{2+}: zinc chloride/zinc ion

References

1.   Watson, J.D. and F.H.C. Crick, *A structure for deoxyribose nucleic acid.* Nature, 1953. **171**: p. 737-738.

2.   Watson, J.D. and F.H. Crick, *Genetical implications of the structure of deoxyribonucleic acid.* Nature, 1953. **171**(4361): p. 964-7.

3.   Postow, L., N.J. Crisona, B.J. Peter, C.D. Hardy, and N.R. Cozzarelli, *Topological challenges to DNA replication: conformations at the fork.* Proc Natl Acad Sci U S A, 2001. **98**(15): p. 8219-26.

4.   Schvartzman, J.B. and A. Stasiak, *A topological view of the replicon.* EMBO Rep, 2004. **5**(3): p. 256-61.

5.   Woldringh, C.L., P.R. Jensen, and H.V. Westerhoff, *Structure and partitioning of bacterial DNA: determined by a balance of compaction and expansion forces?* FEMS Microbiol Lett, 1995. **131**(3): p. 235-42.

6.   Bates, A.D. and A. Maxwell, *DNA Topology.* 2005, Oxford: Oxford University Press.

7.   Dorman, C.J. and M.J. Dorman, *DNA supercoiling is a fundamental regulatory principle in the control of bacterial gene expression.* Biophys Rev, 2016. **8**(3): p. 209-220.

8.   White, J.H. and W.R. Bauer, *Calculation of the twist and the writhe for representative models of DNA.* J Mol Biol, 1986. **189**(2): p. 329-41.

9.   Călugăreanu, G., *Sur les classes d'isotopie des noeuds tridimensionnels et leurs invariants.* Czechoslovak Mathematical Journal, 1961. **11**(4): p. 588-625.

10.   Champoux, J.J., *DNA topoisomerases: structure, function, and mechanism.* Annu. Rev. Biochem., 2001. **70**: p. 369-413.

11.   Bush, N.G., K. Evans-Roberts, and A. Maxwell, *DNA Topoisomerases.* EcoSal Plus, 2015. **6**(2).

12.   Wang, J.C., *Interaction between DNA and an Escherichia coli protein w.* J. Mol. Biol, 1971. **55**: p. 523-533.

13. Corbett, K.D. and J.M. Berger, *Structure, molecular mechanisms, and evolutionary relationships in DNA topoisomerases.* Annu Rev Biophys Biomol Struct, 2004. **33**: p. 95-118.

14. Capranico, G., J. Marinello, and G. Chillemi, *Type I DNA Topoisomerases.* J Med Chem, 2017. **60**(6): p. 2169-2192.

15. Masse, E. and M. Drolet, *Relaxation of transcription-induced negative supercoiling is an essential function of Escherichia coli DNA topoisomerase I.* J Biol Chem, 1999. **274**(23): p. 16654-8.

16. Drolet, M., S. Broccoli, F. Rallu, C. Hraiky, C. Fortin, E. Masse, and I. Baaklini, *The problem of hypernegative supercoiling and R-loop formation in transcription.* Front Biosci, 2003. **8**: p. d210-21.

17. Phoenix, P., M.A. Raymond, E. Masse, and M. Drolet, *Roles of DNA topoisomerases in the regulation of R-loop formation in vitro.* J Biol Chem, 1997. **272**(3): p. 1473-9.

18. Tse-Dinh, Y.C., *Bacterial and archeal type I topoisomerases.* Biochim Biophys Acta, 1998. **1400**(1-3): p. 19-27.

19. Cheng, B., C.X. Zhu, C. Ji, A. Ahumada, and Y.C. Tse-Dinh, *Direct interaction between Escherichia coli RNA polymerase and the zinc ribbon domains of DNA topoisomerase I.* J Biol Chem, 2003. **278**(33): p. 30705-10.

20. Kirkegaard, K. and J.C. Wang, *Bacterial DNA topoisomerase I can relax positively supercoiled DNA containing a single-stranded loop.* J. Mol. Biol., 1985. **185**: p. 625-637.

21. Lima, C.D., J.C. Wang, and A. Mondragón, *Crystallization of a 67 kDa fragment of Escherichia coli DNA topoisomerase I.* J. Mol. Biol., 1993. **232**: p. 1213-1216.

22. Lima, C.D., J.C. Wang, and A. Mondragón, *Three-dimensional structure of the 67K N-terminal fragment of E. coli DNA topoisomerase I.* Nature, 1994. **367**: p. 138-146.

23. Tan, K., Q. Zhou, B. Cheng, Z. Zhang, A. Joachimiak, and Y.C. Tse-Dinh, *Structural basis for suppression of hypernegative DNA supercoiling by E. coli topoisomerase I.* Nucleic Acids Res, 2015. **43**(22): p. 11031-46.

24. Changela, A., R.J. DiGate, and A. Mondragon, *Structural studies of E. coli topoisomerase III-DNA complexes reveal a novel type IA topoisomerase-DNA conformational intermediate.* J Mol Biol, 2007. **368**(1): p. 105-18.

25. Beran-Steed, R.K. and Y.-C. Tse-Dinh, *The carboxyl terminal domain of Escherichia coli DNA topoisomerase I confers higher affinity to DNA.* Proteins: Struct. Funct. Genet., 1989. **6**: p. 249-258.

26. Yang, X., W. Li, E.D. Prescott, S.J. Burden, and J.C. Wang, *DNA topoisomerase IIbeta and neural development.* Science, 2000. **287**(5450): p. 131-4.

27. Li, Z., A. Mondragon, H. Hiasa, K.J. Marians, and R.J. DiGate, *Identification of a unique domain essential for Escherichia coli DNA topoisomerase III-catalysed decatenation of replication intermediates.* Mol Microbiol, 2000. **35**(4): p. 888-95.

28. Kikuchi, A. and K. Asai, *Reverse gyrase - a topoisomerase which introduces positive superhelical turns into DNA.* Nature, 1984. **309**: p. 677-681.

29. Forterre, P., *A hot story from comparative genomics: reverse gyrase is the only hyperthermophile-specific protein.* Trends Genet, 2002. **18**(5): p. 236-7.

30. Confalonieri, F., C. Elie, M. Nadal, C. de La Tour, P. Forterre, and M. Duguet, *Reverse gyrase: a helicase-like domain and a type I topoisomerase in the same polypeptide.* Proc Natl Acad Sci U S A, 1993. **90**(10): p. 4753-7.

31. Rodriguez, A.C. and D. Stock, *Crystal structure of reverse gyrase: insights into the positive supercoiling of DNA.* EMBO J., 2002. **21**(3): p. 418-26.

32. del Toro Duany, Y., S.P. Jungblut, A.S. Schmidt, and D. Klostermeier, *The reverse gyrase helicase-like domain is a nucleotide-dependent switch that is attenuated by the topoisomerase domain.* Nucleic Acids Res, 2008. **36**(18): p. 5882-95.

33. Ganguly, A., Y. del Toro Duany, and D. Klostermeier, *Reverse gyrase transiently unwinds double-stranded DNA in an ATP-dependent reaction.* J Mol Biol, 2013. **425**(1): p. 32-40.

34.   Declais, A.C., J. Marsault, F. Confalonieri, C.B. de La Tour, and M. Duguet, *Reverse gyrase, the two domains intimately cooperate to promote positive supercoiling.* J Biol Chem, 2000. **275**(26): p. 19498-504.

35.   Shibata, T., S. Nakasu, K. Yasui, and A. Kikuchi, *Intrinsic DNA-dependent ATPase Activity of Reverse Gyrase.* J. Biol. Chem., 1987. **262**: p. 10419-10421.

36.   Champoux, J.J. and R. Dulbecco, *An activity from mammalian cells that untwists superhelical DNA--a possible swivel for DNA replication (polyoma-ethidium bromide-mouse-embryo cells-dye binding assay).* Proc Natl Acad Sci U S A, 1972. **69**(1): p. 143-6.

37.   Stewart, L., M.R. Redinbo, X. Qiu, W.G.J. Hol, and J.J. Champoux, *A model for the mechanism of human topoisomerase I.* Science, 1998. **279**: p. 1534-1540.

38.   Stivers, J.T., T.K. Harris, and A.S. Mildvan, *Vaccinia DNA topoisomerase I: evidence supporting a free rotation mechanism for DNA supercoil relaxation.* Biochemistry, 1997. **36**: p. 5212-5222.

39.   Stewart, L., G.C. Ireton, and J.J. Champoux, *Reconstitution of human topoisomerase I by fragment complementation.* J. Mol. Biol., 1997. **269**: p. 355-372.

40.   Leppard, J.B. and J.J. Champoux, *Human DNA topoisomerase I: relaxation, roles, and damage control.* Chromosoma, 2005. **114**(2): p. 75-85.

41.   Alsner, J., J.Q. Svejstrup, E. Kjeldsen, B.S. Sørensen, and O. Westergaard, *Identification of an N-terminal domain of eukaryotic DNA topoisomerase I dispensable for catalytic activity but essential for in vivo function.* J. Biol. Chem., 1992. **267**: p. 12408-12411.

42.   Koster, D.A., V. Croquette, C. Dekker, S. Shuman, and N.H. Dekker, *Friction and torque govern the relaxation of DNA supercoils by eukaryotic topoisomerase IB.* Nature, 2005. **434**(7033): p. 671-4.

43.	Slesarev, A.I., K.O. Stetter, J.A. Lake, M. Gellert, R. Krah, and S.A. Kozyavkin, *DNA topoisomerase V is a relative of eukaryotic topoisomerase I from a hyperthermophilic prokaryote.* Nature, 1993. **364**: p. 735-737.

44.	Rajan, R., A. Osterman, and A. Mondragón, *Methanopyrus kandleri topoisomerase V contains three distinct AP lyase active sites in addition to the topoisomerase active site.* Nucleic acids research, 2016. **44**(7): p. 3464-3474.

45.	Taneja, B., B. Schnurr, A. Slesarev, J.F. Marko, and A. Mondragon, *Topoisomerase V relaxes supercoiled DNA by a constrained swiveling mechanism.* Proc Natl Acad Sci U S A, 2007. **104**(37): p. 14670-5.

46.	Taneja, B., A. Patel, A. Slesarev, and A. Mondragon, *Structure of the N-terminal fragment of topoisomerase V reveals a new family of topoisomerases.* EMBO J, 2006. **25**(2): p. 398-408.

47.	Belova, G.I., R. Prasad, S.A. Kozyavkin, J.A. Lake, S.H. Wilson, and A.I. Slesarev, *A type IB topoisomerase with DNA repair activities.* Proceedings of the National Academy of Sciences, 2001. **98**(11): p. 6015-6020.

48.	Rajan, R., R. Prasad, B. Taneja, S.H. Wilson, and A. Mondragón, *Identification of one of the apurinic/apyrimidinic lyase active sites of topoisomerase V by structural and functional studies.* Nucleic acids research, 2012. **41**(1): p. 657-666.

49.	Forterre, P., *DNA topoisomerase V: a new fold of mysterious origin.* Trends Biotechnol, 2006. **24**(6): p. 245-7.

50.	Lynn, R., G. Giaever, S. Swanberg, and J.C. Wang, *Tandem regions of yeast DNA topoisomerase II share homology with different subunits of bacterial gyrase.* Science, 1986. **233**: p. 647-648.

51.	Mizuuchi, K., L.M. Fisher, M.H. O'Dea, and M. Gellert, *DNA gyrase action involves the introduction of transient double-strand breaks into DNA.* Proc Natl Acad Sci U S A, 1980. **77**(4): p. 1847-51.

52.	Harrison, S.C. and A.K. Aggarwal, *DNA recognition by proteins with the helix-turn-helix motif.* Annu Rev Biochem, 1990. **59**: p. 933-69.

53.   Chang, C.C., Y.R. Wang, S.F. Chen, C.C. Wu, and N.L. Chan, *New insights into DNA-binding by type IIA topoisomerases.* Curr Opin Struct Biol, 2013. **23**(1): p. 125-33.

54.   McKie, S.J., A. Maxwell, and K.C. Neuman, *Mapping DNA Topoisomerase Binding and Cleavage Genome Wide Using Next-Generation Sequencing Techniques.* Genes (Basel), 2020. **11**(1).

55.   Wang, J.C., *DNA topoisomerases.* Annu Rev Biochem, 1996. **65**: p. 635-92.

56.   Dong, K.C. and J.M. Berger, *Structural basis for gate-DNA recognition and bending by type IIA topoisomerases.* Nature, 2007. **450**(7173): p. 1201-5.

57.   Lee, S., S.R. Jung, K. Heo, J.A. Byl, J.E. Deweese, N. Osheroff, and S. Hohng, *DNA cleavage and opening reactions of human topoisomerase IIalpha are regulated via Mg2+-mediated dynamic bending of gate-DNA.* Proc Natl Acad Sci U S A, 2012. **109**(8): p. 2925-30.

58.   Aravind, L., D.D. Leipe, and E.V. Koonin, *Toprim - a conserved catalytic domain in type IA and II topoisomerases, DnaG-type primases, OLD family nucleases and RecR proteins.* Nucleic Acids Res., 1998. **26**: p. 4205-4213.

59.   Sissi, C. and M. Palumbo, *Effects of magnesium and related divalent metal ions in topoisomerase structure and function.* Nucleic Acids Res, 2009. **37**(3): p. 702-11.

60.   Laponogov, I., M.K. Sohi, D.A. Veselkov, X.-S. Pan, R. Sawhney, A.W. Thompson, K.E. McAuley, L.M. Fisher, and M.R. Sanderson, *Structural insight into the quinolone-DNA cleavage complex of type IIA topoisomerases.* Nat. Struct. Mol. Biol., 2009. **16**(6): p. 667-669.

61.   Thomson, N.H., S. Santos, L.A. Mitchenall, T. Stuchinskaya, J.A. Taylor, and A. Maxwell, *DNA G-segment bending is not the sole determinant of topology simplification by type II DNA topoisomerases.* Sci Rep, 2014. **4**: p. 6158.

62.   Wendorff, T.J. and J.M. Berger, *Topoisomerase VI senses and exploits both DNA crossings and bends to facilitate strand passage.* Elife, 2018. **7**: p. e31724.

63. Smith, C.V. and A. Maxwell, *Identification of a residue involved in transition-state stabilization in the ATPase reaction of DNA gyrase.* Biochemistry, 1998. **37**(27): p. 9658-67.

64. Corbett, K.D. and J.M. Berger, *Structure of the topoisomerase VI-B subunit: implications for type II topoisomerase mechanism and evolution.* EMBO J, 2003. **22**(1): p. 151-63.

65. Hu, T., S. Chang, and T. Hsieh, *Identifying Lys[359] as a critical residue for the ATP-dependent reactions of Drosophila DNA topoisomerase II.* J. Biol. Chem., 1998. **273**: p. 9583-9592.

66. Lamour, V., L. Hoermann, J.M. Jeltsch, P. Oudet, and D. Moras, *An open conformation of the Thermus thermophilus gyrase B ATP-binding domain.* J. Biol. Chem., 2002. **277**(21): p. 18947-53.

67. Corbett, K.D. and J.M. Berger, *Structural dissection of ATP turnover in the prototypical GHL ATPase TopoVI.* Structure, 2005. **13**(6): p. 873-82.

68. Bates, A.D., J.M. Berger, and A. Maxwell, *The ancestral role of ATP hydrolysis in type II topoisomerases: prevention of DNA double-strand breaks.* Nucleic Acids Res, 2011. **39**(15): p. 6327-39.

69. Gellert, M., K. Mizuuchi, M.H. O'Dea, and H.A. Nash, *DNA gyrase: an enzyme that introduces superhelical turns into DNA.* Proc Natl Acad Sci U S A, 1976. **73**(11): p. 3872-6.

70. Wall, M.K., L.A. Mitchenall, and A. Maxwell, *Arabidopsis thaliana DNA gyrase is targeted to chloroplasts and mitochondria.* Proc Natl Acad Sci U S A, 2004. **101**(20): p. 7821-6.

71. Dar, M.A., A. Sharma, N. Mondal, and S.K. Dhar, *Molecular cloning of apicoplast-targeted Plasmodium falciparum DNA gyrase genes: unique intrinsic ATPase activity and ATP-independent dimerization of PfGyrB subunit.* Eukaryot Cell, 2007. **6**(3): p. 398-412.

72.   Forterre, P. and D. Gadelle, *Phylogenomics of DNA topoisomerases: their origin and putative roles in the emergence of modern organisms.* Nucleic Acids Res, 2009. **37**(3): p. 679-92.

73.   Yamashiro, K. and A. Yamagishi, *Characterization of the DNA gyrase from the thermoacidophilic archaeon Thermoplasma acidophilum.* J Bacteriol, 2005. **187**(24): p. 8531-6.

74.   Worcel, A. and E. Burgi, *On the structure of the folded chromosome of Escherichia coli.* J Mol Biol, 1972. **71**(2): p. 127-47.

75.   Higgins, N.P., C.L. Peebles, A. Sugino, and N.R. Cozzarelli, *Purification of subunits of Escherichia coli DNA gyrase and reconstitution of enzymatic activity.* Proc Natl Acad Sci U S A, 1978. **75**(4): p. 1773-7.

76.   Kreuzer, K.N. and N.R. Cozzarelli, *Formation and resolution of DNA catenanes by DNA gyrase.* Cell, 1980. **20**: p. 245-254.

77.   Marians, K.J., *DNA gyrase-catalyzed decatenation of multiply linked DNA dimers.* J. Biol. Chem., 1987. **21**: p. 10362-10368.

78.   Ishii, S., T. Murakami, and K. Shishido, *Gyrase inhibitors increase the content of knotted DNA species of plasmid pBR322 in Escherichia coli.* J. Bacteriol., 1991. **173**: p. 5551-5553.

79.   Maxwell, A., *DNA gyrase as a drug target.* Trends Microbiol., 1997. **5**(3): p. 102-9.

80.   Collin, F., S. Karkare, and A. Maxwell, *Exploiting bacterial DNA gyrase as a drug target: current state and perspectives.* Appl Microbiol Biotechnol, 2011. **92**(3): p. 479-97.

81.   Corbett, K.D., R.K. Shultzaberger, and J.M. Berger, *The C-terminal domain of DNA gyrase A adopts a DNA-bending {beta}-pinwheel fold.* Proc Natl Acad Sci U S A, 2004. **101**(19): p. 7293-7298.

82.   Hsieh, T.J., T.J. Yen, T.S. Lin, H.T. Chang, S.Y. Huang, C.H. Hsu, L. Farh, and N.L. Chan, *Twisting of the DNA-binding surface by a beta-strand-bearing proline modulates DNA gyrase activity.* Nucleic Acids Res, 2010. **38**(12): p. 4173-81.

254

83. Ruthenburg, A.J., D.M. Graybosch, J.C. Huetsch, and G.L. Verdine, *A superhelical spiral in the Escherichia coli DNA gyrase A C-terminal domain imparts unidirectional supercoiling bias.* J Biol Chem, 2005. **280**(28): p. 26177-84.

84. Tretter, E.M. and J.M. Berger, *Mechanisms for Defining Supercoiling Set Point of DNA Gyrase Orthologs: II. THE SHAPE OF THE GyrA SUBUNIT C-TERMINAL DOMAIN (CTD) IS NOT A SOLE DETERMINANT FOR CONTROLLING SUPERCOILING EFFICIENCY.* J Biol Chem, 2012. **287**(22): p. 18645-54.

85. Kramlinger, V.M. and H. Hiasa, *The "GyrA-box" Is Required for the Ability of DNA Gyrase to Wrap DNA and Catalyze the Supercoiling Reaction.* J Biol Chem, 2006. **281**(6): p. 3738-42.

86. Lanz, M.A. and D. Klostermeier, *The GyrA-box determines the geometry of DNA bound to gyrase and couples DNA binding to the nucleotide cycle.* Nucleic Acids Res, 2012. **40**(21): p. 10893-903.

