MECHANISM OF REGULATION OF HUMAN RAD51 RECOMBINASE THROUGH POST TRANSLATIONAL MODIFICATIONS & MEDIATOR PROTEINS

BY

SHYAMAL SUBRAMANYAM

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Biochemistry
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Doctoral Committee:

Professor Mary A. Schuler, Chair
Associate Professor Maria Spies, Director of Research (University of Iowa)
Professor Robert B. Gennis
Professor Emad Tajkhorshid

ABSTRACT

RAD51 protein plays an important role in homologous genetic recombination (HR), an essential DNA metabolic process used by cells to faithfully repair the most deleterious forms of DNA damage and maintain genomic integrity. RAD51 along with its bacterial counterpart RecA, bacteriophage UvsX and archaeal RadA have been subjected to genetic and biochemical scrutiny resulting in a plentitude of mechanistic and functional information on formation, regulation and activities of these recombinases. An important disconnect between these two lines of investigation still exists because the recombinase functions of RAD51 are highly regulated through mediator proteins like the BRCA2 recombination mediator, and a host of post translational modifications, namely phosphorylation. The mechanism and biochemical implications of these regulatory processes have not been satisfactorily evaluated in-vitro.

This work characterizes the interaction between RAD51 and the BRCA2 recombination mediator protein using computational methods to generate homology models for this interaction which are validated through experimental data. Using the knowledge gained from our structural model for the RAD51 recombinase, I developed a novel strategy to understand several key mechanisms for the regulation of RAD51 by phosphorylation. RAD51 is phosphorylated by the cABL tyrosine kinase. The mechanistic and functional significance of this event is largely disputed. Using biochemical and single molecule assays reconstituting major activities of RAD51, I have successfully dissected the biochemical mechanism of regulation of RAD51 by the c-Abl kinase. The results of this work strongly correlate with observations made in previous cell based analysis.

Dedicated to My Parents

ACKNOWLEDGEMENTS

The pursuit of knowledge has always been something that has driven me forward. Towards this, I inwardly volunteered to steer my motivation towards constructive research ends and contributing to the global scientific community. Obtaining a PhD, to me, has been the most important step in my path towards academic enlightenment and intellectual freedom. This period of my life, I will always look back on with a feeling of great fulfillment; none of which would be possible without some very important people who have walked me through this process. I deeply appreciate your patience, encouragement and support towards my goals.

I would like to thank my thesis advisor, Dr. Maria Spies for giving me the opportunity to work under her tutelage. I am deeply indebted to her for the support and liberty she gave me, teaching to think like a scientist, critique my work and explore the limits of my research while keeping me soundly grounded towards realistic scientific goals. I attribute my ability work as an independent researcher, entirely to her. I would like to thank her for affording me the opportunity to travel to conferences and talk about my work and build fruitful academic relationships. I would also like to thank Professors Mary A. Schuler, Emad Tajkhorshid, Robert B. Gennis and Taekjip Ha for serving on my committee and for offering support and constructive comments on my research and Jeff Goldberg for helping me jump several administrative hoops during my relocation to the University of Iowa. I appreciate Drs. Masayoshi Honda and Mohammed Ghoneim for their support and many intellectual discussions I have had with them. Using the wealth of their knowledge, I was able to direct my research towards very fruitful gains. Those long nights in the lab would not be fun if it weren't for them. I am appreciative to all the past and current Spies lab members.

In friends, I found a family away from home. I'd like to thank Swabarna Mukherjee for being there when the going got tough. He has supported my ambitions and seen me grow from the very early days. I am grateful to Bijoy Desai, Ranjani Murali, Suchithra Prabhu, Ambika Nadkarni, Rajashekar Iyer & Fletcher Bain for their wonderful friendship. They have been a source of intellectual and personal growth for me through graduate school and continue to be so. I would like to thank Wan-Chian Chen for her love and companionship. Finally, none of this would be possible without the support of my parents who gave me the opportunities to pursue my dreams of becoming a scientist. I hope I make you all proud.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION----------1

CHAPTER 2: CONTRIBUTIONS OF THE RAD51 N-TERMINAL DOMAIN TO BRCA2-RAD51 INTERACTION----------18

CHAPTER 3: TYROSINE PHOSPHORYLATION STIMULATES ACTIVITY OF HUMAN RAD51 RECOMBINASE----------51

CONCLUSIONS & FUTURE DIRECTION----------93

BIBLIOGRAPHY----------94

APPENDIX: IMPORTANT DNA SEQUENCES RELATED TO RAD51 STUDIES----------113

CHAPTER 1: INTRODUCTION

Double stranded DNA (dsDNA) is the carrier of genetic information throughout life. The genomic DNA in all living organisms is constantly subjected to change. Most of these changes are mechanisms evolved to protect the integrity of the DNA or aid in its function as a repository of information storage. However, certain types of modifications are detrimental to the integrity of the genome. These sources of 'DNA Damage' range from external factors such as ultraviolet light (UV), Ionizing radiation (IR), environmental pollutants such as heavy metals as well as chemical substances that are intentionally created to damage DNA (e.g. chemotherapeutic agents), to internal biological processes like DNA hydrolysis, oxidative stress and enzymatic DNA damage due to replication errors and other sources (Hoeijmakers 2009; Lindahl, 1993). It is essential that mechanisms be available to the organisms to be able to correct for these aberrations. The various types of DNA damage include base modifications, abasic sites, inter-strand cross links (ICL's), Single Strand Breaks (SSB's) and DNA Double Strand Breaks (DSB's) (Hoeijmakers 2009; Iyama and Wilson Iii, 2013; Lindahl, 1993) (Figure 1.1). Coupled with environmental factors, DNA damaging agents can lead to ~10^5 lesions per day (Ciccia and Elledge, 2010; Hoeijmakers 2009). This can cause critical mutations in the genome which manifest in the form of cancers, multiple inherited disorders, ageing and even cell death (Kanaar et al., 1998).

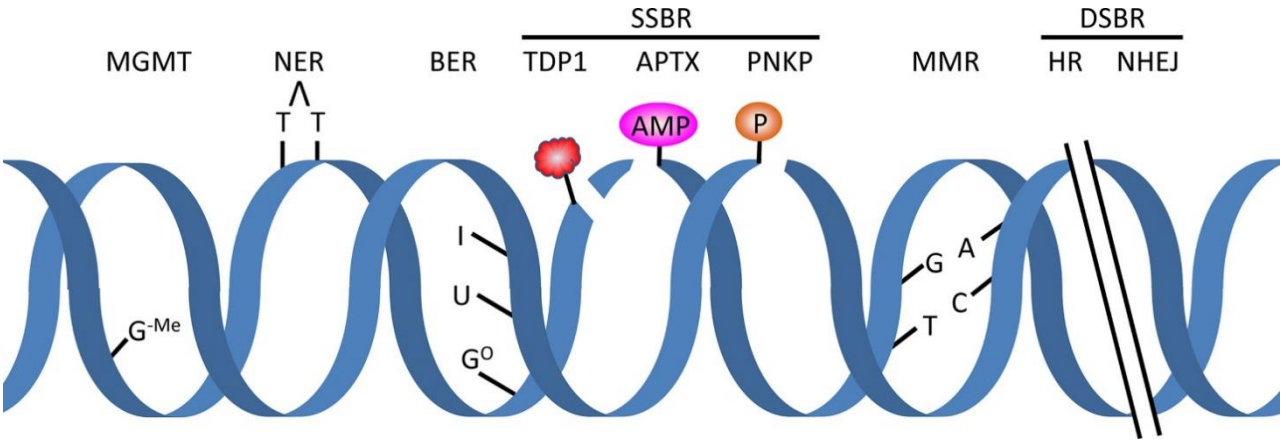


Figure 1.1 DNA damage and repair responses. DNA repair pathways (top) and examples of corresponding DNA damage (bottom). The detailed molecular mechanisms for the repair responses are provided in text. APTX, aprataxin; BER, base excision repair; DSBR, DNA double strand break repair; HR, homologous recombination; MGMT, O6-methylguanine-DNA methyltransferase; MMR, mismatch repair; NER, nucleotide excision repair; NHEJ, nonhomologous end joining; PNKP, polynucleotide kinase 3-phosphatase; SSBR, DNA single strand break repair; SSBs, DNA single strand breaks; TC-NER, transcription-coupled NER; TDP1, tyrosyl-DNA phosphodiesterase 1; G-Me, O6-Methylguanine; T^T, thymine dimer; I, inosine; U, uracil; Go, 8-oxoguanine. Reproduced with permission from Iyama T & Wilson Iii DM (2013), DNA repair 12(8):620-636. Copyright Elsevier B.V.

The mechanisms that preserve the genomic integrity of the cell are collectively referred to as the 'DNA repair' pathways. These pathways, each, focusing on particular type of DNA damage, along with cellular checkpoint mechanisms that coordinate replication, transcription, chromatin remodeling and generating genetic diversity ensure the maintenance of genomic integrity inside the cell (Hoeijmakers 2009).

The choice of DNA repair pathway is largely dependent on the type of DNA damage, however the state of the cell cycle is also an important factor in selection of appropriate pathway as all pathways are tightly regulated and are active during different stages of the cell cycle. Base Excision Repair (BER) is generally the pathway of choice for DNA damage that does not severely affect the helical structure of the DNA. Such damage occurs through base modifications, alkylation, deamination, depurination as well as certain single strand breaks (Figure 1.1) (Krokan and Bjørås, 2013; Wallace, 2014). During BER, the damaged base is first excised, followed by end processing, gap filling and finally ligation which restores the DNA (Krokan and Bjørås, 2013). Nucleotide Excision Repair (NER) is responsible for removal of bulky DNA lesions that destabilize the duplex DNA caused by UV radiation and other mutagens. The NER pathway can be triggered through Global Genome NER (GG-NER) which can occur anywhere in the genome and is generally activated by UV damage or Transcription Coupled NER (TC-NER) which is activated by RNA polymerase stalled at a lesion during transcription (Schärer, 2013). It involves removal of a small tract of DNA ~13 bases long (Kisker et al., 2013) followed by a polymerase that fills in the gap repairing the DNA (Schärer, 2013). Errors made by the replication machinery lead to mismatched bases and are repaired by the Mismatch Repair (MMR) pathway where the mismatched bases are detected, excised and the gap filled in (Honda et al., 2014; Kolodner, 1996). Double strand breaks (DSB's) are the most deleterious forms of DNA damage which can lead to severe loss of genetic information. These are discussed further.

Sources of Double Strand Breaks

DNA damage due to Double Strand Breaks (DSB's) is less frequent compared to other forms of lesions but comprise the most deleterious and cytotoxic forms of DNA damage (Mehta and Haber, 2014) and repairing these forms of damage are of the utmost priority for the cell. This type of damage involves the simultaneous breakdown of the phosphodiester backbones of both strands of

DNA. Agents that cause DSB's include certain anticancer agents (clastogens) and Ionizing Radiation (IR). Chemotherapeutic agents include methyl methanosulfonate (MMS) and temozolomide which alkylate DNA. Agents like Mitomycin C and Cisplatin induce DNA crosslinks while compounds such as Bleomycin mimic radiation damage. Inhibitors such as etoposide trap topoisomerases inducing DSB's. (Mehta and Haber, 2014). Ionizing Radiation (IR) produces reactive oxygen species which attach the sugar phosphate backbone and in high doses can cause breaks in complementary strands of DNA leading to DSB's (Hoeijmakers 2009).

In addition to these exogenous agents, DNA is very vulnerable to the formation of DSB's during the natural replicative cycle of the cell. Unusual DNA structures or collisions with transcription machinery lead to stalled replication forks which can cause DBS's (Aguilera and Gaillard, 2014). Programmed DSB's are initiated by the cell in several instances which lead to genetic diversity during meiosis (Keeney, 2008; Lam and Keeney, 2015) as well as generation of antibody diversity by V(D)J recombination (Soulas-Sprauel et al., 2007).

Repair of DNA Double Strand Breaks

There are two main pathways that the cell uses to repair DSB's, Non Homologous End Joining (NHEJ) and Homologous Recombination (HR) (Kass and Jasin, 2010). NHEJ involves the limited modification of broken DNA ends followed by their ligation (Mehta and Haber, 2014). Alternative NHEJ (alt-NJEJ) pathways are executed when canonical NHEJ is unavailable with the propensity to create mutations, because joints often harbor local deletions with relatively long stretches of microhomology, itself often called MMEJ (Chiruvella et al., 2013). In higher eukaryotes, NHEJ activities are active through all stages of the cell cycle but are generally limited to the G_1 stage where key components of the HR pathway are blocked (Aylon et al., 2004; Ira et al., 2004). It is defined as the Joining of two DSB ends by direct ligation (Chiruvella et al., 2013). Since this process need not result in accurate repair of the DNA, it is recognized as being error prone or an unfaithful mechanism of Double Strand Break Repair (DSBR). Briefly, NHEJ involves the binding of the DSB ends by the Ku protein complex which activate a slew of NHEJ associated proteins through the DNA Protein Kinase (DNA-PK) through association. The ends are then ligated by the DNA Ligase IV and any gaps filled in by the DNA polymerase X (PolX) (Chiruvella et al., 2013) (Figure 1.2).

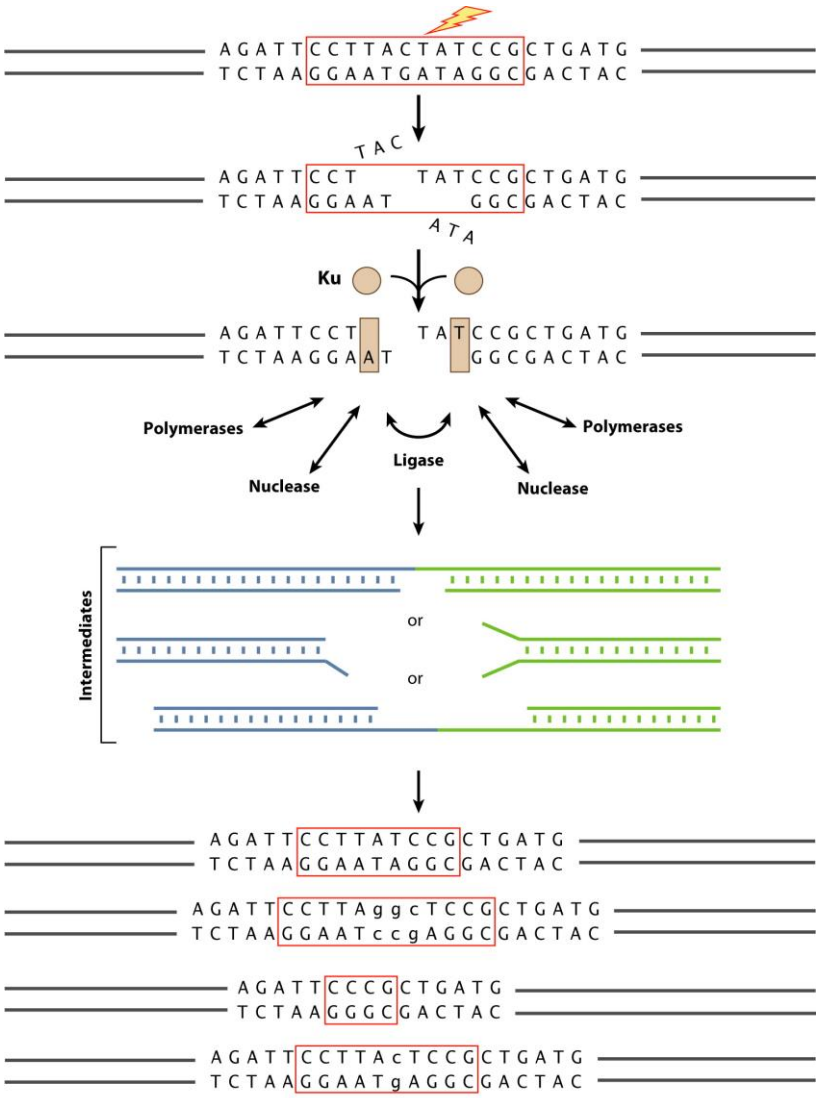


Figure 1.2 General steps of Non Homologous DNA end joining (NHEJ). The lightning arrow indicates ionizing radiation (IR), a reactive oxygen species (ROS), or an enzymatic cause of a DSB. Ku binding to the DNA ends at a double-strand breaks (DSBs) improves binding by nuclease, DNA polymerase, and ligase components. Note that Ku is thought to change conformation upon binding to the DNA end, as depicted by its shape change from a sphere to a rectangle. Flexibility in the loading of these enzymatic components, the option to load repeatedly (iteratively), and independent processing of the two DNA ends all permit mechanistic flexibility for the NHEJ process. This mechanistic flexibility is essential to permit NHEJ to handle a very diverse array of DSB end configurations and to join them. In addition to the overall mechanistic flexibility, each component exhibits enzymatic flexibility and multifunctionality, as discussed in the text. The figure shows that there are many alternative intermediates in the joining process (middle). These intermediates are reflected in a diverse DNA sequences at the junction of the joining process (bottom). Reproduced from, Lieber MR (2010) Annual review of biochemistry 79:181-211, Copyright Annual Reviews.

Accurate repair of DSB's is reliant on the set of homology-directed DNA repair pathways that result in the faithful restoration of the damaged DNA to its original state without the loss of any sequence information. DNA repair through Homologous Recombination (HR) requires a second homologous DNA (donor sequence), from the homologous chromosome or the sister chromatid, as a template for the lost DNA sequence. If the donor sequence is similar to the region surrounding the DSB, the repair is accurate and the DNA sequence is restored to that preceding the DNA damage (Moynahan and Jasin, 2010).

