DNA Topological Stress during DNA Replication in *Saccharomyces cerevisiae*

Nicola E Minchell

University of Sussex

Doctor of Philosophy in Genome Stability

September 2018

STATEMENT

I hereby declare that this thesis has not been previously submitted to this or any other University for a degree.

Signature:

Nicola E Minchell

ACKNOWLEDGMENTS

I would firstly like to thank the University of Sussex and Life Sciences department for funding a PhD scholarship to allow me to carry out this work.

I would like to thank the past and present members of the Baxter Lab for their continual support, pizza and chocolate during my PhD. In particular, I would like to thank my supervisor Jon, who has always supported and believed in me during this four year journey. Thank you for giving your time, invaluable guidance and knowledge, and your care into progressing both my project and myself as a scientist. I would like to thank Stephi for her friendship, for being the fountain of knowledge and help, and for always telling me to apply for jobs, grants and talks, which have been extremely helpful. Thanks to Andrea for her endless help, patiently showing me lab techniques and for putting up with my annoyance! Thank you to Catrina, Sahar and Rose, my fellow PhD's whose 'positive energy' and companionship keep the lab cheerful and fun to be in. Finally, I would like to thank all the members of the GDSC community for making the environment so friendly and easy to work in.

SUMMARY

University of Sussex

Nicola Minchell

Doctor of Philosophy in Genome Stability

DNA Topological Stress during DNA Replication in *Saccharomyces cerevisiae*

DNA topological stress impedes normal DNA replication. If topological stress is allowed to build up in front of the replication fork, the fork rotates to overcome the stress, leading to formation of DNA pre-catenanes. The formation of DNA pre-catenanes is therefore a marker of DNA topological stress. In this study, I have examined how transcription linked DNA topological stress impacts on fork rotation and on endogenous DNA damage. Transcription, similar to replication, affects the topology of the DNA; and collision between the two machineries is likely to lead to high levels of DNA topological stress. I found that the frequency of fork rotation during DNA replication, increases with the number of genes present on a plasmid. Interestingly, I also found that this increase in pre-catenation is dependent on the cohesin complex. Cohesin and transcription are known to be linked, as transcription leads to the translocation of cohesin along budding yeast DNA away from its loading sites. Cohesin plays a major role in establishing chromosomal structure, influencing gene expression and genetic inheritance. In this work, I have analysed the relationship between cohesin and the generation of topological stress and found that topological stress associated with cohesin can lead to DNA replication stress and DNA damage.

CONTENTS

List of Figures

List of Tables

1 INTRODUCTION

1.1 DNA TOPOLOGY

1.1.1 The DNA double Helix

The DNA molecules of a cell provide all the information necessary for the functions of that cell and for its essential ability to proliferate and replicate. However, this vital information is safely kept within the inside of the anti-parallel double stranded helical DNA structure.

The DNA strand is made up of a repeating pattern of nucleotides, each of which comprises a phosphate, a sugar moiety, and one of four separate bases (adenine, guanine, thymidine and cytosine; A, G, T, C respectively) (Figure 1.1). In the Watson-Crick base paring model, two strands of nucleotides, each a mirror image of the other, run antiparallel to one another in a right handed fashion, with hydrogen base pairing between either adenine and thymine or cytosine and guanine (Watson and Crick, 1953b). When the nucleotides are stacked on top of one another each is 36° offset from the last, which leads to the helical three dimensional form of DNA (Figure 1.1).

When Watson and Crick discovered the structure in 1953, they realised even then that the helical nature would potentially provide a major problem for the duplication of the genetic code.

'Since the two chains in our model are intertwined, it is essential for them to untwist if they are to separate… although it is difficult at the moment to see how these processes occur without everything getting tangled, we do not feel that this objection will be insuperable' (Watson and Crick, 1953a).

Indeed, although the DNA replication fork encounters many problems in unwinding the DNA, it is able to overcome this again and again in countless numbers of cells.

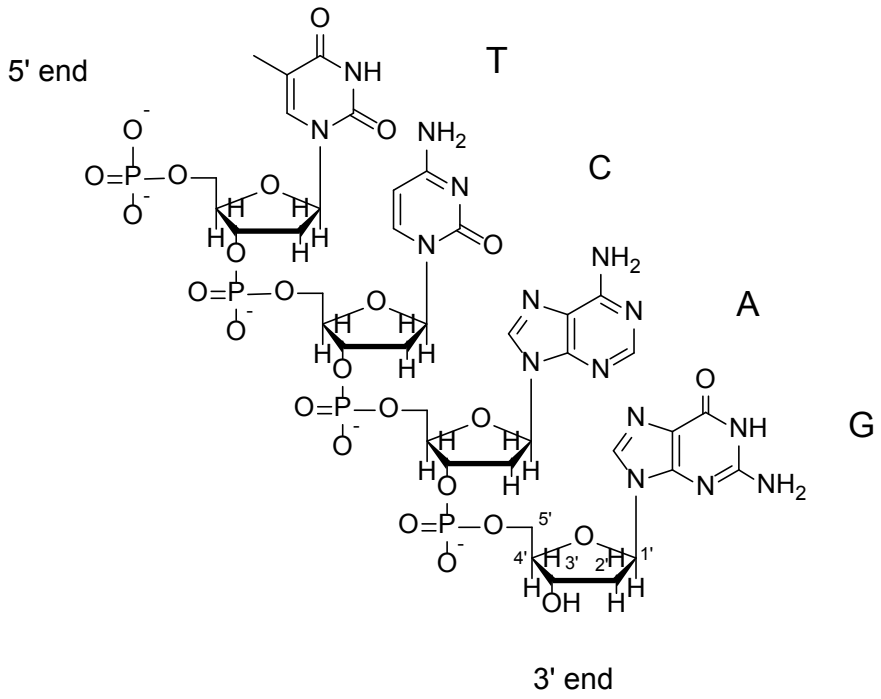


Figure 1.1 Nucleic acid structure

A di-nucleotide chain, showing the phosphodiester bond connecting the nucleotides together and the numbering system for a pentose ring shown in the lower nucleotide. Each nucleotide comprises a sugar (2' deoxyribose), a phosphate, and a base. The bases will hydrogen bond to the opposing bases of the other nucleotide chain of a DNA double strand. Made using ChemDraw software (PerkinElmer Informatics).

1.1.2 Supercoiling and Linking Number (Lk)

The problem that Watson and Crick referred to in their 1953 paper, was that the extreme length of DNA molecules means that pulling them apart in order to duplicate them, will lead to topological entanglements ahead of the replication fork, known as supercoiling. This problem arises due to there being one linkage between the two DNA stands every ~10.4 base pairs.

The extent of supercoiling of a DNA molecule is measured using the linking number of the DNA (Lk). Lk is defined as the number of times one strand of DNA crosses the other strand, if the molecules are imagined in one plane, and the directionality with which it occurs. (Technically linking number only relates to closed circular molecules, however the length and anchorage points of DNA lead to a similar situation, and the properties of linking number can be applied) (Bates and Maxwell, 2005).

Linking number is the sum of the geometric parameters, twist (tw) and writhe (wr).

$$Lk = Tw + Wr$$

Where twist is a function of the rotation of the two strands of DNA around the helical axis (the number of double helical turns), and writhe is a measure of the path of the DNA axis coiling in space. DNA is generally found in the cell as the B form structure first described by Watson and Crick with ~10.4 bp/turn, because it appears to be energetically favourable to stay in this stable form. During DNA replication however, pulling apart the two DNA strands pushes the links between those strands in front of the replication fork. Due to the stability of the B form DNA, an increase in linking number as such, results in an increase in writhe, leading to the formation of supercoiling. This increase in writhe prevents a change in twist occurring, which, when is affected, leads to different forms of DNA structure, such as A or Z, to compensate for linking number change.

Supercoiling, therefore, can be defined as the changes in twist and writhe when the linking number of DNA is not equal to zero (Bates and Maxwell, 2005). The extent of supercoiling of a DNA molecule is measured by linking difference (ΔLk):

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

Lk° is the 'hypothetical' linking number and is exactly equal to the length of the DNA divided by the number of base pairs per turn of helix when linear and relaxed. Therefore, a measure of supercoiling is what remains when the linking number of the relaxed molecule, is taken away from the linking number of the molecule under torsion.

An increase in the linking difference of a DNA molecule will manifest as a change in the writhe of the DNA in a clockwise manner, which is termed positive supercoiling (Figure 1.2). A decrease in linking difference also manifests as a change in writhe however, in an anti-clockwise manner, which is known as negative supercoiling (Figure 1.2). During DNA transcription, part of the DNA template becomes unwound, leading to the formation of positive writhe ahead of the RNA polymerase. Behind the polymerase, where the DNA can reanneal, negative supercoiling forms to compensate for the increase in twist required to form duplex DNA (Liu and Wang, 1987). However, during DNA replication the two strands which are unwound, do not reanneal to one another behind the fork, so do not form compensatory negative supercoiling. However, linking can occur between two sister DNA molecules, and is known as DNA catenation.

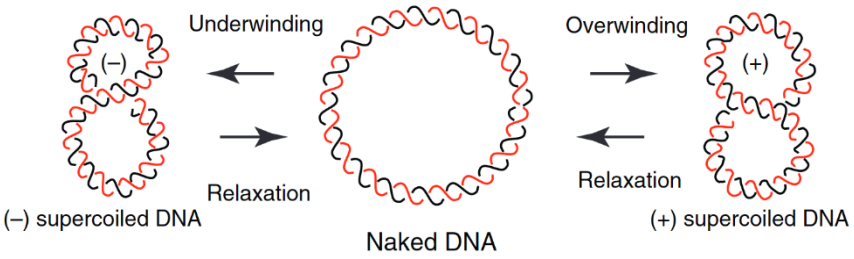


Figure 1.2 Positive and negative supercoiling and relaxation

Modified from Baxter and Aragon (2012). Addition of a linkage leads to overwinding and a clockwise writhe, forming positively supercoiled DNA. Removal of a linkage leads to underwinding and an anti-clockwise writhe, forming negatively supercoiled DNA.

1.1.3 DNA Catenation

A DNA catenane, similarly to supercoiling, is described when thinking about closed circular DNA, and is defined as an interlink between two closed circular DNA molecules (Figure 1.3A). The number of catenanes is measured by directionality of nodes (crossings) when viewed in one plane. A single catenane requires two nodes of either positive or negative directionality (Figure 1.3A). There are three types of DNA catenanes, notated as CatA, CatB and CatC. CatA is a catenane where one of the strands of DNA, on both of the linked duplex DNA's, has been nicked (Figure 1.3A). CatB is a catenane where one of the strands of the DNA on one of the duplex DNA's has been nicked (Figure 1.3B). And CatC describes two intact plasmid DNA molecules interlinked and therefore these can also be supercoiled (Figure 1.3C).

Although the exact description of catenanes requires a covalently closed circle, the word is also used to describe the situation of sister chromatid intertwines (SCIs). This is due to the extreme length of chromosomes, meaning that intertwines between two separate double stranded DNA molecules have similar properties to links between closed circular DNA (i.e. they cannot be resolved without nicking or breaking the DNA strands). Catenanes can be formed via two methods, the first is during DNA replication, when links between the two DNA strands are pushed ahead of the fork and lead to supercoiling. These links can be moved behind the replication fork via replication fork rotation, where they now manifest as a link between two sister chromatids (discussed in detail in section 1.4). The second method is via Topoisomerase action; these are enzymes that are able to cut DNA in order to change the linking number of the DNA, either between strands, or between chromatids.

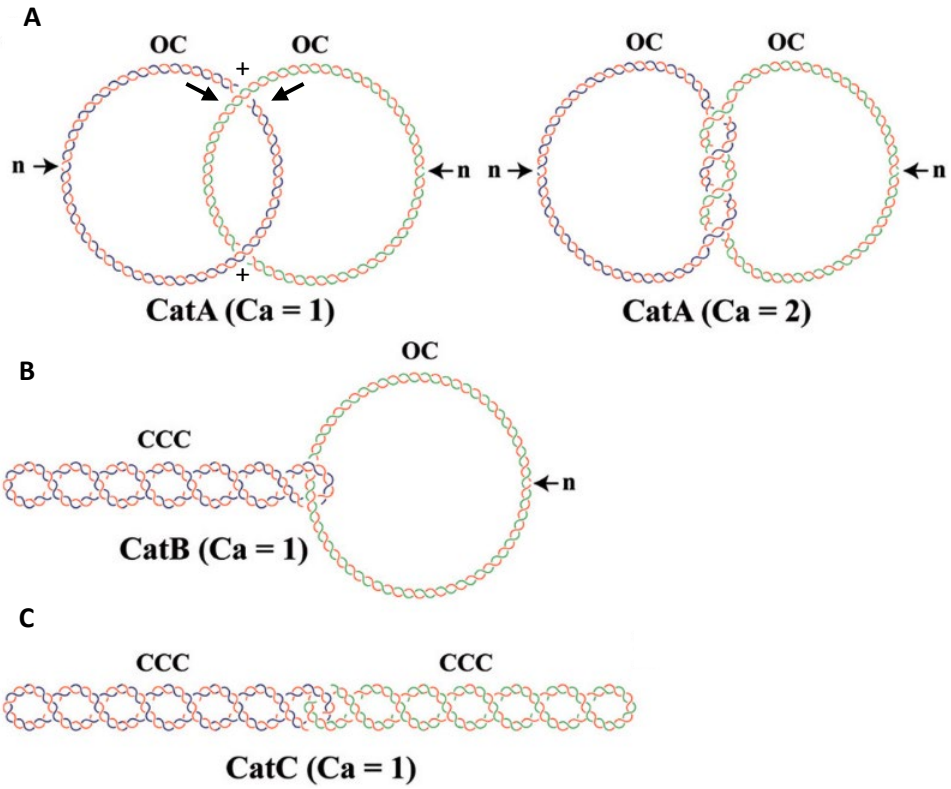


Figure 1.3 Forms of DNA catenanes

Modified from (Martinez-Robles et al., 2009). **A**. Representation of CatA with different linking numbers, either with one linkage or two linkages, n represents a nick in the DNA duplex, the DNA is open circular (OC), the arrows represent an arbitrary directionality given to the DNA circles in order to calculate the node crossings. **B**. CatB with a linking number of 1 where one of the two DNA duplexes has a nick and therefore is in a relaxed Open circular (OC) form and the other is supercoiled covalently closed circular form (CCC). **C**. CatC with a linking number of 1, where neither of the two DNA duplexes has a nick (CCC).

1.2 DNA TOPOISOMERASES

Topoisomerases are enzymes that cut DNA strands in order to change the linking number, leading to addition or resolution of topological stress and entanglements. There are two types of topoisomerases, type I and type II. They are classified as such depending on if they cut a single strand of the DNA duplex or a double strand. Topoisomerases exist across all classes of organisms and are essential. All studied organisms have both type I and type II and they perform the same basic functions.

The type I group of enzymes are able to act upon a single strand of the duplex DNA (Figure 1.4) (Depew et al., 1978; Liu and Wang, 1979). Type I is further sub-grouped into topoisomerase IA and topoisomerase IB (Vos et al., 2011). Topoisomerase IA enzymes act preferentially on single stranded regions, and preferentially relax negatively supercoiled DNA. These enzymes bind to the DNA and cleave one of the stands by a transesterification reaction, which results in a transient 5'phosphotyrosyl bond with the topoisomerase, and a 3' hydroxyl moiety (Figure 1.4A). In the case of single stranded regions, this then gives these enzymes the ability to form or remove catenanes by passing a DNA double helix through the newly formed double stranded gap. In the case of resolving negative supercoiling, the intact DNA strand will be passed through the single stranded gap and the broken ends are then resealed leaving the linking number of the DNA changed by one (Brown and Cozzarelli, 1981; Tse and Wang, 1980).

Type IB topoisomerases can proficiently relax both positively and negatively supercoiled DNA, and they act via a method of swivelling. In this case, a single strand of DNA is cleaved, but the topoisomerase tyrosyl links to a 3'phosphoryl, and leaves a 5'hydroxyl moiety. The free DNA end is then allowed to rotate around its axis thereby returning the DNA to its relaxed form, followed by resealing of the broken ends (Figure 1.4B) (Champoux and Dulbecco, 1972; Koster et al., 2005; Wang, 1971). Type IB enzymes are important during both replication and transcription. They play a key role relaxing positive supercoiling in front of the replication fork, to allow for replication fork progression (Kim and Wang, 1989). During transcription they also relax the positive supercoiling in front of the progressing RNA polymerase, and in eukaryotes, topoisomerase IB also have a role for the relaxation of the negative supercoiling behind the polymerase to prevent DNA melting.

Type II topoisomerases are grouped into two sub-groups: topoisomerase IIA and topoisomerase IIB, and are able to act upon both positive and negative supercoiling as well as DNA catenanes. Both types work by a strand passage mechanism similar to Type IA enzymes and also use the

same transesterification reaction as Type IA. However, instead of cutting single stranded DNA, they cut duplex DNA and utilise ATP in order to pass the second double stranded DNA moiety through the gap created (Figure 1.4C) (Brown and Cozzarelli, 1979; Liu et al., 1979; Wang, 1998). Type II topoisomerases are involved in the relaxation of positive supercoiling in front of the DNA replication fork, and the relaxation of positive and negative supercoiling around an RNA polymerase. Often this function can be compensated for by the action of Type I topoisomerases (Kim and Wang, 1989), however, Type II are essential in removing the DNA catenanes formed during DNA replication to allow for segregation during mitosis (Uemura et al., 1987).

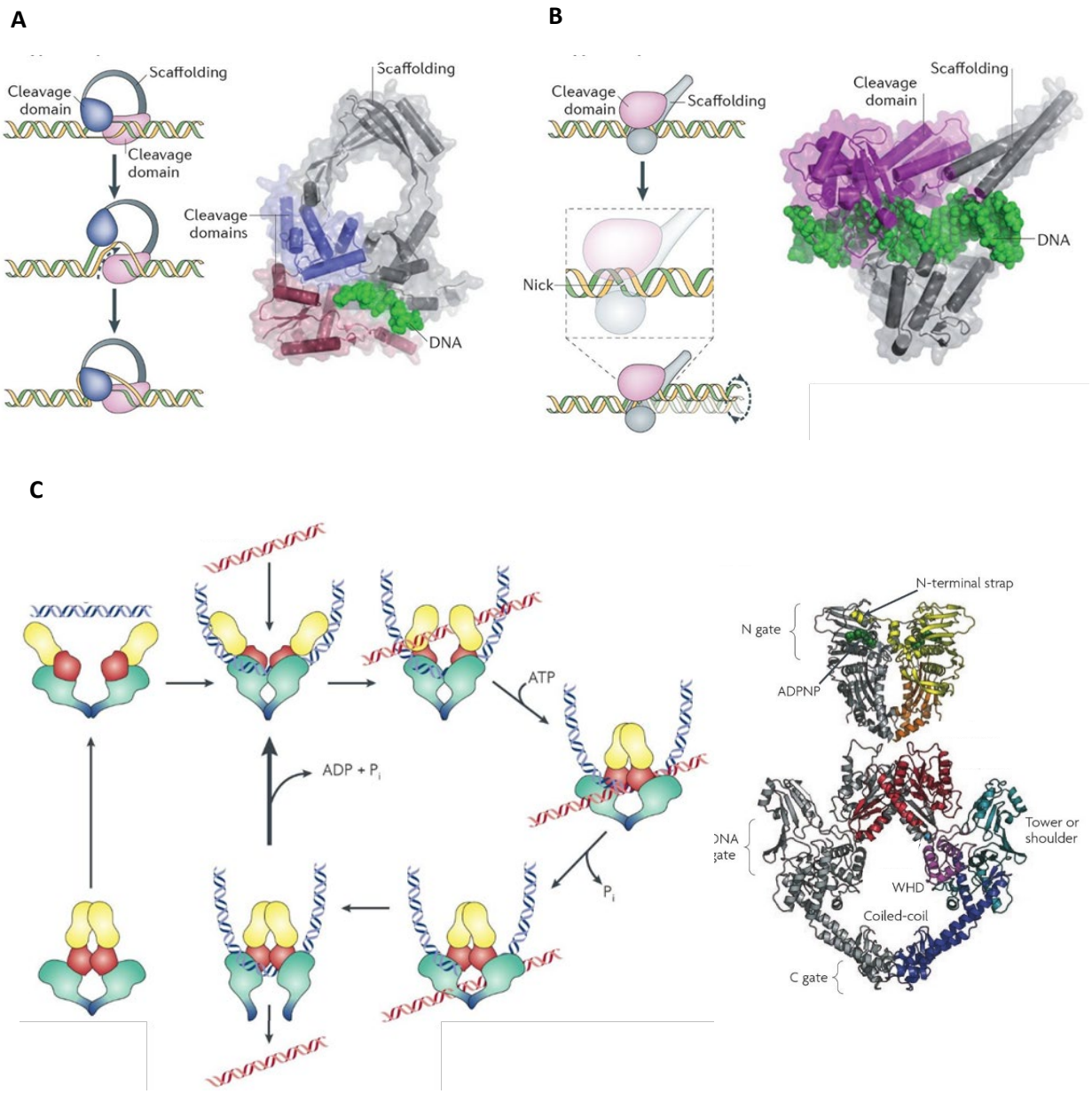


Figure 1.4 Mechanism of action of Type IA, Type IB and Type IIA topoisomerases

Type I topoisomerases taken from Vos et al. (2011), Type IIA topoisomerase taken from Nitiss (2009). **A.** Type IA topoisomerase mechanism and the structure of *E.coli* DNA topoisomerase III bound to DNA. A single stranded DNA break is made (green) and another single stranded DNA segment (yellow) is passed through the gap. **B.** Type IB topoisomerase mechanism and the structure of human topoisomerase IB bound to DNA. A single stranded DNA break is made (yellow) allowing the other DNA strand (green) to rotate with respect to the first. **C.** Top IIA topoisomerase mechanism and the structure of *S. cerevisiae* Top2 based on structures of the ATPase domain and the breakage reunion domain. Both strands of a DNA double helix are cleaved, and a second DNA duplex is passed though the gap. ATP binding allows for a closed clamp conformation, hydrolysis steps then allow for stand passage and clamp opening.

1.2.1 *Saccharomyces cerevisiae* Topoisomerases

The model organism used in this study, *Saccharomyces cerevisiae*, has three topoisomerases proteins; Top1, Top2 and Top3. Top3 is a Type IA Topoisomerase that is able to cut single stranded DNA to resolve sister chromatid intertwines (Wallis et al., 1989). It is thought to foremost act in resolving holiday junctions during homologous recombination in a complex with Sgs1 and Rmi1 (STR complex) (Bizard and Hickson, 2014; Chang et al., 2005; Gangloff et al., 1994; Tang et al., 2015). Its deletion leads to very slow growth, high levels of genome instability, and an inability to sporulate (Gangloff et al., 1999).

Top1 and Top2 are involved in regulating topological changes during replication and transcription. Top1 is a Type IB 'swivelase' enzyme, which, during transcription interacts with the C-terminus of RNA polymerase II (RNAPII) (Phatnani et al., 2004), and when deleted leads to highly negatively supercoiled DNA, suggesting its main role is to remove negative supercoiling during transcription (Brill and Sternglanz, 1988). However, mutants are still viable when *TOP1* is deleted, as Top2 is able to generally compensate for its loss. In replication, a direct interaction was identified between Top1 and Tof1 (topoisomerase 1 interacting factor 1) (Park and Sternglanz, 1999), a protein that travels with the replication fork. Indicating a potential recruitment of Top1 to the replication fork via Tof1, which could help relax supercoiling generated in front of the replication fork.

Yeast Top2 is a Type IIA subfamily member, which appears to have a preference for relaxing positive, over negative supercoiling (Fernandez et al., 2014; French et al., 2011). This leads to the suggestion that, during transcription, Top2 relaxes positive supercoiling ahead of the polymerase and Top1 relaxes negative supercoiling behind (French et al., 2011). A necessity for Top2 during transcription was shown by Joshi et al. (2012) using a *top2-4* temperature sensitive mutant, where transcripts of over 3.5 kb in length were severely decreased in the *top2-4* strain compared to a Top2 active strain (Joshi et al., 2012). However the analysis indicated that Top1 was able to generally compensate for the loss of Top2, with most transcript levels not changing significantly (Joshi et al., 2012). Top2 is however indispensable for mitosis, as was shown by Holm et al. (1985), who found cell death occurred in Top2 mutant cells at the time of mitosis onset. This was later found to be due to decateantion of newly replicated sister chromatids being prevented, further preventing separation and leading to mis-segregation (Baxter and Diffley, 2008).

1.3 DNA REPLICATION IN EUKARYOTES

Unlike in bacteria, where DNA replication and cell division can be ongoing in the same cells, in eukaryotes the cell cycle temporarily divides into four main stages; Gap one phase (G1), Synthesis phase (S), Gap two phase (G2) and Mitosis (M). In order to ensure faithful duplication and segregation, these processes are highly regulated, and temporally segregated. DNA replication takes place during S-phase, which can again be subdivided into three phases: initiation, elongation and termination.

Replication initiation origin licensing actually starts taking place during G1 phase. In eukaryotes the Origin Recognition Complex (ORC) binds to origin sequences in an ATP dependent manner, followed by recruiting Cdc6 and Cdt1 (Bell and Stillman, 1992; Liang et al., 1995). These in turn recruit the essential helicase component minichromosome maintenance (MCM) complex, which completes the formation of the pre-RC (pre-replicative complex) (Evrin et al., 2009; Remus et al., 2009).

Origin formation takes place at multiple points across chromosomes to ensure the full replication of the genome (Blow et al., 2011). In order to prevent re-licensing and genome destabilising re-replication in one cell cycle, the process of origin licensing is highly regulated. The main control of the cell cycle is the phosphorylation of various substrates by CDK (cyclin dependent kinases) in complex with cyclins. Cyclin levels are kept low by the anaphase promoting complex (APC), which is abundant from the metaphase/anaphase transition to G1 phase (Shirayama et al., 1999). This means cyclin levels are kept low from anaphase to G1 phase, and CDK is also inhibited by the action of CKIs (CDK inhibitors). The levels then become high during S and G2 phases, and as CDK activity is inhibitory for pre-RC formation, this leads to cell cycle controlled replication (Diffley, 2004).

The next step in initiation is origin firing, which cannot occur without the sequential recruitment of at least seven other proteins. The recruitment starts with the Cdc45-Sld3-Sld7 complex, followed by the GINS complex (Sld5, Psf1, Psf2, Psf3), the binding of which finishes the formation of the CMG (Cdc45-Mcm2-7-GINS) complex, which is the active form of the helicase (Figure 1.5) (Bauerschmidt et al., 2007; Gambus et al., 2006). This process is controlled by the activities of S-CDK and Dbf4-dependent kinase (DDK), which allows recruitment of the proteins by phosphorylation, and firing of the replication origins during S-phase (Labib, 2010). The minimal required substrates for S-CDK phosphorylation, in *S. cerevisiae* for initiation of replication, are the essential proteins Sld2 and Sld3, which interact with Dpb11 in order to recruit components

of the CMG complex and Pol ε (Kamimura et al., 2001; Muramatsu et al., 2010; Tanaka et al., 2007; Zegerman and Diffley, 2007).

Once an origin is fired, two bi-directional forks will ensue due to the two loaded MCM complexes, and they will then unwind the helix until reaching the next fork, where the termination of replication occurs. Behind the helicase, RPA (replication protein A) coats the single stranded DNA (ssDNA) to prevent re-annealing, followed by DNA polymerase α priming the DNA (Fanning et al., 2006). PCNA is then loaded, followed by recruiting and stabilising both Pol ε and Pol δ (Moldovan et al., 2007). Pol ε functions as the leading strand polymerase and Pol δ as the lagging strand polymerase (Nick McElhinny et al., 2008; Pursell et al., 2007).

Along with the core replisome proteins, other proteins are known to associate with the fork, and form a complex called the replisome protection/progression complex (RPC) (Figure 1.5). This includes the proteins Tof1, Csm3, Mrc1, Ctf4, FACT and Top1, which are thought to perform a major function of linking the helicase to the polymerases (Gambus et al., 2009). They also have a wide range of other functions such as the checkpoint response, relaxation of helical stress and stabilisation of the replication fork (reviewed in Leman and Noguchi (2012)).

At the end of DNA replication elongation, the replisomes generated at every origin will converge with replisomes from neighbouring fired origins, to complete replication. Exactly how this occurs without leaving any unreplicated DNA, and without leading to fork stalling or DNA damage is still unclear. However a recent study in *Xenopus* indicated that the replisomes do not come off the DNA when they converge, they replicate past one another until reaching the replicated DNA of the other strand (Dewar et al., 2015).

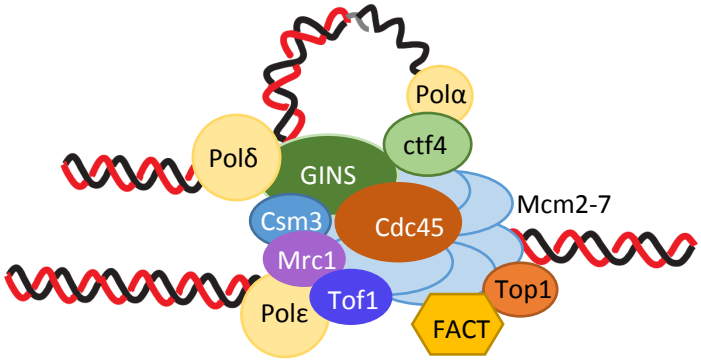


Figure 1.5 Schematic of an elongating *S. cerevisiae* DNA replication fork

The MCM2-7 proteins together with the GINS complex and Cdc45 make up the core CMG complex that represents the active form of the helicase. Polδ is the lagging strand polymerase which is primed by Polα, and Polε is the leading strand polymerase. Other proteins that associate with an elongating fork include various proteins of the RPC.

1.4 DNA Replication Fork Rotation

When the MCM helicases pull apart the duplex DNA and break the hydrogen bonds between the bases, they cannot remove the innate linkages between the DNA strands. These are instead pushed in front of the replication fork as it processes along the DNA. Therefore, as the linking number in front of the fork increases, this leads to a direct increase in DNA topological writhe, which will lead to the formation of positive supercoiling ahead of the replication fork (Figure 1.6A) (Postow et al., 2001).

The increase in topological tension can be detrimental to the progression of the replication fork. If linkages are continually being pushed in front of the fork, the DNA will eventually become highly topologically stressed, which will prevent the helicase from breaking the hydrogen bonds between the bases (Postow et al., 2001). In this situation, the replication fork will stall, it may even lead to replication fork reversal and fork collapse (Baxter, 2014; Bermejo et al., 2007). Therefore, in order to prevent fork collapse and potentially DNA damage, the increase in linking number in front of the fork must be decreased, and there are two pathways to achieve this.

In the first pathway topoisomerases I and II are able to cut the DNA in front of the replication fork, to allow for relaxation of the topological stress, and then re-seal the DNA ends. This relaxes the positive topological stress, preventing a build-up becoming detrimental to the replication fork (Brill et al., 1987).

In the second pathway, first proposed by Champoux and Been (1980), the replication fork itself is able to rotate on the helical axis of the DNA, which moves the linkages between the strands from in front of the replication fork, to behind the fork. This means that the topological stress in front of the fork is removed, relaxing the writhe of the DNA, and allowing the fork to progress unhindered. However the intertwines in this situation are not removed from the DNA entirely, they now manifest behind the replication fork as sister chromatid intertwines known as DNA pre-catenanes (Figure 1.6B). If DNA pre-catenanes are not resolved, they mature into full DNA catenanes on completion of DNA replication. These intertwines must also be removed by the action of topoisomerases, in this case Type II only, in order to allow for the faithful segregation of sister chromatids during Mitosis (Champoux and Been, 1980; Holm et al., 1985; Keszthelyi et al., 2016).

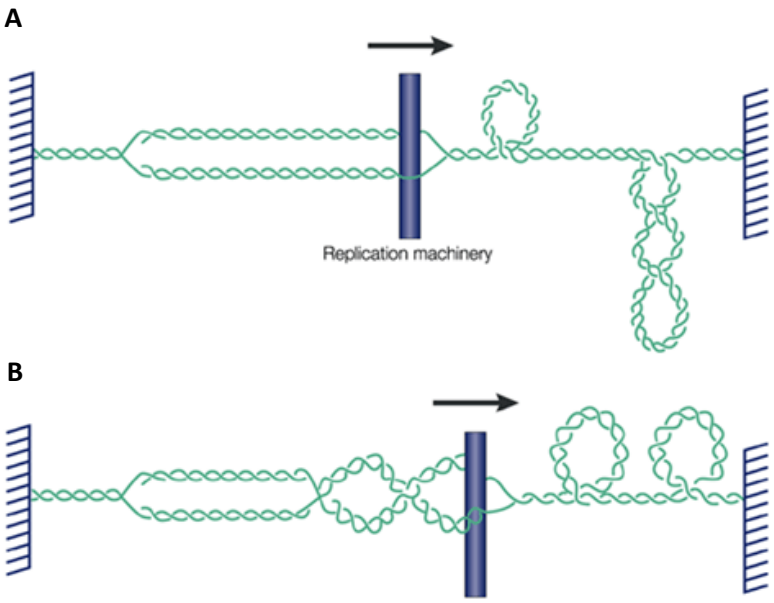


Figure 1.6 Formation of DNA catenanes by replication fork rotation

Taken from Wang (2002). The rod represents the replication fork and the arrow its direction of movement. DNA is envisaged to have barriers preventing extensive diffusion of topological stress. **A**. As the fork progresses along the DNA, the base pairing is broken and the linkages pushed in front of the fork leading to positive helical stress. **B**. positive helical stress in front of the replication fork is diffused via the action of fork rotation, forming sister chromatid intertwines behind the fork.

1.5 Frequency of Fork Rotation

Although two pathways exist, it appears that the pathway of fork rotation is generally restricted. There is a balance between the pathways, and during replication this is greatly tipped towards topoisomerase action ahead of the fork. It is much more energetically favourable for topoisomerases to act ahead of the fork, as oppose to fork rotation occurring. If the action to resolve the supercoiling ahead of the fork is sufficient, then there is no reason for the second pathway to be utilised (Baxter, 2014). In what contexts then does replication fork rotation occur? When topoisomerase action ahead of a replication fork is insufficient to resolve topological stress, fork rotation can be utilised to prevent fork slow down or stalling. The first such situation to be described was during replication termination.

1.5.1 Replication Termination and Fork Rotation

Replication termination occurs when two DNA replication forks converge upon one another at the completion of synthesising sister chromatids. Supercoiling is continually being formed in front of the replication forks, followed by resolution by topoisomerase action. However, as the two forks come together, not only is the diffusion of supercoiling prevented, but the positive supercoiling will converge to create an area of high topological stress. In this situation, it is necessary for the helical stress to be relaxed quickly, to allow for the completion of replication, and to prevent any unreplicated regions prevailing.

In a series of papers by Sundin and Varshavsky (1980) (1981) looking at Simian virus 40 (SV40) plasmid replication, they greatly reduced topoisomerase II activity by using a highly concentrated media. Through this treatment they observed a number of discrete DNA bands by gel electrophoresis, which related to various forms of catenated dimer molecules, and also were able to show electron microscopy images of catenated dimers (Figure 1.7). They then proposed that these catenated dimers are formed during the last 100-200bp of DNA replication termination where the 'replication machinery of both forks occupies much of the surrounding space' and therefore that the 'DNA duplex is no longer readily susceptible to relaxation by nicking-closing enzymes' (Sundin and Varshavsky, 1980). Suggesting that the sterical exclusion of topoisomerases, from binding the DNA between the terminating replisomes, forces fork rotation and DNA catenation to occur (Figure 1.8).

These initial findings in SV40 were followed by an *in vivo* yeast study using temperature sensitive Top2 mutants (DiNardo et al., 1984). This study showed, that in the absence of Top2, cells were only able to survive for one cell cycle due to a defect in segregation during mitosis. They also visualised catenated plasmid DNA, by gel electrophoresis, occurring after one round of replication in the absence of Top2 (DiNardo et al., 1984).

Succeeding these, there were a range of papers aimed at understanding the termination of DNA replication and the roles of topoisomerases. Snapka et al. (1988) started to elucidate the different roles of Top1 and Top2, using topoisomerase inhibitors. They found that Top2 inhibitors did not affect replication progression, and therefore proposed that Top1 generally relaxed positive supercoiling during DNA replication elongation. However, at termination, with a sterical exclusion of Top1, fork rotation would occur followed by resolution of catenanes by Top2 (Snapka et al., 1988).

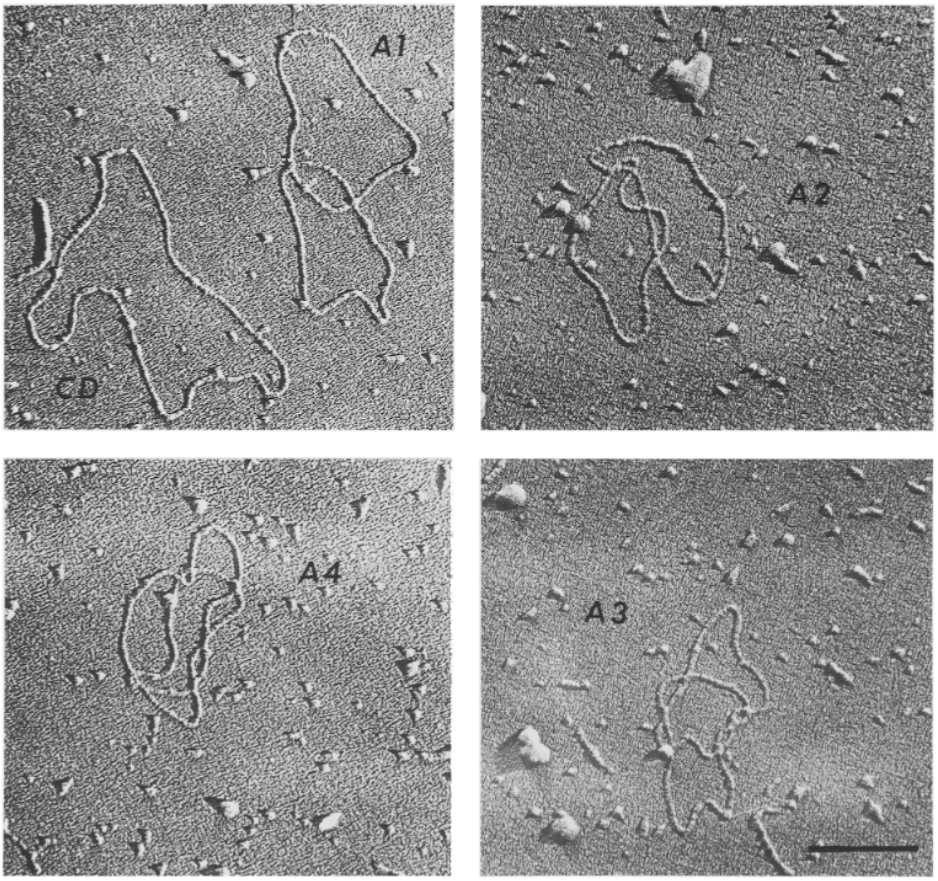


Figure 1.7 Electron Microscopy of Catneated SV40 dimers

Taken from (Sundin and Varshavsky, 1981). Clockwise starting from top left: CatA dimer molecules from a linking number of 1 through to 4.

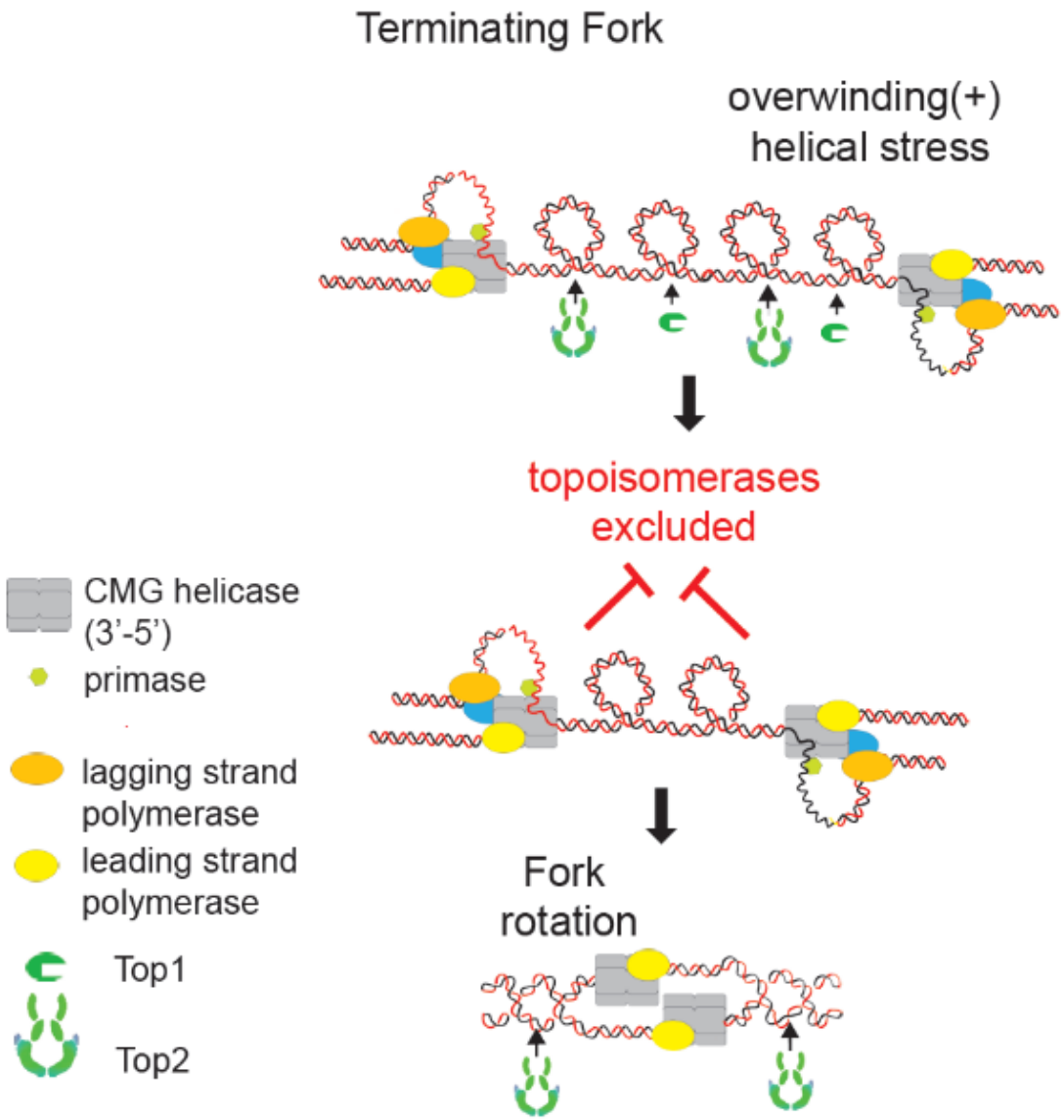


Figure 1.8 Termination of DNA replication and replication fork rotation

Modified from Keszthelyi et al. (2016). As two replisomes converge, overwinding helical stress builds-up between them. When the replisomes come in close proximity of one another, a sterical exclusion of topoisomerases will occur. This leads to fork rotation resolving the helical stress between the replisomes to prevent replication fork stalling.

1.5.2 Fork Rotation is restricted during DNA Replication Elongation

Much less is known about the potential role of replication fork rotation during DNA replication elongation, compared to its role at replication termination. *in vitro,* de-proteinised forks can easily diffuse supercoiling as DNA catenanes behind the fork (Peter et al., 1998). Lucas et al. (2001) also reported that Topo II acts on pre-catenanes during DNA replication in *in vitro Xenopus* egg extracts, and that the formation of DNA pre-catenanes is not restricted to termination. However, the restrictions from *in vivo* replication fork components are likely to be much greater than *in vitro*. *In vivo* it can be imagined that the large size of the replisome, including many accessory factors, hinders its rotation, therefore making action ahead of the fork much more favourable. More recently it has been shown by Schalbetter et al. (2015) that replication fork rotation is generally restricted by the specific replication fork protection/progression components (RPC), Tof1/Csm3 in *S. cerevisiae* (or Timeless/Tipin in humans) (Figure 1.9). It was proposed that this restriction could be due to one of two scenarios. Firstly, Tof1 is known to interact with Top1, potentially bringing it to the replication fork, which presumably allows it to perform its function of resolving the supercoiling ahead of the fork (Park and Sternglanz, 1999). Thus, if Tof1 is not functional, and Top1 is not recruited, then the excess of helical stress will lead to replication fork rotation (Figure 1.9B). The second scenario is that Tof1/Csm3 have a stabilising presence to the replisome, and without them, the coordination of all the components is loosened allowing easy rotation of the fork (Figure 1.9A) (Schalbetter et al., 2015).

This restriction of fork rotation points to the possibility of a detrimental effect due to unrestricted DNA pre-catenane formation (Schalbetter et al., 2015). It is postulated that pre-catenanes are essential for sister chromatid cohesion (SCC) until Mitosis (Farcas et al., 2011; Wang et al., 2008), but conversely that excessive DNA pre-catenation is detrimental due to interrupting processes occurring behind the fork, such as Okazaki fragment maturation, and even lead to cohesion establishment inhibition (Schalbetter et al., 2015). Schalbetter et al. (2015) saw a pronounced G2/M delay following depletion of Top2 in *tof1Δ* cells, and that *TOF1* deletion caused DNA damage through S-phase, which was exacerbated by depletion of Top2. They postulated that the damage occurred during S-phase due to excessive DNA pre-catenation. Therefore, as the presence of catenanes has both positive and negative effects, it is important to strictly regulate their formation.

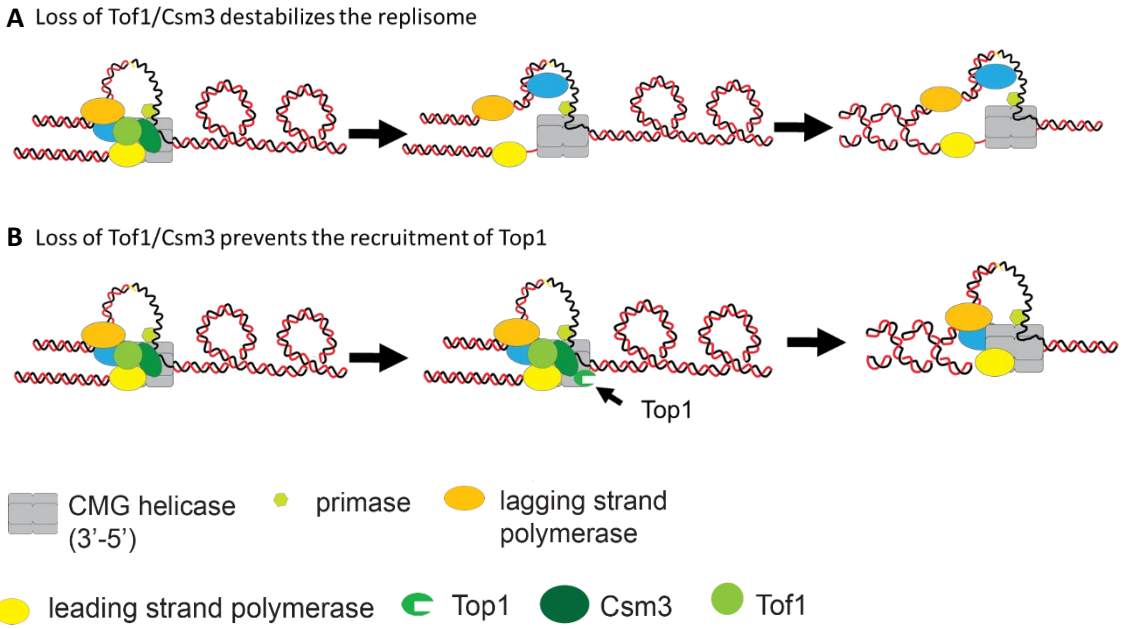


Figure 1.9 Two scenarios for how Tof1/Csm3 inhibit replication fork rotation

A. Tof1/Csm3 have a stabilising effect on the replisome. Loss leads to a mutant replisome that has less resistance to fork rotation. **B**. Tof1/Csm3 recruit Top1. Loss prevents Top1 recruitment, preventing the resolution of supercoiling, leading to a build-up of topological stress which forces fork rotation to occur. Created by J Baxter.

1.5.3 Fork rotation due to Static Block during DNA Replication Elongation

Beyond the Tof1/Csm3 restriction of replication fork rotation, these replication fork protection components have also been seen to allow for replication fork pausing at stable protein-DNA sites. For a long time in yeast it has been understood that replication forks pause at centromeres, due to the binding of kinetochore proteins (Greenfeder and Newlon, 1992). Exactly how the forks eventually overcame the block was discovered by Ivessa et al. (2003), who found the PiF1 family helicase Rrm3 essential to facilitate replication past protein-DNA sites. Not only did Rrm3 help replication past centromeres, but also rRNA genes, tRNA genes, telomeres, inactive replication origins, and transcriptional silencers (Ivessa et al., 2003; Ivessa et al., 2002; Ivessa et al., 2000). Beyond this clear necessity for Rrm3 to help replicate past protein-DNA blocks, Calzada et al. (2005) showed an additional role for Tof1/Csm3 in order for the pausing at these sites to occur. In this study, they were able to slow down the pausing process using RFB barriers that would not be resolved by a fork coming from the opposite direction. They found that when the replisome paused, it was still intact, and that the paused forks did indeed recruit the Rrm3 helicase in order to process through the block. They interestingly saw that both Tof1 and Csm3 but not Mrc1 were necessary for the pausing of these forks, a role which was unrelated to their S-phase checkpoint role (Calzada et al., 2005). This was different to the checkpoint related role they share with Mrc1 at stalled forks due to HU treatment, which requires checkpoint kinases Mec1 and Rad53 to maintain fork integrity (Katou et al., 2003; Lopes et al., 2001).

In the study by Schalbetter et al. (2015), they therefore decided to also look at the effect of various replication fork pausing sites on replication fork rotation. They were able to show that fork rotation is utilised to facilitate unwinding at stable protein-DNA sites, including tRNA genes, inactive origins and potentially centromeres. These sites are thought to lead to fork rotation in a similar manner to the termination of replication; as the fork converges upon the stable protein-DNA site, supercoiling is prevented from diffusing away, and topoisomerases are excluded from in between the fork and the block. This leaves fork rotation as the only pathway available to resolve the topological stress (Figure 1.10) (Schalbetter et al., 2015). In the absence of Tof1/Csm3 this gives the interesting possibility that the reason why replication forks no longer need to pause at stable protein-DNA sites is partially due to a de-restriction or loosening of the fork components, allowing topological stress between the replication fork and the protein-DNA site to readily diffuse.

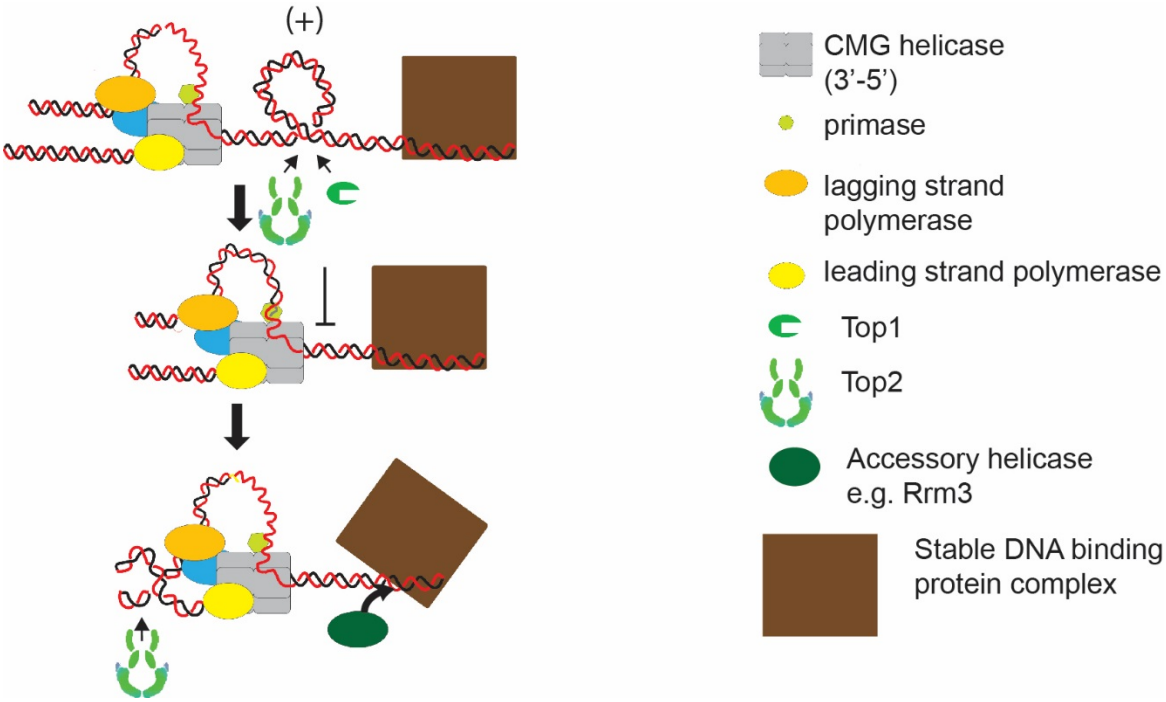


Figure 1.10 Fork rotation due to a static protein-DNA block

Taken from Keszthelyi et al. (2016). As a replication fork converges on a stable DNA binding protein complex the topoisomerases removing positive helical stress in front of the fork will be excluded from acting on the DNA. This leaves fork rotation as an alternative pathway for the resolution of the positive helical stress. The stable DNA binding protein complex can then be removed by an accessory helicase such as Rrm3.

1.6 TRANSCRIPTION AND TOPOLOGICAL STRESS

DNA topological stress is known to affect multiple stages of transcription. DNA supercoiling greatly effects the initiation of transcription; overwound DNA can cause inhibition of transcription (Gartenberg and Wang, 1992), while negative supercoiling facilitates transcription (Dunaway and Ostrander, 1993; Schultz et al., 1992; Tabuchi et al., 1993). This topological landscape is maintained by topoisomerases, especially at highly transcribed genes (Pedersen et al., 2012), and it appears that both Topoisomerase 1 and Topoisomerase 2 promote the recruitment of RNA polymerase II (RNAPII) (Sperling et al., 2011).

Following on from initiation, transcriptional elongation itself affects the local DNA topological environment. The RNA polymerase must have access to the base pair code in order to make nascent RNA transcripts. It appears that the RNA polymerase is prevented from rotating relative to the DNA, possibly due to the immediate processing, and tethering, of nascent RNA transcripts (Liu and Wang, 1987). In the absence of rotation, the unwinding that occurs, leads to what is known as the twin supercoiled domain. This is a localised topological change of positive supercoiling ahead of the transcription bubble and compensatory negative supercoiling behind (Figure 1.11) (Liu and Wang, 1987; Wu et al., 1988).

DNA positive supercoiling forming ahead of the transcription bubble, similarly to DNA replication, would be predicted to impede ongoing transcription. However, topoisomerase action only appears to be required at highly transcribed or long genes (Joshi et al., 2012; Schultz et al., 1992), where it is assumed topological stress is high enough to arrest transcription. In human cells, it is suggested that an interaction between RNAPII and Top1, facilitates the rapid relaxation of positive supercoiling ahead of the transcription bubble (Baranello et al., 2016). However, in yeast, at the rDNA array, Top2 appears to have a preference for action on positive supercoiling ahead of the transcription bubble, while Top1 relaxes negative supercoiling behind (El Hage et al., 2010; French et al., 2011).

The twin domains will disappear when the RNA polymerase comes off the DNA, leading to a quenching of topological stress. However, as positive supercoiling is predominantly relaxed during elongation, when the domains do converge this will lead to a general decrease in linking number. The negative supercoiling behind the transcription bubble, allows for efficient open RNAP-DNA complex formation (Kouzine et al., 2004; Kouzine et al., 2008), but can be problematic if allowed to accumulate. Changes in the twist of the DNA can lead to weak B form DNA, causing non B-DNA structures or DNA:RNA hybrids (R-loops) to form (Drolet et al., 1994).

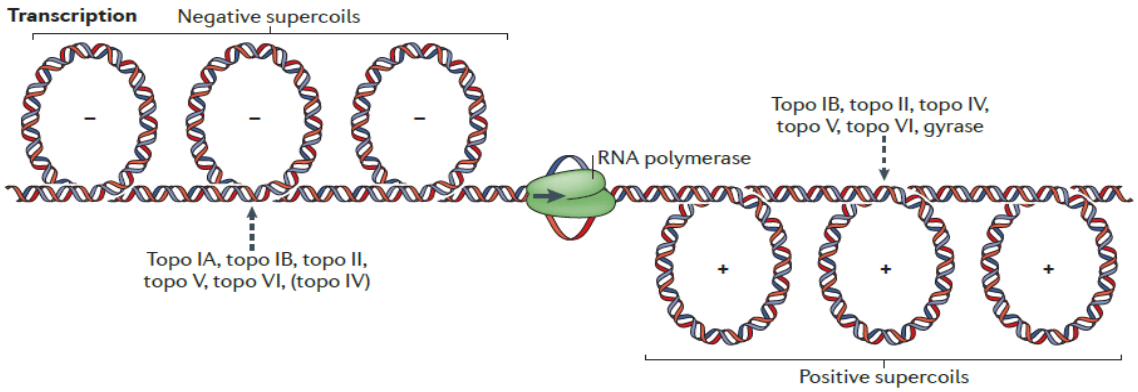


Figure 1.11 Twin supercoiled domain model

Taken from Vos et al. (2011). When the DNA double helix is unwound by the RNA polymerase the linkages between the DNA strands are pushed ahead of the transcription bubble, leading to the formation of positive supercoiling. As the strands re-anneal behind the transcription bubble, negative supercoiling forms to compensate for the twist of the DNA.

1.7 TRANSCRIPTION AS A CAUSE OF DNA REPLICATION STRESS

The term DNA replication stress has been known to encompass a range of issues during DNA replication, and therefore has been hard to define. Most recently it has been termed broadly as 'the slowing or stalling of replication fork progression and/or DNA synthesis' (Zeman and Cimprich, 2014). Replication stress often leads to expression of common fragile sites (CFS), areas that commonly acquire double strand breaks, and are thought to drive oncogenesis (Barlow et al., 2013). Common Fragile sites are found to overlap with very long genes and areas of high replication-transcription collision, implicating this, as well as RNA:DNA hybrids, as one of the main causes of DNA replication stress (Barlow et al., 2013; Helmrich et al., 2013; Helmrich et al., 2011).

1.7.1 Effects of negative supercoiling

Transcription can lead to many problems when it occurs during DNA replication; this is because it is unique in that it presents both a steric barrier to fork progression and also changes the local topology of DNA. Changes to local DNA topology can have different outcomes. Overwound DNA ahead of the transcription bubble could hinder unwinding by a converging replication fork. Or, the formation of non B-DNA structures in the negatively supercoiled region behind RNA polymerases, can lead to genomic instability.

Alternate DNA structures are often facilitated by the negative supercoiling generated behind the RNA polymerase during transcription. They can cause problems during a subsequent S-phase as they must be removed in order for the replication fork to continue unhindered (Mischo et al., 2011; Tuduri et al., 2009). These structures include Z-DNA, G-quadruplex's, or R-loops (Nordheim et al., 1982; Peck et al., 1982; Richardson, 1975). Z-DNA is known to specifically form near transcription start sites (TSS) in eukaryotes, and is stabilised by the negative supercoiling behind the transcription bubble (Schroth et al., 1992; Wittig et al., 1989). G-quadruplex sequences are abundant through the genome, and can be observed to occur *in vivo* at specific sites, such as upstream of promoter regions of genes (Huppert and Balasubramanian, 2007; Lam et al., 2013).