87. Nollmann, M., N.J. Crisona, and P.B. Arimondo, *Thirty years of Escherichia coli DNA gyrase: from in vivo function to single-molecule mechanism.* Biochimie, 2007. **89**(4): p. 490-9.

88. Kirkegaard, K. and J.C. Wang, *Mapping the topography of DNA wrapped around gyrase by nucleolytic and chemical probing of complexes of unique DNA sequences.* Cell, 1981. **23**: p. 721-729.

89. Fisher, L.M., K. Mizuuchi, M.H. O'Dea, H. Ohmori, and M. Gellert, *Site-specific interaction of DNA gyrase with DNA.* Proc. Natl. Acad. Sci. USA, 1981. **78**: p. 4165-4169.

90. Morrison, A. and N.R. Cozzarelli, *Contacts between DNA gyrase and its binding site on DNA: features of symmetry and asymmetry revealed by protection from nucleases.* Proc. Natl. Acad. Sci. USA, 1981. **78**: p. 1416-1420.

91. Wigley, D.B., G.J. Davies, E.J. Dodson, A. Maxwell, and G. Dodson, *Crystal structure of an N-terminal fragment of the DNA gyrase B protein.* Nature, 1991. **351**(6328): p. 624-9.

92.   Tingey, A.P. and A. Maxwell, *Probing the role of the ATP-operated clamp in the strand-passage reaction of DNA gyrase.* Nucleic Acids Res., 1996. **24**: p. 4868-4873.

93.   Fu, G., J. Wu, W. Liu, D. Zhu, Y. Hu, J. Deng, X.E. Zhang, L. Bi, and D.C. Wang, *Crystal structure of DNA gyrase B' domain sheds lights on the mechanism for T-segment navigation.* Nucleic Acids Res, 2009. **37**(17): p. 5908-16.

94.   Piton, J., S. Petrella, M. Delarue, G. Andre-Leroux, V. Jarlier, A. Aubry, and C. Mayer, *Structural insights into the quinolone resistance mechanism of Mycobacterium tuberculosis DNA gyrase.* PLoS ONE, 2010. **5**(8): p. e12245.

95.   Schoeffler, A.J., A.P. May, and J.M. Berger, *A domain insertion in Escherichia coli GyrB adopts a novel fold that plays a critical role in gyrase function.* Nucleic Acids Res, 2010. **38**(21): p. 7830-44.

96.   Chatterji, M., S. Unniraman, A. Maxwell, and V. Nagaraja, *The additional 165 amino acids in the B protein of Escherichia coli DNA gyrase have an important role in DNA binding.* J Biol Chem, 2000. **275**(30): p. 22888-94.

97.   Brino, L., A. Urzhumtsev, M. Mousli, C. Bronner, A. Mitschler, P. Oudet, and D. Moras, *Dimerization of Escherichia coli DNA-gyrase B provides a structural mechanism for activating the ATPase catalytic center.* J Biol Chem, 2000. **275**(13): p. 9468-75.

98.   Vanden Broeck, A., C. Lotz, J. Ortiz, and V. Lamour, *Cryo-EM structure of the complete E. coli DNA gyrase nucleoprotein complex.* Nature Communications, 2019. **10**(1): p. 4935.

99.   Gellert, M., K. Mizuuchi, M.H. O'Dea, T. Itoh, and J.I. Tomizawa, *Nalidixic acid resistance: a second genetic character involved in DNA gyrase activity.* Proc Natl Acad Sci U S A, 1977. **74**(11): p. 4772-6.

100.   Gellert, M., M.H. O'Dea, T. Itoh, and J. Tomizawa, *Novobiocin and coumermycin inhibit DNA supercoiling catalyzed by DNA gyrase.* Proc Natl Acad Sci U S A, 1976. **73**(12): p. 4474-8.

101. Goss, W.A., W.H. Deitz, and T.H. Cook, *Mechanism of action of nalidixic acid on Escherichia coli. II. Inhibition of deoxyribonucleic acid synthesis.* J. Bacteriol., 1965. **89**: p. 1068-1074.

102. Smith, D.H. and B.D. Davis, *Mode of action of novobiocin in Escherichia coli.* J. Bacteriol., 1967. **93**: p. 71-79.

103. Drlica, K. and M. Snyder, *Superhelical Escherichia coli DNA: relaxation by coumermycin.* J. Mol. Biol., 1978. **120**: p. 145-154.

104. Steck, T., G. Pruss, S. Manes, L. Burg, and K. Drlica, *DNA Supercoiling in gyrase mutants.* Journal of Bacteriology, 1984. **158**: p. 397-403.

105. Lockshon, D. and D.R. Morris, *Positively supercoiled plasmid DNA is produced by treatment of Escherichia coli with DNA gyrase inhibitors.* Nucleic Acids Res, 1983. **11**(10): p. 2999-3017.

106. Menzel, R. and M. Gellert, *Regulation of the genes for E. coli DNA gyrase: homeostatic control of DNA supercoiling.* Cell, 1983. **34**(1): p. 105-13.

107. Botchan, P., J.C. Wang, and H. Echols, *Effect of circularity and superhelicity on transcription from bacteriophagelambda DNA.* Proc Natl Acad Sci U S A, 1973. **70**(11): p. 3077-81.

108. Funnell, B.E., T.A. Baker, and A. Kornberg, *Complete enzymatic replication of plasmids containing the origin of the Escherichia coli chromosome.* J Biol Chem, 1986. **261**(12): p. 5616-24.

109. Peter, B.J., J. Arsuaga, A.M. Breier, A.B. Khodursky, P.O. Brown, and N.R. Cozzarelli, *Genomic transcriptional response to loss of chromosomal supercoiling in Escherichia coli.* Genome Biol, 2004. **5**(11): p. R87.

110. Sanzey, B., *Modulation of gene expression by drugs affecting deoxyribonucleic acid gyrase.* J Bacteriol, 1979. **138**(1): p. 40-7.

111. Liu, L.F. and J.C. Wang, *Supercoiling of the DNA template during transcription.* Proc Natl Acad Sci U S A, 1987. **84**(20): p. 7024-7.

257

112. Kreuzer, K.N. and N.R. Cozzarelli, *Escherichia coli mutants thermosensitive for deoxyribonucleic acid gyrase subunit A: effects on deoxyribonucleic acid replication, transcription, and bacteriophage growth.* J. Bacteriol., 1979. **140**: p. 424-435.

113. Ahmed, W., C. Sala, S.R. Hegde, R.K. Jha, S.T. Cole, and V. Nagaraja, *Transcription facilitated genome-wide recruitment of topoisomerase I and DNA gyrase.* PLoS Genet, 2017. **13**(5): p. e1006754.

114. Hiasa, H. and K.J. Marians, *Two distinct modes of strand unlinking during theta-type DNA replication.* J Biol Chem, 1996. **271**(35): p. 21529-35.

115. Kato, J., Y. Nishimura, R. Imamura, H. Niki, S. Hiraga, and H. Suzuki, *New topoisomerase essential for chromosome segregation in E. coli.* Cell, 1990. **63**: p. 393-404.

116. Kato, J., H. Suzuki, and H. Ikeda, *Purification and characterization of DNA topoisomerase IV in Escherichia coli.* J Biol Chem, 1992. **267**(36): p. 25676-84.

117. Corbett, K.D., A.J. Schoeffler, N.D. Thomsen, and J.M. Berger, *The structural basis for substrate specificity in DNA topoisomerase IV.* J. Mol. Biol., 2005. **351**(3): p. 545-61.

118. Crisona, N.J., T.R. Strick, D. Bensimon, V. Croquette, and N.R. Cozzarelli, *Preferential relaxation of positively supercoiled DNA by E. coli topoisomerase IV in single-molecule and ensemble measurements.* Genes Dev, 2000. **14**(22): p. 2881-92.

119. Stone, M.D., Z. Bryant, N.J. Crisona, S.B. Smith, A. Vologodskii, C. Bustamante, and N.R. Cozzarelli, *Chirality sensing by Escherichia coli topoisomerase IV and the mechanism of type II topoisomerases.* Proc Natl Acad Sci U S A, 2003. **100**(15): p. 8654-9.

120. Neuman, K.C., G. Charvin, D. Bensimon, and V. Croquette, *Mechanisms of chiral discrimination by topoisomerase IV.* Proc Natl Acad Sci U S A, 2009. **106**(17): p. 6986-91.

121. Kampranis, S.C. and A. Maxwell, *Conversion of DNA gyrase into a conventional type II topoisomerase.* Proc Natl Acad Sci U S A, 1996. **93**(25): p. 14416-21.

122. Peng, H. and K.J. Marians, *Decatenation activity of topoisomerase IV during oriC and pBR322 DNA replication in vitro.* Proc Natl Acad Sci U S A, 1993. **90**(18): p. 8571-5.

123. Steck, T.R. and K. Drlica, *Bacterial chromosome segregation: evidence for DNA gyrase involvement in decatenation.* Cell, 1984. **36**(4): p. 1081-8.

124. Zechiedrich, E.L. and N.R. Cozzarelli, *Roles of topoisomerase IV and DNA gyrase in DNA unlinking during replication in Escherichia coli.* Genes Dev., 1995. **9**(22): p. 2859-69.

125. El Sayyed, H., L. Le Chat, E. Lebailly, E. Vickridge, C. Pages, F. Cornet, M. Cosentino Lagomarsino, and O. Espeli, *Mapping Topoisomerase IV Binding and Activity Sites on the E. coli Genome.* PLoS Genet, 2016. **12**(5): p. e1006025.

126. McNairn, E., N. Ní Bhriain, and C.J. Dorman, *Overexpression of the Shigella flexneri genes coding for DNA topoisomerase IV compensates for loss of DNA topoisomerase I: effect on virulence gene expression.* Molec. Microbiol., 1995. **15**: p. 507-517.

127. Raji, A., D. Zabel, C. Laufer, and R. Depew, *Genetic Analysis of Mutations That Compnsate for Loss of Escherichia coli DNA Topoisomerase I.* Journal of Bacteriology, 1985. **162**: p. 1173-1179.

128. Reuss, D.R., P. Fasshauer, P.J. Mroch, I. Ul-Haq, B.M. Koo, A. Pohlein, C.A. Gross, R. Daniel, S. Brantl, and J. Stulke, *Topoisomerase IV can functionally replace all type 1A topoisomerases in Bacillus subtilis.* Nucleic Acids Res, 2019. **47**(10): p. 5231-5242.

129. Khodursky, A.B., B.J. Peter, M.B. Schmid, J. DeRisi, D. Botstein, P.O. Brown, and N.R. Cozzarelli, *Analysis of topoisomerase function in bacterial replication fork movement: Use of DNA microarrays.* Proc. Natl. Acad. Sci. USA, 2000. **97**: p. 9419-9424.

130. Nicolas, E., A.L. Upton, S. Uphoff, O. Henry, A. Badrinarayanan, and D. Sherratt, *The SMC Complex MukBEF Recruits Topoisomerase IV to the Origin of Replication Region in Live Escherichia coli.* mBio, 2014. **5**(1): p. e01001-13.

131. Nolivos, S., A.L. Upton, A. Badrinarayanan, J. Müller, K. Zawadzka, J. Wiktor, A. Gill, L. Arciszewska, E. Nicolas, and D. Sherratt, *MatP regulates the coordinated action of*

topoisomerase IV and MukBEF in chromosome segregation. Nature Communications, 2016. **7**(1): p. 10466.

132. Hayama, R. and K.J. Marians, *Physical and functional interaction between the condensin MukB and the decatenase topoisomerase IV in Escherichia coli.* Proceedings of the National Academy of Sciences, 2010. **107**(44): p. 18826-18831.

133. Li, Y., N.K. Stewart, A.J. Berger, S. Vos, A.J. Schoeffler, J.M. Berger, B.T. Chait, and M.G. Oakley, *Escherichia coli condensin MukB stimulates topoisomerase IV activity by a direct physical interaction.* Proc Natl Acad Sci U S A, 2010. **107**(44): p. 18832-7.

134. Espeli, O., C. Levine, H. Hassing, and K.J. Marians, *Temporal regulation of topoisomerase IV activity in E. coli.* Mol Cell, 2003. **11**(1): p. 189-201.

135. Goto, T. and J.C. Wang, *Yeast DNA Topoisomerase II.* J. Biol. Chem., 1982. **257**: p. 5866-5872.

136. Goto, T., P. Laipis, and J. Wang, *The Purification and Characterization of DNA Topoisomerases I and II of the yeast sacchoromyces cereviviae.* Journal of Biological Chemistry, 1984. **259**: p. 10422-10429.

137. Wyckoff, E. and T.S. Hsieh, *Functional expression of a Drosophila gene in yeast: genetic complementation of DNA topoisomerase II.* Proc Natl Acad Sci U S A, 1988. **85**(17): p. 6272-6.

138. Baldi, M.I., P. Benedetti, E. Mattoccia, and G.P. Tocchini-Valentini, *In vitro catenation and decatenation of DNA and a novel eucaryotic ATP-dependent Topoisomerase.* Cell, 1980. **20**: p. 461-467.

139. Drake, F.H., J.P. Zimmerman, F.L. McCabe, H.F. Bartus, S.R. Per, D.M. Sullivan, W.E. Ross, M.R. Mattern, R.K. Johnson, S.T. Crooke, and C.K. Mirabelli, *Purification of topoisomerase II from amsacrine-resistant P388 leukemia cells.* J. Biol. Chem., 1987. **262**: p. 16739-16747.

140. Drake, F., G. Hofmann, H. Bartus, M. Mattern, S. Crooke, and C. Miracelli, *Biochemical and pharmacological properties of p170 and p180 forms of topoisomerase II.* Biochemistry, 1989. **28**: p. 8154-8160.

141. McClendon, A.K. and N. Osheroff, *DNA topoisomerase II, genotoxicity, and cancer.* Mutat Res, 2007. **623**(1-2): p. 83-97.

142. Wasserman, R.A., C.A. Austin, L.M. Fisher, and J.C. Wang, *Use of yeast in the study of anticancer drugs targeting DNA topoisomerases: expression of a functional recombinant human DNA topoisomerase IIa in yeast.* Cancer Res., 1993. **53**: p. 3591-3596.

143. Jensen, S., C.S. Redwood, J.R. Jenkins, A.H. Andersen, and I.D. Hickson, *Human DNA topoisomerases II alpha and II beta can functionally substitute for yeast TOP2 in chromosome segregation and recombination.* Mol Gen Genet, 1996. **252**(1-2): p. 79-86.

144. Bakshi, R.P., S. Galande, and K. Muniyappa, *Functional and regulatory characteristics of eukaryotic type II DNA topoisomerase.* Crit Rev Biochem Mol Biol, 2001. **36**(1): p. 1-37.

145. Goto, T. and J.C. Wang, *Yeast DNA topoisomerase II. An ATP-dependent type II topoisomerase that catalyzes the catenation, decatenation, unknotting, and relaxation of double-stranded DNA rings.* J Biol Chem, 1982. **257**(10): p. 5866-72.

146. Holm, C., T. Goto, J.C. Wang, and D. Botstein, *DNA topoisomerase II is required at the time of mitosis in yeast.* Cell, 1985. **41**(2): p. 553-63.

147. Joshi, R.S., B. Pina, and J. Roca, *Topoisomerase II is required for the production of long Pol II gene transcripts in yeast.* Nucleic Acids Res, 2012. **40**(16): p. 7907-15.

148. Meczes, E.L., K.L. Marsh, L.M. Fisher, M.P. Rogers, and C.A. Austin, *Complementation of temperature-sensitive topoisomerase II mutations in Saccharomyces cerevisiae by a human TOP2b construct allows the study of topoisomerase IIb inhibitors in yeast.* Cancer Chemother. Pharmacol., 1997. **39**: p. 367-375.

149. Nitiss, J.L., *Targeting DNA topoisomerase II in cancer chemotherapy.* Nat Rev Cancer, 2009. **9**(5): p. 338-50.

150. Nitiss, J.L., *Using yeast to study resistance to topoisomerase II-targeting drugs.* Cancer Chemother. Pharmacol., 1994. **34 (Suppl.)**: p. S6-S13.

151. Schmidt, B.H., N. Osheroff, and J.M. Berger, *Structure of a topoisomerase II-DNA-nucleotide complex reveals a new control mechanism for ATPase activity.* Nat Struct Mol Biol, 2012. **19**(11): p. 1147-54.

152. Deweese, J.E., M.A. Osheroff, and N. Osheroff, *DNA Topology and Topoisomerases: Teaching a "Knotty" Subject.* Biochem Mol Biol Educ, 2008. **37**(1): p. 2-10.

153. Capranico, G., S. Tinelli, C.A. Austin, M.L. Fisher, and F. Zunino, *Different patterns of gene expression of topoisomerase II isoforms in differentiated tissues during murine development.* Biochim Biophys Acta, 1992. **1132**(1): p. 43-8.

154. Woessner, R.D., M.R. Mattern, C.K. Mirabelli, R.K. Johnson, and F.H. Drake, *Proliferation- and cell cycle-dependent differences in expression of the 170 kilodalton and 180 kilodalton forms of topoisomerase II in NIH-3T3 cells.* Cell Growth Differ, 1991. **2**(4): p. 209-14.

155. Yu, X., J.W. Davenport, K.A. Urtishak, M.L. Carillo, S.J. Gosai, C.P. Kolaris, J.A.W. Byl, E.F. Rappaport, N. Osheroff, B.D. Gregory, and C.A. Felix, *Genome-wide TOP2A DNA cleavage is biased toward translocated and highly transcribed loci.* Genome Res, 2017. **27**(7): p. 1238-1249.

156. McClendon, A.K., A.C. Rodriguez, and N. Osheroff, *Human topoisomerase IIalpha rapidly relaxes positively supercoiled DNA: implications for enzyme action ahead of replication forks.* J Biol Chem, 2005. **280**(47): p. 39337-45.

157. McClendon, A.K., A.C. Gentry, J.S. Dickey, M. Brinch, S. Bendsen, A.H. Andersen, and N. Osheroff, *Bimodal recognition of DNA geometry by human topoisomerase II alpha: preferential relaxation of positively supercoiled DNA requires elements in the C-terminal domain.* Biochemistry, 2008. **47**(50): p. 13169-78.

158. Seol, Y., A.C. Gentry, N. Osheroff, and K.C. Neuman, *Chiral discrimination and writhe-dependent relaxation mechanism of human topoisomerase IIalpha.* J Biol Chem, 2013. **288**(19): p. 13695-703.

159. Li, Y., H. Hao, E. Tzatzalos, R.K. Lin, S. Doh, L.F. Liu, Y.L. Lyu, and L. Cai, *Topoisomerase IIbeta is required for proper retinal development and survival of postmitotic cells.* Biol Open, 2014. **3**(2): p. 172-84.

160. Zhang, Y.L., C. Yu, S.Y. Ji, X.M. Li, Y.P. Zhang, D. Zhang, D. Zhou, and H.Y. Fan, *TOP2beta is essential for ovarian follicles that are hypersensitive to chemotherapeutic drugs.* Mol Endocrinol, 2013. **27**(10): p. 1678-91.

161. Meyer-Ficca, M.L., J.D. Lonchar, M. Ihara, M.L. Meistrich, C.A. Austin, and R.G. Meyer, *Poly(ADP-ribose) polymerases PARP1 and PARP2 modulate topoisomerase II beta (TOP2B) function during chromatin condensation in mouse spermiogenesis.* Biol Reprod, 2011. **84**(5): p. 900-9.

162. Daev, E., N. Chaly, D.L. Brown, B. Valentine, J.E. Little, X. Chen, and P.R. Walker, *Role of topoisomerase II in the structural and functional evolution of mitogen-stimulated lymphocyte nuclei.* Exp Cell Res, 1994. **214**(1): p. 331-42.

163. 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, *Topoisomerase II beta interacts with cohesin and CTCF at topological domain borders.* Genome Biol, 2016. **17**(1): p. 182.