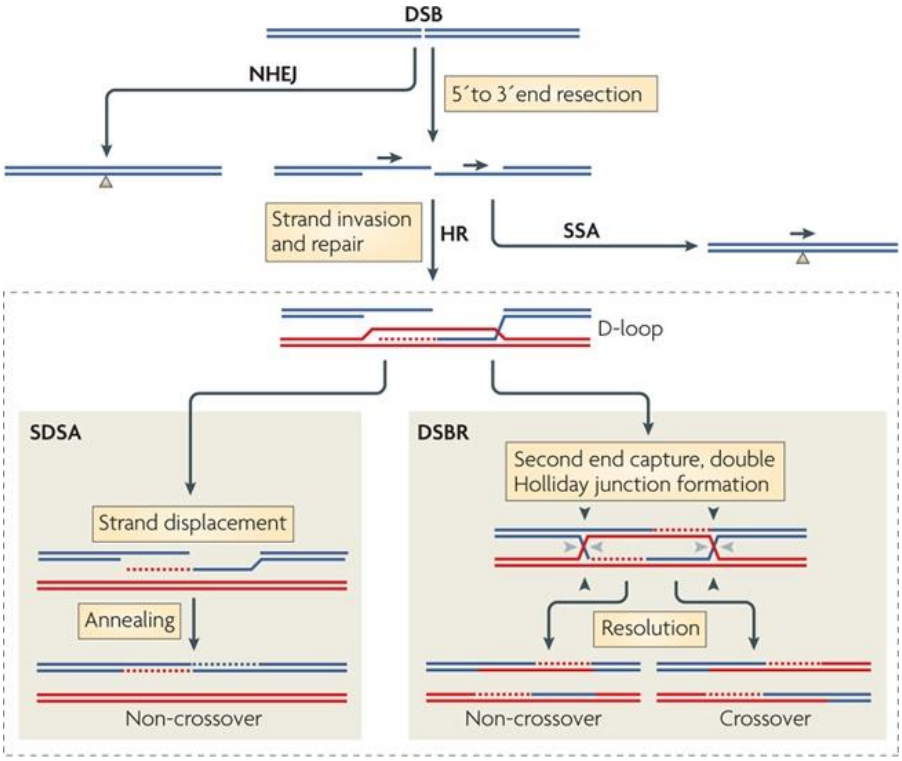


Figure. 1.3 Pathways of DNA DSB repair. Double-strand breaks (DSBs) are efficiently repaired in mammalian cells by homologous recombination (HR) and non-homologous end joining (NHEJ). HR initiates with end resection, which produces a 3′ single-stranded end that can invade a homologous template to initiate repair. Alternative HR pathways can ensue from the displacement loop (D-loop) intermediate: synthesis-dependent strand annealing (SDSA) and DSB repair (DSBR). In SDSA, the newly synthesized strand is displaced to anneal to the other DNA end, resulting in a non-crossover outcome with no change to the template DNA. In DSBR, the second DNA end is 'captured' by the D-loop to form a double Holliday junction, which in principle can result in a non-crossover (cleavage at black or grey arrowheads) or a crossover (cleavage at black arrowheads on one side and grey arrowheads) outcome. NHEJ involves the joining of non-homologous DNA ends. It can be imprecise and lead to deletions and other mutations through numerous end-processing steps (not shown). Single-strand annealing takes place when end resection occurs at sequence repeats (arrowheads) to provide complementary single strands that anneal, giving rise to a product with a single copy of the repeat and a deletion of intervening sequences. Reproduced with permission from, Moynahan ME & Jasin M (2010) Nature reviews. Molecular cell biology 11(3):196-207, Copyright Macmillan Publishers Limited.

The central step in homologous recombination involves the formation of a nucleoprotein filament which invades the homologous DNA template searching for homology ending in the forming a D-Loop intermediate. This intermediate is then funneled into different pathways which lead to the repair of the damaged DNA (Figure 1.3) (Moynahan and Jasin, 2010). Like all DNA repair pathways, DSBR is initiated by a DNA damage sensing mechanism which involves the immediate recruitment of the MRN (MRE11-RAD50-NBS1) complex, which binds dsDNA ends. This leads to the activation of the ATM (ataxia-telangiectasia mutated) and ATR (ATM- and Rad3-Related) kinases (Lee and Paull, 2007; Maréchal and Zou, 2013), while simultaneously initiating DNA resection (generation of 3` ssDNA ends) through the exonuclease activity of the MRE11 protein (Symington, 2014). The resection process is then handed to the CtIP/EXO1/BLM or BLM/DNA2/RPA complex (Nimonkar et al., 2011), which generate ssDNA ends that are bound by the ssDNA binding protein RPA. The ATM/ATR kinases activated here in turn activates phosphorylates a number of proteins involved in cell cycle checkpoint control, apoptotic responses and DNA repair. Phosphorylation of these substrates by ATM initiates cell-cycle arrest at G_1/S, intra-S and G_2/M checkpoints and promotes DNA repair (Lee and Paull, 2007). The active intermediate for the strand invasion step is the RAD51-ssDNA nucleoprotein filament. This filament is formed by displacement of the RPA coated ssDNA by the BRCA2 protein which binds RAD51 through its BRC repeats (Carreira et al., 2009; Shahid et al., 2014) and helps nucleoprotein filament assembly along with several other mediator proteins (Suwaki et al., 2011; Zelensky et al., 2014a) which stabilize the RAD51 filaments. The active nucleoprotein filament then invades the homologous DNA template forming a Displacement Loop (D-Loop) intermediate (Figure 1.3). The invaded strand is then extended at the 3`-OH end with the help of the RAD54 protein and DNA polymerases (Li and Heyer, 2009; Murakami and Trakselis, 2014). At this stage the D-Loop can be disassembled by structure selective helicases in which the complementary strand anneals with the single-stranded tail of the other end of the DSB. After fill-in synthesis and ligation, this pathway generates non-crossover products and is referred to as synthesis-dependent strand annealing (SDSA) (Prakash et al., 2009) (Figure 1.4).

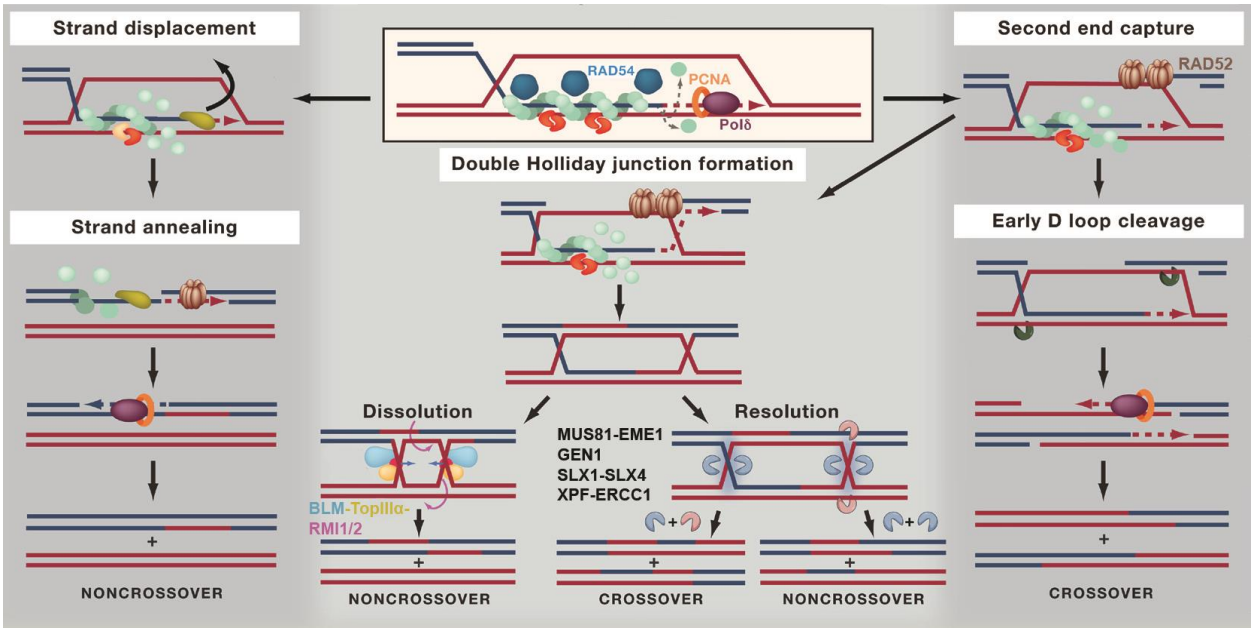


Figure 1.4 Series of events following formation of D-Loops in Double Strand Break Repair. Once formed, the complex of RAD51 and single-stranded DNA searches for a homologous sequence in double-stranded DNA and then promotes invasion of the single-stranded DNA into donor double-stranded DNA to form a joint molecule with a displaced strand (D loop). RAD54 displaces RAD51 from double-stranded DNA in vitro, uncovering the 3′ end of paired intermediates to allow initiation of DNA synthesis. DNA polymerase δ extends the 3′ end from the broken chromosome using the donor strand as a template and replacing nucleotides lost by end resection. To resolve the intermediate by synthesis-dependent strand annealing, the invading strand that has been extended by DNA synthesis is displaced (strand displacement) and anneals to complementary sequences exposed by 5′–3′ resection of the other side of the break forming non-crossover products exclusively (strand annealing). In addition to preventing initiation of inappropriate recombination events by disrupting RAD51 nucleoprotein filaments, the FBH1 DNA helicase can disrupt D loop intermediates to promote non-crossover. The FBH1 and RTEL1 helicases also dissociate D loop intermediates to facilitate synthesis-dependent strand annealing. In the canonical DNA double-strand break repair model, the other end of the break interacts with the displaced strand of the strand invasion intermediate (second end capture) and the 3′ end primes DNA synthesis, forming a double Holliday junction (dHJ) intermediate (double Holliday junction formation). The dHJ intermediates can be dissolved or resolved to yield separate intact duplex molecules. Alternatively, the extended D loop structure could be cleaved by the MUS81-EME1 nuclease prior to formation of a mature dHJ intermediate (early D loop cleavage). MUS81-EME1 exhibits higher activity toward D loop and nicked Holliday junction intermediates than intact Holliday junctions and could promote crossovers by cleaving the strand invasion intermediate directly. Dissolution of dHJ intermediates requires the combined activity of the BLM helicase, which drives migration of the constrained Holliday junctions, and the TOPIIIα-RMI1/2 complex, which decatenates the interlinked strands between the two Holliday junctions eventually leading to non-crossover products. On the other hand, resolution through nucleolytic cleavage of the Holliday junctions can yield crossover (cutting inner strands of one Holliday junction and outer strands of the other) or non-crossover (cutting both junctions in the same plane) products. Adapted with permission from, Mazon G, Mimitou EP, & Symington LS (2010) SnapShot: Homologous recombination in DNA double-strand break repair. Cell 142(4):646, 646 e641. Copyright Elsevier Inc.

An alternative mechanism involves formation of joint-molecules such as single or double Holliday Junctions (HJ's) which can either be dissolved by the BTR, (BLM-TOPIIIα-RMI1-RMI2) complex leading to a non-crossover event (Bizard and Hickson, 2014), or be resolved by the structure-selective nucleases, such as MUS81-EME1, SLX1-SLX4, and GEN1 proteins which can produce crossover or non-crossover products (Wyatt and West, 2014) (Figure 1.4).

Homologous Recombination in DNA Replication

In addition to faithfully repairing damaged DNA, the cell has to ensure the faithful and complete replication of the chromosomes to maintain genomic integrity during cell division. Replication forks are susceptible to stalling or collapse when the replication machinery encounters secondary structure in DNA, DNA bound proteins or DNA lesions. Several chemicals like hydroxyurea and aphidicolin inhibit replication leading to stalled or collapsed replication forks.

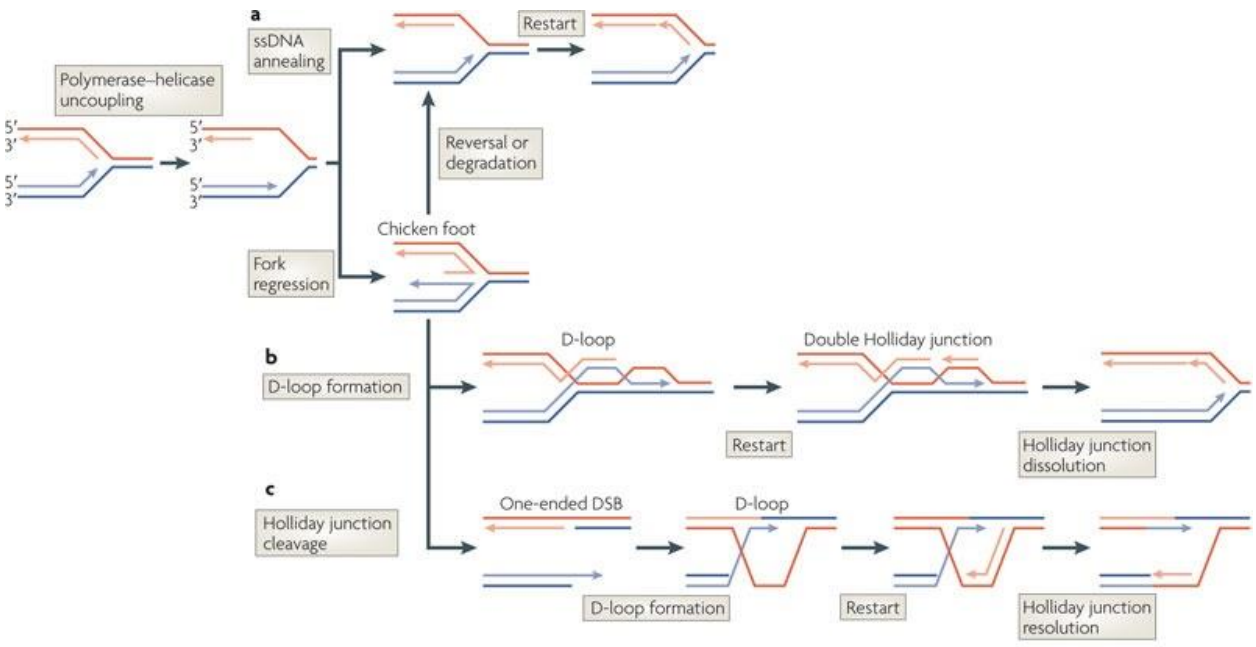


Figure 1.5 Models of replication fork restart. a | Restart by fork remodeling. A stalled replication fork might be stabilized by the re-annealing of single-strand DNA (ssDNA) generated by excessive unwinding of the template, or might undergo regression and pairing of the newly synthesized strands to form a Holliday junction in a structure termed a 'chicken foot'. Restart after Holliday junction formation may be difficult if it requires the removal and subsequent re-loading of the replication machinery. **b** | Holliday junction-mediated fork restart. The double-stranded DNA (dsDNA) end of the Holliday junction is recombined into the template through strand invasion, forming a displacement loop (D-loop). In E. coli, the D-loop allows re-loading of the replication machinery. The invading strand re-anneals with the template, forming a double Holliday junction that can be removed by Holliday junction resolution, which leads to sister chromatid exchanges, or Holliday junction dissolution, which avoids the generation of

recombination products. **c** | Double-strand break (DSB)-mediated restart. The Holliday junction is processed into a one-ended DSB and fork restart is achieved through homologous recombination repair of the DSB in a mechanism analogous to break-induced replication. The resulting single Holliday junction would be resolved by Holliday junction resolution. The leading strand and leading-strand template are shown in light and dark blue, respectively, and the lagging strand and lagging-strand template are shown in light and dark red, respectively. Reproduced with permission from Petermann E & Helleday T (2010) Nature reviews. Molecular cell biology 11(10):683-687, Copyright Macmillan Publishers Limited.

Restarting stalled or collapsed replication forks occurs through several pathways (Petermann and Helleday, 2010) (Figure 1.5), most of which require processing by the DSBR machinery which is dependent on recombination and hence necessitates the participation of the RAD51 protein. Collapsed replication forks leads to generation of single ended DSB's which are processed by the end resection machinery (Franchitto and Pichierri, 2002), which helps loading of RAD51 onto the resected ssDNA ends. The nucleoprotein filament then invades the other replicating strand forming D-Loops which eventually lead to replication fork restart (Figure 1.5c). In addition to DSB mediated fork restart, RAD51 is also implicated in Holliday Junction mediated restart (Figure 1.5b) which does not require the presence of a double strand break (Petermann et al., 2010).

The RAD51 Recombinase

RAD51 recombinase, is the eukaryotic homolog of the *Escherichia coli* RecA protein. The yeast *RAD51* gene is a member of the RAD52 epistasis group, which includes *RAD50, RAD51, RAD52, RAD54, RAD55, and RAD57* genes (Game, 1983), which were initially identified as mutants defective in the repair of DNA damage caused by ionizing irradiation and were subsequently shown to be deficient in both genetic recombination and the repair of DNA lesions (Baumann et al., 1996). Homozygous deletions for the *RAD5l* null mutation (*RAD5* $^{-/-}$), in mice were found to be embryonically lethal (Lim and Hasty, 1996; Tsuzuki et al., 1996); suggesting an important role for RAD51 in cell proliferation and development. RAD51 protein was found to catalyze DNA strand exchange through the recombinational "D-Loop" intermediate in an ATP dependent manner (Baumann et al., 1996; Sung, 1994; Sung and Robberson, 1995). These biochemical studies cemented the role of RAD51 as a recombinase protein which enabled homologous recombination in the eukaryotic cell.

The RAD51 protein has been investigated very extensively through the years, which has led to a wealth of information regarding its biochemical properties. The Human RAD51 is a 339 amino acid long, 37.5KDa protein. The structure of the RAD51 protein is highly conserved among eukaryotes, and the human RAD51 homolog exhibits 67% sequence identity to its yeast counterpart (Shinohara et al., 1993; Yoshimura et al., 1993). RAD51 has considerably slow ATP hydrolytic activity compared to the *E. coli* RecA protein (Tombline and Fishel, 2002), but binds ssDNA similarly, in an ATP dependent manner (Sung and Robberson, 1995) forming a right handed helical filament on ssDNA with a mean helical pitch of ~96Å, extending its contour length to 1.5 times of the B-Form DNA which corresponds to an axial rise of -18.6 bases of ssDNA per helical repeat of the nucleoprotein filament (Sung and Robberson, 1995) (Figure 1.6). RAD51 also interacts with dsDNA in a similar ATP dependent manner with a pitch of ~80Å extending its contour length to 1.35 times of the B-Form DNA which corresponds to an axial rise of -17.6 base pairs per helical repeat of the dsDNA nucleoprotein filament (Ogawa et al., 1993; Sung and Robberson, 1995) (Figure 1.6). Studies have shown that in an ssDNA-RAD51 nucleoprotein filament, each monomer binds ~3 bases with axial rise per ssDNA base extended to nearly 5 Å with the interconversion of sugar puckers inducing horizontal base rotation (Lee et al., 2015; Masuda et al., 2009; Nishinaka et al., 1998; Shibata et al., 2001). While RAD51 protein binds ssDNA as well as dsDNA substrates similarly, it is the ssDNA-RAD51 filament that is productive in strand exchange and homology search, whereas, the RAD51 filament on dsDNA is not only incapable of initiating pairing and strand exchange, but is in fact strongly inhibitory to these reactions (Sung and Robberson, 1995).