DNA:RNA hybrids or R-loops, also occur abundantly behind RNA polymerases. They occur in regions of negative supercoiling, where the new RNA formed will base pair with its complimentary DNA strand (Drolet et al., 1994). R-loops can cause problems in subsequent DNA

replication by causing slowing and stalling of replication forks, which could lead to DNA breaks (Gan et al., 2011; Tuduri et al., 2009). The interference of transcription and R-loops during DNA replication has also been implicated in causing genomic instability via transcription-associated recombination (TAR) (Tuduri et al., 2009; Wellinger et al., 2006). TAR occurs because non-template single stranded DNA (ssDNA) can be more accessible to recombination machinery, or more easily damaged, resulting in repair by homologous recombination (HR) (Gottipati and Helleday, 2009). For DNA:RNA hybrids to be removed and processed, a number of different proteins are required. RNase H enzymes are well known to degrade the hybrids, the Senataxin helicase is implicated in preventing genomic instability and for facilitating replication fork progression through genes, and various replication fork proteins have also been implicated in this process such as BRCA, FACT and Pif1 (Alzu et al., 2012; Bhatia et al., 2017; Mischo et al., 2011). Clearly, R-loops are problematic for the replication fork, and therefore it is important that topoisomerases act to remove the negative supercoiling behind the polymerase, inhibiting the formation of R-loops (Brill and Sternglanz, 1988; Masse and Drolet, 1999). Therefore R-loops are a cause of replication fork problems during S–phase, however, they are also a consequence of replication–transcription collisions (Helmrich et al., 2011).

1.7.2 Replication -Transcription Collisions

Replication-transcription collisions have been implicated in causing genomic instability due to the interference between the two processes (Azvolinsky et al., 2009; Dutta et al., 2011; Helmrich et al., 2011; Merrikh et al., 2011; Prado and Aguilera, 2005). Collisions can be categorised as either co-directional or head-on. Co-directional collisions occur when the direction of DNA transcription and the direction of DNA replication are the same as one another. Head-on collisions occur where the direction of DNA transcription is towards that of the direction of DNA replication (French, 1992).

In *Escherichia. coli,* collisions occur by physical interaction of the replication and transcription machineries (Boubakri et al., 2010; Mirkin and Mirkin, 2005). In bacteria there is a large preference for, in particular, essential genes to be on the leading strand of the DNA, causing mostly co-directional collisions (Guy and Roten, 2004). In *E.coli* 70% of essential genes are co-directionally transcribed and in *Bacillus. subtilis* 90% are co transcribed (Blattner et al., 1997; Kunst et al., 1997). Suggesting that it is much more favourable for collisions to be in a co-directional orientation than head-on. However, this preference is much less obvious in

eukaryotes, for example, budding yeast appears to have no bias for transcription to be in the same direction as replication (McGuffee et al., 2013), which could be due to the directionality of the helicases. In *E. coli*, DnaB, the replicative helicase, processes with a 5' to 3' directionality, associating with the lagging strand (Kornberg and Baker, 1992). However the CMG helicase in eukaryotes, processes with a 3' to 5' directionality, associating with the leading strand (Fu et al., 2011). This leads to differences in collisions; in *E. coli* the helicase and the RNA polymerase will converge together on the same strand, presumably increasing the negative consequences of a head-on collision. Whereas, in eukaryotes they will converge upon one another on opposite strands, and it can be envisioned that the two could process past one another in a manner analogous to the termination of DNA replication (Figure 1.12A).

Although eukaryotes don't have a preference for genes on one DNA strand, there is regulation by spatial and temporal separation of transcription and replication, in order to prevent genomic instability due to collisions (Wei et al., 1998). There is a trend towards genes that replicate early being transcribed later, and late replicating genes transcribed early (Meryet-Figuiere et al., 2014), indicating a preference against the collisions of replication and transcription. As well as temporal regulation, highly transcribed genes such as rDNA have been seen to be spatially separated during replication (Dimitrova, 2011; Smirnov et al., 2014). Topologically associating domains (TAD) are mostly entirely early or late replicating, indicating a spatial separation from DNA replication due to DNA domain (Pope et al., 2014).

1.8 Transcription and Replication Fork rotation

Head-on replication-transcription collisions lead to high levels of topological stress between the replication fork and the RNA polymerase (Figure 1.12A). Replication fork stalling occurs at the most highly transcribed genes, where there would presumably be the highest amount of topological stress (Azvolinsky et al., 2009). As the replication fork progresses, positive supercoiling forms ahead of it, and this will be prevented from diffusing away from the replication fork by the RNA polymerase. It will also converge on the positive supercoiling forming in front of the transcription bubble as part of the twin supercoiled domain. This situation is similar to that of the termination of DNA replication, and therefore may lead to similar consequences. As the machineries converge, topoisomerases may be excluded from in-between them, leading to fork rotation being utilised as a method for relaxing the high topological stress, and allowing the continuation of replication (Figure 1.12A). Another solution to this problem

would be for the RNA polymerase to fall off the DNA and therefore allow for replication to continue unhindered. However, this is not always possible, for example in the case of very long transcripts that will take longer than one cell cycle to transcribe (Helmrich et al., 2011).

In eukaryotes, co-directional collisions also lead to fork stalling and genomic instability. In this situation, the RNA polymerase itself may act as a barrier to the diffusion of supercoiling ahead of the replication fork. For example, when an RNA polymerase is paused, it would act in a similar manner to a stable protein-DNA block, and therefore have the same consequence of fork rotation occurring to diffuse the supercoiling stress (Figure 1.12B) (Keszthelyi et al., 2016). Again a barrier such as an R-loop formed behind a transcription bubble could lead to fork rotation by acting as a barrier to the diffusion of supercoils (Tuduri et al., 2009).

The environment around genes is also important when considering if fork rotation is likely to occur. The frequency of nucleosomes, as well as potentially the prevalence of topologically associating domains (TADs), will prevent the spreading of topological stress (Salceda et al., 2006). Gene gating is the physical tethering of transcribed genes to the nuclear envelope, preventing the diffusion of supercoils due to the prevention of rotation of the DNA, therefore giving a topologically confined situation (Blobel, 1985; Postow et al., 2001; Postow et al., 2004). When gene gating occurs during DNA replication, the local build-up of super-helical stress could generate a barrier for the replication fork and potentially lead to fork stalling. It has been suggested that mediation of replication through these situations is facilitated by checkpoint proteins, which are able to release the DNA from the nuclear pore (Bermejo et al., 2011). In this study, the authors further suggested that in checkpoint deficient cells, fork reversal occurs due to the energetics of the topologically constrained DNA forcing the fork backwards (Bermejo et al., 2011). However, in other areas of constrained topological tension, such as at the termination of DNA replication, the pathway of fork rotation is used to diffuse the built up supercoiling, and potentially this could be utilised in the situation of gated genes.

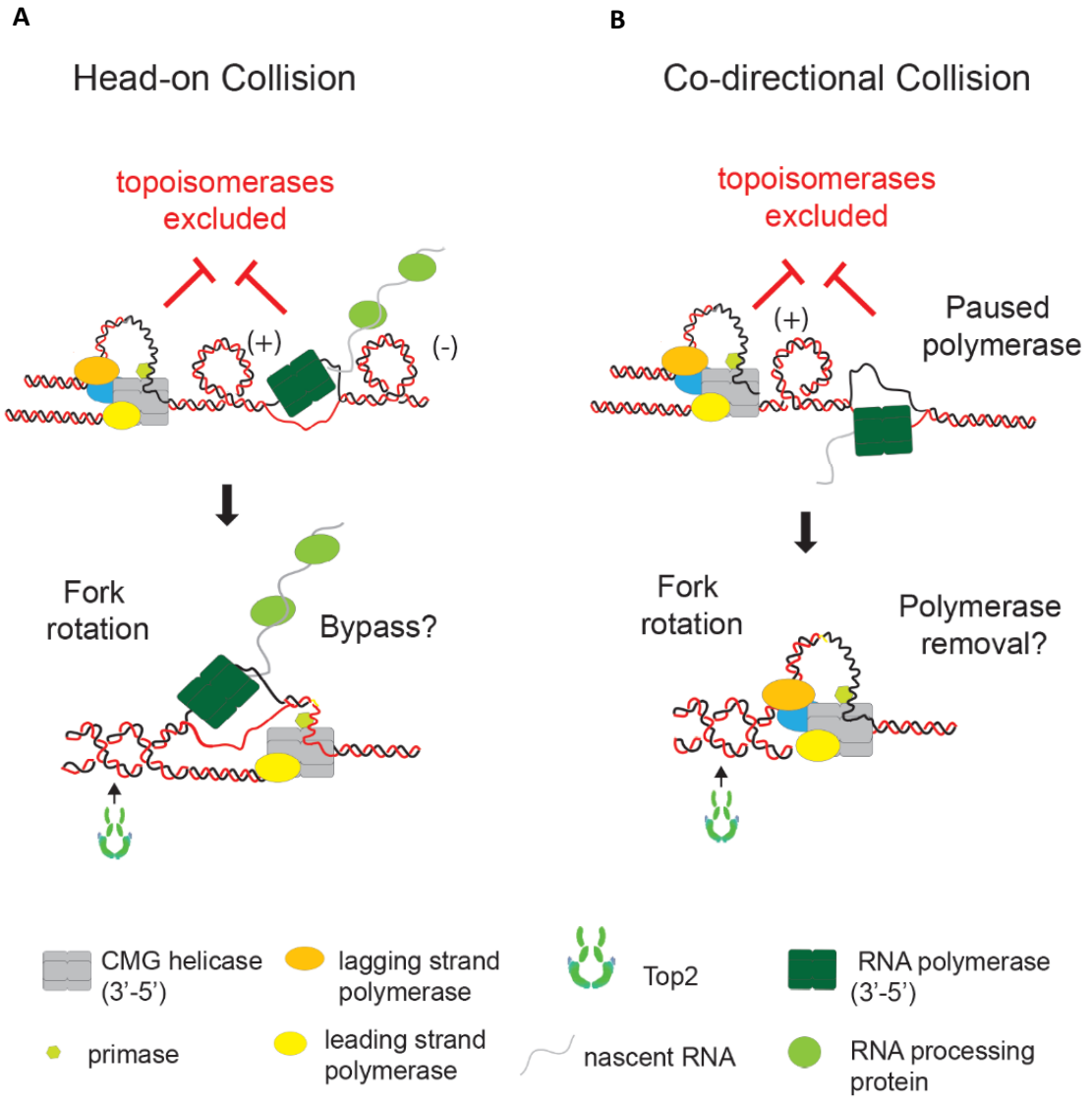


Figure 1.12 Two hypothetical scenarios at eukaryotic replication-transcription collisions

Modified from Keszthelyi et al. (2016). **A**. At a head-on replication-transcription collision, high levels of positive (+) superhelical stress will build up ahead of the fork. Negative (-) superhelical stress will arise behind the transcription bubble. A sterical exclusion of topoisomerases will occur, which will lead to fork rotation occurring, to relax the positive helical stress. In eukaryotes the replication and transcription machineries could potentially bypass one another in a manner analogous to the termination of replication. **B**. At a co-directional collision due to a paused polymerase, high levels of positive (+) superhelical stress will build up ahead of the fork. A sterical exclusion of topoisomerases will occur, which will lead to fork rotation occurring, to relax the positive helical stress. The paused polymerase could potentially be removed from the DNA to allow passage for the replication fork.

1.8.1 rDNA array during DNA replication

The rDNA array is a particularly difficult region to replicate through; in *S. cerevisiae* 150 – 200 repeats of a 9 kb rDNA unit make up the rDNA array. Each unit comprises a 35S region transcribed by RNA polymerase I, and a 5S region transcribed by RNA polymerase III (Figure 1.13) (Petes, 1979). Each unit also has two non-transcribed regions, IGS1 and IGS2, which contain an origin of replication, a replication fork barrier (RFB), a non-coding promoter and elements of *HOT1* (Figure 1.13) (Kobayashi, 2014). The RFB is placed preventing replication in the opposite direction of RNA polymerase I transcription (Figure 1.13) (Brewer and Fangman, 1988; Linskens and Huberman, 1988). Unidirectional replication is only blocked at the RFB sites when fork blocking protein 1 (Fob1) binds there, and Tof1/Csm3 are present (Bairwa et al., 2010; Mohanty et al., 2006). The role of Tof1/Csm3 at these barriers, similarly to other protein-DNA replication fork pause sites, is separate from that of Mrc1 and checkpoint proteins. Tof1/Csm3 action is counteracted by the helicase Rrm3, which releases the forks from the block (Ivessa et al., 2000; Mohanty et al., 2006). The rDNA is an intrinsically unstable region of DNA, prone to high levels of recombination, leading to highly changeable numbers of repeats. This is due to the stalling of replication forks and the repetitive sequence, which subsequently leads to high levels of double strand breaks arising and recombination occurring (Kobayashi and Horiuchi, 1996; Mohanty et al., 2006). The levels of Pol1 transcription are also important for recombination events, with studies indicating that rDNA copy number is generally stabilised by Pol1 transcription (Kobayashi et al., 1998).

The rDNA region is also unique in its compaction in *S. cerevisiae*. In late anaphase arrested cells, DNA in general appears decondensed, however the rDNA remains compacted (Guacci et al., 1994). Pre-anaphase it appears to be compacted by a combination of the SMC proteins, condensin and cohesin (Guacci et al., 1997; Lavoie et al., 2004; Schalbetter et al., 2017). However, post anaphase it undergoes a further level of longitudinal compaction by the condensin complex (D'Amours et al., 2004; Sullivan et al., 2004). It was suggested by Sullivan et al. (2004) that the reason for this condensin focused action, is that the rDNA is a difficult region for segregation due to the high levels of sister chromatid intertwines and extreme length of the region. Why then is sister chromatid intertwining so abundant at the rDNA?

It is possible that this region is a unique barrier to the DNA replication fork, leading to higher than usual levels of DNA replication fork rotation. Indeed, each rDNA repeat contains multiple Rrm3 dependent replication fork pause sites, including the replication fork barrier (Ivessa et al., 2000). Stable protein-DNA replication fork pause sites are known sites of replication fork

rotation, and it is possible that the pause sites in the rDNA array, including the RFB, could similarly lead to fork rotation. As a replication fork converges with a RFB site, helical stress in front of the fork may be prevented from diffusing away, leading to high levels of topological stress. Furthermore, not only does the rDNA region represent a substantial static block to the replication fork, but also a non-static transcriptional block. The rDNA is highly transcribed throughout the replication cycle, and Top1 is specifically recruited to the rDNA region by an interaction with Tof2, indicating that this is a region of constitutive topological stress (Krawczyk et al., 2014). This constitutive stress may be exacerbated by replication through the array, and therefore as well as requiring Top1 activity may require replication fork rotation in order to be resolved. Moreover, high levels of transcription through the array gives a high potential for co-directional replication-transcription collisions to occur. As outlined in the previous section, co-directional collisions are areas likely to lead to a build-up of topological stress and therefore also, replication fork rotation.

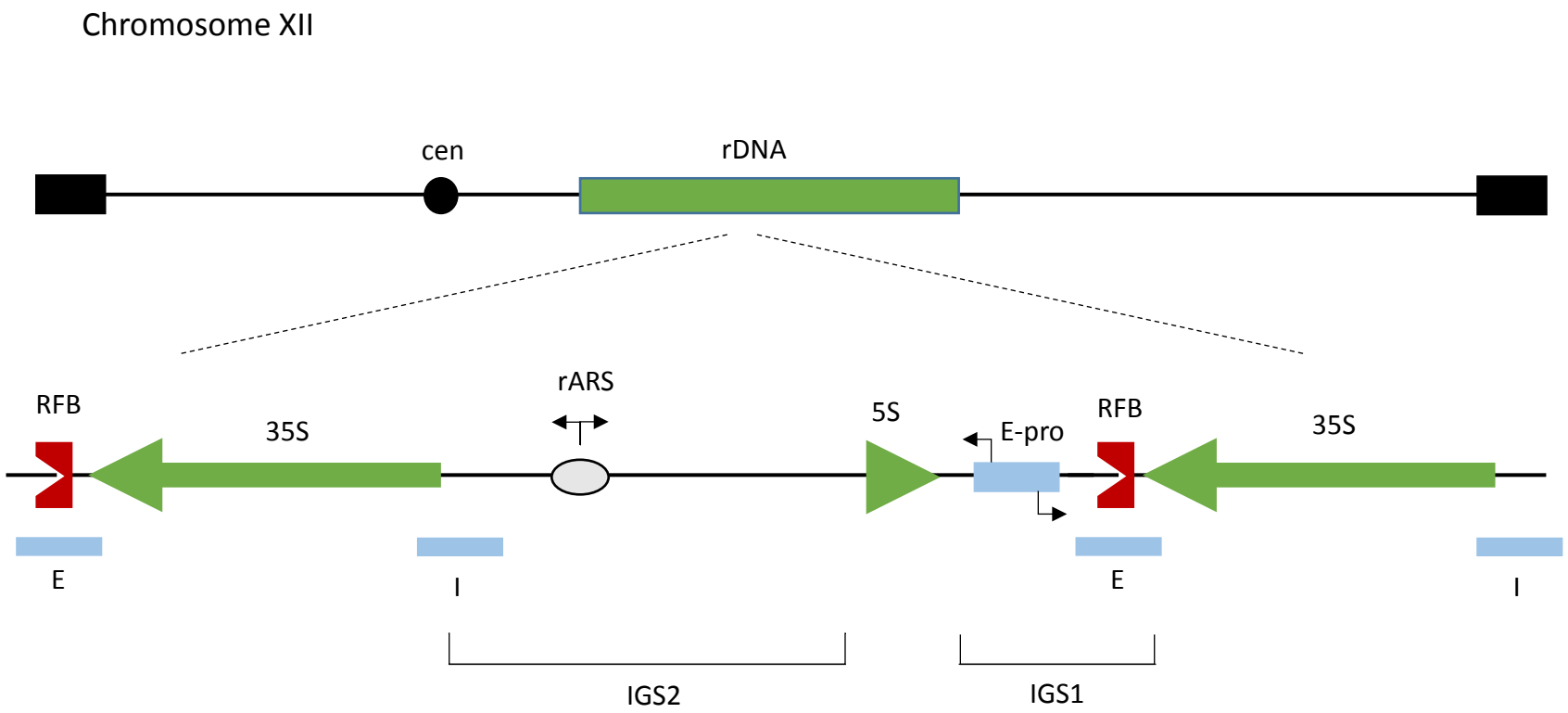


Figure 1.13 Schematic of *S. cerevisiae* chromosome XII rDNA region

RNA pol I 35S elements are flanked by E and I elements of *HOT1*, the RFB is situated in the non-transcribed IGS1 region at the end of the 35S element preventing replication in the opposite direction of transcription. The non-coding promoter E-pro is also situated in the IGS1 region. The 5S element transcribes in a directionality towards the 35S element. The origin is situated in the IGS2 region.

1.9 SMC PROTEINS

Static and non-static barriers to DNA replication have been discussed above, however little is known about the DNA topological effects of a group of non-static proteins from the structural maintenance of chromosomes family (SMC). SMC proteins are essential, highly conserved and extremely important factors in chromosome compaction, homologous recombination, and transcription (Hirano, 2012; Soppa, 2001).

The eukaryotic SMC family comprises three complexes: condensin, cohesin and SMC5/6. Each of which have two 50 nm SMC coiled-coil arms, joined at a hinge region in the middle, to give a characteristic 'V' shape. The ATPase Walker A (N-terminus) and Walker B (C-terminus) motifs are then joined together by the binding of a kleisin-type protein at the ATPase head domains (Figure 1.14) (Schleiffer et al., 2003). This combination of proteins allows each of the complexes to form a characteristic ring like shape that has been proposed to encapsulate DNA to allow for each of their individual functions (for reviews see (Diaz and Pecinka, 2018; Hirano, 2016)). The complexes also recruit other proteins, predominantly via interaction with the kleisin protein. These proteins are members of the HAWK (Heat proteins associated with kleisins) or KITE (kleisin interacting winged-helix tandem elements) families (Wells et al., 2017).

Here the focus will be on the cohesin complex, a complex first known as essential for sister chromatid cohesion (Guacci et al., 1997; Michaelis et al., 1997), and more recently has been found to be a key player in the organisation and compaction of chromosomes (Fudenberg et al., 2016; Hadjur et al., 2009; Schalbetter et al., 2017). Although it carries out essential cellular functions, there is still a wide ranging debate as to how it carries out these functions, how it is regulated, and the extent of its roles on DNA. Some of these aspects will be discussed below.

Figure 1.14 A composite structure of the yeast cohesin complex

Adapted from Uhlmann (2016), a composite model incorporating crystal structures from various sources. The cohesin complex comprises a heterodimer of Smc1 and Smc3, which join at a hinge region and at their ATPase head domains. Scc1 and Scc3 are also situated at, and bridge between, the heads.

1.9.1 Structure of the Cohesin Complex

Cohesin in yeast, comprises the characteristic SMC coiled-coils, Smc1 and Smc3, with two other proteins, the kinesin subunit known as Scc1/Mcd1 (Rad21) which contains the separase (Esp1) cleavage sites, and a HEAT repeat-like containing subunit (HAWK) Scc3 (SA1/SA2) (Figure 1.14) (Guacci et al., 1997; Michaelis et al., 1997; Neuwald and Hirano, 2000; Uhlmann et al., 1999). Smc1 and Smc3 interact via their hinge domains and head domains; the globular N domain containing the ATP binding site and C domain, come together to form the head domains that contain the ABC family ATPase activity (Lowe et al., 2001). Scc1/Mcd1 finishes the tripartite ring by binding the nucleotide binding domain of Smc1, via the Scc1/Mcd1 C terminus winged helix domain, and Smc3 via a four helix bundle just above the head domain at its N terminus (Gligoris et al., 2014; Haering et al., 2004).

Two other proteins, Pds5 and Scc2, which both interact with Scc1/Mdc1, have also been implicated as being essential cohesin factors. They are both of the HEAT repeat family of proteins (recently named HAWKs), along with Scc3 and two essential condensin proteins, Ycg1 and Ycs4 (Kikuchi et al., 2016; Lee et al., 2016; Wells et al., 2017). Pds5 and Scc2 appear to compete for binding to Scc1/Mdc1 (Kikuchi et al., 2016; Murayama and Uhlmann, 2015; Petela et al., 2018).

1.10 Cohesin Loading

In budding yeast, loading of Cohesin starts at the end of G1 phase and increases through S phase, until cleavage at the metaphase/anaphase transition (Hu et al., 2015). Loading is thought to be mediated by the Scc2-Scc4 complex, and thought to topologically entrap DNA (Ciosk et al., 2000; Haering et al., 2008). In a reconstituted *Schizosaccharomyces pombe* system, Murayama and Uhlmann (2014) showed that Scc2 (Mis4) bound double stranded DNA (dsDNA), and was required for the topological entrapment of cohesin onto DNA in an ATP hydrolysis dependent manner (Murayama and Uhlmann, 2014). It has recently been further argued that Scc2 stimulated ATP hydrolysis is also required for the translocation of cohesin along DNA (Petela et al., 2018; Rhodes et al., 2017).

In order for cohesin to topologically entrap DNA, the ring must open, although exactly how this occurs is still under debate. It was first proposed to load via the hinge domain, by Gruber et al.

(2006), via a series of crosslinking experiments hindering the hinge domain from opening, which prevented cohesin from loading DNA. However, it has also been proposed that the interface suggested to release topologically entrapped DNA from the ring, the Smc3-Scc1 interface, is also the route for cohesin loading onto DNA (Figure 1.15) (Chan et al., 2012; Murayama and Uhlmann, 2015). In this model the protein Wapl (wings apart-like protein homologue), in complex with Pds5, disrupts the Scc1-Smc3 interface to allow entry and exit of cohesin from the DNA, both in an ATP dependent manner. It has further been recently put forward that a subset of cohesin, namely that required for the function of loop extrusion, may not topologically entrap the DNA (Srinivasan et al., 2018), although this remains to be seen.

Cohesin is mostly found to accumulate at the centromeres and at certain places along chromosome arms (Blat and Kleckner, 1999; Hu et al., 2011; Hu et al., 2015; Tanaka et al., 1999). It appears that cohesin is loaded at centromeres by the Scc2-Scc4 complex, which preferentially binds here, and also at highly transcribed genes (Lengronne et al., 2004). Following this it is thought that either cohesin can translocate down chromosome arms from its loading at the centromeres, or it loads with lower affinity without the aid of Scc2-Scc4 along the arms (Hinshaw et al., 2015; Hu et al., 2011).

Although the exact mechanism by how the Scc2-Scc4 complex is recruited is under debate, it appears that certain proteins recruit it to different parts of the chromosome. At the centromeres, localisation of Scc2 appears to rely on the Ctf19 complex, namely the Chl4 protein (Fernius et al., 2013). Whereas certain chromatin remodelling complexes, such as the RSC complex (Remodelling the Structure of Chromosome) and Irc5, have been implicated in cohesin/ Scc2-Scc4 loading along chromosome arms in budding yeast (Litwin et al., 2017; Lopez-Serra et al., 2014; Natsume et al., 2013).

Other important ongoing questions surround whether or not cohesin loads onto single or double stranded DNA, and following this, whether it acts as a monomer or as a dimer molecule. There has been some evidence for oligomerisation of cohesin (reviewed in Skibbens (2016)) however the more favoured model appears to be a single ring. ssDNA capture has been previously proposed for SMC5/6 binding and appears to have some role in the binding of other SMC proteins (Alt et al., 2017; Roy and D'Amours, 2011; Roy et al., 2011). A further study in fission yeast by Murayama et al. (2018) using a reconstituted system, found that a single cohesin ring appears to initially bind double stranded DNA via Scc2-Scc4 using ATP, followed by a second capture by the same ring of single stranded DNA, again dependent on Scc2-Scc4 and ATP (Murayama et al., 2018).

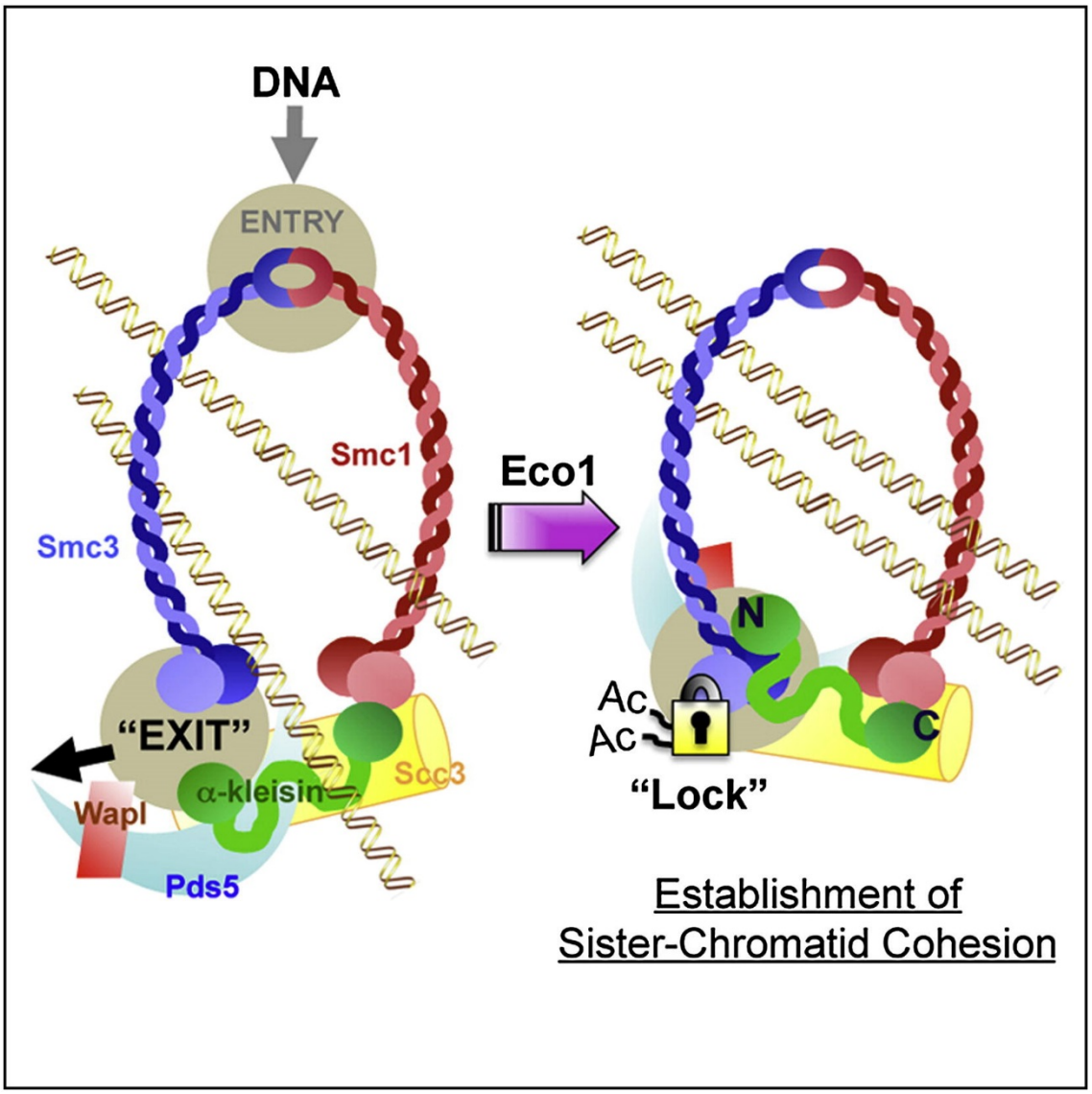


Figure 1.15 Proposed DNA entry and exit gates and establishment of sister chromatid cohesion

Taken from Chan et al (2012). DNA has been proposed to enter via the cohesin hinge domain and exit via the Smc3-Smc1 interface with the help of Wapl/Pds5 or Separase. Eco1 acetylation of Smc3 allows for sister chromatid cohesion and prevents removal of cohesin from DNA by Wapl.

1.11 SISTER CHROMATID COHESION

The first known role of cohesin was that of sister chromatid cohesion. This was suggested by Michaelis et al. (1997) when looking at a genetic screen for mutants showing precocious separation of sister chromatids during mitosis in *S.cerevisiae,* and also Guacci et al. (1997) who used a temperature sensitive mutant of the Scc1/Mcd1 subunit of cohesin and found it is required for sister chromatid cohesion (Guacci et al., 1997; Michaelis et al., 1997). These initial observations were quickly backed up by subsequent studies in *Xenopus* (Losada et al., 1998). These studies identified the cohesin subunits as essential, and the loss of a subunit causing aneuploidy and cell death in mitosis.

1.11.1 Establishing sister chromatid cohesion

When cohesin is loaded onto DNA it has a high turnover, as it can be quickly removed by the actions of Wapl (Rad61) and Pds5 (Figrue 1.15) (Chan et al., 2012; Kueng et al., 2006; Sutani et al., 2009). However, during S-phase the protein Eco1 (Esco1/Esco2 in humans) acetylates Smc3 on K112 and K113 residues (human K105, K106), which appears to counteract the effect of Wapl, allowing for stable and long lived binding of cohesin on DNA (Figure 1.15) (Ivanov et al., 2002; Rolef Ben-Shahar et al., 2008; Rowland et al., 2009; Unal et al., 2008; Zhang et al., 2008). Furthermore a deletion of *ECO1* in yeast leads to severe segregation defects and lethality which can be rescued by a further deletion of *RAD61,* indicating a main role for Eco1 in counteracting the effect of Wapl (Rowland et al., 2009; Skibbens et al., 1999).

After loading, cohesin translocates along chromosome arms, for example in budding yeast it is highly abundant in-between convergent genes (Hu et al., 2011; Jeppsson et al., 2014; Lengronne et al., 2004). Although, how cohesin translocates, in the context of chromatin and nucleosomes is mostly unknown. From single molecule studies, cohesin does not appear to be able to translocate past nucleosomes easily, although some diffusion was possible, suggesting it requires a nucleosome remodelling complex to help re-position it (Stigler et al., 2016). Various remodellers have been implicated in this process, the RSC chromatin remodelling complex and the Scc2-Scc4 loader (Lopez-Serra et al., 2014). However, with respect to chromatin hindrance more work is needed to fully understand this.

An interesting question for cohesion establishment is the extent of the interaction of cohesin with the replication fork. As cohesin is mainly proposed to be loaded at the centromeres before

translocating along chromosomes, how does the replication fork deal with cohesin in front of it? And how does cohesion become properly established? Cohesion behind the replication fork may become established due to the cohesion establishment protein, Eco1, being closely linked to the replication fork, supposedly via an interaction with PCNA (Lengronne et al., 2006; Moldovan et al., 2006). It has also been proposed that this interaction between Eco1 and PCNA is what is important for the correct establishment of cohesion, without which, the acetylation of cohesin is not sufficient (Song et al., 2012). It was once thought that the fork could simply processes through the cohesin ring, leading to establishment of cohesion behind (Lengronne et al., 2006), however it is hard to imagine this as a physically viable option due to the number of replisome components and subsequent size (Stigler et al., 2016). So does an established cohesin ring lead to problems with replication fork progression? Can the fork push the cohesin ring in front of it, as it is proposed for transcription? Or do the accessory factors Wapl and Pds5, or a currently unknown factor associated with the fork, lead to removal of the cohesin ring from in front of the fork?

Recently, proposals have stipulated an idea that cohesin is removed when a replication fork encounters it due to the close association of cohesin with its removal protein Wapl (Terret et al., 2009). It can then be quickly loaded behind the fork and acetylated by Eco1. Alternatively, a pool of cohesin molecules near the replication fork could be used to establish cohesion behind the fork, while being replenished by cohesin removal in front of the fork. A study by Frattini et al. (2017) established a role for Wapl in removing cohesin from in front of forks with the aid of checkpoint mediated cohesin ubiquitylation, followed by Eco1 dependent entrapment of sister chromatids. However, this leads to questions as whether there is a pathway that is independent of stalled replication forks and checkpoint activation. These data indicate that the fork is hindered and even stalls, due to cohesin establishment in front of the replication fork (Frattini et al., 2017). This promotes the idea that cohesin could lead to an increase in topological stress around replication forks, which could be through preventing the dissipation of supercoiling, or by being a physical barrier to topoisomerase action in front of the replication fork. It has been proposed that cohesin protects sister chromatid intertwines behind the replication fork, possibly by preventing topoisomerase action here (Farcas et al., 2011; Jeppsson et al., 2014), therefore could cohesin be performing a similar role in front of the replication fork leading to fork rotation occurring and sister chromatid intertwines forming?

1.11.2 Cohesin accessory factors

Although there appears to be an established role for Eco1, Wapl (Rad61) and Pds5 there is still a lot of controversy about the exact ways these proteins function in respect to cohesin. Wapl and Pds5 are said to work in complex to remove cohesin from DNA by opening up the cohesin exit gate (the interface between Scc1 and Smc3 (Huis in 't Veld et al., 2014)). They are known to form a complex together and associate with cohesin on chromosomes (Sutani et al., 2009). However when not in complex with one another, Pds5 appears to have sister cohesion promoting effect by binding cohesin, promoting Eco1 acetylation and preventing de-acetylation (Chan et al., 2013) and in human cells by recruiting Sororin, the Wapl antagonising factor (Nishiyama et al., 2010).

Wapl also appears to have contradictory effects on sister chromatid cohesion. In human cells, where cohesin is extensively loaded in G1 phase, the depletion of Wapl leads to re-localisation of cohesin along chromosome arms, increased cohesin residence time on DNA, and the "vermicelli" condensed structured chromosomes (Tedeschi et al., 2013). In budding yeast, it is established that Wapl promotes the release of cohesin from DNA in the absence of Smc3 acetylation (Kueng et al., 2006), however, there have been a number of contradictory findings over the years. In a genetic screen, a range of replication fork proteins were implicated in having sister chromatid cohesion defects when deleted in *S. cerevisiae,* including Rad61 (Wapl) (Warren et al., 2004). This was again seen by Rowland et al. (2009) who used a GFP system, bound to the *URA3* locus in budding yeast, and assayed sister chromatid cohesion by counting the number of times two GFP dots appeared as oppose to one, indicating sister chromatid separation. Using this method they surprisingly found that *rad61Δ* cells increased the number of loci that were not cohesed, and that further deletion of *ECO1* increased this phenotype. This has also been seen by Kulemzina et al. (2012) and by Guacci and Koshland (2012) who tested sister chromatid cohesion by using a similar mechanism of LacO-GFP either at *LYS4* or *TRP1* loci to compare chromosome arm to centromere proximal cohesion. They again found that *wpl1Δ* cells, arrested in nocodazole, had a chromosome segregation defect at both loci, the opposite of what would be expected if its sole role is to remove cohesin. However, the segregation defect in *eco1-ts* cells and *eco1Δwpl1Δ* was much more severe, indicating that *RAD61* deletion can rescue the lethality of *eco1*Δ cells, but it doesn't lead to the rescue of loss of cohesion through loss of acetylation (Rowland et al., 2009).

In the screen in *S. cerevisiae* by Warren et al. (2004), as well as looking at Wapl, they found sister chromatid cohesion defects in *xrs2Δ, kar3Δ, srs2Δ, mrc1Δ, rrm3Δ, tof1Δ, sgs1Δ, csm3Δ*, *rad27Δ*

and *chl4Δ* cells. It had also been previously reported that deletion of *CTF18-RFC*, *CTF4* and *CHL1* lead to defects in sister chromatid cohesion (Hanna et al., 2001; Mayer et al., 2001; Skibbens, 2004). Several of these genes encode proteins that are replication fork components, Tof1/Csm3 are known to promote replication fork stability and complex with Mrc1 which also promotes fork stability and is a major checkpoint player (Pardo et al., 2017). Ctf18-RFC appears to load the PCNA clamp, and Ctf4 links the CMG helicase to polymerase α (Bermudez et al., 2003; Simon et al., 2014). How they contribute to sister chromatid cohesion however is largely unknown, although their involvement implicates the replication fork as an important factor in cohesion establishment.

1.11.3 Cohesion anti-establishment

Cohesin appears to be removed from DNA by two separate methods; the role of Wapl which can disrupt the Scc1-Smc3 interface when it is not acetylated, or the pathway of separase cleavage of Scc1 which breaks apart the tripartite ring. In yeast, the main pathway of removal is via cleavage by the protease, separase (Esp1), activating the onset of anaphase, which raises the question of how important Wapl (Rad61) really is for cohesin removal in this organism (Uhlmann et al., 2000). However, in human cells the two pathways work together to remove cohesin from particular areas at specific time points. During prophase and prometaphase, Wapl is able to remove cohesin in a chromosome arm specific manner, to allow for de-catenation (Haarhuis et al., 2013; Kueng et al., 2006; Waizenegger et al., 2000). Leaving a subset of cohesin protected by Sgo1 bound to centromeres, which is then cleaved by separase, triggering anaphase (Kitajima et al., 2004). This two-step mechanism appears to promote the resolution of sister chromatids from one another to prevent a delay in anaphase onset, where presumably with the less complexity of the yeast genome the two-step mechanism is redundant.

1.12 Cohesin, Loop extrusion and Transcription

It is fast emerging that cohesin is an extremely important factor in the 3D interphase organisation of the genome, and that this arrangement allows for the correct regulation of transcriptional activation (Merkenschlager and Nora, 2016; Seitan et al., 2013). Conversely, it appears transcription drives the interphase organisation of the genome by its re-localisation of cohesin away from sites of loading (Davidson et al., 2016).

In recent years the emergence of TADs (topologically associating domains) and compartments, made out of loops, defined by the Hi-C technique, has revolutionised the SMC field (Dixon et al., 2012; Fudenberg et al., 2016; Lieberman-Aiden et al., 2009; Rao et al., 2014). Cohesin's role and mechanism in genome organisation is still debated, however, it appears that cohesin has an important role during interphase for the formation of TADs formed due to peaks/loops (Schwarzer et al., 2017). Loops appear to be defined by CCCTC-binding factor (CTCF) sites, where 90% of CTCF sites at loop anchors are in a convergent orientation (Rao et al., 2014). CTCF is also known to strongly interact with cohesin on chromosomes (Parelho et al., 2008; Rubio et al., 2008; Stedman et al., 2008). Therefore, this gives the possibility that cohesin tracks along the DNA, extruding in loops of a defined size due to a bidirectional cohesin block at CTCF sites (Fudenberg et al., 2016). What is currently unknown, is mechanistically how cohesin performs this function. The molecular mechanism by how condensin performs loop extrusion was recently analysed by microscopy of purified condensin acting on DNA, showing an ATP-dependent one sided extrusion mechanism (Ganji et al., 2018). It has been suggested that cohesin has a similar method for loop extrusion, however, since the ATPase activity appears to be necessary for loading of the complex it is difficult to separate these mechanisms (Hu et al., 2011; Vian et al., 2018).

There is evidence for the passive movement of cohesin, in studies in budding yeast it has clearly been seen that cohesin localises to areas of converging genes, and was postulated that this was due to pushing of the cohesin complex by the RNA polymerase (Glynn et al., 2004; Hu et al., 2011; Jeppsson et al., 2014; Lengronne et al., 2004). It could be that in yeast, an organism where CTCF does not exist, the positioning of cohesin at sites of convergent transcription contributes to the distribution of TAD boundaries (Mizuguchi et al., 2014; Sun et al., 2015; Wendt et al., 2008). Recently in human cells, the same transcriptional positioning has been seen when CCCTC-binding factor (CTCF) sites, and Wapl, were removed (Busslinger et al., 2017). Busslinger et al. (2017) showed that a major determinant of cohesin localisation was through transcription; at a convergent gene pair the distribution of cohesin between the pair was asymmetric, depending on the strength of each gene (Busslinger et al., 2017). This recent study adds to the growing information about the roles of both transcription and Wapl, in positioning cohesin. However, how the transcription dependent movement mechanistically works is still a mystery, some think that the supercoiling in front of the polymerase pushes cohesin to the ends of genes (Bjorkegren and Baranello, 2018). However, if this occurs, then does the supercoiling in front of the replication fork push cohesin? Maybe cohesion associated replication fork proteins can regulate this.

1.13 THESIS AIMS

In this work I have aimed to study topological stress during DNA replication and the impact on the replication fork. I have looked at distinct contexts during replication that lead to changes in the topology of the DNA. I use two main assays to analyse this; a plasmid replication fork rotation assay and genome wide DNA damage assay analysing damage due to topological stress by γH2A ChIP-seq.

My main aims were to look at the incidence of transcription during DNA replication and its impact on replication fork rotation. I have also analysed the impact of the cohesin complex on replication fork rotation and the association between cohesin and topological stress generated by transcription. I further aimed to analyse the specific contexts in which topological stress leads to endogenous DNA damage during DNA replication, and how cohesin generated replication stress impacts this.

2 MATERIALS AND METHODS

2.1 MEDIA

2.1.1 YP rich media

1% w/v Bacto-yeast extract (Melford, Y1333)

2% w/v Bacto-peptone (Melford, P1328)

For Plates 2% w/v bacto-agar was added

pH 5.5

Carbon sources (2% w/v glucose, 2% w/v raffinose (Sigma-Aldrich, R0250), 2% w/v galalactose (Sigma-Aldrich, G0625)) were filter sterilised (Millipore) and added as indicated in section 2.6.7 of the material and methods after autoclave of the YP. 40 mg/l adenine sulphate was added to all liquid YP media when using strains without the *ADE2* gene.

2.1.2 Minimal Media Plates

1x YNB (Yeast Nitrogen Base, Melford, Y2004)

2% w/v Bacto-agar

2% w/v glucose

Supplemented with 0.004% w/v adenine sulphate, L-histidine, L-leucine, uracil and L-tryptophan. For selection of strains that have specific amino acid genes this amino acid was left out of the minimal media plates.

2.1.3 Rich Sporulation Media (RSM) Plates

0.25% w/v Bacto-yeast extract (Melford, Y1333)

1.5% w/v Potassium Acetate

0.1% w/v glucose

2% w/v Bacto-agar

80 mg/l adenine sulphate, uracil

40 mg/l L-histidine, L-leucine, L-lysine, L-tryptophan, L-methionine, L-arginine

16 mg/l L-tyrosine

200 mg/l L-phenylalanine

Amino acid mix was added after filter sterilization (Millipore) and after autoclave of media

2.1.4 Dropout media

1x YNB (Yeast Nitrogen Base)

1.92 g/l Yeast Synthetic Drop-out Medium Supplements without leucine or uracil (Sigma, Y1376, Y1501)

2% w/v glucose or 2% w/v raffinose

Filter sterilised (Millipore)

2.2 DRUGS USED IN THIS STUDY

Nourseothricin (NAT) (Jena Bioscience, AB-102L) 100mg/ml stock – 100 µg/ml final

Geneticin disulphite (G-418) (Melford, G0175) 100mg/ml stock - 200 µg/ml final

Hygromyocin B (HygB) (invitrogen ,1068701050) 45 mg/ml stock - 300 µg/ml final

Ampicillin sodium salt (Amp) (Sigma, A9518) 100mg/ml stock – 100 µg/ml final

Doxycycline (Dox) (Sigma, D9891) 50mg/ml stock – 50 µg/ml final

Alpha Factor (synthesised by GenScript) 5mg/ml stock – both 5 µg/ml and 10 µg/ml final

Nocodazole (Noco) (Sigma, M1404) 1.5mg/ml stock - 10 µg/ml final

2.3 STRAINS

All numbers refer to the Baxter lab strain database

Number	Strain name	Genotype
1	w303 wildtype	*Mat a ade2-1 his3-11 leu2-3 trp1-1 ura3-1 can1-100*
12	Degron background wildtype pRS316	*Mat a ade2-1 his3-11 leu2-3 trp1-1 ura3-1 can1-100 UBR1::GAL1-10-Ubiquitin-M-LacI fragment-Myc-UBR1 (HIS3) leu2-3::pCM244 (CMVp-tetR'-SSN6, LEU2) x3 pRS316*

13	*top2-td* pRS316	*Mat a ade2-1 his3-11 leu2-3 trp1-1 ura3-1 can1-100* *UBR1::GAL1-10-Ubiquitin-M-LacI* *fragment-Myc-UBR1 (HIS3)* *leu2-3::pCM244 (CMVp-tetR'-SSN6, LEU2) x3* *top2-td TOP2 5' upstream -100 to -1 replaced with kanMX-tTA (tetR-VP16)-tetO2 - Ub -DHFRts - Myc –linker)* *pRS316*
171	*scc1-73 top2-td* pRS316	*Mat a ade2-1 his3-11 leu2-3 trp1-1 ura3-1 can1-100* *UBR1::GAL1-10-Ubiquitin-M-LacI* *fragment-Myc-UBR1 (HIS3)* *leu2-3::pCM244 (CMVp-tetR'-SSN6, LEU2) x3* *top2-td TOP2 5' upstream -100 to -1 replaced with kanMX-tTA (tetR-VP16)-tetO2 - Ub -DHFRts - Myc -linker* *scc1-73* *pRS316*
211	*scc1-73*	*Mat a ade2-1 his3-11 leu2-3 ura3-1 can1-100* *scc1-73* *trp1Δ:: hphNT1*
275	*top2-4* pRS316	*Mat a his4-539 lys2-801 ura3-52* *top2-4* *pRS316*
365	*ndc10-1 top2-td*	*Mat a ade2-1 his3-11 leu2-3 trp1-1 ura3-1 can1-100* *UBR1::GAL1-10-Ubiquitin-M-LacI* *fragment-Myc-UBR1 (HIS3)* *leu2-3::pCM244 (CMVp-tetR'-SSN6, LEU2) x3* *top2-td TOP2 5' upstream -100 to -1 replaced with kanMX-tTA (tetR-VP16)-tetO2 - Ub -DHFRts - Myc -linker* *ndc10-1*
367	*ndc10-1*	*Mat a ade2-1 his3-11 leu2-3 trp1-1 ura3-1 can1-100* *UBR1::GAL1-10-Ubiquitin-M-LacI* *fragment-Myc-UBR1 (HIS3)* *leu2-3::pCM244 (CMVp-tetR'-SSN6, LEU2) x3* *ndc10-1*
479	*top2-4*	*Mat a his4-539 ura3-52*

		top2-4
484	*top2-4* 3x *tRNA* pRS316	*Mat a his4-539 lys2-801 ura3-52* *top2-4* *pRS316-3x tRNA*
1307	*scc1-73 top2-4*	*Mat a ade2-1 leu2-3 his4-539/his3-11 ura3-52/ura3-1* *can1-100* *top2-4* *scc1-73* *trp1Δ:: hphNT1*
1313	*top2-4 scc1-73 SEC53* *FMP32*-pRS315 (converging)	*Mat a ade2-1 his4-539/his3-11 lys2-801 ura3-52/ura3-1* *can1-100* *top2-4* *scc1-73* *trp1Δ::hphNT1* *leuΔ::natNT2* *SEC53 FMP32 pRS315 (converging)*
1314	*top2-4 scc1-73 SEC53* *FMP32*-pRS315 (unidirectional)	*Mat a ade2-1 his4-539/his3-11 lys2-801 ura3-52/ura3-1* *can1-100* *top2-4* *scc1-73* *trp1Δ::hphNT1* *leuΔ::natNT2* *SEC53 FMP32 pRS315 (unidirectional)*
1315	*top2-4 scc1-73* *mukB*-pRS315 (genes removed)	*Mat a ade2-1 his4-539/his3-11 lys2-801 ura3-52/ura3-1* *can1-100* *top2-4* *scc1-73* *trp1Δ::hphNT1* *leuΔ::natNT2* *mukB pRS315 (genes removed using bglII)*
1317	*top2-4 scc1-73* pRS316-3*tRNA*	*Mat a ade2-1 his4-539/his3-11 lys2-801 ura3-52/ura3-1* *can1-100* *top2-4* *scc1-73*

		trp1Δ::hphNT1 *leuΔ::natNT2* *pRS316-3tRNA*
1321	*top2-4 SEC53 FMP32*-pRS315 (unidirectional)	*Mat a ade2-1 his3-11 leu2-3,112 trp1-1 ura3-1 can1-100* *top2-4* *SEC53 FMP32 pRS315 (unidirectional)*
1322	*top2-4* pRS315 (genes removed)	*Mat a ade2-1 his3-11 leu2-3,112 trp1-1 ura3-1 can1-100* *top2-4* *SEC53 FMP32 pRS315 (genes removed using bglII)*
1323	*top2-4 mukB* pRS315 (genes removed)	*Mat a ade2-1 his3-11 leu2-3,112 trp1-1 ura3-1 can1-100* *top2-4* *mukB pRS315 (genes removed using bglII)*
1325	*top2-4 mukB* pRS315	*Mat a ade2-1 his3-11 leu2-3,112 trp1-1 ura3-1 can1-100* *top2-4* *mukB pRS315*
1326	*top2-4* pRS315	*Mat a ade2-1 his3-11 leu2-3,112 trp1-1 ura3-1 can1-100* *top2-4* *pRS315*
1396	*top2-4 rad61Δ SEC53 FMP32* pRS315 (converging)	*Mat a ade2-1 his3-11 leu2-3,112 trp1-1 ura3-1 can1-100* *top2-4* *rad61Δ::natNT2* *SEC53 FMP32 pRS315 (converging)*
1398	*top2-4 rad61Δ mukB* pRS315 (genes removed)	*Mat a ade2-1 his3-11 leu2-3,112 trp1-1 ura3-1 can1-100* *top2-4* *rad61Δ::natNT2* *mukB pRS315 (genes removed using bglII)*
1401	*top2-4 sgs1Δ top3Δ SEC53 FMP32* pRS315 (converging)	*Mat a ade2-1 ura3-1* *leu2-3::pCM244 (CMVp-tetR'-SSN6, LEU2) x3* *top2-4* *sgs1Δ::TRP1* *top3Δ::hphMX* *SEC53 FMP32 pRS315 (converging)*

1403	*top2-4 sgs1Δ top3Δ mukB* pRS315 (genes removed)	*Mat a ade2-1 ura3-1* *leu2-3::pCM244 (CMVp-tetR'-SSN6, LEU2) x3* *top2-4* *sgs1Δ::TRP1* *top3Δ::hphMX* *mukB pRS315 (genes removed using bglII)*
1412	*eco1-1*	*Mat a ade2-1 his3-11 leu2-3 trp1-1 ura3-1 can1-100* *UBR1::GAL1-10-Ubiquitin-M-LacI* *fragment-Myc-UBR1 (HIS3)* *leu2-3::pCM244 (CMVp-tetR'-SSN6, LEU2) x3* *eco1-1 (G211H)*
1436	*rad61Δ*	*Mat a ade2-1 his3-11 leu2-3 trp1-1 ura3-1 can1-100* *UBR1::GAL1-10-Ubiquitin-M-LacI* *fragment-Myc-UBR1 (HIS3)* *leu2-3::pCM244 (CMVp-tetR'-SSN6, LEU2) x3* *rad61Δ::hphmx4*
1448	*top2-4 eco1-1 SEC53 FMP32* pRS315 (converging)	*Mat a ade2-1 his3-11 leu2-3,112 trp1-1 ura3-1 can1-100* *top2-4* *eco1-1 (G211H)* *SEC53 FMP32* pRS315 (converging)
1452	*top2-4 eco1-1 mukB* pRS315 (genes removed)	*Mat a ade2-1 his3-11 leu2-3,112 trp1-1 ura3-1 can1-100* *top2-4* *eco1-1 (G211H)* *mukB* pRS315 (genes removed using bglII)
1455	*top2-4 SEC53 FMP32* pRS315 cenmut (converging)	*Mat a ade2-1 his3-11 leu2-3,112 trp1-1ura3-1 can1-100* *top2-4* *SEC53 FMP32* pRS315 *cen-mut* (converging)
1456	*top2-4 mukB* pRS315 cenmut (genes removed)	*Mat a ade2-1 his3-11 leu2-3,112 trp1-1 ura3-1 can1-100* *top2-4* *mukB* pRS315 *cen-mut* (genes removed using bglII)
1472	*top2-4 ndc10-1 SEC53 FMP32* pRS315 (converging)	*Mat a* *top2-4* *ndc10-1*

		SEC53 FMP32 pRS315 (converging)
1473	*top2-4 ndc10-1 mukB* pRS315 (genes removed)	*top2-4* *ndc10-1* *mukB* pRS315 (genes removed using bglII)
1475	*top2-4 GALL-SCC1 SEC53 FMP32* pRS315 (converging)	*Mat a ade2-1 his3-11 leu2-3,112 trp1-1 ura3-1 can1-100* *top2-4* *GALL-SCC1* *SEC53 FMP32* pRS315 (converging)
1476	*top2-4 GALL-SCC1 mukB* pRS315 (genes removed)	*Mat a ade2-1 his3-11 leu2-3,112 trp1-1 ura3-1 can1-100* *top2-4* *GALL-SCC1* *mukB* pRS315 (genes removed using bglII)
1480	*top2-4 rad61Δ SEC53 FMP32* pRS315 (converging)	*Mat a ade2-1 his3-11 leu2-3,112 trp1-1 ura3-1 can1-100* *top2-4* *rad61Δ::hphmx4* *SEC53 FMP32 pRS315 (*converging)
1481	*top2-4 rad61Δ mukB* pRS315 (genes removed)	*Mat a ade2-1 his3-11 leu2-3,112 trp1-1 ura3-1 can1-100* *top2-4* *rad61Δ::hphmx4* *mukB* pRS315 (genes removed using bglII)
1486	*rad61Δ top2-td*	*Mat a ade2-1 his3-11 leu2-3 trp1-1 ura3-1 can1-100* *UBR1::GAL1-10-Ubiquitin-M-LacI* *fragment-Myc-UBR1 (HIS3)* *leu2-3::pCM244 (CMVp-tetR'-SSN6, LEU2) x3* *top2-td TOP2 5' upstream -100 to -1 replaced with kanMX-tTA (tetR-VP16)-tetO2 - Ub -DHFRts - Myc -linker)* *rad61Δ::hphmx4*
1487	*rad61Δ top2-td*	*Mat a ade2-1 his3-11 leu2-3 trp1-1 ura3-1 can1-100* *UBR1::GAL1-10-Ubiquitin-M-LacI* *fragment-Myc-UBR1 (HIS3)* *leu2-3::pCM244 (CMVp-tetR'-SSN6, LEU2) x3* *top2-td TOP2 5' upstream -100 to -1 replaced with kanMX-tTA (tetR-VP16)-tetO2 - Ub -DHFRts - Myc -linker)* *rad61Δ::hphmx4*

1489	*eco1-1 top2-td*	*Mata ade2-1 his3-11 leu2-3 trp1-1 ura3-1 can1-100* *UBR1::GAL1-10-Ubiquitin-M-LacI* *fragment-Myc-UBR1 (HIS3)* *leu2-3::pCM244 (CMVp-tetR'-SSN6, LEU2) x3* *top2-td TOP2 5' upstream -100 to -1 replaced with kanMX-tTA (tetR-VP16)-tetO2 - Ub -DHFRts - Myc -linker)* *eco1-1 (G211H)*
1490	*eco1-1 top2-td*	*Mat a ade2-1 his3-11 leu2-3 trp1-1 ura3-1 can1-100* *UBR1::GAL1-10-Ubiquitin-M-LacI* *fragment-Myc-UBR1 (HIS3)* *leu2-3::pCM244 (CMVp-tetR'-SSN6, LEU2) x3* *top2-td TOP2 5' upstream -100 to -1 replaced with kanMX-tTA (tetR-VP16)-tetO2 - Ub -DHFRts - Myc -linker)* *eco1-1 (G211H)*
1496	*top2-4 SEC53 FMP32* pRS315 (convergent)	*Mat a his4-539 ura3-52* *top2-4* *SEC53 FMP32 pRS315* (converging)
1629	*top2-4 eco1Δ rad61Δ* *SEC53 FMP32* pRS315 (convergent)	*Mat a* *top2-4* *eco1Δ::natNT2* *rad61Δ::hphmx4* *SEC53 FMP32 pRS315* (converging)
1630	*top2-4 eco1Δ rad61Δ* *mukB* pRS315 (genes removed)	*Mat a* *top2-4* *eco1Δ::natNT2* *rad61Δ::hphmx4* *mukB* pRS315 (genes removed using bglII)

Table 2.1 List of Strains used in the study

2.4 LIST OF OLIGONUCLEOTIDES FOR GENETIC MANIPULATION

SCC1_S1	AAGAAAAGACAACTCAATTGCACAATTACTTTACAAGAAACACGACAATGCGTACGCTGCAGGTCGAC
SCC1_S4	TTATTGGTGGCAAGTCTTAAAACAGTAAGACGTTGAGGATTTTCTGTAACCATCGATGAATTCTCTGTCG
RAD61_S1	AAAACGAAACCATCTTCTTACCCTAAAGCATCCTGTTTCTGAAAAAATGCGTACGCTGCAGGTCGAC
RAD61_S2	TGCCAGCAGGGTGAAGATGAAGCCAGGCTATGTTCAATGTATGCTTTCTTTAATCGATGAATTCGAGCTCG
cen_mut_F1	AAGAAATTAAAGAAAAAATAGTTTTTGTTTTCATAAGATGTAAAAGACTCTAGGGGGATCG
cen_mut_R1	CGATCCCCCTAGAGTCTTTTACATCTTATGAAAACAAAAACTATTTTTTCTTTAATTTCTT
RGD2_Hind_F	ATTCAAAGCTTCCCGGGCTGCAGACCTTTTGCA
SEC53_Bam_R	ATTGAGGATCCTGAAAAAGGAATTCCCAGACTACGGTTTGACT

Table 2.2 List of oligonucleotides

2.5 MOLECULAR CLONING TECHNIQUES

2.5.1 *E.coli* transformation

Competent DH5α cells stored at -80°C were thawed on ice. DNA is added to 50 µl of the cells and incubated for 30 min on ice. DNA was then heat-shocked at 42°C for 30 seconds followed by being on ice for 2 min. 800ul of 37°C LB was then added and the cells were put for 1 h shaking (8000 rpm) at 37°C. 80 µl and 720 µl were plated onto LB plates with Ampicillin (100 µg/ml) and incubated at 37°C overnight.