164. Linka, R.M., A.C. Porter, A. Volkov, C. Mielke, F. Boege, and M.O. Christensen, *C-terminal regions of topoisomerase IIalpha and IIbeta determine isoform-specific functioning of the enzymes in vivo.* Nucleic Acids Res, 2007. **35**(11): p. 3810-22.

165. Gadelle, D., M. Krupovic, K. Raymann, C. Mayer, and P. Forterre, *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**(13): p. 8578-91.

166. Takahashi, T.S., V. Da Cunha, M. Krupovic, C. Mayer, P. Forterre, and D. Gadelle, *Expanding the type IIB DNA topoisomerase family: identification of new topoisomerase*

and topoisomerase-like proteins in mobile genetic elements. NAR Genomics and Bioinformatics, 2020. **2**(1): p. lqz021.

167. Bergerat, A., D. Gadelle, and P. Forterre, *Purification of a DNA topoisomerase II from the hyperthermophilic archaeon Sulfolobus shibatae. A thermostable enzyme with both bacterial and eucaryal features.* J Biol Chem, 1994. **269**(44): p. 27663-9.

168. Gardner, M.J., N. Hall, E. Fung, O. White, M. Berriman, R.W. Hyman, J.M. Carlton, A. Pain, K.E. Nelson, and S. Bowman, *Genome sequence of the human malaria parasite Plasmodium falciparum.* Nature, 2002. **419**(6906): p. 498.

169. Sugimoto-Shirasu, K., N.J. Stacey, J. Corsar, K. Roberts, and M.C. McCann, *DNA topoisomerase VI is essential for endoreduplication in Arabidopsis.* Curr Biol, 2002. **12**(20): p. 1782-6.

170. Nichols, M.D., K. DeAngelis, J.L. Keck, and J.M. Berger, *Structure and function of an archaeal topoisomerase VI subunit with homology to the meiotic recombination factor Spo 11.* EMBO J., 1999. **18**: p. 6177-6188.

171. Bergerat, A., B. de Massy, D. Gadelle, P.C. Varoutas, A. Nicolas, and P. Forterre, *An atypical topoisomerase II from Archaea with implications for meiotic recombination.* Nature, 1997. **386**(6623): p. 414-7.

172. Keeney, S., C.N. Giroux, and N. Kleckner, *Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family.* Cell, 1997. **88**(3): p. 375-84.

173. Robert, T., A. Nore, C. Brun, C. Maffre, B. Crimi, H.M. Bourbon, and B. de Massy, *The TopoVIB-Like protein family is required for meiotic DNA double-strand break formation.* Science, 2016. **351**(6276): p. 943-9.

174. Vrielynck, N., A. Chambon, D. Vezon, L. Pereira, L. Chelysheva, A. De Muyt, C. Mezard, C. Mayer, and M. Grelon, *A DNA topoisomerase VI-like complex initiates meiotic recombination.* Science, 2016. **351**(6276): p. 939-43.

175. Hartung, F. and H. Puchta, *Molecular characterization of homologues of both subunits A (SPO11) and B of the archaebacterial topoisomerase 6 in plants.* Gene, 2001. **271**(1): p. 81-6.

176. Hartung, F. and H. Puchta, *Molecular characterisation of two paralogous SPO11 homologues in Arabidopsis thaliana.* Nucleic Acids Res, 2000. **28**(7): p. 1548-54.

177. Grelon, M., D. Vezon, G. Gendrot, and G. Pelletier, *AtSPO11-1 is necessary for efficient meiotic recombination in plants.* Embo J, 2001. **20**(3): p. 589-600.

178. Hartung, F., K.J. Angelis, A. Meister, I. Schubert, M. Melzer, and H. Puchta, *An archaebacterial topoisomerase homolog not present in other eukaryotes is indispensable for cell proliferation of plants.* Curr Biol, 2002. **12**(20): p. 1787-91.

179. Sugimoto-Shirasu, K. and K. Roberts, *"Big it up": endoreduplication and cell-size control in plants.* Curr Opin Plant Biol, 2003. **6**(6): p. 544-53.

180. Corbett, K.D. and J.M. Berger, *Emerging Roles for Plant Topoisomerase VI.* Chem Biol, 2003. **10**(2): p. 107-11.

181. Sugimoto-Shirasu, K., G.R. Roberts, N.J. Stacey, M.C. McCann, A. Maxwell, and K. Roberts, *RHL1 is an essential component of the plant DNA topoisomerase VI complex and is required for ploidy-dependent cell growth.* Proc Natl Acad Sci U S A, 2005. **102**(51): p. 18736-41.

182. Breuer, C., N.J. Stacey, C.E. West, Y. Zhao, J. Chory, H. Tsukaya, Y. Azumi, A. Maxwell, K. Roberts, and K. Sugimoto-Shirasu, *BIN4, a novel component of the plant DNA topoisomerase VI complex, is required for endoreduplication in Arabidopsis.* Plant Cell, 2007. **19**(11): p. 3655-68.

183. Rogers, S. and R. Wells, *Rechsteiner. M. 1986.* Amino acid sequences common to rapidly degraded proteins: The PEST hypothesis. Science. **234**: p. 364-368.

184. Austin, C.A. and K.L. Marsh, *Eukaryotic DNA topoisomerase IIb.* BioEssays, 1998. **20**: p. 215-226.

185.   Yin, Y., H. Cheong, D. Friedrichsen, Y. Zhao, J. Hu, S. Mora-Garcia, and J. Chory, *A crucial role for the putative Arabidopsis topoisomerase VI in plant growth and development.* Proc Natl Acad Sci U S A, 2002. **99**(15): p. 10191-6.

186.   Kirik, V., A. Schrader, J.F. Uhrig, and M. Hulskamp, *MIDGET Unravels Functions of the Arabidopsis Topoisomerase VI Complex in DNA Endoreduplication, Chromatin Condensation, and Transcriptional Silencing.* Plant Cell, 2007.

187.   Mittal, A., R. Balasubramanian, J. Cao, P. Singh, S. Subramanian, G. Hicks, E.A. Nothnagel, N. Abidi, J. Janda, D.W. Galbraith, and C.D. Rock, *TOPOISOMERASE 6B is involved in chromatin remodelling associated with control of carbon partitioning into secondary metabolites and cell walls, and epidermal morphogenesis in Arabidopsis.* J Exp Bot, 2014. **65**(15): p. 4217-39.

188.   Jain, M., A.K. Tyagi, and J.P. Khurana, *Overexpression of putative topoisomerase 6 genes from rice confers stress tolerance in transgenic Arabidopsis plants.* FEBS J, 2006. **273**(23): p. 5245-60.

189.   Simkova, K., F. Moreau, P. Pawlak, C. Vriet, A. Baruah, C. Alexandre, L. Hennig, K. Apel, and C. Laloi, *Integration of stress-related and reactive oxygen species-mediated signals by Topoisomerase VI in Arabidopsis thaliana.* Proc Natl Acad Sci U S A, 2012. **109**(40): p. 16360-5.

190.   Aravind, L., L.M. Iyer, T.E. Wellems, and L.H. Miller, *Plasmodium biology: genomic gleanings.* Cell, 2003. **115**(7): p. 771-85.

191.   Malik, S.B., M.A. Ramesh, A.M. Hulstrand, and J.M. Logsdon, Jr., *Protist homologs of the meiotic Spo11 gene and topoisomerase VI reveal an evolutionary history of gene duplication and lineage-specific loss.* Mol Biol Evol, 2007. **24**(12): p. 2827-41.

192.   Chalapareddy, S., S. Chakrabarty, M.K. Bhattacharyya, and S. Bhattacharyya, *Radicicol-Mediated Inhibition of Topoisomerase VIB-VIA Activity of the Human Malaria Parasite Plasmodium falciparum.* mSphere, 2016. **1**(1).

193. Buhler, C., D. Gadelle, P. Forterre, J.C. Wang, and A. Bergerat, *Reconstitution of DNA topoisomerase VI of the thermophilic archaeon Sulfolobus shibatae from subunits separately overexpressed in Escherichia coli.* Nucleic Acids Res, 1998. **26**(22): p. 5157-62.

194. Buhler, C., J.H. Lebbink, C. Bocs, R. Ladenstein, and P. Forterre, *DNA topoisomerase VI generates ATP-dependent double-strand breaks with two-nucleotide overhangs.* J. Biol. Chem., 2001. **276**(40): p. 37215-22.

195. Corbett, K.D., P. Benedetti, and J.M. Berger, *Holoenzyme assembly and ATP-mediated conformational dynamics of topoisomerase VI.* Nat Struct Mol Biol, 2007. **14**(7): p. 611-9.

196. Graille, M., L. Cladiere, D. Durand, F. Lecointe, D. Gadelle, S. Quevillon-Cheruel, P. Vachette, P. Forterre, and H. van Tilbeurgh, *Crystal structure of an intact type II DNA topoisomerase: insights into DNA transfer mechanisms.* Structure, 2008. **16**(3): p. 360-70.

197. Roca, J. and J.C. Wang, *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, 1992. **71**: p. 833-840.

198. Roca, J. and J.C. Wang, *DNA transport by a type II DNA topoisomerase: evidence in favor of a two-gate mechanism.* Cell, 1994. **77**: p. 609-616.

199. Klapholz, S., C.S. Waddell, and R.E. Esposito, *The Role of the Spo11 Gene in Meiotic Recombination in Yeast.* Genetics, 1985. **110**(2): p. 187-216.

200. Esposito, M.S., R.E. Esposito, M. Arnaud, and H.O. Halvorson, *Acetate Utilization and Macromolecular Synthesis during Sporulation of Yeast.* Journal of Bacteriology, 1969. **100**(1): p. 180-+.

201. Neale, M.J., J. Pan, and S. Keeney, *Endonucleolytic processing of covalent protein-linked DNA double-strand breaks.* Nature, 2005. **436**(7053): p. 1053-7.

202. Lam, I. and S. Keeney, *Mechanism and regulation of meiotic recombination initiation.* Cold Spring Harbor perspectives in biology, 2015. **7**(1): p. a016634.

203. Keeney, S., *Spo11 and the Formation of DNA Double-Strand Breaks in Meiosis.* Genome Dyn Stab, 2008. **2**: p. 81-123.

204. Keeney, S. and M.J. Neale, *Initiation of meiotic recombination by formation of DNA double-strand breaks: mechanism and regulation.* Biochem Soc Trans, 2006. **34**(Pt 4): p. 523-5.

205. Pan, J., M. Sasaki, R. Kniewel, H. Murakami, H.G. Blitzblau, S.E. Tischfield, X. Zhu, M.J. Neale, M. Jasin, N.D. Socci, A. Hochwagen, and S. Keeney, *A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation.* Cell, 2011. **144**(5): p. 719-31.

206. Hartung, F., R. Wurz-Wildersinn, J. Fuchs, I. Schubert, S. Suer, and H. Puchta, *The Catalytically Active Tyrosine Residues of Both SPO11-1 and SPO11-2 Are Required for Meiotic Double-Strand Break Induction in Arabidopsis.* Plant Cell, 2007.

207. Stacey, N.J., T. Kuromori, Y. Azumi, G. Roberts, C. Breuer, T. Wada, A. Maxwell, K. Roberts, and K. Sugimoto-Shirasu, *Arabidopsis SPO11-2 functions with SPO11-1 in meiotic recombination.* Plant J, 2006. **48**(2): p. 206-16.

208. Fu, M., C. Wang, F. Xue, J. Higgins, M. Chen, D. Zhang, and W. Liang, *The DNA Topoisomerase VI-B Subunit OsMTOPVIB Is Essential for Meiotic Recombination Initiation in Rice.* Mol Plant, 2016. **9**(11): p. 1539-1541.

209. Xue, Z., Y. Li, L. Zhang, W. Shi, C. Zhang, M. Feng, F. Zhang, D. Tang, H. Yu, and M. Gu, *OsMTOPVIB promotes meiotic DNA double-strand break formation in rice.* Molecular plant, 2016. **9**(11): p. 1535-1538.

210. Bouuaert, C.C., S.E. Tischfield, S. Pu, E.P. Mimitou, E. Arias-Palomo, J.M. Berger, and S. Keeney, *Structural and functional characterization of the Spo11 core complex.* bioRxiv, 2020.

211. Sugino, A., C.L. Peebles, K.N. Kreuzer, and N.R. Cozzarelli, *Mechanism of action of nalidixic acid: purification of Escherichia coli nalA gene product and its relationship to*

DNA gyrase and a novel nicking-closing enzyme. Proc Natl Acad Sci U S A, 1977. **74**(11): p. 4767-71.

212. van Maanen, J.M., J. Retel, J. de Vries, and H.M. Pinedo, *Mechanism of action of antitumor drug etoposide: a review.* J Natl Cancer Inst, 1988. **80**(19): p. 1526-33.

213. Maxwell, A. and D.M. Lawson, *The ATP-binding site of type II topoisomerases as a target for antibacterial drugs.* Curr. Top. Med. Chem., 2003. **3**(1): p. 283-303.

214. Heide, L. and S.-M. Li, *Aminocoumarin antibiotics*, in *Microbial Secondary metabolites: Biosynthesis, Genetics and Regulation*. 2002.

215. Gadelle, D., C. Bocs, M. Graille, and P. Forterre, *Inhibition of archaeal growth and DNA topoisomerase VI activities by the Hsp90 inhibitor radicicol.* Nucleic Acids Res, 2005. **33**(7): p. 2310-7.

216. Taylor, J.A., L.A. Mitchenall, M. Rejzek, R.A. Field, and A. Maxwell, *Application of a novel microtitre plate-based assay for the discovery of new inhibitors of DNA gyrase and DNA topoisomerase VI.* PLoS One, 2013. **8**(2): p. e58010.

217. Rudolph, M.G. and D. Klostermeier, *Mapping the Spectrum of Conformational States of the DNA- and C-Gates in Bacillus subtilis Gyrase.* J Mol Biol, 2013. **425**(15): p. 2632-40.

218. Gubaev, A., M. Hilbert, and D. Klostermeier, *The DNA-gate of Bacillus subtilis gyrase is predominantly in the closed conformation during the DNA supercoiling reaction.* Proc Natl Acad Sci U S A, 2009. **106**(32): p. 13278-83.

219. Tadesse, S. and P.L. Graumann, *Differential and dynamic localization of topoisomerases in Bacillus subtilis.* J Bacteriol, 2006. **188**(8): p. 3002-11.

220. Neuman, K.C., *Single-molecule measurements of DNA topology and topoisomerases.* J Biol Chem, 2010. **285**(25): p. 18967-71.

221. Seol, Y. and K.C. Neuman, *Magnetic Tweezers for Single-Molecule Manipulation*, in *Single Molecule Analysis: Methods and Protocols*, E.J.G. Peterman and G.J.L. Wuite, Editors. 2011, Humana Press: Totowa, NJ. p. 265-293.

222. Seol, Y. and K.C. Neuman, *Single-molecule measurements of topoisomerase activity with magnetic tweezers*, in *Single molecule enzymology*. 2011, Springer. p. 229-241.

223. Charvin, G., A. Vologodskii, D. Bensimon, and V. Croquette, *Braiding DNA: experiments, simulations, and models.* Biophysical journal, 2005. **88**(6): p. 4124-4136.

224. Mills, M., Y.C. Tse-Dinh, and K.C. Neuman, *Direct observation of topoisomerase IA gate dynamics.* Nat Struct Mol Biol, 2018. **25**(12): p. 1111-1118.

225. Litwin, T.R., M. Sola, I.J. Holt, and K.C. Neuman, *A robust assay to measure DNA topology-dependent protein binding affinity.* Nucleic Acids Res, 2014.

226. Maxwell, A. and M. Gellert, *The DNA dependence of the ATPase activity of DNA gyrase.* J Biol Chem, 1984. **259**(23): p. 14472-80.

227. Seol, Y., M.-P. Strub, and K.C. Neuman, *Single molecule measurements of DNA helicase activity with magnetic tweezers and t-test based step-finding analysis.* Methods, 2016. **105**: p. 119-127.

228. Crooks, G.E., G. Hon, J.M. Chandonia, and S.E. Brenner, *WebLogo: a sequence logo generator.* Genome Res, 2004. **14**(6): p. 1188-90.

229. Hardin, A.H., S.K. Sarkar, Y. Seol, G.F. Liou, N. Osheroff, and K.C. Neuman, *Direct measurement of DNA bending by type IIA topoisomerases: implications for non-equilibrium topology simplification.* Nucleic Acids Res, 2011. **39**(13): p. 5729-43.

230. Rybenkov, V.V., C. Ullsperger, A.V. Vologodskii, and N.R. Cozzarelli, *Simplification of DNA topology below equilibrium values by type II topoisomerases.* Science, 1997. **277**: p. 690-693.

231. Stuchinskaya, T., L.A. Mitchenall, A.J. Schoeffler, K.D. Corbett, J.M. Berger, A.D. Bates, and A. Maxwell, *How do type II topoisomerases use ATP hydrolysis to simplify DNA topology beyond equilibrium? Investigating the relaxation reaction of nonsupercoiling type II topoisomerases.* J Mol Biol, 2009. **385**(5): p. 1397-408.

232. Tan, S., *A modular polycistronic expression system for overexpressing protein complexes in Escherichia coli.* Protein Expr Purif, 2001. **21**(1): p. 224-34.

270

233. Some, D., H. Amartely, A. Tsadok, and M. Lebendiker, *Characterization of Proteins by Size-Exclusion Chromatography Coupled to Multi-Angle Light Scattering (SEC-MALS).* J Vis Exp, 2019(148).

234. Ortega-Arroyo, J. and P. Kukura, *Interferometric scattering microscopy (iSCAT): new frontiers in ultrafast and ultrasensitive optical microscopy.* Phys Chem Chem Phys, 2012. **14**(45): p. 15625-36.

235. Bolivar, F., R.L. Rodriguez, P.J. Greene, M.C. Betlach, H.L. Heyneker, and H.W. Boyer, *Construction and characterization of new cloning vehicles. II. A multipurpose cloning system.* Gene, 1977. **2**(2): p. 95-113.

236. Boros, I., G. Pósfai, and P. Venetianer, *High-copy-number derivatives of the plasmid cloning vector pBR322.* Gene, 1984. **30**: p. 257-260.

237. McClendon, A.K. and N. Osheroff, *The geometry of DNA supercoils modulates topoisomerase-mediated DNA cleavage and enzyme response to anticancer drugs.* Biochemistry, 2006. **45**(9): p. 3040-50.

238. Battumur, U., Y. Yoon, G.S. Bae, and C.H. Kim, *Isolation and characterization of new Methanosarcina mazei strains KOR-3, -4, -5, and -6 from an anaerobic digester using pig slurry.* Asian-Australas J Anim Sci, 2017. **30**(8): p. 1198-1205.

239. Rosenberg, E., E.F. DeLong, S. Lory, E. Stackebrandt, and F. Thompson, *The prokaryotes: other major lineages of Bacteria and the Archaea*. 2014: Springer.

240. Stewart, L., G.C. Ireton, L.H. Parker, K.R. Madden, and J.J. Champoux, *Biochemical and biophysical analyses of recombinant forms of human topoisomerase I.* J. Biol. Chem., 1996. **271**: p. 7593-7601.

241. Stivers, J.T., S. Shuman, and A.S. Mildvan, *Vaccinia DNA topoisomerase I: single-turnover and steady state kinetic analysis of the DNA strand cleavage and ligation reactions.* J. Biol. Chem., 1994. **33**: p. 327-339.

242. Reece, R.J. and A. Maxwell, *Tryptic fragments of the Escherichia coli DNA gyrase A protein.* J Biol Chem, 1989. **264**(33): p. 19648-53.

243. Kampranis, S.C. and A. Maxwell, *Conformational changes in DNA gyrase revealed by limited proteolysis.* J. Biol. Chem., 1998. **273**: p. 22606-22614.