Attempts have also been made to describe the structures of the Human RAD51 but have so far been unsuccessful with only one partial structure of its core domain available (Pellegrini et al., 2002). Since the recombinases are highly conserved across species, crystal structures of *E. coli* RecA (Chen et al., 2008) (Figure 1.7A) and the *Saccharomyces cerevisiae* Rad51 protein (Conway et al., 2004) (Figure 1.7B) have been used to describe the function of the human RAD51 homolog. However, due to slight differences in mechanism, filament structure as well as biochemical properties and regulation, information from these sources need not be entirely relevant to the human RAD51 recombinase.

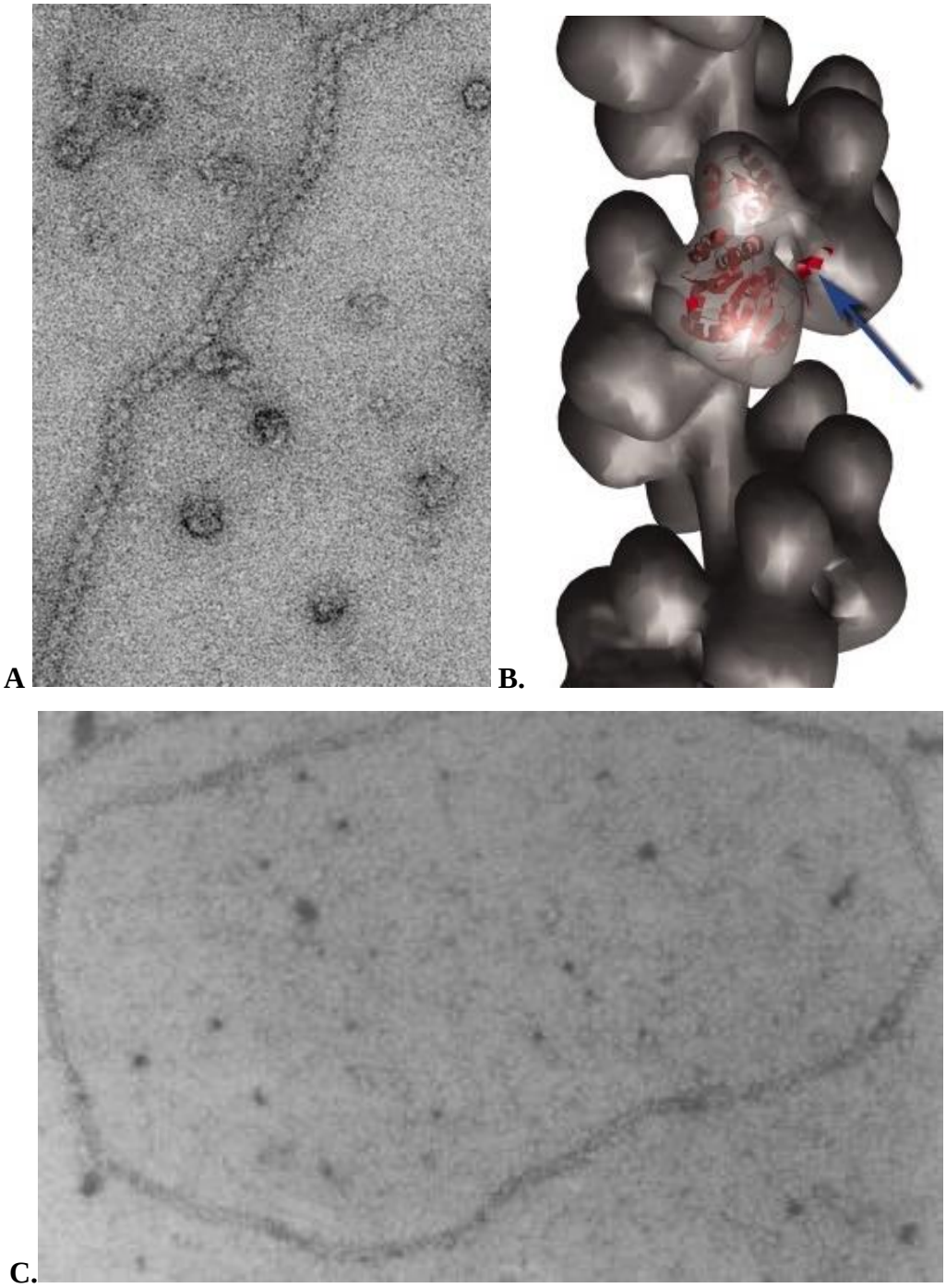
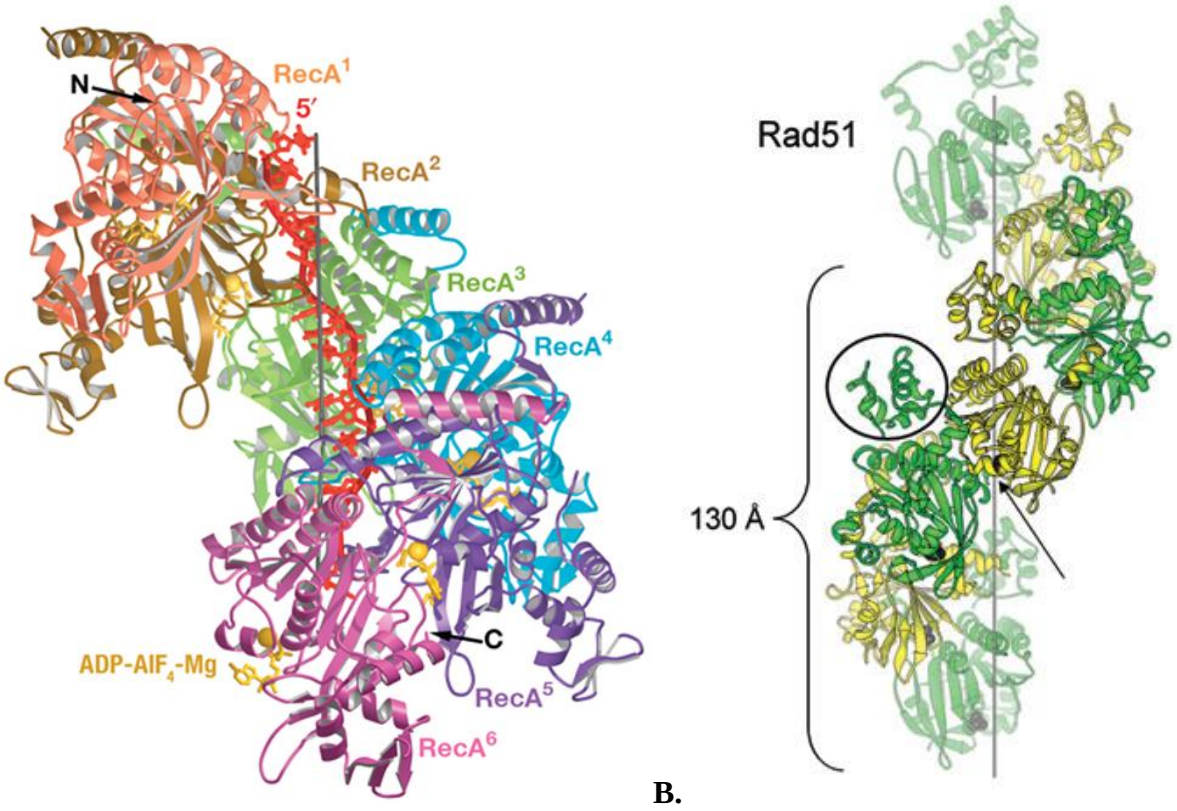


Figure 1.6 Structure of the RAD51 nucleoprotein filament. (A) Electron micrographs of a representative RAD51 filament formed on ssDNA in the presence of ATP**, (B)** 3-Dimensional reconstruction of an ssDNA nucleoprotein filament showing a single monomer of RAD51. Adapted from (Galkin et al., 2005), Copyright National Academy of Sciences. **(C)** Electron micrographs of a representative RAD51 filament formed on dsDNA in the presence of ATP. Adapted with permission from Sung P & Robberson DL (1995), Cell 82(3):453-461. Copyright Elsevier Inc.



A. **B.**

Figure 1.7. Crystal Structures of the *E. coli* RecA and yeast Rad51 Filaments. (A) Structure of the RecA6–(ADP-AlF4-Mg)6–(dT)18 complex. The six RecA protomers are numbered from the N-terminal RecA of the fusion protein and are coloured pink, brown, green, cyan, purple and magenta, respectively. Only 15 of the 18 nucleotides are ordered (red). TheDNAbackbone is traced by a red coil. The six ADP-AlF4-Mg molecules are coloured gold. The five individual rotation/translation axes that relate adjacent RecA protomers are shown as grey vertical lines. Adapted with permission from, Chen Z, Yang H, & Pavletich NP (2008), Nature 453(7194):489-484. Copyright Nature Publishing Group. (B) The Rad51 filament found in these crystals has a helical pitch of 130 Å and is composed of two crystallographically independent monomers (yellow and green) that alternate to form a filament with exact three-fold but only approximate six-fold screw symmetry. A sulfate (black spheres) mimics the binding of phosphate in the ATPase site, which is nestled directly at the interface between two protomers (arrow). One of the N-terminal domains that line the upper surface of the filament is circled. Adapted with permission from, Conway AB, et al. (2004), Nature structural & molecular biology 11(8):791-796. Copyright Nature Publishing Group.

One of the most important characteristics of the RAD51 nucleoprotein filament is its ability to perform homology search. i.e. on invading the homologous duplex DNA, the nucleoprotein filament searches for regions of homology forming the heteroduplex DNA (Morrical, 2015). This is defined as a dsDNA product that arises from a recombination event. The strands of the heteroduplex formed this way can be completely complementary or can contain small regions of

non-complementarity. This is stabilized by Watson and Crick pairing that ensures that the heteroduplex remains as stable DNA once the recombination machinery dissociates from it. The ability to form heteroduplex DNA using strands from two different DNA molecules is central to the processes that control genome stability. It allows inheritance of genetic information and genetic diversity within the resulting populations (Amunugama et al., 2012a). During meiosis, the RAD51 protein along with its meiotic homolog DMC1 forms heteroduplex DNA facilitating crossing-over and allelic exchange between homologous chromosomes; this process ensures that progeny are not identical clones of their parents and ensures a genetically diverse population (Lam and Keeney, 2015). The mechanism for homology search has intrigued the field for a long time and has been a debated topic. How the RAD51 protein is able to find a small region of homology compared to the expanse of the genome is staggering. However, research using state-of-the-art single molecule methods are beginning to shed new light on this topic. Experiments from these studies show a three dimensional search strategy where the nucleoprotein filament initiates intersegmental contact between coiled regions of DNA (Forget and Kowalczykowski, 2012) (Figure 1.8A) while searching for short regions (~8 nucleotides long) of microhomology between the DNA. The nucleoprotein filament can undergo exchange with other regions of dsDNA bearing the same microhomology, but resists exchange with unrelated sequences.

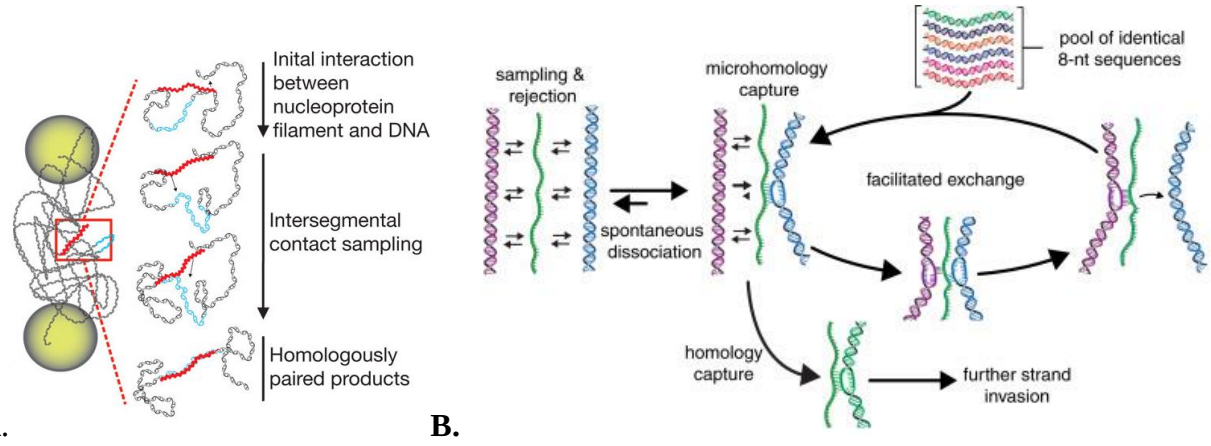


Figure 1.8 Mechanism of 3-Dimensional Homology Search (A). Model for RecA homology search by intersegmental contact sampling; for simplicity, only two simultaneous points of interaction are depicted. Adapted with permission from, Forget AL & Kowalczykowski SC (2012) Nature 482(7385):423-427. Copyright Macmillan Publishers Limited. **(B).** Model depicting a homology search mechanism involving rapid sampling and rejection of DNA lacking microhomology, followed by eventual capture of an 8-nt tract of microhomology and facilitated exchange allowing for an iterative search through sequence space. Adapted with permission from Qi Z, et al. (2015), Cell 160(5):856-869. Copyright Elsevier Inc.

Moreover, shorter tracts of microhomology are more readily exchanged with longer tracts, reflecting the higher stability of intermediates held together by longer tracts of Watson-Crick pairing which might in turn funnel the nucleoprotein filament through progressively smaller pools of sequences leading to the homologous target (Qi et al., 2015) (Figure 1.8B).

Mediators in Recombination

Due to the critical nature of the process of Homologous Recombination (HR), it is imperative that it be tightly regulated in the cell. Excessive DNA rearrangements also need to be avoided where they would be disastrous rather than beneficial. Thus, there need to be checkpoints at which the reaction can either be driven forward or reversed, an important feature required to attain quality control (Holthausen et al., 2011). Recombination mediators are proteins that have evolved to ensure quality control of the process of HR. These mediators could have a positive or negative effect on HR (Daley et al., 2014a). Proteins like RPA act in both ways paving the way for nucleoprotein filament formation by removing secondary structure from resected ssDNA while promoting the activities of the DNA resection machinery (Symington, 2014). However, due to its high affinity for ssDNA, it kinetically impedes the formation of the RAD51 nucleoprotein filament (New et al., 1998). To facilitate the formation of the RAD51 nucleoprotein filament, the BRCA2 protein sequesters RAD51 monomers and help loading them onto RPA coated ssDNA (Jensen et al., 2010; Shahid et al., 2014) (Figure 1.9).

The BRCA2 tumor suppressor, mutations in which have been linked to breast and ovarian cancers among other oncogenic conditions (Tavtigian et al., 1996), is a 3418 amino acids long protein, containing DNA- and protein-binding domains, and interacts with a number of other proteins. The most notable of these interactions are with BRCA1 and PALB2, as deficiencies in these proteins manifest similarly to the BRCA2 deficiency (Chen et al., 1999). BRCA2 interacts with RAD51 through a set of eight BRC repeats. This interaction has been studied extensively (Carreira et al., 2009; Chen et al., 1998; Pellegrini et al., 2002) and has been elaborated on further in my research *(vida infra)*. In addition to these mediators, there are a set of proteins in the cell that have detectible sequence similarities with the RAD51 protein (Lin et al., 2006; Wiese et al., 2007).

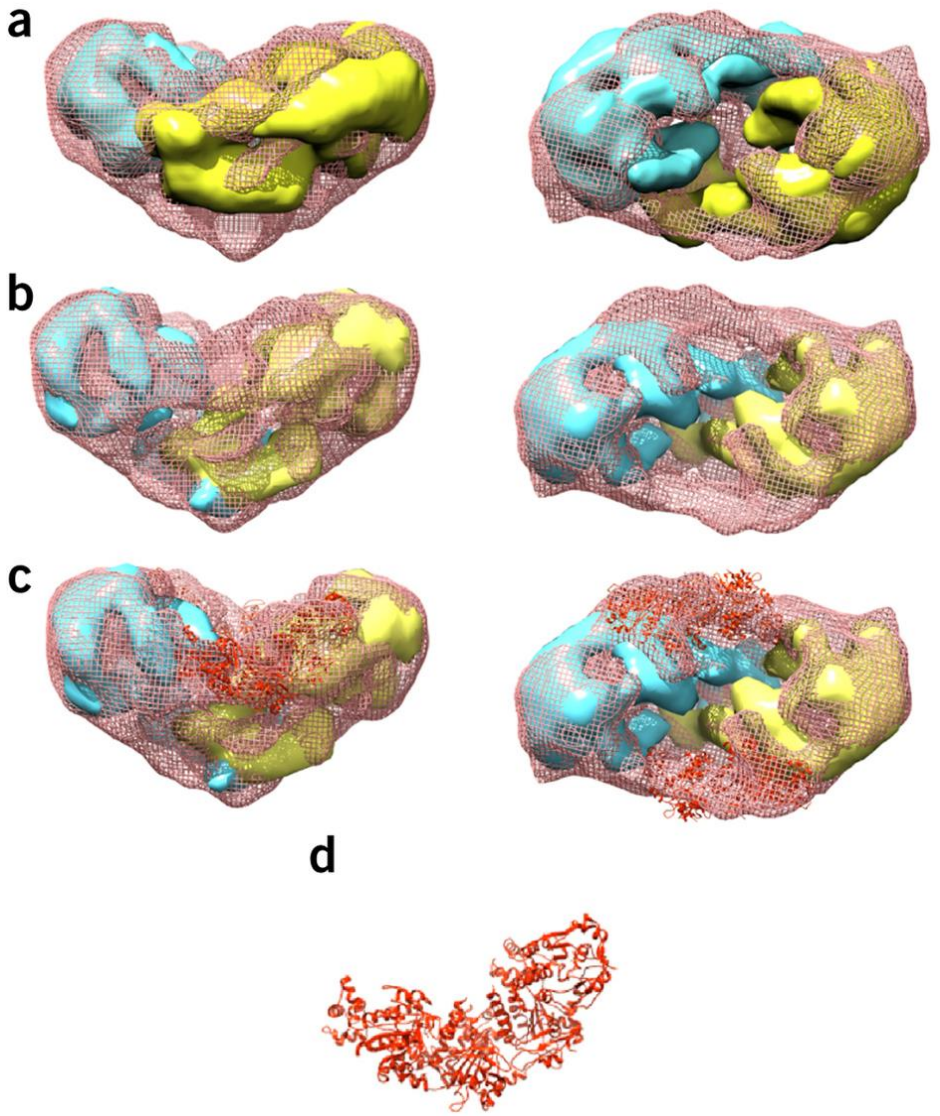


Figure 1.9 RAD51 binding to BRCA2. **(a)** Overlay of BRCA2 dimer (yellow and cyan) and BRCA2–RAD51 (pink mesh), highlighting the differences in their shape. **(b)** Rearranged BRCA2 dimer fitted into the BRCA2–RAD51 complex. **(c)** As in b. Four RAD51 monomers (orange ribbon) are fitted into the additional density in BRCA2–RAD51 not accounted for by BRCA2 density. **(d)** Four RAD51 monomers, arranged as in filaments. Reproduced with permission from, Shahid T, et al. (2014), Nature structural & molecular biology 21(11):962-968. Copyright Nature Publishing Group.