2.5.2 Miniprep

To extract plasmid from *E.coli* transformations a Miniprep kit (Quiagen, 27106) was used as per manufacturer's instructions. E. coli cells were grown overnight in 8 ml LB with Ampicillin (100 µg/ml) at 37°C. DNA was eluted in 30-50 µl of water.

2.5.3 Restriction digests

Restriction digests were carried out according to the manufacturer's instructions. The enzymes were purchased from New England Biolabs (NEB) or Thermo Fisher Scientific.

2.5.4 End processing

FastAP Thermosensitive Alkaline Phosphatase (Thermo Fisher Scientific, EF0654) and Mung Bean Nuclease (NEB, M0250S) were used for end processing techniques as per the manufacturer's instructions. FastAP was used for catalysing the release of 5'- and 3'-phosphate groups and Mung Bean was used for producing blunt end products.

2.5.5 Ligation

50 ng of vector DNA was incubated on ice with insert DNA in a ratio of 1:2 and 1:3 vector to insert with negative controls. T4 Rapid DNA ligase, (Thermo Fisher Scientific, k1422), was then used as per manufacturer's instructions (15 min at room temperature).

2.5.6 PCR (Polymerase Chain Reaction)

2.5.6.1 KOD hot start DNA polymerase (Novagen, 71086)

Used with: 10x buffer, 25 mM MgS04, 2 mM each dNTPs, 5' Primer 10 µm, 3' Primer 10 µm, template DNA 10 ng plasmid 100 ng genomic, H_2O. Cycling conditions: activation 95°C for 2 min, 1 cycle. Denaturation 95°C for 20 s, annealing 55°C for 10 s, extension 70°C with the time dependent on the amount of template normally 15-20 s/kb, 2X 29 cycles. 70°C for 10 min, 1 cycle.

2.5.6.2 Bacterial colony PCR

Taq DNA polymerase (Thermo Fisher Scientific, 18038042) used with: 10x buffer IV, 25 mM $MgCL_2$, 2 mM each dNTPs, 5' Primer 10 µm, 3' Primer 10 µm, denatured template DNA, H_2O. Cycling conditions: activation 95°C for 2 min, 1 cycle. Denaturation 95°C for 20 s, annealing 55°C for 30 s, extension 70°C 1 min/kb, 2X 29 cycles. 70°C for 5 min, 1 cycle.

Phusion High-Fidelity PCR Master Mix (NEB, M0531S) used: 5' Primer 10 µm, 3' Primer 10 µm, 5 µl Master Mix, 5 µl Denatured DNA. Cycling conditions: activation 98°C for 30 s, 1 cycle. Denaturation 98°C for 10 s, annealing 55°C for 30 s, extension 72°C 30 s/kb, 2X 29 cycles. 72°C for 5 min, 1 cycle.

2.5.7 Site-Directed Mutagenesis

Performed using the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent technologies, 210518). Overlapping primers were designed with the desired mutation. PCR reaction: 10x buffer, 10 ng template plasmid, 125 ng 5' Primer, 125 ng 3' Primer, 1 µl dNTP mix, 1.5 µl of QuikSolution reagent, H_2O, 1 µl QuikChange Lightning Enzyme. Cycling conditions: activation 95°C for 2 min, 1 cycle. Denaturation 95°C for 20 s, annealing 60°C for 10 s, extension 68°C 30 s/kb, 2X 18 cycles. 68°C for 5 min, 1 cycle. Digestion of the parental DNA: 2 µl DpnI restriction enzyme added and left at 37°C for 5 min.

2.5.8 Agarose Gel electrophoresis

DNA was resolved using 1% Agarose (Fisher Scientific, BP1356), 0.5xTBE, 0.5 mg/ml Ethidium Bromide (Sigma, E1510). Samples were loaded using 6X Loading Dye (Thermo Fisher Scientific; R0611) and run next to a 1 kb DNA ladder (Bioline, BIO-33053) at 100 V. Gels were visualised with a UV illuminator (Syngene InGenious Gel Analysis System).

2.5.9 DNA purification by gel extraction

DNA was resolved by the method outlined in 2.5.8. DNA bands were visualised using a UV illuminator and select DNA band sizes were exised from the Agarose gel. These were dissolved and the DNA was extracted and purified using NucleoSpin Gel and PCR clean-up kit as per the manufacterers instructions (MACHEREY-NAGEL, 740609.50).

2.6 YEAST TECHNIQUES

2.6.1 Genetic Crosses

MATa strains were crossed with MATα strains by growing each up individually from a -80 60% glycerol stock on YPD agar. They were mixed together and incubated on YPD agar or selection media at 25°C for 24-48 hours. Then streaked out to single colonies and grown on RSM sporulation media for up to 5 days until sporulated.

2.6.2 Tetrad dissection

Sporulated cells were resuspended in 250 µl H_2O and incubated for 5 min with 1 µl of Zymolyase (AMS Biotech, 120493-1). 10 µl of digested cells were spread down one side of the plate and individual tetrads were dissected using the singer tetrad dissector (Singer, MSM400).

2.6.3 Yeast Colony PCR

A colony was re-suspended in 50 µl H_2O and heated at 95°C for 10 min. 5 µl was used for each PCR reaction. Taq DNA polymerase (Thermo Fisher Scientific, 18038042) or Phusion High-Fidelity PCR Master Mix (NEB, M0531S) was used as in 2.5.6.2.

2.6.4 Yeast Transformation

Cells were grown overnight in YPD and diluted in the morning to 2-4 x 10^5 in a 50 ml culture. When the cells were in exponential phase (100-120 x 10^5) they were spun down 3000 rpm 2 min and washed in 10 ml H_20. Cells re-suspended in 1 ml H_20 and transferred to eppendorfs. Then spun at 8000 rpm for 1 min and washed once in 1 ml 0.1 M LiOAc- 1xTE pH 7.5, before re-suspending in 250 µl LiOAc-TE.

For each transformation 50 µl of cells was used with 5 µl (10 mg/ml) denatured salmon sperm (Life technologies, 15632-011) (boiled for 5min and put on ice), 1 µg DNA, 300 µl 40% PEG in 0.1M LiOAc- 1xTE pH7.5. The mix was then vortexed thoroughly and put at 25°C shaking for 45 min. Then 40 µl of Dimethyl sulfoxide (DMSO) (Fisher scientific, D/4121/PB08) was added and mixed and heat shocked for 15 min at 42°C (For some sick or temperature sensitive strains the heat shock step was missed out). Then put on ice for 2 min. The cells were allowed to recover at

25°C for 1 hour when plating onto minimal plates or overnight when plating onto HygB, NAT or G418 (see section 2.2). Cells were re-suspended in 1x TE before plating.

2.6.5 Gene knockout

The method designed by (Janke et al., 2004) was used for homology-mediated recombination of the desired genomic loci. S1 and S2 primers were designed specifically for the gene to be removed and ended in 5′ -CGTACGCTGCAGGTCGAC-3′ or 5′-ATCGATGAATTCGAGCTCG-3′ respectively (primers shown in Table 2.2). These were then recombined with regions of homology from the plasmid outlined in Janke et al. (2004).

2.6.6 GALL Gene promoter insertion

The method designed by (Janke et al., 2004) was used for homology-mediated recombination of the N terminal promoter region of the gene. S1 and S4 primers were designed specifically for the gene and ended in 5′-CGTACGCTGCAGGTCGAC-3′or 5′- CATCGATGAATTCTCTGTCG -3′ respectively (primers shown in Table 2.2). These were then recombined with GALL DNA from the PYM-N27 plasmid from Janke et al. (2004).

2.6.7 Cell cycle synchronisation timecourses

10 ml liquid cultures of either dropout or YP media with 40 mg/l adenine, 2% glucose or 2% raffinose were grown up from the morning by looping cells from a minimal or YPD plate (for media recipes see section 2.1). In the evening cell numbers were counted using a sample number under the microscope, and then diluted in a 50 ml culture and agitated at 25°C in a waterbath. The next morning cell number was counted (estimated from a sample) and If grown in dropout media moved to YP media with either 2% glucose or 2% raffinose and 40 mg/l adenine when at 80 x 10^5 cells/ml. With cultures grown in YP media and with those transferred to YP media, alpha factor was added to a concentration of 10 µg/ml when they were exponentially growing ~120 x 10^5 cells/ml. Just before the alpha factor pheromone addition, an exponential Flow cytometry sample was taken (see 2.6.8).

The alpha factor addition lead to a block in Gap 1 (G1) phase of the cell cycle. 5 µg/ml of alpha factor was added as a supplement 1 h 30 min after the first addition. The budding index was

then also taken and if > 90% of the cells were in G1 phase (not budded) the timecourse could proceed. If using a degron system, cultures were grown in YP with 40 mg/l adenine and 2% raffinose and at this point 2% galactose was also added. 20 min later doxycycline (dox) at a concentration of 50 µg/ml was added. Then 10 min after the dox addition the temperature of the waterbath was set to 37°C and allowed to increase slowly. At this point, the flow cytometry sample 2 was also taken. If not using a degron system, when the budding index had been verified the temperature was increased straight away to 37°C and Flow cytometry sample 2 taken.

1 hour after reaching 37°C Flow cytometry sample 3 was taken, and washes from alpha factor undertaken. There were three washes in 25 ml of depletion media followed by resuspension in 50 ml depletion media. 45 min after the first wash nocodazole was added at a concentration of 10 µg/ml in order to lead to a Gap 2/Mitosis (G2/M) arrest. 75 min after the first wash the budding index was checked. If > 80% of cells were budded, samples, and Flow cytometry sample 4, were taken from between 80 and 100 min post release (collection of samples for various experiments are explained in sections 2.7.2, 2.8.1 and 2.9.2).

2.6.8 Flow Cytometry

500 µl of culture samples were spun down and re-suspended in 70% ethanol for fixing. They were then spun down and re-suspended in 1 ml of 50 mM Tris-HCl pH8 with 5 mg/ml RNaseA (Sigma, R4875). These tubes were left shaking at 37°C overnight.

Samples were pelleted and re-suspended in 5 mg/ml pepsin (Sigma, P6887) 5 µl/ml concentrated HCl. They were then incubated at 37°C for 30 min. Samples were pelleted and washed in 50 mM Tris-HCl pH8 (Sigma, T3038) before being re-suspended in 50 mM Tris-HCl pH8 with 0.5 mg/ml Propidium iodide (Sigma, P4170) and sonicated ready for analysing using the BD Accuri C6 sampler and analysed using FCS express 4 flow software.

2.6.9 DNA extraction without phenol

A colony of cells were picked from a plate and re-suspended in 200 µl lysis buffer (50 mM Tris-HCl (pH 8.0), 100 mM NaCl, 10 mM EDTA, 1% SDS), with 40 units lyticase (Sigma Aldrich, L2524) and 5 µl 2-mercaptoethanol (Sigma Aldrich, 63689) at 37°C for 5 min. 120 µl of 5 M KAc was then added and samples were left on ice for 10 min before a 10 min centrifugation at 4°C 13000 rpm.

Supernatant was removed to new eppendorf's and 700 µl isopropanol added. Tubes were then centrifuged for 15 min at 4°C, 13000 rpm. Pellets washed with 70% ethanol before being air dried and re-solubilized in 100 µl 1x TE with RNaseA.

2.7 CATENATION ANALYSIS

2.7.1 Solutions

1 x TBE	1 L: Tris-base 10.8 g, Boric Acid 5.5 g, 0.5 mM EDTA pH8 4 ml, H_2O
Depurination buffer	0.125 M HCl
Denaturation buffer	0.5 M NaOH, 1.5 M NaCl
Neutralisation buffer	0.5 M Tris-HCl, 1.5 M NaCl pH 7.5
20 x SSC	1L: 175.3 g NaCl, 88.2 g Sodium citrate pH 7.0, H_2O

Table 2.3 List of solutions for 2D gels and Southern blotting

2.7.2 Collection of cell samples

At the end of the timecourse 10 ml samples were spun down 3500 rpm for 2 min, supernatant removed and pellets frozen in dry ice before storage at -80.

2.7.3 Preparation of Plasmid DNA by Phenol DNA Extraction and Nicking

The 10 ml frozen pellets were re-suspended in 400 µl lysis buffer (50 mM Tris-HCl (pH 8.0), 100 mM NaCl, 10 mM EDTA, 1% SDS), with 40 units lyticase (Sigma, L2524) and 5 µl 2-mercaptoethanol (Sigma Aldrich, 63689) at 37°C for 5 min. Then 450 µl phenol/chloroform/iso-amylalcohol (25:24:1-Sigma) was added and mixed by rotation. The aqueous layer was collected using phase lock tubes (5 prime, 2302800), 5 min 12000 rpm. The DNA was precipitated with 1 ml of 100% ethanol (~2x concentration) followed by 500 µl 70% ethanol before being air dried and re-solubilized in 100 µl 10 mM Tris (pH 8.0).

The nicking of the plasmids was done using 25 µl of the solubilised DNA (depending on amount of cells collected) with either Nb.Bsm1 (NEB, R0706S) or Nb.BsrDI enzymes (NEB, R0648S) used

according to the manufacturers instructions and made up to a 90 µl reaction volume. Each enzyme was used depending on the plasmid being assayed; in order to achieve greater than one nick per plasmid.

After nicking the process of precipitation of DNA was again repeated however this time with the addition of 10 µl of 3 M sodium acetate (pH 5.2). Again a 2X concentration of 100% ethanol was used (250 µl) followed by 200 µl of 70% ethanol. The precipitated DNA then was air dried and re-solubilized in 18 µl 10 mM Tris (pH 8.0). Before 6X loading dye (Thermo Fisher Scientific, R0611) was added.

2.7.4 Two Dimension Gels for catenation analysis

DNA was separated in the first dimension by 0.4% MegaSieve agarose (Flowgen, H15608) in 250 ml 1X TBE. The gel was run at 30 V at room temperature in 1X TBE, (see table 2.4 for running times of various plasmids). 2 µl was loaded separately to the rest of the sample using loading dye next to DNA ladder (Bioline, 33053). After running, the lanes with 2 µl sample and ladder were removed and incubated in 1X TBE + 0.5 µg/ml ethidium bromide. This allowed visualisation of how far the DNA had run into the gel and therefore accurate excision of the plasmid DNA for analysis (Figure 2.1). The excised lane was set in MegaSieve agarose (Table 2.4) and again run for a period of time to separate the DNA (Table 2.4).

Plasmid size	Running conditions 1st dimension	Running conditions second dimension
5 kb	0.4% agarose, 30 V, 16 h, rt, up to 4 samples per gel	1.2% agarose, 120 V, 18-20 h, 4˚C, up to 2 samples per gel
6 kb	0.4% agarose, 30 V, 16 h, rt, up to 4 samples per gel	1% agarose, 100 V, 20 h, 4˚C, up to 2 samples per gel
8 – 9 kb	0.4% agarose, 30 V, 24 h, rt, up to 4 samples per gel	0.8% agarose, 45 V, 24 h, rt, up to 1 sample per gel

Table 2.4 2D gel running conditions

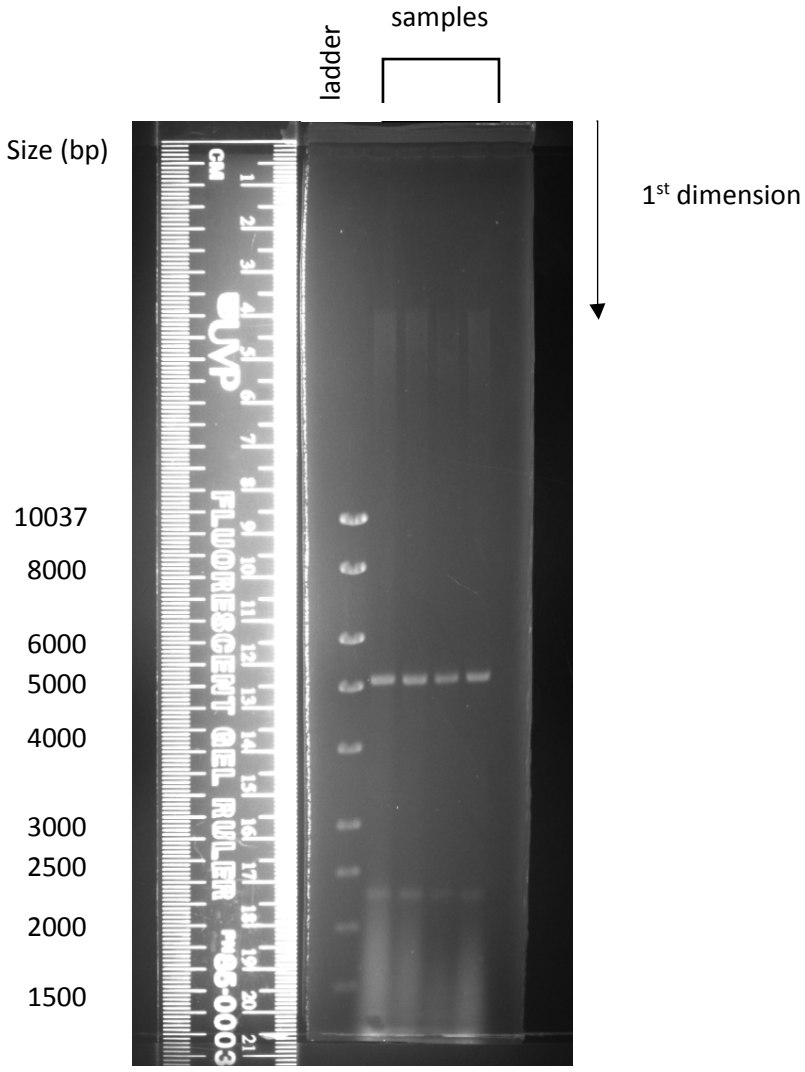


Figure 2.1 Example of a first dimension gel

Run with four 2 µl DNA samples and DNA ladder. A slow running species is visualised at 4 cm from the wells. A strong band can be seen just above 5 kb which relates to the rDNA. The fast running species is fragmented genomic DNA. The gel tank was 14.6 cm in width and therefore a section of this length was cut from the gel, including the catenated DNA forms, to run in the second dimension.

2.7.5 Southern blotting

To prepare the gel for Southern blotting it was washed as indicated in Table 2.5 with gentle shaking (buffers from Table 2.3). The DNA was transferred onto a Hybond N+ transfer membrane (GE Healthcare Amersham, RPN203B) by capillary action in 20X SCC. After transfer, DNA was ultraviolet cross-linked to the membrane using a UV Stratalinker 1800 (Stratagen) at 1200 J/m. The membrane was then gently shaken in 5X SSC for at least 4 h. The membrane was blocked at 60°C (5X SSC, 5% Dextran Sulphate (Sigma Aldrich, D8906), 0.2% Tropix I-Block (Applied Biosystems, T2015), 0.1% SDS) for at least 1 h. DNA probe (see section 2.7.6) was boiled for 5 min before cooling for 2 min and added to the blocking liquid at 60°C, before being left overnight.

The membrane was washed at 60°C twice in 1X SSC with 0.1% SDS, followed by 2 washes with 0.5X SSC and 0.1% SDS. It was then blocked in AB buffer (100 mM Tris-HCl, 150 mM NaCl (pH 7.5)) with 1% milk for at least 1 h, before incubating in AB buffer with 0.5% milk and alkaline phosphatase Anti-fluorescein-AP Fab fragments (Roche, 11426338910) for 1 h. Following this, the membrane was washed 3 times in AB buffer with 0.2% Tween 20 (Sigma Aldrich, P1379). DNA was visualised using CDP-Star detection agent (GE Healthcare, RPN3682), it was incubated for 5 min, before sealing and incubating in the dark for 15 min. Exposures of the membrane were taken using ImageQuant LAS 4000 (GE healthcare).

Buffer (600 ml)	Time (min)
H_2O	5
Depurination	10
H_2O	10
Denaturation	45
H_2O	2
Neutralisation	30

Table 2.5 Gel preparation washes

2.7.6 Southern blotting FldUTP Probe

To make the probe, a section of plasmid DNA > 1 kb was first amplified, and 200 ng was diluted to 40 µl using H_2O (Primers used for amplification are shown in Table 2.6). This was then boiled for 5 min, following this 60 µM random nonamer primers (invitrogen) were added and boiled

for a further 1 min, then put on ice for 5 min. Following this a mixture was made of: 10X NEB buffer 2.1 (NEB, B7201), 2 µl 5 mM dATP (Invitrogen), 2 µl 5 mM dCTP, 2 µl 5 mM dGTP, 2 µl 0.35 mM Flourescein-12-dUTP (Roche, 11373242910)/0.65 mM dTTP, 30 µl H_2O. This mix was then added to the DNA, and 2 µl klenow (NEB, M0212L) was added. This was left at 37°C for 2 h.

Primer Name	Primer Sequence	Plasmid amplified
URA3 Forward	GCTACTCATCCTAGTCCTGTT	pRS316
pRS Reverse	CGCATCTGTGCGGTATTTCA	pRS316/ pRS315
LEU2 Forward	CTTAACTTCTTCGGCGACAGCAT	pRS315

Table 2.6 Primers used to amplify the *URA3* and *LEU2* genes

2.7.7 Quantification of catenane numbers across a population

ImageQuant TL software was used to carry out densitometry analysis, where the strength of the signal for each catenated node, 1 – 27 was taken and could be represented as a percentage in a histogram format. The first 5 nodes should be below 0.5% of the population for each, which is due to the termination of DNA replication requiring fork rotation to unwind the final base pairs of DNA. If it was much higher than 0.5% this indicated a reactivation of Top2, and data is excluded where the percentage for these nodes increased above 2%. Above 27, the nodes merged together and became more faint, due to low numbers of the population being very highly catenated (Figure 3.1D). Therefore, a range of signal was taken, one range including the catenated nodes from 1 to 20, and the other range from 20 and above. These could be used as an additional indicator of the levels of catenanes and were used to help calculate the mean catenated node. Error bars, and percentage population having over 20 catenanes error, were calculated using the average deviation of the repeats.

2.8 SUPERCOILING ANALYSIS

2.8.1 Collection of samples and preparation of plasmid DNA

10 ml samples were collected 20 min post alpha factor release, they were spun down 3500 rpm for 2 min, supernatant removed and pellets frozen in dry ice before storage at -80.

Preparation of plasmid DNA occurred as in 2.7.3, without the further nicking process.

2.8.2 Two dimension gels for plasmid supercoiling analysis

DNA was separated in the first dimension by 0.4% Megasieve agarose in 250 ml 1X TBE +0.5 µg/ml chloroquine (Sigma Aldrich, C6628). The gel was run at 30 V for 20 h, rt, in the dark, in 1x TBE +0.5 µg/ml chloroquine. 2 µl DNA was loaded separately to the rest of the sample using loading dye next to DNA ladder. After running, the lanes with 2 µl sample and ladder were removed and incubated in 1X TBE + 0.5 µg/ml ethidium bromide. These lanes were visualised to see the running length of the DNA. The other lanes were incubated with 1X TBE + 1 µg/ml chloroquine for 3 h. The second dimension was set in 1.2% MegaSieve agarose in 250 ml 1X TBE + 1 µg/ml chloroquine. The second dimension running conditions were: 1X TBE +1 µg/ml chloroquine, at 120 V, 4°C, for 10 h.

Southern blotting was carried out as in 2.7.5

2.8.3 Analysis of plasmid supercoiling

The chloroquine intercalater was used in the gels to separate supercoiling by sign. In the first dimension the plasmids were separated by size, with more highly supercoiled monomers running the furthest. In a native gel, supercoiling of either +/- 14 ran at the same distance. In a chloroquine gel of 0.5 µg/ml +7 supercoils were introduced meaning a +14 plasmid ran father, equivalent to +21, and conversely a -14 plasmid ran slower at -7 which then overlapped with plasmids that started at 0 and were then at +7 (Figure 2.2). Then in the second dimension, run at a chloroquine concentration of 1 µg/ml, an extra +7 supercoils were introduced. This means that the overlapping plasmids of -7 and +7 were separated as the -7 ran as if it has 0 supercoils and +7 will run as +14 supercoils (Figure 2.2) (Baxter et al., 2011).

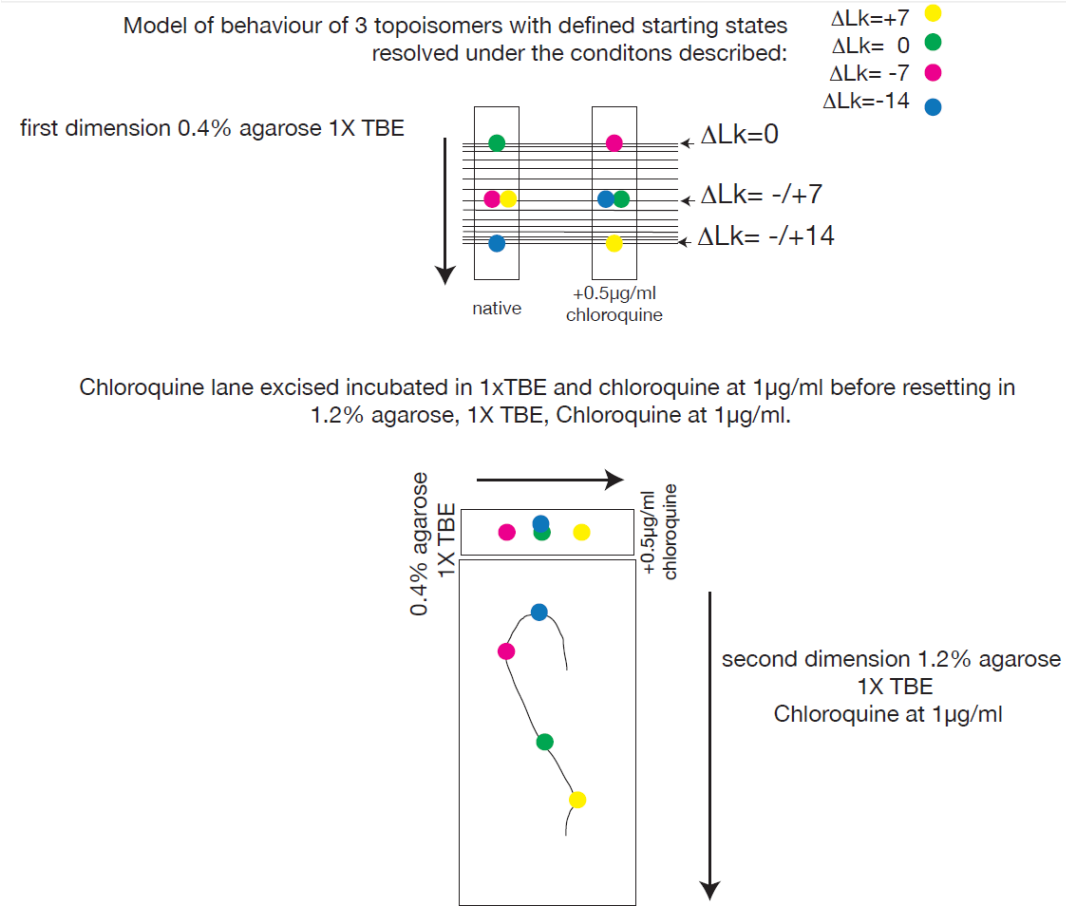


Figure 2.2 Chloroquine intercalation 2D gel electrophoresis

From J. Baxter. In a first dimension agarose gel of 0.5 µg/ml chloroquine +7 supercoils are introduced. The lane is exised and run in a second dimension of 1 µg/ml chloroquine adding a further +7 supercoils to separate the supercoiling state by sign.

2.9 ChIP-seq analysis

2.9.1 Solutions

SDS buffer	1% SDS, 10 mM EDTA, 5M Tris HCl, 1 tablet/10-50 ml protease inhibitor, 1 tablet/10-50 ml phosphatase inhibitor, H_2O keep at 4°C
IP buffer	0.1% SDS, 1.1% Triton-X-100, 1.2 mM EDTA, 16.7 mM TRIS HCl (pH8), 1 tablet/50 ml protease inhibitor, 1 tablet/50 ml phosphatase inhibitor, H_2O
Elution Buffer	1% SDS, 0.1 M $NaHC0_3$, H_2O
TSE-150	1% Triton-X-100, 0.1% SDS, 2 mM EDTA, 20 mM Tris HCl (pH8), 150 mM NaCl, H_2O
TSE-500	1% Triton-X-100, 0.1% SDS, 2 mM EDTA, 20 mM Tris HCl (pH8), 500 mM NaCl, H_2O
LiCl wash	0.25 M LiCl, 1% NP-40, 1% dioxycholate, 1 mM EDTA, 10 mM Tris HCl (pH8), H_2O

Table 2.7 Solutions for ChIP

2.9.2 Collection of cell samples

At the end of the timecourse (2.6.7), 50 ml cultures were spun down and re-suspended in YP media 25°C. A final concentration of 1% formaldehyde (Sigma Aldrich, F8775) was added and the cultures were incubated at 25°C shaking for 45 min. The cross-linking reaction was then quenched with 125 mM glycine for 5 min at 25°C shaking. Cultures were spun down and washed twice with 10 ml cold Phosphate buffered saline (PBS) (NaCl 0.138 M, KCl 0.0027 M, pH 7.4). The pellets were then re-suspended in 1 ml cold PBS and moved to ribolyser tubes. They were pelleted again, the supernatant removed, and the pellet snap frozen in liquid nitrogen before storage at -80.

2.9.3 Chromatin Immunoprecipitation (ChIP)

Pellets were resuspended in 500 µl SDS buffer (Table 2.7), and 200 µl Zirconia/Silica beads were added (BioSpec Products, 11079105z). Cells were lysed using the FastPrep-24 (M.P. Biomedicals, 116004500), 5X 1 min with 3 min on ice in between. Tubes were then needle pierced and lysed cells transferred to new ribolyser tubes by centrifugation at 2000 rpm, 1 min 4°C. The new tubes were centrifuged at 13,000 rpm for 1 min and the supernatant was transferred to new pre-cooled tubes, and IP buffer (Table 2.7) was added to make a final volume of 1.1 ml. Tubes were again centrifuged at 13,000 rpm for 1 min, and 1 ml supernatant transferred to a milliTube (Covaris, 520130). These steps were taken to minimise cell debris, which reduces the efficiency of sonication. Tubes were sonicated using the Focused-Ultrasonicator (Covaris, M220) (Average incident power – 7.5 Watts, Peak Incident Power – 75 Watts, Duty Factor – 10 %, Cycles/Burst – 200, Duration – 20 min). After sonication, samples were checked for the efficiency of the sonication. 10 µl of the sonicated sample was added to 190 µl of elution buffer (Table 2.7), and 20 µl of 5 M NaCl, followed by incubation at 95°C for 15 min. Then 10 µl of DNase-free RNase (0.5 mg/ml, Roche, 11119915001) was added and tubes were incubated at 37°C for 30 min. The DNA was then purified and eluted in 50 µl, 1 µl was run on the bioanalyser (Agilent Technologies, G2939B) to measure DNA content and size (Figure 2.3). If DNA was not efficiently sonicated, the sonication and checking procedures were repeated. When the DNA was efficiently sonicated, the sample was transferred to an eppendorf 1.5 ml tube and centrifuged for 20 min at 13,000 rpm at 4°C. Supernatant was then diluted to 1:10 (5 ml total) and added to a preparation of Dynabeads. The preparation of Dynabeads consisted of: 50 µl protein A Dynabeads (Invitrogen, 10002D) and 50 µl protein G Dynabeads (Invitrogen, 10004D), added to 1 ml IP buffer (Table 2.7) and incubated for 3 min on a wheel at room temperature (rt). The supernatant was removed using a magnetic rack and the wash was repeated twice. The mixture of Dynabeads and sample was left for 2 h at 4°C. Supernatant was split, with 2X 2 ml being taken to 15 ml falcon tubes, and 1 ml being kept at -20°C as an input sample. To the two 2 ml samples antibody was added, either H2A 1:500 (active motif, 39235) or 1.6 µg/ml yH2A (Abcam, ab15083), and these were placed on a rotating wheel at 4°C for 15 – 20 h. Again a preparation of Dynabeads, Protein A (30 µl) and Protien G (30 µl), was made up and washed 3 times in IP buffer. This mix was added to each of the samples and was then incubated at 4°C for 4 h. Supernatant was removed and beads were washed at 4°C for 6 min in TSE-150, followed by TSE-500, followed by LiCl wash (Table 2.7) and finally Tris-EDTA (TE pH8). Beads were resuspended in 200 µl elution buffer (Table 2.7), and left at room temperature on a wheel for 30 min. Supernatant was transferred to a new tube and the elution step was repeated. At the same time 50 µl from the input sample was added to 150

µl of elution buffer. 20 µl of 5 M NaCl and 10 µl of 10 mg/ml proteinase K (Invitrogen, 10124532) was then added to the input and 40 µl and 20 µl to the IP samples respectively. These were incubated at 65°C overnight. Then 10 µl of DNase-free RNase was added to the input and 20 µl to the IP samples, and they were left at 37°C for 30 min. All DNA was purified with a Qiagen PCR pufirication kit (28106) and eluted in 50 µl. DNA amount was measured using the Qubit 2.0 Fluorometer (Life technologies) as per the manufacturer's instructions.

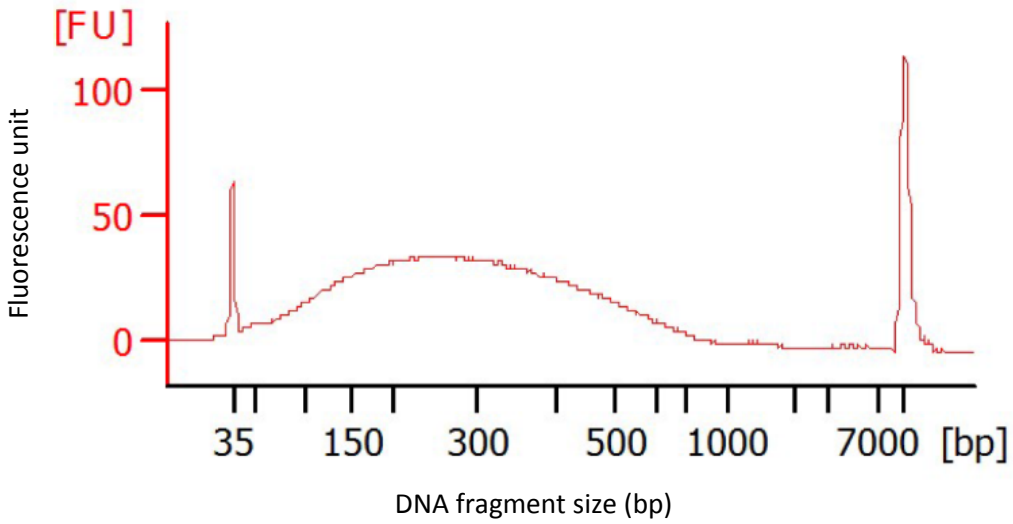


Figure 2.3 An example of efficient sonication of DNA

DNA should be sheered to a level of 300 bp. If inefficient sonication occurs, (where the fluorescence is ~1000 bp) sonication is repeated. When marker is run, two sharp marker spikes appear at either end of the graph at 35 bp and 10380 bp for reference.

2.9.4 ChIP-seq Library Preparation

Library preparation was carried out using the NEBNext Ultra II library kit (NEB, E7645S). 1 ng of DNA was made up to 50 µl with H_2O, followed by adding 7 µl NEBNext Ultra II End Prep Reaction Buffer and 3 µl NEBNext Ultra II End Prep Enzyme Mix. The tubes were incubated for 30 min at 20°C, then for 30 min at 65°C. Adapters were then ligated by adding:

- NEBNext Ultra II Ligation Master Mix 30 µl
- NEBNext Ligation Enhancer 1 µl
- 0.6 µM NEBNext Adaptor 2.5 µl

Which were then incubated for 15 min at 20°C. Following this, 3 µl USER enzyme was added and tubes were incubated at 37°C for 15 min. Adaptor ligated DNA was then size selected for 250 bp insert size. The volume of the DNA was made up to 100 µl using H_2O, followed by addition of 45 µl of AMPure XP beads (Beckman Coulter, A63881), and the mixture was left for 5 min at rt. Supernatant was removed and 25 µl of AMPure XP beads was added and then left for 5 min at rt. Supernatant was discarded and beads were washed 3X with 80% ethanol. Beads were then air dried for 10 min, and DNA was eluted from the beads into 20 µl of H_2O.

To perform the PCR library enrichment, primers were used from any of NEBNext Multiplex Oligos for Illumina sets 1-4 (NEB, E7335S, E7500S, E7710S, E7730S). Components added were:

- 2X NEBNext High-Fidelity PCR Master Mix 25 µl
- 10 µM Universal PCR primer 2.5 µl
- 10 µM Index Primer 2.5 µl

PCR conditions were:

Cycle step	Temperature °C	Time	Cycles
Initial denaturation	98	30 sec	1
Denaturation	98	10 sec	13
Annealing/extension	65	75 sec	
Final extension	65	5 min	1
Hold	4		

Table 2.8 Cycling conditions for PCR

The PCR reaction was then purified using AMPure XP beads. 50 µl beads were added to the PCR reaction, and incubated for 5 min rt. Supernatant was discarded and beads were washed twice with 80% ethanol. Beads were dried for 10 min, and the DNA eluted from the beads into 25 µl H_2O. 25 µl beads were then added to the DNA, and incubated for 5 min rt. Supernatant was discarded and beads were washed twice with 80% ethanol. Beads were dried for 10 min, and the DNA eluted from the beads into 20 µl H_2O. The library DNA was then measured by the Qubit 2.0 Fluorometer and bioanalyser. DNA was diluted to make a 4 nM library to run using the MiSeq system (illumina) as per the manufacturer's instructions. Each strain had 3 samples, H2A ChIP, γH2A ChIP and input sample. 12 samples (4 strains) were run at one time using the MiSeq system to gain a high level of reads per sample.

2.9.5 Analysing the ChIP-seq data

FASTQ files were generated by illumina basespace (https://basespace.illumina.com/home/index). This generates two files representing 75 bp paired end reads for each sample. These files were then aligned to a reference genome using Bowtie 2, which generates a SAM output file for each sample (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml). Command:

```
bowtie2 -p 14 -x [path to index folder]  --trim3 25 --trim5 1 -1 [Path and name of R1 fastq file] -2 [Path and name of R2 fastq file] -S [name of the resulting .sam file]
```

Each of the reads aligned were trimmed 25 bp from the 3' end and 1 bp from the 5' end, to allow for any initial sequencing error. SAM files were then converted into BAM files by using SAMtools (http://samtools.sourceforge.net/). Command:

```
samtools sort [name of the .sam file generated with bowtie2] -o [name for the resulting .bam file] -O bam -T [name for resulting .bam file wo .bam]
```

BAM files were used for Model-based Analysis of ChIP-Seq (MACS2). MACS2 was used to analyse ChIP-seq data to identify binding sites. To do this the program uses a 'call peak' function and generates genome wide score data for these peaks. Following this, the samples were compared

to one another, for example, the γH2A sample to the H2A sample. Which then generated a fold enrichment track. Command:

```
macs2 callpeak -t [sorted BAM file from yh2a data]-c  [sorted BAM file
from h2a data]-f BAMPE -g 12100000 -n [name for output file] -B -q
0.01 --SPMR
```

Following this, the data was binned into 50 bp regions and WIG file format generated in order to allow visualisation on the Integrative Genomics Viewer (IGV) software, CSV file formats were generated to further analyse the data using the R programme.

3 ANALYSIS OF FORK ROTATION IN THE CONTEXT OF TRANSCRIPTION

3.1 INTRODUCTION

DNA replication fork rotation appears to occur at situations of high topological stress, such as stable protein-DNA sites (Schalbetter et al., 2015), and the termination of DNA replication (Champoux and Been, 1980). At these sites, the exclusion of topoisomerases and subsequent build-up of topological stress, force the replication fork to rotate upon its axis (Keszthelyi et al., 2016).

An area of high topological stress that has yet to be analysed in the context of fork rotation, is that of DNA transcription. Although topoisomerases are known to resolve topological stress arising due to transcription, they do not have the ability to resolve all of the stress generated (Koster et al., 2005; Wang and Droge, 1996). Replication forks slow as they converge with an RNA polymerase, and this interference was proposed to be due to high levels of topological stress (Azvolinsky et al., 2009). In human cells, Top1 is important for relaxing positive stress in front of both DNA replication and transcriptional machinery (Baranello et al., 2016). In the absence of Top1, replication slowing and stalling and subsequent breaks, primarily occur at transcribed regions (Tuduri et al., 2009). Does topological stress, occurring due to transcription during DNA replication, lead to replication fork rotation?

In the case of a head-on replication–transcription collision, the positive supercoiling forming ahead of the replication fork and the positive supercoiling forming ahead of the transcription machinery, are likely to converge leading to high levels of topological stress. Furthermore, as the machineries come together, they could physically prevent the topoisomerases from binding to the supercoiled DNA, preventing relaxation via any other pathway than fork rotation. Co-directional collisions can occur when the polymerase is priming at the beginning of genes, or at genes of short repeats such as at the rDNA (Hamperl and Cimprich, 2016). In this situation, the RNA polymerase could act as a physical barrier to the diffusion of supercoiling in front of the replication fork, and again lead to the exclusion of topoisomerases from the DNA. In addition, DNA melting behind an RNA polymerase, that subsequently leads to the formation of non B-form DNA, could be a barrier to DNA replication and to the diffusion of positive supercoiling. For example, R-loops are known to hinder replication progression, causing fork stalling (Gan et al., 2011). Thus, since fork rotation occurs at places of replication fork stalling, this could be a potential situation for fork rotation to occur.

Therefore, both the RNA polymerase itself, and the supercoiling changes it generates, have the potential to lead to high topological stress during DNA replication. This chapter aims to understand the effects of DNA transcription on replication fork rotation during S-phase. I used a plasmid system in budding yeast with a direct assay for fork rotation, to analyse how much rotation occurs during one round of replication. By changing the numbers and orientation of transcriptional units on a plasmid, this system allowed me to determine fork rotation in the context of transcription.

3.2 ANALYSIS OF REPLICATION FORK ROTATION BY 2D GEL ELECTROPHORESIS AND SOUTHERN BLOTTING

To analyse fork rotation in *S. cerevisiae* I used an *in vivo* plasmid assay as previously described by Schalbetter et al. (2015). For every rotation of the replication fork one intertwine is moved from in front of the fork, to behind the fork forming one DNA pre-catenane. Therefore, the number of catenanes formed during S-phase is a direct readout of the number of rotations of the replication fork. Top2 is required for the resolution of DNA pre-catenanes, therefore a complete depletion is required to visualise the DNA catenanes. I depleted Top2 using a *top2-4* temperature sensitive allele, first isolated in a screen by (Holm et al., 1985) which has an amino acid change Pro820Gln. A depletion of Top2 prevents the resolution of catenanes, but still allows for some relaxation of supercoiling via Type I Topoisomerases.

To analyse fork rotation over one round of replication, a population of yeast cells were blocked in G1 phase using the alpha factor pheromone, followed by a *top2-4* mediated inactivation of Top2. The cultures were then released from G1 into S-phase, with a sustained depletion and blocked in the following G2 phase using the drug nocodazole; where samples were collected (materials and methods 2.6.7) (figure 3.1A). Cell cycle arrest and release was analysed by taking the budding index of the cultures and fluorescence activated cell sorting (FACS) samples at four time points through each timecourse (Figure 3.1A) (materials and methods 2.6.8). Using the *top2-4* allele the release from the G1 block is sometimes prevented in ~10% of the cell population.

After the G2 samples were collected, DNA was extracted and nicked to resolve supercoiling (materials and methods 2.7.3). DNA was then size resolved using gel electrophoresis, and run in a second dimension to resolve by geometry (Figure 3.1B, C) (materials and methods 2.7.4). Southern blotting with detection of the DNA by a cold probe was then used (materials and

methods 2.7.5, 2.7.6). Genomic DNA was also highlighted, but runs at different speeds to the plasmid DNA (Figure 3.1D). The relative signal of each catenated node, 1 - 27, was then quantified, and presented as a histogram (Figure 3.1D) (materials and methods 2.7.7). Above each of the histograms is specified the median of distribution, and the percentage of the population that has over 20 catenanes with the average deviation of this number.

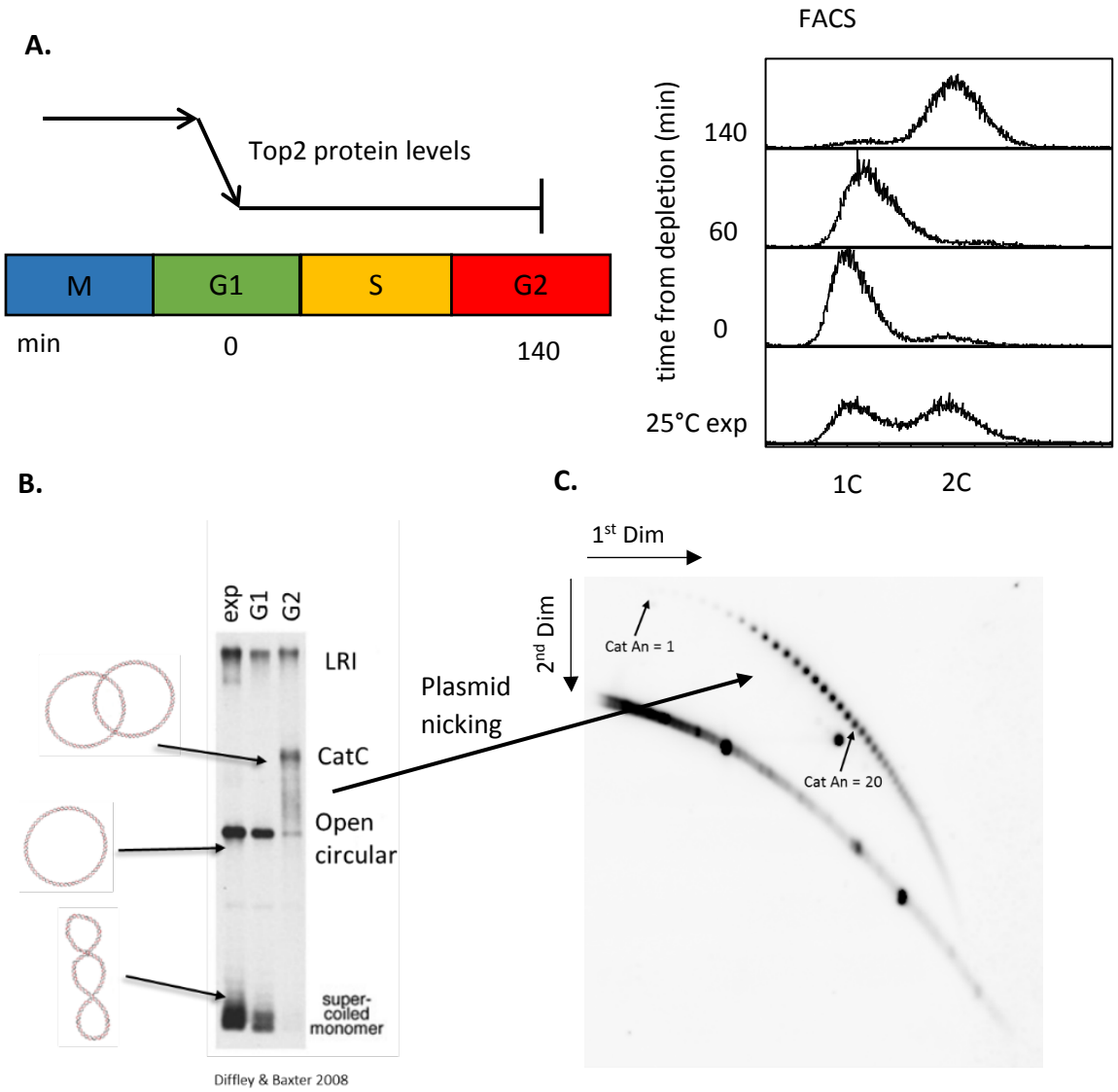

A.
Top2 protein levels
M
G1
S
G2
min
0
140
FACS
time from depletion (min)
140
60
0
25°C exp
1C
2C
B.
exp
G1
G2
LRI
CatC
Open circular
super-coiled monomer
Diffley & Baxter 2008
C.
1st Dim
2nd Dim
Cat An = 1
Cat An = 20
Plasmid nicking
D.
Region of signal from CatAn=1 to CatAn=20
Region of signal of CatAn=21 and above
genomic DNA
nicked monomer plasmid
% CatA signal
12
10
8
6
4
2
0
1 3 5 7 9 11 13 15 17 19 21 23 25 27
No. of catenated nodes

Figure 3.1 Analysis of replication fork rotation by 2D gel electrophoresis and Southern blotting

A. Schematic representation and FACS analysis of the procedure to deplete topoisomerase II during one round of replication. The cultures were blocked in G1 phase at time 0 min where Top2 was depleted. After 60 min the cultures were released and blocked in G2 phase by the addition of nocodazole, whereupon 10 ml samples were taken at 140 min. FACS shows peaks at 1C (compliments), indicating a G1 culture, 2C indicating a G2/M arrested culture or in-between, indicating S-phase. **B.** Taken from Baxter and Diffley (2008), a 1st dimensional gel showing the motilities of plasmids during the timecourse. In the exponential sample, there are late replicating intermediates (LRI), open circular monomer plasmids, and the fastest species is a range of supercoiled monomer plasmids. Top2 was still active preventing accumulation of catenated dimers. In the G1 sample, there are more monomer plasmids and less LRI. In the G2 sample there is a new slower running species which represents a range of catenated and supercoiled dimer molecules. **C.** The catenanes can be distinguished from one another by a nicking process that removes the supercoiling followed by running the gels in two dimensions. **D.** Quantification of the catenanes requires comparing the intensity of the signal for each catenane and to look at the total signal from CatAn=1 to CatAn=20 and for CatAn=21 and above. This can then be represented as a histogram of the percentage of the population against the number of catenated nodes. See materials and methods section 2.7.

3.3 Increasing the number of Transcriptional Units on a Plasmid Replicon Increases Replication Fork Rotation

Firstly, I wanted to understand if increasing the number of transcriptional units on a plasmid would increase the levels of fork rotation. To analyse this I chose to use a pRS315 plasmid with two additional housekeeping genes inserted; *SEC53* and *FMP32*. The pRS315 plasmid has an ARS, CEN 6, AmpR bacterial gene (not expected to be expressed), and the *LEU2* gene used for selection of the plasmid (Figure 3.2A). The direction of transcription of *LEU2* is towards the origin and centromere. I firstly needed to know the base level of fork rotation events that occurred using pRS315, without additional genes. This plasmid was analysed for levels of fork rotation in one round of replication, in *top2-4* cells, using the method outlined in Figure 3.1. The result was a distribution of catenanes with a peak at 11, and a slight tail towards the higher numbers resulting in a median of 12, and 11% of the population having over 20 catenanes (Figure 3.2A).

I then went on to understand whether adding genes to this plasmid would increase fork rotation events. I used the pRS315 plasmid with an addition of genomic DNA containing two genes from chromosome VI: 41735..44978. This plasmid was isolated by Dr Jon Baxter. This particular plasmid was chosen for two main reasons; firstly, both the genes *SEC53* and *FMP32* are housekeeping genes that are active throughout the cell cycle (Cherry et al., 2012). Secondly, these two genes are transcribing towards one another, and they transcribe with this directionality in the genome. This directionality gives the possibility of high levels of topological stress converging at the 3' ends of the genes, due to positive supercoiling forming ahead of each

of the transcription bubbles. High levels of topological stress have previously been seen to lead to high levels of replication fork rotation due to topoisomerase action being insufficient to fully resolve the intertwines (Schalbetter et al., 2015).

On the cartoon plasmid map (figure 3.2B), the pink gene represents *SEC53*, the Green gene represents *FMP32*, the blue section of DNA represents a 3' part of the *OTU1* ORF and finally the purple section of DNA represents a 3' section of the *RGD2* ORF, however neither the *OTU1* or *RGD2* sections of DNA involves their promoter (Figure 3.2B) (The full plasmid map is shown in Figure 3.3A). This convergent transcription plasmid showed a dramatic increase in the overall levels of fork rotation compared to pRS315 alone. The overall curve of the graph increased and flattened removing the peak at 11 catenanes, the median shifted up to 18, and 36% of the population had over 20 catenanes, compared to 11% with pRS315.

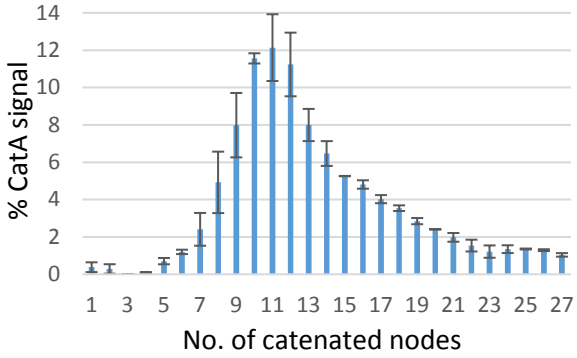
This large increase in DNA catenanes following DNA replication could be due to either of two scenarios: 1. Fork rotation occurs stochastically during replication, therefore the increase in plasmid size leads to longer replication fork elongation and more replication fork rotation or 2. Fork rotation occurs at specific areas of high topological stress; therefore, the additional section of DNA containing two extra genes causes a form of topological stress leading to more replication fork rotation. To understand which of these cases was leading to the large increase in fork rotation, I inserted a non-transcribed piece of DNA into pRS315 to increase the size from 6 kb to 8 kb. With this new plasmid there was no increases in fork rotation compared to pRS315 (Figure 3.2C). It had a normal distribution with median of 12 catenanes, and a percentage population over 20 catenanes of 8%, which is even a slight decrease on the pRS315 result of 11%, reflecting a slight sharpening of the peak (Figure 3.2).

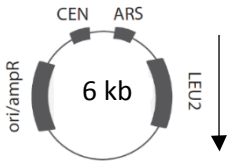
A.

top2-4 pRS315

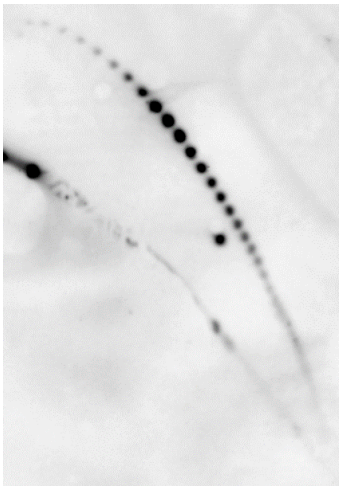
median of distribution: n = 12

% CatA >20 =11 +/- 3





Background plasmid

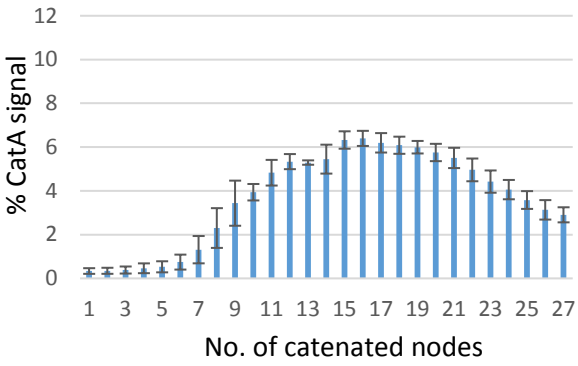


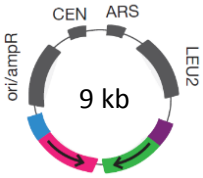
B.

top2-4 SEC53 FMP32 pRS315

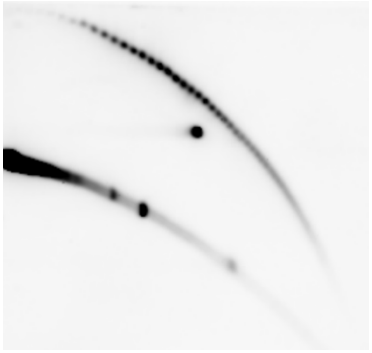
median of distribution: n = 18

% CatA >20 =36 +/- 5





Converging genes

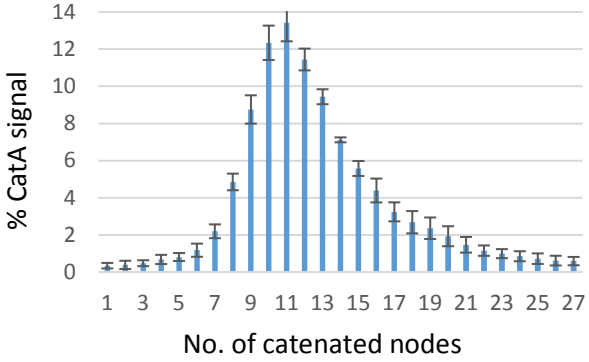


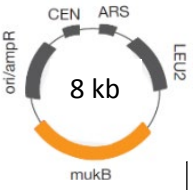
C.

top2-4 mukb pRS315

median of distribution: n = 12

% CatA >20 =8 +/- 4





Background plasmid with size control

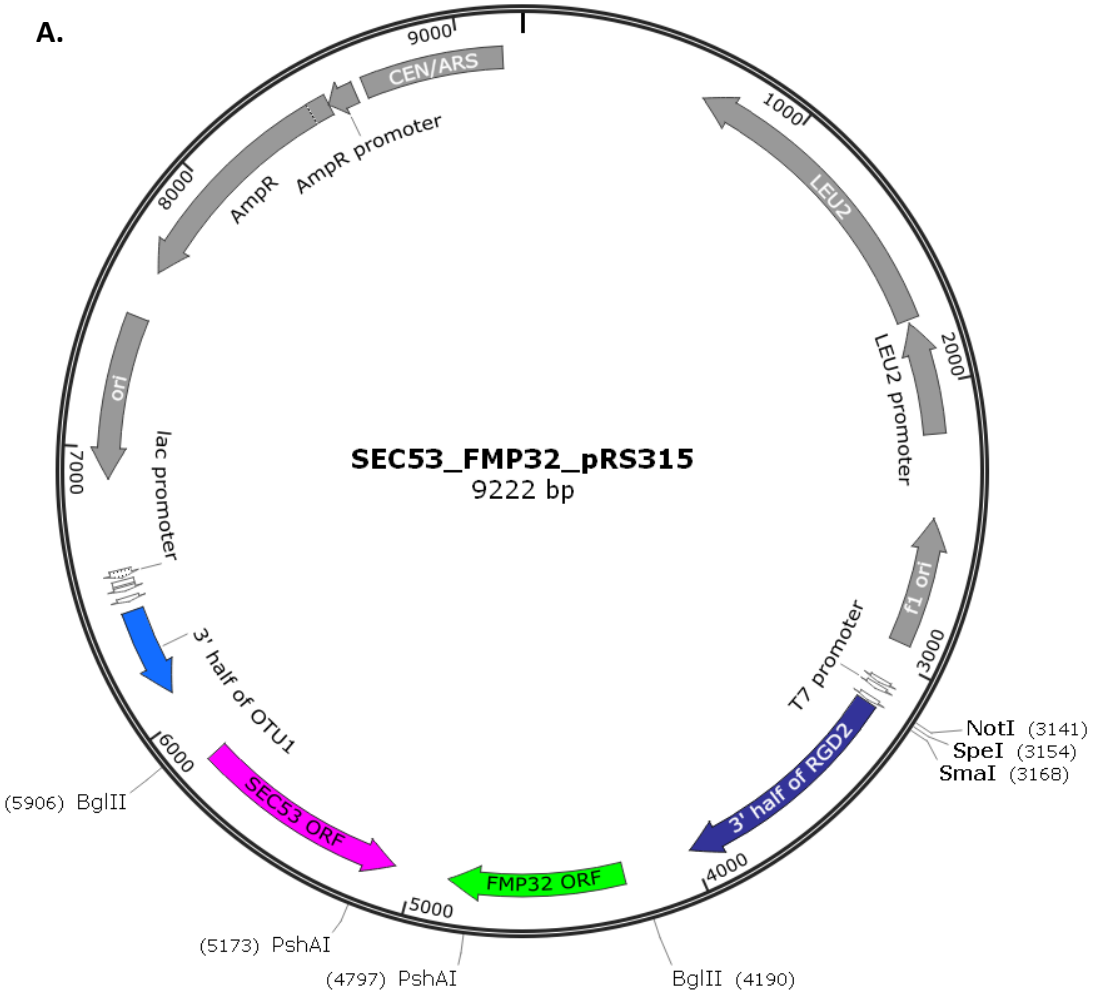
Figure 3.2 Additional genes on a plasmid increases replication fork rotation

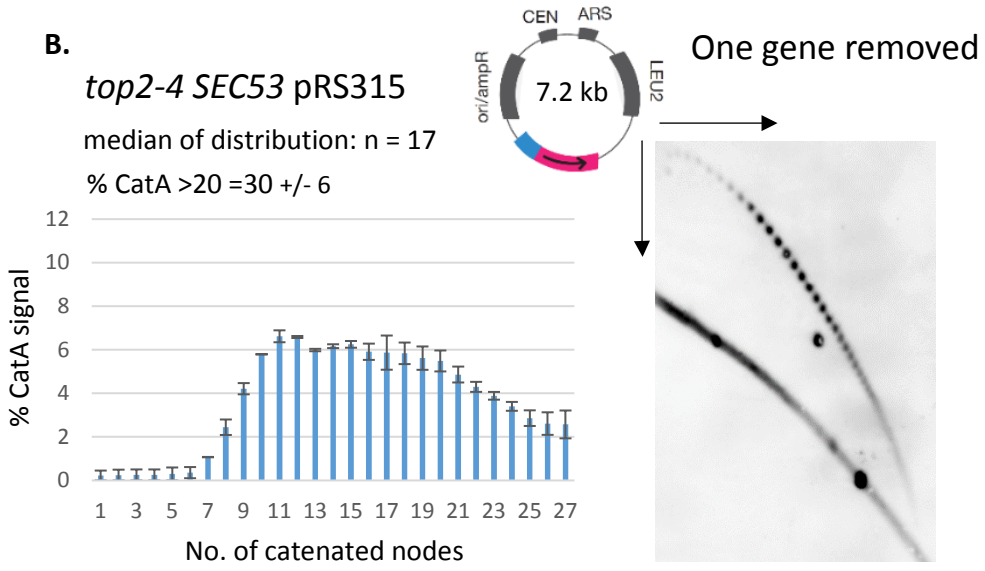
A. Analysis of fork rotation during S-phase with plasmid pRS315 6018 bp in a *top2-4* background. The analysis was done as in Figure 3.1. The histogram represents the average of 2 repeats and error bars represent the average deviation. **B**. Analysis of the pRS315 plasmid with an addition of *SEC53* and *FMP32* genes making the plasmid 9222 bp in size in a *top2-4* background. The histogram represents an average of 4 repeats. **C.** The size of a plasmid replicon makes no difference to fork rotation. Analysis of the pRS315 plasmid with an additional 2 kb of DNA inserted to make it 8120 bp in size in a *top2-4* background. The histogram represents the average of 3 repeats.