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

245. Gilroy, K.L. and C.A. Austin, *The impact of the C-terminal domain on the interaction of human DNA topoisomerase II alpha and beta with DNA.* PLoS ONE, 2011. **6**(2): p. e14693.

246. Aubry, A., L.M. Fisher, V. Jarlier, and E. Cambau, *First functional characterization of a singly expressed bacterial type II topoisomerase: the enzyme from Mycobacterium tuberculosis.* Biochem Biophys Res Commun, 2006. **348**(1): p. 158-65.

247. Osheroff, N., E.L. Zechiedrich, and K. Gale, *Catalytic function of DNA topoisomerase II.* BioEssays, 1991. **13**: p. 269-275.

248. Shuman, S., M. Golder, and B. Moss, *Characterization of vaccinia virus DNA topoisomerase I expressed in Escherichia coli.* J Biol Chem, 1988. **263**(31): p. 16401-7.

249. Osheroff, N., E.R. Shelton, and D.L. Brutlag, *DNA topoisomerase II from Drosophila melangaster. Relaxation of supercoiled DNA.* J. Biol. Chem., 1983. **258**: p. 9536-9543.

250. Baldwin, E.L., J.A. Byl, and N. Osheroff, *Cobalt enhances DNA cleavage mediated by human topoisomerase IIalpha in vitro and in cultured cells.* Biochemistry, 2004. **43**(3): p. 728-35.

251. Shuman, R.M. and M.A. McBride, *Resistance of chickens to Rous sarcoma virus challenge following immunization with a recombinant avian leukosis virus.* Avian Dis, 1988. **32**(3): p. 410-5.

252. Hartwig, A. and T. Schwerdtle, *Interactions by carcinogenic metal compounds with DNA repair processes: toxicological implications.* Toxicol Lett, 2002. **127**(1-3): p. 47-54.

253.   Kawanishi, S., Y. Hiraku, M. Murata, and S. Oikawa, *The role of metals in site-specific DNA damage with reference to carcinogenesis.* Free Radic Biol Med, 2002. **32**(9): p. 822-32.

254.   Westermann, P., B.K. Ahring, and R.A. Mah, *Acetate production by methanogenic bacteria.* Appl Environ Microbiol, 1989. **55**(9): p. 2257-61.

255.   Westermann, P., B.K. Ahring, and R.A. Mah, *Threshold acetate concentrations for acetate catabolism by aceticlastic methanogenic bacteria.* Appl Environ Microbiol, 1989. **55**(2): p. 514-5.

256.   Shapiro, T.A. and P.T. Englund, *The structure and replication of kinetoplast DNA.* Annu Rev Microbiol, 1995. **49**: p. 117-43.

257.   Charvin, G., D. Bensimon, and V. Croquette, *Single-molecule study of DNA unlinking by eukaryotic and prokaryotic type-II topoisomerases.* Proc Natl Acad Sci U S A, 2003. **100**(17): p. 9820-5.

258.   Burden, D.A., P.S. Kingma, S.J. Froelich-Ammon, M.-A. Bjornsti, M.W. Patchan, R.B. Thompson, and N. Osheroff, *Topoisomerase IIŠetoposide interactions direct the formation of drug-induced enzyme-DNA cleavage complexes.* J. Biol. Chem., 1996. **271**: p. 29238-29244.

259.   Pommier, Y., E. Leo, H. Zhang, and C. Marchand, *DNA topoisomerases and their poisoning by anticancer and antibacterial drugs.* Chem Biol, 2010. **17**(5): p. 421-33.

260.   Drlica, K. and X. Zhao, *DNA gyrase, topoisomerase IV, and the 4-quinolones.* Microbiol. Mol. Biol. Revs., 1997. **61**: p. 377-392.

261.   Mueller-Planitz, F. and D. Herschlag, *Coupling between ATP binding and DNA cleavage by DNA topoisomerase II: A unifying kinetic and structural mechanism.* J Biol Chem, 2008. **283**(25): p. 17463-76.

262.   Germe, T., J. Voros, F. Jeannot, T. Taillier, R.A. Stavenger, E. Bacque, A. Maxwell, and B.D. Bax, *A new class of antibacterials, the imidazopyrazinones, reveal structural transitions involved in DNA gyrase poisoning and mechanisms of resistance.* Nucleic Acids Res, 2018. **46**(8): p. 4114-4128.

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

264. Deweese, J.E. and N. Osheroff, *Coordinating the two protomer active sites of human topoisomerase IIalpha: nicks as topoisomerase II poisons.* Biochemistry, 2009. **48**(7): p. 1439-41.

265. Lloyd, D.R. and D.H. Phillips, *Oxidative DNA damage mediated by copper(II), iron(II) and nickel(II) fenton reactions: evidence for site-specific mechanisms in the formation of double-strand breaks, 8-hydroxydeoxyguanosine and putative intrastrand cross-links.* Mutat Res, 1999. **424**(1-2): p. 23-36.

266. Xiao, H., T.K. Li, J.M. Yang, and L.F. Liu, *Acidic pH induces topoisomerase II-mediated DNA damage.* Proc Natl Acad Sci U S A, 2003. **100**(9): p. 5205-10.

267. Barry, M.A. and A. Eastman, *Identification of deoxyribonuclease II as an endonuclease involved in apoptosis.* Arch Biochem Biophys, 1993. **300**(1): p. 440-50.

268. Pingoud, A., M. Fuxreiter, V. Pingoud, and W. Wende, *Type II restriction endonucleases: structure and mechanism.* Cell Mol Life Sci, 2005. **62**(6): p. 685-707.

269. Shevelev, I.V. and U. Hubscher, *The 3' 5' exonucleases.* Nat Rev Mol Cell Biol, 2002. **3**(5): p. 364-76.

270. Shen, B., W. Zhang, J. Zhang, J. Zhou, J. Wang, L. Chen, L. Wang, A. Hodgkins, V. Iyer, X. Huang, and W.C. Skarnes, *Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects.* Nat Methods, 2014. **11**(4): p. 399-402.

271. Higgins, N.P. and N.R. Cozzarelli, *The binding of gyrase to DNA: analysis by retention by nitrocellulose filters.* Nucleic Acids Res., 1982. **10**: p. 6833-6847.

272. Subramanian, D., E. Kraut, A. Staubus, D.C. Young, and M.T. Muller, *Analysis of topoisomerase I/DNA complexes in patients administered topotecan.* Cancer Res, 1995. **55**(10): p. 2097-103.

274

273. Sullivan, D.M., M.D. Latham, T.C. Rowe, and W.E. Ross, *Purification and characterization of an altered topoisomerase II from a drug-resistant Chinese hamster ovary cell line.* Biochemistry, 1989. **28**(13): p. 5680-7.

274. Shuman, S. and J. Prescott, *Specific DNA cleavage and binding by vaccinia virus DNA topoisomerase I.* J Biol Chem, 1990. **265**(29): p. 17826-36.

275. Marians, K.J. and H. Hiasa, *Mechanism of quinolone action. A drug-induced structural perturbation of the DNA precedes strand cleavage by topoisomerase IV.* J. Biol. Chem., 1997. **272**: p. 9401-9409.

276. Sander, M., T. Hsieh, A. Udvardy, and P. Schedl, *Sequence dependence of Drosophila topoisomerase II in plasmid relaxation and DNA binding.* J Mol Biol, 1987. **194**(2): p. 219-29.

277. Yang, Z. and J.J. Champoux, *Assays for the preferential binding of human topoisomerase I to supercoiled DNA.* Methods Mol Biol, 2009. **582**: p. 49-57.

278. Mitchenall, L.A., R.E. Hipkin, M.M. Piperakis, N.P. Burton, and A. Maxwell, *A rapid high-resolution method for resolving DNA topoisomers.* BMC Res Notes, 2018. **11**(1): p. 37.

279. Bates, A.D. and A. Maxwell, *Energy Coupling in Type II Topoisomerases: Why Do They Hydrolyze ATP?* Biochemistry, 2007. **46**: p. 7929-7941.

280. Bates, A.D. and A. Maxwell, *The role of ATP in the reactions of type II DNA topoisomerases.* Biochem Soc Trans, 2010. **38**(2): p. 438-42.

281. Sifaoui, F., V. Lamour, E. Varon, D. Moras, and L. Gutmann, *ATP-bound conformation of topoisomerase IV: a possible target for quinolones in Streptococcus pneumoniae.* J Bacteriol, 2003. **185**(20): p. 6137-46.

282. Gubaev, A. and D. Klostermeier, *DNA-induced narrowing of the gyrase N-gate coordinates T-segment capture and strand passage.* Proc Natl Acad Sci U S A, 2011. **108**(34): p. 14085-90.

283. Ban, C., M. Junop, and W. Yang, *Transformation of MutL by ATP binding and hydrolysis: a switch in DNA mismatch repair.* Cell, 1999. **97**(1): p. 85-97.

284. Ban, C. and W. Yang, *Crystal structure and ATPase activity of MutL: implications for DNA repair and mutagenesis.* Cell, 1998. **95**: p. 541-552.

285. Baird, C.L., M.S. Gordon, D.M. Andrenyak, J.F. Marecek, and J.E. Lindsley, *The ATPase reaction cycle of yeast DNA topoisomerase II. Slow rates of ATP resynthesis and P(i) release.* J Biol Chem, 2001. **276**(30): p. 27893-8.

286. Boonyalai, N., P. Sittikul, N. Pradidphol, and N. Kongkathip, *Biophysical and molecular docking studies of naphthoquinone derivatives on the ATPase domain of human topoisomerase II.* Biomed Pharmacother, 2013. **67**(2): p. 122-8.

287. Hockings, S.C. and A. Maxwell, *Identification of four GyrA residues involved in the DNA breakage-reunion reaction of DNA gyrase.* J. Mol. Biol., 2002. **318**(2): p. 351-359.

288. Heddle, J.G., T. Lu, X. Zhao, K. Drlica, and A. Maxwell, *gyrB-225, a mutation of DNA gyrase that compensates for topoisomerase I deficiency: investigation of its low activity and quinolone hypersensitivity.* J. Mol. Biol., 2001. **309**: p. 1219-1231.

289. Bergant, K., M. Janezic, and A. Perdih, *Bioassays and In Silico Methods in the Identification of Human DNA Topoisomerase IIalpha Inhibitors.* Curr Med Chem, 2018. **25**(28): p. 3286-3318.

290. Ali, J.A., A.P. Jackson, A.J. Howells, and A. Maxwell, *The 43-kilodalton N-terminal fragment of the DNA gyrase B protein hydrolyzes ATP and binds coumarin drugs.* Biochemistry, 1993. **32**(10): p. 2717-24.

291. Cullis, P.M., A. Maxwell, and D.P. Weiner, *Energy coupling in DNA gyrase: a thermodynamic limit to the extent of DNA supercoiling.* Biochemistry, 1992. **31**: p. 9642-9646.

292. Williams, N.L. and A. Maxwell, *Locking the DNA gate of DNA gyrase: investigating the effects on DNA cleavage and ATP hydrolysis.* Biochemistry, 1999. **38**(43): p. 14157-64.

293. Williams, N.L. and A. Maxwell, *Probing the two-gate mechanism of DNA gyrase using cysteine cross-linking.* Biochemistry, 1999. **38**(41): p. 13502-11.

294. Pierrat, O.A. and A. Maxwell, *Evidence for the role of DNA strand passage in the mechanism of action of microcin B17 on DNA gyrase.* Biochemistry, 2005. **44**(11): p. 4204-15.

295. Sugino, A. and N.R. Cozzarelli, *The intrinsic ATPase of DNA gyrase.* J Biol Chem, 1980. **255**(13): p. 6299-306.

296. Kampranis, S.C. and A. Maxwell, *The DNA gyrase-quinolone complex. ATP hydrolysis and the mechanism of DNA cleavage.* J Biol Chem, 1998. **273**(35): p. 22615-26.

297. Sugino, A., N.P. Higgins, P.O. Brown, C.L. Peebles, and N.R. Cozzarelli, *Energy coupling in DNA gyrase and the mechanism of action of novobiocin.* Proc. Natl. Acad. Sci. USA, 1978. **75**: p. 4838-4842.

298. Corbett, K.D. and J.M. Berger, *Structural basis for topoisomerase VI inhibition by the anti-Hsp90 drug radicicol.* Nucleic Acids Res, 2006. **34**(15): p. 4269-77.

299. Crick, F.H.C. and A.F.W. Hughes, *The physical properties of cytoplasm: A study by means of the magnetic particle method Part I. Experimental.* Experimental Cell Research, 1950. **1**(1): p. 37-80.

300. Smith, S.B., L. Finzi, and C. Bustamante, *Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads.* Science, 1992. **258**(5085): p. 1122-6.

301. Strick, T.R., J.F. Allemand, D. Bensimon, A. Bensimon, and V. Croquette, *The elasticity of a single supercoiled DNA molecule.* Science, 1996. **271**(5257): p. 1835-7.

302. Yang, Y., M. Bai, W.S. Klug, A.J. Levine, and M.T. Valentine, *Microrheology of highly crosslinked microtubule networks is dominated by force-induced crosslinker unbinding.* Soft Matter, 2013. **9**(2): p. 383-393.

303. Adhikari, A.S., J. Chai, and A.R. Dunn, *Mechanical load induces a 100-fold increase in the rate of collagen proteolysis by MMP-1.* J Am Chem Soc, 2011. **133**(6): p. 1686-9.

304. Chen, H., H. Fu, X. Zhu, P. Cong, F. Nakamura, and J. Yan, *Improved high-force magnetic tweezers for stretching and refolding of proteins and short DNA.* Biophys J, 2011. **100**(2): p. 517-23.

305. del Rio, A., R. Perez-Jimenez, R. Liu, P. Roca-Cusachs, J.M. Fernandez, and M.P. Sheetz, *Stretching single talin rod molecules activates vinculin binding.* Science, 2009. **323**(5914): p. 638-41.

306. Kilinc, D. and G.U. Lee, *Advances in magnetic tweezers for single molecule and cell biophysics.* Integr Biol (Camb), 2014. **6**(1): p. 27-34.

307. Haber, C. and D. Wirtz, *Magnetic tweezers for DNA micromanipulation.* Review of Scientific instruments, 2000. **71**(12): p. 4561-4570.

308. Mosconi, F., J.F. Allemand, D. Bensimon, and V. Croquette, *Measurement of the torque on a single stretched and twisted DNA using magnetic tweezers.* Phys Rev Lett, 2009. **102**(7): p. 078301.

309. Dittmore, A., S. Brahmachari, Y. Takagi, J.F. Marko, and K.C. Neuman, *Supercoiling DNA Locates Mismatches.* Phys Rev Lett, 2017. **119**(14): p. 147801.

310. Hodeib, S., S. Raj, M. Manosas, W. Zhang, D. Bagchi, B. Ducos, J.F. Allemand, D. Bensimon, and V. Croquette, *Single molecule studies of helicases with magnetic tweezers.* Methods, 2016. **105**: p. 3-15.

311. Klaue, D., D. Kobbe, F. Kemmerich, A. Kozikowska, H. Puchta, and R. Seidel, *Fork sensing and strand switching control antagonistic activities of RecQ helicases.* Nat Commun, 2013. **4**: p. 2024.

312. Dulin, D., J.J. Arnold, T. van Laar, H.S. Oh, C. Lee, A.L. Perkins, D.A. Harki, M. Depken, C.E. Cameron, and N.H. Dekker, *Signatures of Nucleotide Analog Incorporation by an RNA-Dependent RNA Polymerase Revealed Using High-Throughput Magnetic Tweezers.* Cell Rep, 2017. **21**(4): p. 1063-1076.

313. Herbert, K.M., W.J. Greenleaf, and S.M. Block, *Single-molecule studies of RNA polymerase: motoring along.* Annu Rev Biochem, 2008. **77**: p. 149-76.

314. Lipfert, J., D.A. Koster, I.D. Vilfan, S. Hage, and N.H. Dekker, *Single-molecule magnetic tweezers studies of type IB topoisomerases.* Methods Mol Biol, 2009. **582**: p. 71-89.

315. Neuman, K.C. and A. Nagy, *Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy.* Nat Methods, 2008. **5**(6): p. 491-505.

316. Strick, T.R., J.-F. Allemand, D. Bensimon, A. Bensimon, and V. Croquette, *The elasticity of a single supercoiled DNA molecule.* Science, 1996. **271**: p. 1835-1837.

317. van Loenhout, M.T., M.V. de Grunt, and C. Dekker, *Dynamics of DNA supercoils.* Science, 2012. **338**(6103): p. 94-7.

318. Nollmann, M., M.D. Stone, Z. Bryant, J. Gore, N.J. Crisona, S.C. Hong, S. Mitelheiser, A. Maxwell, C. Bustamante, and N.R. Cozzarelli, *Multiple modes of Escherichia coli DNA gyrase activity revealed by force and torque.* Nat Struct Mol Biol, 2007. **14**(4): p. 264-71.

319. Basu, A., A.J. Schoeffler, J.M. Berger, and Z. Bryant, *ATP binding controls distinct structural transitions of Escherichia coli DNA gyrase in complex with DNA.* Nat Struct Mol Biol, 2012. **19**(5): p. 538-46.

320. Hays, F.A., A. Teegarden, Z.J. Jones, M. Harms, D. Raup, J. Watson, E. Cavaliere, and P.S. Ho, *How sequence defines structure: a crystallographic map of DNA structure and conformation.* Proc Natl Acad Sci U S A, 2005. **102**(20): p. 7157-62.

321. Charvin, G., T.R. Strick, D. Bensimon, and V. Croquette, *Topoisomerase IV bends and overtwists DNA upon binding.* Biophys J, 2005. **89**(1): p. 384-92.

322. Sanger, F., S. Nicklen, and A.R. Coulson, *DNA sequencing with chain-terminating inhibitors.* Proc Natl Acad Sci U S A, 1977. **74**(12): p. 5463-7.

323. Liu, L., Y. Li, S. Li, N. Hu, Y. He, R. Pong, D. Lin, L. Lu, and M. Law, *Comparison of next-generation sequencing systems.* J Biomed Biotechnol, 2012. **2012**: p. 251364.

324. International Human Genome Sequencing, C., *Finishing the euchromatic sequence of the human genome.* Nature, 2004. **431**(7011): p. 931-45.

325. van Dijk, E.L., H. Auger, Y. Jaszczyszyn, and C. Thermes, *Ten years of next-generation sequencing technology.* Trends Genet, 2014. **30**(9): p. 418-26.

326. Margulies, M., M. Egholm, W.E. Altman, S. Attiya, J.S. Bader, L.A. Bemben, J. Berka, M.S. Braverman, Y.J. Chen, Z. Chen, S.B. Dewell, L. Du, J.M. Fierro, X.V. Gomes, B.C.

Godwin, W. He, S. Helgesen, C.H. Ho, G.P. Irzyk, S.C. Jando, M.L. Alenquer, T.P. Jarvie, K.B. Jirage, J.B. Kim, J.R. Knight, J.R. Lanza, J.H. Leamon, S.M. Lefkowitz, M. Lei, J. Li, K.L. Lohman, H. Lu, V.B. Makhijani, K.E. McDade, M.P. McKenna, E.W. Myers, E. Nickerson, J.R. Nobile, R. Plant, B.P. Puc, M.T. Ronan, G.T. Roth, G.J. Sarkis, J.F. Simons, J.W. Simpson, M. Srinivasan, K.R. Tartaro, A. Tomasz, K.A. Vogt, G.A. Volkmer, S.H. Wang, Y. Wang, M.P. Weiner, P. Yu, R.F. Begley, and J.M. Rothberg, *Genome sequencing in microfabricated high-density picolitre reactors.* Nature, 2005. **437**(7057): p. 376-80.

327.   Mardis, E.R., *The impact of next-generation sequencing technology on genetics.* Trends Genet, 2008. **24**(3): p. 133-41.