The best characterized of these so-called RAD51 paralogs are RAD51B, RAD51C, RAD51D, XRCC2, and XRCC3. These proteins mediate HR by interacting with the RAD51 protein and aiding as well as stabilizing the nucleoprotein filament. Deficiency in any of these proteins results in a homologous recombination phenotype, which is nearly as severe as RAD51 loss (Amunugama et al., 2013; Suwaki et al., 2011; Zelensky et al., 2014a). Research on these RAD51 paralogs is an ongoing and intense area in the DNA repair field. Proteins which act

antagonistically to RAD51 function are only being investigated recently. However some of these proteins have already been classified. The RAD54 protein which has been studied extensively in yeast, has been shown to stabilize yeast Rad51 filaments on ssDNA and to mediate invasion into the nucleosomal DNA, as well as displace Rad51 from dsDNA indicating a role for the clearance of Rad51 after strand invasion (Li et al., 2007; Mazin et al., 2010). However, its role in higher eukaryotes is not very clear and may be limited to a developmental role since mouse RAD54−/− embryonic stem cells are sensitive to IR but the adult RAD54−/− mice are no more IR-sensitive than the wild type animals, because their IR-sensitivity is rescued by the non-homologous end-joining (NHEJ) repair pathway (Essers et al., 2000). Proteins like FBH1 and RTEL are helicases that are thought to have antirecombinogenic activity by their ability to dismantle RAD51 mediated D-Loops thereby preventing excessive recombination (Simandlova et al., 2013; Vannier et al., 2012).

Perspective and Overview of my Research

Importance of RAD51 protein was highlighted by DNA repair defects and embryonic lethality of murine *RAD51* deletions. RAD51 along with its bacterial counterpart RecA, bacteriophage UvsX and archaeal RadA have been subjected to genetic and biochemical scrutiny resulting in a plentitude of mechanistic and functional information on formation, regulation and activities of these recombinases. An important disconnect between the two lines of investigation still exists because the recombinase functions RAD51 are highly regulated through mediator proteins like the BRCA2 recombination mediator, and a host of post translational modifications, namely phosphorylation. The mechanism and biochemical implications of these regulatory processes have not been satisfactorily evaluated *in-vitro*.

Due to the lack of extensive and complete structural information available for the RAD51 protein, it has been largely difficult to speculate and verify the mechanistic basis for the regulation of the RAD51 recombinase. Secondly, RAD51 is an attractive target for development of novel anticancer therapeutics. Of a particular value is its complex with BRCA2 tumor suppressor protein, whose recombination mediator activity is critical for "loading" of RAD51 onto ssDNA coated with ssDNA binding protein RPA. To facilitate rational design of such molecules, collaborating with experts in computational biology (M. Ashley Spies' Lab, University of Iowa, MNPC), we used the

Saccharomyces cerevisiae Rad51 crystal structure (PDB 1SZP) as a template to build a homology model of Human RAD51. The model suggested previously unknown interactions between the RAD51 N-terminal domain and BRC4 peptide, which I was able to validate experimentally using biochemical assays to measure BRC4 peptide binding affinity, extension of ssDNA within RAD51 nucleoprotein filament, and ATP hydrolysis.

Using the knowledge gained from our structural model for the RAD51 recombinase, I developed a novel strategy to understand several key mechanisms for the regulation of RAD51 by phosphorylation. RAD51 is phosphorylated by the cABL tyrosine kinase at tyrosine's 54 & 315. The mechanistic and functional significance of this event is largely disputed. My strategy employed producing RAD51 incorporating an artificial phosphomimetic (pCMF) amino acid using amber suppressor technology, which more accurately represents tyrosine phosphorylation on RAD51. The resulting "phosphorylated" protein was analyzed using biochemical and single molecule assays reconstituting major activities of RAD51. Using the optimized RAD51 expression system, I have successfully dissected the biochemical mechanism of regulation of RAD51 by the c-Abl kinase. The results of my studies strongly correlate with observations made in previous cell based analysis which was missing until now. These studies also allowed me to propose a mechanism of filament nucleation and extension by the RAD51 protein.

By integrating *in-vivo, in-singulo* and *in-silico* approaches, my goal is to provide a coherent description of how post translational modifications affect RAD51 filament formation, dynamics and strand exchange activities and thereby to generate insights into one of the most critical steps in DNA repair and to find an 'Achilles Heel' in RAD51 function and exploit it to develop novel anticancer therapies.

CHAPTER 2: CONTRIBUTIONS OF THE RAD51 N-TERMINAL DOMAIN TO BRCA2-RAD51 INTERACTION

Abstract

RAD51 DNA strand exchange protein catalyzes the central step in homologous recombination, a cellular process fundamentally important for accurate repair of damaged chromosomes, preservation of the genetic integrity, restart of collapsed replication forks and telomere maintenance. BRCA2 protein, a product of the breast cancer susceptibility gene, is a key recombination mediator which interacts with RAD51 and facilitates RAD51 nucleoprotein filament formation on ssDNA generated at the sites of DNA damage. An accurate atomistic level description of this interaction, however, is limited to a partial crystal structure of the RAD51 core fused to BRC4 peptide. Here, by integrating homology modeling and molecular dynamics we generated a structure of the full-length RAD51 in the complex with BRC4 peptide. Our model predicted previously unknown hydrogen bonding patterns involving the N-terminal domain of RAD51. These interactions guide positioning of the BRC4 peptide within a cavity between the core and the N-terminal domains; the peptide binding separates the two domains and restricts internal dynamics of RAD51 protomers. The model's depiction of the RAD51-BRC4 complex was validated by free energy calculations and *in vitro* functional analysis of rationally designed mutants. All generated mutants, $RAD51^{E42A}$, $RAD51^{E59A}$ and $RAD51^{E237A}$ maintained basic biochemical activities of the wild type RAD51, but displayed reduced affinities for the BRC4 peptide. Strong correlation between the calculated and experimental binding energies confirmed the predicted structure of the RAD51-BRC4 complex and highlighted the importance of RAD51 N-terminal domain in RAD51-BRCA2 interaction.

This chapter appeared in its entirety in the journal Nucleic Acids Research and is referred to later in this dissertation as "Subramanyam et al., 2013". Subramanyam S, Jones WT, Spies M, & Spies MA (2013) Contributions of the RAD51 N-terminal domain to BRCA2-RAD51 interaction. Nucleic acids research 41(19):9020-9032. This article is reprinted with the permission of the publisher and is available from http://nar.oxfordjournals.org using DOI: 10.1093/nar/gkt691.

Introduction

Human RAD51 and BRCA2 proteins are the key contributors to genomic integrity and are of a paramount importance to the cell. The intricate choreography of molecular events orchestrated by these two proteins ensures accurate and timely progression of homologous recombination (HR) (Amunugama et al., 2012a; Holloman, 2011) and has an important additional function at the replication fork (reviewed in (Costanzo, 2011)). HR repairs genotoxic DNA lesions via precisely coordinated DNA transactions that lead to exchange of information between two homologous DNA molecules (Li and Heyer, 2008b) and play a prominent role in faithful duplication of the genome (Li and Heyer, 2008b) and telomere maintenance (Oganesian and Karlseder, 2011). DNA Repair by HR depends on assembly of the RAD51 recombinase into a continuous nucleoprotein filament on ssDNA generated at the site of damage or collapsed replication (Ciccia and Elledge, 2010; Holthausen et al., 2011; Li and Heyer, 2008b; Moynahan and Jasin, 2010). Once formed, the nucleoprotein filament sequesters template dsDNA, searches for homology and exchanges DNA strands, forming joints between recombining molecules. Assembly of the RAD51 filament is tightly regulated and requires assistance from a recombination mediator. The 3148 amino acid (aa) BRCA2 is a tumor suppressor protein which mediates HR by recruiting RAD51 to sites of the DNA double strand breaks (DSBs) and facilitates loading of the RAD51 protein onto resected single stranded DNA (ssDNA) coated with RPA ssDNA binding protein. Once formed, the pre-synaptic RAD51 nucleoprotein filament promotes homology search and the DNA strand exchange reaction (Holloman, 2011; Jensen et al., 2010; Liu et al., 2010a; Thorslund et al., 2010a). Mutations in human BRCA2 protein predispose to breast and ovarian cancers, and increase susceptibility to other tumorigenic conditions (Nathanson et al., 2001; Turner et al., 2004).

BRCA2 interacts with RAD51 through a series of eight motifs called BRC repeats and a separate binding site located in the C-terminal region (Chen et al., 1998; Heyer et al., 2010; Holloman, 2011; Moynahan and Jasin, 2010). The highly conserved BRC repeats, approximately 35 amino acids in length, are variably spaced in a segment of the protein encoded in exon 11 of the *BRCA2* gene (Bignell et al., 1997; Bork et al., 1996). Although all BRC repeats of BRCA2 promote RAD51 nucleoprotein formation, the modes of their interaction are split among two classes based upon their affinities for RAD51 versus the RAD51−ssDNA filament (Carreira and

Kowalczykowski, 2011). Among eight BRC repeats involved in RAD51 binding, BRC4 displays the highest affinity for RAD51 (Carreira and Kowalczykowski, 2011).

The interaction between BRC peptides and RAD51 family proteins likely represents a universal means of regulating (positively or negatively) the recombinase assembly and activity. A variant of BRC motif recently found in RECQL5 helicase plays a critical role in its anti-recombinogenic activity (Islam et al., 2012), whereas loading of bacterial RecA recombinase on the ssDNA by RecBCD helicase/nuclease involves interaction with the structural elements on the RecA/RAD51 core also critical for BRC4-RAD51 interaction (Spies and Kowalczykowski, 2006). While the interaction between BRCA2 and RAD51 is of a paramount importance, its structural understanding is limited to the crystal structure of the core domain of RAD51 fused to the BRC4 peptide (Pellegrini et al., 2002). The structure features two areas of hydrophobic interactions involving F1524 and F1546 of the BRCA2 and the core of RAD51 and has been exploited in design of the peptide inhibitors of BRC4-RAD51 interaction (Nomme et al.) and, more recently, in the identification of low-molecular-weight fragments that display mM affinity with a goal of utilizing them in the fragment-based approach (Scott et al., 2013). As only the core of RAD51 is present in the structure, potential involvement of the missing N-terminal domain (NTD) is unclear.

Here, we combined homology modeling and molecular dynamics (MD) simulations to build an accurate atomistic description of the full length RAD51 protein in complex with the BRC4 peptide. Our computational studies predicted previously unknown interactions between BRC4 peptide and the RAD51 N-terminal domain. Moreover, the structure and the position of the NTD differed significantly from that of the yeast Rad51 (Conway et al., 2004) used as a template for our model. The model's rendition of the complex was validated by *in silico* binding studies and free energy calculations, and *in vitro* by the functional analysis of mutants designed based on the results of the computational studies.

Results

RAD51 Homology Model and BRC4 peptide placement

Saccharomyces cerevisiae Rad51 structure (PDB ID: 1SZP) (Conway et al., 2004) was used as a template to build a homology model of two adjacent protomers within human RAD51 filament (Figure 2.1).

Following all atom MD simulations with the knowledge-based force field, YASARA, approximately 60% of the RAD51 model (chain B) overlapped with the yeast Rad51 (1SZP-B) with the average RMSD between Cα atoms under 1.23 Å. Several parts of the structure, however, deviated significantly between the yeast and human proteins (Figure 2.2) and therefore were not recognized by the alignment algorithm, MUSTANG, which was used for the structure comparison and has a cutoff of 3.75 Å for the structurally similar residues to be matched (Konagurthu et al., 2006). The most profound difference between the two structures involved the linker region between the NTD and the conserved core domain (Figures 2.2A and 2.3A), which in the yeast Rad51 consists of two alpha helices connected by a flexible loop, while the N-terminal domain of human RAD51 is connected to the core by a long rigid helix. Although the overall 4 helix bundle structure of the N-terminal domain is preserved in the human protein, two of these helices display a slightly different orientation. Note that these differences in the NTD and the linker region affect mainly chain B of RAD51 model whose NTD is located near the protomer-protomer interface. Structural change in the linker region and NTD also affected the overall orientation of the N terminal domain resulting in the wider cavity between the NTD and the core, which in RAD51 accommodates BRC4 peptide after a relatively small conformational change (Figure 2.1C-F; Figure 2.2C; Figure 2.3C shows structural overlap between the peptide-free and the peptide bound RAD51). Notably, we observed a remarkable overlap between our model and the crystal structure of the core RAD51 (1N0W), fused to BRC4 peptide of BRCA2 (Pellegrini et al., 2002), which was not used in the model construction. Most of the structure (1N0W) overlapped with the homology model to yield an RMSD under 1Å, with exception of flexible areas around the DNA binding loops, and several residues preceding the Walker A box. The overall RMSD for Cα atoms was 1.22 Å (Figures 2.2B and 2.3B). This provides valuable orthogonal information on the quality of the model as 1N0W was not used to construct the homology model. The convergence between our model and 1N0W RAD51 core structure is especially remarkable since these two structures

share higher similarity than that observed between our modeled RAD51 and the structure of the yeast protein, which was used as a template to build the model.

Our RAD51 model scored very well on a number of physical metrics. The Z-score, which evaluates the homology model for a set of physical parameters such as 3D packing, bond lengths, bond angles, etc (Krieger et al., 2009), after MD was -0.89, which means that the normality of 3D- and 1D-packing, as well as Ramachandran coordinates are less than 1 standard deviation from the gold standards from high-resolution crystal structures (Krieger et al., 2009). Many other checks were also performed to ensure normality, including normality of van der Waals and Columbic energies = 0.557; planarity of peptide bonds = -0.215; normality of dihedral bonds = 0.08; normality of bond angles = 0.291; normality of bond lengths = 0.48; normality of water positions = -0.392; isomers = zero wrong isomers.

The placement of the BRC4 peptide in the RAD51 model was a critical step as the previously published crystal structure of RAD51 core in complex with BRC4 peptide had the peptide fused to the RAD51 core domain via a flexible linker (Pellegrini et al., 2002). This posed some uncertainty with regard to the native complex in solution, and how it should be represented in MD simulations. Furthermore, absence of the NTD of RAD51 from the crystal structure allowed BRC4 peptide to partially occupy the spatial domain that should be occluded by the NTD.

First, we placed the peptide in the chain B. Due to the inherent flexibility of the peptide, an exhaustive conformational search was performed using LowMode MD while keeping the protein coordinates static. LowMode MD is an accelerated MD method allowing configurational searches outside of the usually restrictive classical MD time scales, concentrating the kinetic energy on low frequency vibrational modes (see methods) (Labute, 2010).

Finally, simulated annealing energy minimization and MD was performed with the YASARA knowledge-based force field (KBFF), as described above, (see methods) (Krieger et al., 2009). The second peptide was placed in the chain A by structural superposition of the subunit A of the homology model and the peptide, followed by simulated annealing energy minimization, as described in the MD methods section. Thus we obtained a system of two monomers, each with a bound peptide. The motivation for placing the second peptide was to provide a greater degree of structural similarity to a multimeric system. Only the first peptide (placed in the chain B), which

interfaces with a structurally complete binding site was used in the free energy calculations and pocket analysis, and yielded insights that guided the experimental part of this study. Notably, the overall structure and position of this BRC4 peptide in the homology model was remarkably similar to that observed in the crystal structure (Figure 2.3B). This was primarily achieved through a slight movement of the RAD51 NTD relative to the peptide free model (Figures 2.2C and 2.3C).

Model of the RAD51-BRC4 complex predicts that both the core and NTD of RAD51 participate in the peptide positioning

Some structural rearrangements within RAD51 model were necessary to accommodate the BRC4 peptide (Figure 2.1C-F, Figure 2.2C & 2.3C), which primarily involved movement of the NTD and the DNA binding loops. The average RMSD between Cα atoms of the RAD51 core in the absence and presence of the peptide were 2 Å, while the average RMSD between Cα atoms of the NTDs were 4.9 Å. The NTD and the RAD51 core formed a cleft, which accommodates the C-terminal part of the BRC4 peptide (Figure 2.4A).

Several previously unknown interactions were predicted between RAD51 and the BRC4 peptide (Figure 2.4A). Notably, while most of these interactions involved the RAD51 monomer to which the peptide is placed (chain B), the model reveals additional contacts between BRC4 peptide and the adjacent monomer (chain A). The carboxylate of E237 defines the turn between the dsDNA binding Loop 1 and the N-terminus of α-helix 5 (in the nomenclature from (Pellegrini et al., 2002)), by receiving a hydrogen bond from the backbone amide of S233. This feature results in the β-hydroxyl of S233 pointing into the peptide binding pocket and acting as a hydrogen bond donor for the backbone carbonyl of E1548 and a hydrogen bond recipient of the backbone amide of this same residue. Simulations of the E237A mutant showed that the turn region of the adjacent monomer became less defined than the wild-type, resulting in the S233 group pointing away from the peptide binding cavity, and thus contributing to a decrease in the peptide binding free energy.

The N-terminal domain of RAD51 participates in the interaction as well: the carboxylate of E59 forms hydrogen bonds with the amide of N1544 and the amide of K1549 on the BRC4 peptide; the amide of K1543 bonds with the backbone carbonyls of A44 and E42. Within the same region, E42 interacts through hydrogen bonding with S26 which in turn stabilizes the secondary

structure of the RAD51 N-terminal domain (Figure 2.4A). The average distance between the heteroatoms involved in these hydrogen bonds consistently remained below 3Å over the course of a 3ns MD simulation (Figure 2.4B).

In silico mutagenesis of three key glutamates (E42, E59 and E237) suggested their importance for correct positioning of the BRC4 peptide and predicted a network of interactions that the RAD51 NTD contributes toward the stable binding of the BRC4 peptide to RAD51 protein (Figure 2.4). MD simulations of RAD51 E42A and E237A mutants were used to analyze disruption of the hydrogen bonding interactions mentioned earlier. As expected, the distances between the involved heteroatoms increased markedly after the 4ns MD of the mutants. Simulation results for E42A showed that the only hydrogen bonding partners left for K1543 are L41 and A42 backbone carbonyls, with heteroatom distance of ~ 3.0 A; this is substantially different than the wild-type hydrogen bonding pattern in which every available hydrogen of the ε-nitrogen of K1543 is engaged in hydrogen bonding: 2.8 Å to carbonyl of A44, 2.8 Å to carbonyl of E42 and 3.0 Å to carbonyl of F46. Therefore, both E237A and E42A mutants yield predictable decreases in binding affinities toward the BRC4 peptide and predict a series of collective interactions that the RAD51 NTD contributes toward the stable binding of the BRC4 peptide to RAD51 protein (Figure 2.4A). These observations also suggested the mutant candidates for *in vitro* analysis of the contributions of the RAD51 NTD as well as for validation of our homology model.