I further went on to confirm this result by making two modifications to the plasmid containing *SEC53* and *FMP32* (materials and methods 2.5). The first modification was to remove the *FMP32* gene including the 3' part of *RGD2* (Figure 3.3A). The second plasmid was made by removing both *SEC53* and *FMP32* with their promoters, and inserting a size control 2 kb section of *mukb* DNA (the same as used in Figure 3.2C) (Figure 3.3A).

Using the first new plasmid with *FMP32* removed, I aimed to see if the number of fork rotation events correlated with the number of genes on a plasmid. I found, compared to the plasmid with both genes intact, that there appeared to be a shift down to lower levels of catenanes in a subpopulation, where a new peak at around 12 catenanes can be seen (Figure 3.3B). The overall distribution drops slightly, falling from a median of 18 in the plasmid with two genes, to a median of 17 in the plasmid with one gene and likewise from 36% of the population having over 20 catenanes to 30% (comparison of Figure 3.2B with Figure 3.3B). However, fork rotation is still generally much higher than the original pRS315 plasmid, indicating that the *SEC53* gene alone appears to have a large effect on fork rotation.

The next experiment was to confirm the roles of the intact genes as oppose to the partial genes, in the fork rotation events. I removed both of the genes and their promoters to make a plasmid of size 7217 bp, followed by inserting a non-transcribed size control to increase the size to 9609 bp, making this the largest plasmid tested thus far. As expected the distribution decreases significantly with a clear return to a peak as oppose to a flatter distribution (Figure 3.3C) (statistical significance can be found for all results in Figure 3.6). The median for this transcriptional control plasmid is 13, and the percentage catenanes over 20 is 20%, although this is not completely back to control pRS315 levels of 12 catenanes and 11%, the peak of the distribution is again at 11 catenanes, with a larger tail. Overall, these results support the results from figure 3.2 by showing that the increase in fork rotation is a specific effect of the additional genes on the plasmid.


A.
SEC53_FMP32_pRS315
9222 bp
9000
CEN/ARS
AmpR promoter
AmpR
8000
1000
LEU2
LEU2 promoter
2000
ori
7000
lac promoter
3' half of OTU1
6000
(5906) BglII
SEC53 ORF
FMP32 ORF
5000
(5173) PshAI
(4797) PshAI
BglII (4190)
4000
T7 promoter
3' half of RGD2
f1 ori
3000
NotI (3141)
SpeI (3154)
SmaI (3168)


B.
top2-4 SEC53 pRS315
median of distribution: n = 17
% CatA >20 =30 +/- 6
CEN
ARS
ori/ampR
LEU2
7.2 kb
One gene removed
% CatA signal
12
10
8
6
4
2
0
1 3 5 7 9 11 13 15 17 19 21 23 25 27
No. of catenated nodes

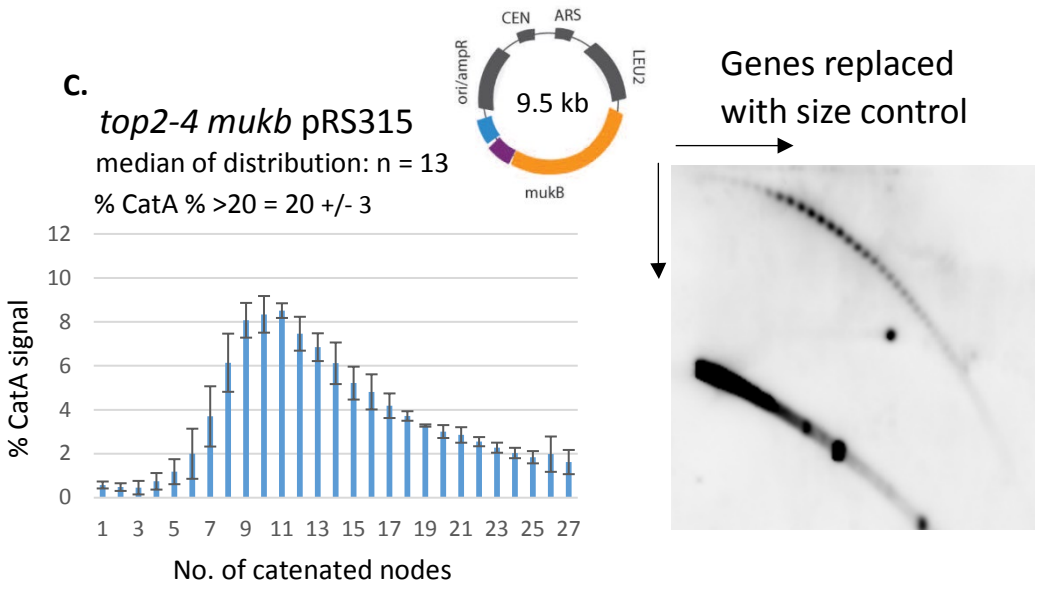


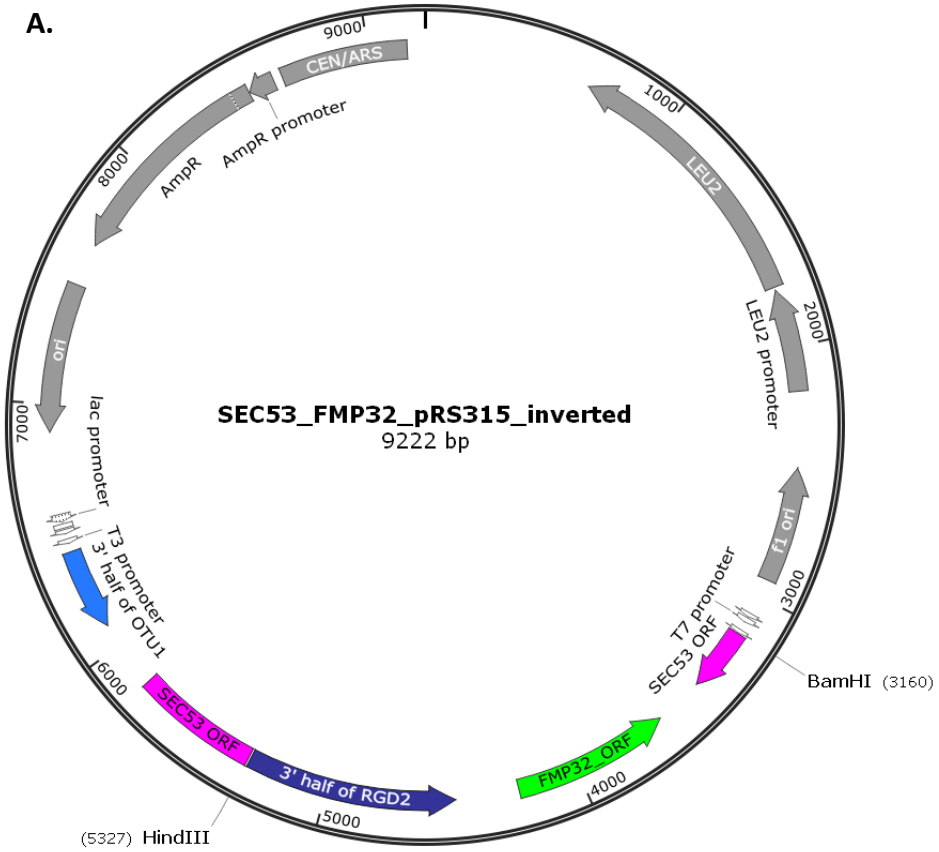
Figure 3.3 Removing genes from a plasmid decreases the number of fork rotation events

A. Plasmid map of *SEC53 FMP32* pRS315 made using the Software SnapGene Viewer. Restriction enzyme cutting points for the new plasmids are shown (materials and methods 2.5.3). PshaI and SmaI (NEB) were used to cut out *FMP32*, the smaller clone of which was isolated. BglII (NEB) enzyme was used to remove both of the genes. NotI and SpeI were the insert sites for the *mukb* DNA. **B.** Removing a gene from the plasmid decreases fork rotation. Analysis of the SEC53 pRS315 7217 bp plasmid in a *top2-4* background. The histogram represents an average of 2 repeats. **C.** Removing the two genes with promoters decreases fork rotation back to a distribution similar to pRS315 levels. Analysis of the plasmid with both *SEC53* and *FMP32* removed and *mukb* DNA inserted giving a plasmid of size 9608 bp in a *top2-4* background. The histogram represents an average of 3 repeats.

3.4 CONVERGING OR UNIDIRECTIONAL GENES HAVE THE SAME EFFECT ON FORK ROTATION

The plasmid that lead to the highest levels of fork rotation contained a pair of genes in a converging orientation. It is predicted that converging genes will have areas of high topological stress at their 3' ends, therefore I next strived to understand if the position of these genes relative to one another had an effect on fork rotation. I used the plasmid with *SEC53 FMP32* inserted and changed the directionality of the *FMP32* gene to give a new plasmid as outlined in figure 3.4A, with each of the *SEC53*, *FMP32* and *LEU2* promoters facing in the same orientation as one another (Figure 3.4A).

The orientation of the gene promoters appears to have little to no effect on fork rotation levels with this plasmid (Figure 3.4B). The distribution of catenanes when compared to the convergent transcription plasmid result (Figure 3.2B), has no change at all, the median decreases from 18 to 17, and the percentage of the population that has over 20 catenanes goes from 36% to 33%; a non-significant change (Figures 3.4B, 3.6). This indicates the orientation of the promoters on a plasmid has no great effect on fork rotation, as measured by this assay.


A.
SEC53_FMP32_pRS315_inverted
9222 bp
9000
CEN/ARS
AmpR promoter
AmpR
8000
1000
LEU2
LEU2 promoter
2000
ori
7000
lac promoter
T3 promoter
3' half of OTU1
f1 ori
3000
T7 promoter
SEC53 ORF
BamHI (3160)
6000
SEC53 ORF
3' half of RGD2
FMP32_ORF
4000
5000
(5327) HindIII

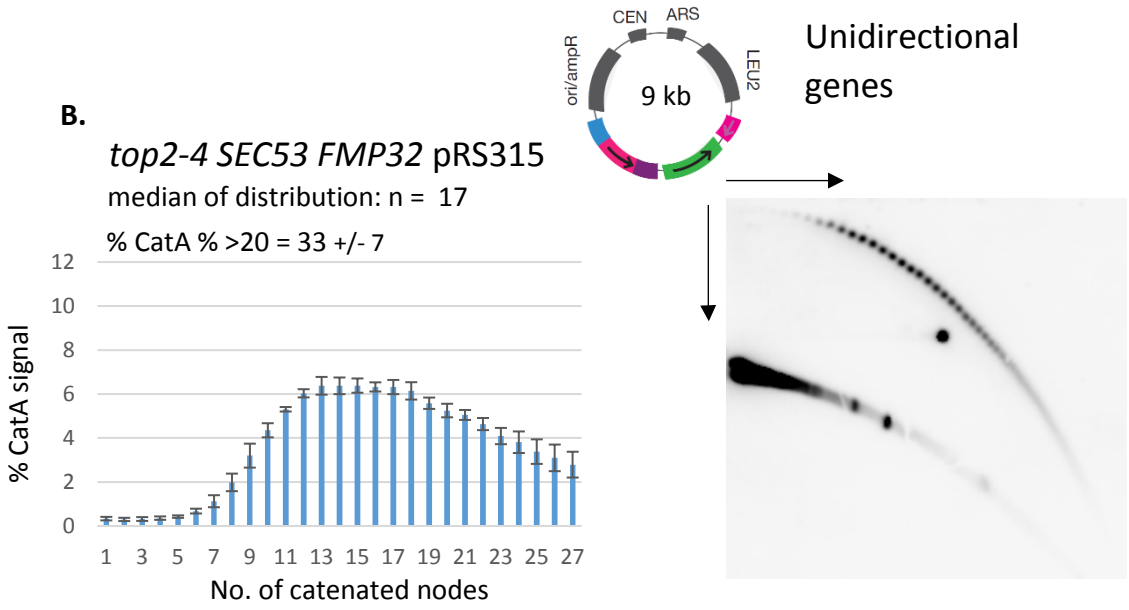

B.
top2-4 SEC53 FMP32 pRS315
median of distribution: n = 17
% CatA % >20 = 33 +/- 7
% CatA signal
No. of catenated nodes
0
2
4
6
8
10
12
1 3 5 7 9 11 13 15 17 19 21 23 25 27
CEN
ARS
ori/ampR
LEU2
9 kb
Unidirectional genes

Figure 3.4 Direction of transcription does not affect levels of fork rotation

A. Plasmid map of *SEC53 FMP32* pRS315 with all genes facing the same orientation made using Software SnapGene Viewer. Restriction enzymes HindIII and BamHI (NEB) were used to create the new plasmid. This section of DNA was removed, re-orientated and re-inserted into the vector. **B**. Unidirectional orientation of genes leads to high levels of replication fork rotation. The assay was carried out in *top2-4* background. The histogram represents an average of 4 repeats.

3.5 THE ROLE OF TOP3 IN RESOLVING DNA CATENANES

Saccharomyces cerevisiae Top3 is a Type 1A topoisomerase, which means it can cut a single strand of DNA to resolve supercoiling, or in the case of single stranded gaps, sister chromatid intertwines. For the results shown here to be valid and comparable to one another, there must be no resolution of intertwines during S-phase; this is achieved by a complete depletion of Top2. However, transcription promotes DNA underwinding behind the RNA polymerase, which could potentially allow Top3 access for resolution to negative supercoiling or pre-catenanes. Therefore, to rule out the involvement of Top3 I constructed a *top2-4 sgs1Δ top3Δ* strain. The deletion of *SGS1* is required in a *TOP3* mutant to prevent the slow growing and genome instability phenotypes of this deletion (Gangloff et al., 1994).

With this strain, I performed the same assay as outlined in figure 3.1 using two plasmids previously used in a *top2-4* background in figures 3.2B and 3.3C. These plasmids were: *SEC53 FMP32* pRS315 with genes in a convergent orientation, and the same plasmid with both of these genes removed and *mukb* size control inserted.

When the assay was repeated with the additional deletion of *TOP3* there were no significant changes to a Top2 depletion alone (Figure 3.5A and B, Figure 3.6). The statistics of the two experiments are almost identical which, as well as ruling out Top3 involvement, also indicates the reproducibility of this assay.

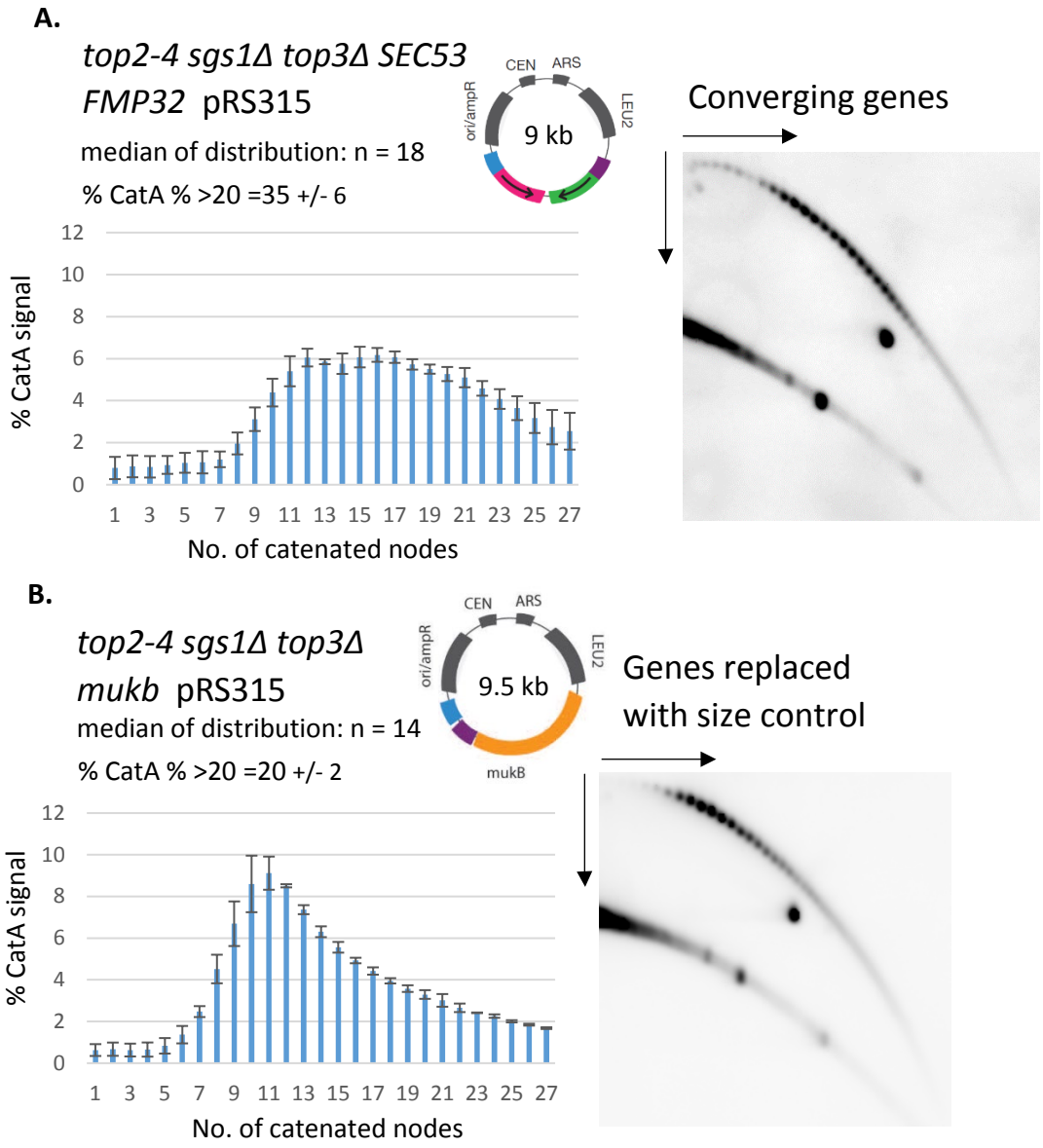
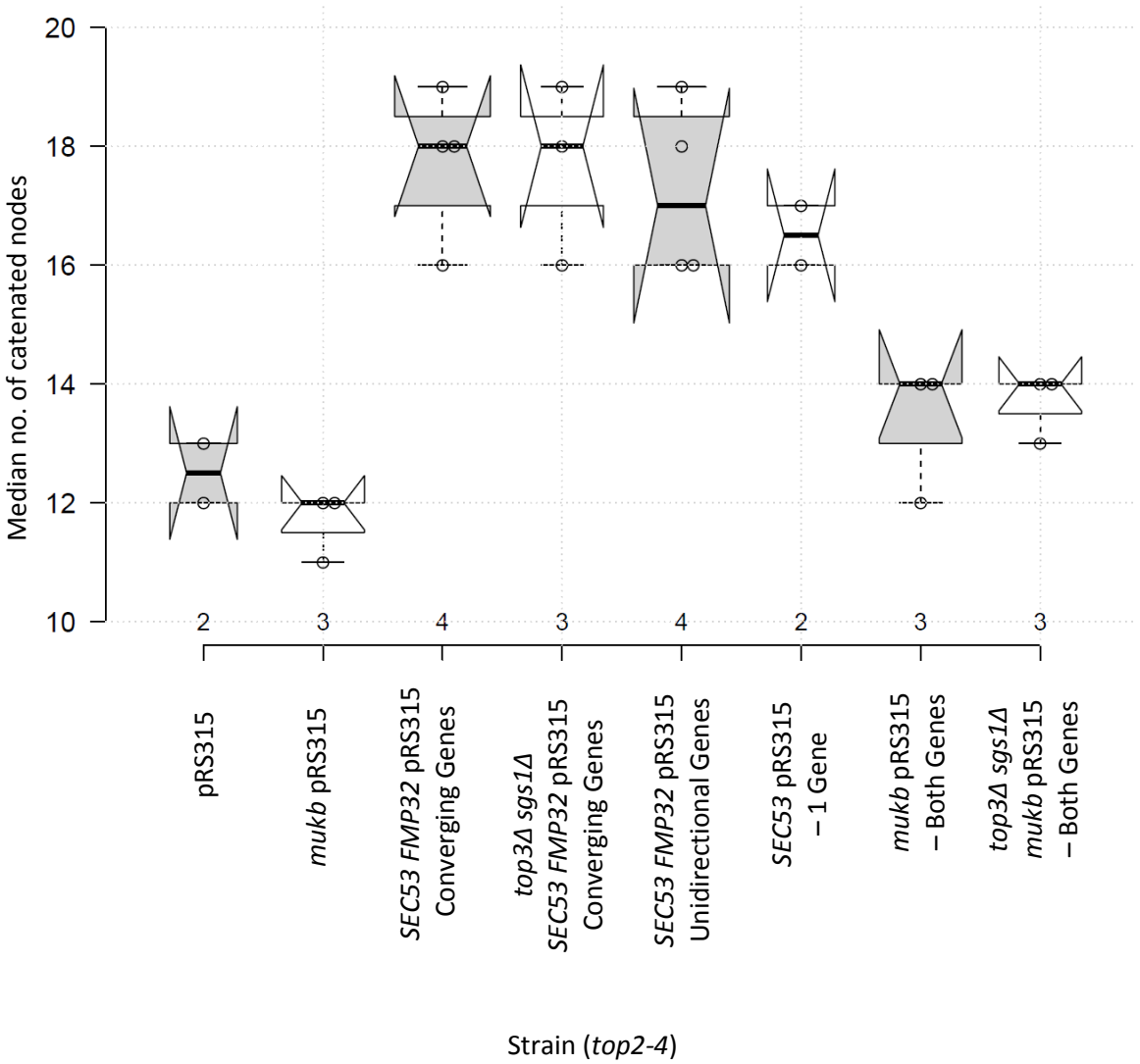


Figure 3.5 Top3 is not able to resolve catenanes in this assay

A. Using the converging genes plasmid there is no difference in catenation level in a *top2-4* background alone or *top2-4 top3Δ*. The histogram represents an average of 3 repeats. **B.** Using the control plasmid with genes removed and *mukb* inserted, there is no difference in catenation level in a *top2-4* background alone or *top2-4 top3Δ*. The histogram represents an average of 3 repeats.

A.


Median no. of catenated nodes
20
18
16
14
12
10
2
3
4
3
4
2
3
3
pRS315
mukb pRS315
SEC53 FMP32 pRS315 Converging Genes
top3Δ sgs1Δ SEC53 FMP32 pRS315 Converging Genes
SEC53 FMP32 pRS315 Unidirectional Genes
SEC53 pRS315 – 1 Gene
mukb pRS315 – Both Genes
top3Δ sgs1Δ mukb pRS315 – Both Genes
Strain (top2-4)

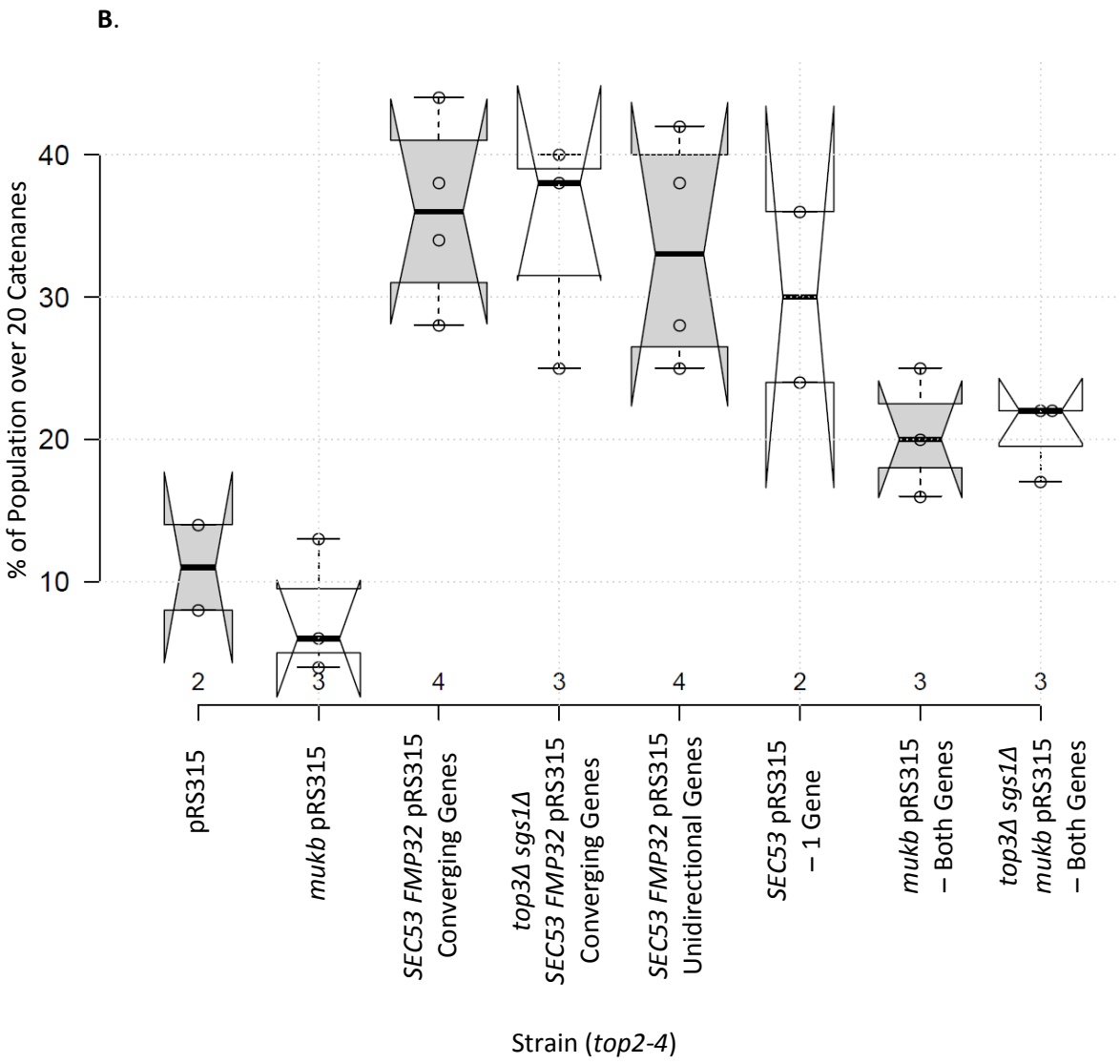


Figure 3.6 Summary Boxplots of Fork rotation Results

Made using BoxPlotR (Spitzer et al., 2014). The width of each box is proportional to the square-root of the number of observations. The number above each strain represents the number of repeats. The notches (narrowing at the median) are defined by:

$$\pm(1.58 \times \frac{IQR}{\sqrt{n}})$$

Where IQR represents the interquartile range, n represents the number of repeats, and 1.58 is a constant representing the confidence interval. Therefore this is a calculation of the confidence interval multiplied by the standard deviation of the median (McGill et al., 1978). Non-overlapping notches give roughly 95% confidence that two medians differ. **A.** Boxplots representing the median number of fork rotation events in each repeat used in this chapter. Strains are listed on the X axis with median number along the Y axis. All strains are in a *top2-4* background. **B**. Boxplots representing the percentage of each population that had over 20 catenanes for all repeats used in this chapter. Strains are listed on the X axis in the same order as in 3.6A, percentage of the population is represented on the Y axis. All strains are in a *top2-4* background.

3.6 DISCUSSION

3.6.1 DNA replication fork rotation during elongation

Fork rotation is well known to occur during the termination of DNA replication, however what is less well defined is, if and when it occurs during replication elongation. In a previous study Schalbetter et al. (2015) were able to show that fork rotation does indeed occur during elongation, at hard to replicate regions where static protein complexes impede progression of DNA replication. This study looked at situations of fork stalling known to be replicated with the use of accessory helicase Rrm3, which has also been implicated in resolving DNA:RNA hybrids (Gadaleta and Noguchi, 2017; Ivessa et al., 2003), and found that tRNA genes, inactive origins, and possibly centromeres increased rotation events. They then speculated that this was due to the stability of the protein-DNA complexes (Schalbetter et al., 2015). I have made use of this system to test the impact of replication-transcription collisions on replication fork rotation. I have been able to show that increasing transcriptional units on a plasmid and potentially increasing the frequency of collision events with a replication fork, increases the number of fork rotation events that occur.

Firstly, I analysed the normal level of fork rotation during one S-phase for a pRS315 plasmid (Figure 3.2A). The level of rotation occurring (median of 12 catenanes), is slightly higher than expected from one termination event alone. Using a 5.2 kb SV40 plasmid, a termination event was thought to lead to around 10 rotations of the fork (Sundin and Varshavsky, 1980, 1981). It is known that the centromere/kinetochore can possibly lead to some fork rotation events (Schalbetter et al., 2015), and another possibility is that the presence of the *LEU2* gene on this plasmid is leading to these rotation events. However, it appears that *LEU2* is only having a minimal effect, possibly reflecting a low level of transcription, due to cultures being grown in non-selective media. Unfortunately, the necessity of this gene for the plasmid selection in pre-cultures, prevents using a plasmid without the gene as a control.

I was then able to show a dramatic increase in fork rotation from pRS315 to a plasmid with two additional genes (*SEC53* and *FMP32*) (Figure 3.2). However, these two plasmids had a 3 kb size difference, meaning that if a rotation of the replication fork occurs after a certain number of base pairs, it is likely overall rotation number would increase in a larger plasmid. I was able to conclusively rule out larger plasmids leading to higher levels of fork rotation by increasing the size of both pRS315 and the plasmid with the additional genes removed (Figure 3.2C, 3.3C). I increased plasmid size by 2 kb and saw no increases in replication fork rotation. As well as this, Schalbetter et al. (2015) also reported seeing no increase in fork rotation with an increase in

plasmid size. The authors used the same assay and increased the size of a ARS/CEN 4.8 kb pRS316 plasmid to 8 kb and 12 kb. They saw no changes in the median fork rotation, although a spreading of the distribution (Schalbetter et al., 2015). I conclude from these results that fork rotation occurs at specific areas of high topological stress, as oppose to stochastically through elongation.

3.6.2 The impact of directionality of genes and build-up of topological stress

In this chapter I saw a high level of replication fork rotation occurring with a plasmid that had a convergent orientation of genes. This plasmid presumably had both a high likelihood of a replication-transcription collision, and high levels of topological stress at the 3' ends of the genes (Figure 3.2B). However when these genes were re-orientated to be in the same direction as one another, replication fork rotation levels remained high (figure 3.4B). This was unexpected because, with the genes facing in the same direction, it might be thought that there would be less likelihood of a replication-transcription collision. This is because the two machineries travel at the same pace, however if the RNA polymerase is priming at the beginning of a gene this would result in a collision. There also wouldn't be the possibility of having a build-up of topological stress at the 3' ends of genes. However, due to the closed nature of a circular plasmid, dissipation of DNA supercoiling will be limited on the plasmid replicon. Therefore, these experiments indicate, that the general number of genes on a plasmid during DNA replication is important for replication fork rotation. Removing one of the genes from the plasmid, decreased the level of replication fork rotation to an intermediate amount, however also lead to the formation of a binomial distribution (Figure 3.3B). This phenomenon could be caused either by having less transcriptional units on the plasmid, or by changing the distance between the replication origin and the gene. This could be tested in future by having plasmids with genes at different distances from the origin of replication.

3.6.3 Conclusions and future directions

In this chapter I have shown that adding transcriptional units onto a plasmid, in itself, increases replication fork rotation. I have seen that the orientation of these genes doesn't appear to effect fork rotation. I have also shown that Top3 is not able to resolve catenanes during normal replication elongation, even when higher than normal levels of fork rotation occur.

In the future, an alternative, and more conclusive way, of analysing the directionality of replication-transcription collisions could be to use a block to replication next to the origin on one side. I propose the plasmid with genes in one orientation could be analysed, allowing only one fork to proceed, and then compare the differences between each directionality. Furthermore, it might also be interesting to look at the effects of divergently orientated genes and R-loop formation. A recent study by Pannunzio and Lieber (2016) in *S. cerevisiae* looking at chromosomal breaks, found that convergent transcription did not increase chromosomal breaks compared to having only one gene on the chromosome. However, they saw a large increase in breaks with a divergent gene pair, which they concluded was due to the negative supercoiling domain between the pair leading to non-B form DNA structures. Therefore, in the future, the current assay could be expanded to include a pair of divergently orientated genes on a plasmid. It would also be interesting to specifically study if R-loop formation has any effect on replication fork rotation. In the current study, fork rotation appears unlikely to be due to R-loops, although it cannot be conclusively ruled out. R-loops are known to cover 8% of the budding yeast genome and cause severe impediments to DNA replication, however they generally occur at specific places in the genome, in particular at highly transcribed genes (Chan et al., 2014; Wahba et al., 2016). None of the three genes used here came up as particular sites for R-loop formation in the DRIP-seq (DNA:RNA hybrid immunoprecipitation) studies, although the known fork rotation site of tRNA was highly enriched.

Unfortunately, it cannot be distinguished if the topological stress leading to fork rotation arises specifically due to a sterical hindrance, or due to transcriptional helical stress. It is likely that both lead to fork rotation together; a block to the diffusion of supercoiling will cause a build-up, leading to high levels of topological stress. A caveat to this assay of course, is the lack of understanding around the transcription of the genes on the plasmid. As these yeast cells have both an endogenous copy of *SEC53* and *FMP32* and a copy on a yeast episomal plasmid it would be difficult to assay the level of transcription occurring on the plasmid. Potentially, Quantitative Reverse Transcription PCR (RT-qPCR) could be used to analyse the RNA expression but only if the endogenous copies of the genes are first deleted, and as they are both essential genes this is a difficulty. What would be useful, would be to look at other genes that have differing levels of transcription, to see if fork rotation levels are affected. However, even though we do not know the absolute levels of transcription by these genes in these experiments, even understanding this would not elucidate whether the fork rotation events are occurring because of the topological stress alone, because of the sterical block to the fork, or because of a mixture of the two.

4 COHESIN AS AN IMPEDIMENT TO THE DNA REPLICATION FORK

4.1 INTRODUCTION

In chapter 3, I have described how transcription during DNA replication leads to an increased frequency of fork rotation compared to unhindered DNA replication. In this chapter, I further explore this system, examining the role of the cohesin complex in transcription dependent replication fork rotation.

The cohesin complex is intimately linked with transcription, it is highly abundant at genes, and is thought to regulate gene expression by regulating the 3D organisation of DNA in the nucleus (Merkenschlager and Nora, 2016). There is considerable evidence in *S. cerevisiae* for the translocation of cohesin along DNA by the action of RNA polymerases pushing the complex (Glynn et al., 2004; Hu et al., 2011; Jeppsson et al., 2014; Lengronne et al., 2004; Racko et al., 2018). It has also been indirectly observed that cohesin may lead to the formation of excess catenanes. In a paper by Jeppsson et al. (2014), the authors predicted that sister chromatid intertwines are marked by the chromatin localisation of SMC5/6. They observed that a depletion of cohesin, often found at the ends of genes, leads to a depletion of SMC5/6, which they inferred was due to loss of sister chromatid intertwines. The results described in this thesis so far provide an excellent model system to test the hypothesis that cohesin promotes the formation of sister chromatid intertwines (SCI). Using this system, it can be explored, whether a promotion of SCIs is not only via their protection from topoisomerase activity, but also via promoting their formation.

From this knowledge of the link between transcription, cohesin, and DNA topology, I postulated that cohesin could have a role in the transcriptional dependent increase in replication fork rotation. In budding yeast, cohesin abundantly localises to the termination regions of convergent genes (Jeppsson et al., 2014). Cohesin is loaded at centromeres, as analysed by ChIP-seq (Hu et al., 2011), so cohesin will be in high abundance on these plasmids, where it will interact, and likely be translocated by, the actively transcribed genes contained within them.

Furthermore, it is not clear how the replication fork interacts with the cohesin complex when they converge during DNA replication. Some cohesin complexes are thought to be altered to enforce sister chromatid cohesion between the replicated chromatids. However, is this the only response when DNA replication and cohesin collide? Does the fork push cohesin in front of it? Does cohesin hinder the progression of the replication fork? How is it then removed and what

consequences must arise? In this chapter, I specifically look at the effects of cohesin on the replication fork. I use a system where replication-transcription collisions are expected, and could potentially be more potent, due to cohesin localisation and prevention of the removal of DNA topological stress.

4.2 COHESIN PROMOTES REPLICATION FORK ROTATION

I firstly tested if cohesin activity affects levels of replication fork rotation. To inactivate cohesin I used the *scc1-73* allele, which contains a point mutation allowing for inactivation by heat at 37°C (for strain genotypes see materials and methods 2.3). When combined with the *top2-4* allele, to prevent the resolution of DNA catenanes, numbers of fork rotation events after one round of replication could be analysed in various different plasmids. In this chapter I primarily compare two plasmids constructed in chapter 3; the *SEC53 FMP32* pRS315 plasmid, now named the convergent transcription plasmid. And the same plasmid with the two genes removed and a size control inserted (*mukb* pRS315 genes removed), now named the transcriptional control plasmid. To understand effects of differing transcription situations, I also analyse the unidirectional transcription plasmid (*SEC53 FMP32* pRS315), and a plasmid with tRNA genes inserted (3*tRNA* pRS316).

When cohesin and Top2 were inactivated over one S-phase, I found no significant changes in the level of replication fork rotation of the transcriptional control plasmid (which does not contain the *SEC53* and *FMP32* genes) compared to Top2 inactivation alone (Figure 4.1A, 4.9). Both the distributions looked similar, with peaks at 11 catenanes. Following cohesin inactivation the distribution of catenated states slightly sharpens at the peak which is reflected in the change of 13 to 12 catenanes and the percentage CatA greater than 20 changed from 20% down to 13%. Due to this insignificant difference, I therefore found that cohesin does not appear to generally affect replication fork rotation. Next, I went on to see if cohesin effects transcription-dependent replication fork rotation. Using the *scc1-73* allele, there was a significant decrease in fork rotation levels when analysing the convergent transcription plasmid (Figure 4.1B, 4.9). The distribution clearly shifts down to lower numbers of rotations which is reflected in the drop in median from 18 to 15 and a drop in the number of the population that has over 20 catenanes; from 36% to 21%. Cohesin is known to localise to areas of convergent transcription in budding yeast (Lengronne et al., 2004). I therefore tested if this loss of fork rotation was a specific convergent genes effect, or a more general transcriptional effect. I analysed the unidirectional genes plasmid and found that there was also a significant decrease in fork rotation when cohesin

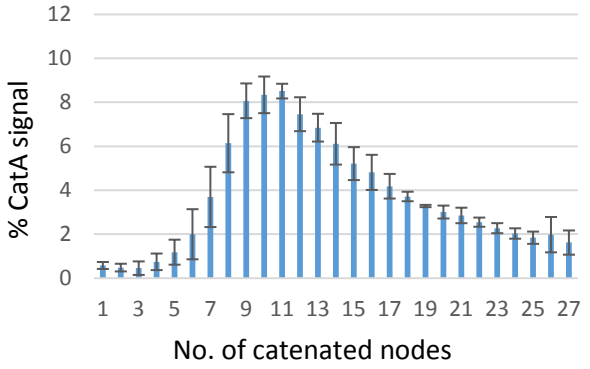
was inactivated (Figure 4.1C, 4.9). Both the median and population percentage over 20 catenanes were exactly the same as the convergent transcription result, 15 and 21% respectively (figure 4.1C). I conclude that cohesin activity generally leads to more frequent fork rotation on plasmids containing active genes.

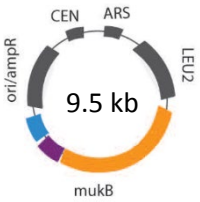
Ai.

top2-4 mukb pRS315

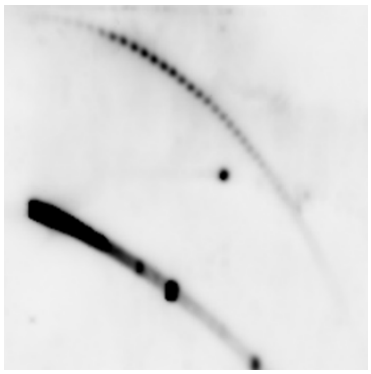
median of distribution: n = 13

% CatA % >20 = 20 +/- 3





Genes replaced with size control

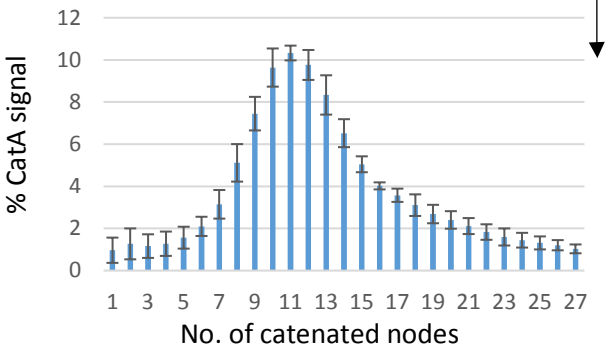


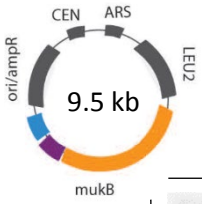
Aii.

top2-4 scc1-73 *mukb* pRS315

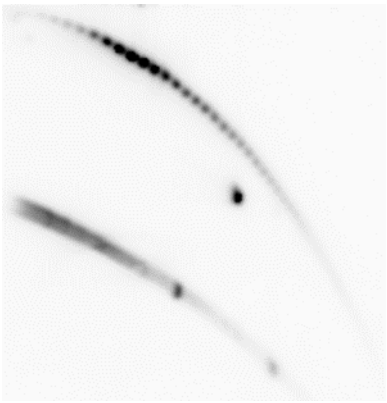
median of distribution: n = 12

% CatA >20 =13 +/- 2





Genes replaced with size control

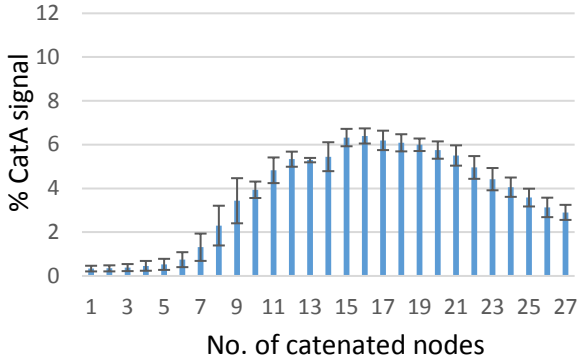


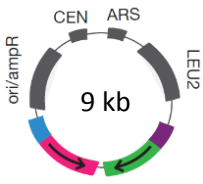
Bi.

top2-4 SEC53 FMP32 pRS315

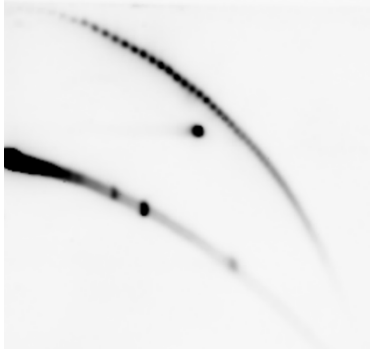
median of distribution: n = 18

% CatA >20 =36 +/- 5





Converging genes

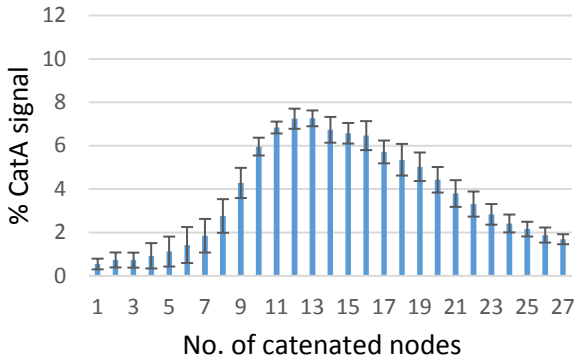


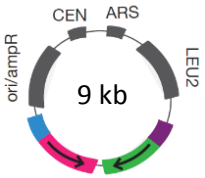
Bii.

top2-4 scc1-73 *SEC53 FMP32* pRS315

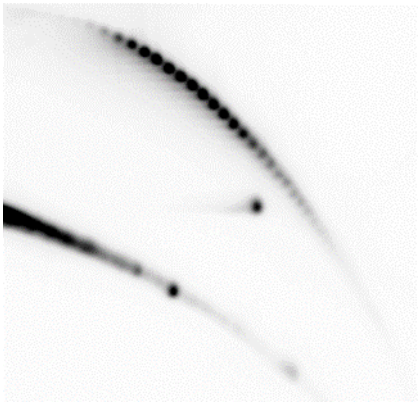
median of distribution: n = 15

% CatA >20 =21 +/- 4





Converging genes

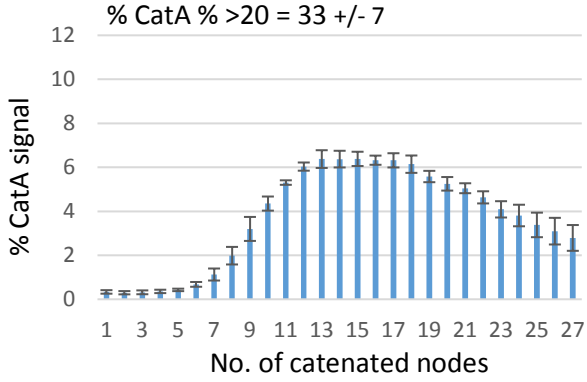


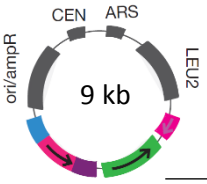
Ci.

top2-4 SEC53 FMP32 pRS315

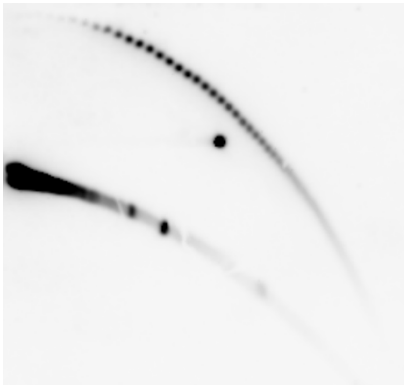
median of distribution: n = 17

% CatA % >20 = 33 +/- 7





Unidirectional genes

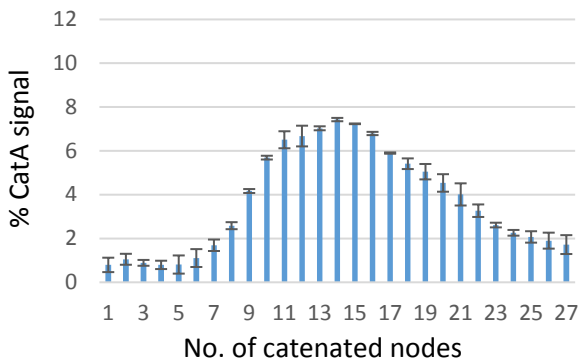


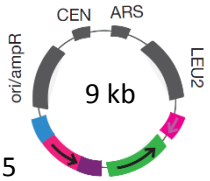
Cii.

top2-4 scc1-73 *SEC53 FMP32* pRS315

median of distribution: n = 15

% CatA >20 =21 +/- 3





Unidirectional genes

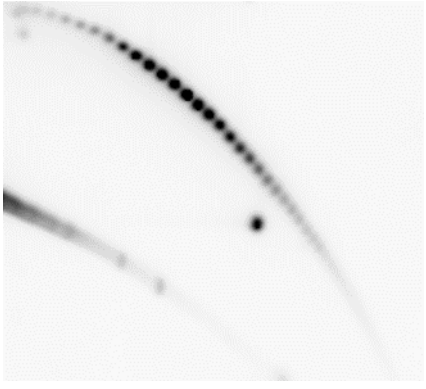


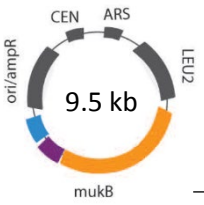
Figure 4.1 Cohesin depletion leads to a decrease in replication fork rotation when additional genes are present on a plasmid

DNA catenation assay in a *top2-4* background with **A.** 9608 bp transcriptional control plasmid **Ai**. Active cohesin (repeat of Figure 3.3D) **Aii.** Inactive cohesin, the histogram represents 4 repeats **B.** 9222 bp convergent transcription plasmid **Bi**. Active cohesin (repeat of Figure 3.2B) **Bii**. Inactive cohesin, the histogram represents 4 repeats **C.** 9222bp unidirectional transcription plasmid **Ci**. Active cohesin (repeat of Figure 3.4B) **Cii**. Inactive cohesin, the histogram represents 2 repeats.

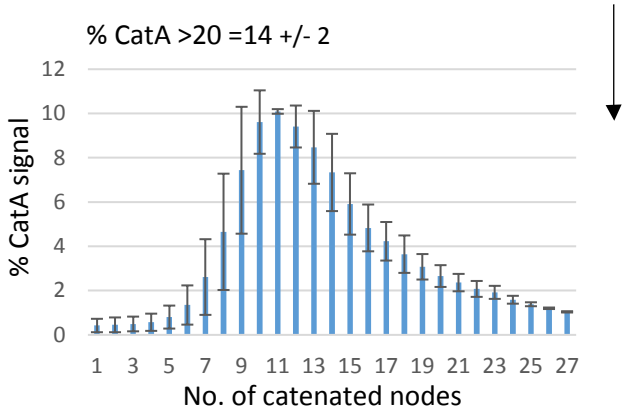
One potential issue with these data is that the *scc1-73 top2-4* stain may have adapted to not completely ablating Top2 activity at the restricted temperature. In order to rule out any possibility of residual Top2 activity in the *scc1-73 top2-4* strains causing the reduced levels of detected catenanes in these assays, I next sampled each population at two time points; once straight after replication (50 min post release) and once after a further hour in a nocodazole block (110 min post release). If residual Top2 activity was resolving the catenanes, there would be a decrease in the numbers of catenanes between the first and the second time points. However, I found no change between the two time points in either of the two plasmid replicons tested (Figure 4.2). Therefore, I can conclude the decrease in fork rotation is solely due to the inactivation of cohesin as oppose to any residual Topoisomerase II activity.

A.

top2-4 scc1-73 *mukB*
pRS315
median of distribution: n = 12

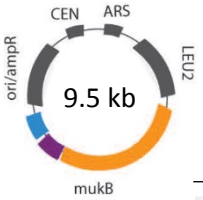


50 min post alpha factor release

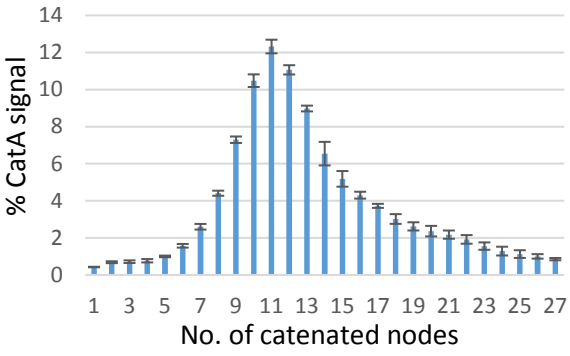


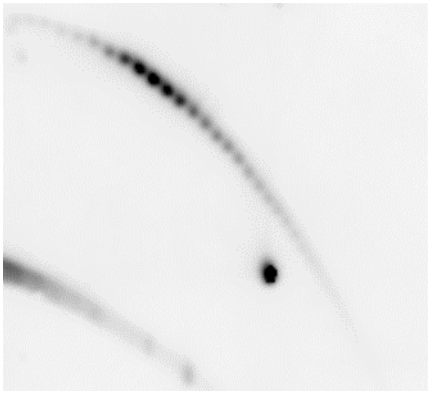
B.

top2-4 scc1-73 *mukB*
pRS315
median of distribution: n = 12



110 min post alpha factor release





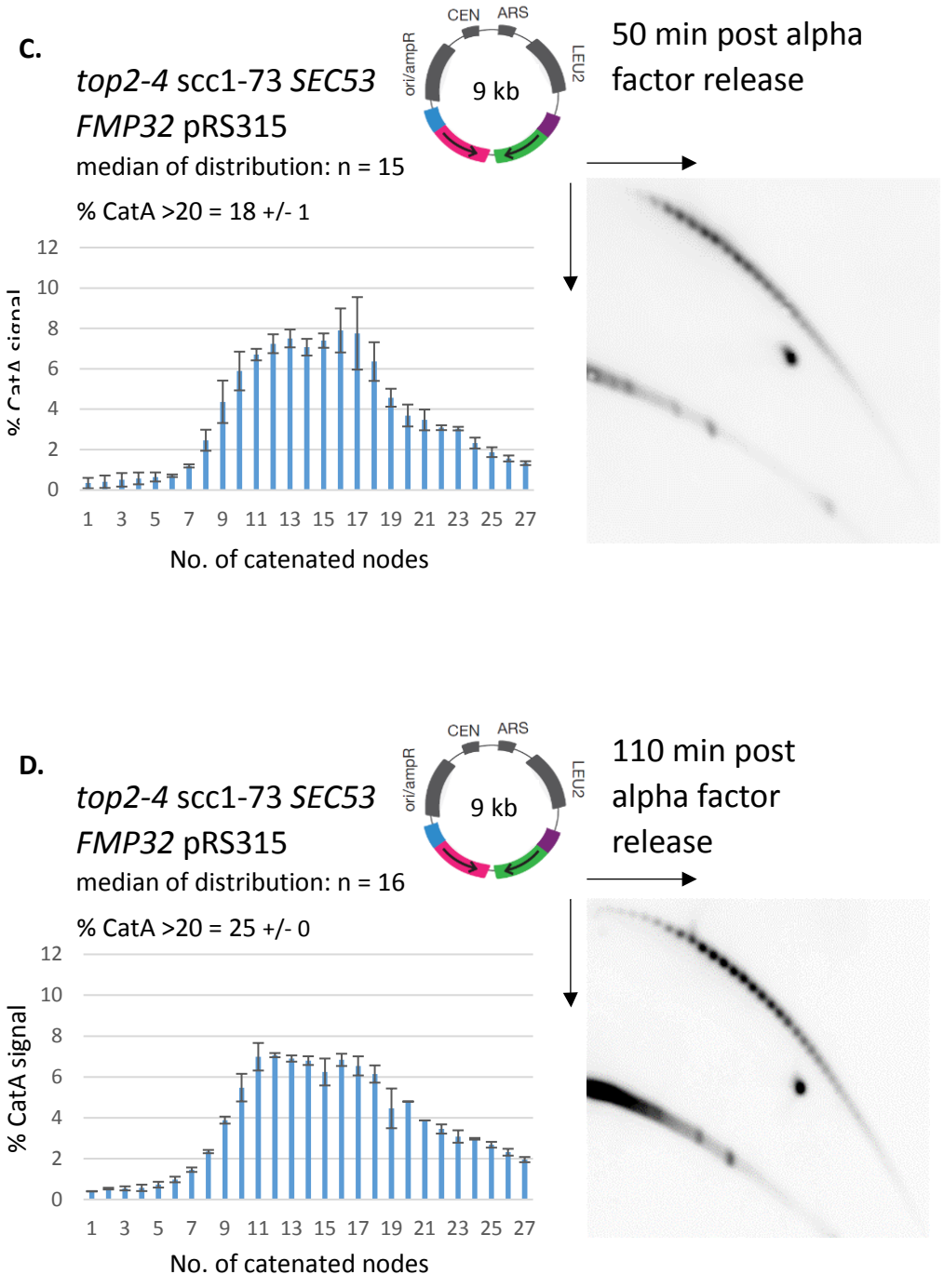


Figure 4.2 Residual Top2 activity is not affecting the assay

A comparison of catenation levels 50 min post G1 release, **A** and **C**, versus 110 min post G1 release **B** and **D**, with the convergent transcription plasmid **A** and **B** or the transcriptional control plasmid, **C** and **D**. Each of the histograms represents two repeats.

To confirm the results gained from inactivation of cohesin, I used two separate approaches to remove cohesin activity from the tested plasmids. The first method I used involved an alternative way of preventing cohesin activity during S-phase. It also was a control for the slow growth phenotype seen at permissive temperatures with *scc1-73 top2-4* cells. I took advantage of a system based on inducible repression of gene expression, the GaLL promoter (Janke et al., 2004), to specifically repress SCC1 expression before DNA replication. GALL is a truncated version of the highly transcribed GAL1 galactose inducible promoter, and has a lower general gene expression activity than GAL1 (Janke et al., 2004). In galactose, this helps prevent overexpression of the gene it is controlling relative to the gene's endogenous expression (Mumberg et al., 1994). In glucose, the GALL promoter completely supresses the expression of the targeted gene (Mumberg et al., 1994). The timecourse used here involved growth of the cultures in galactose until the alpha factor block where glucose was added to repress transcription of *SCC1* (materials and methods 2.6.7). I again analysed the transcriptional control plasmid, and the convergent transcription plasmid. The results were very similar to those using the *scc1-73* allele. There were no clear decreases in fork rotation levels following repression of *SCC1* when analysing the transcriptional control plasmid (Figure 4.3A). The median was 13, the same as when cohesin was active, and the level of the population with over 20 catenanes was 14% down from 20% (Figure 4.3A). However, the *SCC1* repression prior to DNA replication led to a large drop in fork rotation levels with the convergent transcription plasmid (Figure 4.3B). The median of fork rotation decreased from 18 to 16 and the percentage population with over 20 catenanes, was down from 36% to 23%.