328.   Friedberg, E.C., *DNA damage and repair.* Nature, 2003. **421**(6921): p. 436-40.

329.   Jackson, S.P. and J. Bartek, *The DNA-damage response in human biology and disease.* Nature, 2009. **461**(7267): p. 1071-8.

330.   Crosetto, N., A. Mitra, M.J. Silva, M. Bienko, N. Dojer, Q. Wang, E. Karaca, R. Chiarle, M. Skrzypczak, K. Ginalski, P. Pasero, M. Rowicka, and I. Dikic, *Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing.* Nat Methods, 2013. **10**(4): p. 361-5.

331.   Canela, A., S. Sridharan, N. Sciascia, A. Tubbs, P. Meltzer, B.P. Sleckman, and A. Nussenzweig, *DNA Breaks and End Resection Measured Genome-wide by End Sequencing.* Mol Cell, 2016. **63**(5): p. 898-911.

332.   Holliday, R., *A mechanism for gene conversion in fungi.* Genet Res, 2007. **89**(5-6): p. 285-307.

333.   Baranello, L., F. Kouzine, D. Wojtowicz, K. Cui, T.M. Przytycka, K. Zhao, and D. Levens, *DNA break mapping reveals topoisomerase II activity genome-wide.* Int J Mol Sci, 2014. **15**(7): p. 13111-22.

334.   Rigby, P.W., M. Dieckmann, C. Rhodes, and P. Berg, *Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I.* J Mol Biol, 1977. **113**(1): p. 237-51.

335.   Muslimovic, A., S. Nystrom, Y. Gao, and O. Hammarsten, *Numerical analysis of etoposide induced DNA breaks.* PLoS ONE, 2009. **4**(6): p. e5859.

336.   Schmidt, D., M.D. Wilson, C. Spyrou, G.D. Brown, J. Hadfield, and D.T. Odom, *ChIP-seq: using high-throughput sequencing to discover protein-DNA interactions.* Methods, 2009. **48**(3): p. 240-8.

337.   Gittens, W.H., D.J. Johnson, R.M. Allison, T.J. Cooper, H. Thomas, and M.J. Neale, *A nucleotide resolution map of Top2-linked DNA breaks in the yeast and human genome.* Nat Commun, 2019. **10**(1): p. 4846.

338.   Gomez-Herreros, F., J.H. Schuurs-Hoeijmakers, M. McCormack, M.T. Greally, S. Rulten, R. Romero-Granados, T.J. Counihan, E. Chaila, J. Conroy, S. Ennis, N. Delanty, F. Cortes-Ledesma, A.P. de Brouwer, G.L. Cavalleri, S.F. El-Khamisy, B.B. de Vries, and K.W. Caldecott, *TDP2 protects transcription from abortive topoisomerase activity and is required for normal neural function.* Nat Genet, 2014. **46**(5): p. 516-21.

339.   Sutormin, D., N. Rubanova, M. Logacheva, D. Ghilarov, and K. Severinov, *Single-nucleotide-resolution mapping of DNA gyrase cleavage sites across the Escherichia coli genome.* Nucleic Acids Res, 2019. **47**(3): p. 1373-1388.

340.   Canela, A., Y. Maman, S.N. Huang, G. Wutz, W. Tang, G. Zagnoli-Vieira, E. Callen, N. Wong, A. Day, J.M. Peters, K.W. Caldecott, Y. Pommier, and A. Nussenzweig, *Topoisomerase II-Induced Chromosome Breakage and Translocation Is Determined by Chromosome Architecture and Transcriptional Activity.* Mol Cell, 2019. **75**(2): p. 252-266 e8.

341.   El Sayyed, H. and O. Espeli, *Mapping E. coli Topoisomerase IV Binding and Activity Sites.* Methods Mol Biol, 2018. **1703**: p. 87-94.

342. Kuempel, P.L., J.M. Henson, L. Dircks, M. Tecklenburg, and D.F. Lim, *dif, a recA-independent recombination site in the terminus region of the chromosome of Escherichia coli.* New Biol, 1991. **3**(8): p. 799-811.

343. Glaziou, P., K. Floyd, and M.C. Raviglione, *Global Epidemiology of Tuberculosis.* Semin Respir Crit Care Med, 2018. **39**(3): p. 271-285.

344. Nagaraja, V., A.A. Godbole, S.R. Henderson, and A. Maxwell, *DNA topoisomerase I and DNA gyrase as targets for TB therapy.* Drug Discov Today, 2017. **22**(3): p. 510-518.

345. Tsao, Y.P., H.Y. Wu, and L.F. Liu, *Transcription-driven supercoiling of DNA: direct biochemical evidence from in vitro studies.* Cell, 1989. **56**(1): p. 111-8.

346. Pruss, G.J. and K. Drlica, *Topoisomerase I mutants: the gene on pBR322 that encodes resistance to tetracycline affects plasmid DNA supercoiling.* Proc Natl Acad Sci U S A, 1986. **83**(23): p. 8952-6.

347. Manjunatha, U.H., M. Dalal, M. Chatterji, D.R. Radha, S.S. Visweswariah, and V. Nagaraja, *Functional characterisation of mycobacterial DNA gyrase: an efficient decatenase.* Nucleic Acids Res, 2002. **30**(10): p. 2144-53.

348. Rani, P. and V. Nagaraja, *Genome-wide mapping of Topoisomerase I activity sites reveal its role in chromosome segregation.* Nucleic Acids Res, 2019. **47**(3): p. 1416-1427.

349. Bhat, A.G., M.N. Leelaram, S.M. Hegde, and V. Nagaraja, *Deciphering the distinct role for the metal coordination motif in the catalytic activity of Mycobacterium smegmatis topoisomerase I.* J Mol Biol, 2009. **393**(4): p. 788-802.

350. Godbole, A.A., W. Ahmed, R.S. Bhat, E.K. Bradley, S. Ekins, and V. Nagaraja, *Targeting Mycobacterium tuberculosis topoisomerase I by small-molecule inhibitors.* Antimicrob Agents Chemother, 2015. **59**(3): p. 1549-57.

351.   Bhaduri, T., T.K. Bagui, D. Sikder, and V. Nagaraja, *DNA topoisomerase I from Mycobacterium smegmatis. An enzyme with distinct features.* J. Biol. Chem., 1998. **273**(22): p. 13925-13932.

352.   Sikder, D. and V. Nagaraja, *Determination of the recognition sequence of Mycobacterium smegmatis topoisomerase I on mycobacterial genomic sequences.* Nucleic Acids Res, 2000. **28**(8): p. 1830-7.

353.   Sikder, D. and V. Nagaraja, *A novel bipartite mode of binding of M. smegmatis topoisomerase I to its recognition sequence.* J Mol Biol, 2001. **312**(2): p. 347-57.

354.   Baranello, L., D. Wojtowicz, K. Cui, B.N. Devaiah, H.J. Chung, K.Y. Chan-Salis, R. Guha, K. Wilson, X. Zhang, H. Zhang, J. Piotrowski, C.J. Thomas, D.S. Singer, B.F. Pugh, Y. Pommier, T.M. Przytycka, F. Kouzine, B.A. Lewis, K. Zhao, and D. Levens, *RNA Polymerase II Regulates Topoisomerase 1 Activity to Favor Efficient Transcription.* Cell, 2016. **165**(2): p. 357-71.

355.   Berger, J.M., *Type II DNA topoisomerases.* Curr. Op. Struct. Biol., 1998. **8**: p. 26-32.

356.   Liu, J., T.C. Wu, and M. Lichten, *The location and structure of double-strand DNA breaks induced during yeast meiosis: evidence for a covalently linked DNA-protein intermediate.* EMBO J, 1995. **14**(18): p. 4599-608.

357.   Morrison, A. and N.R. Cozzarelli, *Site-specific cleavage of DNA by E. coli DNA gyrase.* Cell, 1979. **17**(1): p. 175-84.

358.   Fisher, L.M., H.A. Barot, and M.E. Cullen, *DNA gyrase complex with DNA: determinants for site-specific DNA breakage.* EMBO J., 1986. **5**: p. 1411-1418.

359.   O'Connor, M. and M. Malamy, *Mapping of DNA Gyrase Cleavage Sites in Vivo Oxolinic Acid Induced Cleavages in Plasmid pBR322.* J Mol Biol, 1984. **181**: p. 545-550.

360.   Lockshon, D. and D.R. Morris, *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**(1): p. 63-74.

361. Kreuzer, K.N. and B.M. Alberts, *Site-specific recognition of bacteriophage T4 DNA by T4 Type II DNA topoisomerase and Escherichia coli DNA gyrase.* J. Biol. Chem., 1984. **259**: p. 5339-5346.

362. Sander, M. and T. Hsieh, *Double strand DNA cleavage by type II DNA topoisomerase from Drosophila melanogaster.* J. Biol. Chem., 1983. **258**: p. 8421-8428.

363. Sander, M. and T. Hsieh, *Drosophila topoisomerase II Double-strand DNA cleavage: analysis of DNA Sequence homology at the cleavage site.* Nucleic Acids Research, 1985. **13**: p. 1057-1072.

364. Liu, L.F., T.C. Rowe, L. Yang, K.M. Tewey, and G.L. Chen, *Cleavage of DNA by mammalian DNA topoisomerase II.* Journal of Biological Chemistry, 1983. **258**(24): p. 15365-15370.

365. Pommier, Y., G. Capranico, A. Orr, and K.W. Kohn, *Local base sequence preferences for DNA cleavage by mammalian topoisomerase II in the presence of amsacrine or teniposide.* Nucleic Acids Res., 1991. **19**: p. 5973-5980.

366. Capranico, G., F. Zunino, K.W. Kohn, and Y. Pommier, *Sequence-selective topoisomerase II inhibition by anthracycline derivatives in SV40: relationship with DNA binding affinity and cytotoxicity.* Biochemistry, 1990. **29**: p. 562-569.

367. Condemine, G. and C. Smith, *Transcription regulates oxolinic acid-induced DNA gyrase cleavage at specific sites on the E.coli Chromosome.* Nucleic Acids Research, 1990. **18**: p. 7389-7396.

368. Snyder, M. and K. Drlica, *DNA gyrase on the bacterial chromosome: DNA cleavage induced by oxolinic acid.* J Mol Biol, 1979. **131**(2): p. 287-302.

369. Richter, S.N., G. Giaretta, V. Comuzzi, E. Leo, L.A. Mitchenall, L.M. Fisher, A. Maxwell, and M. Palumbo, *Hot-spot consensus of fluoroquinolone-mediated DNA cleavage by Gram-negative and Gram-positive type II DNA topoisomerases.* Nucleic Acids Res, 2007. **35**(18): p. 6075-85.

370. Leo, E., K.A. Gould, X.S. Pan, G. Capranico, M.R. Sanderson, M. Palumbo, and L.M. Fisher, *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**(14): p. 14252-63.

371. Deininger, P.L., *Random subcloning of sonicated DNA: application to shotgun DNA sequence analysis.* Anal Biochem, 1983. **129**(1): p. 216-23.

372. Il'icheva, I.A., D.Y. Nechipurenko, and S.L. Grokhovsky, *Ultrasonic cleavage of nicked DNA.* J Biomol Struct Dyn, 2009. **27**(3): p. 391-8.

373. Brock, T.D., K.M. Brock, R.T. Belly, and R.L. Weiss, *Sulfolobus: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature.* Arch Mikrobiol, 1972. **84**(1): p. 54-68.

374. Forterre, P., A. Bergerat, and P. Lopez-Garcia, *The unique DNA topology and DNA topoisomerases of hyperthermophilic archaea.* FEMS Microbiol Rev, 1996. **18**(2-3): p. 237-48.

375. Murakami, H. and A. Nicolas, *Locally, meiotic double-strand breaks targeted by Gal4BD-Spo11 occur at discrete sites with a sequence preference.* Mol Cell Biol, 2009. **29**(13): p. 3500-16.

376. Diaz, R.L., A.D. Alcid, J.M. Berger, and S. Keeney, *Identification of residues in yeast Spo11p critical for meiotic DNA double-strand break formation.* Mol Cell Biol, 2002. **22**(4): p. 1106-15.

377. Wahle, E. and A. Kornberg, *The partition locus of plasmid pSC101 is a specific binding site for DNA gyrase.* EMBO, 1988. **7**: p. 1889-1895.

378. Oram, M., A.J. Howells, A. Maxwell, and M.L. Pato, *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**(1): p. 333-47.

379. Forterre, P., C. Brochier, and H. Philippe, *Evolution of the Archaea.* Theor Popul Biol, 2002. **61**(4): p. 409-22.

380. Forterre, P., S. Gribaldo, D. Gadelle, and M.C. Serre, *Origin and evolution of DNA topoisomerases.* Biochimie, 2007. **89**(4): p. 427-46.

381. Lee, C.M., G.S. Wang, A. Pertsinidis, and K.J. Marians, *Topoisomerase III Acts at the Replication Fork To Remove Precatenanes.* Journal of Bacteriology, 2019. **201**(7).

382. Bizard, A., F. Garnier, and M. Nadal, *TopR2, the Second Reverse Gyrase of Sulfolobus solfataricus, Exhibits Unusual Properties.* Journal of Molecular Biology, 2011. **408**(5): p. 839-849.

383. Nadal, M., *Reverse gyrase: an insight into the role of DNA-topoisomerases.* Biochimie, 2007. **89**(4): p. 447-55.

384. Bizard, A.H., X. Yang, H. Debat, J.M. Fogg, L. Zechiedrich, T.R. Strick, F. Garnier, and M. Nadal, *TopA, the Sulfolobus solfataricus topoisomerase III, is a decatenase.* Nucleic Acids Res, 2018. **46**(2): p. 861-872.

385. Gadelle, D., M. Graille, and P. Forterre, *The HSP90 and DNA topoisomerase VI inhibitor radicicol also inhibits human type II DNA topoisomerase.* Biochem Pharmacol, 2006. **72**(10): p. 1207-16.

386. Taylor, J.A., N.P. Burton, and A. Maxwell, *High-Throughput Microtitre Plate-Based Assay for DNA Topoisomerases*, in *Functional Genomics: Methods and Protocols*, M. Kaufmann and C. Klinger, Editors. 2012, Humana Press: New York. p. 229-239.
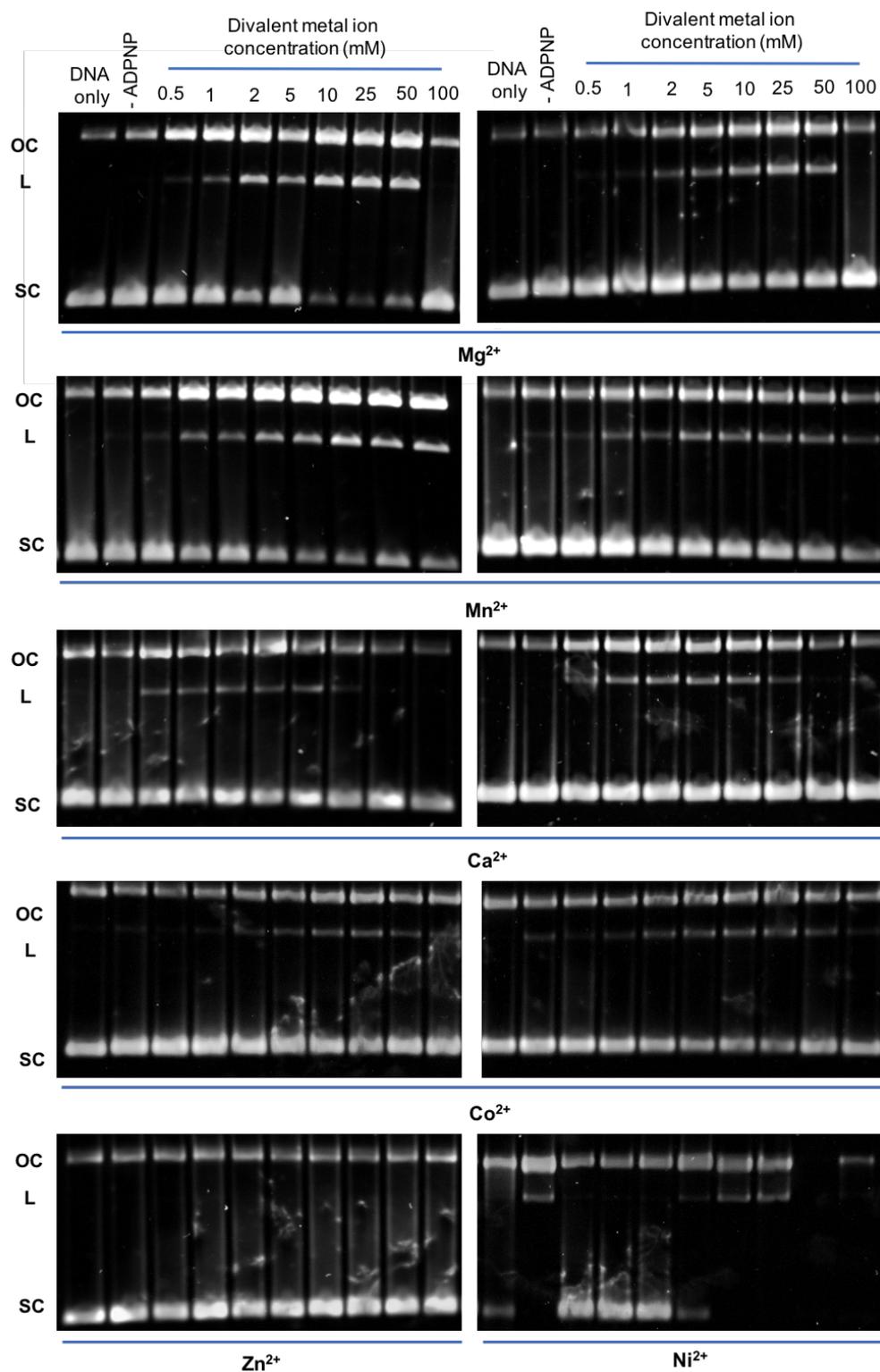
Appendix I

Figure I: Supporting data for Figures 3.17-3.18, exploring the effects of varying divalent metal ion identity and concentration on the DNA cleavage activity of MmT6. MmT6 (160 nM) was incubated

with 1 mM ADPNP, at 37°C for 30 min in the presence of different divalent metal ions. Divalent metal ions explored were magnesium chloride (MgCl$_2$ or Mg^{2+}), manganese chloride (MnCl$_2$ or Mn^{2+}), zinc chloride (ZnCl$_2$ or Zn^{2+}), calcium chloride (CaCl$_2$ or Ca^{2+}), cobalt chloride (CoCl$_2$ or Co^{2+}), and nickel chloride (NiCl$_2$ or Ni^{2+}), with each varied in concentration from 0.5 to 100 mM. The DNA-only sample was incubated in the absence of MmT6 and in the presence of 100 mM of the respective metal cation salt. All samples were run on a 1% (w/v) native agarose gel for 15 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination. OC: open circular or nicked, L: linear, and SC: supercoiled.
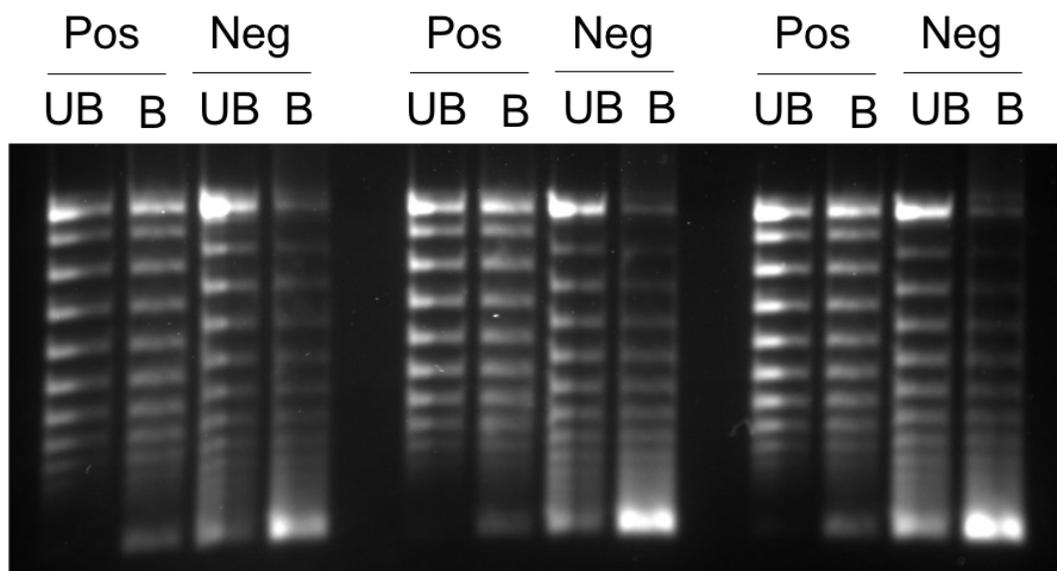