RAD51 mutants retained structure and basic biochemical properties of the wild type protein

Based on the predictions from the model, alanine mutants of three residues E42, E59 and E237 were constructed, purified and analyzed to validate the role of these residues in the RAD51-BRC4 complex (Figures 2.5 and 2.6). If our model is sound, these amino acids should display reduced affinities toward BRC4 peptide without perturbing other biochemical activities and properties of RAD51. E237A was used to indirectly study the effect of S233 which is located in the dsDNA binding loop (L1) and may be critical for RAD51 function.

Analysis by circular dichroism spectroscopy confirmed that the three mutants maintained the secondary structure contents characteristic of wild type RAD51 (Figure 2.5C & D).

Formation of the recombination competent RAD51 nucleoprotein filament causes ~1.5 fold ssDNA extension over the B-form. Such filaments contain one RAD51 monomer per 3 nucleotides of ssDNA and can be formed in the presence of ATP and Ca^{2+} ions (Bugreev and Mazin, 2004; Ristic et al., 2005). Both DNA extension and binding stoichiometry are indicators of the active nucleoprotein filament formation. DNA binding and extension activity was measured using a FRET based assay wherein, we observed and quantified the RAD51-mediated extension of a 60-mer ssDNA substrate labeled with the FRET donor (Cy3) and acceptor (Cy5) fluorophores separated by 25 nucleotides (Grimme et al., 2010; Grimme and Spies, 2011; Masuda-Ozawa et al., 2013). Under the stoichiometric binding conditions, RAD51 titration results in gradual FRET decrease due to the spatial separation of the Cy3 and Cy5 fluorophores until the substrate is saturated with RAD51 and no further extension can be achieved. The inflection point in the titration curve reports on the binding stoichiometry, while the amplitude of the FRET decrease reports on the DNA extension (see methods). Similar to the wild type RAD51, all three mutants showed capacity to bind and extend DNA (Figure 2.6A) with the characteristic ~1:3 protein:nucleotide binding stoichiometries.

RAD51 is a DNA-dependent ATPase (Bugreev and Mazin, 2004). ATP hydrolysis plays a role in the nucleoprotein filament dynamics and in particular in the protein turnover (Amunugama et al., 2012b; Bugreev and Mazin, 2004; Tombline and Fishel, 2002). E42A and E59A mutants hydrolyzed ATP similar to wild type RAD51 with k_{cat} values of 0.26±0.01 min^{-1}, 0.27±0.01 min^{-1} and 0.27±0.01 min^{-1} (mean ± standard error) for the E42A, E59A and wild type respectively. E237A mutant had a slightly slower rate of ATP hydrolysis with k_{cat} of 0.13±0.02 min^{-1} (Figure 2.6B). This was still within the range previously reported for human RAD51 (Tombline and Fishel, 2002). Moreover, since the mutant was able to form an extended nucleoprotein filament, E237A substitution likely affects ATP hydrolysis, but not ATP binding resulting in a more stable filament and lower protein turnover.

RAD51 Mutants display reduced affinity for BRC4

Since the three mutants retained structure and biochemical activities of the wild type RAD51, they can be directly compared for their ability to bind BRC4 peptide. Fluorescence Polarization

Anisotropy (FPA) was used to characterize binding of RAD51 mutants to the BRC4 peptide. Protein was incrementally titrated into buffer containing FITC-labeled BRC4 peptide. Increase in FPA reflected RAD51-BRC4 complex formation (LiCata and Wowor, 2008). In the presence of ATP, RAD51 bound BRC4 peptide with K_d of 33.6 ± 6.9nM. As expected, all alanine mutants displayed reduced affinity for BRC4 peptide and bound with K_d values of 174 ± 31nM, 959 ± 274nM and 560 ± 131nM for E42A, E59A and E237A mutants respectively (Figure 2.7). The magnitude of the change in the FPA signal reflects the size of the complex containing fluorescence peptide. The E59A and E237A mutants displayed higher FPA changes compared to the wild type RAD51 and the E42 mutant. This is likely due to binding of higher RAD51 oligomeric species to BRC4 and highlights importance of the contacts between the C-terminal end of BRC peptide and both adjacent monomers.

Correlation between *in silico* and *in vitro* results reaffirm the role of RAD51 N-terminal domain

Endpoint free energy calculations were used to determine the relative $\Delta G_{binding}$ for peptide binding to RAD51, and were measured over the final 3 ns of the MD simulations, after an initial 1 ns pre-equilibration period (see Methods section). These binding endpoint free energy calculations exclude a number of terms such as ligand and receptor binding entropies and the nonpolar interactions with solvent (i.e., cavitation and van der Waals interaction with solvent). Furthermore, the use of an implicit solvent model to calculate the changes in the solvation energies necessitates the use of a uniform dielectric constant, which strongly affects the magnitude of the calculated binding energy. However, such endpoint free energy calculations have been useful in accessing the relative changes in binding free energy (Brown and Muchmore, 2009; Shirts et al., 2010; Steinbrecher and Labahn, 2010).

The ΔG_{bind} trajectories for each protein (Figure 2.8A) were analyzed and converted into frequency histograms to yield the distributions of ΔG_{bind} values over the course of MD simulation (Figure 2.8B). These representations allow following small, but discernible differences in the computed binding energies. Relative ΔG_{bind} values derived from these distributions were compared to the respective relative ΔG values calculated from experimentally derived *K*ds. All three mutants,

E42A, E59A and E237A, showed average changes in ΔG_{bind} rank-ordered similarly to the experimentally derived values. The changes in computational binding energies correlate extremely well with free energy changes obtained empirically with an R value of 0.964 (Figure 2.9). An agreement between relative changes in free energy of binding validates the protein-peptide interface of the homology model and the contributions of the RAD51 NTD toward BRC4 peptide binding.

Discussion

Proteins and enzymes orchestrating key steps in DNA repair are emerging as promising new targets in anticancer drug discovery, with RAD51 being one of the most attractive targets. Crucial for maintenance of genomic integrity in normal cells, RAD51 allows the transformed or cancerous cells to develop resistance to radiation and DNA-damaging drugs used in chemotherapy. Elevated levels of RAD51 lead to rapid accumulation of genetic variation, genomic instability, acquisition of invasiveness, drug and radiation resistance and disease progression in many cancers including Barrett's adenocarcenoma (Pal et al.), multiple myeloma (Shammas et al., 2009), recurrence of chronic myeloid leukemia (Slupianek et al., 2011), high grade gliomas (Short et al., 2011), and lung cancer (Qiao et al., 2005). Targeting RAD51 may therefore allow chemo- and radio-sensitization of cancerous cells as an adjuvant in standard combination anticancer regimens (Huang et al., 2012).

Several recent HTS campaigns and rational design of inhibitors/effectors of RAD51 were only mildly successful yielding an inhibitor of unknown mode of action with IC_{50} of 27.4 μM and poor drug-like properties (*log*P = 5 and low ligand efficiency of -0.23 kcal/heavy atom) (Huang et al., 2011); a small molecule that mildly stimulated RAD51-mediated strand exchange activity (Jayathilaka et al., 2008); DNA aptamers (Martinez et al.); peptide inhibitors (Nomme et al.) and most recently, an inhibitor of RAD51 filament formation that covalently binds to the RAD51 (Budke et al., 2012).

Whereas the overall extended filamentous structure and DNA strand exchange function is highly conserved within the RecA/RAD51 family of recombinases, the important features of their structures and mechanisms differ between species. Understanding of the RAD51 structure within

the dynamic nucleoprotein filament may greatly facilitate rational discovery of small-molecule scaffolds, which can be developed into potentially effective anticancer treatments and highly specific molecular probes. It also will improve our understanding of the mechanisms by which the recombination mediators and anti-recombinases affect RAD51 interaction with ssDNA and dsDNA, nucleoprotein filament assembly, disassembly and dynamics.

The existing high resolution structure of the RAD51 core fused to the BRC4 peptide of BRCA2 tumor suppressor protein identified several key contacts between the two proteins, suggested the determinants of BRC4 peptide affinity for RAD51 and the mechanism by which it may affect the nucleoprotein filament (Pellegrini et al., 2002). The identified contact areas on the RAD51 core surface, however, display poor druggability due to a featureless interaction surface (Surade and Blundell, 2012) with only small pockets that bind phenylalanine (Scott et al., 2013). In contrast, the interface between the two adjacent RAD51 monomers within the filament and the cleft between the NTD and the RAD51 core may contain numerous loci that can be targeted.

Here, we report a model of RAD51, which includes both the NTD and the conserved core. Four ns MD simulations with knowledge based force field were employed to relax the RAD51 model to a stable form, whose core was similar to the structure of yeast Rad51 gain of function mutant (PDB ID: 1SZP) used as a template. More remarkably, the post-MD model converged with the structure of the RAD51 core (PDB ID: 1N0W) which was not used in the model building.

The presence of the two monomers in the model allowed us to realistically represent the monomer-monomer interface within the RAD51 filament as well as the position of the NTD of one of the monomers (chain B). Our model also predicted interactions between BRC4 peptide and the two RAD51 monomers adjacent in the filament. It, therefore, represents the initial stage of the RAD51-BRCA2 complex formation where the interface between adjacent monomers of RAD51 is slightly perturbed, but not yet completely abolished and replaced by the interactions with N-terminal part of the BRC4 peptide. While the BRC4 peptides were placed in both monomers, contacts made by only one peptide placed in the chain B and contacting the monomer-monomer interface were evaluated. The backbone of the BRC4 peptide in our model overlapped well with the peptide in the 1N0W structure; the orientations of several side chains, however, were different reflecting their interactions with the NTD and the adjacent monomer. This was somewhat expected since the position of the peptide in the crystal structure might have been constrained by fusion of

the peptide to the RAD51 core and by extraordinary dense crystal packing. In our model, on the other hand, the NTD partially overlapped with the space occupied by the peptide in the crystal structure. Modeling of the RAD51-BRC4 complex suggested that the largest deviation between the RAD51 structures with and without the BRC4 peptide will involve movement of the NTD (Figure 2.1C&D, Figures 2.2C and 2.3C).

To confirm the validity of the model we identified several previously unknown contacts between the BRC4 peptide and RAD51. In particular, we focused on the contacts that involved the NTD or the adjacent RAD51 monomer. The binding energy calculations of the BRC4 complexes with the 3 RAD51 glutamate to alanine mutants carried out over 4 ns MD simulations suggested that the three following mutants E42A, E59A, and E237A should display compromised affinities for the BRC4 peptide, but retain the biochemical properties of the wild type RAD51. Indeed the three purified mutants had lower than the wild type affinity for the BRC4 peptide with the rank order predicted by the binding energy calculations. While the contributions from each of the identified residues to the overall energy of the RAD51-BRC4 complex formation are relatively modest, as expected for the hydrogen bond disruptions in a protein-peptide contact surface, collectively they highlight the importance of the NTD of RAD51 in the BRC4 positioning and suggest the mechanism by which the interaction between BRC4 peptide and the adjacent monomer within the RAD51 oligomer contributes to the RAD51 oligomer destabilization and selectivity for ssDNA.

We envision that the interaction between K1543 of the BRC4 and E42 and A44 of the NTD (Figure 2.4) guide the peptide into the cleft between NTD and the RAD51 core and simultaneously shift the position of the NTD (Figure 2.1 C & D, Figures 2.2C and 2.3C). This allows placement of the F1546 of the peptide in the hydrophobic pocket within the RAD51 core revealed by the crystal structure. Two hydrogen bonds accepted by E59 from N1544 and K1549 further stabilize the peptide in the cleft between the NTD and the RAD51 core. The peptide then interacts with the S233, which belongs to the Loop1 (dsDNA binding loop) of the adjacent RAD51 monomer. This interaction may constrain the dsDNA binding and at least in part be responsible for enforcing selectivity of the RAD51-BRCA2 for ssDNA over dsDNA. Together with the interaction between E1548 of the BRC4 and R250 observed both in the 1N0W structure and in our model, the S233-

BRC4 interaction may also affect relative orientation of the two adjacent RAD51 monomers and thereby destabilize the monomer-monomer interface.

Another notable interaction is the hydrogen bond between Q1551 of the peptide and P56 within the NTD of RAD51. This interaction positions the C-terminus of the BRC4 peptide near the N terminus of the peptide docked in the second monomer.

The difference between the yeast and human RAD51 proteins in the position and structure of the NTD provides a glimpse into co-evolution of the recombinase and recombination mediator. The more rigid connector between the two domains of human RAD51 shifts the position of NTD relative to the core and creates the binding site that has enough flexibility to accommodate diverse BRC peptides.

Materials and Methods

RAD51 Homology Model

The crystal structure of yeast Rad51 gain of function mutant (PDB ID: 1SZP, (Conway et al., 2004)) representing an active conformation of the Rad51 filament was used to build a homology model of two RAD51 monomers adjacent in the nucleoprotein filament. The model was constructed using The Chemical Computing Group's Molecular Operating Environment (MOE) (Molecular Operating Environment (MOE) 2010.09, 2012). Ten intermediate homology models resulting from permutational selection of different loop candidates and side chain rotamers were built for RAD51, each subjected to a degree of energy minimization using the force field MMFF94x, with a distance-dependent dielectric. The model of the monomer (chain B) was constructed in the presence of the adjacent monomer in the 1SZP-A template, in order to optimize the monomer-monomer interface. The intermediate model which scored best according to the packing evaluation function was chosen for the next level of refinement: the RAD51 dimer was constructed by superposition of the RAD51 homology modeled monomer onto the 1SZP dimer crystal structure, followed by simulated annealing energy minimization and 4 ns MD simulations with the knowledge-based YASARA force field (see below).

The Mg-ATP substrate was placed by first building and then docking with flexible ligand docking into a region that corresponds to the canonical ATP binding site of the *E. coli* RecA protein (PDB ID: 1XMS, (Xing and Bell, 2004)). The AutoDock implementation in YASARA Structure was employed. AutoDock 4 employs a Lamarckian genetic algorithm to sample ligand conformations and binding modes. It uses a semi-empirical free energy force field to predict free energies of binding which accounts for intermolecular and intramolecular energies, as well as charge-based desolvation. The following general docking parameters were used: 25 independent docking runs, each with a total of 2.5 x 106 energy evaluations, a torsional degrees of freedom value of 8, grid point spacing was left at the default of 0.375 Å, and the force field selected was AMBER03. Specific to the genetic algorithm, the following parameters were used: a population size of 150, 2.7 x 104 generations, an elitism value of 1, a mutation rate of 0.02, and a crossover rate of 0.8. Final poses were considered distinct if they varied by > 5 Å RMSD. All atom energy minimization was then performed on the docked structure. This represented a starting point for MD simulations, and no constraints were placed on the Mg-ATP.

Placement and Conformational Search of BRCA4 Peptide

The partial structure of human Rad51 with BRACA4 peptide (PDB 1N0W) was superposed onto the dimer Rad51 homology model using the superpose utility of MOE, in order to initially place the BRACA4 peptide at the interface of the two monomers. The complex was then subjected to a specialized stochastic conformational search protocol called LowModeMD (Labute, 2010) within the MOE package. This method concentrates kinetic energy on low-frequency vibrational modes, in order to populate conformations in multiple low-energy states with high computational efficiency, and is particularly appropriate for complex systems with large numbers of nonbonded interactions, such as peptides, peptide loops and macrocyles. The LowModeMD conformational search procedure includes an iterative process of initial energy minimization, filtering of high frequency vibrational modes, a short (~0.5 ps) MD and saving distinct structures in a database. The energy minimization gradient threshold was 0.001 kcal/mol/Å, and searches were configured to terminate after 200 contiguous failed attempts to generate novel conformations, with up to 10,000 iterations. Conformations were identified as unique if their root-mean-square-distance was

above a threshold value of 0.75 Å. RAD51 protein remained frozen throughout the LowModeMD conformational search, and a single low-energy peptide conformation was identified.

Classical MD Simulations with two Adjacent RAD51 Monomers

The molecular dynamics simulations on the homology model of RAD51 containing ATP and in the presence and absence of BRC4 peptide were performed with the YASARA Structure package version 12.4.1 (YASARA Biosciences) (Krieger). A periodic simulation cell with dimensions 104.94 Å, 73.68 Å, and 78.23 Å was used with explicit solvent. The YASARA KBFF was used with long-range electrostatic potentials calculated with the Particle Mesh Ewald (PME) method (Darden et al., 1993; Essmann et al., 1995), with a van der Waals cutoff of 7.864 Å. This force field has been highly successful for use with homology modeling and protein structure prediction, in that it limits the damage (i.e. drifting into structurally unrealistic protein phase space) that often results from energy minimization and MD using empirical force fields.

ATP force field parameters were generated with the AutoSMILES utility (Jakalian et al., 2002) which employs semi-empirical AM1geometry optimization and assignment of charges, followed by assignment of AM1BCC atom and bond types with refinement using RESP charges, and finally the assignments of general AMBER force field atom types. The hydrogen bond network of RAD51 was optimized using the method of Hooft and coworkers (Hooft et al., 1996), in order to address ambiguities from multiple side chain conformations and protonation states that are not resolved by the electron density of the template. YASARA's pKa utility was used to assign pKa values at pH 7.0 (Krieger et al., 2006). The box was filled with water, with a maximum sum of all bumps per water of 1.0 Å, and a density of 0.997 g/ml. The simulation cell was neutralized with NaCl (0.9% w/v final concentration). Excessive water molecules were deleted to readjust the solvent density to 0.997 g/ml. A short MD was run on the solvent only. The entire system was then energy minimized using first a steepest descent minimization to remove conformational stress, followed by a simulated annealing minimization until convergence (<0.05 kJ/mol/200 steps). The MD simulation was then initiated, using the NVT ensemble at 298 K, and integration time steps for intramolecular and intermolecular forces were calculated every 1.25 fs and 2.5 fs,

respectively. This procedure was conducted after each *in silico* mutation as well. The structural alignments were performed with the MUSTANG method (Konagurthu et al., 2006).