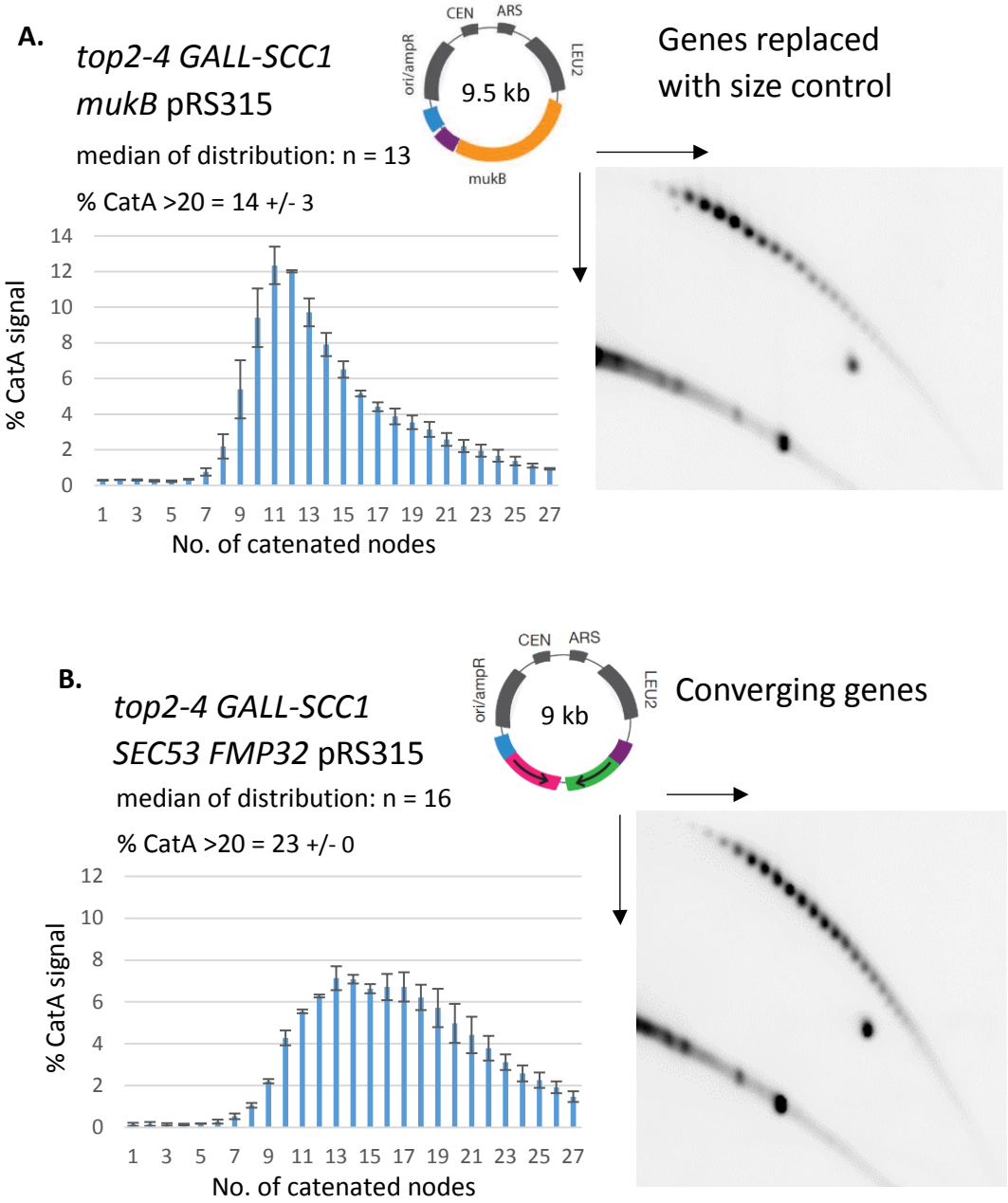


Figure 4.3 Transcription dependent fork rotation is promoted by cohesin

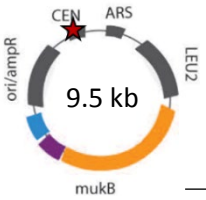
DNA catenation assay in *GALL-SCC1 top2-4* cells with **A**. Transcriptional control plasmid, the histogram represents 2 repeats. **B**. Convergent transcription plasmid, the histogram represents 2 repeats.

The second method used to confirm that cohesin activity on the plasmid was promoting replication fork rotation was to prevent cohesin loading onto the plasmids. Cohesin loads abundantly at the centromeres in budding yeast (Hu et al., 2011), therefore inactivating the centromere on the plasmid should prevent most cohesin loading onto the plasmid. This method would also allow the cells to grow normally without any adverse effects from an inactivation of cohesin. I inactivated the plasmid centromere 6 by two methods, firstly making a double point mutant at bp 4-A and bp 5-T in the CDE III region (based on a paper by Jehn et al. (1991)) (see primers used in section 2.4). Secondly, preventing the assembly of the kinetochore by inactivating the inner kinetochore protein Ndc10 using a temperature sensitive mutation (Ma et al., 2012). Each of the transcriptional control and convergent transcription plasmids was analysed in these conditions. When these CEN mutated plasmids were analysed, I found that they produced similar results to a complete cohesin inactivation (Figure 4.4A, B). With a centromere mutation, the transcriptional control plasmid had a median of 12 and a percentage population over 20 catenanes of 14% (Figure 4.4A), which was almost exactly the same as the *scc1-73* mutant result of 12 and 13%, respectively. The convergent transcription centromere mutant plasmid had a large decrease in fork rotation as compared to the intact centromere result (Figure 4.4B compared to 4.1Bi). The distribution dropped and had a sharper peak, with a median of 16, and population number over 20 catenanes of 22% (Figure 4.4B). The *scc1-73* inactivation had a median of 15 and population number over 20 catenanes of 21%, a very similar result (Figure4.1Bii) (See Figure 4.9 for full comparisons).

However when Ndc10 was downregulated there were no discernible changes compared to when the kinetochore was intact. The cell population for the transcriptional control plasmid had a median of 13 catenanes and a percentage population over 20 catenanes of 14% (Figure 4.4C). The cell population for convergent transcription plasmid had a median of 18 catenanes and a percentage population over 20 catenanes of 38% (Figure 4.4D). Which was very similar to when the kinetochore was intact; a median of 18 and a percentage population over 20 catenanes of 36% (Figure 4.1Bi).

These differences between the methods for inactivating the centromere, point towards a residual accumulation of cohesin in the Ndc10 mutant cells. Before Ndc10 is downregulated, the cells are held for a long period in a G1 arrest. It is possible that the levels of cohesin that accumulate on the DNA during this arrest are sufficient for the phenotype seen.

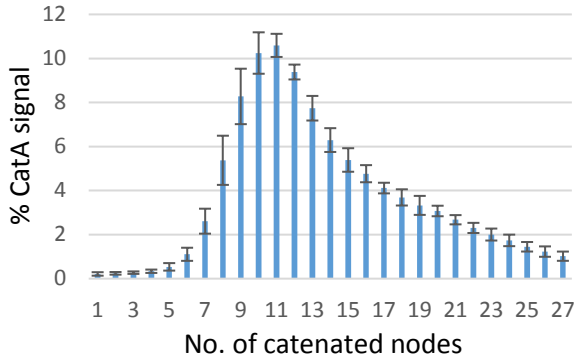
A. *top2-4 mukb* pRS315 cenmut



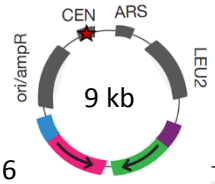
Centromere point mutant

median of distribution: n = 12

% CatA >20 = 14 +/- 2



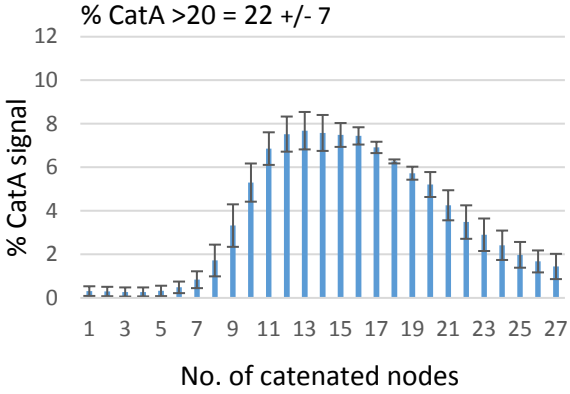
B. *top2-4 SEC53 FMP32* pRS315 cenmut

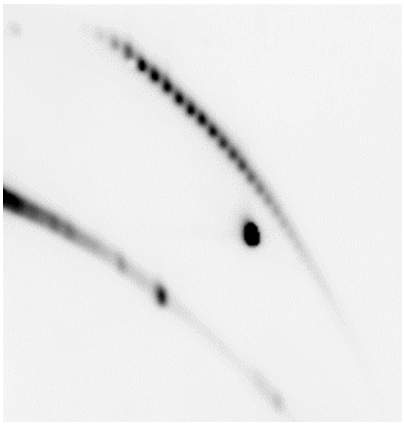


Centromere point mutant

median of distribution: n = 16

% CatA >20 = 22 +/- 7





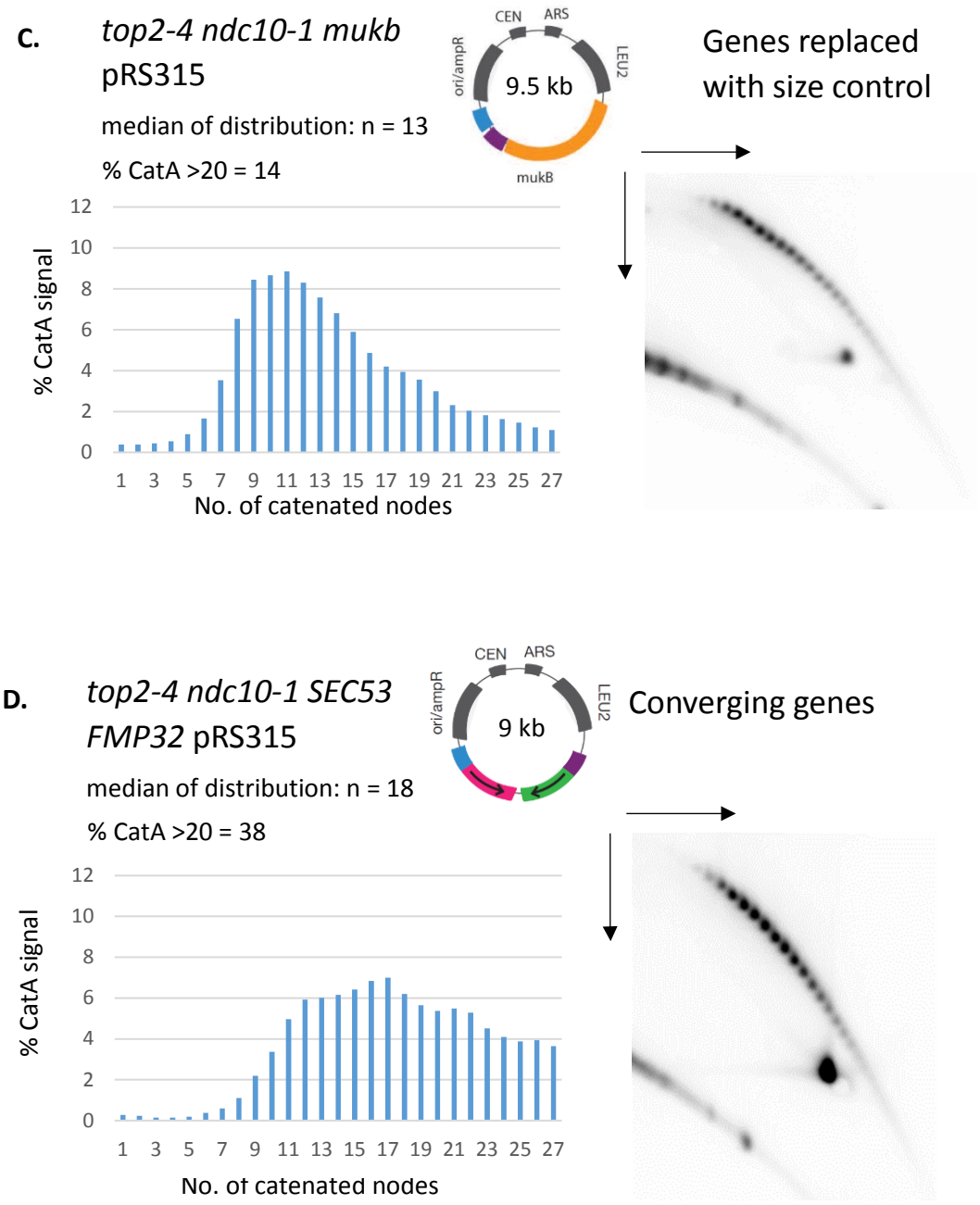


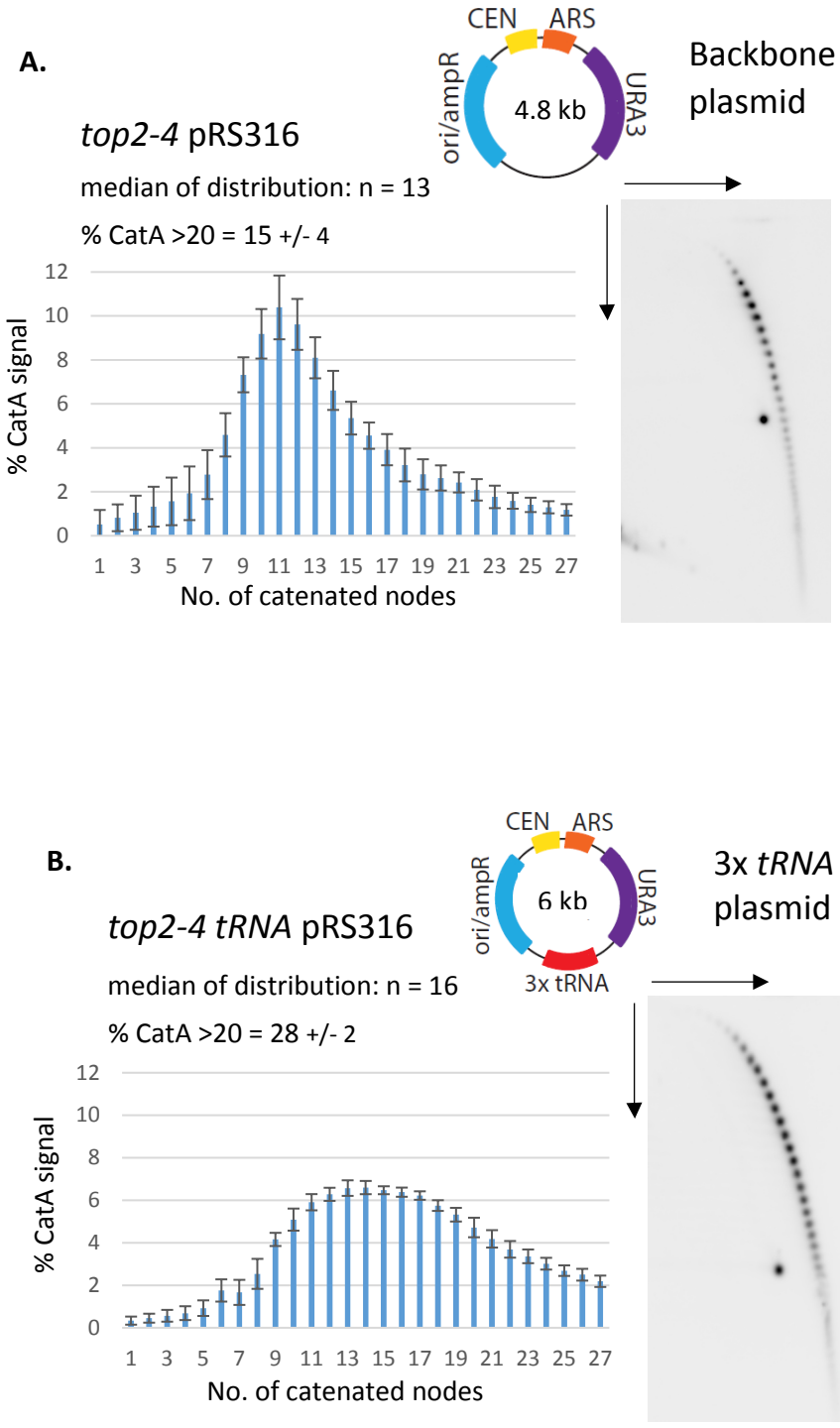
Figure 4.4 Inactivation of the centromere through the cell cycle causes a transcription dependent decrease in replication fork rotation

DNA catenation assay with **A.** Transcriptional control plasmid containing a centromere point mutant, the histogram represents 4 repeats. **B.** Convergent transcription plasmid containing a centromere point mutant, the histogram represents 4 repeats. **C.** *top2-4 ndc10-1* cells with the transcriptional control plasmid, the histogram represents 1 experiment. **D.** *top2-4 ndc10-1* cells with the convergent transcription plasmid, the histogram represents 1 experiment.

4.3 *tRNA* GENES POTENTIALLY HAVE A COMPLEX EFFECT ON COHESIN ACTIVITY

From the combined results of Figures 4.1 – 4.4, I concluded that cohesin activity does affect fork rotation on a plasmid containing active genes, and indicate that this affect required cohesin to be specifically loaded onto the plasmid replicon. I next wanted to see if cohesin promotes replication fork rotation at short transcripts, and therefore examined a plasmid with 3 *tRNA* genes. The addition of *tRNA* genes on a plasmid increases the frequency of fork rotation during DNA replication, which is thought to be due to a physical block to the replication fork (Schalbetter et al., 2015). *tRNA* genes also appear to increase the recruitment of the cohesin loader Scc2 (D'Ambrosio et al., 2008). However, when cohesin was inactivated I found there was a complex but inconsistent effect (Figure 4.5C).

For reference, the backbone plasmid (pRS316) and plasmid with 3 *tRNA* motifs are shown, with permission from Schalbetter et al. (2015), in Figures 4.5A and 4.5B. When cohesin was inactivated using the pRS316 3x *tRNA* plasmid, I saw no change in the median of 16 or the percentage of cells with over 20 catenanes. However, the distribution looked different, with a range of medians between repeats, giving a binomial distribution (figure 4.5C). *tRNA* genes are potentially having a complex effect on cohesin activity, however the variability in the experiments means I cannot draw any firm conclusions at this point.


A.
CEN
ARS
ori/ampR
URA3
4.8 kb
Backbone plasmid
top2-4 pRS316
median of distribution: n = 13
% CatA >20 = 15 +/- 4
% CatA signal
No. of catenated nodes
B.
CEN
ARS
ori/ampR
URA3
6 kb
3x tRNA
3x tRNA plasmid
top2-4 tRNA pRS316
median of distribution: n = 16
% CatA >20 = 28 +/- 2
% CatA signal
No. of catenated nodes

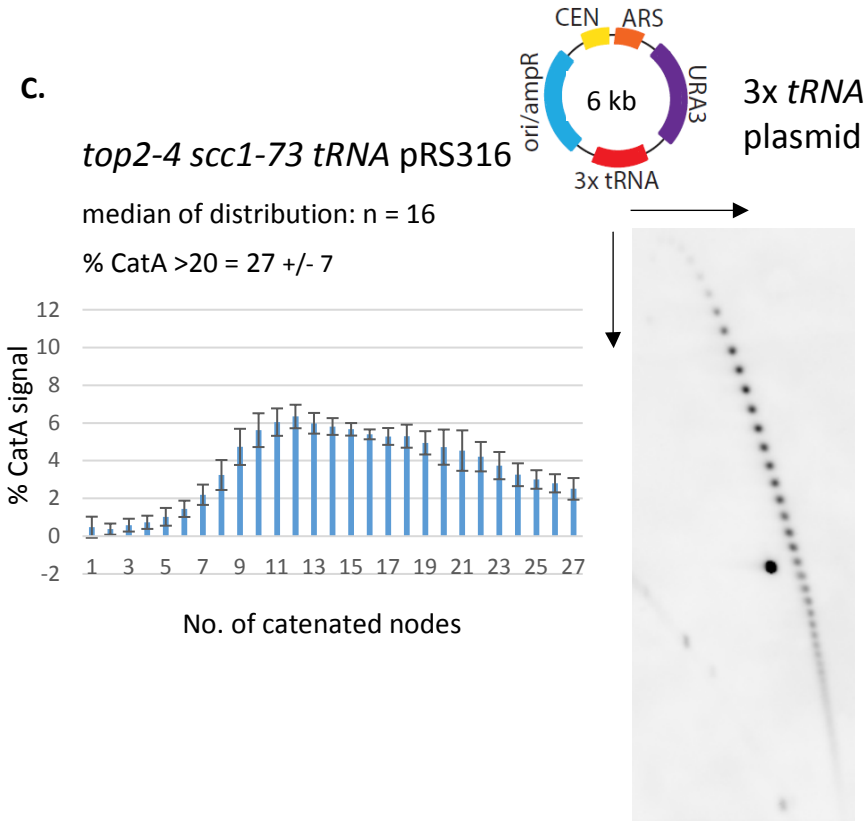


Figure 4.5 Cohesin inactivation has varying effects on fork rotation at *tRNA* genes

A. Taken from Schalbetter et al. (2015), catenation assay with plasmid pRS316 in a *top2-4* background, the histogram represents 5 repeats. **B.** Taken from Schalbetter et al. (2015), catenation assay with plasmid pRS316 3*tRNA* in a *top2-4* background, the histogram represents 7 repeats. **C.** Catenation assay with plasmid pRS316 3*tRNA* in a *top2-4 scc1-73* background, the histogram represents 6 repeats.

4.4 COHESIN ACTION IS NOT ALTERING THE SUPERCOILING STATE OF PLASMIDS GOING INTO S-PHASE

Following on from this, I next aimed to understand what specific effect cohesin was having on DNA topology. Potentially the effect could be due to the combination of cohesin and transcription acting as a topological block to DNA replication. In this case topological stress would only arise transiently as the replisome converges on the transcription cohesin complex. Alternatively, cohesin action could be constitutively altering the DNA topology of the plasmid. For example, maintaining or protecting a constitutively negatively supercoiled state could help RNA polymerase binding to promoters, similar to a protection of SCI's. However, the change in rotation events could be directly due to the decrease in active cohesin leading to a decrease in transcription. This would affect the supercoiling state of the plasmid going into S-phase, and subsequently change the rate of DNA replication fork rotation. To differentiate between these possibilities of transient versus constitutive DNA topological stress on the plasmid I went back to using the transcriptional control and convergent transcription plasmids and analysed their DNA supercoiling in the presence and absence of cohesin.

To look at levels of DNA supercoiling going into S-phase, the timecourse was shortened so that samples were taken 20 min after the release from G1 phase (figure 4.6A). Depletion of Top2 is started at 0 min when cultures were blocked in alpha factor, 60min after the start of depletion the cultures were washed from alpha factor and at 80 min the samples were taken. The FACS shows that the cells had not gone through S phase. Figure 4.6B shows the same 1D gel from Figure 3.1B, however, for supercoiling analysis, the section of the gel being looked at was the supercoiled monomer plasmids which were not nicked to remove the supercoiling. In this case chloroquine intercalation was used in the gels to separate supercoiling by sign (+/-). This allowed for a separation of the supercoiled plasmids over two dimensions, according to how positively or negatively supercoiled they were (Figure 4.6C) (See materials and methods 2.8 for full protocol). The supercoiled forms are less easily resolvable than catenanes, however there is still a range of running species visible, with the most highly negatively supercoiled running the furthest (figure 4.6C).

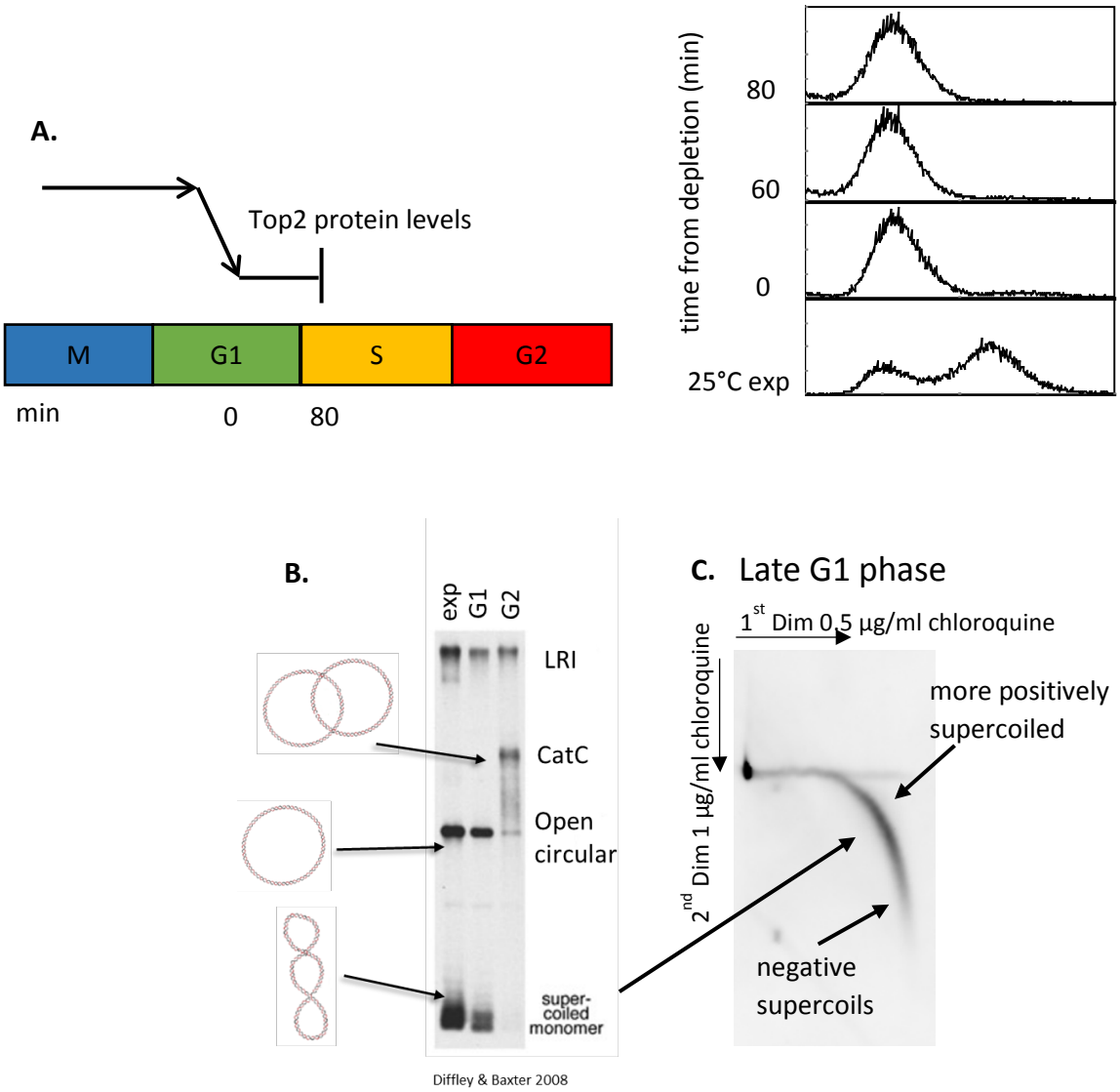


Figure 4.6 Experimental setup to analyse supercoiling levels by 2D gel electrophoresis and southern blotting

A. Supercoiling timecourse and FACS analysis, cultures were blocked in G1 phase at time 0 min where the temperature was increased from 25°C to 37°C, which leads to depletion of the temperature sensitive alleles. After 60 min the cultures were released from the G1 block and samples were taken at 80 min before the cells have gone through S-phase. **B**. A 1st dimension gel Taken from Baxter and Diffley (2008) showing the motilities of plasmids during a timecourse. In this case cultures are harvested as they enter S-phase and the supercoiled monomer plasmids are run in a second dimension. **C.** 2nd dimension gel with chloroquine intercalation to allow for distinguishing plasmids by sign. In this case the further run plasmids are more negatively supercoiled.

Using this assay I firstly analysed the differences in supercoiling level between the two plasmids (transcriptional control and convergent transcription). I found that the supercoiling levels were very different, with the convergent transcription plasmid in a more negatively supercoiled state than the transcriptional control plasmid (Figure 4.7A and B). I then went on to analyse the effect of cohesin on the convergent transcription plasmid. However, following inactivation of cohesin, I saw no major changes to the supercoiling state (Figure 4.7C). This argues that cohesin action is not constitutively altering the DNA topology, or the transcriptional state of the plasmids.

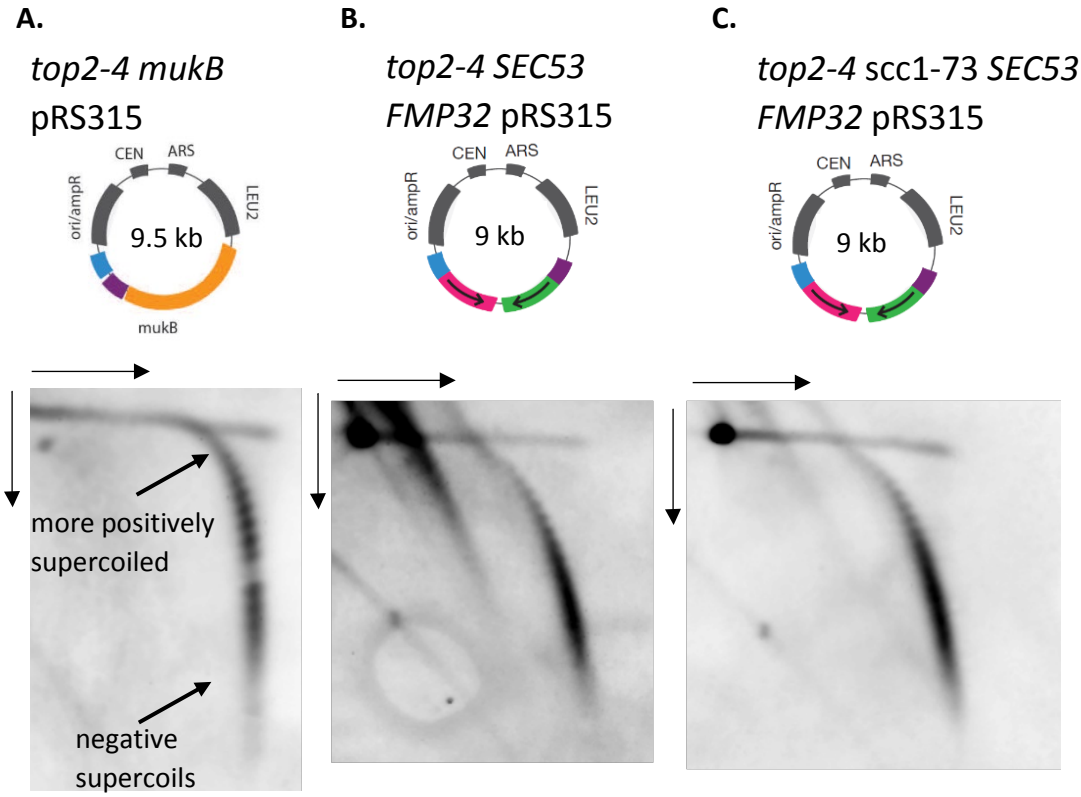


Figure 4.7 Representative blots of plasmid supercoiling state going into S-phase

Analysis of plasmid supercoiling, the first dimension was run with 0.5 µg/ml chloroquine, the second dimension was run with 1 µg/ml chloroquine. *top2-4* cells with the **A.** transcriptional control plasmid, one out of three repeats shown. **B.** convergent transcription plasmid, one out of four repeats shown. **C.** convergent transcription plasmid with a cohesin inactivation, one out of three repeats shown.

4.5 THE EFFECT OF ACCESSORY PROTEINS TO COHESIN

To better understand the effects of cohesin, I next aimed to elucidate whether the topological changes were a particular effect of stably bound cohesive cohesin, or cohesin molecules on DNA in general. I analysed the effects of downregulating both Eco1 and Rad61. Eco1 is known to acetylate cohesin on the Smc3 residues K112 and K113, preventing the removal of cohesin by Rad61 (Wapl) (Rolef Ben-Shahar et al., 2008; Rowland et al., 2009; Unal et al., 2008; Zhang et al., 2008). This function is specifically associated with sister chromatid cohesion, and it allows for long residency times of the cohesin complex on DNA, from S-phase through to cleavage at anaphase onset (Borrie et al., 2017). The removal of Rad61 has more contrary effects in yeast. Rad61 acts to remove non-acetylated cohesin from the DNA and therefore its downregulation should lead to increased residency of cohesin on DNA. In this vane, its deletion rescues the null phenotype of an *ECO1* deletion (Rowland et al., 2009). However, there have been a number of reports showing that a deletion of *RAD61* actually leads to sister chromatid cohesion defects (Rowland et al., 2009). It has also been seen that a *rad61Δ* leads to lowered levels of *SCC1* transcription (Chan et al., 2012; Sutani et al., 2009).

To analyse the effects of an Eco1 depletion I used the temperature sensitive mutant *eco1-1* (Toth et al., 1999) because an *ECO1* deletion is lethal (Skibbens et al., 1999). Analysing the catenation profile of an S-phase Eco1 depletion, I found no major changes in fork rotation for either of the plasmids used (transcriptional control and convergent transcription) (Figure 4.8A, B). This indicates that sister chromatid cohesion is not sufficient for promoting replication fork rotation. This result lead on to the hypothesis that a *RAD61* deletion, preventing the removal of non-acetylated cohesin, would lead to higher levels of fork rotation. However, this was not the case, the convergent transcription plasmid showed a decrease in fork rotation levels with a deletion of *RAD61* (Figure 4.8C). The distribution of catenanes was similar to that of a full cohesin depletion, with a broad peak around 12 catenanes. The median result using *top2-4 rad61Δ* cells was at 16 as oppose to 15 with a *scc1-73* inactivation, and the percentage of the population with plasmids of over 20 catenanes was 30% and 21% respectively (Figure 4.8D and 4.1A). When the transcriptional control plasmid was analysed in the same *top2-4 rad61Δ* cells there was a slight decrease from *top2-4* alone, but no significant changes to the fork rotation levels (Figure 4.8C, 4.9). The median number of catenanes was at 12 and amount of the population with over 20 catenanes was 14%, whereas *top2-4* alone had a median of 13 and percentage of 20% (Figure 4.8C, 4.1Ai). It is a possibility that the *rad61Δ,* which leads to lowered levels of *SCC1*

transcription, effects the amount of cohesin that binds to the plasmid DNA (Chan et al., 2012; Petela et al., 2018; Sutani et al., 2009).

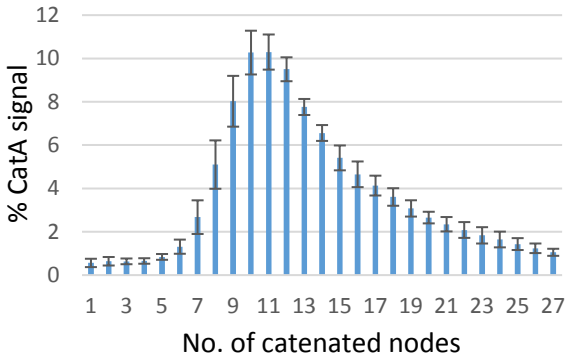
These two contradictory results lead on to the final experiment of using a *top2-4 eco1Δ rad61Δ* strain to look at the fork rotation events using both plasmids; transcriptional control and convergent transcription. A deletion of both *ECO1* and *RAD61* supposedly balances out the effects of either deletion alone, leading to the restoration of sister chromatid cohesion (Rowland et al., 2009). *ECO1* deletion was used in this instance, because a deletion of *RAD61* combined with a deletion of *ECO1,* rescues the null phenotype seen in the *ECO1* delete cells (Rowland et al., 2009). The distributions of catenanes for both of the plasmids looked very similar to both the *top2-4* alone and the *top2-4 eco1-1* backgrounds (Figure 4.8E, F). Indicating that a further deletion of *ECO1* on top of *RAD61* can rescue the lowered level of fork rotation in that background. It was also previously observed that this double deletion restored *SCC1* transcription levels to normal (Chan et al., 2012).

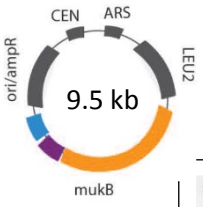
A.

top2-4 eco1-1 *mukB* pRS315

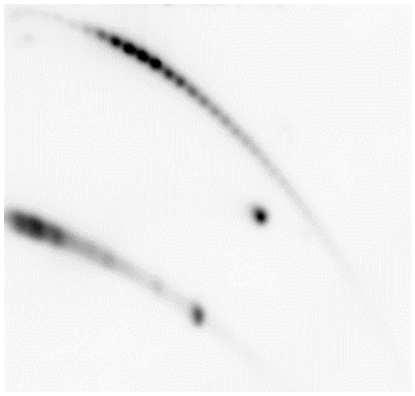
median of distribution: n = 13

% CatA >20 =14 +/- 2





Genes replaced with size control

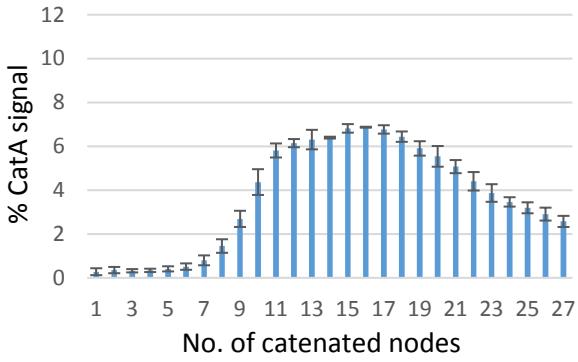


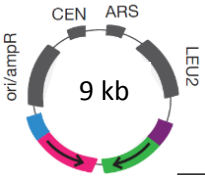
B.

top2-4 eco1-1 SEC53 FMP32 pRS315

median of distribution: n = 17

% CatA >20 = 30 +/- 2





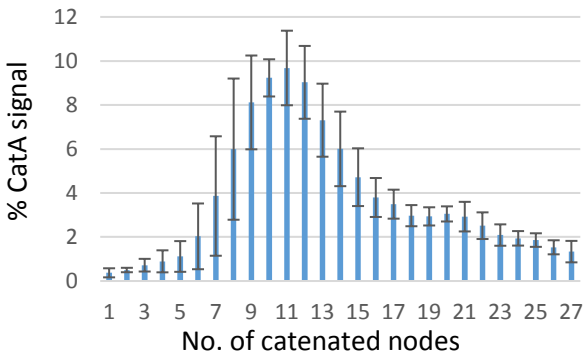
Converging genes

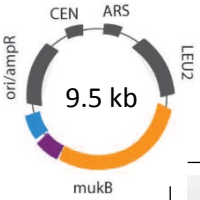
C.

top2-4 rad61Δ mukb pRS315

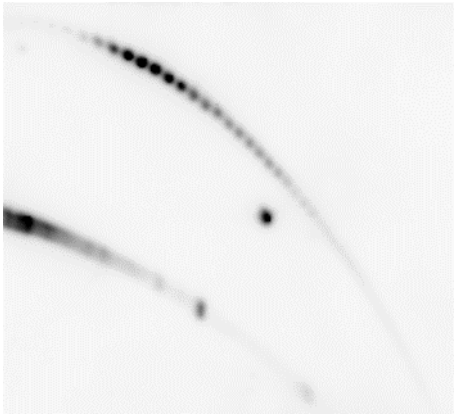
median of distribution: n = 12

% CatA >20 = 14 +/- 7





genes replaced with size control

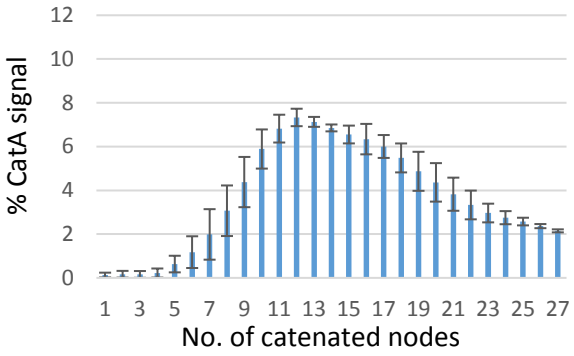


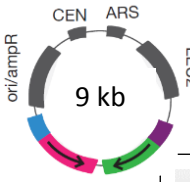
D.

top2-4 rad61Δ SEC53 FMP32 pRS315

median of distribution: n = 16

% CatA >20 = 30 +/- 3





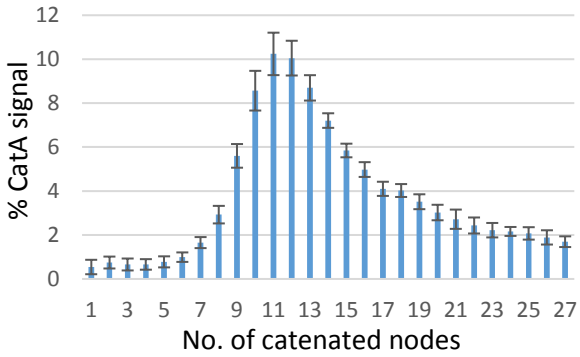
Converging genes

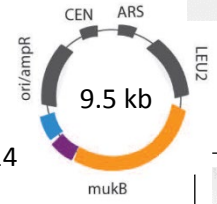
E.

top2-4 rad61Δ eco1Δ mukb pRS315

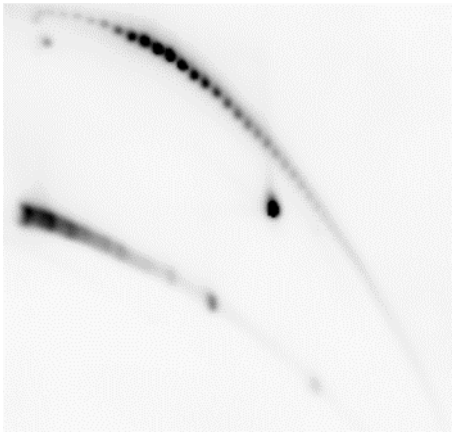
median of distribution: n = 14

% CatA >20 = 23 +/- 3





genes replaced with size control

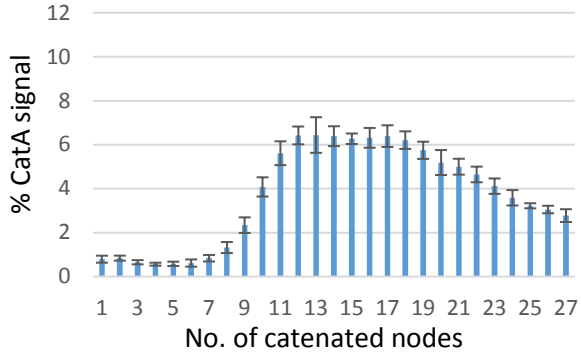


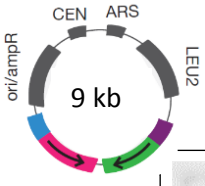
F.

top2-4 rad61Δ eco1Δ SEC53 FMP32 pRS315

median of distribution: n = 18

% CatA >20 = 39 +/- 4

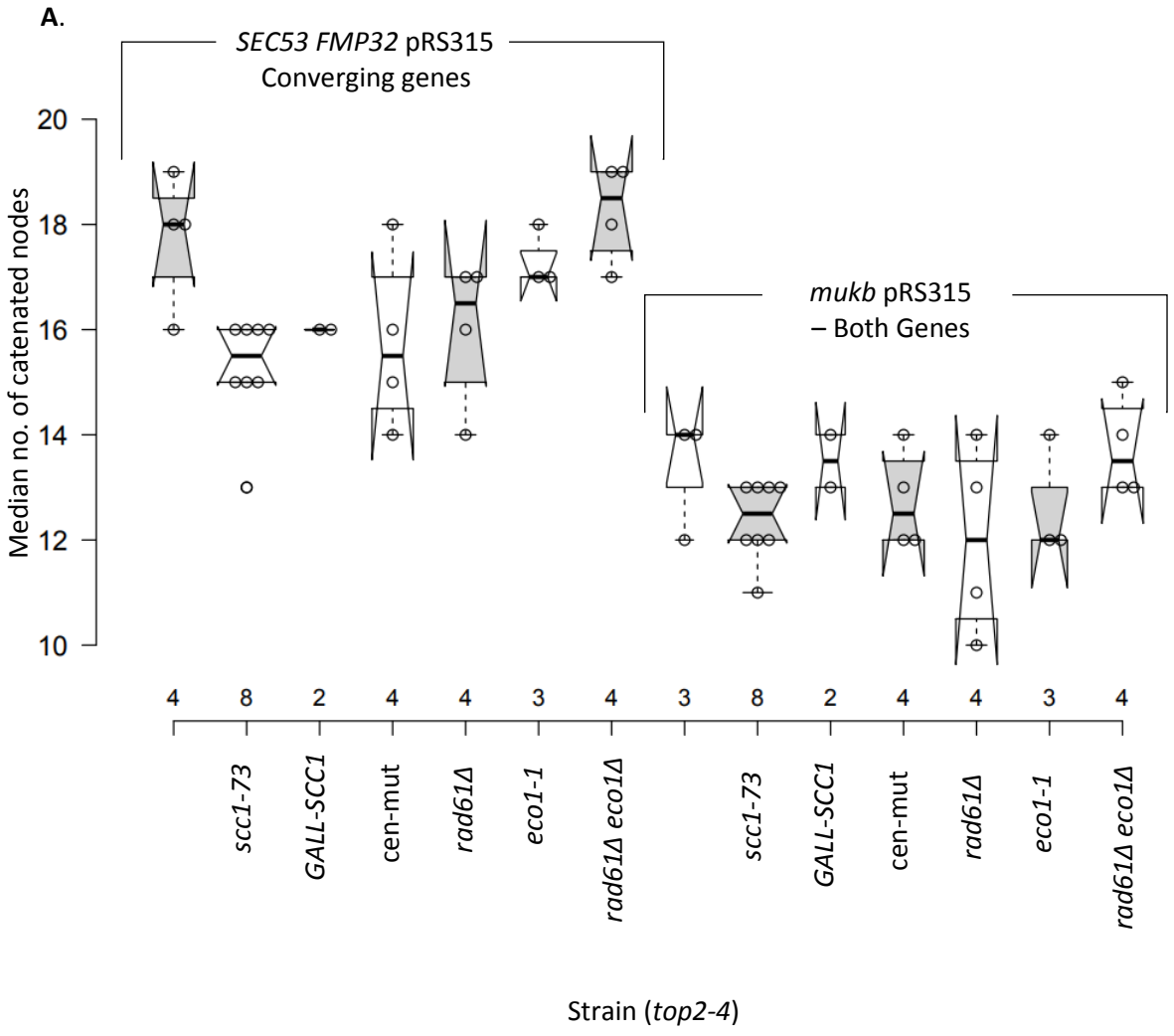




Converging genes

Figure 4.8 Rad61 is upstream of Eco1 in promoting cohesin-dependent topological changes at genes

Catenation assay of, *top2-4 eco1-1* cells **A** and **B**, *top2-4 rad61Δ* cells **C** and **D**, *top2-4 eco1Δ rad61Δ* background **E** and **F**. With the transcriptional control plasmid, **A, C** and **E** and the convergent transcription plasmid **B, D** and **F**. The histograms in **A** and **B** represent 3 repeats each, the histograms in **C- F** represent 4 repeats each.



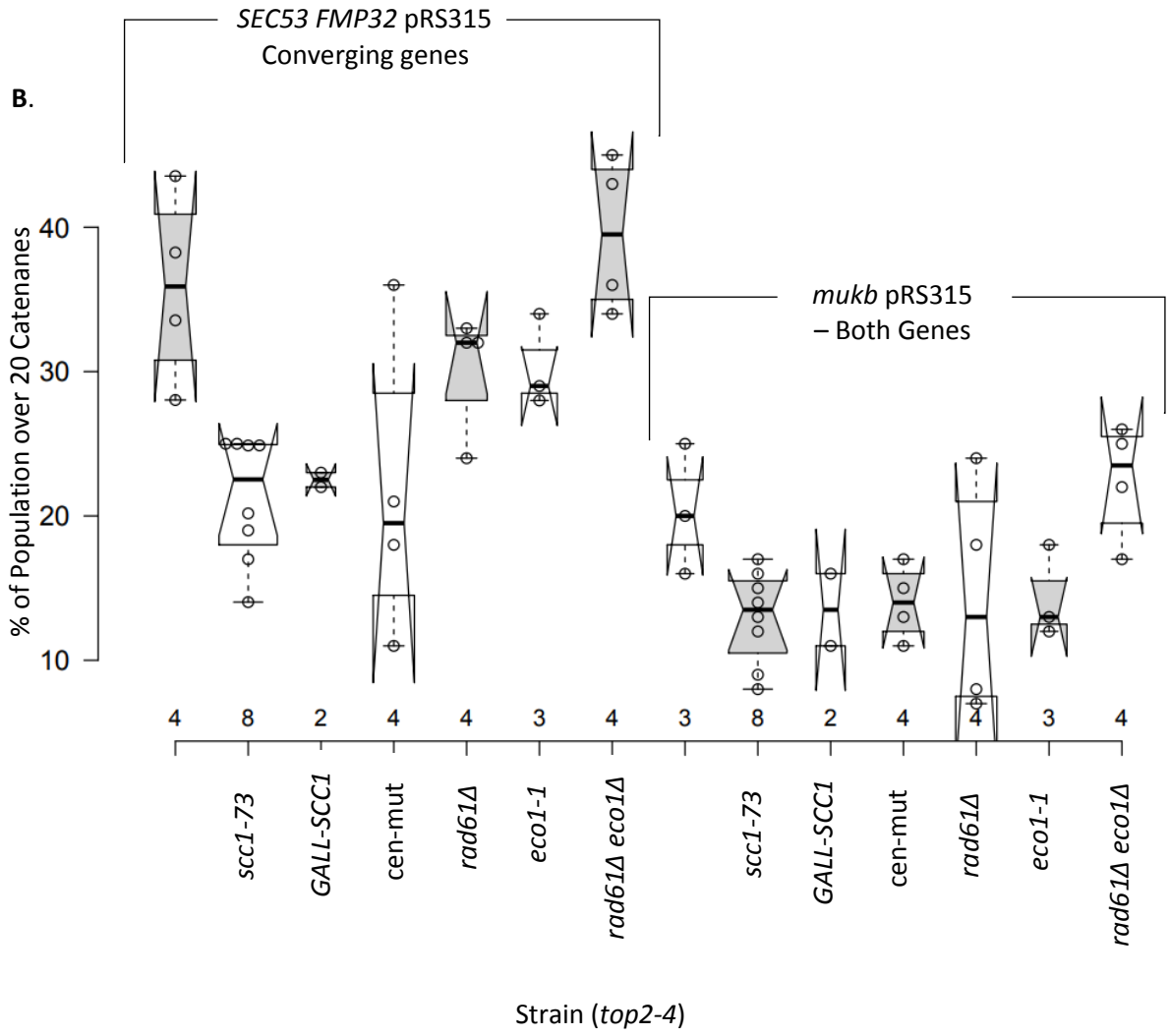


Figure 4.9 Summary boxplots of fork rotation results

Made using BoxPlotR (Spitzer et al., 2014). The width of each box is proportional to the square-root of the number of observations. The number above each strain represents the number of repeats. The notches (narrowing at the median) are defined by:

$$\pm(1.58 \times \frac{IQR}{\sqrt{n}})$$

Where IQR represents the interquartile range, n represents the number of repeats, and 1.58 is a constant representing the confidence interval. Therefore this is a calculation of the confidence interval multiplied by the standard deviation of the median (McGill et al., 1978). Non-overlapping notches give roughly 95% confidence that two medians differ. All strains listed are also in a *top2-4* background. **A**. Boxplots representing the median number of fork rotation events in each repeat for each of the specified strains. Strains are listed on the X axis with median number along the Y axis. **B**. Boxplots representing the percentage of each population that had over 20 catenanes for all repeats in each of the specified strains. All strains listed are also in a *top2-4* background.

4.6 DISCUSSION

4.6.1 Cohesin during S-phase promotes transcription-dependent replication fork rotation

In this chapter I have aimed to understand what factors are involved in the transcription-dependent increase in topological stress and fork rotation during S-phase. This chapter focused on extensively analysing the effects of cohesin, which was initially chosen knowing that cohesin is intimately related to transcription and that, as a non-static tracking protein, it could impede DNA replication itself.

In this chapter I focus on two of the plasmids from chapter 3, the convergent transcription plasmid (*SEC53 FMP32* pRS315), and the transcriptional control plasmid (*mukB* pRS315 - genes). As discussed in chapter 3, the control plasmid contains a *LEU2* gene, however in rich media, as was used here, it is likely this would be transcribed at a very low rate. I also analysed the unidirectional transcription plasmid (*SEC53 FMP32* pRS315), and a plasmid with tRNA genes inserted (3*tRNA* pRS316).

I firstly inactivated the cohesin complex by depleting the subunit Scc1, which is the kleisin subunit of the complex and required as part of the tripartite ring structure (Gruber et al., 2003). I did this via two methods; using a temperature sensitive *scc1-73* allele or a glucose induced transcriptional repression using *GALL-SCC1* (Figures 4.1, 4.3). The *scc1-73 top2-4* cells had a slow growth phenotype even at permissive temperatures, therefore *GALL-SCC1* which did not have slow growth could be used as a control for any untoward effects during the previous cell cycle. In both of these strains the inactivation of the cohesin complex during S-phase lead to a clear decrease in the level of fork rotation when transcription was present on the plasmid (Figure 4.9).

4.6.2 The cohesin-dependent increase in fork rotation is not specific to convergent gene sites

I next aimed to see if the cohesin effect was dependent on the convergent nature of the genes. I inactivated the cohesin complex and analysed the unidirectional transcriptional plasmid (Figure 4.1). I found that cohesin was also promoting fork rotation in the context of unidirectional genes.

It must be noted that I have not done any ChIP analyses on cohesin to see where the complex is localising on the plasmid and therefore cannot be sure that it is being pushed to the site of convergent transcription. This would be a helpful analysis as converging gene pairs in *S. cerevisiae* have been shown to have differing levels of cohesin enrichment. Jeppsson et al. (2014) showed that if the pairs of converging genes were close to centromeres they were likely to have

cohesin enrichment; which is likely due to cohesin being highly enriched around centromeres (Hu et al., 2011; Jeppsson et al., 2014). The two genes in the current study, being very far away from *CEN 6* on the endogenous chromosome, did not come up as highly enriched for cohesin as measured by a Top2 inactivation dependent accumulation of Smc5/6 (Jeppsson et al., 2014). However since these two genes were now located on a plasmid of less than 10 kb in size, they are within the distance limits of converging genes that have cohesin enrichment. In my analysis, I found that changing the direction of the genes on the plasmid does not affect the functionality of the cohesin on DNA in respect to the replication fork. This could reflect the fact that the distance from the centromere is close in both situations and the enrichment of cohesin and transcription-dependent effects could be high in both.

In a recent study by Borrie et al. (2017), looking at G2/M phase, they increased transcription of a gene 2.5 X above normal levels and found that this increase displaced functional cohesin, leading to loss of sister chromatid cohesion. The loss of cohesion was rescued either if the DNA was circularised or if convergent genes were inserted. On a chromosome, it therefore may be necessary to have convergent transcription to control the spread of cohesin. However, in my study this directionality does not appear to effect the actual topological change the cohesin is having in association with transcription and replication fork rotation.

4.6.3 Does cohesin promote fork rotation outside of transcription?

The transcriptional control plasmid does not appear to be greatly affected by cohesin (Figure 4.9). This can be seen in the low general level of fork rotation during S-phase (a median of 13). One termination event is expected to lead to a median of around 10 fork rotation events, and as discussed in Chapter 3 the increase to 13 could be due to having a centromeric replication fork pause site or the presence of the *LEU2* gene (See chapter 3.6.1). When cohesin is removed, there appears to be a slight, but not major effect, with a sharpening of the peak and a decrease in the median to 12 (figure 4.1, 4.2, 4.3). This argues that without transcription/ very low transcription, cohesin does not appear to have any significant effects on replication fork rotation (figure 4.9).

4.6.4 The role of the centromere

Using Scc1 ChIP-seq analysis Hu et al. (2011) identified a 10 kb region either side of the centromere where cohesin density was highly affected by the presence of the CEN region. Since, in this assay, the plasmids themselves are less than 10 kb in size, I presumed that the major effector of cohesin accumulation, is the centromere. To understand if this is true for a plasmid situation, and whether or not cohesin loaded at a centromere was important for the increase in topological stress seen, I inactivated the centromere using a point mutant to prevent the loading of cohesin. Plasmid specific depletion also prevented potential adverse general effects on cellular metabolism arising from a general cohesin inactivation. However, inactivating the centromere itself could possibly lead to lower levels of fork rotation, as it is a known pause site for replication (Ivessa et al., 2003). This possibility has been looked at in Schalbetter et al. (2015), but they found only a very slight decrease (median of 12 compared to 13). Since I expected a lowered level of cohesin to be loaded onto the plasmid, and for the centromere to be very important in the enrichment over the entire plasmid, I expected a large decrease in fork rotation rather than a small decrease. This was exactly the case, the distribution of catenanes for both of the plasmids looked very similar to the distributions when cohesin was inactive (Figure 4.4). This, not only strengthens the evidence of cohesin as an effector of fork rotation, but also suggests that on these plasmids the main loading site for cohesin is at the centromere rather than stochastically near the genes.

However, following on from this result I also aimed to inactivate the centromere using an Ndc10 mutation (Figure 4.4). This mutation would prevent the formation of the kinetochore and its associated components (Cho and Harrison, 2011). However, when this method was utilised, I found that there was no large decrease in fork rotation as was expected with the lowered levels of cohesin molecules on the plasmid. I concluded that this could possibly be due to the differences in the conditional depletion of Ndc10 during late G1 phase, compared to the constitutive inactivation of the centromere. The long arrest in G1 phase triggered by alpha factor is known to lead to accumulation of cohesin on chromosomes (Lengronne et al., 2006) and could lead to the accumulation of cohesin on DNA in the alpha factor arrested cells before the centromere was inactivated (>2 hours after addition of alpha factor). This population then could be sufficient to promote replication fork rotation on the convergent transcription plasmid.

4.6.5 Transcription of short genes

tRNA genes are short, highly transcribed genes, that are known fork pausing sites and lead to high levels of replication fork rotation (Figure 4.5B) (Ivessa et al., 2003; Schalbetter et al., 2015). They also increase the recruitment of Scc2, the cohesin loader (D'Ambrosio et al., 2008). When analysing a plasmid with 3x *tRNA* genes inserted, there was no clear effect of the cohesin inactivation as compared to the effect on the convergent transcription or unidirectional transcription plasmids (Figure 4.5). The difference between tRNA genes and the *SEC53* and *FMP32* genes could be due to the length of the transcripts; greater topological stress builds up over longer transcripts, as is seen by the necessity for Top2 activity at the longest genes (Joshi et al., 2012). Alternatively, it could indicate a fundamental difference between the interaction of cohesin with RNA polymerase II and RNA polymerase III action.

Potentially, the high levels of fork rotation seen at *tRNA* genes, could be due to a different non-static SMC protein rather than cohesin. Cohesin does not accumulate at such high levels at *tRNA* genes compared to the rDNA or centromere (Hu et al., 2011). However, the cohesin loader is proposed to facilitate condensin loading, and it appears condensin accumulates at *tRNA* genes (D'Ambrosio et al., 2008). Therefore, an interesting future experiment would be to test a condensin inactivation with a plasmid containing *tRNA* genes.

4.6.6 How is cohesin leading to topological stress during S-phase

I firstly looked at supercoiling levels when the cells entered S-phase, to understand if the initial supercoiling state of the plasmid affects fork rotation levels during replication. It was clear that the transcriptional control plasmid (without additional genes) was less negatively supercoiled than the convergent transcription plasmid (Figure 4.7). This may be due to a negatively supercoiled state being favourable for RNA polymerase (RNAP) activation and transcription to occur. However, it appears that the supercoiling state going into S-phase is not directly related to levels of fork rotation, as negative supercoiling would only be likely to lead to a decrease as opposed to an increase in replication fork rotation. When comparing supercoiling levels following cohesin inactivation, compared to active cohesin I saw no discernible differences. This result argues against the possibility that the inactivation of cohesin causes a lower level of transcription on the plasmid (since transcription is presumably regulating DNA supercoiling) and thus causing the rescue of the transcription dependent fork rotation.