Figure II: Supporting data for Figure 3.26, MmT6 DNA binding affinity as a function of the number of supercoils in pBR322*. MmT6 (8nM) was incubated with 5 nM pBR322*, either positively (pos) and negatively (neg) supercoiled, and 10 mM MgCl$_2$, for 30 min at 37°C. The MmT6-bound DNA (B) was separated from the unbound DNA (UB) using 0.45µm nitrocellulose Centrex MF filters (Whatman). The bound and unbound DNA samples were run on a 1% (w/v) agarose gel for 13 hr at ~2 Vcm^{-1}, stained with 0.5 µg/mL ethidium bromide and imaged under UV illumination.
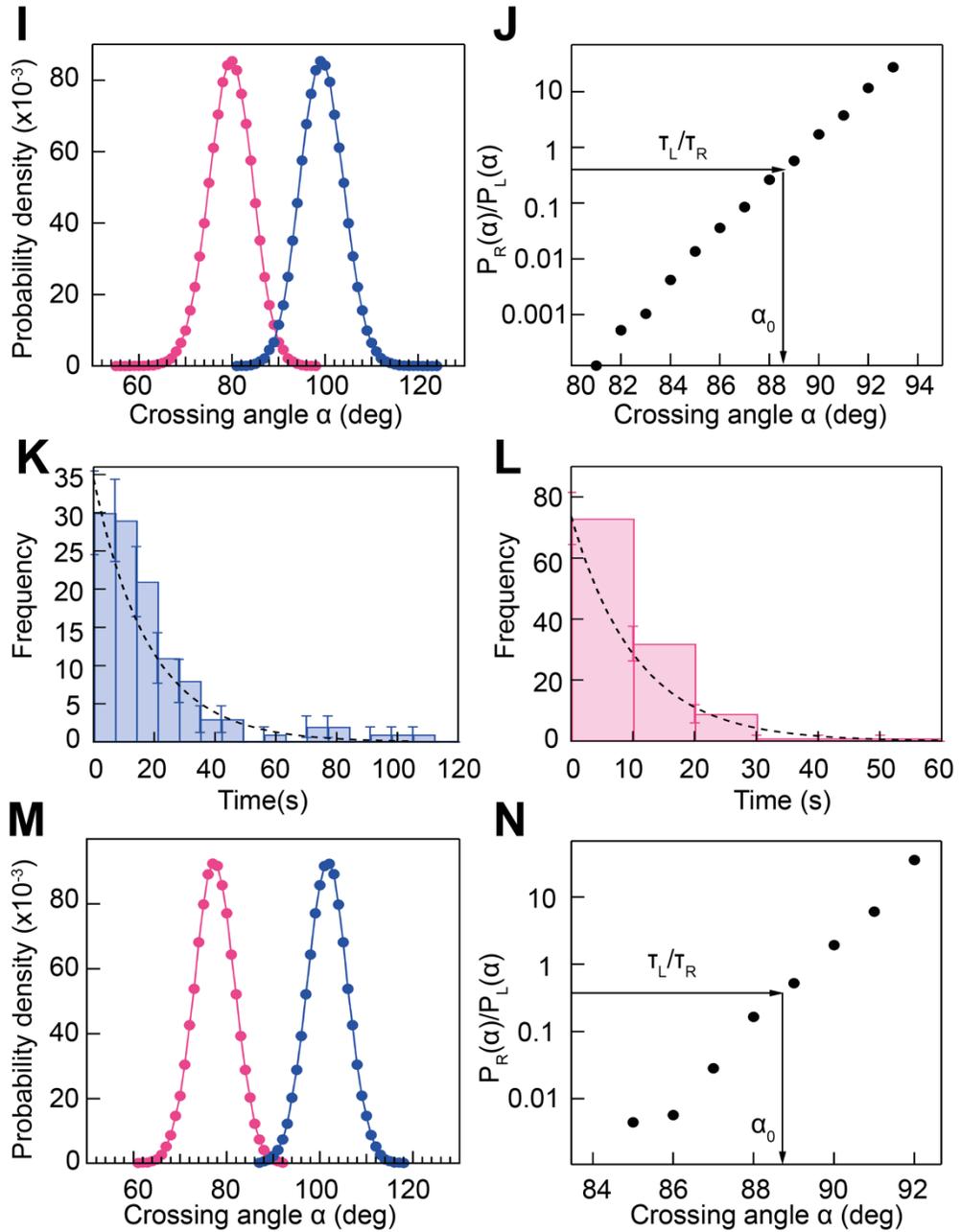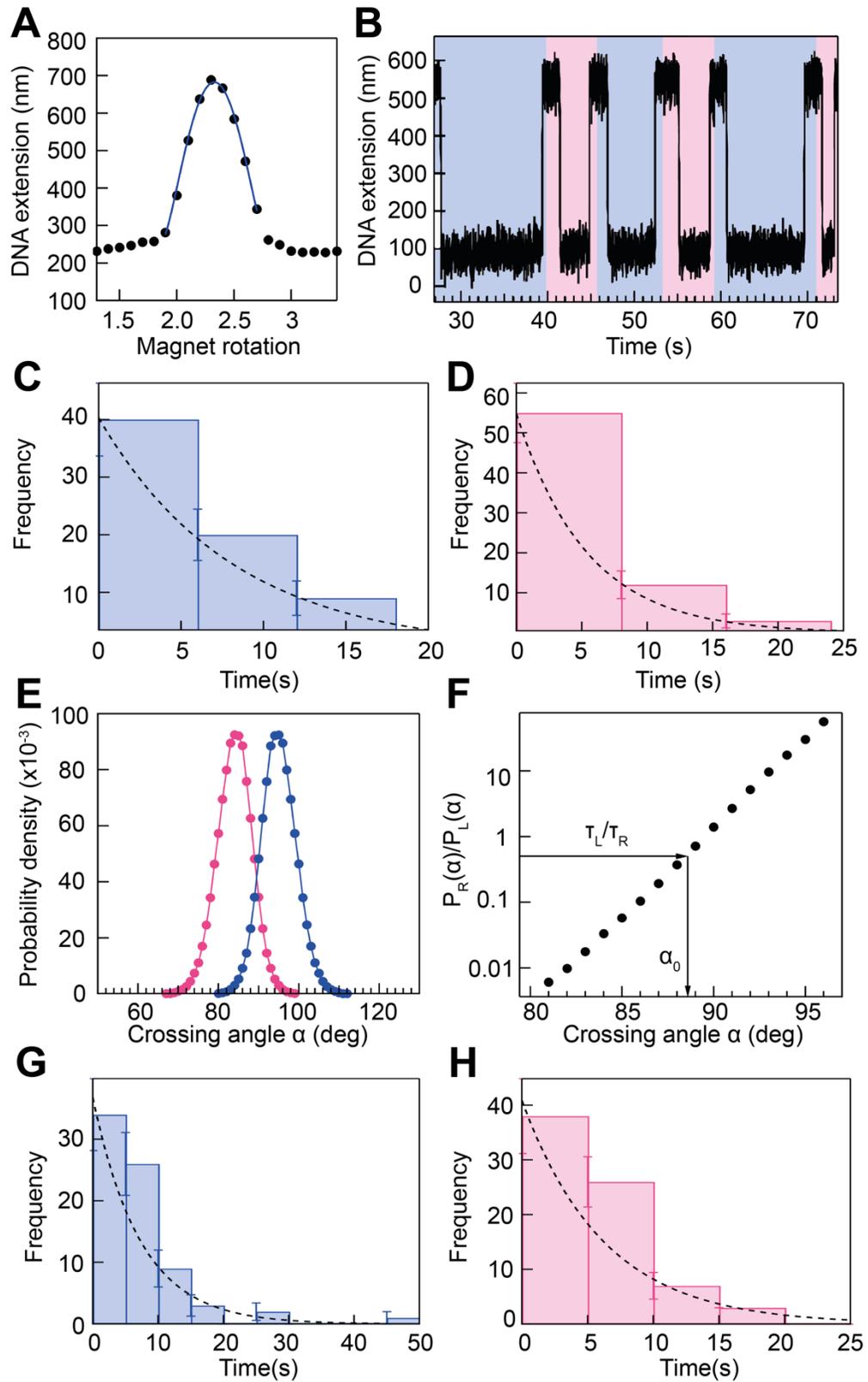
Appendix II
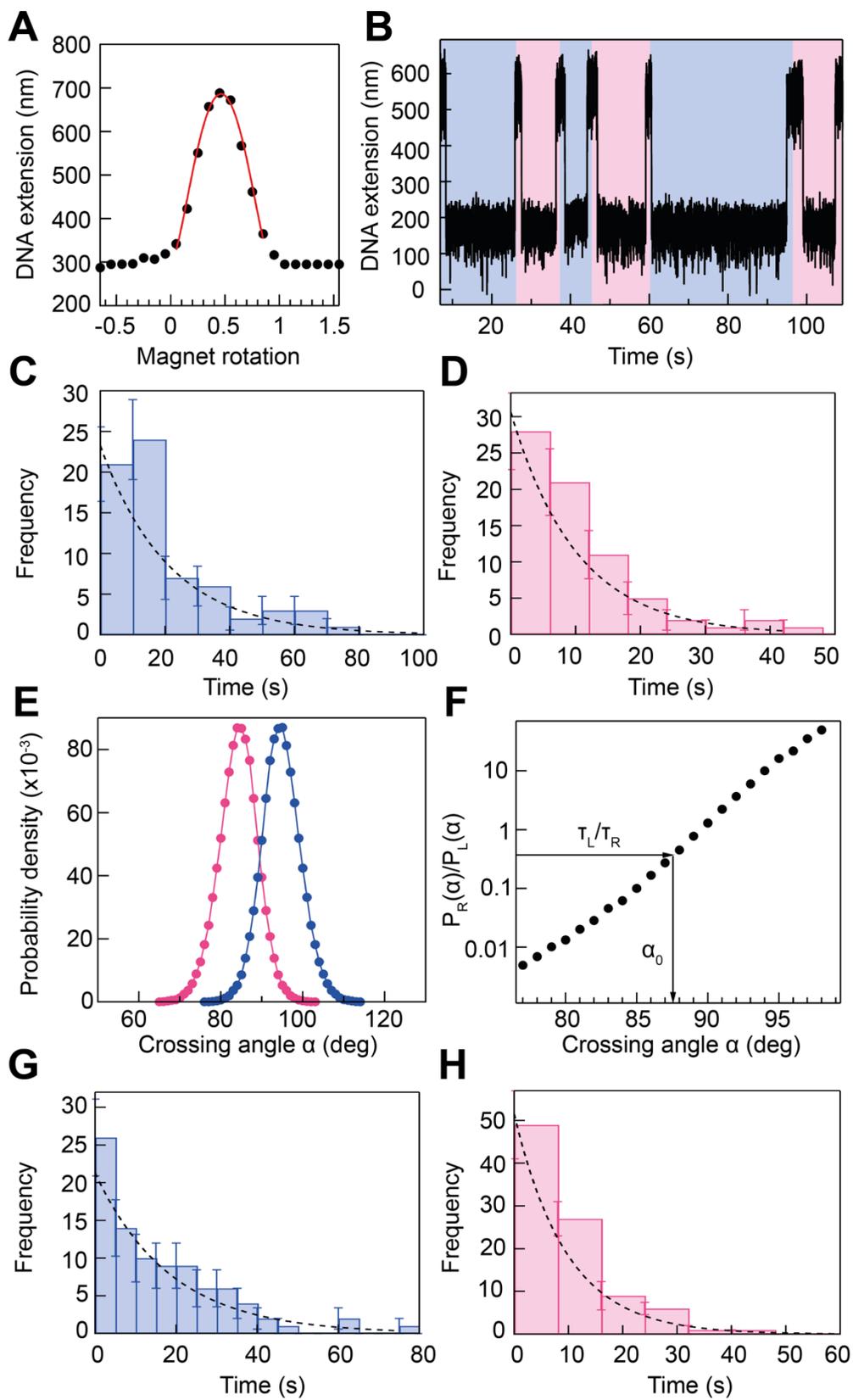
Figure III: Single crossing data analysis for braid 2 (see table I, appendix II). A: Calibration curve of DNA extension as a function of magnet rotations (black circles), fit the geometric function, $L = \sqrt{L_0^2 - 4e^2 sin^2(n\pi)} - r + \sqrt{r^2 - e^2}, |n| < 0.5$ (red line), calculating the DNA spacing on the bead to 660 nm. B: Example trace of positive (pink) and negative (blue) single crossings being input mechanically (DNA extension decrease), and removed by 900 nM MmT6 (DNA extension increase), under 0.5 pN, at 21°C. C: Unlinking times for negative crossings under 0.5 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_R) 21.06 ± 3.5 s. D: Unlinking times for positive crossings under 0.5 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_L) 10.2 ± 1.4 s. E: Crossing angle probability distributions for single positive ($P_L(\alpha)$, pink) and negative ($P_R(\alpha)$, blue) crossings, from