Free Energy Binding Calculations with the Fast Boundary Element Method (BEM)

The method employed here is called the Boundary Element Method (BEM) (Juffer et al., 1991; Zauhar et al., 1985). It falls under the class of free energy calculations known as Endpoint Methods, which also includes the popular MM-PBSA approach (also called the finite difference method) (Steinbrecher and Labahn, 2010). Endpoint Methods calculate the ΔG_{bind} from constituent parts of a thermodynamic cycle that involve solvation of the individual components. The binding energy expression is:

$$\Delta G_{Bind,Solv} = \Delta G_{Complex,Vaccum} + \Delta G_{Complex,Solv} - (\Delta G_{Ligand,Solv} + \Delta G_{Receptor,solv}) + \Delta G_{np}$$

In the BEM much of the focus is placed on accurately representing the boundary between the two dielectrics, in which a very accurate boundary charge distribution is used to represent a uniform dielectric at the interface between the low and the high dielectric continuum. From this boundary region of uniform dielectric strength, Coulomb's Law is used to calculate the electrostatic potentials.

A major difficulty in Endpoint methods is assigning an internal dielectric (Schutz and Warshel, 2001). Since the BRC-peptide binding cavity of RAD51 is relatively solvated, we chose to employ a protein dielectric ε of 13; values of 4 to 20 are routinely employed, often using mixed values (Ravindranathan et al., 2011; Schutz and Warshel, 2001). It is important to note that ΔG_{bind} values obtained from Endpoint methods, such as MM-PBSA or BEM, should be viewed as enhanced scoring functions, which have enhanced rank-ordering value, rather than as metrics of accurate absolute binding free energy (Sotriffer and Matter, 2011). For the current study, using BEM, the boundary between solvent (dielectric constant 78) and the solute (dielectric constant 13) was formed by the latter's molecular surface, constructed with a solvent probe radius of 1.4 Å and the following radii for the solute elements: polar hydrogens 0.32 Å, other hydrogens 1.017 Å, carbon 1.8 Å, oxygen 1.344 Å, nitrogen 1.14 Å, sulfur 2.0 Å. The solute charges were assigned based on the AMBER03 force field (Cornell et al., 1995). The term for the hydrophobic component

of peptide binding, ΔG_{np} was not included in these calculations, since this value is not expected to change in the mutated complexes being considered. The peptide binding entropy was not included in the relative binding energy calculations, and is not expected to significantly contribute to this value. After a 1 ns equilibration period, BEM ΔG_{bind} value was calculated every 7 ps, for duration of 3 ns. The ensembles of the endpoint peptide binding free energy values were obtained for the wild type RAD51, E42A, E59A and E237A mutants (Figure 2.8A), and transformed into the histograms (Figure 2.8B) using GraphPad Prism 4.

RAD51 Protein Expression and Purification

RAD51 protein was expressed in *E. coli* Acella™ strain in the presence of pLysSRARE and pChaperone (generous gift from Dr. Alex Mazin, Drexel University) plasmids in LB medium containing Carbenicillin (50µg/ml), Kanamycin (40µg/ml) and Chloramphenicol (34µg/ml). The cells were grown at 37˚C. After OD_{600} reached 0.6 RAD51 expression was induced with 0.1mM IPTG (Calbiochem). Induced cells were further incubated at 37˚C for 4 hours, pelleted by centrifugation and lysed by sonication in the lysis buffer containing 100mM Tris-OAc (pH 7.5), 2mM EDTA, 10% Glycerol, 1mM DTT, Lysozyme (0.5mg/ml), 0.1% Triton X-100 and Complete, mini, EDTA Free Protease Inhibitor Tablets (Roche). The clarified lysate was then dialyzed overnight against 3 changes of 0.5 liters of 20mM Tris-OAc pH 7.5, 7mM Spermidine, 10% Glycerol, 0.1mM DTT. The precipitate was collected by centrifugation, re-suspended in T-75 buffer (50mM Tris-HCl pH 7.5, 10% Glycerol, 75mM NaCl, 0.1mM DTT) and then centrifuged again. The Pellet and supernatant were collected and the process was repeated by re-suspending the pellet in T-150, T-250, T-500 and T-600 buffers (containing 150mM, 250mM, 500mM and 600mM NaCl, respectively). Fractions containing RAD51 were pooled and loaded onto HiTrap Blue column equilibrated with Buffer BA (100mM Potassium Phosphate pH 7.0, 5% Glycerol, 300mM NaCl, 1mM EDTA, 1mM DTT). Protein was eluted using a 0 – 2M NaSCN gradient and dialyzed overnight in Buffer HA (20mM HEPES pH 7.5, 5% Glycerol, 150mM NaCl, 1mM EDTA, 1mM DTT). RAD51 containing fractions were then loaded onto a Heparin column and protein eluted with a 150mM – 2M NaCl gradient followed by overnight dialysis in Buffer HA. Finally, the RAD51-containing fractions were concentrated on a MonoQ column using a steep 150mM – 1.2M NaCl gradient elution. The purified RAD51 was then dialyzed overnight in

modified Buffer HA (0.1mM EDTA), aliquoted and stored at -80˚C. RAD51 concentration was determined using absorption at A_{280} with an extinction coefficient of 12800$M^{-1}cm^{-1}$ (Baumann et al., 1997). Percentage of glycerol in buffers could be varied between 5% and 10%. Human RPA was purified as previously described (Henricksen et al., 1994).

RAD51 E42A, E59A and E237A mutants were produced using the QuikChange II XL site-directed mutagenesis kit (Agilent) using the following oligonucleotide primers (Supplementary Table S1). All mutants were purified using the protocol described for the wild type RAD51.

Fluorescence Polarization Anisotropy based RAD51-BRC4 Binding Assay

Binding affinity of RAD51 for the BRC4 peptide was measured by following fluorescence polarization anisotropy (FPA) of FITC labeled BRC4 peptide (LiCata and Wowor, 2008). Fifteen nM of FITC-BRC4 peptide (FITC-KEPTLLGFHTASGKKVKIAKESLDKVKNLFDEKEQ) (Carreira et al., 2010) was incubated at 37˚C in the reaction buffer containing (20mM HEPES pH 7.0, 2mM $CaCl_2$, 10mM $MgCl_2$, 1mM DTT). FITC fluorescence was measured using excitation and emission wavelengths of 490 nm and 518 nm, respectively using Cary Eclipse Fluorimeter. Experiments were carried out in the presence of 1mM ATP. Dissociation constant (K_d) was measured by fitting the data to the binding isotherm –

$$\Delta A = Ao + \frac{\Delta Amax\left\{(Kd + [RAD51] + [BRC4]) - \sqrt{(Kd + [RAD51] + [BRC4])^2 - 4[RAD51][BRC4]}\right\}}{2[BRC4]}$$

Where ΔA is change in anisotropy, A_o is initial anisotropy in the absence of RAD51 and ΔA_{max} is maximum change in anisotropy, [RAD51] is the total RAD51 concentration at each point of the titration and [BRC4] is the total peptide concentration. GraphPad Prism 4 software was used for data analysis.

Circular Dichroism Spectroscopy

The CD spectra of 3µM of RAD51 was recorded between 190 nm and 260 nm at room temperature in 50 mM Potassium Borate buffer pH 8.0 using J-750 Spectropolarimeter (JASCO). The

secondary structure contents for the wild type and mutant RAD51 proteins (α-helix:β-strand:Random coil ratios) were determined using the DichroWeb K2D program (Whitmore and Wallace, 2004).

RAD51 ATPase Assay

ATP hydrolysis activity of RAD51 was measured using coupled reactions assay (Kowalczykowski and Krupp, 1987). Briefly, 5μM of RAD51 was incubated with 30μM (nucleotides) Poly dT_{100} in reaction buffer (20mM HEPES pH 7.5, 6mM $MgCl_2$, 2mM ATP, 1mM DTT, 7.5mM PEP, 0.2mg/mL NADH, 3.6-6.0U/mL pyruvate kinase, 5.4-8.4U/mL lactate dehydrogenase) at 37˚C; oxidation of NADH was measured by recording decrease in absorbance at A_{340}. The slope of the absorbance change was then converted into the rate of ATP hydrolysis in μM ATP hydrolyzed per min per μM RAD51.

DNA Binding and Extension Assay

The ability of RAD51 to bind and extend ssDNA was measured using the FRET based assay described previously (Grimme and Spies, 2011). Six hundred nM (nucleotides) of dT_{60} oligo labeled with the Cy3 and Cy5 dyes separated by 25 nt was titrated with RAD51 protein in FRET Reaction Buffer (20mM HEPES pH 7.5, 5mM $CaCl_2$, 5mM $MgCl_2$, 1mM ATP, 1mM DTT) at 37˚C. Cy3 and Cy5 fluorescence was recorded using Cary Eclipse Fluorimeter. The FRET efficiency was calculated as a fraction of acceptor intensity relative to the total donor and acceptor intensity adjusted by correction factors as described in (Grimme and Spies, 2011).

RAD51 ability to extend DNA beyond the contour length was evaluated by comparing to the DNA extension by human RPA.

Images

Molecular graphics images were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081; (Pettersen et al., 2004)).

Figures

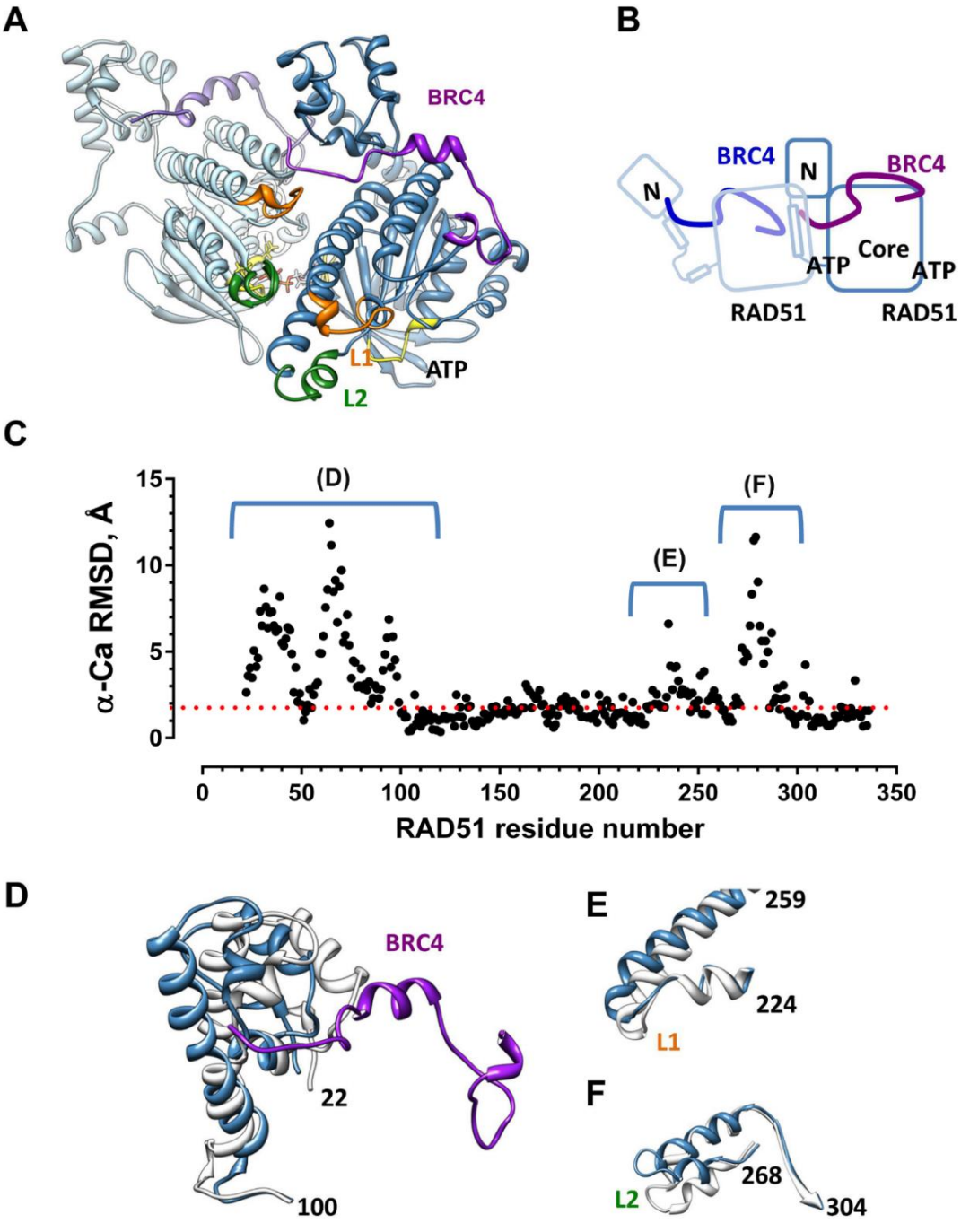


Figure 2.1. RAD51-BRC4 model and BRC4-induced structural rearrangements in RAD51 filament.

Figure 2.1. RAD51-BRC4 model and BRC4-induced structural rearrangements in RAD51 filament. **A.** Ribbon diagram of the homology model of two adjacent RAD51 protomers containing BRC4 peptides: Chain B is shown in a dark blue; chain A is in a lighter blue; the two peptides are shown in purple. The RAD51 DNA binding loops L1 and L2 are shown in orange and green, respectively, and the ATP binding region is shown in yellow. **B.** Schematic representation of the BRC4 peptide bound RAD51. **C.** α −Carbon RMSDs for the peptide-free and BRC4-bound RAD51 (chain B) as a function of the residue number. Letters D-F above the graph indicate the most divergent structural elements. The respective structural overlaps are shown in the panels **D-F: D.** NTD and the linker; **E**. dsDNA binding Loop L1; **F.** ssDNA binding loop L2.

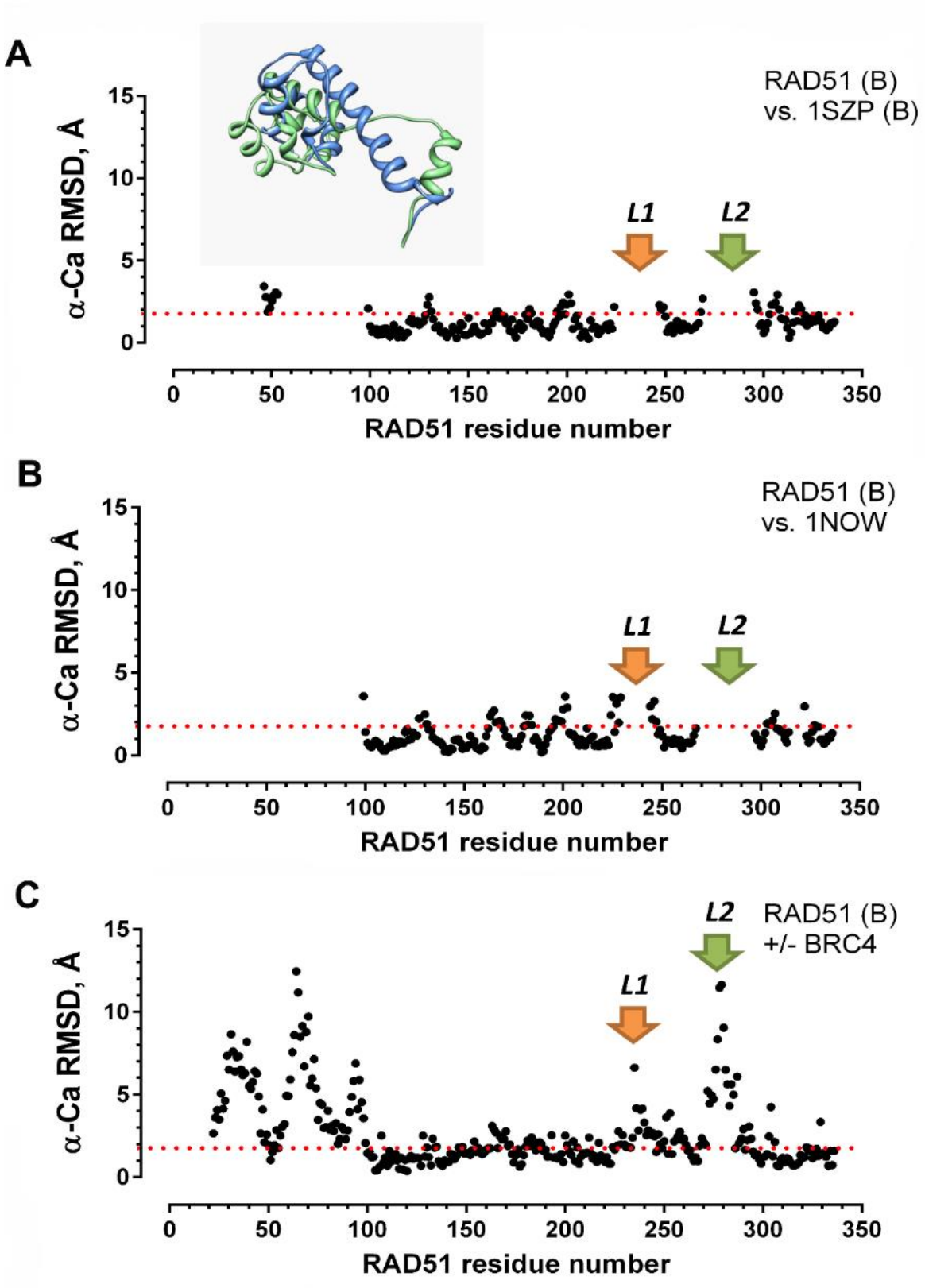


Figure 2.2. Structural comparison of the RAD51 homology model and known RAD51 and Rad51 structures.

Figure 2.2. Structural comparison of the RAD51 homology model and known RAD51 and Rad51 structures. **A.** $\alpha-$Cα RMSDs derived from the structural overlap between RAD51 homology model (Chain B, no peptide) and yeast Rad51 (1SZP, Chain B). $\alpha-$Cα RMSDs for the most of the residues within the NTD were larger than the 3.75Å, and therefore could not be paired by the alignment algorithm (MUSTANG) used to compare the two structures. The inset shows overlap between the NTD of RAD51 (blue) and the NTD of Rad51 (green). **B.** $\alpha-$Cα RMSDs derived from the structural overlap between RAD51 homology model (chain B with BRC4 peptide) and the RAD51 core crystal structure (1NOW). Residues of the NTD and the DNA binding loops were absent from the 1NOW structure and therefore could not be compared. **C.** $\alpha-$Cα RMSDs derived from the structural overlap between RAD51 homology models (chain B) with and without BRC4 peptide.