Following on from this I looked at two proteins that are closely related to cohesin function, Eco1 and Rad61 (Wapl). Eco1 can travel with the replication fork, it is known to acetylate cohesin throughout S-phase and counteract Rad61 in removing cohesin from DNA (Rowland et al., 2009). Eco1 acetylation of Smc3 promotes sister chromatid cohesion and longevity of cohesin on DNA. Using an Eco1 depletion strain I aimed to understand if the cohesive nature of cohesin was a factor effecting its topological associations with the replication fork. I saw no significant changes using the *top2-4 eco1-1* background compared to *top2-4* alone (Figure 4.8A, B). The convergent transcription plasmid catenane distribution changed from a median of 18 to 17 and a percentage over 20 of 36% to 30% (Figure 4.8A). If there is a decrease here, this could be due to the high turnover of cohesin molecules on the DNA that occurs in the absence of Eco1 due to Rad61 activity. However, from this experiment it looks as if the acetylation of cohesin per se is not necessary for the promotion of replication fork rotation. This might be because Eco1 promotes sister chromatid cohesion behind the replication fork, which may not interfere with replication fork rotation.

I went on to analyse the antagonist of Eco1, Rad61 and found that the distribution of fork rotation events using the convergent transcription plasmid in *top2-4 rad61Δ* cells was reduced compared to *top2-4* alone. A clear peak formed at 12 catenanes, although the percentage of the population that had over 20 catenanes was still high (30%) (Figure 4.8C). These results were at first surprising, as Rad61 is known to remove non-acetylated cohesin from DNA, increasing the cohesin turnover. However, in budding yeast it appears the role of Rad61 is not fully elucidated. Although *rad61Δ* increases the cohesin residency time on DNA, it also has been shown to cause a large decrease in the total number of cohesin molecules on DNA (Chan et al., 2012; Rowland et al., 2009). This is thought to be because a deletion leads to a decrease in *SCC1* transcription about 2 fold (Sutani et al., 2009). Therefore, if there is less cohesin loading onto the plasmids, it is likely that a decrease in fork rotation would result. This would particularly effect the G1 phase population of cohesin binding DNA. In late G1 phase there is a small population of cohesin that binds DNA (Petela et al., 2018). It appears from the results using an *ndc10-1* mutation that the amount of cohesin that accumulates during an alpha factor arrest, holding cells in late G1 phase, is sufficient to promote fork rotation on the convergent transcription plasmid (Figure 4.4D). Therefore a mutation that results in lowered *SCC1* transcription when it appears to be required, suggests that this may be the cause of the rescue in replication fork rotation in this strain. In order to prove this point it would be possible to overexpress *SCC1* during G1-phase and analyse the effects of *rad61Δ* subsequently. Chan et al. (2012) went on to further analyse the number of cohesin molecules on DNA in a *rad61Δ eco1Δ* strain, and found that normal levels of cohesin

was bound to DNA. When I deleted both *ECO1* as well as *RAD61* this also rescued the *RAD61* phenotype and brought fork rotation levels back to normal *top2-4* levels (Figure 4.7E, F).

These results further add to the evidence that cohesin is promoting replication fork rotation, specifically in the context of transcription on a plasmid. How though does cohesin promote fork rotation? One theory is that cohesin increases the potency of a replication–transcription collision. Cohesin is known to be pushed by the RNAP and therefore could be situated between the replication fork and transcription machinery when a collision occurs, leading to high topological stress and fork rotation. In this situation, cohesin could lead to the sterical exclusion of topoisomerases from acting on the positive helical stress, forcing fork rotation to occur. Or, it could prevent the passing of the two machineries by acting as a physical block. If cohesin is more generally affecting the topological stress of the DNA, it could be that it acts as a block to the diffusion of supercoiling, similar to heterochromatin, causing more fork rotation to be required where populations of cohesin exist. It is likely that a build-up of cohesin molecules in front of the replication fork could result in higher levels of fork rotation, as they would need to be removed for the diffusion of topological stress and the continuation of replication to occur.

5 COHESIN PROMOTES DNA TOPOLOGICAL STRESS RELATED DNA DAMAGE

5.1 INTRODUCTION

In Chapters 3 and 4 I have explored how topological stress and the factors involved in promoting topological stress affect the replication fork, causing replication fork rotation. In this chapter, I explore the possibility that high topological stress during S-phase leads to DNA damage.

DNA supercoiling ahead of a replication fork is overcome by either the removal of the putative block to diffusion of DNA topological stress by specialist helicases (Ivessa et al., 2003), or by replication fork rotation and removal of DNA catenanes by topoisomerase II (Keszthelyi et al., 2016). When these pathways fail, the replication fork may arrest and fork reversal may occur. Inappropriate processing of reversed forks can lead to DNA damage. However, fork rotation in itself may also be a detrimental process. In the paper by Schalbetter et al. (2015), it was observed that the derestriction of replication fork rotation due to a deletion of *TOF1,* lead to DNA damage during S-phase, as measured by increased levels of H2A S129P (γH2A). This indicated that the reason replication fork rotation may be restricted is to prevent DNA damage occurring behind the replication fork. They suggested that DNA damage may be a consequence of the braiding of the sister chromatids preventing processes such as Okazaki fragment maturation (Schalbetter et al., 2015).

Whether it is the inappropriate processing of replication forks or the consequences of high levels of catenanes that leads to DNA damage, it is important to understand where DNA damage is occurring in the genome following S-phase induced DNA topological stress. In this chapter, I analyse topologically associated DNA damage, using a Chromatin Immunoprecipitation Sequencing (ChIP-seq) method for antibodies against H2A S129P (γH2A) during S-phase. Histone 2A is phosphorylated by Tel1/Mec1 in budding yeast at sites of double stranded DNA (dsDNA) breaks (Downs et al., 2000; Shroff et al., 2004). It can therefore be used as a marker for DNA damage. Previous studies on DNA damage in Top2 depleted budding yeast cells have only looked at DNA damage that occurs following the mis-segregation of chromsomes during Mitosis (Bermejo et al., 2009). Here, I follow on from the work in Chapter 4, by analysing how the activity of the SMC cohesin complex affects topological stress-associated DNA damage, and the relation of cohesin linked DNA damage to transcription.

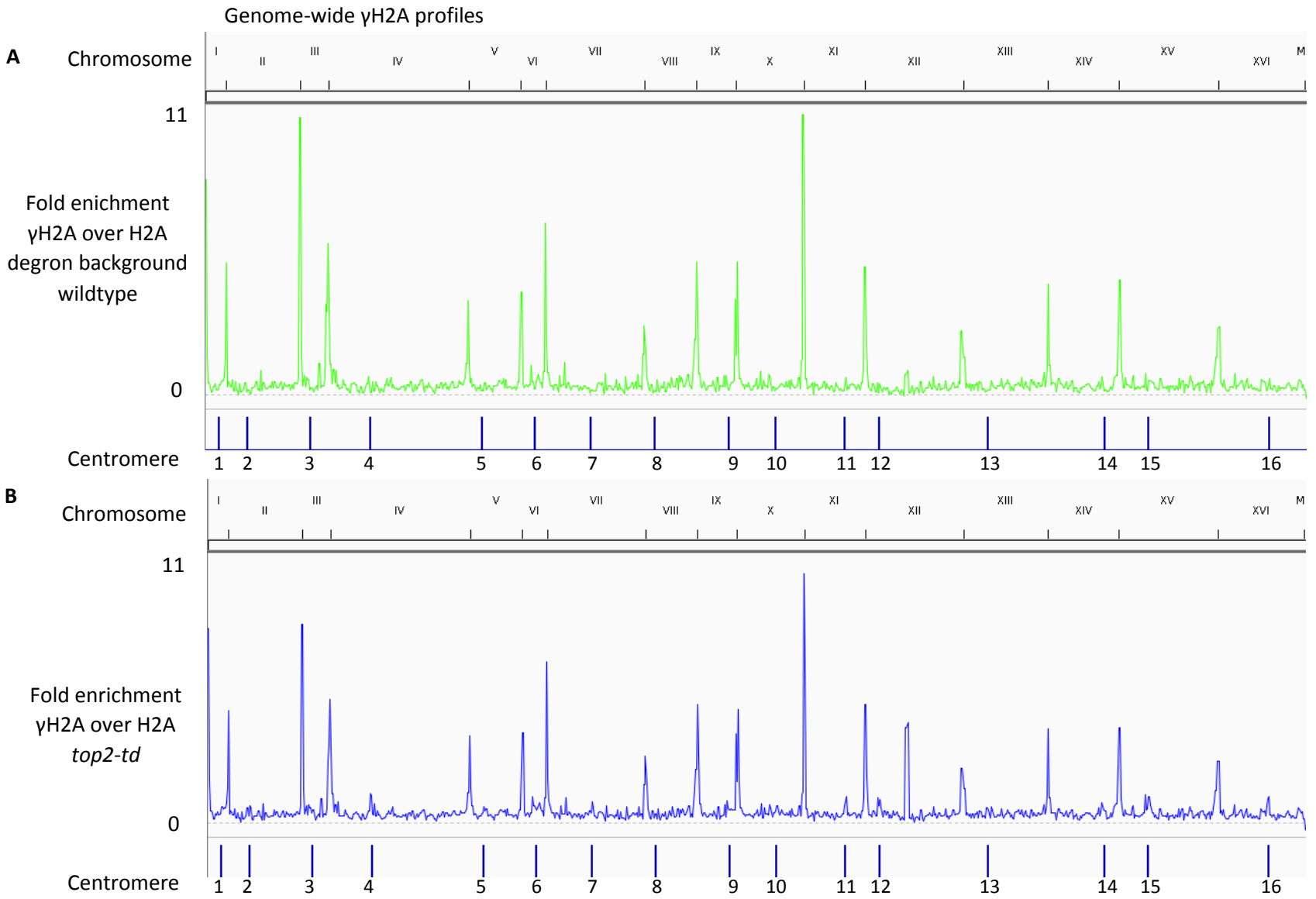
5.2 DNA TOPOLOGICAL STRESS DURING S-PHASE RESULTS IN DNA DAMAGE AT CENTROMERES AND THE RDNA ARRAY

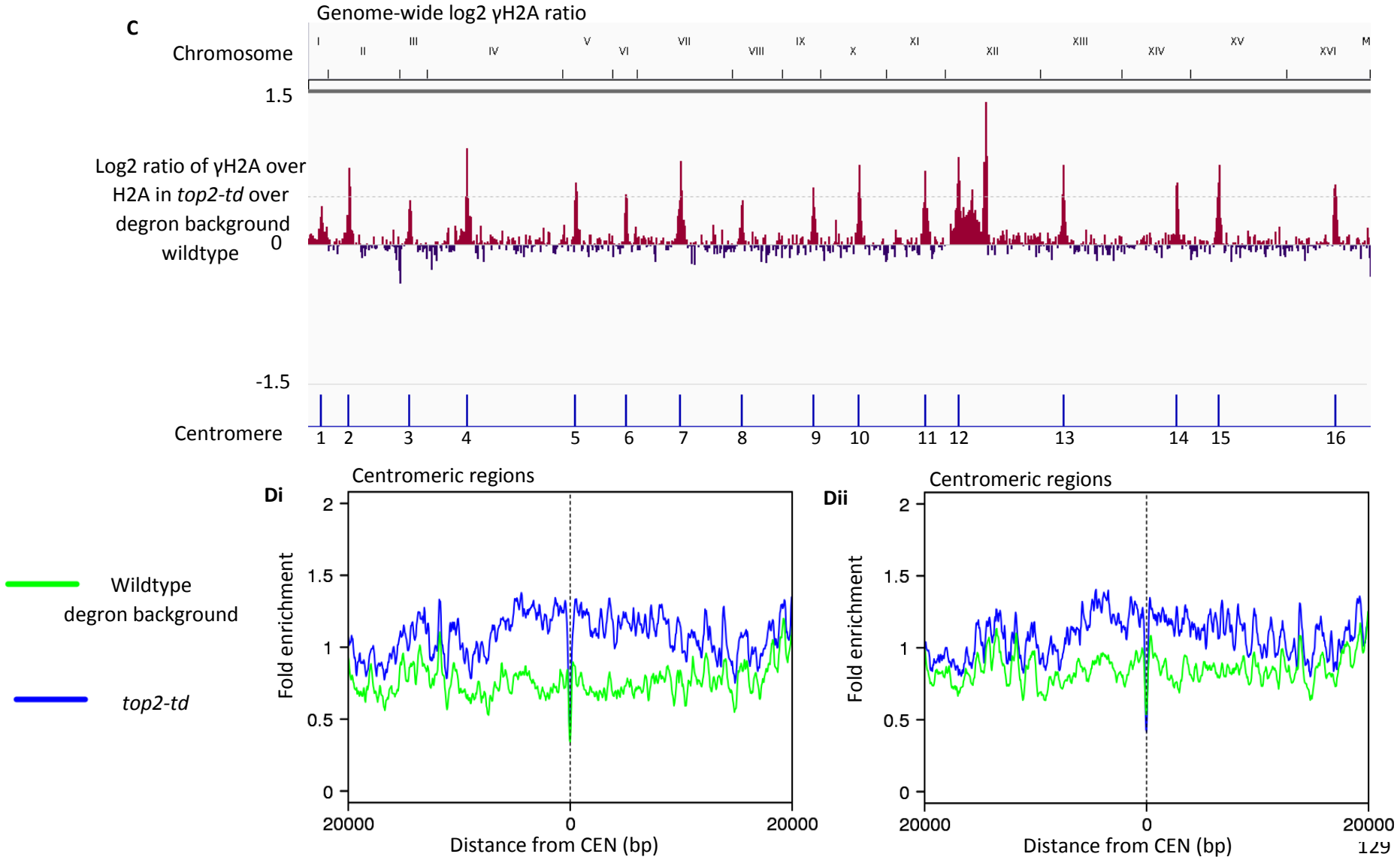
My first aim was to analyse genome wide, topological stress-associated, DNA damage during S-phase. To analyse DNA damage, I looked at phosphorylation levels of H2A (γH2A). I used a ChIP-seq method set up in the lab by a co-worker, Dr Andrea Keszthelyi, for full methodology see materials and methods section 2.9. The scripts for this analysis were also written by Dr Andrea Keszthelyi (materials and methods 2.9.5). To induce topological stress, Top2 can be depleted during S-phase using a *top2-td* degron strain and the DNA damage profiles compared against a Top2 active control (wildtype degron background) (Figure 5.1). When Dr Andrea Keszthelyi set up this technique, she looked at the profiles for these two strains, the results of which I recapitulated here. The DNA damage profile for the wildtype (degron background) strain appeared fairly even across the bulk of each chromosome (Figure 5.1A). However, there were peaks of γH2A at sub-telomeric regions and the rDNA array on chromosome XII (Figure 5.1A). Across chromosome arms the profile of γH2A following *top2-td* depletion in S-phase looked similar to the wildtype. There was a generally consistent level of γH2A across each chromosome with large peaks at some sub-telomeric regions (Figure 5.1B). However, there were two genomic contexts that showed clearly increased levels of γH2A. The rDNA region and at the centromeres of each chromosome had much higher γH2A enrichment in the *top2-td* cells (Figure 5.1B). This difference between the two strains is shown by a log2 ratio map, comparing *top2-td* over wildtype degron background (Figure 5.1C). The positive peaks represent an increase in γH2A enrichment over H2A in *top2-td* compared to the degron background, whereas the negative peaks represent an increase in γH2A enrichment over H2A in the degron background compared to *top2-td*. This clearly shows an increase in DNA damage at the centromeres and the rDNA regions in the absence of Top2 (Figure 5.1C).

To better visualise the increased level of DNA damage at the centromeres I used meta-data analysis to compile γH2A enrichment at the 16 centromeres (Figure 5.1D, E). I used an R script written by Dr Keszthelyi, where I looked at the γH2A fold enrichment 20 kb either side of the compiled centromeres. *top2-td* had a much higher overall level of DNA damage 10 kb either side of the point centromeres compared to the wildtype degron background strain (Figure 5.1D). This result was clearly reproducible as can be seen by the repeat (figure 5.D, E).

At the rDNA, different strains contain different numbers of repeats, however, in my analysis I looked at the ratio of γH2A over H2A, which therefore averages out the repeat number. The level of DNA damage at the rDNA appears much higher than at the centromeres, even in the

wildtype degron background. The damage is further elevated in *top2-td* strains, with the increase in DNA damage occurring primarily over the highly transcribed 5S and 35S regions (Figure 5.1E).


Genome-wide γH2A profiles
A
Chromosome
I II III IV V VI VII VIII IX X XI XII XIII XIV XV XVI M
11
Fold enichment
γH2A over H2A
degron background
wildtype
0
Centromere
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
B
Chromosome
I II III IV V VI VII VIII IX X XI XII XIII XIV XV XVI M
11
Fold enrichment
γH2A over H2A
top2-td
0
Centromere
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16


C
Genome-wide log2 γH2A ratio
Chromosome
I
II
III
IV
V
VI
VII
VIII
IX
X
XI
XII
XIII
XIV
XV
XVI
M
1.5
Log2 ratio of γH2A over H2A in top2-td over degron background wildtype
0
-1.5
Centromere
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
Di
Centromeric regions
Fold enrichment
2
1.5
1
0.5
0
20000
0
20000
Distance from CEN (bp)
Dii
Centromeric regions
Fold enrichment
2
1.5
1
0.5
0
20000
0
20000
Distance from CEN (bp)
Wildtype degron background
top2-td

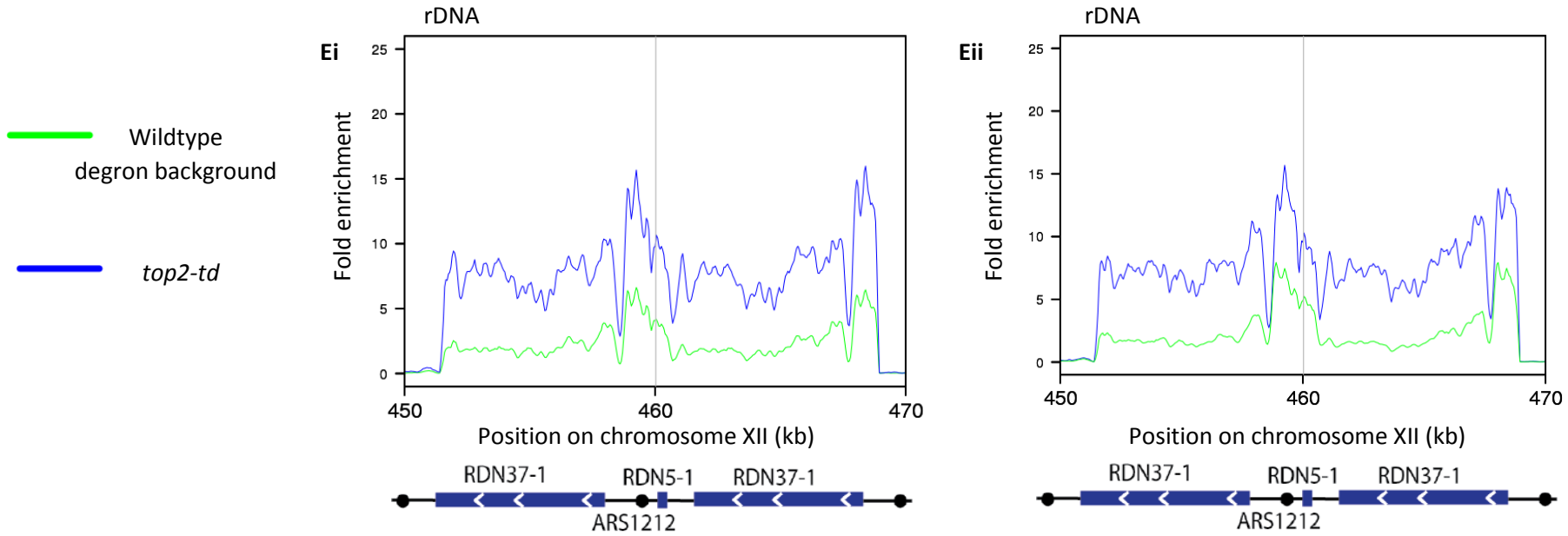
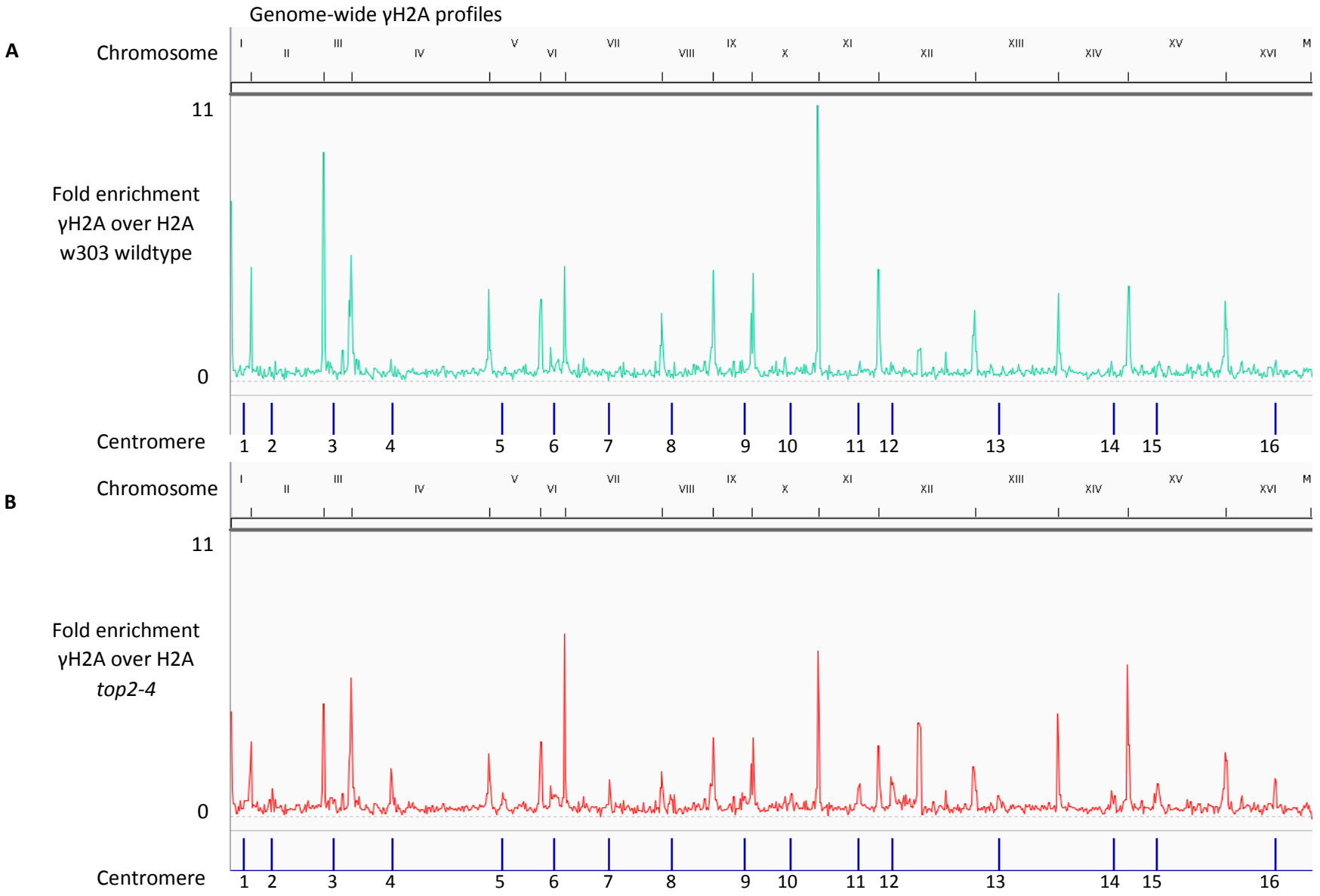


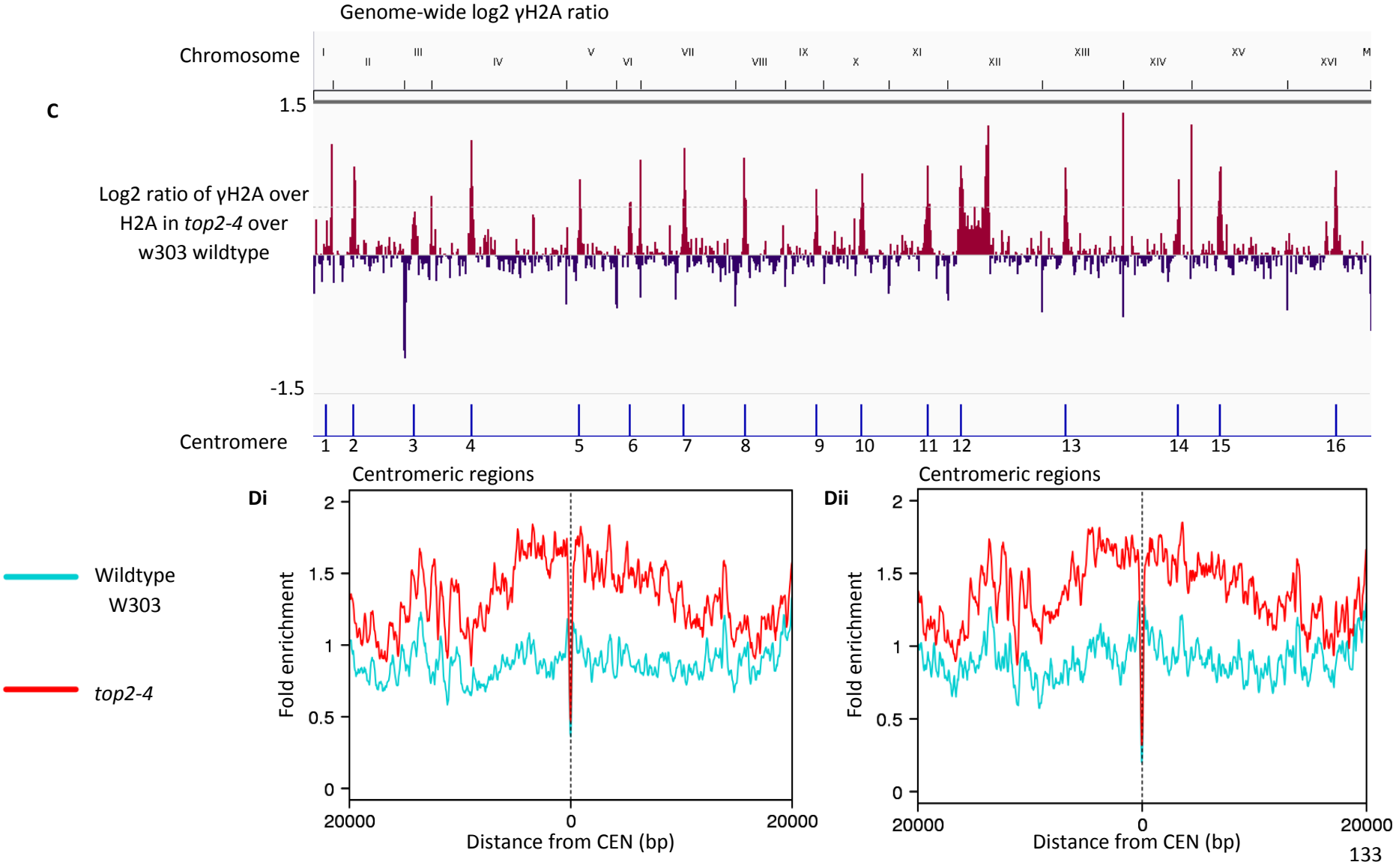
Figure 5.1 Topological stress leads to specific S-phase accumulated DNA damage at centromeres and rDNA

A. Genome wide wildtype (degron background) levels of γH2A (H2AS129P) fold enrichment over H2A during S-phase, the centromere of each chromosome is marked **B.** genome wide *top2-td* levels of γH2A fold enrichment over H2A during S-phase. **C.** Log2 ratio of *top2-td* vs wildtype levels of γH2A fold enrichment over H2A during S-phase **D**i degron background wildtype and *top2-td* γH2A fold enrichment over H2A, for a 1 – 16 centromere pile up of a 40 kb region around the centromeres, using the same data as in A and B. **Dii.** Repeat of experiment from Di. **Ei** wildtype (degron background) and *top2-td* γH2A fold enrichment over H2A, for a 20 kb rDNA array from Chromosome XII, same data as used in A and B. **Eii.** Repeat of experiment from Ei.

As a control for the degron method for depleting Top2, and for continuity between the ChIP-seq method and catenane experiments described in chapters 3 and 4, I went on to analyse a *top2-4* strain against a w303 wildtype control (Figure 5.2). I found the wildtype had a consistently low level of γH2A across the chromosome body with peaks at the same telomeres and the rDNA as the degron background wildtype (Figure 5.2A). Similarly to the results seen in *top2-td,* when the topological stress was increased through S-phase using the *top2-4* allele, damage was increased at the centromeres and the rDNA (Figure 5.2). At centromeres, this increase even exceeded the increase seen in the *top2-td* cells (Figure 5.2C).

The meta-data analysis for centromeres in a *top2-4* background confirmed that there was an even greater difference in DNA damage between *top2-4* and the wildtype, than *top2-td* and the wildtype *(*Figure 5.2D*)*. There was a large increase in DNA damage at the rDNA array between *top2-4* and w303 wildtype, with a peak at the 5S region and lower levels at the 35S region, very similar to the *top2-td* profile (Figure 5.2E).


Genome-wide γH2A profiles
A
Chromosome
I II III IV V VI VII VIII IX X XI XII XIII XIV XV XVI M
11
Fold enrichment
γH2A over H2A
w303 wildtype
0
Centromere
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
B
Chromosome
I II III IV V VI VII VIII IX X XI XII XIII XIV XV XVI M
11
Fold enrichment
γH2A over H2A
top2-4
0
Centromere
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16


Genome-wide log2 γH2A ratio
C
Chromosome
I
II
III
IV
V
VI
VII
VIII
IX
X
XI
XII
XIII
XIV
XV
XVI
M
1.5
Log2 ratio of γH2A over H2A in top2-4 over w303 wildtype
-1.5
Centromere
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
Di
Centromeric regions
Fold enrichment
2
1.5
1
0.5
0
20000
0
20000
Distance from CEN (bp)
Dii
Centromeric regions
Fold enrichment
2
1.5
1
0.5
0
20000
0
20000
Distance from CEN (bp)
Wildtype W303
top2-4

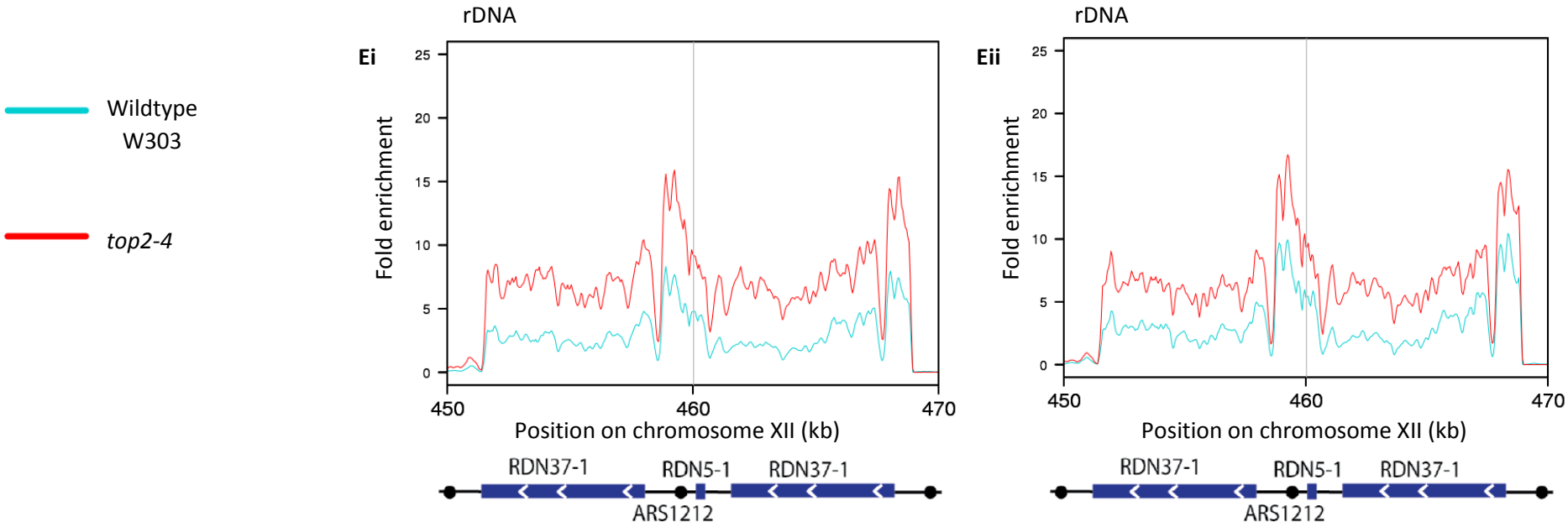
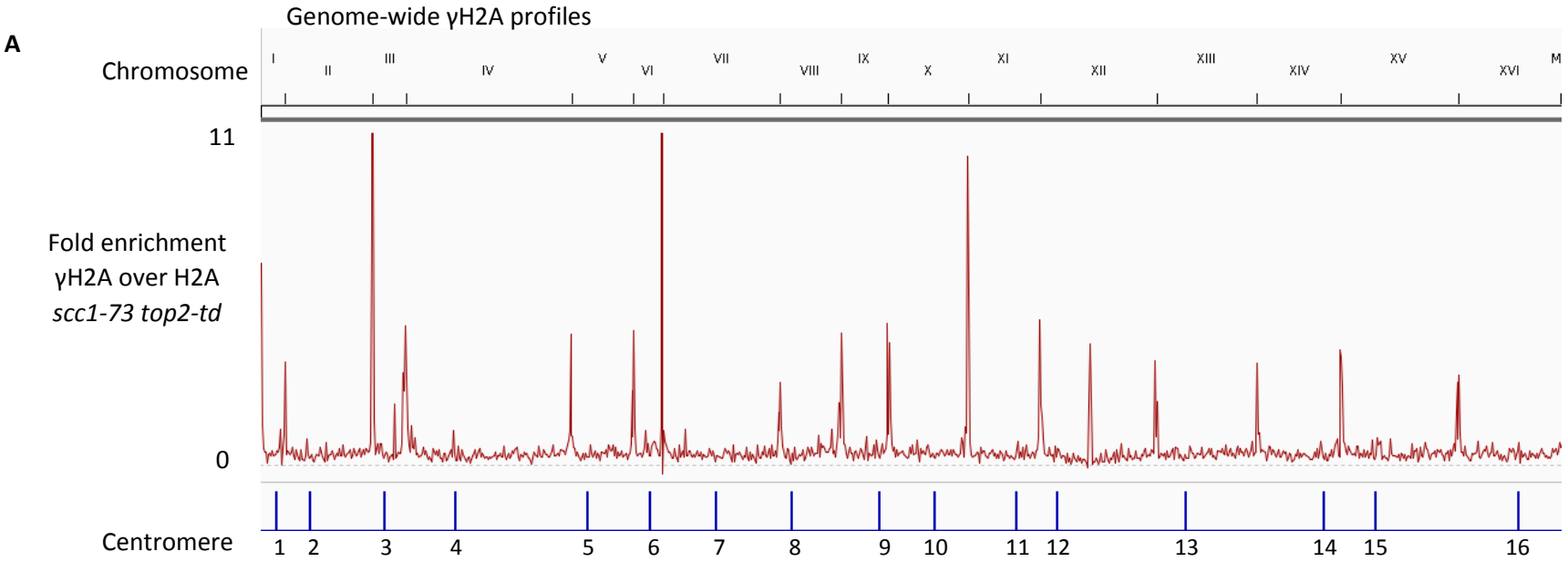


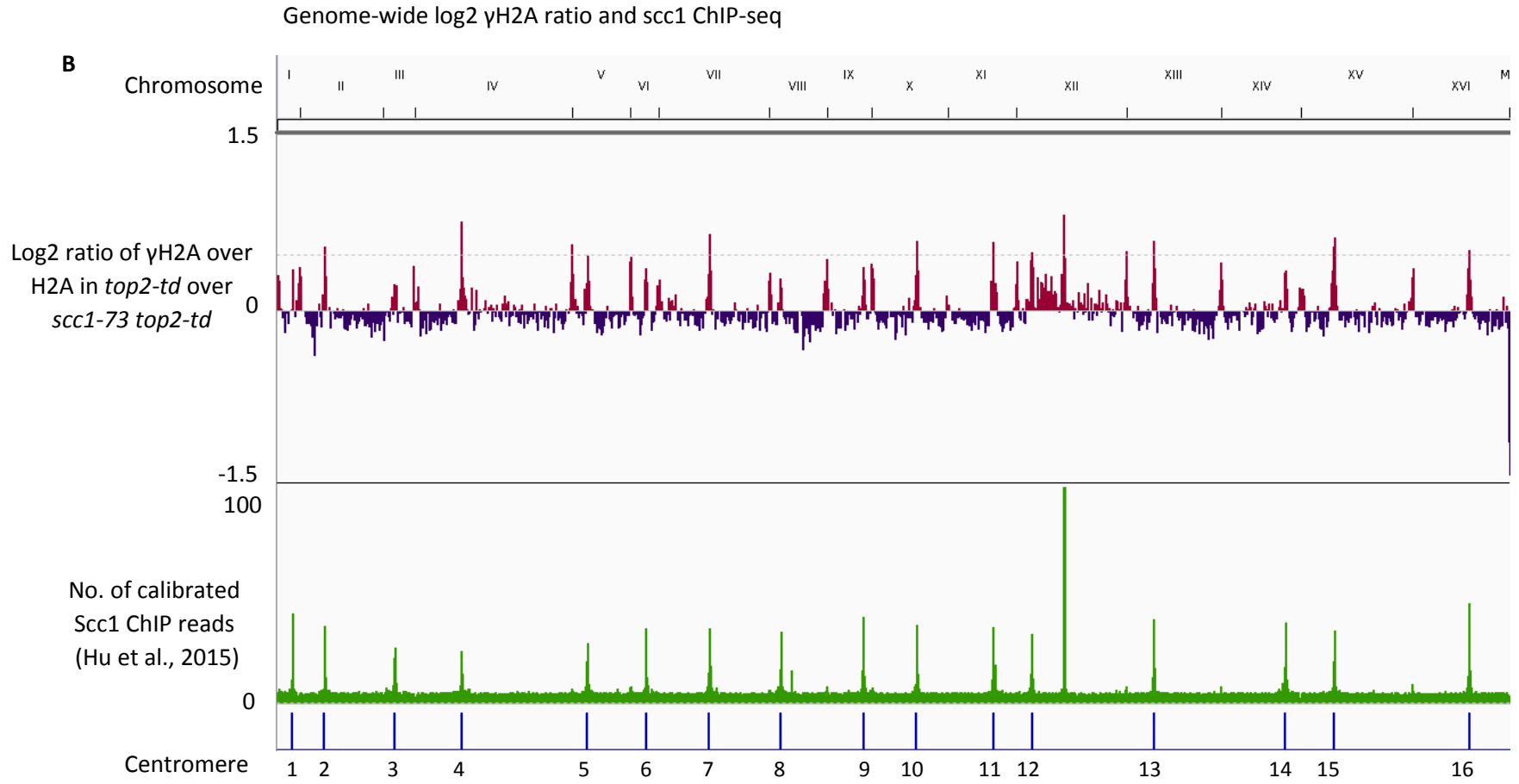
Figure 5.2 In the absence of Top2, DNA damage is prevalent across centromeres and the rDNA during S-phase

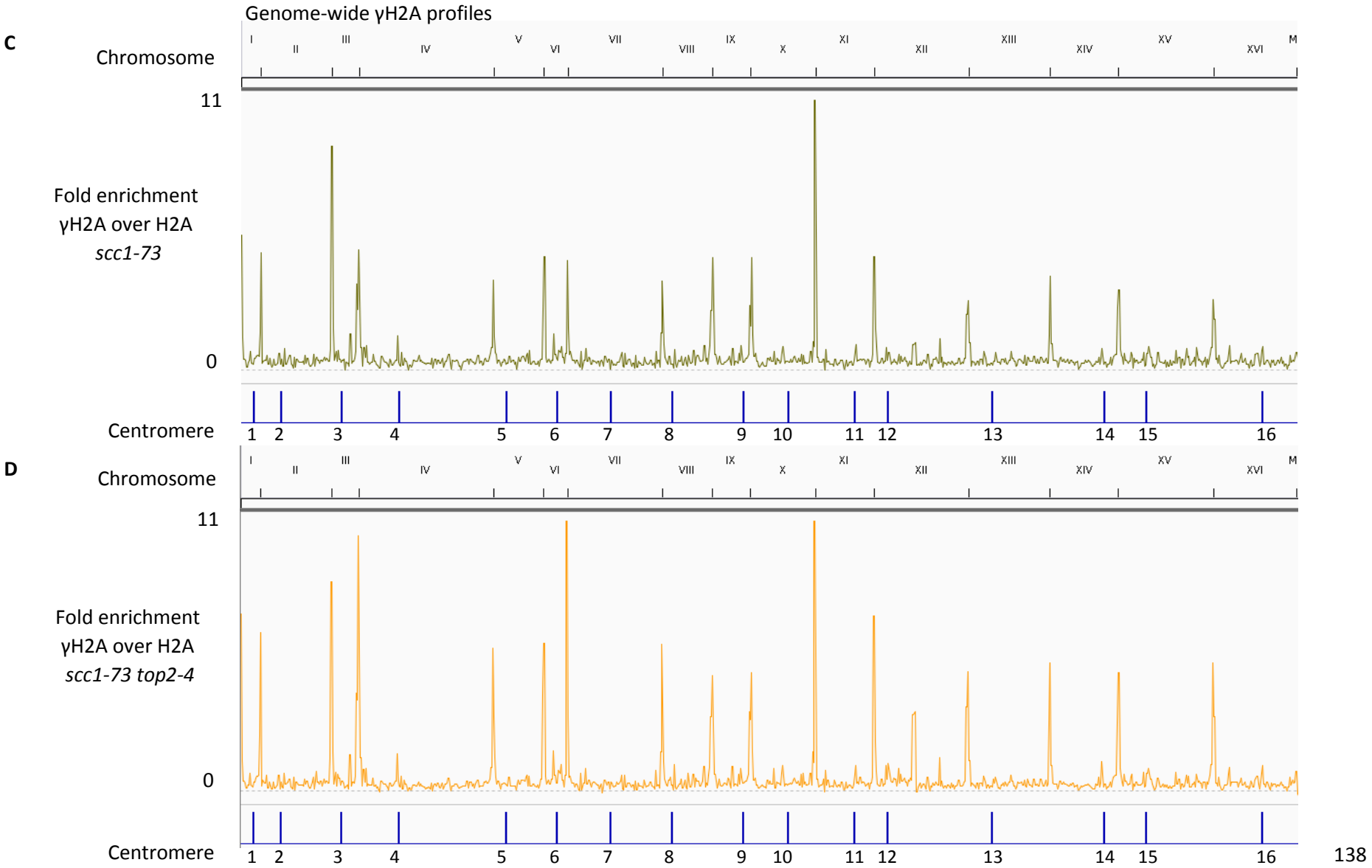
A. Genome wide wildtype (w303) levels of γH2A fold enrichment over H2A during S-phase, the centromere of each chromosome is marked **B.** Genome wide *top2-4* levels of γH2A fold enrichment over H2A during S-phase. **C.** Log2 ratio of *top2-4* vs wildtype (w303) levels of γH2A fold enrichment over H2A during S-phase **D**i wildtype and *top2-4* γH2A fold enrichment over H2A, for a 1 – 16 centromere pile up of a 40 kb region around the centromeres, using the same data as in A and B. **Dii.** Repeat of experiment from Di. **Ei** wildtype (w303) and *top2-4* γH2A fold enrichment over H2A, for a 20 kb rDNA array from Chromosome XII, same data as used in A and B. **Eii.** Repeat of experiment from Ei.

5.3 COHESIN LEADS TO TOPOLOGICAL STRESS ASSOCIATED DNA DAMAGE

I next sought to understand why particular regions are susceptible to DNA damage occurring due to topological stress from a Top2 depletion. The SMC complex cohesin is seen to be enriched at both the centromere and the rDNA regions during S-phase (Hu et al., 2015). It is known to be loaded at the core centromere and the 5S region of the rDNA. In Chapter 4 I showed that cohesin promotes replication fork rotation, and therefore topological stress, when analysing a centromeric plasmid with high levels of fork rotation due to transcription. I therefore went on to analyse the impact of cohesin on DNA damage linked to topological stress. I inactivated cohesin, either on its own or in addition to a Top2 depletion using the *scc1-73* allele. In both situations, the DNA damage profiles looked very similar to the wild type profiles (Figure 5.3). This indicated a rescue of the Top2 depletion-dependent DNA damage with an inactivation of cohesin. This rescue can be seen in the log2 ratios of top2 depletion over cohesin inactivation with top2 depletion (Figure 5.3B, E).


A
Genome-wide γH2A profiles
Chromosome
I
II
III
IV
V
VI
VII
VIII
IX
X
XI
XII
XIII
XIV
XV
XVI
M
11
Fold enrichment
γH2A over H2A
scc1-73 top2-td
0
Centromere
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16


Genome-wide log2 γH2A ratio and scc1 ChIP-seq
B
Chromosome
I
II
III
IV
V
VI
VII
VIII
IX
X
XI
XII
XIII
XIV
XV
XVI
M
1.5
Log2 ratio of γH2A over H2A in top2-td over scc1-73 top2-td
0
-1.5
100
No. of calibrated Scc1 ChIP reads (Hu et al., 2015)
0
Centromere
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16


Genome-wide γH2A profiles
C
Chromosome
11
Fold enrichment
γH2A over H2A
scc1-73
0
Centromere
I II III IV V VI VII VIII IX X XI XII XIII XIV XV XVI M
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
D
Chromosome
11
Fold enrichment
γH2A over H2A
scc1-73 top2-4
0
Centromere
I II III IV V VI VII VIII IX X XI XII XIII XIV XV XVI M
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

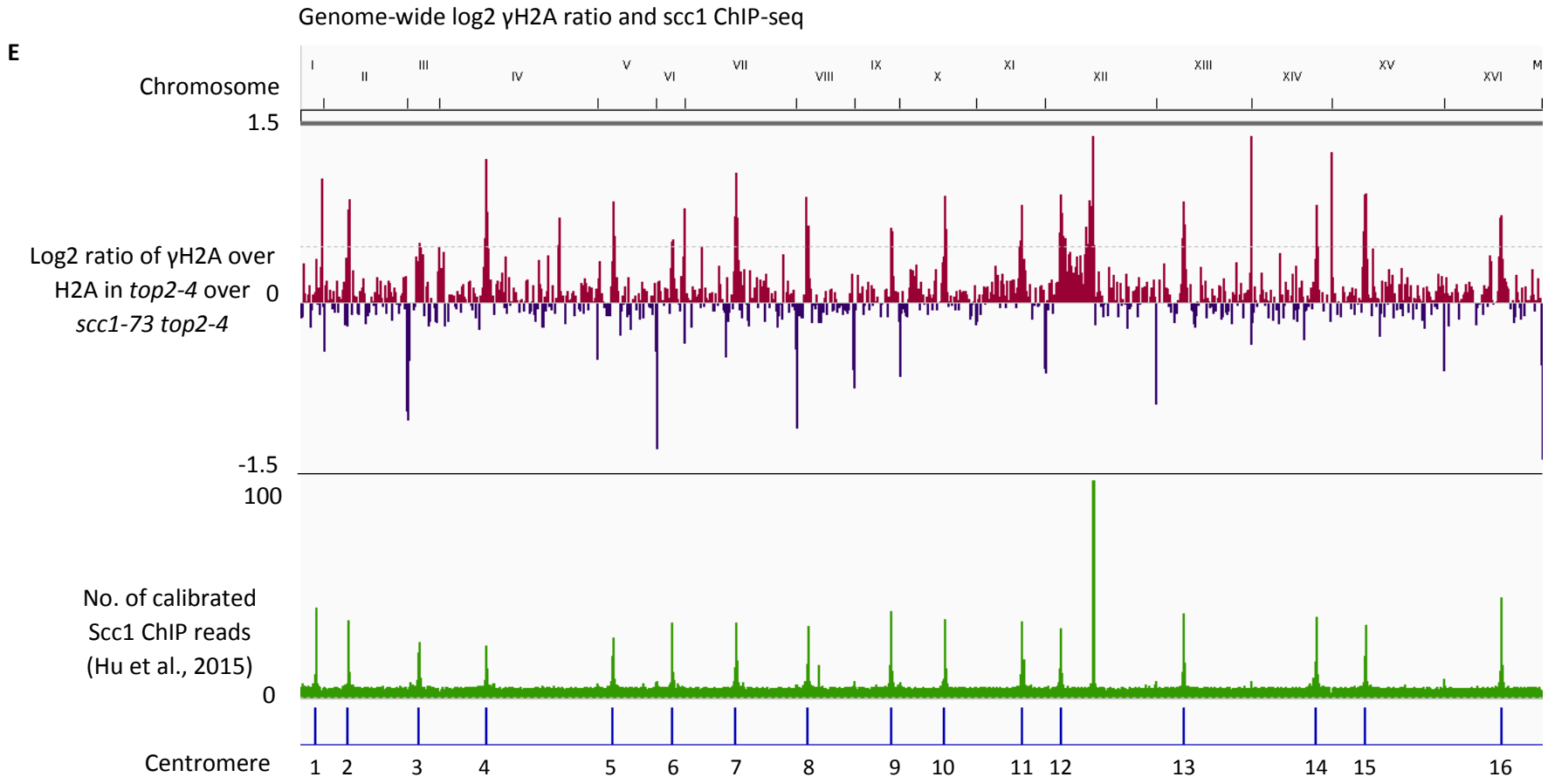
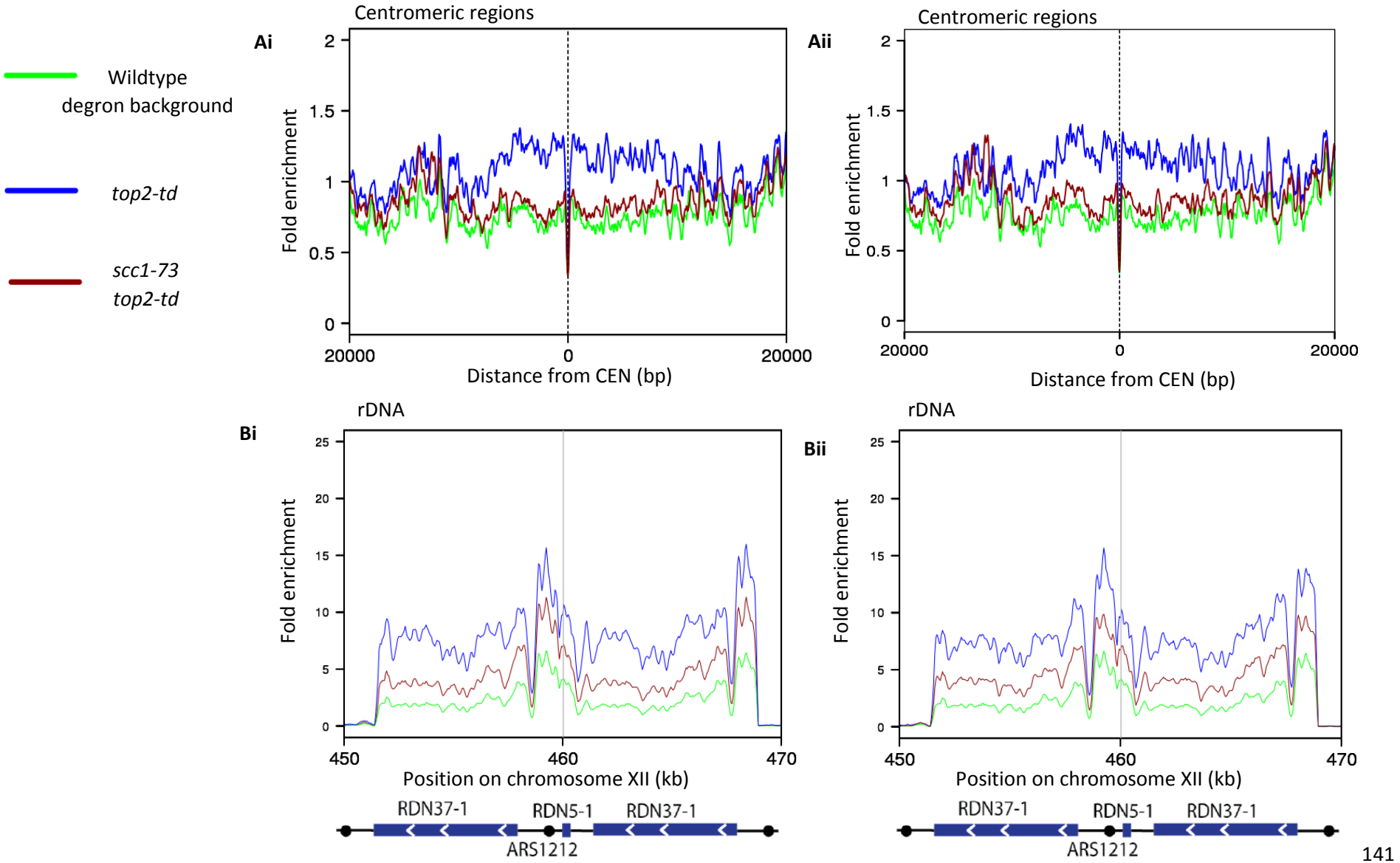


Figure 5.3 Cohesin inactivation rescues topological stress induced DNA damage during S-phase

A. Genome wide *scc1-73 top2-td* levels of γH2A fold enrichment over H2A during S-phase **B.** Log2 ratio of *top2-td* against *scc1-73 top2-td* levels of γH2A fold enrichment over H2A during S-phase, compared to 15min post alpha factor release Scc1 ChIP-seq taken from Hu et al. (2015) **C.** Genome wide *scc1-73* levels of γH2A fold enrichment over H2A during S-phase. **D.** Genome wide *scc1-73 top2-4* levels of γH2A fold enrichment over H2A during S-phase E**.** Log2 ratio of *top2-4* against *scc1-73 top2-4* levels of γH2A fold enrichment over H2A during S-phase, compared to 15min post alpha factor release Scc1 ChIP-seq taken from Hu et al. (2015)

To clearly see the specific areas affected by the topological stress, I looked closely at the rDNA arrays and 20 kb either side of a 16 centromere pile up (Figure 5.4). A complete rescue of the DNA damage was apparent at centromeres, with either *scc1-73 top2-td* or *scc1-73 top2-4* compared to a Top2 mutant alone (Figure 5.4A, 5.4C). There were also no increases in DNA damage with a cohesin inactivation alone compared to wildtype (Figure 5.4C). The rDNA however, had only a partial rescue of DNA damage, when cohesin was inactivated in the presence of Top2 depletion (Figure 5.4B, D). In both *scc1-73 top2-td* and *scc1-73 top2-4* there was a partial rescue of the Top2 mutant DNA damage at the 35S regions of the rDNA. However the 5S region DNA damage had some rescue in *scc1-73 top2-td* cells, but no rescue in *scc1-73 top2-4* cells; in one of the two repeats DNA damage is even higher at this point (Figure 5.4B, D). The *scc1-73* mutation alone caused no increases in DNA damage compared to the wildtype (Figure 5.4D).