Monte Carlo (MC) simulations [223] for the tether geometry seen in A, under 0.5 pN. The positive angle probability distributions were attained directly, with the negative distributions derived from the relationship $P_R(\alpha) = P_L(180°-\alpha)$. F: Ratio of negative to positive probability from Figure 1E plotted on a semilogarithmic axis. Using the relationship $\tau_L/\tau_R = P_R(\alpha_0)/P_L(\alpha_0)$, the preferred angle ($\alpha_0$) can be attained, illustrated by the black arrow. For this tether geometry and force (0.5 pN), $\tau_L/\tau_R = 0.48 \pm 0.09$, which gave α_0 of 88.11° ± 4.4° when using MC simulations, and 85° ± 4.3° when using BD simulations (BD plots not shown). G: Unlinking times for negative crossings under 1 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_R) 18.49 ± 2.8 s. H: Unlinking times for positive crossings under 1 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_L) 9.6 ± 1.1 s. I: Crossing angle probability distributions for single positive ($P_L(\alpha)$, pink) and negative ($P_R(\alpha)$, blue) crossings, from Monte Carlo (MC) simulations [223] for the tether geometry seen in A, under 1 pN. The positive angle probability distributions were attained directly, with the negative distributions derived from the relationship $P_R(\alpha) = P_L(180°-\alpha)$. J: Ratio of negative to positive probability from Figure 1I plotted on a semilogarithmic axis. Using the relationship $\tau_L/\tau_R = P_R(\alpha_0)/P_L(\alpha_0)$, the preferred angle ($\alpha_0$) can be attained, illustrated by the black arrow. For this tether geometry and force (1 pN), $\tau_L/\tau_R = 0.52 \pm 0.08$, which gave α_0 of 88.82° ± 4.4° when using MC simulations, and 87.75° ± 4.4° when using BD simulations (BD plots not shown). K: Unlinking times for negative crossings under 1.5 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_R) 18.32 ± 2.2 s. L: Unlinking times for positive crossings under 1.5 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_L) 10.6 ± 0.6 s. M: Crossing angle probability distributions for single positive ($P_L(\alpha)$, pink) and negative ($P_R(\alpha)$, blue) crossings, from Monte Carlo (MC) simulations [223] for the tether geometry seen in A, under 1.5 pN. The positive angle probability distributions were attained directly, with the negative distributions derived from the relationship $P_R(\alpha) = P_L(180°-\alpha)$. N: Ratio of negative to positive probability from Figure 1M plotted on a semilogarithmic axis. Using the relationship $\tau_L/\tau_R = P_R(\alpha_0)/P_L(\alpha_0)$, the preferred angle ($\alpha_0$) can be attained, illustrated by the black arrow. For this tether geometry and force (1.5 pN), $\tau_L/\tau_R = 0.58 \pm 0.08$, which gave α_0 of 89.04° ± 4.5° when using MC simulations, and 88.5° ± 4.4° when using BD simulations (BD plots not shown). The error associated with α_0 is the combination of the statistical and systematic error.
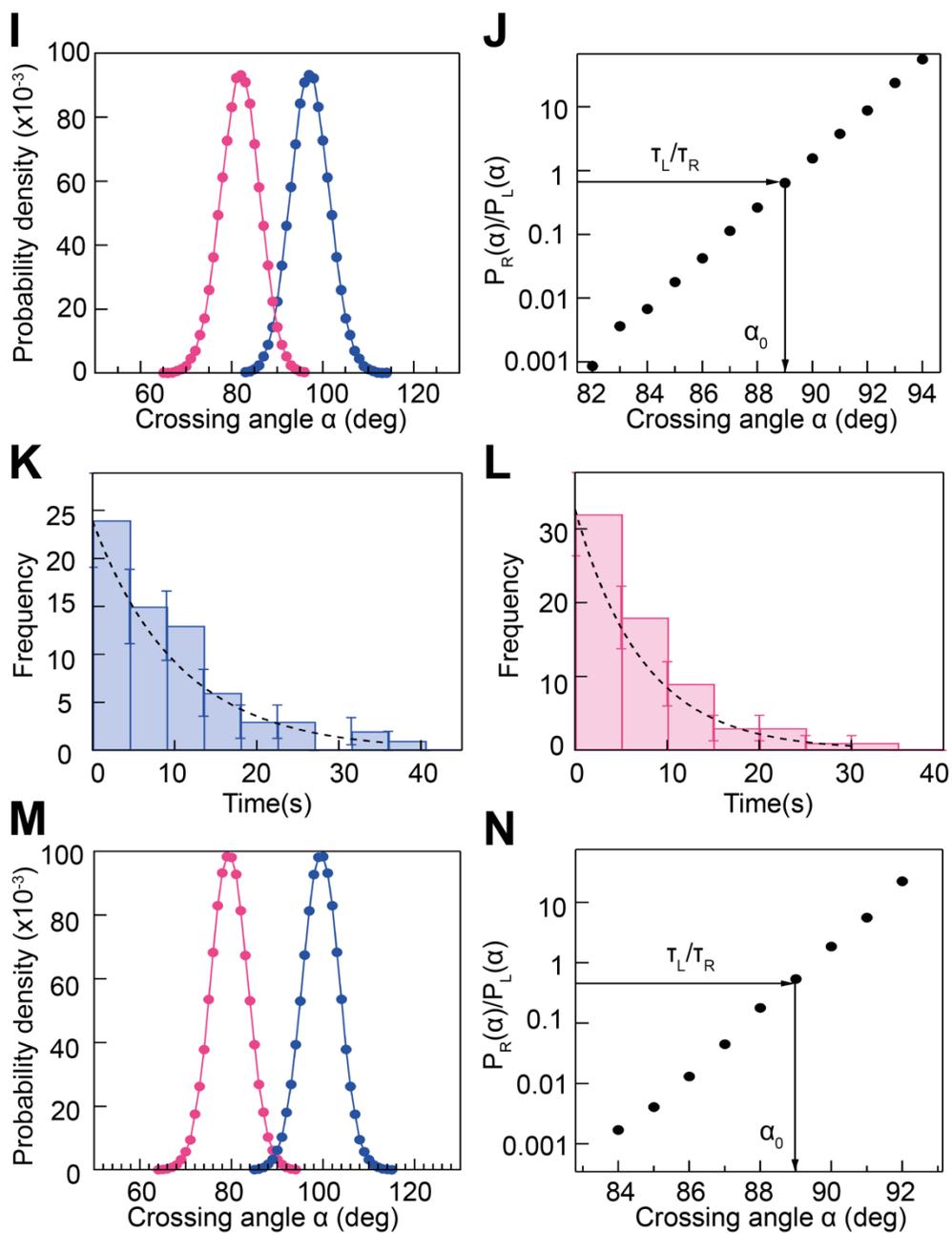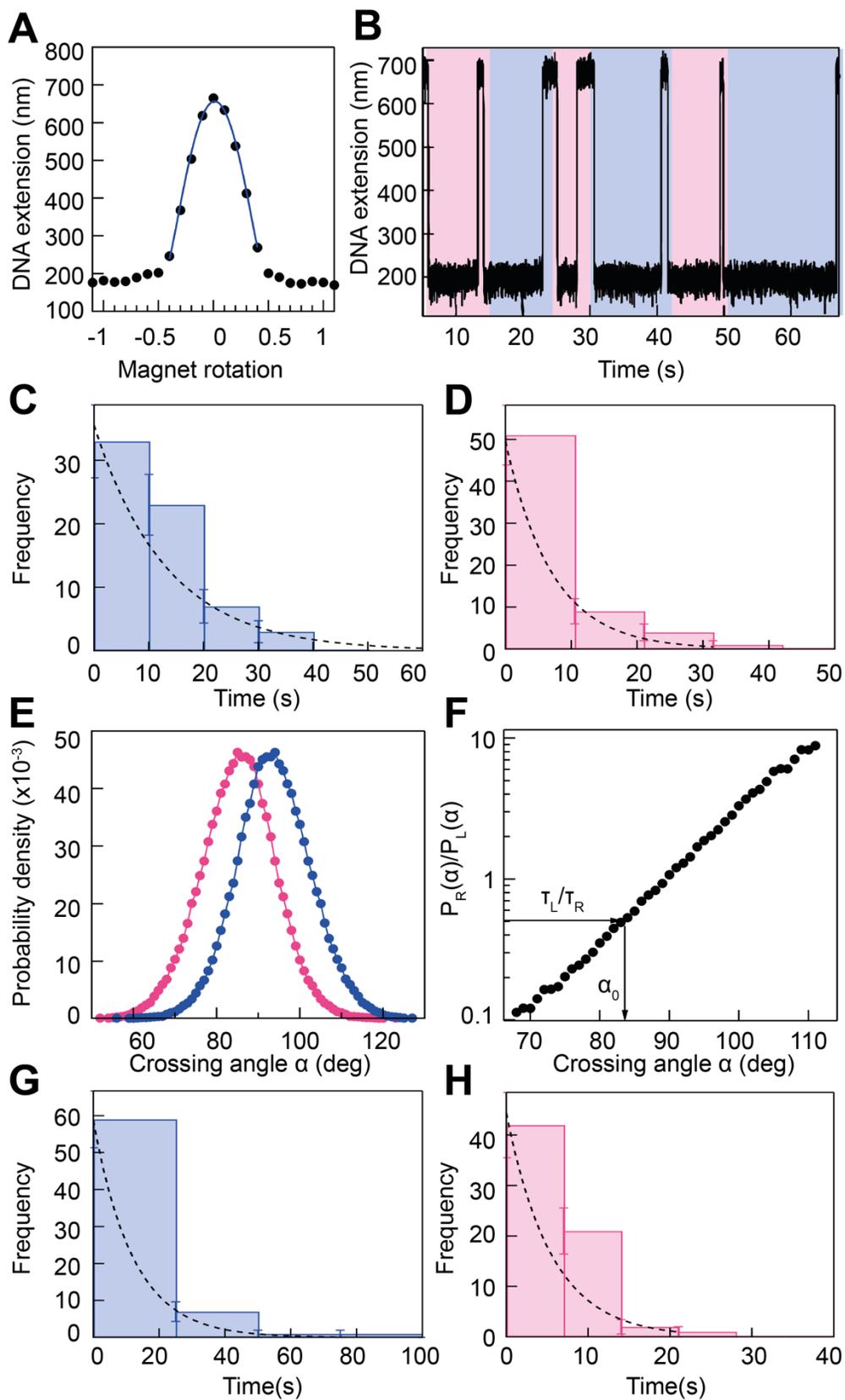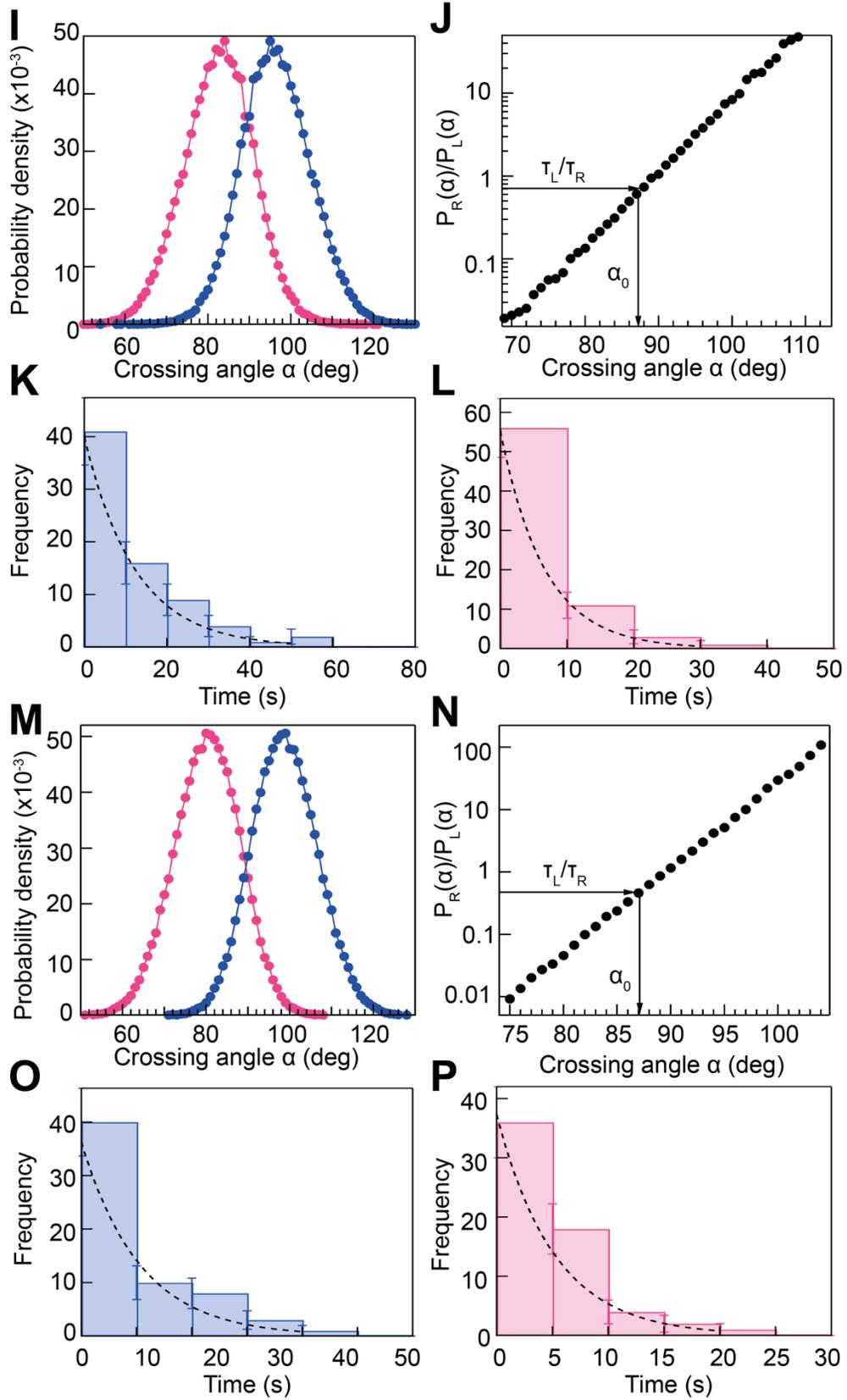
Figure IV: Single crossing data analysis for braid 3 (see table I, appendix II). A: Calibration curve of DNA extension as a function of magnet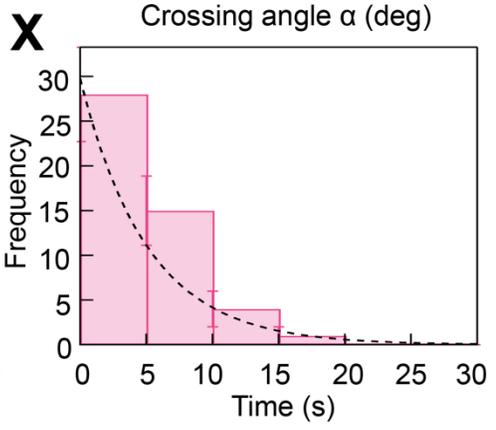 rotations (black circles), fit the geometric function, $L = \sqrt{L_0^2 - 4e^2 sin^2(n\pi)} - r + \sqrt{r^2 - e^2}, |n| < 0.5$ (red line), calculating the DNA spacing on the bead to 660 nm. B: Example trace of positive (pink) and negative (blue) single crossings being input mechanically (DNA extension decrease), and removed by 900 nM MmT6 (DNA extension increase), under 0.7 pN, at 21°C. C: Unlinking times for negative crossings under 0.7 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_R) 8.18 ± 1.9 s. D: Unlinking times for positive crossings under 0.7 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_L) 5.4 ± 0.8 s. E: Crossing angle probability distributions for single positive ($P_L(\alpha)$, pink) and negative ($P_R(\alpha)$, blue) crossings, from

294

Monte Carlo (MC) simulations [223] for the tether geometry seen in A, under 0.7 pN. The positive angle probability distributions were attained directly, with the negative distributions derived from the relationship $P_R(\alpha) = P_L(180°-\alpha)$. F: Ratio of negative to positive probability from Figure 1E plotted on a semilogarithmic axis. Using the relationship $\tau_L/\tau_R = P_R(\alpha_0)/ P_L(\alpha_0)$, the preferred angle ($\alpha_0$) can be attained, illustrated by the black arrow. For this tether geometry and force (0.7 pN), $\tau_L/\tau_R = 0.65 \pm 0.17$, which gave α_0 of 88.81° ± 4.4° when using MC simulations, and 85.75° ± 4.3° when using BD simulations (BD plots not shown). G: Unlinking times for negative crossings under 1 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_R) 7.2 ± 1.1 s. H: Unlinking times for positive crossings under 1 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_L) 6.6 ± 1.3 s. I: Crossing angle probability distributions for single positive ($P_L(\alpha)$, pink) and negative ($P_R(\alpha)$, blue) crossings, from Monte Carlo (MC) simulations [223] for the tether geometry seen in A, under 1 pN. The positive angle probability distributions were attained directly, with the negative distributions derived from the relationship $P_R(\alpha) = P_L(180°-\alpha)$. J: Ratio of negative to positive probability from Figure 1I plotted on a semilogarithmic axis. Using the relationship $\tau_L/\tau_R = P_R(\alpha_0)/ P_L(\alpha_0)$, the preferred angle ($\alpha_0$) can be attained, illustrated by the black arrow. For this tether geometry and force (1 pN), $\tau_L/\tau_R = 0.9 \pm 0.19$, which gave α_0 of 89.28° ± 4.5° when using MC simulations, and 86.25° ± 4.3° when using BD simulations (BD plots not shown). K: Unlinking times for negative crossings under 1.5 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_R) 10.5 ± 1.7 s. L: Unlinking times for positive crossings under 1.5 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_L) 7.4 ± 1.1 s. M: Crossing angle probability distributions for single positive ($P_L(\alpha)$, pink) and negative ($P_R(\alpha)$, blue) crossings, from Monte Carlo (MC) simulations [223] for the tether geometry seen in A, under 1.5 pN. The positive angle probability distributions were attained directly, with the negative distributions derived from the relationship $P_R(\alpha) = P_L(180°-\alpha)$. N: Ratio of negative to positive probability from Figure 1M plotted on a semilogarithmic axis. Using the relationship $\tau_L/\tau_R = P_R(\alpha_0)/ P_L(\alpha_0)$, the preferred angle ($\alpha_0$) can be attained, illustrated by the black arrow. For this tether geometry and force (1.5 pN), $\tau_L/\tau_R = 0.7 \pm 0.15$, which gave α_0 of 89.12° ± 4.5° when using MC simulations, and 87.55° ± 4.4° when using BD simulations (BD plots not shown). The error associated with α_0 is the combination of the statistical and systematic error.

Figure V: Single crossing data analysis for braid 4 (see table I, appendix II). A: Calibration curve of DNA extension as a function of magnet rotations (bla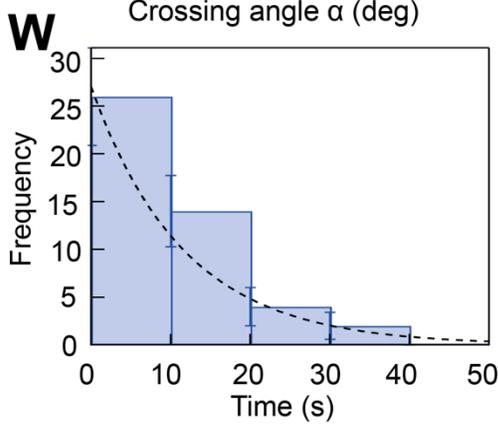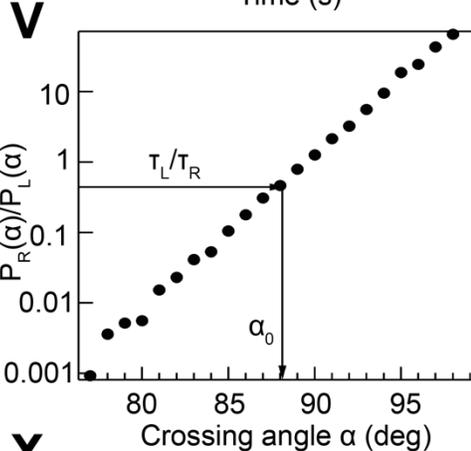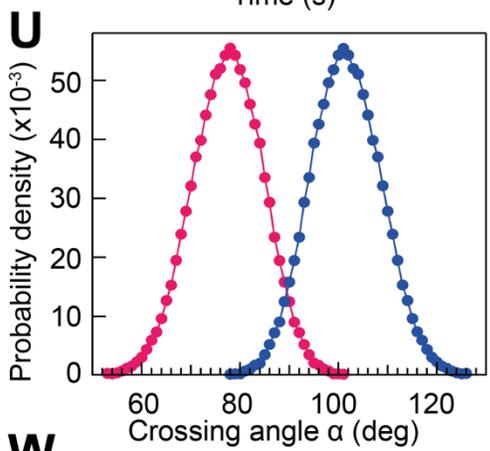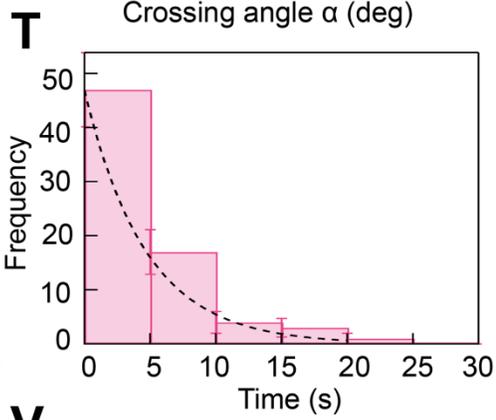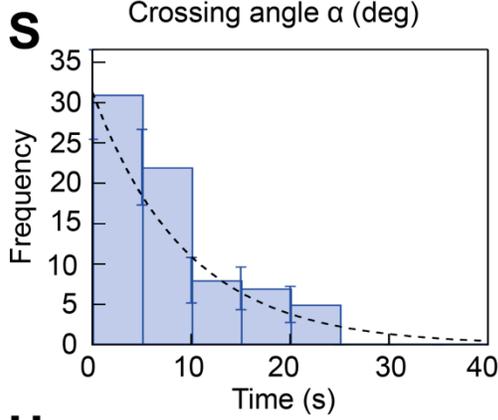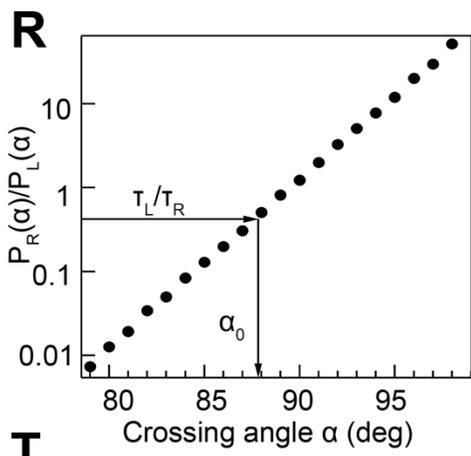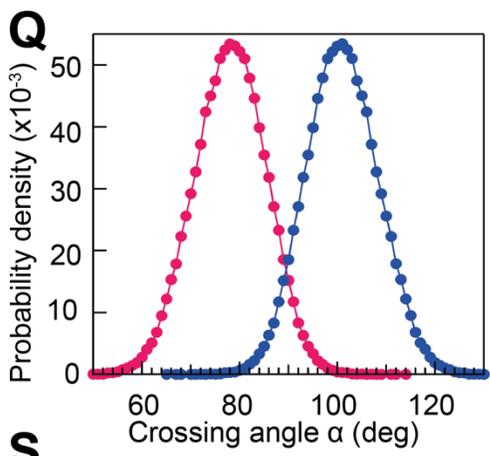ck circles), fit the geometric function, $L = \sqrt{L_0^2 - 4e^2 sin^2(n\pi)} - r + \sqrt{r^2 - e^2}, |n| < 0.5$ **(red line), calculating the DNA spacing on the bead to 660 nm. B: Example trace of positive (pink) and negative (blue) single crossings being input mechanically (DNA extension decrease), and removed by 900 nM MmT6 (DNA extension increase), under 0.5 pN, at 21°C. C: Unlinking times for negative crossings under 0.5 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_R) 11.42 ± 1.64 s. D: Unlinking times for positive crossings under 0.5 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_L) 6.6 ± 1.2 s. E: Crossing angle probability distributions for single positive ($P_L(\alpha)$, pink) and negative ($P_R(\alpha)$, blue) crossings, from**