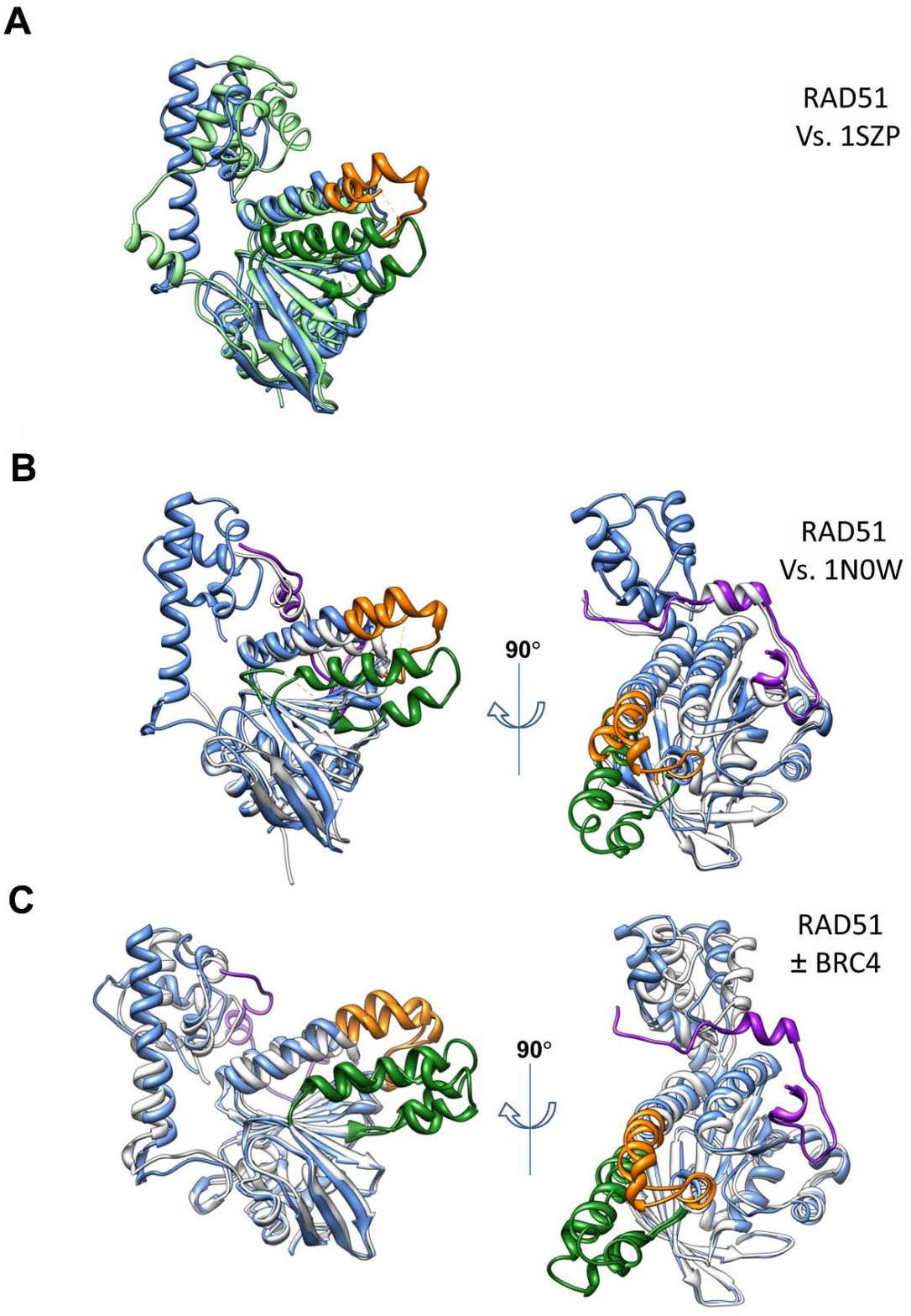


Figure 2.3. Ribbon representation of the structural overlaps that yielded RMSD values shown in Figure 2.2.

Figure 2.3. Ribbon representation of the structural overlaps that yielded RMSD values shown in Figure 2.2. **A.** RAD51 chain B (blue) vs. yeast Rad51 1SZP-B (green). **B.** RAD51 chain B with peptide (RAD51 is shown in blue, peptide is in purple) vs. the core RAD51 and BRC4 from 1N0W (both are shown in white). **C.** RAD51 chain B with peptide (RAD51 is shown in blue, peptide is in purple) vs. RAD51 Chain B in the absence of the peptide (white). In all structures, the DNA binding loops L1 & L2 are shown in orange and green, respectively.

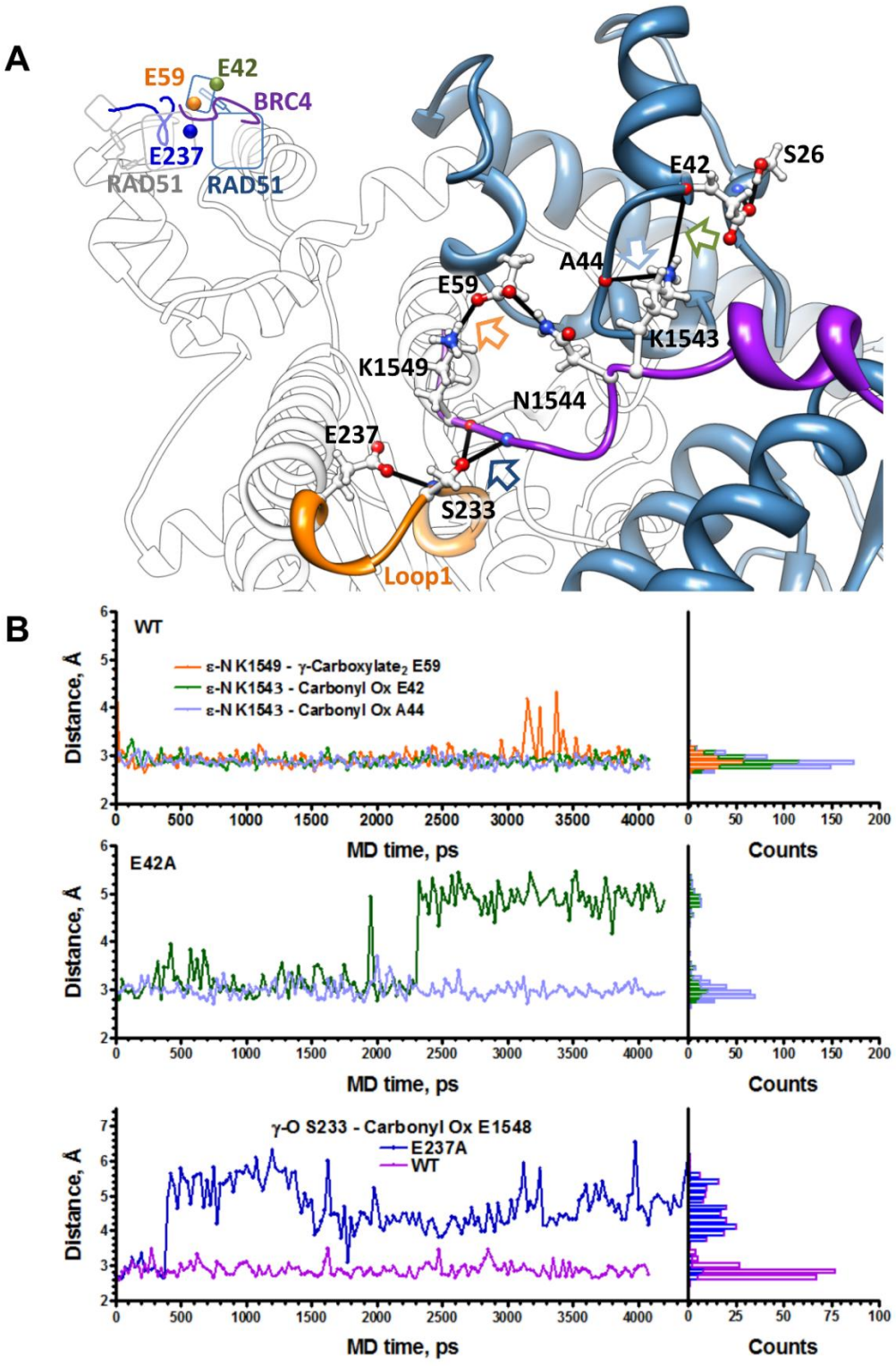
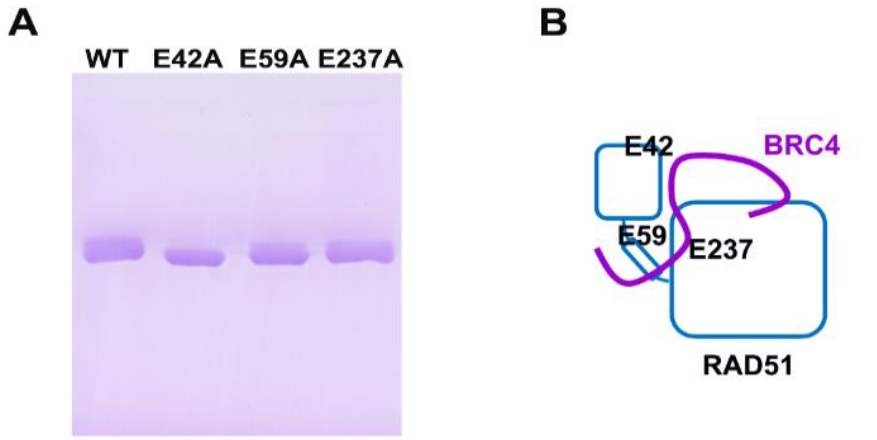
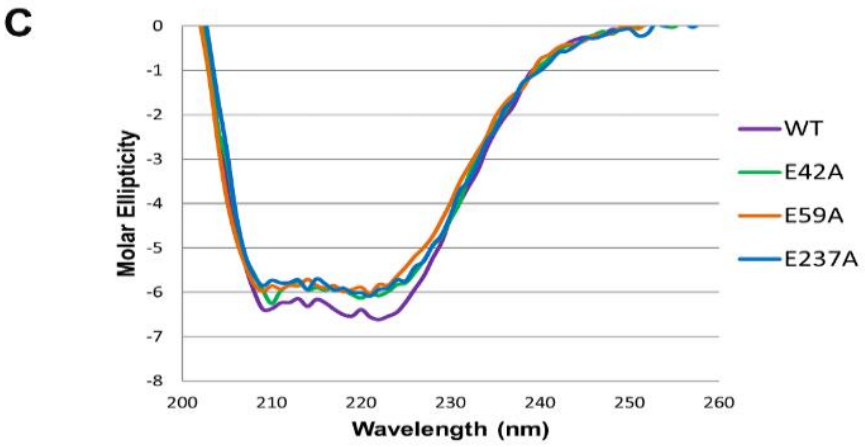


Figure 2.4. Non-obvious interactions involved in BRC4 positioning between the RAD51 core and NTD revealed by the model.

Figure 2.4. Non-obvious interactions involved in BRC4 positioning between the RAD51 core and NTD revealed by the model. **A.** Interface between RAD51 (chain B is shown in blue, chain A is shown in white) and BRC4 peptide (purple). The key residues are represented in ball in stick. The key hydrogen bonds described in the text are shown as black lines. The colored arrows indicate hydrogen bonds followed over the MD simulation as shown in the panel B. **B.** Change in the distances between heteroatoms participating in the hydrogen bonds as functions of MD simulation time. The distances were extracted from the wild type structure (top) and E42A mutant (middle); comparison of the wild type and E237A (bottom). The distributions of the distances are also shown as histograms on the right.





D

	Alpha	Beta	Random
WT	28%	19%	52%
E42A	28%	20%	52%
E59A	27%	20%	52%
E237A	26%	23%	50%

Figure 2.5. Purified RAD51 mutants retained secondary structure of the wild type RAD51 protein. **A.** SDS-PAGE gel of the purified wild type RAD51, RAD51^{E42A}, RAD51^{E59A}, and RAD51^{E237A}. **B.** Schematic representation of the RAD51 monomer showing locations of each mutant within RAD51-BRC4 complex. RAD51 is shown in blue and BRC4 peptide is in purple. **C.** CD Spectra of RAD51 Wt and alanine mutants were similar suggesting preservation of protein secondary structure. **D.** The secondary structure compositions determined using the DichroWeb K2D program.

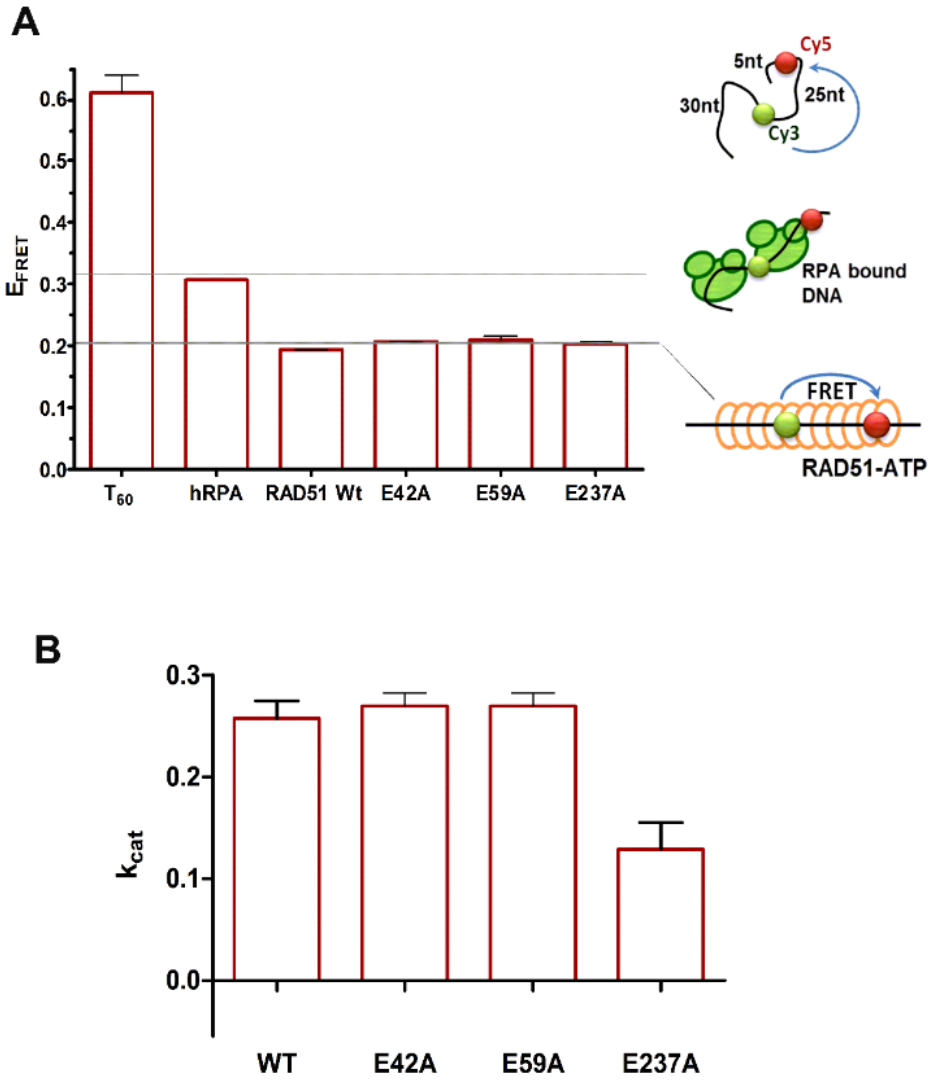


Figure 2.6. RAD51 mutants retained basic biochemical activities characteristic to the wild type protein. **A.** Formation of the active RAD51 nucleoprotein filament was monitored by following extension of the 60-mer oligonucleotide poly(dT)-60 containing Cy3 (FRET donor) and Cy5 (FRET acceptor) fluorophores separated by 25 nucleotides. Under the selected buffer conditions, the protein-free DNA has a characteristic FRET value of ~0.6; the same DNA molecule extended to its counter length upon complexation with RPA give a FRET signal of ~0.3; 1.5 times extension of the ssDNA bound by RAD51 decreases FRET to ~0.2. **B.** The k_{cat}s for the ssDNA-dependent ATP hydrolysis by the wild-type and mutant RAD51 proteins. ATP hydrolysis was measured using coupled reactions assay as described in the Supplementary Methods.

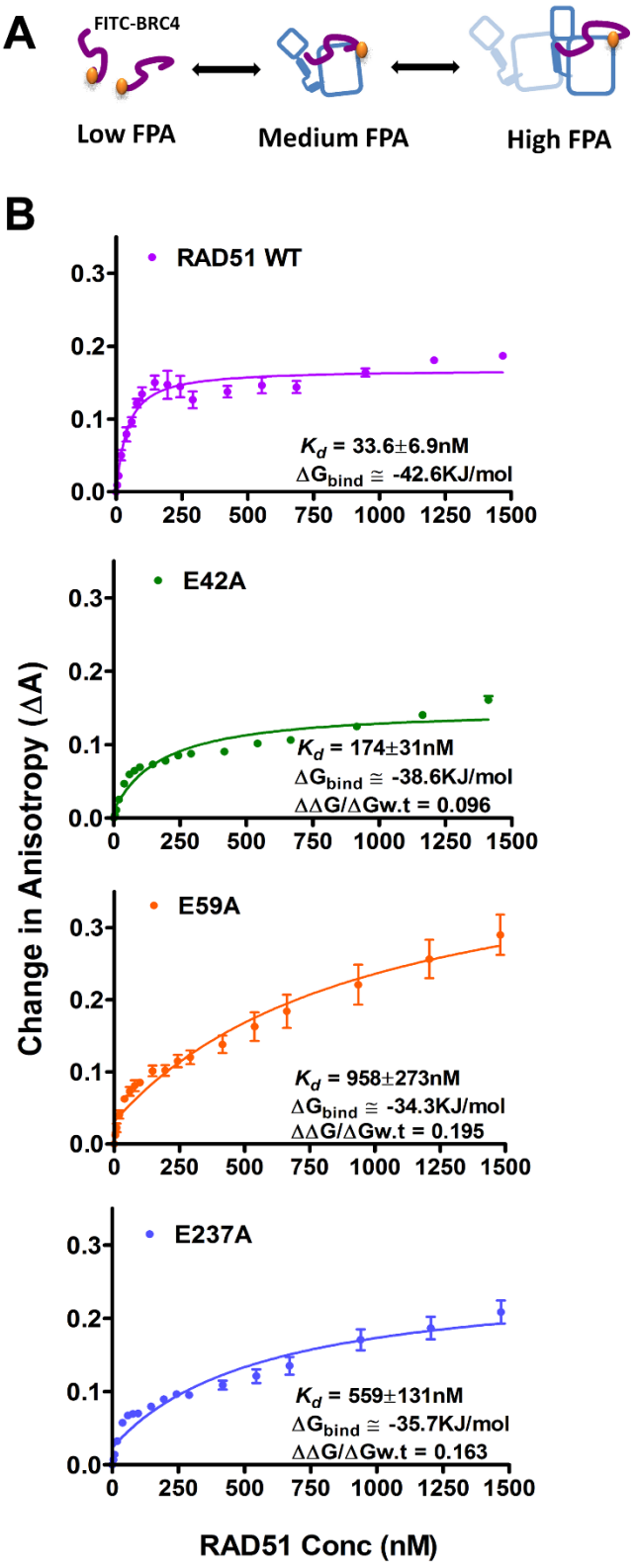


Figure 2.7. RAD51 mutants display compromised affinities for BRC4 peptide. **A.** schematic representation of the FPA-based peptide binding assay. BRC4 peptide is shown in purple; FITC dye is in orange; RAD51 is in blue **B.** Binding isotherms show titration of the 15 nM BRC4 peptide with the indicated concentrations of the wild type and mutant RAD51 proteins. The K_d values are indicated by the respective isotherms. The error bars represent standard error (SEM) for three independent experiments; error bars are smaller than the data points where not seen.