Wildtype
degron background
top2-td
scc1-73
top2-td
Ai
Centromeric regions
Fold enrichment
Distance from CEN (bp)
Aii
Centromeric regions
Fold enrichment
Distance from CEN (bp)
Bi
rDNA
Fold enrichment
Position on chromosome XII (kb)
RDN37-1
RDN5-1
RDN37-1
ARS1212
Bii
rDNA
Fold enrichment
Position on chromosome XII (kb)
RDN37-1
RDN5-1
RDN37-1
ARS1212

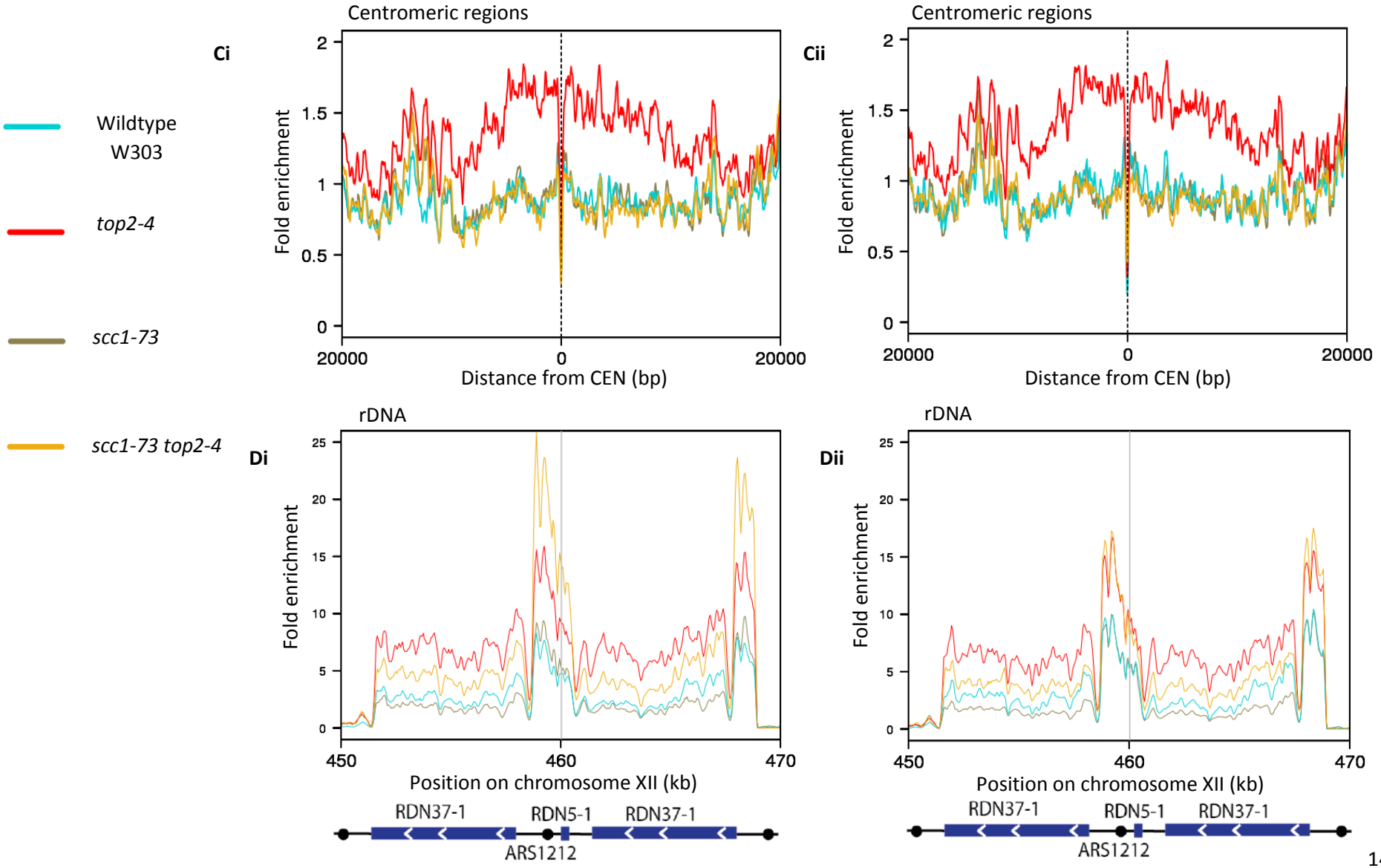

Wildtype
W303
top2-4
scc1-73
scc1-73 top2-4
Ci
Centromeric regions
Fold enrichment
Distance from CEN (bp)
Cii
Centromeric regions
Fold enrichment
Distance from CEN (bp)
Di
rDNA
Fold enrichment
Position on chromosome XII (kb)
RDN37-1
RDN5-1
RDN37-1
ARS1212
Dii
rDNA
Fold enrichment
Position on chromosome XII (kb)
RDN37-1
RDN5-1
RDN37-1
ARS1212

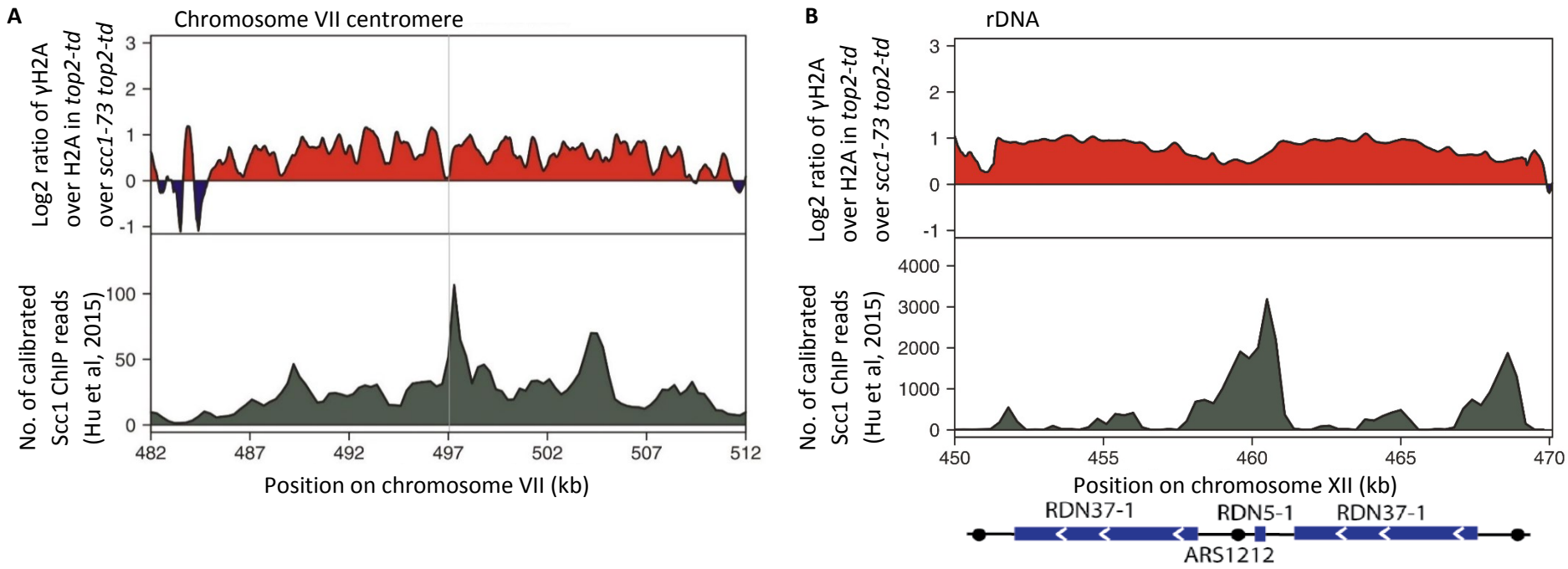
Figure 5.4 Cohesin inactivation rescues topological stress induced DNA damage at centromeres and the 35S rDNA regions, during S-phase

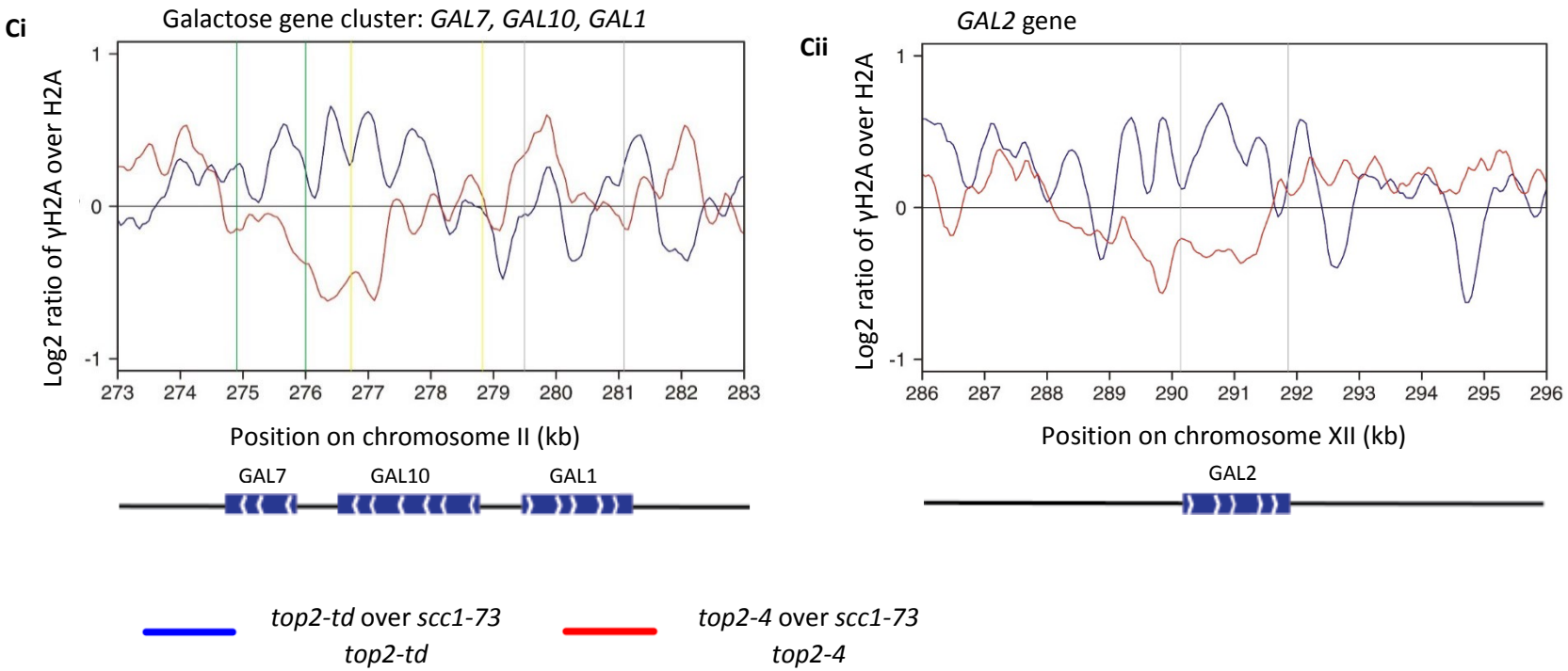
Ai. *scc1-73 top2-td* shown with degron background wildtype and *top2-td* for comparison, γH2A fold enrichment over H2A, for a 1 – 16 centromere pile up of a 40 kb region around the centromeres A**ii.** Repeat of *scc1-73 top2-td* B**i.** *top2-td scc1-73* shown with degron background wildtype and *top2-td* for comparison, γH2A fold enrichment over H2A, for a 20 kb rDNA array from Chromosome XII B**ii.** Repeat of *scc1-73 top2-td* **Ci.** *scc1-73* and *scc1-73 top2-4* shown with w303 wildtype and *top2-4* for comparison, γH2A fold enrichment over H2A, for a 1 – 16 centromere pile up of a 40 kb region around the centromeres **Cii.** Repeat of *scc1-73* and *scc1-73 top2-4* **Di.** *scc1-73* and *scc1-73 top2-4* shown with w303 wildtype and *top2-4* for comparison, γH2A fold enrichment over H2A, for a 20 kb rDNA array from Chromosome XII **Dii.** Repeat of *scc1-73* and *scc1-73 top2-4*

5.4 COHESIN TRANSLOCATION IS IMPORTANT FOR ITS EFFECT ON DNA DAMAGE

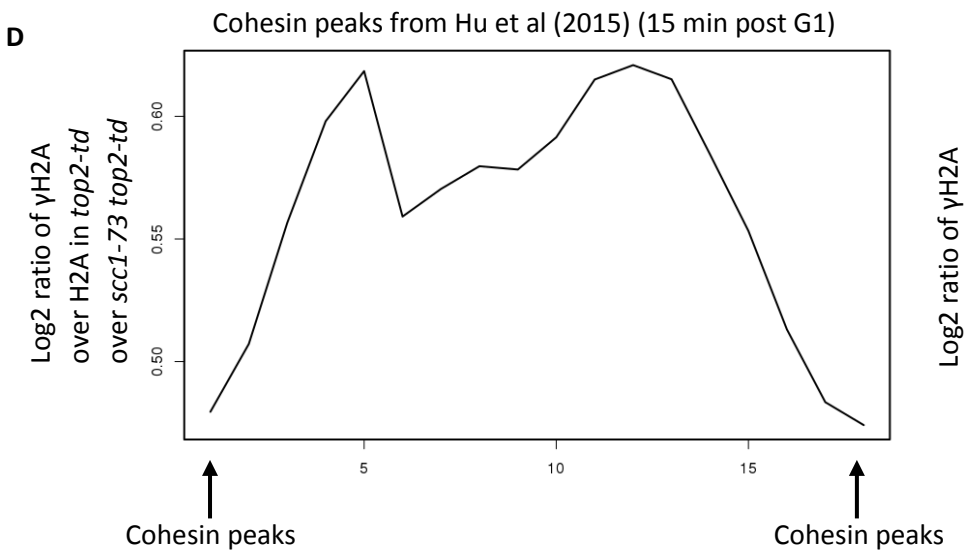
As there was only a rescue of DNA damage at the rDNA 35S region and not at the 5S region, this lead to the hypothesis that it is the translocation of cohesin that is important for topological stress associated DNA damage. Cohesin is known to load at the 5S region, and then translocate away from that region, due to the high levels of transcription. Cohesin also loads at the centromere and translocates away, spreading out during S-phase. To demonstrate this, Dr Keszthelyi took S-phase cohesin ChIP data from Hu et al. (2015) and compared it to the cohesin dependent, Top2 depleted, DNA damage profiles. The calibrated Scc1 ChIP-seq from Hu et al. (2015) was taken at 15 min post alpha factor release. The places of high cohesin enrichment, at both the centromere (CEN VII) and at the rDNA, are the places of the lowest levels of cohesin dependent *top2-td* DNA damage (Figure 5.5A, B). Whereas the places of cohesin translocation are the places of highest cohesin dependent DNA damage (Figure 5.5B). To further understand the link between cohesin translocation, transcription and DNA damage, cohesin dependent *top2-td* damage was analysed at highly expressed genes. These genes, not only had topological changes due to transcription, but presumably also had high levels of cohesin translocation. For this analysis we were able to take advantage of the differences between *top2-td* and *top2-4*. The former was grown in galactose media for the induction of the degron system, whereas the latter was grown in glucose. We could therefore analyse the difference in DNA damage with the difference of gene expression at the most highly expressed galactose promoters (*GAL1, GAL2, GAL7* and *GAL10*) ((Bash and Lohr, 2001). The results show a high level of cohesin dependent DNA damage in the *top2-td* cells, but contrastingly there was a decrease in DNA damage in *top2-4* cells compared to *scc1-73 top2-4* (Figure 5.5C, created by Dr Keszthelyi). Furthermore, this was seen across the genome; the regions in between cohesin enrichment peaks were seen to be the highest areas of cohesin dependent DNA damage (Figure 5.5D, 5.5E). This indicates, the areas of

cohesin translocation, which are presumably difficult to detect via ChIP-seq, are the areas of the most cohesin-dependent damage. Whereas cohesin domains, which are known to form to promote DNA integrity (Strom et al., 2004; Unal et al., 2004), were found at the lowest areas of topological stress related DNA damage.


A
Chromosome VII centromere
Log2 ratio of γH2A over H2A in top2-td over scc1-73 top2-td
No. of calibrated Scc1 ChIP reads (Hu et al, 2015)
Position on chromosome VII (kb)
B
rDNA
Position on chromosome XII (kb)
RDN37-1
RDN5-1
RDN37-1
ARS1212


Ci
Galactose gene cluster: GAL7, GAL10, GAL1
Log2 ratio of γH2A over H2A
1
0
-1
273 274 275 276 277 278 279 280 281 282 283
Position on chromosome II (kb)
GAL7
GAL10
GAL1
Cii
GAL2 gene
Log2 ratio of γH2A over H2A
1
0
-1
286 287 288 289 290 291 292 293 294 295 296
Position on chromosome XII (kb)
GAL2
top2-td over scc1-73
top2-td
top2-4 over scc1-73
top2-4

D


Cohesin peaks from Hu et al (2015) (15 min post G1)
Log2 ratio of γH2A over H2A in top2-td over scc1-73 top2-td
0.50
0.55
0.60
5
10
15
Cohesin peaks
Cohesin peaks

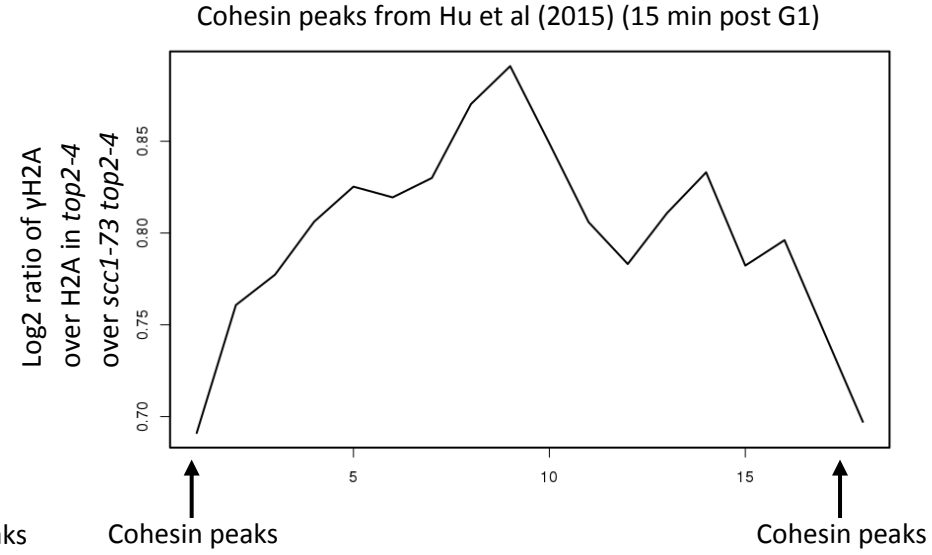

Cohesin peaks from Hu et al (2015) (15 min post G1)
Log2 ratio of γH2A over H2A in top2-4 over scc1-73 top2-4
0.70
0.75
0.80
0.85
5
10
15
Cohesin peaks
Cohesin peaks

Figure 5.5 S-phase topological stress associated DNA damage depends on cohesin translocation and can be related to transcription

A. Log2 ratio of γH2A over H2A in *top2-td* over *scc1-73 top2-td* at chromosome VII centromere. Compared against Scc1 ChIP-seq taken from Hu et al. (2015) at 15 min post alpha factor release B**.** Log2 ratio of γH2A over H2A in *top2-td* over *scc1-73 top2-td* at the rDNA array. Compared against Scc1 ChIP-seq taken from Hu et al. (2015) at 15 min post alpha factor release **Ci**. Log2 ratio of γH2A over H2A, in blue: *top2-td* over *scc1-73 top2-td,* in red: *top2-4* over *scc1-73 top2-4,* at a galactose gene cluster on chromosome II of *GAL7 (*green)*, GAL10* (yellow) and *GAL1* (grey) **Cii**. Log2 ratio of γH2A over H2A, in blue: *top2-td* over *scc1-73 top2-td,* in red: *top2-4* over *scc1-73 top2-4* at *GAL2* (grey) on chromosome XII **D.** Log2 ratio of γH2A over H2A in *top2-td* over wildtype and *top2-4* over wildtype between cohesin peaks.

5.5 PREVENTING COHESIN ACCUMULATION AT CENTROMERES RESCUES TOPOLOGICAL STRESS ASSOCIATED DNA DAMAGE

In order to understand the role of the centromere in cohesin accumulation I next aimed to prevent cohesin binding to centromeres. To do this I disrupted the formation of the kinetochore by using an *ndc10-1* mutation, which prevents the accumulation of the inner kinetochore protein. In chapter 4 the downregulation of this protein did not have any effects, compared to *top2-4* alone, on replication fork rotation during S-phase (Figure 4.4D). However, on endogenous chromosomes this depletion clearly leads to a partial rescue of topological stress-associated DNA damage both at the centromere and at the rDNA array (Figure 5.6). It is possible that the long arrest in G1 phase before Ndc10 depletion allows for the low level G1 accumulation of cohesin at centromeres. This could account for the fact that the rescue at the centromeres is not as great as with a cohesin mutant. It was surprising that *ndc10-1 top2-td* also resulted in a partial rescue across the 35S region of the rDNA, since this mutation should not affect cohesin loading at the 5S regions. It is possible that this effect could be related to the interaction between Ndc10 and Cbf5, another kinetochore protein, whose downregulation causes a substantial decrease in rDNA synthesis (Cadwell et al., 1997).

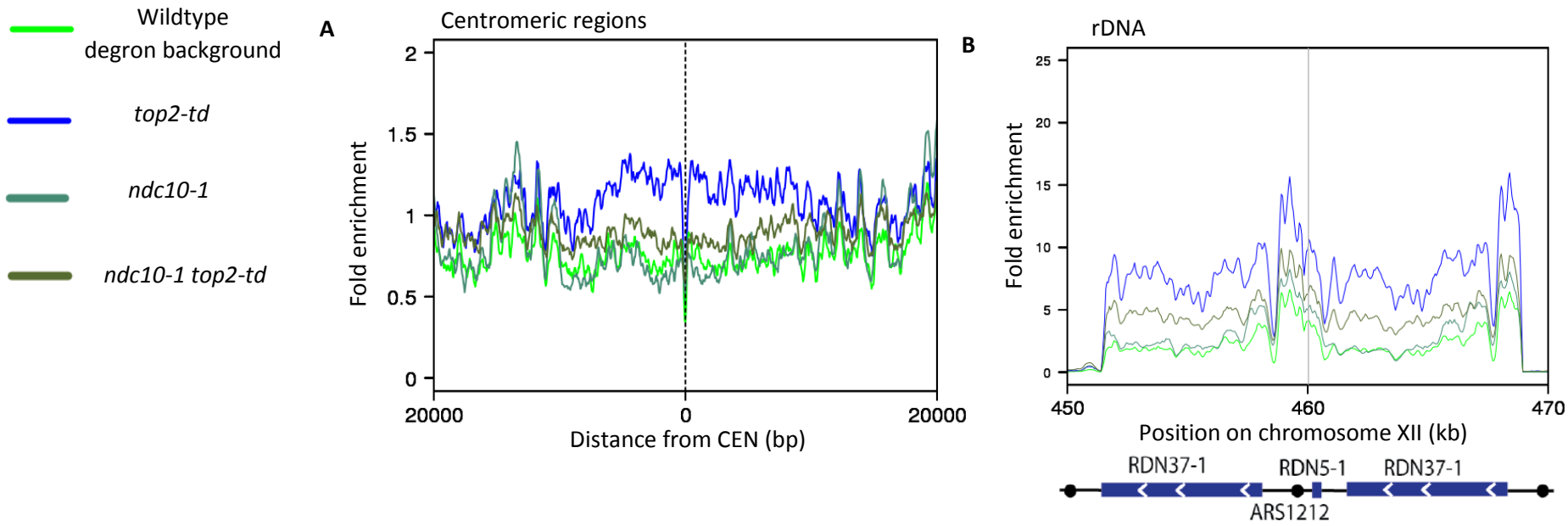


Figure 5.6 Inactivating the kinetochore partially rescues the topological stress-associated DNA damage

A. *ndc10-1* and *ndc10-1 top2-td* shown with degron background wildtype and *top2-td* for comparison, γH2A fold enrichment over H2A, for a 1 – 16 centromere pile up of a 40 kb region around the centromeres **B.** *ndc10-1* and *ndc10-1 top2-td* shown with degron background wildtype and *top2-td* for comparison, γH2A fold enrichment over H2A, for Chromosome XII 20 kb rDNA array

5.6 ECO1 AND RAD61 MAY MODIFY THE EXTENT OF COHESIN ASSOCIATED DNA DAMAGE

Following on from the extensive cohesin analysis, I further wanted to understand how the different functions of cohesin are related to the DNA damage induced by DNA topological stress. Similarly to chapter 4, I aimed to separate the cohesive nature of cohesin from its other aspects. To do this, I downregulated Eco1 during S-phase using an *eco1-1* allele, in a *top2-td* background or in the wildtype degron background. Eco1 acetylates cohesin on its Smc3 subunit during S-phase to allow for the establishment of sister chromatid cohesion (Rolef Ben-Shahar et al., 2008; Rowland et al., 2009; Unal et al., 2008; Zhang et al., 2008). In Chapter 4 I found that Eco1-dependent acetylation does not appear to effect the increase in topological stress due to cohesin on plasmids. To see if this also applies on chromosomes, and if an Eco1 depletion leads to a rescue of DNA damage caused by Top2 depletion, I carried out a centromere pile-up analysis showing degron background and *top2-td* alone for reference. Taking out Eco1 alone showed either no change from wildtype levels of DNA damage, or a slight increase in DNA damage levels (Figure 5.7). When Eco1 was removed in conjunction with Top2, there was a partial rescue 5 kb either side of the centromeres (Figure 5.7A). However, at the rDNA, DNA damage levels are very similar in *eco1-1 top2-td* to that of *top2-td* alone (Figure 5.7B).

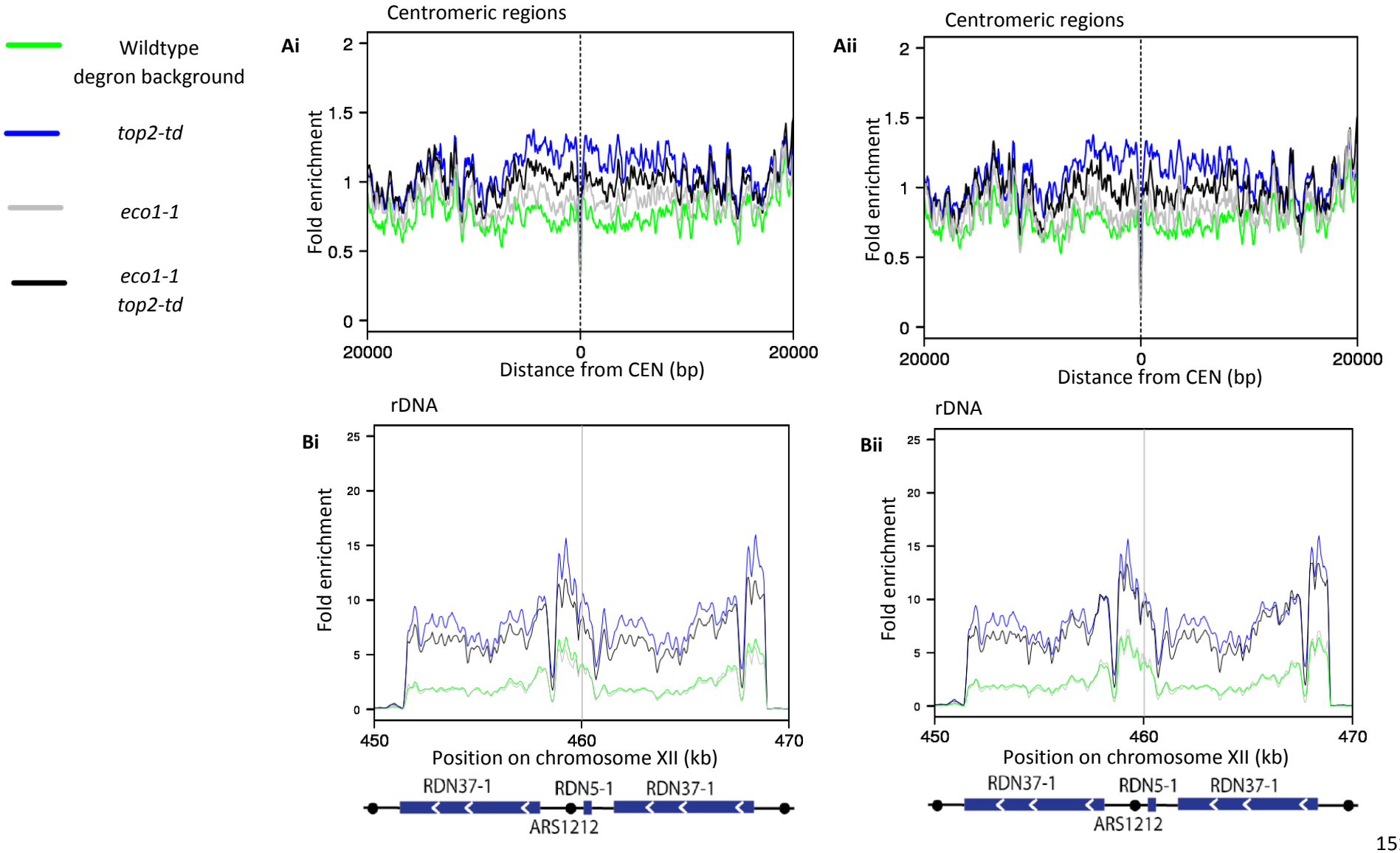

Wildtype
degron background
top2-td
eco1-1
eco1-1
top2-td
Ai
Centromeric regions
Fold enrichment
Distance from CEN (bp)
Aii
Centromeric regions
Fold enrichment
Distance from CEN (bp)
Bi
rDNA
Fold enrichment
Position on chromosome XII (kb)
RDN37-1
RDN5-1
RDN37-1
ARS1212
Bii
rDNA
Fold enrichment
Position on chromosome XII (kb)
RDN37-1
RDN5-1
RDN37-1
ARS1212

Figure 5.7 Smc3 acetylation can modify the extent of cohesin associated DNA damage

Ai. *eco1-1* and *eco1-1 top2-td* shown with degron background wildtype and *top2-td* for comparison, γH2A fold enrichment over H2A, for a 1 – 16 centromere pile up of a 40 kb region around the centromeres A**ii.** Repeat of Ai *eco1-1* and *eco1-1 top2-td* **Bi.** *eco1-1* and eco1-1 *top2-td* shown with degron background wildtype and *top2-td* for comparison, γH2A fold enrichment over H2A, for Chromosome XII 20 kb rDNA array B**ii.** Repeat of Bi *eco1-1* and *eco1-1 top2-td*

I finally went on to analyse the DNA damage profile when knocking out the antagonist of Eco1, *RAD61* (Wapl). In chapter 4, I found that a *RAD61* deletion lead to a lowering of DNA replication fork rotation on a plasmid, here I used *rad61Δ top2-td* cells to look at topological stress-associated DNA damage. Surprisingly I found no decreases in DNA damage around the centromeres when *RAD61* was deleted compared to *top2-td* cells (Figure 5.8A). At the rDNA, *rad61Δ top2-td* had a slightly lowered level of DNA damage across the 35S regions of the rDNA array, compared to *top2-td* alone (Figure 5.8B).

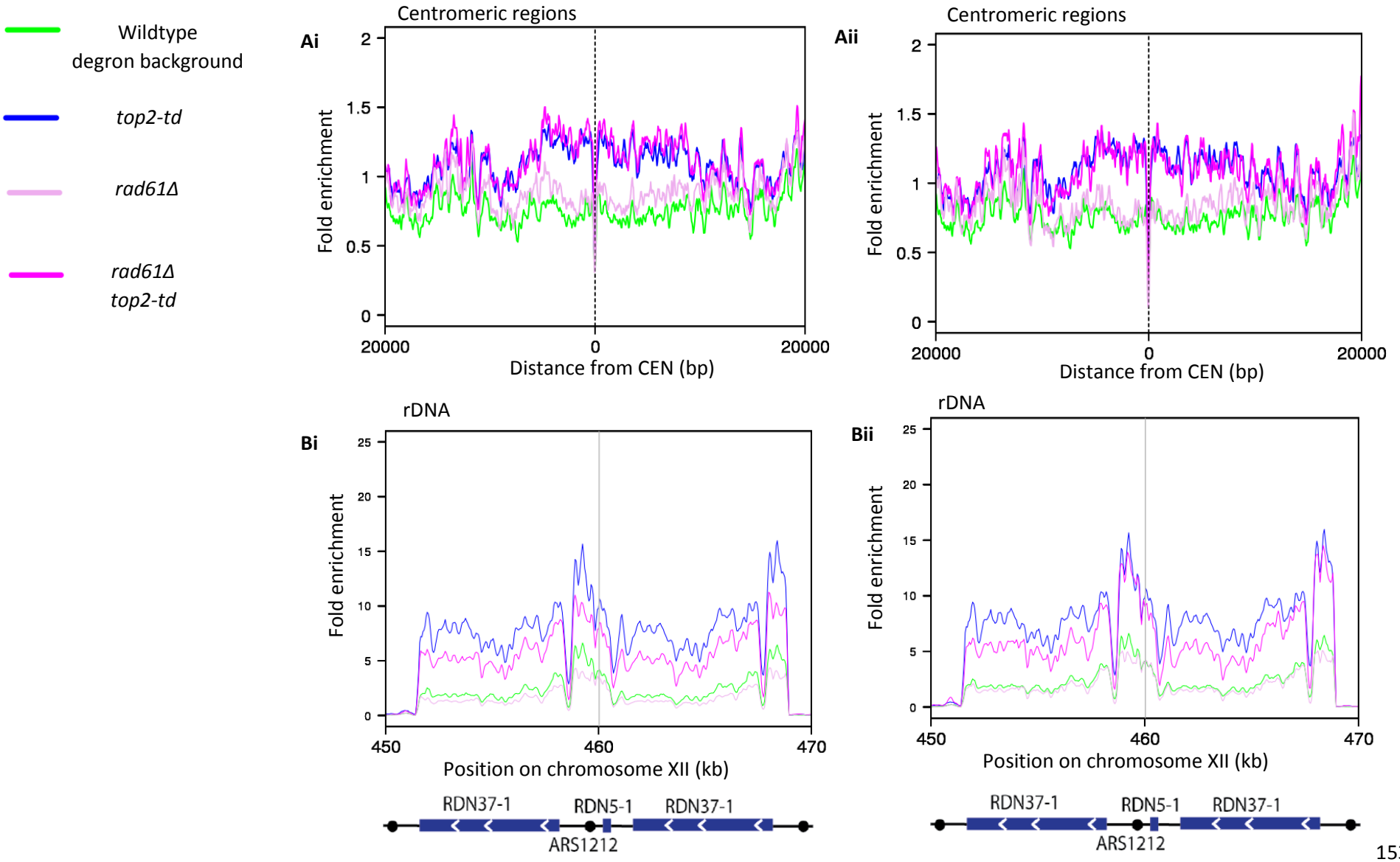

Wildtype
degron background
top2-td
rad61Δ
rad61Δ
top2-td
Ai
Centromeric regions
Fold enrichment
Distance from CEN (bp)
Aii
Centromeric regions
Fold enrichment
Distance from CEN (bp)
Bi
rDNA
Fold enrichment
Position on chromosome XII (kb)
RDN37-1
RDN5-1
RDN37-1
ARS1212
Bii
rDNA
Fold enrichment
Position on chromosome XII (kb)
RDN37-1
RDN5-1
RDN37-1
ARS1212

Figure 5.8 *RAD61* deletion leads to no rescue of DNA damage at centromeres and very slight rescue at the rDNA

Ai. *rad61Δ* and *rad61Δ top2-td* shown with degron background wildtype and *top2-td* for comparison, γH2A fold enrichment over H2A, for a 1 – 16 centromere pile up of a 40 kb region around the centromeres **Aii.** Repeat of Bi *rad61Δ* and *rad61Δ top2-td* **Bi.** *rad61Δ* and *rad61Δ top2-td* shown with degron background wildtype and *top2-td* for comparison, γH2A fold enrichment over H2A, for Chromosome XII 20 kb rDNA array B**ii.** Repeat of *rad61Δ* and *rad61Δ top2-td*

5.7 DISCUSSION

5.7.1 DNA damage during S-phase

In this chapter I have aimed to understand the S-phase specific DNA damage that occurs due to topological stress. I have shown that cohesin activity causes DNA topological stress during S-phase, and that accumulation of this stress in the absence of Top2 leads to endogenous DNA damage.

I started by analysing the levels of DNA damage though S-phase by analysing wildtype γH2A ChIP profiles. I found that γH2A levels were generally low across the genome, however, there were a small number of peaks at particular loci. There were clear increases at certain sub-telomeric regions, which appeared to be the same in each repeat, and was also previously seen in asynchronous budding yeast cells (Figure 5.1) (Kim et al., 2007; Szilard et al., 2010) (Figure 5.1A).). It was suggested by Kim et al. (2007) that telomeres are transiently recognised as double strand breaks, and the γH2A signal accrued, persists through cell cycles due to the telomere ends not being 'repaired'. The γH2A accumulation could also be due to telomeres being hard to replicate sites; the specialist helicase Rrm3 is required for replication through telomeres (Ivessa et al., 2002). However, even wildtype replication forks slow at telomeric regions, and replication fork pausing can begin 100 bp upstream of the telomere tracts (Makovets et al., 2004). This replication stress could be one factor contributing to the S-phase γH2A signal seen at these regions. Apart from the telomeric regions, the rDNA stands out as a region with high levels of endogenous DNA damage, which has again been seen previously (Figure 5.1) (Szilard et al., 2010). This high level of DNA damage could be due to replication stress from the replication fork barrier and repeated collisions with stalled RNA polymerases (Labib and Hodgson, 2007).

5.7.2 Topological stress associated DNA damage during S-phase

To understand DNA damage due to S-phase topological stress, I further removed Top2, a protein that can relax DNA supercoiling and resolve DNA catenation, in order to resolve topological stress or DNA entanglements (Vos et al., 2011). I surprisingly found that removing Top2 during S-phase did not generally increase DNA damage across the genome (Figures 5.1, 5.2). This suggests that Top1 is generally sufficient to compensate for loss of Top2 during S-phase. However, at two specific loci there were large increases in γH2A signal; the centromeres and the rDNA (Figures 5.1, 5.2). Which lead me to conclude that the loss of specific activities of Top2 (for which Top1 cannot fully compensate), increases topological stress, and leads to DNA replication-dependent DNA damage.

5.7.3 Cohesin promotes DNA damage associated with topological stress

As was outlined in chapter 4, the cohesin complex appears to have a role in promoting topological stress during S-phase. It was seen to promote replication fork rotation on plasmids. The cohesin complex is also known to be loaded at both the core centromeres and the 5S region of the rDNA (Laloraya et al., 2000). Which lead me on to test whether cohesin was having a role in topological stress-associated DNA damage. I found that an inactivation of cohesin alone, did not result in any clear differences in γH2A signal across the genome, compared to wildtype (Figures 5.3, 5.4). However, I observed that inactivation of cohesin in conjunction with a Top2 depletion, suppressed the accumulation of the Top2 depletion-dependent γH2A signal at yeast centromeres. Moreover, I found that preventing cohesin loading at centromeres via *ndc10-1 top2-td* mutant cells partially rescued the topological stress associated DNA damage (Figure 5.6).

At the rDNA array, a cohesin inactivation with a Top2 depletion resulted in a rescue of the Top2-depletion dependent damage at the 35S region of the rDNA arrays, although not to wildtype levels. However at the rDNA 5S region, where cohesin loads on the DNA, there was no clear rescue of the Top2-depletion dependent DNA damage (Figures 5.3, 5.4). This result suggested that the cohesin dependent DNA damage is at places of cohesin translocation, due to the regions of high transcriptional activity, rather than the areas of cohesin loading. This hypothesis was looked at in Figure 5.5, where a cohesin S-phase ChIP was compared to the centromere and rDNA cohesin dependent DNA damage profile. The lowest levels of cohesin dependent damage were the sites of loading, where the highest were in places of cohesin translocation (Figure 5.5A, B). Across the rDNA this correlates with highly expressed genic regions, however next to the

centromere it does not. The common factor to the centromeres and the rDNA is that cohesin dependent DNA damage occurs across regions where cohesin translocates, either when promoted by transcription (at the rDNA) or otherwise (centromeres). There has also been some evidence that loss of cohesin leads to a downregulation of transcription of rDNA genes (Gartenberg and Smith, 2016), which could be affecting the levels of topological stress here and therefore leading to lower levels of DNA damage. However, the transcriptional dependent damage was easily seen at the highly expressed *GAL* genes (Figure 5.5B, C). Which is further evidence of both the importance of the link between transcription and cohesin, and the translocation of cohesin for topological stress.

5.7.4 Role of cohesin accessory proteins in introducing topological stress

I next wanted to understand the roles of two accessory proteins to cohesin, Eco1 and Rad61. Eco1 acetylation of cohesin during S-phase is known to regulate replication fork dynamics in human cells (Terret et al., 2009), thereby showing the importance of the interaction of cohesin and the replication fork. When analysing DNA damage during S-phase, I found that the removal of Eco1 with Top2, partially suppressed the γH2A accumulation across the centromeres (figure 5.7). However, *eco1-1 top2-td* showed no rescue across the rDNA array. This is in line with the results from chapter 4, which showed a small non-significant decrease in replication fork rotation when Eco1 was depleted in conjunction with Top2. Clearly, Smc3 acetylation, which affects only a small subset of cohesin molecules, can modify the extent of cohesin-associated DNA damage. However, it is not specifically required for the cohesin-dependent damage to accumulate.

A *RAD61* deletion surprisingly lead to no rescue of DNA damage across the centromeres, and only a slight rescue at the rDNA (Figure 5.8). This is in contrast to the effects on replication fork rotation on plasmids seen in chapter 4; when *RAD61* was deleted, there was significantly less fork rotation during one round of replication (Figure 4.8). It was concluded that the lowering of the numbers of cohesin molecules that occur with a *RAD61* deletion lead to this phenotype. However, I saw almost no rescue of the topological stress associated DNA damage. This paradox may be due to the effects of Rad61 only affecting centromeres on plasmid DNA and not centromeres on endogenous chromosomes. This is reflected in the difference in the effect of Ndc10 in either situation as well as Rad61. When Ndc10 was depleted in a plasmid situation, the level of catenanes formed during S-phase was unchanged. However, when Ndc10 was depleted,

and DNA damage analysed along endogenous chromosomes, there was a clear rescue of the Top2-dependent DNA damage (Figure 5.6). Potentially the plasmid replicon, which will have lowered levels of cohesin localisation compared to endogenous chromosomes, is more greatly affected by small changes in the number of cohesin molecules available. Arguably, the plasmid replicon represents a pericentromeric region, whereas endogenous chromosomes also contain arm regions, which build up cohesin in a distinct manner to the centromeres (Kogut et al., 2009). This distinction between the two situations could be affecting the difference in the pervasiveness of *rad61Δ*.

These data could indicate a threshold level of cohesin related topological stress required to lead to DNA damage, separate to the levels required for DNA replication fork rotation to occur. Since cohesin is loaded at centromeres and rDNA during S-phase, and increases until anaphase, even with lowered SCC1 transcription levels, enough may load at the centromeres and rDNA to lead to a level of topological stress causing DNA damage. Cohesin is loaded much more abundantly at the rDNA than at the centromeres (Hu et al., 2015), and the DNA damage at the rDNA is much greater (Figures 5.1, 5.2). The still high, but slightly lowered DNA damage phenotype when Rad61 is deleted may be reflecting a slightly lower abundance of cohesin loading here. In the case of the lowered topological stress on a plasmid, there is a possibility that much less cohesin loads here in the absence of Rad61, as compared to at centromeres and rDNA on chromosomes.

5.7.5 Conclusions and future directions

This chapter has followed on from the work in chapters 3 and 4, discerning the role of transcription and cohesin in topological stress during S-phase. In this chapter I focused on DNA damage as marked by γH2A (H2A S129P) accumulation related to topological stress during S-phase. I found that topological stress associated DNA damage was restricted in most part to 10 kb either side of the centromeres and at the rDNA array. I went on to show that this DNA damage was dependent on the cohesin complex. I also looked at transcription and cohesin translocation, and found that at highly expressed genes transcription was important for cohesin dependent DNA damage. However, transcription was shown not to be absolutely required, since there was high levels of cohesin dependent DNA damage at the non-transcribed centromere regions. At the transcribed regions, it appeared that transcription may be exacerbating the cohesin-dependent effect by both adding to the topological stress, and leading to high rates of cohesin translocation.

In the future, further work could be undertaken to look at the links between transcription and cohesin, and whether the translocation of cohesin along DNA, or the accumulation of cohesin is what is important for the DNA dependent damage. It would also be interesting to understand the role of Rad61 in cohesin-dependent topological changes due to the conflicting results seen in chapters 4 and 5.

6 DISCUSSION

During this thesis I have aimed to understand the impact of DNA topological stress during DNA replication. I have looked at how transcription during S-phase impacts on DNA replication fork rotation, how cohesin is involved in promoting topological stress during S-phase, and how cohesin induced topological stress can lead to DNA replication stress and DNA damage.

6.1 TRANSCRIPTION DURING DNA REPLICATION INCREASES TOPOLOGICAL STRESS, LEADING TO REPLICATION FORK ROTATION

I started my analysis of DNA topological stress during S-phase, by looking at transcription during DNA replication, an area well known for causing replication stress. Replication–transcription collisions have a highly genotoxic nature (Helmrich et al., 2011; Prado and Aguilera, 2005), and there are many strategies in place to minimise numbers of collisions (Meryet-Figuiere et al., 2014, Pope et al., 2014, Wei et al., 1998), clearly indicating that the meeting of two forks causes many problems. DNA transcription itself causes topological changes around the transcription bubble; it forms the twin supercoiled domain, where positive supercoiling forms ahead of the bubble and negative supercoiling forms behind. I wanted to further understand the role of topological stress in collisions; and to do this I utilised a method used in our lab to identify places of high topological stress causing DNA replication fork rotation.

I initially found that increasing numbers of genes on a plasmid, by adding a convergent gene pair, led to high levels of replication fork rotation in the absence of Top2 (Figure 3.2). This could be due to the RNA polymerase acting as a block to DNA replication, the diffusion of DNA supercoiling, and the access of DNA topoisomerases to relieve the stress. In the case of a head-on replication-transcription collision, the RNA polymerase and DNA replication fork converge similarly to the termination of DNA replication. When they converge, the supercoiling domains in front of each would also join leading to high levels of topological stress, and topoisomerases are likely to be sterically excluded from relaxing the supercoiling. In a situation of a co-directional replication-transcription collision the replication fork would converge upon a paused RNA polymerase or a non B-form DNA structure, leading to replication fork rotation in a manner analogous to that at a stable protein-DNA structure.

In addition to the topological stress generated by collisions of the RNA polymerases and the replication fork, the convergent gene pair could be leading to high levels of topological stress between their 3' ends. However, I ruled this out as a major factor in the transcription induced increase in replication fork rotation. I found that changing the directionality of the genes on the plasmid, so that all were in the same orientation as one another, still lead to increased levels of replication fork rotation during S-phase (Figure 3.4). Therefore, in chapter 3 I showed that an increase in gene units on a plasmid, increases the levels of topological stress-associated replication fork rotation. I found that this appears to directly correlate to the number of genes on the plasmid as oppose to the orientation of the collision.

6.2 COHESIN TRANSLOCATION PROMOTES TRANSCRIPTION-DEPENDENT REPLICATION FORK ROTATION

In Chapter 4, I went on to analyse the role of the cohesin complex on topological stress during S-phase. I chose to look at the cohesin complex both because of the possibility of the tracking SMC complexes being a barrier to the replication fork, but also due to cohesin's link with transcription and being translocated in an RNA polymerase dependent manner. Therefore, following on from the finding that transcription itself increases the level of replication fork rotation on a centromeric plasmid, I found that this is exacerbated by the presence of cohesin (Figures 4.1 – 4.4). Since transcription is closely linked to the positioning and movement of cohesin molecules, this argues that it is the translocation of cohesin per se that is leading to the topological stress phenotype. Additionally, a plasmid with very low transcription did not appear to have much, if any, decrease in topological stress without cohesin, compared to plasmids with multiple active genes. In this case, I suggest that cohesin still loads onto the low transcription plasmid, but it does not translocate in an RNA polymerase manner. My data indicates that it is the combined topological changes resulting from multiple transcription units and cohesin, which leads to the high levels of replication fork rotation.

In the case of the plasmid with additional *tRNA* genes, there was an inconsistent effect of removing cohesin. In some experiments it lowered levels of fork rotation and in some experiments it did not (Figure 4.5). Again, the translocation of cohesin could be important for the increases in topological stress, and here, although there are high levels of transcription, the transcripts are very short. Meaning that less translocation would be occurring, reflecting the inconsistent changes in replication fork rotation. I also suggested that the SMC complex

condensin may be an important modulator here, as it is known to be recruited to *tRNA* genes. Potentially cohesin could be modifying an effect of condensin at *tRNA* genes. This is consistent with the ChIP-seq data, as *tRNA* genes did not arise as areas of cohesin-dependent topological stress that leads to DNA damage in the absence of Top2 (Figures 5.1, 5.2). Perhaps looking at the condensin complex at these genes may further elucidate what is occurring here.

The idea of cohesin translocation being important for its effect on DNA topology can also be argued from the ChIP-seq data in Chapter 5. The highest cohesin-dependent DNA damage is at areas proximal to the loading sites, rather than exactly at the point of loading. This was seen at the rDNA, where there is an obvious change in DNA damage from the RDN5 region and 35S regions (Figure 5.5). The highly transcribed 35S regions have high levels of topological stress-associated cohesin-dependent damage. In contrast, the rDNA 5S (RDN5) area of loading appeared to have relatively little cohesin-dependent damage. This is most easily seen when a cohesin ChIP (Hu et al., 2015) is compared to the damage ChIP data (Figure 5.5). Where cohesin is situated, are the areas of lowest cohesin-dependent damage, but the regions of high transcription, thus areas of cohesin translocation are the areas of highest cohesin-dependent DNA damage (figure 5.5). This was not only seen at sites of loading, but also at arm transcription sites and a general correlation with areas in-between cohesin accumulation points (Figure 5.5).

If it is the transcription-driven translocation of cohesin that is important for DNA topological stress, this suggests that cohesin dependent extrusion could be involved, rather than cohesin complexes performing the function of sister chromatid cohesion. It is known that the numbers of cohesin molecules required for sister chromatid cohesion are low, compared to the overall number of cohesin complexes in the cell, suggesting that most cohesin is engaged in processes not directly related to sister chromatid cohesion, including chromosome compaction (Heidinger-Pauli et al., 2010). Additionally, loop extrusion in *S. cerevisiae* has been shown to be mainly driven by cohesin on chromosomes (Schalbetter et al., 2017). It could be, that in the process of extruding, large topological entanglements build up at the anchor points of cohesin binding, which would then require Top2 activity to be alleviated, and allow for continuing loop extrusion. In the absence of Top2, large entanglements build up where cohesin is translocating, leading to cohesin associated DNA damage. It has already been suggested in human cells that loop extrusion may lead to DNA topological stress, and was found that Top2B generated DNA double strand breaks at cohesin loop anchors (CTCF sites) (Figure 6.1) (Canela et al., 2017).

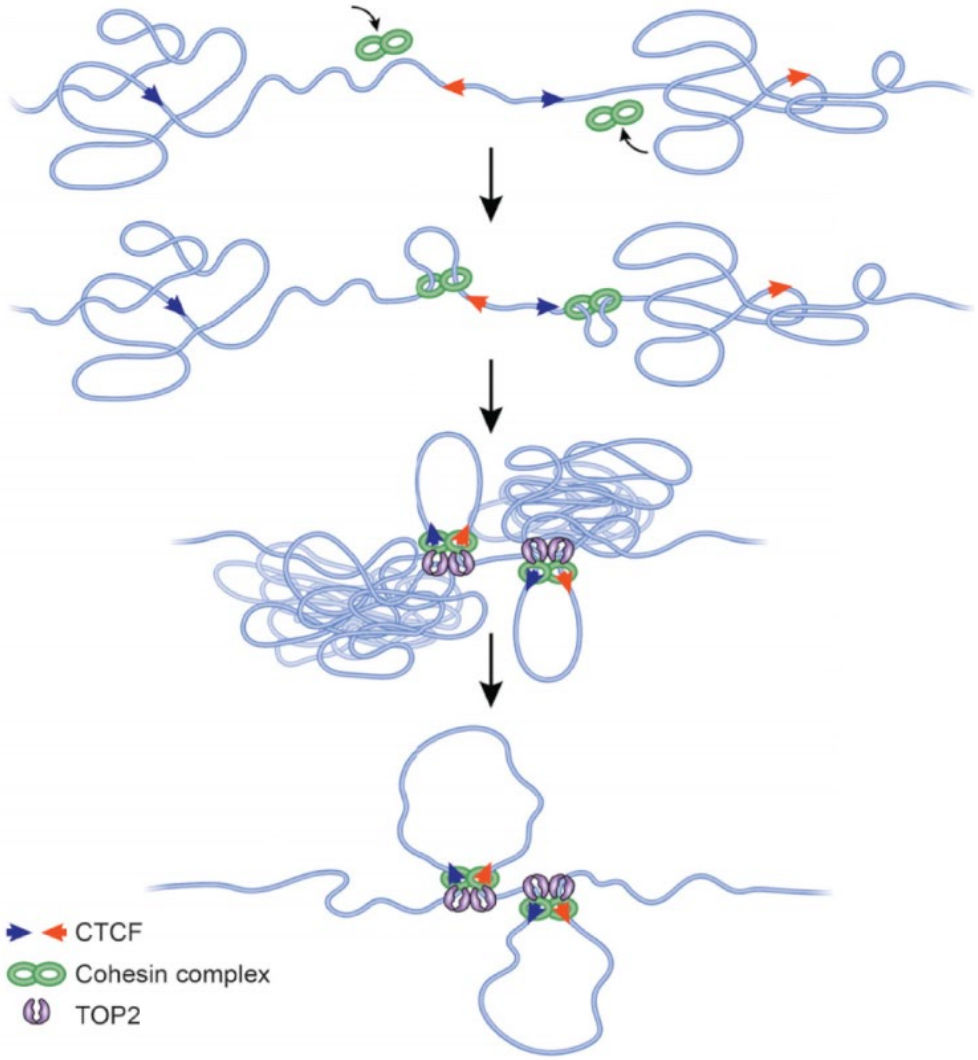


Figure 6.1 A model for cohesin promoted topological stress

Modified from Canela et al (2017). Cohesin extrudes DNA to create interphase chromosome organisation. The extrusion leads to DNA topological entanglements, which can be resolved by the action of Top2.

6.3 COHESIN AS A CAUSE OF DNA REPLICATION STRESS

Cohesin is required during S-phase to promote sister chromatid cohesion, whereas the other SMC complexes are less abundant. Condensin binding to DNA is very low in G1 phase and only binds strongly after DNA replication (Leonard et al., 2015), SMC5/6 binding is also very low in G1 phase and requires previous binding of cohesin for localisation (Jeppsson et al., 2014). Therefore, it appears that where possible, SMC complexes are kept very low during S-phase. Could this be to prevent interference with the replication fork? I propose that levels are kept low to prevent SMC associated replication stress. DNA replication stress is a complex variety of different problems that lead to slowing or stalling of replication forks, which can lead to DNA breakage at common fragile sites (CFS), and genome instability (Zeman and Cimprich, 2014). Replication stress is well known to be a major driver of tumorigenesis, it leads to genomic instability, causing further mutations which can drive tumour progression (Hills and Diffley, 2014). Common Fragile sites are found to be abundant at areas of high replication-transcription collision, implicating this as one of the main causes of replication stress (Barlow et al., 2013; Helmrich et al., 2013). However what have not, until this point, been directly associated with replication stress are the SMC complexes.

A study in *S.cerevisiae* showed that cohesin levels that were reduced to 13% of their original levels, caused no loss of sister chromatid cohesion (Heidinger-Pauli et al., 2010). This poses the possibility that cohesin is kept at low levels at the start of S-phase to minimise DNA damage, but still can maintain its cohesive functionality. Additionally, centromeric DNA in budding yeast is one of the earliest replicated regions, which could potentially aid to prevent DNA damage occurring due to a build-up of cohesin molecules here (McCarroll and Fangman, 1988). In Chapter 5 of this thesis I saw specific DNA damage at the centromeres and rDNA, the places in the genome which have the very highest level of cohesin accumulation during S-phase (Hu et al., 2015) (Figures 5.1, 5.2). Suggesting that there is a threshold level of cohesin associated topological stress that will lead to DNA damage in the absence of Top2.

6.4 HOW DOES COHESIN PROMOTE TOPOLOGICAL STRESS AND DNA DOUBLE STRAND BREAKS?

I propose three different models for how cohesin is promoting topological stress, and leads to Top2 dependent DNA damage. Canela et al. (2017) proposed that Cohesin associated topological

stress is via loop extrusion generated DNA entanglements (Figure 6.1). They looked at cancer translocation clusters susceptible to DNA breakage, and found Top2 promotes DNA damage because it is required to cause breaks to resolve the DNA entanglements. However, here my data suggests that the cause of the DNA damage comes from the topological stress itself, and that the resolution of this by Top2 prevents DNA breaks. Although, in a cancerous situation, where Top2 poisons are being used therapeutically, Top2 dependent DNA breaks are also likely to lead to DNA damage. My first model fits with the idea suggested by Canela et al. (2017), I propose that cohesin loop extrusion creates topological stress as it is pulling DNA though, which is therefore generated at the base of DNA loops. Top2 will then resolve the topological stress, preventing DNA damage occurring (Figure 6.1). In this case, the DNA damage occurring when Top2 is removed, would occur due to collisions with the replication fork. It would both act as a physical barrier to replication fork progression, and as a barrier to the diffusion of replication associated supercoiling. This could lead to a situation such as replication fork rotation, creating high numbers of DNA pre-catenanes behind the fork and leading to DNA damage. Or it could lead to a situation such as slowing and stalling of the replication fork, leading to fork breakdown and subsequent DNA damage.

My second proposed model, is a transcription dependent build-up of DNA topological stress. Here I propose that cohesin translocation via transcriptional elongation, leads to DNA topological stress, due to the prevention of diffusion of transcription dependent supercoiling (Figure 6.2). The positive supercoiling forming ahead of the transcription bubble can mechanistically drive cohesin along the DNA, allowing for the movement of the complexes away from their abundant loading sites. I propose that the prevention of the diffusion of positive supercoiling and subsequent prevention of resolution of this same supercoiling by sterical hindrance, leads to high DNA topological stress (Figure 6.2). If this is then encountered by a replication fork, it will again provide a barrier and high levels of topological stress which can lead to similar situations of DNA damage as in the first model.

The final model comes from the idea that SMC complexes themselves introduce DNA topological stress. It has been shown *in vitro* that cohesin and condensin binding to DNA, changes the topology of DNA (Kimura and Hirano, 1997; Sun et al., 2013). Therefore, could cohesin tracking along DNA lead to a build-up of topological stress due to its intrinsic mechanism? My inability to observe cohesin dependent topological change in S-phase (in contrast to condensin dependent action in mitosis) argues against this possibility (Figure 4.7).

Why none of these models lead to DNA damage generally, during DNA replication, would be because of Topoisomerase action. The levels of topological stress would not reach the threshold required to lead to DNA damage due to topoisomerases resolving this stress. In these models the removal of Top2 leads to higher than normal levels of topological stress, or it leads to unresolved DNA pre-catenanes, leading to DNA damage.

Further work will be required to elucidate which of these models is the likeliest to lead to cohesin dependent topological stress, and how widespread this novel source of endogenous replication stress is in cells.

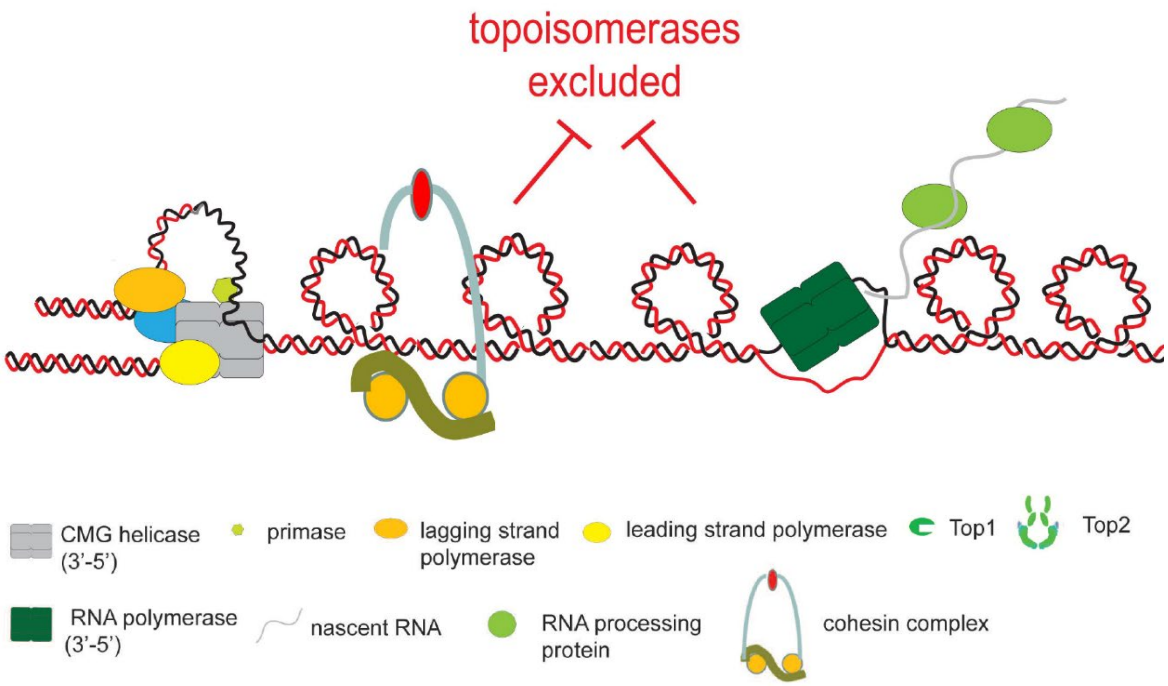


Figure 6.2 Model of a build-up of topological stress due to transcriptional elongation and cohesin

A theoretical view on the interplay between transcription, cohesin and replication stress. Cohesin prevents the dissipation of topological stress due to transcriptional elongation. This leads to high levels of topological stress. If encountered by a DNA replication fork it will represent a barrier to DNA replication. This situation could lead to replication fork rotation to diffuse topological stress, or it could lead to replication fork pausing. In the situation of a Top2 depletion, DNA damage could occur due to either high levels of DNA catenanes, or fork stalling and misprocessing leading to DNA damage.

Bibliography

Alt, A., Dang, H.Q., Wells, O.S., Polo, L.M., Smith, M.A., McGregor, G.A., Welte, T., Lehmann, A.R., Pearl, L.H., Murray, J.M*., et al.* (2017). Specialized interfaces of Smc5/6 control hinge stability and DNA association. Nat Commun *8*, 14011.

Alzu, A., Bermejo, R., Begnis, M., Lucca, C., Piccini, D., Carotenuto, W., Saponaro, M., Brambati, A., Cocito, A., Foiani, M*., et al.* (2012). Senataxin associates with replication forks to protect fork integrity across RNA-polymerase-II-transcribed genes. Cell *151*, 835-846.

Azvolinsky, A., Giresi, P.G., Lieb, J.D., and Zakian, V.A. (2009). Highly transcribed RNA polymerase II genes are impediments to replication fork progression in Saccharomyces cerevisiae. Molecular cell *34*, 722-734.

Bairwa, N.K., Zzaman, S., Mohanty, B.K., and Bastia, D. (2010). Replication fork arrest and rDNA silencing are two independent and separable functions of the replication terminator protein Fob1 of Saccharomyces cerevisiae. The Journal of biological chemistry *285*, 12612-12619.