Monte Carlo (MC) simulations [223] for the tether geometry seen in A, under 0.5 pN. The positive angle probability distributions were attained directly, with the negative distributions derived from the relationship $P_R(\alpha) = P_L(180°-\alpha)$. F: Ratio of negative to positive probability from Figure 1E plotted on a semilogarithmic axis. Using the relationship $\tau_L/\tau_R = P_R(\alpha_0)/ P_L(\alpha_0)$, the preferred angle ($\alpha_0$) can be attained, illustrated by the black arrow. For this tether geometry and force (0.5 pN), $\tau_L/\tau_R = 0.57 \pm 0.12$, which gave α_0 of 84.04° ± 4.2° when using MC simulations, and 88.25° ± 4.4° when using BD simulations (BD plots not shown). G: Unlinking times for negative crossings under 0.7 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_R) 9.2 ± 1.5 s. H: Unlinking times for positive crossings under 0.7 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_L) 6.1 ± 0.9 s. I: Crossing angle probability distributions for single positive ($P_L(\alpha)$, pink) and negative ($P_R(\alpha)$, blue) crossings, from Monte Carlo (MC) simulations [223] for the tether geometry seen in A, under 0.7 pN. The positive angle probability distributions were attained directly, with the negative distributions derived from the relationship $P_R(\alpha) = P_L(180°-\alpha)$. J: Ratio of negative to positive probability from Figure 1I plotted on a semilogarithmic axis. Using the relationship $\tau_L/\tau_R = P_R(\alpha_0)/ P_L(\alpha_0)$, the preferred angle ($\alpha_0$) can be attained, illustrated by the black arrow. For this tether geometry and force (0.7 pN), $\tau_L/\tau_R = 0.7 \pm 0.13$, which gave α_0 of 85.57° ± 4.3° when using MC simulations, and 88.25° ± 4.4° when using BD simulations (BD plots not shown). K: Unlinking times for negative crossings under 1 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_R) 12.3 ± 1.7 s. L: Unlinking times for positive crossings under 1 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_L) 6.6 ± 1 s. M: Crossing angle probability distributions for single positive ($P_L(\alpha)$, pink) and negative ($P_R(\alpha)$, blue) crossings, from Monte Carlo (MC) simulations [223] for the tether geometry seen in A, under 1 pN. The positive angle probability distributions were attained directly, with the negative distributions derived from the relationship $P_R(\alpha) = P_L(180°-\alpha)$. N: Ratio of negative to positive probability from Figure 1M plotted on a semilogarithmic axis. Using the relationship $\tau_L/\tau_R = P_R(\alpha_0)/ P_L(\alpha_0)$, the preferred angle ($\alpha_0$) can be attained, illustrated by the black arrow. For this tether geometry and force (1 pN), $\tau_L/\tau_R = 0.54 \pm 0.1$, which gave α_0 of 87.45° ± 4.4° when using MC simulations, and 88.25° ± 4.4° when using BD simulations (BD plots not shown). The error associated with α_0 is the comb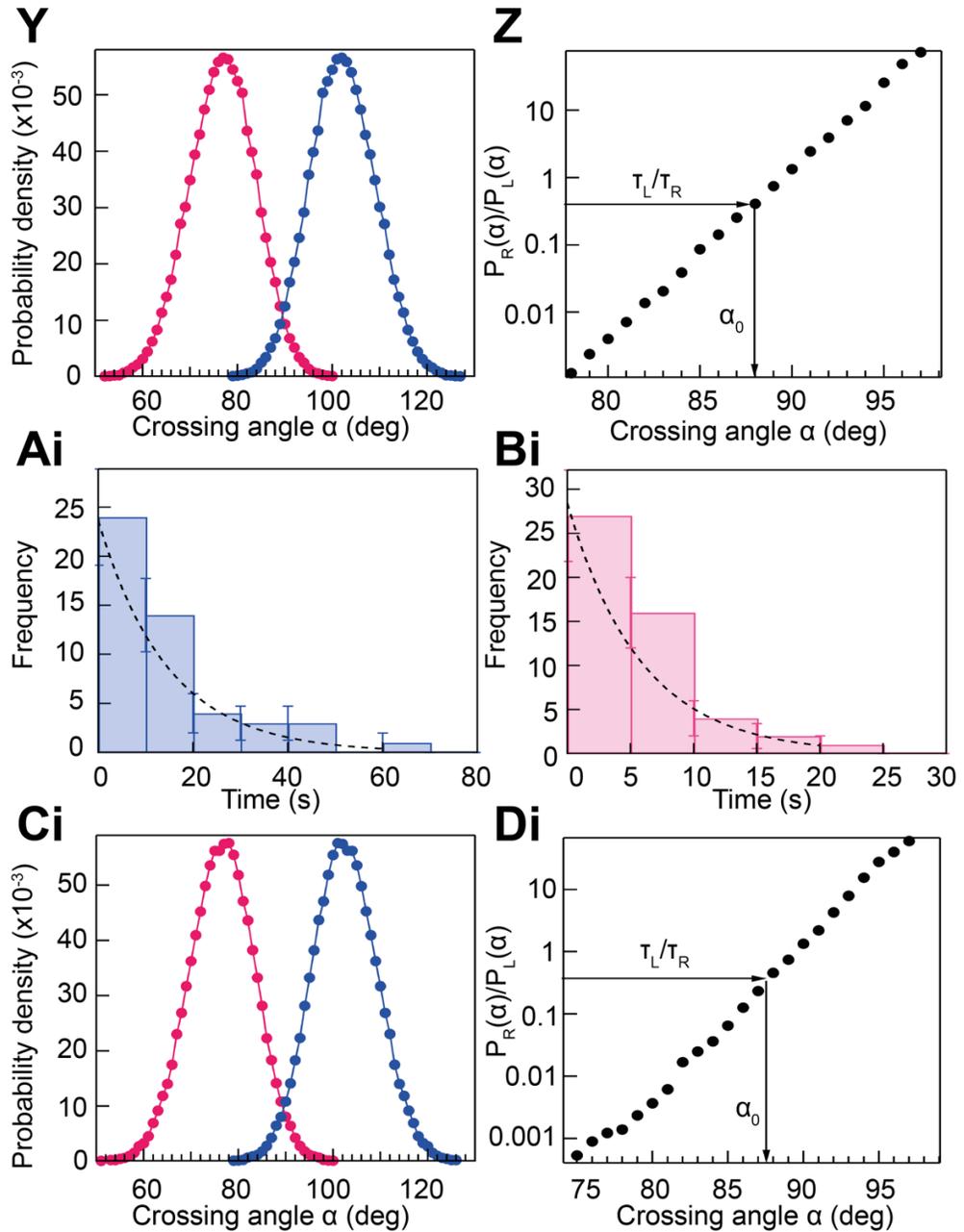ination of the statistical and systematic error. O: Unlinking times for negative crossings under 1.5 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_R) 10.6 ± 1.8 s. P: Unlinking times for positive crossings under 1.5 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_L) 5.2 ± 0.8 s. Q: Crossing angle probability distributions for single positive ($P_L(\alpha)$, pink) and negative ($P_R(\alpha)$, blue) crossings, from Monte Carlo (MC) simulations [223] for the tether geometry seen in A, under 1.5 pN. The positive angle probability distributions were attained directly, with the negative distributions derived from the relationship $P_R(\alpha) = P_L(180°-\alpha)$. R: Ratio of negative to positive probability from Figure 1Q plotted on a semilogarithmic axis. Using the relationship $\tau_L/\tau_R = P_R(\alpha_0)/ P_L(\alpha_0)$, the preferred angle ($\alpha_0$) can be attained, illustrated by the black arrow. For this tether geometry and force (1.5 pN),

$\tau_L/\tau_R = 0.49 \pm 0.11$, which gave α_0 of 87.88° ± 4.4° when using MC simulations, and 87.75° ± 4.4° when using BD simulations (BD plots not shown). S: Unlinking times for negative crossings under 1.7 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_R) 9.5 ± 1.8 s. T: Unlinking times for positive crossings under 1.7 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_L) 4.7 ± 0.7 s. U: Crossing angle probability distributions for single positive ($P_L(\alpha)$, pink) and negative ($P_R(\alpha)$, blue) crossings, from Monte Carlo (MC) simulations [223] for the tether geometry seen in A, under 1.7 pN. The positive angle probability distributions were attained directly, with the negative distributions derived from the relationship $P_R(\alpha) = P_L(180°\text{-}\alpha)$. V: Ratio of negative to positive probability from Figure 1I plotted on a semilogarithmic axis. Using the relationship $\tau_L/\tau_R = P_R(\alpha_0)/ P_L(\alpha_0)$, the preferred angle ($\alpha_0$) can be attained, illustrated by the black arrow. For this tether geometry and force (1.7 pN), $\tau_L/\tau_R = 0.49 \pm 0.12$, which gave α_0 of 88.07° ± 4.4° when using MC simulations, and 88.25° ± 4.4° when using BD simulations (BD plots not shown). W: Unlinking times for negative crossings under 2 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_R) 11.6 ± 2.3 s. X: Unlinking times for positive crossings under 2 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_L) 5.08 ± 0.9 s. Y: Crossing angle probability distributions for single positive ($P_L(\alpha)$, pink) and negative ($P_R(\alpha)$, blue) crossings, from Monte Carlo (MC) simulations [223] for the tether geometry seen in A, under 2 pN. The positive angle probability distributions were attained directly, with the negative distributions derived from the relationship $P_R(\alpha) = P_L(180°\text{-}\alpha)$. Z: Ratio of negative to positive probability from Figure 1Y plotted on a semilogarithmic axis. Using the relationship $\tau_L/\tau_R = P_R(\alpha_0)/ P_L(\alpha_0)$, the preferred angle ($\alpha_0$) can be attained, illustrated by the black arrow. For this tether geometry and force (2 pN), $\tau_L/\tau_R = 0.44 \pm 0.11$, which gave α_0 of 88.07° ± 4.5° when using MC simulations, and 89.75° ± 4.5° when using BD simulations (BD plots not shown). Ai: Unlinking times for negative crossings under 2.2 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_R) 14.6 ± 2.3 s. Bi: Unlinking times for positive crossings under 2.2 pN, plot as a histogram and fit to a single exponential resulting in a characteristic unlinking time (τ_L) 5.8 ± 1 s. Ci: Crossing angle probability distributions for single positive ($P_L(\alpha)$, pink) and negative ($P_R(\alpha)$, blue) crossings, from Monte Carlo (MC) simulations [223] for the tether geometry seen in A, under 2.2 pN. The positive angle probability distributions were attained directly, with the negative distributions derived from the relationship $P_R(\alpha) = P_L(180°\text{-}\alpha)$. Di: Ratio of negative to positive probability from Figure 1M plotted on a semilogarithmic axis. Using the relationship $\tau_L/\tau_R = P_R(\alpha_0)/ P_L(\alpha_0)$, the preferred angle ($\alpha_0$) can be attained, illustrated by the black arrow. For this tether geometry and force (2.2 pN), $\tau_L/\tau_R = 0.4 \pm 0.1$, which gave α_0 of 87.74° ± 4.3° when using MC simulations, and 88.75° ± 4.4° when using BD simulations (BD plots not shown). The error associated with α_0 is the combination of the statistical and systematic error.

Table I: Single crossing data used to determine the MmT6 preferred DNA crossing angle.

Braid No.	Force (pN)	DNA spacing (nm)	Unlinking Time on Positive (τ_L)	Unlinking Time on Negative (τ_R)	Unlinking Time Ratio (τ_L/τ_R)	Preferred Angle (MC)	Preferred Angle (BD)
1	1	624	8.43	43.85	0.19	87.87	83.5
2	0.5	660	10.17	21.06	0.48	88.11	85
	1		9.64	18.49	0.52	88.82	87.75
	1.5		10.59	18.32	0.58	89.04	88.5
3	0.7	679	5.38	8.18	0.66	88.81	85.75
	1		6.56	7.27	0.90	89.28	86.25
	1.5		7.39	10.50	0.70	89.12	87.75
4	0.5	677	6.57	11.42	0.57	84.04	88.25
	0.7		6.05	9.19	0.66	85.57	88.25
	1		6.61	12.34	0.54	87.45	88.25
	1.5		5.15	10.61	0.49	87.88	87.75
	1.7		4.65	9.51	0.49	88.07	88.25
	2		5.08	11.61	0.44	88.07	89.75
	2.2		5.82	14.63	0.40	87.74	88.75

Appendix III

Table II: Sequence information of MmT6 cleavage sites on pBR322*, detected using the Illumina NGS platform. The sequences have been characterised as strong and weak.

No.	Position (bp)	Top Strand sequence	Position (bp)	Bottom Strand sequence
		Strong topo VI cleavage sites		
1	132-167	TGGATGCTGTAGGCATAGGCT TGGTTATGCCGGTA	133-98	CAGGGTGACGGTGCCGAGGATG ACGATGAGCGCAT
2	918-953	CAGGCCATTATCGCCGGCATG GCGGCCGACGCGCT	919-884	CTGCTTCTCGCCGAAACGTTTG GTGGCGGGACCAG
3	995-1030	TTCCCCATTATGATTCTTCTC GCTTCCGGCGGCAT	996-961	AAGGCCATCCAGCCTCGCGTCG CGAACGCCAGCAA
4	1503-1538	TTGCCTTACTGGTTAGCAGAA TGAATCACCGATAC	1504-1469	AACCCCGCCAGCCTAGCCGGGT CCTCAACGACAGG
5	1683-1718	ATGCTGCTGGCTACCCTGTGG AACACCTACATCTG	1684-1649	ATCCTGCGATGCAGATCCGGAA CATAATGGTGCAG
6	1907-1942	AGGCATCAGTGACCAAACAGG AAAAAACCGCCCTT	1908-1873	CTCCGTGTAAGGGGGATTTCTG TTCATGGGGGTAA
7	2019-2054	AGACATCTGTGAATCGCTTCA CGACCACGCTGATG	2020-1985	CTGCCTGTTCATCCGCGTCCAG CTCGTTGAGTTTC
8	2292-2327	CACCATATGCGGTGTGAAATA CCGCACAGATGCGT	2293-2258	TGCACTCTCAGTACAATCTGCT CTGATGCCGCATA
9	2656-2691	CTCCTGTTCCGACCCTGCCGC TTACCGGATACCTG	2657-2622	AGAGCGCACGAGGGAGCTTCCA GGGGGAAACGCCT
10	3394-3429	AGGGCTTACCATCTGGCCCCA GTGCTGCAATGATA	3395-3360	CTCCCGTATCGTAGTTATCTAC ACGACGGGGAGTC
11	3478-3513	CAGCCGGAAGGGCCGAGCGCA GAAGTGGTCCTGCA	3479-3444	TGGCTGGTTTATTGCTGATAAA TCTGGAGCCGGTG
12	3611-3646	AGGCATCGTGGTGTCACGCTC GTCGTTTGGTATGG	3612-3577	CTGCAGCAATGGCAACAACGTT GCGCAAACTATTA
13	3814-3849	CATCCGTAAGATGCTTTTCTG TGACTGGTGAGTAC	3815-3780	TGGCATGACAGTAAGAGAATTA TGCAGTGCTGCCA
14	4097-4132	AAGGCAAAATGCCGCAAAAAA GGGAATAAGGGCGA	4098-4063	TTCCTGTTTTTGCTCACCCAGA AACGCTGGTGAAA
15	4341-4376	AGGCCCTTTCGTCTTCAAGAA TTCTCATGTTTGAC	4342-4307	CTCGTGATACGCCTATTTTTAT AGGTTAATGTCAT
		Weak topo VI cleavage sites		
1	403-438	CGGCATCACCGGCGCCACAGG TGCGGTTGCTGGCG	404-369	CGGCCACGATGCGTCCGGCGTA GAGGATCCACAGG
2	523-558	AGGCCCCGTGGCCGGGGGACT GTTGGGCGCCATCT	524-489	CTGCCACCATACCCACGCCGAA ACAAGCGCTCATG
3	866-901	CTCGCTCAAGCCTTCGTCACT GGTCCCGCCACCAA	867-832	AGGGCGTGCAAGATTCCGAATA CCGCAAGCGACAG
4	1035-1070	TGCCCGCGTTGCAGGCCATGC TGTCCAGGCAGGTA	1036-1001	CATCCCGATGCCGCCGGAAGCG AGAAGAATCATAA
5	1704-1739	AACACCTACATCTGTATTAAC GAAGCGCTGGCATT	1705-1670	TTCCACAGGGTAGCCAGCAGCA TCCTGCGATGCAG
6	2465-2500	AGGAAAGAACATGTGAGCAAA AGGCCAGCAAAAGG	2466-2431	CTGCGTTATCCCCTGATTCTGT GGATAACCGTATT
7	2742-2777	AGGTATCTCAGTTCGGTGTAG GTCGTTCGCTCCAA	2743-2708	CTACAGCGTGAGCTATGAGAAA GCGCCACGCTTCC
8	2949-2984	TGGCCTAACTACGGCTACACT AGAAGGACAGTATT	2950-2915	CACCACTTCAAGAACTCTGTAG CACCGCCTACATA
9	3524-3559	CTCCATCCAGTCTATTAATTG TTGCCGGGAAGCTA	3525-3490	AGGCGGATAAAGTTGCAGGACC ACTTCTGCGCTCG
10	3557-3592	TAGAGTAAGTAGTTCGCCAGT TAATAGTTTGCGCA	3558-3523	TAGCTTCCCGGCAACAATTAAT AGACTGGATGGAG

No.	Position (bp)	Top strand sequence	Position (bp)	Top strand sequence
11	3647-3682	TTCATTCAGCTCCGGTTCCCA ACGATCAAGGCGAG	3648-3613	AAGCCATACCAAACGACGAGCG TGACACCACGATG
12	3996-4031	TTACCGCTGTTGAGATCCAGT TCGATGTAACCCAC	3997-3962	AAGATCCTTGAGAGTTTTCGCC CCGAAGAACGTTT
13	4040-4075	CAACTGATCTTCAGCATCTTT TACTTTCACCAGCG	4041-4006	TGGGTGCACGAGTGGGTTACAT CGAACTGGATCTC
14	4201-4236	TCATGAGCGGATACATATTTG AATGTATTTAGAAA	4202-4167	GAGACAATAACCCTGATAAATG CTTCAATAATATT

Table III: Sequence information of EcT4 cleavage sites on pBR322*, detected using the Illumina NGS platform. The sequences have been characterised as strong and weak.

No.	Position (bp)	Top strand sequence
		Strong topo IV cleavage sites
1	1414-1470	ACGCGGCGCATCTCGGGCAGCGTTGGGTCCTGGCCACGGGTGCGCATGATCGTGCT
2	1846-1902	AGCATCCTCTCTCGTTTCATCGGTATCATTACCCCCATGAACAGAAATCCCCCTTA
3	2809-2865	CCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACAC
4	3890-3946	TTGCTCTTGCCCGGCGTCAACACGGGATAATACCGCGCCACATAGCAGAACTTTAA
		Weak topo IV cleavage sites
1	379-434	CCTCTACGCCGGACGCATCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTG
2	493-549	GAGCGCTTGTTTCGGCGTGGGTATGGTGGCAGGCCCCGTGGCCGGGGGACTGTTGG
3	861-916	ACGCCCTCGCTCAAGCCTTCGTCACTGGTCCCGCCACCAAACGTTTCGGCGAGAAG
4	930-985	CCGGCATGGCGGCCGACGCGCTGGGCTACGTCTTGCTGGCGTTCGCGACGCGAGGC
5	1169-1225	TCGGCGAGAAGCAGGCCATTATCGCCGGCATGGCGGCCGACGCGCTGGGCTACGTC
6	1569-1625	AACGTCTGCGACCTGAGCAACAACATGAATGGTCTTCGGTTTCCGTGTTTCGTAAA
7	2180-2236	CGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCATGACCCAGTCACGTAGCGAT
8	2413-2469	CAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGA
9	2661-2717	GTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGT
10	2980-3036	TATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGC
11	3375-3431	TAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACC

Table IV: Sequence information of EcT4 cleavage sites on pBR322*, detected using the Illumina NGS platform. The last column is the GC content (%) of each fragment (pBR322* GC content is 53%). Fragment number 2 was topology-sensitive and only cleaved when negatively supercoiled pBR322* was used.

No.	Position (bp)	Top strand sequence	GC content (%)
1	467-607	GATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTCG GCGTGGGTATGGTGGCAGGCCCCGTGGCCGGGGGACTGTT GGGCGCCATCTCCTTGCATGCACCATTCCTTGCGGCGGCG GTGCTCAACGGCCTCAACCTA	66.0
2	1353-1493	GAATGCGCAAACCAACCCTTGGCAGAACATATCCATCGCG TCCGCCATCTCCAGCAGCCGCACGCGGCGCATCTCGGGCA	64.5

		GCGTTGGGTCCTGGCCACGGGTGCGCATGATCGTGCTCCT GTCGTTGAGGACCCGGCTAGG	
3	2637-2777	GGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGC TTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGT GGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCG GTGTAGGTCGTTCGCTCCAAG	59.6