Baranello, L., Wojtowicz, D., Cui, K., Devaiah, B.N., Chung, H.J., Chan-Salis, K.Y., Guha, R., Wilson, K., Zhang, X., Zhang, H*., et al.* (2016). RNA Polymerase II Regulates Topoisomerase 1 Activity to Favor Efficient Transcription. Cell *165*, 357-371.

Barlow, J.H., Faryabi, R.B., Callen, E., Wong, N., Malhowski, A., Chen, H.T., Gutierrez-Cruz, G., Sun, H.W., McKinnon, P., Wright, G*., et al.* (2013). Identification of early replicating fragile sites that contribute to genome instability. Cell *152*, 620-632.

Bash, R., and Lohr, D. (2001). Yeast chromatin structure and regulation of GAL gene expression. Prog Nucleic Acid Res Mol Biol *65*, 197-259.

Bates, A.D., and Maxwell, A. (2005). DNA Topology (Oxford University Press).

Bauerschmidt, C., Pollok, S., Kremmer, E., Nasheuer, H.P., and Grosse, F. (2007). Interactions of human Cdc45 with the Mcm2-7 complex, the GINS complex, and DNA polymerases delta and epsilon during S phase. Genes Cells *12*, 745-758.

Baxter, J. (2014). "Breaking Up Is Hard to Do": The Formation and Resolution of Sister Chromatid Intertwines. J Mol Biol.

Baxter, J., and Aragon, L. (2012). A model for chromosome condensation based on the interplay between condensin and topoisomerase II. Trends in genetics : TIG *28*, 110-117.
Baxter, J., and Diffley, J.F. (2008). Topoisomerase II inactivation prevents the completion of DNA replication in budding yeast. Molecular cell *30*, 790-802.

Baxter, J., Sen, N., Martinez, V.L., De Carandini, M.E.M., Schvartzman, J.B., Diffley, J.F.X., and Aragon, L. (2011). Positive Supercoiling of Mitotic DNA Drives Decatenation by Topoisomerase II in Eukaryotes. Science *331*, 1328-1332.

Bell, S.P., and Stillman, B. (1992). ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. Nature *357*, 128-134.

Bermejo, R., Capra, T., Gonzalez-Huici, V., Fachinetti, D., Cocito, A., Natoli, G., Katou, Y., Mori, H., Kurokawa, K., Shirahige, K.*, et al.* (2009). Genome-organizing factors Top2 and Hmo1 prevent chromosome fragility at sites of S phase transcription. Cell *138*, 870-884.

Bermejo, R., Capra, T., Jossen, R., Colosio, A., Frattini, C., Carotenuto, W., Cocito, A., Doksani, Y., Klein, H., Gomez-Gonzalez, B.*, et al.* (2011). The replication checkpoint protects fork stability by releasing transcribed genes from nuclear pores. Cell *146*, 233-246.

Bermejo, R., Doksani, Y., Capra, T., Katou, Y.M., Tanaka, H., Shirahige, K., and Foiani, M. (2007). Top1- and Top2-mediated topological transitions at replication forks ensure fork progression and stability and prevent DNA damage checkpoint activation. Genes & development *21*, 1921-1936.

Bermudez, V.P., Maniwa, Y., Tappin, I., Ozato, K., Yokomori, K., and Hurwitz, J. (2003). The alternative Ctf18-Dcc1-Ctf8-replication factor C complex required for sister chromatid cohesion loads proliferating cell nuclear antigen onto DNA. Proceedings of the National Academy of Sciences of the United States of America *100*, 10237-10242.

Bhatia, V., Herrera-Moyano, E., Aguilera, A., and Gomez-Gonzalez, B. (2017). The Role of Replication-Associated Repair Factors on R-Loops. Genes (Basel) *8*.

Bizard, A.H., and Hickson, I.D. (2014). The dissolution of double Holliday junctions. Cold Spring Harbor perspectives in biology *6*, a016477.

Bjorkegren, C., and Baranello, L. (2018). DNA Supercoiling, Topoisomerases, and Cohesin: Partners in Regulating Chromatin Architecture? International journal of molecular sciences *19*.
Blat, Y., and Kleckner, N. (1999). Cohesins bind to preferential sites along yeast chromosome III, with differential regulation along arms versus the centric region. Cell *98*, 249-259.

Blattner, F.R., Plunkett, G., 3rd, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F.*, et al.* (1997). The complete genome sequence of Escherichia coli K-12. Science *277*, 1453-1462.

Blobel, G. (1985). Gene gating: a hypothesis. Proceedings of the National Academy of Sciences of the United States of America *82*, 8527-8529.

Blow, J.J., Ge, X.Q., and Jackson, D.A. (2011). How dormant origins promote complete genome replication. Trends in biochemical sciences *36*, 405-414.

Borrie, M.S., Campor, J.S., Joshi, H., and Gartenberg, M.R. (2017). Binding, sliding, and function of cohesin during transcriptional activation. Proceedings of the National Academy of Sciences of the United States of America *114*, E1062-E1071.

Boubakri, H., de Septenville, A.L., Viguera, E., and Michel, B. (2010). The helicases DinG, Rep and UvrD cooperate to promote replication across transcription units in vivo. The EMBO journal *29*, 145-157.

Brewer, B.J., and Fangman, W.L. (1988). A replication fork barrier at the 3' end of yeast ribosomal RNA genes. Cell *55*, 637-643.

Brill, S.J., DiNardo, S., Voelkel-Meiman, K., and Sternglanz, R. (1987). Need for DNA topoisomerase activity as a swivel for DNA replication for transcription of ribosomal RNA. Nature *326*, 414-416.

Brill, S.J., and Sternglanz, R. (1988). Transcription-dependent DNA supercoiling in yeast DNA topoisomerase mutants. Cell *54*, 403-411.

Brown, P.O., and Cozzarelli, N.R. (1979). A sign inversion mechanism for enzymatic supercoiling of DNA. Science *206*, 1081-1083.

Brown, P.O., and Cozzarelli, N.R. (1981). Catenation and knotting of duplex DNA by type 1 topoisomerases: a mechanistic parallel with type 2 topoisomerases. Proceedings of the National Academy of Sciences of the United States of America *78*, 843-847.

Busslinger, G.A., Stocsits, R.R., van der Lelij, P., Axelsson, E., Tedeschi, A., Galjart, N., and Peters, J.M. (2017). Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. Nature *544*, 503-507.

Cadwell, C., Yoon, H.J., Zebarjadian, Y., and Carbon, J. (1997). The yeast nucleolar protein Cbf5p is involved in rRNA biosynthesis and interacts genetically with the RNA polymerase I transcription factor RRN3. Molecular and cellular biology *17*, 6175-6183.

Calzada, A., Hodgson, B., Kanemaki, M., Bueno, A., and Labib, K. (2005). Molecular anatomy and regulation of a stable replisome at a paused eukaryotic DNA replication fork. Genes & development *19*, 1905-1919.

Canela, A., Maman, Y., Jung, S., Wong, N., Callen, E., Day, A., Kieffer-Kwon, K.R., Pekowska, A., Zhang, H., Rao, S.S.P*., et al.* (2017). Genome Organization Drives Chromosome Fragility. Cell *170*, 507-521 e518.

Champoux, J.J., and Been, M.D. (1980). TOPOISOMERASES AND THE SWIVEL PROBLEM1. In Mechanistic Studies of DNA Replication and Genetic Recombination, B. Alberts, ed. (Academic Press), pp. 809-815.

Champoux, J.J., and Dulbecco, R. (1972). An activity from mammalian cells that untwists superhelical DNA--a possible swivel for DNA replication (polyoma-ethidium bromide-mouse-embryo cells-dye binding assay). Proceedings of the National Academy of Sciences of the United States of America *69*, 143-146.

Chan, K.L., Gligoris, T., Upcher, W., Kato, Y., Shirahige, K., Nasmyth, K., and Beckouet, F. (2013). Pds5 promotes and protects cohesin acetylation. Proceedings of the National Academy of Sciences of the United States of America *110*, 13020-13025.

Chan, K.L., Roig, M.B., Hu, B., Beckouet, F., Metson, J., and Nasmyth, K. (2012). Cohesin's DNA exit gate is distinct from its entrance gate and is regulated by acetylation. Cell *150*, 961-974.
Chan, Y.A., Aristizabal, M.J., Lu, P.Y., Luo, Z., Hamza, A., Kobor, M.S., Stirling, P.C., and Hieter, P. (2014). Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-chip. PLoS Genet *10*, e1004288.

Chang, M., Bellaoui, M., Zhang, C., Desai, R., Morozov, P., Delgado-Cruzata, L., Rothstein, R., Freyer, G.A., Boone, C., and Brown, G.W. (2005). RMI1/NCE4, a suppressor of genome

instability, encodes a member of the RecQ helicase/Topo III complex. The EMBO journal *24*, 2024-2033.

Cherry, J.M., Hong, E.L., Amundsen, C., Balakrishnan, R., Binkley, G., Chan, E.T., Christie, K.R., Costanzo, M.C., Dwight, S.S., Engel, S.R.*, et al.* (2012). Saccharomyces Genome Database: the genomics resource of budding yeast. Nucleic Acids Res *40*, D700-705.

Cho, U.S., and Harrison, S.C. (2011). Ndc10 is a platform for inner kinetochore assembly in budding yeast. Nat Struct Mol Biol *19*, 48-55.

Ciosk, R., Shirayama, M., Shevchenko, A., Tanaka, T., Toth, A., Shevchenko, A., and Nasmyth, K. (2000). Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins. Molecular cell *5*, 243-254.

D'Ambrosio, C., Schmidt, C.K., Katou, Y., Kelly, G., Itoh, T., Shirahige, K., and Uhlmann, F. (2008). Identification of cis-acting sites for condensin loading onto budding yeast chromosomes. Genes & development *22*, 2215-2227.

D'Amours, D., Stegmeier, F., and Amon, A. (2004). Cdc14 and condensin control the dissolution of cohesin-independent chromosome linkages at repeated DNA. Cell *117*, 455-469.

Davidson, I.F., Goetz, D., Zaczek, M.P., Molodtsov, M.I., Huis In 't Veld, P.J., Weissmann, F., Litos, G., Cisneros, D.A., Ocampo-Hafalla, M., Ladurner, R.*, et al.* (2016). Rapid movement and transcriptional re-localization of human cohesin on DNA. The EMBO journal *35*, 2671-2685.

Depew, R.E., Liu, L.F., and Wang, J.C. (1978). Interaction between DNA and Escherichia coli protein omega. Formation of a complex between single-stranded DNA and omega protein. The Journal of biological chemistry *253*, 511-518.

Dewar, J.M., Budzowska, M., and Walter, J.C. (2015). The mechanism of DNA replication termination in vertebrates. Nature *525*, 345-350.

Diaz, M., and Pecinka, A. (2018). Scaffolding for Repair: Understanding Molecular Functions of the SMC5/6 Complex. Genes (Basel) *9*.

Diffley, J.F. (2004). Regulation of early events in chromosome replication. Current biology : CB *14*, R778-786.

Dimitrova, D.S. (2011). DNA replication initiation patterns and spatial dynamics of the human ribosomal RNA gene loci. Journal of cell science *124*, 2743-2752.

DiNardo, S., Voelkel, K., and Sternglanz, R. (1984). DNA topoisomerase II mutant of Saccharomyces cerevisiae: topoisomerase II is required for segregation of daughter molecules at the termination of DNA replication. Proceedings of the National Academy of Sciences of the United States of America *81*, 2616-2620.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376-380.

Downs, J.A., Lowndes, N.F., and Jackson, S.P. (2000). A role for Saccharomyces cerevisiae histone H2A in DNA repair. Nature *408*, 1001-1004.

Drolet, M., Bi, X., and Liu, L.F. (1994). Hypernegative supercoiling of the DNA template during transcription elongation in vitro. The Journal of biological chemistry *269*, 2068-2074.
Dunaway, M., and Ostrander, E.A. (1993). Local domains of supercoiling activate a eukaryotic promoter in vivo. Nature *361*, 746-748.

Dutta, D., Shatalin, K., Epshtein, V., Gottesman, M.E., and Nudler, E. (2011). Linking RNA polymerase backtracking to genome instability in E. coli. Cell *146*, 533-543.

El Hage, A., French, S.L., Beyer, A.L., and Tollervey, D. (2010). Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. Genes & development *24*, 1546-1558.

Evrin, C., Clarke, P., Zech, J., Lurz, R., Sun, J., Uhle, S., Li, H., Stillman, B., and Speck, C. (2009). A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. Proceedings of the National Academy of Sciences of the United States of America *106*, 20240-20245.

Fanning, E., Klimovich, V., and Nager, A.R. (2006). A dynamic model for replication protein A (RPA) function in DNA processing pathways. Nucleic Acids Res *34*, 4126-4137.

Farcas, A.M., Uluocak, P., Helmhart, W., and Nasmyth, K. (2011). Cohesin's concatenation of sister DNAs maintains their intertwining. Molecular cell *44*, 97-107.

Fernandez, X., Diaz-Ingelmo, O., Martinez-Garcia, B., and Roca, J. (2014). Chromatin regulates DNA torsional energy via topoisomerase II-mediated relaxation of positive supercoils. The EMBO journal *33*, 1492-1501.

Fernius, J., Nerusheva, O.O., Galander, S., Alves Fde, L., Rappsilber, J., and Marston, A.L. (2013). Cohesin-dependent association of scc2/4 with the centromere initiates pericentromeric cohesion establishment. Current biology : CB *23*, 599-606.

Frattini, C., Villa-Hernandez, S., Pellicano, G., Jossen, R., Katou, Y., Shirahige, K., and Bermejo, R. (2017). Cohesin Ubiquitylation and Mobilization Facilitate Stalled Replication Fork Dynamics. Molecular cell *68*, 758-772 e754.

French, S. (1992). Consequences of replication fork movement through transcription units in vivo. Science *258*, 1362-1365.

French, S.L., Sikes, M.L., Hontz, R.D., Osheim, Y.N., Lambert, T.E., El Hage, A., Smith, M.M., Tollervey, D., Smith, J.S., and Beyer, A.L. (2011). Distinguishing the roles of Topoisomerases I and II in relief of transcription-induced torsional stress in yeast rRNA genes. Molecular and cellular biology *31*, 482-494.

Fu, Y.V., Yardimci, H., Long, D.T., Ho, T.V., Guainazzi, A., Bermudez, V.P., Hurwitz, J., van Oijen, A., Scharer, O.D., and Walter, J.C. (2011). Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase. Cell *146*, 931-941.

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., and Mirny, L.A. (2016). Formation of Chromosomal Domains by Loop Extrusion. Cell reports *15*, 2038-2049.
Gadaleta, M.C., and Noguchi, E. (2017). Regulation of DNA Replication through Natural Impediments in the Eukaryotic Genome. Genes (Basel) *8*.

Gambus, A., Jones, R.C., Sanchez-Diaz, A., Kanemaki, M., van Deursen, F., Edmondson, R.D., and Labib, K. (2006). GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. Nature cell biology *8*, 358-366.

Gambus, A., van Deursen, F., Polychronopoulos, D., Foltman, M., Jones, R.C., Edmondson, R.D., Calzada, A., and Labib, K. (2009). A key role for Ctf4 in coupling the MCM2-7 helicase to DNA polymerase alpha within the eukaryotic replisome. The EMBO journal *28*, 2992-3004.

Gan, W., Guan, Z., Liu, J., Gui, T., Shen, K., Manley, J.L., and Li, X. (2011). R-loop-mediated genomic instability is caused by impairment of replication fork progression. Genes & development *25*, 2041-2056.

Gangloff, S., de Massy, B., Arthur, L., Rothstein, R., and Fabre, F. (1999). The essential role of yeast topoisomerase III in meiosis depends on recombination. The EMBO journal *18*, 1701-1711.

Gangloff, S., McDonald, J.P., Bendixen, C., Arthur, L., and Rothstein, R. (1994). The yeast type I topoisomerase Top3 interacts with Sgs1, a DNA helicase homolog: a potential eukaryotic reverse gyrase. Molecular and cellular biology *14*, 8391-8398.

Ganji, M., Shaltiel, I.A., Bisht, S., Kim, E., Kalichava, A., Haering, C.H., and Dekker, C. (2018). Real-time imaging of DNA loop extrusion by condensin. Science *360*, 102-105.

Gartenberg, M.R., and Smith, J.S. (2016). The Nuts and Bolts of Transcriptionally Silent Chromatin in Saccharomyces cerevisiae. Genetics *203*, 1563-1599.

Gartenberg, M.R., and Wang, J.C. (1992). Positive supercoiling of DNA greatly diminishes mRNA synthesis in yeast. Proceedings of the National Academy of Sciences of the United States of America *89*, 11461-11465.

Gligoris, T.G., Scheinost, J.C., Burmann, F., Petela, N., Chan, K.L., Uluocak, P., Beckouet, F., Gruber, S., Nasmyth, K., and Lowe, J. (2014). Closing the cohesin ring: structure and function of its Smc3-kleisin interface. Science *346*, 963-967.

Glynn, E.F., Megee, P.C., Yu, H.G., Mistrot, C., Unal, E., Koshland, D.E., DeRisi, J.L., and Gerton, J.L. (2004). Genome-wide mapping of the cohesin complex in the yeast Saccharomyces cerevisiae. PLoS biology *2*, E259.

Gottipati, P., and Helleday, T. (2009). Transcription-associated recombination in eukaryotes: link between transcription, replication and recombination. Mutagenesis *24*, 203-210.
Greenfeder, S.A., and Newlon, C.S. (1992). Replication forks pause at yeast centromeres. Molecular and cellular biology *12*, 4056-4066.

Gruber, S., Arumugam, P., Katou, Y., Kuglitsch, D., Helmhart, W., Shirahige, K., and Nasmyth, K. (2006). Evidence that loading of cohesin onto chromosomes involves opening of its SMC hinge. Cell *127*, 523-537.

Gruber, S., Haering, C.H., and Nasmyth, K. (2003). Chromosomal cohesin forms a ring. Cell *112*, 765-777.

Guacci, V., Hogan, E., and Koshland, D. (1994). Chromosome condensation and sister chromatid pairing in budding yeast. The Journal of cell biology *125*, 517-530.

Guacci, V., and Koshland, D. (2012). Cohesin-independent segregation of sister chromatids in budding yeast. Molecular biology of the cell *23*, 729-739.

Guacci, V., Koshland, D., and Strunnikov, A. (1997). A direct link between sister chromatid cohesion and chromosome condensation revealed through the analysis of MCD1 in S. cerevisiae. Cell *91*, 47-57.

Guy, L., and Roten, C.A. (2004). Genometric analyses of the organization of circular chromosomes: a universal pressure determines the direction of ribosomal RNA genes transcription relative to chromosome replication. Gene *340*, 45-52.

Haarhuis, J.H., Elbatsh, A.M., van den Broek, B., Camps, D., Erkan, H., Jalink, K., Medema, R.H., and Rowland, B.D. (2013). WAPL-mediated removal of cohesin protects against segregation errors and aneuploidy. Current biology : CB *23*, 2071-2077.

Hadjur, S., Williams, L.M., Ryan, N.K., Cobb, B.S., Sexton, T., Fraser, P., Fisher, A.G., and Merkenschlager, M. (2009). Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. Nature *460*, 410-413.

Haering, C.H., Farcas, A.M., Arumugam, P., Metson, J., and Nasmyth, K. (2008). The cohesin ring concatenates sister DNA molecules. Nature *454*, 297-301.

Haering, C.H., Schoffnegger, D., Nishino, T., Helmhart, W., Nasmyth, K., and Lowe, J. (2004). Structure and stability of cohesin's Smc1-kleisin interaction. Molecular cell *15*, 951-964.
Hamperl, S., and Cimprich, K.A. (2016). Conflict Resolution in the Genome: How Transcription and Replication Make It Work. Cell *167*, 1455-1467.

Hanna, J.S., Kroll, E.S., Lundblad, V., and Spencer, F.A. (2001). Saccharomyces cerevisiae CTF18 and CTF4 are required for sister chromatid cohesion. Molecular and cellular biology *21*, 3144-3158.

Heidinger-Pauli, J.M., Mert, O., Davenport, C., Guacci, V., and Koshland, D. (2010). Systematic reduction of cohesin differentially affects chromosome segregation, condensation, and DNA repair. Current biology : CB *20*, 957-963.

Helmrich, A., Ballarino, M., Nudler, E., and Tora, L. (2013). Transcription-replication encounters, consequences and genomic instability. Nat Struct Mol Biol *20*, 412-418.

Helmrich, A., Ballarino, M., and Tora, L. (2011). Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. Molecular cell *44*, 966-977.

Hills, S.A., and Diffley, J.F. (2014). DNA replication and oncogene-induced replicative stress. Current biology : CB *24*, R435-444.

Hinshaw, S.M., Makrantoni, V., Kerr, A., Marston, A.L., and Harrison, S.C. (2015). Structural evidence for Scc4-dependent localization of cohesin loading. Elife *4*, e06057.

Hirano, T. (2012). Condensins: universal organizers of chromosomes with diverse functions. Genes & development *26*, 1659-1678.

Hirano, T. (2016). Condensin-Based Chromosome Organization from Bacteria to Vertebrates. Cell *164*, 847-857.

Holm, C., Goto, T., Wang, J.C., and Botstein, D. (1985). DNA topoisomerase II is required at the time of mitosis in yeast. Cell *41*, 553-563.

Hu, B., Itoh, T., Mishra, A., Katoh, Y., Chan, K.L., Upcher, W., Godlee, C., Roig, M.B., Shirahige, K., and Nasmyth, K. (2011). ATP hydrolysis is required for relocating cohesin from sites occupied by its Scc2/4 loading complex. Current biology : CB *21*, 12-24.

Hu, B., Petela, N., Kurze, A., Chan, K.L., Chapard, C., and Nasmyth, K. (2015). Biological chromodynamics: a general method for measuring protein occupancy across the genome by calibrating ChIP-seq. Nucleic Acids Res *43*, e132.

Huis in 't Veld, P.J., Herzog, F., Ladurner, R., Davidson, I.F., Piric, S., Kreidl, E., Bhaskara, V., Aebersold, R., and Peters, J.M. (2014). Characterization of a DNA exit gate in the human cohesin ring. Science *346*, 968-972.

Huppert, J.L., and Balasubramanian, S. (2007). G-quadruplexes in promoters throughout the human genome. Nucleic Acids Res *35*, 406-413.

Ivanov, D., Schleiffer, A., Eisenhaber, F., Mechtler, K., Haering, C.H., and Nasmyth, K. (2002). Eco1 is a novel acetyltransferase that can acetylate proteins involved in cohesion. Current biology : CB *12*, 323-328.

Ivessa, A.S., Lenzmeier, B.A., Bessler, J.B., Goudsouzian, L.K., Schnakenberg, S.L., and Zakian, V.A. (2003). The Saccharomyces cerevisiae helicase Rrm3p facilitates replication past nonhistone protein-DNA complexes. Molecular cell *12*, 1525-1536.

Ivessa, A.S., Zhou, J.Q., Schulz, V.P., Monson, E.K., and Zakian, V.A. (2002). Saccharomyces Rrm3p, a 5' to 3' DNA helicase that promotes replication fork progression through telomeric and subtelomeric DNA. Genes & development *16*, 1383-1396.

Ivessa, A.S., Zhou, J.Q., and Zakian, V.A. (2000). The Saccharomyces Pif1p DNA helicase and the highly related Rrm3p have opposite effects on replication fork progression in ribosomal DNA. Cell *100*, 479-489.

Janke, C., Magiera, M.M., Rathfelder, N., Taxis, C., Reber, S., Maekawa, H., Moreno-Borchart, A., Doenges, G., Schwob, E., Schiebel, E*., et al.* (2004). A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. Yeast *21*, 947-962.

Jehn, B., Niedenthal, R., and Hegemann, J.H. (1991). In vivo analysis of the Saccharomyces cerevisiae centromere CDEIII sequence: requirements for mitotic chromosome segregation. Molecular and cellular biology *11*, 5212-5221.

Jeppsson, K., Carlborg, K.K., Nakato, R., Berta, D.G., Lilienthal, I., Kanno, T., Lindqvist, A., Brink, M.C., Dantuma, N.P., Katou, Y*., et al.* (2014). The chromosomal association of the Smc5/6 complex depends on cohesion and predicts the level of sister chromatid entanglement. PLoS Genet *10*, e1004680.

Joshi, R.S., Pina, B., and Roca, J. (2012). Topoisomerase II is required for the production of long Pol II gene transcripts in yeast. Nucleic Acids Res *40*, 7907-7915.

Kamimura, Y., Tak, Y.S., Sugino, A., and Araki, H. (2001). Sld3, which interacts with Cdc45 (Sld4), functions for chromosomal DNA replication in Saccharomyces cerevisiae. The EMBO journal *20*, 2097-2107.

Katou, Y., Kanoh, Y., Bando, M., Noguchi, H., Tanaka, H., Ashikari, T., Sugimoto, K., and Shirahige, K. (2003). S-phase checkpoint proteins Tof1 and Mrc1 form a stable replication-pausing complex. Nature *424*, 1078-1083.

Keszthelyi, A., Minchell, N.E., and Baxter, J. (2016). The Causes and Consequences of Topological Stress during DNA Replication. Genes (Basel) *7*.

Kikuchi, S., Borek, D.M., Otwinowski, Z., Tomchick, D.R., and Yu, H. (2016). Crystal structure of the cohesin loader Scc2 and insight into cohesinopathy. Proceedings of the National Academy of Sciences of the United States of America *113*, 12444-12449.

Kim, J.A., Kruhlak, M., Dotiwala, F., Nussenzweig, A., and Haber, J.E. (2007). Heterochromatin is refractory to gamma-H2AX modification in yeast and mammals. The Journal of cell biology *178*, 209-218.

Kim, R.A., and Wang, J.C. (1989). Function of DNA topoisomerases as replication swivels in Saccharomyces cerevisiae. Journal of Molecular Biology *208*, 257-267.

Kimura, K., and Hirano, T. (1997). ATP-dependent positive supercoiling of DNA by 13S condensin: a biochemical implication for chromosome condensation. Cell *90*, 625-634.

Kitajima, T.S., Kawashima, S.A., and Watanabe, Y. (2004). The conserved kinetochore protein shugoshin protects centromeric cohesion during meiosis. Nature *427*, 510-517.

Kobayashi, T. (2014). Ribosomal RNA gene repeats, their stability and cellular senescence. Proc Jpn Acad Ser B Phys Biol Sci *90*, 119-129.

Kobayashi, T., Heck, D.J., Nomura, M., and Horiuchi, T. (1998). Expansion and contraction of ribosomal DNA repeats in Saccharomyces cerevisiae: requirement of replication fork blocking (Fob1) protein and the role of RNA polymerase I. Genes & development *12*, 3821-3830.

Kobayashi, T., and Horiuchi, T. (1996). A yeast gene product, Fob1 protein, required for both replication fork blocking and recombinational hotspot activities. Genes Cells *1*, 465-474.

Kogut, I., Wang, J., Guacci, V., Mistry, R.K., and Megee, P.C. (2009). The Scc2/Scc4 cohesin loader determines the distribution of cohesin on budding yeast chromosomes. Genes & development *23*, 2345-2357.

Kornberg, A., and Baker, T.A. (1992). DNA replication, 2nd edn (W. H. Freeman and Co., New York).

Koster, D.A., Croquette, V., Dekker, C., Shuman, S., and Dekker, N.H. (2005). Friction and torque govern the relaxation of DNA supercoils by eukaryotic topoisomerase IB. Nature *434*, 671-674.

Kouzine, F., Liu, J., Sanford, S., Chung, H.J., and Levens, D. (2004). The dynamic response of upstream DNA to transcription-generated torsional stress. Nat Struct Mol Biol *11*, 1092-1100.
Kouzine, F., Sanford, S., Elisha-Feil, Z., and Levens, D. (2008). The functional response of upstream DNA to dynamic supercoiling in vivo. Nat Struct Mol Biol *15*, 146-154.

Krawczyk, C., Dion, V., Schar, P., and Fritsch, O. (2014). Reversible Top1 cleavage complexes are stabilized strand-specifically at the ribosomal replication fork barrier and contribute to ribosomal DNA stability. Nucleic Acids Res *42*, 4985-4995.

Kueng, S., Hegemann, B., Peters, B.H., Lipp, J.J., Schleiffer, A., Mechtler, K., and Peters, J.M. (2006). Wapl controls the dynamic association of cohesin with chromatin. Cell *127*, 955-967.

Kulemzina, I., Schumacher, M.R., Verma, V., Reiter, J., Metzler, J., Failla, A.V., Lanz, C., Sreedharan, V.T., Ratsch, G., and Ivanov, D. (2012). Cohesin rings devoid of Scc3 and Pds5 maintain their stable association with the DNA. PLoS Genet *8*, e1002856.

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., Bertero, M.G., Bessieres, P., Bolotin, A., Borchert, S.*, et al.* (1997). The complete genome sequence of the gram-positive bacterium Bacillus subtilis. Nature *390*, 249-256.

Labib, K. (2010). How do Cdc7 and cyclin-dependent kinases trigger the initiation of chromosome replication in eukaryotic cells? Genes & development *24*, 1208-1219.
Labib, K., and Hodgson, B. (2007). Replication fork barriers: pausing for a break or stalling for time? EMBO Rep *8*, 346-353.

Laloraya, S., Guacci, V., and Koshland, D. (2000). Chromosomal addresses of the cohesin component Mcd1p. The Journal of cell biology *151*, 1047-1056.

Lam, E.Y., Beraldi, D., Tannahill, D., and Balasubramanian, S. (2013). G-quadruplex structures are stable and detectable in human genomic DNA. Nat Commun *4*, 1796.

Lavoie, B.D., Hogan, E., and Koshland, D. (2004). In vivo requirements for rDNA chromosome condensation reveal two cell-cycle-regulated pathways for mitotic chromosome folding. Genes & development *18*, 76-87.

Lee, B.G., Roig, M.B., Jansma, M., Petela, N., Metson, J., Nasmyth, K., and Lowe, J. (2016). Crystal Structure of the Cohesin Gatekeeper Pds5 and in Complex with Kleisin Scc1. Cell reports *14*, 2108-2115.

Leman, A.R., and Noguchi, E. (2012). Local and global functions of Timeless and Tipin in replication fork protection. Cell cycle *11*, 3945-3955.

Lengronne, A., Katou, Y., Mori, S., Yokobayashi, S., Kelly, G.P., Itoh, T., Watanabe, Y., Shirahige, K., and Uhlmann, F. (2004). Cohesin relocation from sites of chromosomal loading to places of convergent transcription. Nature *430*, 573-578.

Lengronne, A., McIntyre, J., Katou, Y., Kanoh, Y., Hopfner, K.P., Shirahige, K., and Uhlmann, F. (2006). Establishment of sister chromatid cohesion at the S. cerevisiae replication fork. Molecular cell *23*, 787-799.

Leonard, J., Sen, N., Torres, R., Sutani, T., Jarmuz, A., Shirahige, K., and Aragon, L. (2015). Condensin Relocalization from Centromeres to Chromosome Arms Promotes Top2 Recruitment during Anaphase. Cell reports *13*, 2336-2344.

Liang, C., Weinreich, M., and Stillman, B. (1995). ORC and Cdc6p interact and determine the frequency of initiation of DNA replication in the genome. Cell *81*, 667-676.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O.*, et al.* (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science *326*, 289-293.

Linskens, M.H., and Huberman, J.A. (1988). Organization of replication of ribosomal DNA in Saccharomyces cerevisiae. Molecular and cellular biology *8*, 4927-4935.

Litwin, I., Bakowski, T., Maciaszczyk-Dziubinska, E., and Wysocki, R. (2017). The LSH/HELLS homolog Irc5 contributes to cohesin association with chromatin in yeast. Nucleic Acids Res *45*, 6404-6416.

Liu, L.F., Liu, C.-C., and Alberts, B.M. (1979). T4 DNA topoisomerase: a new ATP-dependent enzyme essential for initiation of T4 bacteriophage DNA replication. Nature *281*, 456.

Liu, L.F., and Wang, J.C. (1979). Interaction between DNA and Escherichia coli DNA topoisomerase I. Formation of complexes between the protein and superhelical and nonsuperhelical duplex DNAs. The Journal of biological chemistry *254*, 11082-11088.

Liu, L.F., and Wang, J.C. (1987). Supercoiling of the DNA template during transcription. Proceedings of the National Academy of Sciences of the United States of America *84*, 7024-7027.

Lopes, M., Cotta-Ramusino, C., Pellicioli, A., Liberi, G., Plevani, P., Muzi-Falconi, M., Newlon, C.S., and Foiani, M. (2001). The DNA replication checkpoint response stabilizes stalled replication forks. Nature *412*, 557-561.

Lopez-Serra, L., Kelly, G., Patel, H., Stewart, A., and Uhlmann, F. (2014). The Scc2-Scc4 complex acts in sister chromatid cohesion and transcriptional regulation by maintaining nucleosome-free regions. Nature genetics *46*, 1147-1151.

Losada, A., Hirano, M., and Hirano, T. (1998). Identification of Xenopus SMC protein complexes required for sister chromatid cohesion. Genes & development *12*, 1986-1997.

Lowe, J., Cordell, S.C., and van den Ent, F. (2001). Crystal structure of the SMC head domain: an ABC ATPase with 900 residues antiparallel coiled-coil inserted. J Mol Biol *306*, 25-35.

Lucas, I., Germe, T., Chevrier-Miller, M., and Hyrien, O. (2001). Topoisomerase II can unlink replicating DNA by precatenane removal. The EMBO journal *20*, 6509-6519.

Ma, L., Ho, K., Piggott, N., Luo, Z., and Measday, V. (2012). Interactions between the kinetochore complex and the protein kinase A pathway in Saccharomyces cerevisiae. G3 (Bethesda) *2*, 831-841.

Makovets, S., Herskowitz, I., and Blackburn, E.H. (2004). Anatomy and dynamics of DNA replication fork movement in yeast telomeric regions. Molecular and cellular biology *24*, 4019-4031.

Martinez-Robles, M.L., Witz, G., Hernandez, P., Schvartzman, J.B., Stasiak, A., and Krimer, D.B. (2009). Interplay of DNA supercoiling and catenation during the segregation of sister duplexes. Nucleic acids research *37*, 5126-5137.

Masse, E., and Drolet, M. (1999). Relaxation of transcription-induced negative supercoiling is an essential function of Escherichia coli DNA topoisomerase I. The Journal of biological chemistry *274*, 16654-16658.

Mayer, M.L., Gygi, S.P., Aebersold, R., and Hieter, P. (2001). Identification of RFC(Ctf18p, Ctf8p, Dcc1p): an alternative RFC complex required for sister chromatid cohesion in S. cerevisiae. Molecular cell *7*, 959-970.

McCarroll, R.M., and Fangman, W.L. (1988). Time of replication of yeast centromeres and telomeres. Cell *54*, 505-513.

McGill, R., Tukey, J.W., and Larsen, W.A. (1978). Variations of Box Plots. The American Statistician *32*, 12-16.

McGuffee, S.R., Smith, D.J., and Whitehouse, I. (2013). Quantitative, genome-wide analysis of eukaryotic replication initiation and termination. Molecular cell *50*, 123-135.

Merkenschlager, M., and Nora, E.P. (2016). CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation. Annual review of genomics and human genetics *17*, 17-43.

Merrikh, H., Machon, C., Grainger, W.H., Grossman, A.D., and Soultanas, P. (2011). Co-directional replication-transcription conflicts lead to replication restart. Nature *470*, 554-557.

Meryet-Figuiere, M., Alaei-Mahabadi, B., Ali, M.M., Mitra, S., Subhash, S., Pandey, G.K., Larsson, E., and Kanduri, C. (2014). Temporal separation of replication and transcription during S-phase progression. Cell cycle *13*, 3241-3248.

Michaelis, C., Ciosk, R., and Nasmyth, K. (1997). Cohesins: chromosomal proteins that prevent premature separation of sister chromatids. Cell *91*, 35-45.

Mirkin, E.V., and Mirkin, S.M. (2005). Mechanisms of transcription-replication collisions in bacteria. Molecular and cellular biology *25*, 888-895.

Mischo, H.E., Gomez-Gonzalez, B., Grzechnik, P., Rondon, A.G., Wei, W., Steinmetz, L., Aguilera, A., and Proudfoot, N.J. (2011). Yeast Sen1 helicase protects the genome from transcription-associated instability. Molecular cell *41*, 21-32.

Mizuguchi, T., Fudenberg, G., Mehta, S., Belton, J.M., Taneja, N., Folco, H.D., FitzGerald, P., Dekker, J., Mirny, L., Barrowman, J.*, et al.* (2014). Cohesin-dependent globules and heterochromatin shape 3D genome architecture in S. pombe. Nature *516*, 432-435.

Mohanty, B.K., Bairwa, N.K., and Bastia, D. (2006). The Tof1p-Csm3p protein complex counteracts the Rrm3p helicase to control replication termination of Saccharomyces

cerevisiae. Proceedings of the National Academy of Sciences of the United States of America *103*, 897-902.

Moldovan, G.L., Pfander, B., and Jentsch, S. (2006). PCNA controls establishment of sister chromatid cohesion during S phase. Molecular cell *23*, 723-732.

Moldovan, G.L., Pfander, B., and Jentsch, S. (2007). PCNA, the maestro of the replication fork. Cell *129*, 665-679.

Mumberg, D., Muller, R., and Funk, M. (1994). Regulatable promoters of Saccharomyces cerevisiae: comparison of transcriptional activity and their use for heterologous expression. Nucleic Acids Res *22*, 5767-5768.

Muramatsu, S., Hirai, K., Tak, Y.S., Kamimura, Y., and Araki, H. (2010). CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol (epsilon}, and GINS in budding yeast. Genes & development *24*, 602-612.

Murayama, Y., Samora, C.P., Kurokawa, Y., Iwasaki, H., and Uhlmann, F. (2018). Establishment of DNA-DNA Interactions by the Cohesin Ring. Cell *172*, 465-477 e415.

Murayama, Y., and Uhlmann, F. (2014). Biochemical reconstitution of topological DNA binding by the cohesin ring. Nature *505*, 367-371.

Murayama, Y., and Uhlmann, F. (2015). DNA Entry into and Exit out of the Cohesin Ring by an Interlocking Gate Mechanism. Cell *163*, 1628-1640.

Natsume, T., Muller, C.A., Katou, Y., Retkute, R., Gierlinski, M., Araki, H., Blow, J.J., Shirahige, K., Nieduszynski, C.A., and Tanaka, T.U. (2013). Kinetochores coordinate pericentromeric cohesion and early DNA replication by Cdc7-Dbf4 kinase recruitment. Molecular cell *50*, 661-674.

Neuwald, A.F., and Hirano, T. (2000). HEAT repeats associated with condensins, cohesins, and other complexes involved in chromosome-related functions. Genome research *10*, 1445-1452.

Nick McElhinny, S.A., Gordenin, D.A., Stith, C.M., Burgers, P.M., and Kunkel, T.A. (2008). Division of labor at the eukaryotic replication fork. Molecular cell *30*, 137-144.

Nishiyama, T., Ladurner, R., Schmitz, J., Kreidl, E., Schleiffer, A., Bhaskara, V., Bando, M., Shirahige, K., Hyman, A.A., Mechtler, K.*, et al.* (2010). Sororin mediates sister chromatid cohesion by antagonizing Wapl. Cell *143*, 737-749.

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

Nordheim, A., Lafer, E.M., Peck, L.J., Wang, J.C., Stollar, B.D., and Rich, A. (1982). Negatively supercoiled plasmids contain left-handed Z-DNA segments as detected by specific antibody binding. Cell *31*, 309-318.

Pannunzio, N.R., and Lieber, M.R. (2016). Dissecting the Roles of Divergent and Convergent Transcription in Chromosome Instability. Cell reports *14*, 1025-1031.

Pardo, B., Crabbe, L., and Pasero, P. (2017). Signaling pathways of replication stress in yeast. FEMS yeast research *17*.

Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H.C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T.*, et al.* (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. Cell *132*, 422-433.

Park, H., and Sternglanz, R. (1999). Identification and characterization of the genes for two topoisomerase I-interacting proteins from Saccharomyces cerevisiae. Yeast *15*, 35-41.

Peck, L.J., Nordheim, A., Rich, A., and Wang, J.C. (1982). Flipping of cloned d(pCpG)n.d(pCpG)n DNA sequences from right- to left-handed helical structure by salt, Co(III), or negative supercoiling. Proceedings of the National Academy of Sciences of the United States of America *79*, 4560-4564.

Pedersen, J.M., Fredsoe, J., Roedgaard, M., Andreasen, L., Mundbjerg, K., Kruhoffer, M., Brinch, M., Schierup, M.H., Bjergbaek, L., and Andersen, A.H. (2012). DNA Topoisomerases maintain promoters in a state competent for transcriptional activation in Saccharomyces cerevisiae. PLoS Genet *8*, e1003128.

Petela, N.J., Gligoris, T.G., Metson, J., Lee, B.G., Voulgaris, M., Hu, B., Kikuchi, S., Chapard, C., Chen, W., Rajendra, E.*, et al.* (2018). Scc2 Is a Potent Activator of Cohesin's ATPase that Promotes Loading by Binding Scc1 without Pds5. Molecular cell *70*, 1134-1148 e1137.

Peter, B.J., Ullsperger, C., Hiasa, H., Marians, K.J., and Cozzarelli, N.R. (1998). The structure of supercoiled intermediates in DNA replication. Cell *94*, 819-827.

Petes, T.D. (1979). Yeast ribosomal DNA genes are located on chromosome XII. Proceedings of the National Academy of Sciences of the United States of America *76*, 410-414.

Phatnani, H.P., Jones, J.C., and Greenleaf, A.L. (2004). Expanding the functional repertoire of CTD kinase I and RNA polymerase II: novel phosphoCTD-associating proteins in the yeast proteome. Biochemistry *43*, 15702-15719.

Pope, B.D., Ryba, T., Dileep, V., Yue, F., Wu, W., Denas, O., Vera, D.L., Wang, Y., Hansen, R.S., Canfield, T.K.*, et al.* (2014). Topologically associating domains are stable units of replication-timing regulation. Nature *515*, 402-405.

Postow, L., Crisona, N.J., Peter, B.J., Hardy, C.D., and Cozzarelli, N.R. (2001). Topological challenges to DNA replication: conformations at the fork. Proceedings of the National Academy of Sciences of the United States of America *98*, 8219-8226.

Postow, L., Hardy, C.D., Arsuaga, J., and Cozzarelli, N.R. (2004). Topological domain structure of the Escherichia coli chromosome. Genes & development *18*, 1766-1779.

Prado, F., and Aguilera, A. (2005). Impairment of replication fork progression mediates RNA polII transcription-associated recombination. The EMBO journal *24*, 1267-1276.

Pursell, Z.F., Isoz, I., Lundstrom, E.B., Johansson, E., and Kunkel, T.A. (2007). Yeast DNA polymerase epsilon participates in leading-strand DNA replication. Science *317*, 127-130.

Racko, D., Benedetti, F., Dorier, J., and Stasiak, A. (2018). Transcription-induced supercoiling as the driving force of chromatin loop extrusion during formation of TADs in interphase chromosomes. Nucleic Acids Res *46*, 1648-1660.

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S.*, et al.* (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell *159*, 1665-1680.

Remus, D., Beuron, F., Tolun, G., Griffith, J.D., Morris, E.P., and Diffley, J.F. (2009). Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. Cell *139*, 719-730.

Rhodes, J., Mazza, D., Nasmyth, K., and Uphoff, S. (2017). Scc2/Nipbl hops between chromosomal cohesin rings after loading. Elife *6*.

Richardson, J.P. (1975). Attachment of nascent RNA molecules to superhelical DNA. J Mol Biol *98*, 565-579.

Rolef Ben-Shahar, T., Heeger, S., Lehane, C., East, P., Flynn, H., Skehel, M., and Uhlmann, F. (2008). Eco1-dependent cohesin acetylation during establishment of sister chromatid cohesion. Science *321*, 563-566.

Rowland, B.D., Roig, M.B., Nishino, T., Kurze, A., Uluocak, P., Mishra, A., Beckouet, F., Underwood, P., Metson, J., Imre, R.*, et al.* (2009). Building sister chromatid cohesion: smc3 acetylation counteracts an antiestablishment activity. Molecular cell *33*, 763-774.

Roy, M.A., and D'Amours, D. (2011). DNA-binding properties of Smc6, a core component of the Smc5-6 DNA repair complex. Biochemical and biophysical research communications *416*, 80-85.

Roy, M.A., Siddiqui, N., and D'Amours, D. (2011). Dynamic and selective DNA-binding activity of Smc5, a core component of the Smc5-Smc6 complex. Cell cycle *10*, 690-700.

Rubio, E.D., Reiss, D.J., Welcsh, P.L., Disteche, C.M., Filippova, G.N., Baliga, N.S., Aebersold, R., Ranish, J.A., and Krumm, A. (2008). CTCF physically links cohesin to chromatin. Proceedings of the National Academy of Sciences of the United States of America *105*, 8309-8314.

Salceda, J., Fernandez, X., and Roca, J. (2006). Topoisomerase II, not topoisomerase I, is the proficient relaxase of nucleosomal DNA. The EMBO journal *25*, 2575-2583.

Schalbetter, S.A., Goloborodko, A., Fudenberg, G., Belton, J.M., Miles, C., Yu, M., Dekker, J., Mirny, L., and Baxter, J. (2017). SMC complexes differentially compact mitotic chromosomes according to genomic context. Nature cell biology *19*, 1071-1080.

Schalbetter, S.A., Mansoubi, S., Chambers, A.L., Downs, J.A., and Baxter, J. (2015). Fork rotation and DNA precatenation are restricted during DNA replication to prevent chromosomal instability. Proceedings of the National Academy of Sciences of the United States of America *112*, E4565-4570.

Schleiffer, A., Kaitna, S., Maurer-Stroh, S., Glotzer, M., Nasmyth, K., and Eisenhaber, F. (2003). Kleisins: a superfamily of bacterial and eukaryotic SMC protein partners. Molecular cell *11*, 571-575.

Schroth, G.P., Chou, P.J., and Ho, P.S. (1992). Mapping Z-DNA in the human genome. Computer-aided mapping reveals a nonrandom distribution of potential Z-DNA-forming sequences in human genes. The Journal of biological chemistry *267*, 11846-11855.

Schultz, M.C., Brill, S.J., Ju, Q., Sternglanz, R., and Reeder, R.H. (1992). Topoisomerases and yeast rRNA transcription: negative supercoiling stimulates initiation and topoisomerase activity is required for elongation. Genes & development *6*, 1332-1341.

Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N.A., Huber, W., C, H.H., Mirny, L.*, et al.* (2017). Two independent modes of chromatin organization revealed by cohesin removal. Nature *551*, 51-56.

Seitan, V.C., Faure, A.J., Zhan, Y., McCord, R.P., Lajoie, B.R., Ing-Simmons, E., Lenhard, B., Giorgetti, L., Heard, E., Fisher, A.G.*, et al.* (2013). Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. Genome research *23*, 2066-2077.

Shirayama, M., Toth, A., Galova, M., and Nasmyth, K. (1999). APC(Cdc20) promotes exit from mitosis by destroying the anaphase inhibitor Pds1 and cyclin Clb5. Nature *402*, 203-207.

Shroff, R., Arbel-Eden, A., Pilch, D., Ira, G., Bonner, W.M., Petrini, J.H., Haber, J.E., and Lichten, M. (2004). Distribution and dynamics of chromatin modification induced by a defined DNA double-strand break. Current Biology *14*, 1703-1711.

Simon, A.C., Zhou, J.C., Perera, R.L., van Deursen, F., Evrin, C., Ivanova, M.E., Kilkenny, M.L., Renault, L., Kjaer, S., Matak-Vinkovic, D.*, et al.* (2014). A Ctf4 trimer couples the CMG helicase to DNA polymerase alpha in the eukaryotic replisome. Nature *510*, 293-297.

Skibbens, R.V. (2004). Chl1p, a DNA helicase-like protein in budding yeast, functions in sister-chromatid cohesion. Genetics *166*, 33-42.

Skibbens, R.V. (2016). Of Rings and Rods: Regulating Cohesin Entrapment of DNA to Generate Intra- and Intermolecular Tethers. PLoS Genet *12*, e1006337.

Skibbens, R.V., Corson, L.B., Koshland, D., and Hieter, P. (1999). Ctf7p is essential for sister chromatid cohesion and links mitotic chromosome structure to the DNA replication machinery. Genes & development *13*, 307-319.

Smirnov, E., Borkovec, J., Kovacik, L., Svidenska, S., Schrofel, A., Skalnikova, M., Svindrych, Z., Krizek, P., Ovesny, M., Hagen, G.M.*, et al.* (2014). Separation of replication and transcription domains in nucleoli. J Struct Biol *188*, 259-266.

Snapka, R.M., Powelson, M.A., and Strayer, J.M. (1988). Swiveling and decatenation of replicating simian virus 40 genomes in vivo. Molecular and cellular biology *8*, 515-521.

Song, J., Lafont, A., Chen, J., Wu, F.M., Shirahige, K., and Rankin, S. (2012). Cohesin acetylation promotes sister chromatid cohesion only in association with the replication machinery. The Journal of biological chemistry *287*, 34325-34336.

Soppa, J. (2001). Prokaryotic structural maintenance of chromosomes (SMC) proteins: distribution, phylogeny, and comparison with MukBs and additional prokaryotic and eukaryotic coiled-coil proteins. Gene *278*, 253-264.

Sperling, A.S., Jeong, K.S., Kitada, T., and Grunstein, M. (2011). Topoisomerase II binds nucleosome-free DNA and acts redundantly with topoisomerase I to enhance recruitment of RNA Pol II in budding yeast. Proceedings of the National Academy of Sciences of the United States of America *108*, 12693-12698.

Spitzer, M., Wildenhain, J., Rappsilber, J., and Tyers, M. (2014). BoxPlotR: a web tool for generation of box plots. Nature methods *11*, 121-122.

Srinivasan, M., Scheinost, J.C., Petela, N.J., Gligoris, T.G., Wissler, M., Ogushi, S., Collier, J.E., Voulgaris, M., Kurze, A., Chan, K.L*., et al.* (2018). The Cohesin Ring Uses Its Hinge to Organize DNA Using Non-topological as well as Topological Mechanisms. Cell *173*, 1508-1519 e1518.

Stedman, W., Kang, H., Lin, S., Kissil, J.L., Bartolomei, M.S., and Lieberman, P.M. (2008). Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. The EMBO journal *27*, 654-666.

Stigler, J., Camdere, G.O., Koshland, D.E., and Greene, E.C. (2016). Single-Molecule Imaging Reveals a Collapsed Conformational State for DNA-Bound Cohesin. Cell reports *15*, 988-998.

Strom, L., Lindroos, H.B., Shirahige, K., and Sjogren, C. (2004). Postreplicative recruitment of cohesin to double-strand breaks is required for DNA repair. Molecular cell *16*, 1003-1015.

Sullivan, M., Higuchi, T., Katis, V.L., and Uhlmann, F. (2004). Cdc14 phosphatase induces rDNA condensation and resolves cohesin-independent cohesion during budding yeast anaphase. Cell *117*, 471-482.

Sun, M., Nishino, T., and Marko, J.F. (2013). The SMC1-SMC3 cohesin heterodimer structures DNA through supercoiling-dependent loop formation. Nucleic Acids Res *41*, 6149-6160.

Sun, X., Huang, L., Markowitz, T.E., Blitzblau, H.G., Chen, D., Klein, F., and Hochwagen, A. (2015). Transcription dynamically patterns the meiotic chromosome-axis interface. Elife *4*.

Sundin, O., and Varshavsky, A. (1980). Terminal stages of SV40 DNA replication proceed via multiply intertwined catenated dimers. Cell *21*, 103-114.

Sundin, O., and Varshavsky, A. (1981). Arrest of segregation leads to accumulation of highly intertwined catenated dimers: dissection of the final stages of SV40 DNA replication. Cell *25*, 659-669.

Sutani, T., Kawaguchi, T., Kanno, R., Itoh, T., and Shirahige, K. (2009). Budding yeast Wpl1(Rad61)-Pds5 complex counteracts sister chromatid cohesion-establishing reaction. Current biology : CB *19*, 492-497.

Szilard, R.K., Jacques, P.E., Laramee, L., Cheng, B., Galicia, S., Bataille, A.R., Yeung, M., Mendez, M., Bergeron, M., Robert, F*., et al.* (2010). Systematic identification of fragile sites via genome-wide location analysis of gamma-H2AX. Nat Struct Mol Biol *17*, 299-305.

Tabuchi, H., Handa, H., and Hirose, S. (1993). Underwinding of DNA on binding of yeast TFIID to the TATA element. Biochemical and biophysical research communications *192*, 1432-1438.

Tanaka, S., Umemori, T., Hirai, K., Muramatsu, S., Kamimura, Y., and Araki, H. (2007). CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. Nature *445*, 328-332.

Tanaka, T., Cosma, M.P., Wirth, K., and Nasmyth, K. (1999). Identification of cohesin association sites at centromeres and along chromosome arms. Cell *98*, 847-858.

Tang, S., Wu, M.K., Zhang, R., and Hunter, N. (2015). Pervasive and essential roles of the Top3-Rmi1 decatenase orchestrate recombination and facilitate chromosome segregation in meiosis. Molecular cell *57*, 607-621.

Tedeschi, A., Wutz, G., Huet, S., Jaritz, M., Wuensche, A., Schirghuber, E., Davidson, I.F., Tang, W., Cisneros, D.A., Bhaskara, V.*, et al.* (2013). Wapl is an essential regulator of chromatin structure and chromosome segregation. Nature *501*, 564-568.

Terret, M.E., Sherwood, R., Rahman, S., Qin, J., and Jallepalli, P.V. (2009). Cohesin acetylation speeds the replication fork. Nature *462*, 231-234.

Toth, A., Ciosk, R., Uhlmann, F., Galova, M., Schleiffer, A., and Nasmyth, K. (1999). Yeast cohesin complex requires a conserved protein, Eco1p(Ctf7), to establish cohesion between sister chromatids during DNA replication. Genes & development *13*, 320-333.

Tse, Y., and Wang, J.C. (1980). E. coli and M. luteus DNA topoisomerase I can catalyze catenation of decatenation of double-stranded DNA rings. Cell *22*, 269-276.

Tuduri, S., Crabbe, L., Conti, C., Tourriere, H., Holtgreve-Grez, H., Jauch, A., Pantesco, V., De Vos, J., Thomas, A., Theillet, C.*, et al.* (2009). Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. Nature cell biology *11*, 1315-1324.

Uemura, T., Ohkura, H., Adachi, Y., Morino, K., Shiozaki, K., and Yanagida, M. (1987). DNA topoisomerase II is required for condensation and separation of mitotic chromosomes in S. pombe. Cell *50*, 917-925.

Uhlmann, F. (2016). SMC complexes: from DNA to chromosomes. Nature reviews Molecular cell biology.

Uhlmann, F., Lottspeich, F., and Nasmyth, K. (1999). Sister-chromatid separation at anaphase onset is promoted by cleavage of the cohesin subunit Scc1. Nature *400*, 37-42.

Uhlmann, F., Wernic, D., Poupart, M.A., Koonin, E.V., and Nasmyth, K. (2000). Cleavage of cohesin by the CD clan protease separin triggers anaphase in yeast. Cell *103*, 375-386.

Unal, E., Arbel-Eden, A., Sattler, U., Shroff, R., Lichten, M., Haber, J.E., and Koshland, D. (2004). DNA damage response pathway uses histone modification to assemble a double-strand break-specific cohesin domain. Molecular cell *16*, 991-1002.

Unal, E., Heidinger-Pauli, J.M., Kim, W., Guacci, V., Onn, I., Gygi, S.P., and Koshland, D.E. (2008). A molecular determinant for the establishment of sister chromatid cohesion. Science *321*, 566-569.

Vian, L., Pekowska, A., Rao, S.S.P., Kieffer-Kwon, K.R., Jung, S., Baranello, L., Huang, S.C., El Khattabi, L., Dose, M., Pruett, N.*, et al.* (2018). The Energetics and Physiological Impact of Cohesin Extrusion. Cell *173*, 1165-1178 e1120.

Vos, S.M., Tretter, E.M., Schmidt, B.H., and Berger, J.M. (2011). All tangled up: how cells direct, manage and exploit topoisomerase function. Nature reviews Molecular cell biology *12*, 827-841.

Wahba, L., Costantino, L., Tan, F.J., Zimmer, A., and Koshland, D. (2016). S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. Genes & development *30*, 1327-1338.

Waizenegger, I.C., Hauf, S., Meinke, A., and Peters, J.M. (2000). Two distinct pathways remove mammalian cohesin from chromosome arms in prophase and from centromeres in anaphase. Cell *103*, 399-410.

Wallis, J.W., Chrebet, G., Brodsky, G., Rolfe, M., and Rothstein, R. (1989). A hyper-recombination mutation in S. cerevisiae identifies a novel eukaryotic topoisomerase. Cell *58*, 409-419.

Wang, J.C. (1971). Interaction between DNA and an Escherichia coli protein omega. J Mol Biol *55*, 523-533.

Wang, J.C. (1998). Moving one DNA double helix through another by a type II DNA topoisomerase: the story of a simple molecular machine. Quarterly reviews of biophysics *31*, 107-144.

Wang, J.C. (2002). Cellular roles of DNA topoisomerases: a molecular perspective. Nature reviews Molecular cell biology *3*, 430-440.

Wang, X., Reyes-Lamothe, R., and Sherratt, D.J. (2008). Modulation of Escherichia coli sister chromosome cohesion by topoisomerase IV. Genes & development *22*, 2426-2433.

Wang, Z., and Droge, P. (1996). Differential control of transcription-induced and overall DNA supercoiling by eukaryotic topoisomerases in vitro. The EMBO journal *15*, 581-589.

Warren, C.D., Eckley, D.M., Lee, M.S., Hanna, J.S., Hughes, A., Peyser, B., Jie, C., Irizarry, R., and Spencer, F.A. (2004). S-phase checkpoint genes safeguard high-fidelity sister chromatid cohesion. Molecular biology of the cell *15*, 1724-1735.

Watson, J.D., and Crick, F.H. (1953a). Genetical implications of the structure of deoxyribonucleic acid. Nature *171*, 964-967.

Watson, J.D., and Crick, F.H. (1953b). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature *171*, 737-738.

Wei, X., Samarabandu, J., Devdhar, R.S., Siegel, A.J., Acharya, R., and Berezney, R. (1998). Segregation of transcription and replication sites into higher order domains. Science *281*, 1502-1506.

Wellinger, R.E., Prado, F., and Aguilera, A. (2006). Replication fork progression is impaired by transcription in hyperrecombinant yeast cells lacking a functional THO complex. Molecular and cellular biology *26*, 3327-3334.

Wells, J.N., Gligoris, T.G., Nasmyth, K.A., and Marsh, J.A. (2017). Evolution of condensin and cohesin complexes driven by replacement of Kite by Hawk proteins. Current biology : CB *27*, R17-R18.

Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T.*, et al.* (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature *451*, 796-801.

Wittig, B., Dorbic, T., and Rich, A. (1989). The level of Z-DNA in metabolically active, permeabilized mammalian cell nuclei is regulated by torsional strain. The Journal of cell biology *108*, 755-764.

Wu, H.Y., Shyy, S.H., Wang, J.C., and Liu, L.F. (1988). Transcription generates positively and negatively supercoiled domains in the template. Cell *53*, 433-440.

Zegerman, P., and Diffley, J.F. (2007). Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. Nature *445*, 281-285.

Zeman, M.K., and Cimprich, K.A. (2014). Causes and consequences of replication stress. Nature cell biology *16*, 2-9.

Zhang, J., Shi, X., Li, Y., Kim, B.J., Jia, J., Huang, Z., Yang, T., Fu, X., Jung, S.Y., Wang, Y.*, et al.* (2008). Acetylation of Smc3 by Eco1 is required for S phase sister chromatid cohesion in both human and yeast. Molecular cell *31*, 143-151.