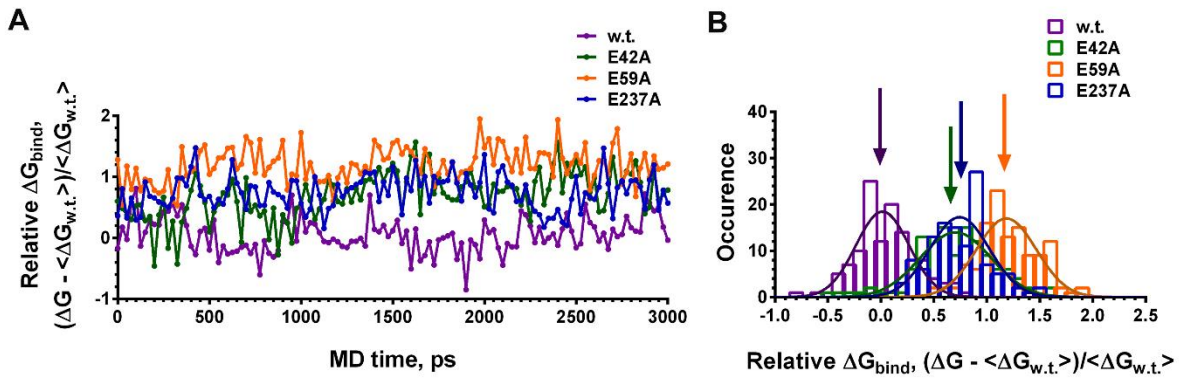


Figure 2.8. Computational binding energies ΔG_{bind} for the BRC4 interaction with wild type RAD51, E42A, E59A and E237A. **A.** ΔG_{bind} trajectories for the BRC4 complex with each protein over the course of MD simulation. **B.** Trajectories from 34 were converted into frequency histograms to yield the distributions of ΔG_{bind} values.

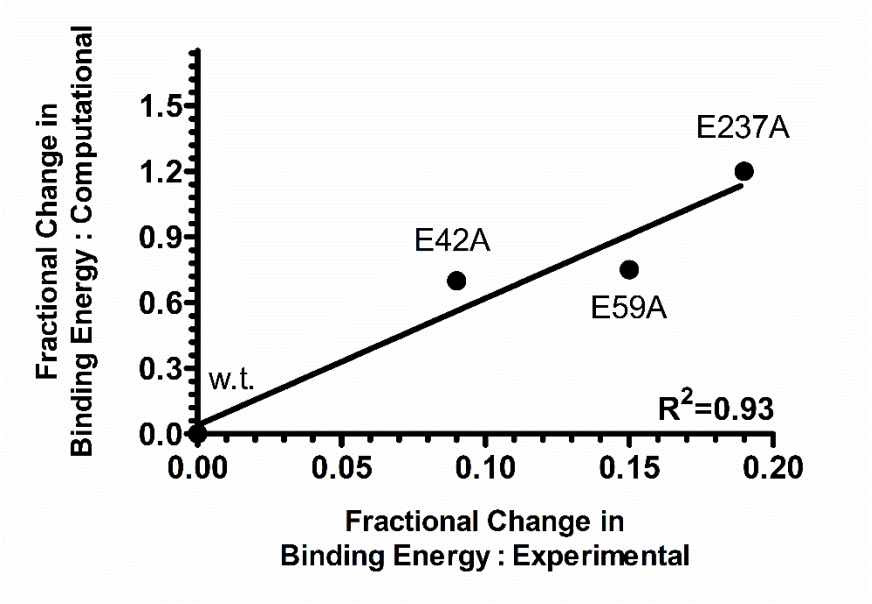


Figure 2.9. Correlation between computational and experimental relative binding free energies. Computational and experimental relative binding free energies show a high degree of correlation (R=0.96). This validates the protein-peptide interface of the RAD51-BRC4 homology model and the contributions of the RAD51 N-Terminal Domain toward BRC4 binding. Experimental fractional changes in binding energy were calculated from K_d values (Figure 2.7B). Computational fractional changes in binding energy were calculated using endpoint, boundary element method as $\langle \Delta G_{bind} \rangle = \left(\Delta G - \langle \Delta G_{w.t} \rangle\right) / \langle \Delta G_{w.t} \rangle$ (Figure 2.8).

CHAPTER 3: TYROSINE PHOSPHORYLATION STIMULATES ACTIVITY OF HUMAN RAD51 RECOMBINASE

Abstract

The DNA strand exchange protein RAD51 facilitates the central step in homologous recombination, a process fundamentally important for accurate repair of damaged chromosomes, restart of collapsed replication forks and telomere maintenance. The c-Abl tyrosine kinase and its oncogenic BCR-ABL fusion phosphorylate RAD51 on tyrosine residues 54 and 315. We combined biochemical reconstitutions with Total Internal Reflection Fluorescence Microscopy (TIRFM) to determine how the two phosphorylation events affect the biochemical activities of RAD51 and properties of the RAD51 nucleoprotein filament. By mimicking RAD51 tyrosine phosphorylation using a non-natural amino acid, p-Caboxymethyl-L-Phenylalanine (pCMF) we demonstrated that Y54 phosphorylation enhances the RAD51 recombinase activity, modifies the RAD51 nucleoprotein filament formation and allows RAD51 to efficiently compete with ssDNA binding protein RPA. In contrast, Y315 phosphorylation has little effect on the RAD51 activities. Based on our work and previous cellular studies we propose a mechanism underlying RAD51 activation by c-Abl/BCR-ABL kinases.

Introduction

The DNA in the human genome is constantly subjected to damage. This damage is a byproduct of normal cellular metabolic activities, exposure to radiation and chemical mutagens present in the environment (Ciccia and Elledge, 2010). Homologous Recombination (HR) and the pathways that employ the machinery of HR are responsible for the most accurate repair of the most deleterious DNA lesions including double-stranded DNA breaks (DSBs), inter-strand DNA crosslinks (ICLs), and damaged replication forks and thereby contributes to maintenance of the stable genome (Couedel et al., 2004; Head, 2010; Heyer, 2015; Jasin and Rothstein, 2013b; Li and Heyer, 2008a; Moynahan and Jasin, 2010; Schipler and Iliakis, 2013). HR also plays an important role in telomere maintenance (Doksani and de Lange, 2014; Oganesian and Karlseder, 2011). HR, a process highly conserved throughout the evolution, is carried out through a precisely coordinated and tightly regulated series of events. The key step in HR is the assembly of a RecA-family recombinase (phage UvsX, bacterial RecA, archaeal RadA, or eukaryotic RAD51 DNA strand exchange protein) onto resected single stranded DNA (Symington, 2014). The recombinase forms a nucleoprotein filament which then invades homologous duplex DNA resulting in a displacement loop structure that can be used as a primer for synthesis using the intact duplex as a template (Jasin and Rothstein, 2013a). Similarly to its bacterial and yeast homologs, human RAD51 binds single stranded DNA (ssDNA) in an ATP dependent manner (Tombline et al., 2002). The nucleoprotein filament formed by the ATP-bound RAD51, is arranged such that each RAD51 monomer binds three nucleotides, forming the pairing unit in these reactions (Lee et al., 2015). Beyond these basic attributes, the characteristics of RAD51 protein differs significantly from its archaeal and bacterial homologues. Subtle differences in ATP hydrolysis (Tombline and Fishel, 2002), DNA binding and nucleoprotein filament formation and extension (Holthausen et al., 2011) influence the mechanism by which the RecA homologs function as well as how they are regulated.

The critical role of HR requires all steps of this process to be tightly regulated to avoid untimely or illegitimate recombination that may cause carcinogenic genome rearrangements or result in the cytotoxic intermediates. Inactive HR causes loss of faithful DNA repair and leads to

genetic instability whereas excessive HR interferes with cellular processes such as replication, transcription and telomere maintenance and also can lead to gross chromosomal rearrangements (Heyer, 2015). In human cells, the assembly of RAD51 nucleoprotein filament is aided by the recombination mediator BRCA2 and RAD51 paralogs (Prakash et al., 2015; Zelensky et al., 2014b), antagonized by the antirecombinases (Daley et al., 2014b) and the heteroduplex rejection machinery (Spies and Fishel, 2015), and is also regulated by post translational modifications. Activities of RAD51 in the cell are influenced by two types of phosphorylation. RAD51 protein is phosphorylated at threonine residue 309 by the Chk1 checkpoint kinase (Sorensen et al., 2005) which regulates the DNA damage response. It is also phosphorylated at tyrosine residues 54 and 315 by the c-Abl/BCR-ABL tyrosine kinase (Slupianek et al., 2001; Yuan et al., 1998). Ionizing radiation and other genotoxic agents activate c-Abl kinase in an ATM and DNA-PK dependent manner (Baskaran et al., 1997; Shafman et al., 1997). There have been several studies on how c-Abl regulates RAD51 with a major debate regarding the site targeted by this tyrosine kinase. Recent studies suggested that RAD51 is phosphorylated in two steps with Y315 phosphorylation being a prerequisite for the Y54 phosphorylation (Popova et al., 2009). There remains, however, an uncertainty about the biochemical effects of RAD51 phosphorylation by c-Abl. Results of the cell-based studies suggested that the Y54 and Y315 phosphorylation enhances the RAD51 nuclear foci formation and resistance to DNA damaging agents (Slupianek et al., 2001; Yuan et al., 2003). In contrast, biochemical studies which used aspartate and glutamate amino acids to mimic Y315 phosphorylation, or used *s. cerevisiae* Rad51 (*sce*Rad51) phosphorylated *in vitro* by c-Abl kinase showed a decrease in the DNA strand exchange activity (Takizawa et al., 2004; Yuan et al., 1998). The discrepancy between the results of the biochemical and cellular studies may stem from the fact that *Saccharomyces cerevisiae* lacks tyrosine phosphorylation (Lim and Pawson, 2010) and from a substantial structural difference between yeast and human proteins (Subramanyam et al., 2013). Also, representation of phosphotyrosine by a negatively charged amino acid is often inaccurate. In fact, we have recently shown that Y/D substitution incorrectly represents tyrosine phosphorylation in another c-Abl target, human RAD52 (Honda et al., 2011).

Our studies aim to provide a mechanistic model that reconciles biochemical data with that observed in the cell-based studies and to identify the properties of RAD51 protein and its nucleoprotein filament that are altered to enhance or diminish HR. By incorporating a non-natural amino acid that accurately mimics phosphorylated tyrosine residues, we provide a realistic

representation of the phosphorylated RAD51. Moreover, the impact of the two individual phosphorylation events can be easily distinguished. Our model of human RAD51 (Subramanyam et al., 2013) predicted that Y54 participates in the inter-subunit stacking interaction with F195 of the adjacent monomer within the RAD51 filament similar to the interaction revealed in *sce*Rad51 structure (Conway et al., 2004) (Figure 3.1A). Introduction of a large negative charge by Y54 phosphorylation is expected to break the stacking and may alter the RAD51-RAD51 interface. Even slight deviations towards either a more or a less stable RAD51 filament due to this alteration may have significant effect of the RAD51 nucleoprotein filament assembly and function. Y315 is also near the RAD51-RAD51 interface and is directly adjacent to D316, which was proposed to form a salt-bridge to the γ-phosphate of ATP and to function as a conformational sensor that enhances nucleoprotein filament turnover (Amunugama et al., 2012b). It is more difficult to anticipate the effect of placing an additional negative charge next to this charged aspartate.

Using biochemical assays that report on the recombinase function of phosphomimetic RAD51, we observed elevated DNA strand exchange activity of Y54-phosphomimetic ($RAD51^{Y54pCMF}$), but not the Y-315-phosphomimetic protein ($RAD51^{Y315pCMF}$) consistent with previously published cellular studies that reported the enhanced HR and RAD51 localization at the sites of the DNA damage (Shimizu et al., 2009; Slupianek et al., 2001). Moreover, we showed that $RAD51^{Y54pCMF}$ can efficiently carry out the so-called "RPA-first" DNA strand exchange reactions that would normally require the presence of a recombination mediator. To elaborate on the mechanism of RAD51 function and regulation, we employed Forster Resonance Energy Transfer (FRET) based experiments at ensemble and single molecule levels. Analyzing single molecule data by Hidden Markov Modeling (HMM's), we showed differences in DNA binding on phosphorylation that explain the basis of nucleoprotein filament formation and robust DNA strand exchange activity of $RAD51^{Y54pCMF}$. Finally, we provide a model for the regulation of the RAD51 protein by sequential phosphorylation, highlighting the mechanism of regulation by the c-Abl/BCR-ABL kinases.

Results

Mimetics of the RAD51 Y54 and Y315 Phosphorylation

Protein phosphorylation is commonly mimicked by charged amino acids Aspartate (D) and Glutamate (E). These substitutions may inaccurately represent phosphotyrosine when the precise geometries or the aromatic ring of the tyrosine are important (Honda et al., 2011; Xie et al., 2007). For a more accurate representation of the effects of tyrosine phosphorylation on RAD51, we used a non-natural amino acid, p-carboxmethyl-L-phenylalanine (pCMF), which mimics all aspects of a phosphorylated tyrosine residue (Figure 3.1B) (Honda et al., 2011; Xie et al., 2007). RAD51^{Y54pCMF} and RAD51^{Y315pCMF} mutants were produced using expression in *E. coli* Acella™ Cells transformed with the pCH1-RAD51 vector (see methods), along with the pSUPT-UaaRS vector encoding the amber tRNA and pCMF-RS (Figure 3.2). Induction of the cells in the presence of IPTG produced large amounts of truncated proteins which could easily be distinguished from the full length protein which was expressed in the presence of pCMF (Figure 3.1C, D & E). The uniform incorporation of the amino acid for both RAD51^{Y54pCMF} and RAD51^{Y315pCMF} mutants was confirmed through western blot analysis using α-RAD51 (3C10) antibodies (Figure 3.1D) and by MALDITOF-MS as a ~42Da increase in molecular weight (Figure 3.2). Protein yields were around 20% compared to the wild type RAD51 expression owing to loss to insoluble fractions and truncated products. Mutants replacing the RAD51 Y54 and Y315 residues with Aspartate (D), Glutamate (E) and phenylalanine (F) residues were also purified (Figure 3.2) to illustrate the differences between these and the pCMF mutants.

Tyrosine 54 Phosphorylation stimulates RAD51 DNA Strand Exchange activity

A critical HR function of the RAD51 nucleoprotein filament is to perform an ATP-dependent homologous DNA pairing and the DNA strand exchange reaction (Ciccia and Elledge, 2010; Mehta and Haber, 2014). This activity can be reconstituted *in-vitro* in the three strand exchange reaction where a RAD51 nucleoprotein filament preformed on circular φX174 virion ssDNA pairs with and invades ϕX174 linear double stranded DNA (dsDNA) (Baumann et al.,

1996; Sigurdsson et al., 2001). The DNA strand exchange activity of RAD51 leads to the displacement of the complementary linear ssDNA from the dsDNA substrate eventually leading to the formation of a nicked circular product through a series of joint molecule intermediates (Figure 3.3). The reaction depends on the presence of the ssDNA binding protein RPA, which, when added after the RAD51 nucleoprotein filament is formed, stimulates the DNA strand exchange by destabilizing remaining secondary structure in the ssDNA and by sequestering the displaced DNA strand. The nicked circular product along with the joint molecule intermediates are separated on an agarose gel which tracks the progress of the reaction over time.

There are two general regimes under which human RAD51 promotes an efficient DNA strand exchange reaction. The reaction can be carried out in the presence of Ca^{2+} and Mg^{2+} ions, which stabilize the RAD51 nucleoprotein filament by preventing ATP hydrolysis, leading to efficient DNA strand exchange (Bugreev and Mazin, 2004). To measure the DNA strand exchange activity while permitting ATP hydrolysis, our reaction conditions contained Mg^{2+} and Ammonium Sodium Phosphate ($NaNH_4PO_4$) (see methods) (Shim et al., 2006). RAD51 Y54 phosphomimetic mutants showed an enhanced DNA strand exchange activity compared to the wild type protein (Figure 3.3). Nicked circular products indicating completion of the reaction, were observed earlier in reaction facilitated by $RAD51^{Y54pCMF}$ than in the reaction the wild type protein. We also observed an increase in product formed over a time course of 180 minutes from ~60% to ~80% of the input dsDNA (Figure 3.3). The $RAD51^{Y54D}$ and $RAD51^{Y54E}$ mutants were observed to have a higher stimulatory effect compared to $RAD51^{Y54pCMF}$ whereas $RAD51^{Y54F}$ behaved identical to the wild type protein (Figure 3.4). Two distinct species of the joint molecule intermediates are formed during the strand exchange reaction (Tham et al., 2013). The early joint molecule intermediates are known to migrate faster in the agarose gel compared to those formed further along reaction progression towards the nicked circular product. The stimulation of strand exchange activity can be attributed to the efficient conversion of early joint molecule intermediates which were observed in much higher amounts for the wild type protein compared to $RAD51^{Y54pCMF}$ (Figure3.3), $RAD51^{Y54D}$ or $RAD51^{Y54E}$ (Figure 3.4) mutants. Although mild stimulation of strand exchange was observed in the $RAD51^{Y315D}$ or $RAD51^{Y315E}$ mutants, the $RAD51^{Y315pCMF}$ mutant showed strand exchange activity identical to the wild type protein (Figure 3.3). Considering the

pCMF mutants have the closest structural resemblance to phosphorylated proteins, it is likely that the glutamate and aspartate mutations, strongly exaggerate the effects of tyrosine phosphorylation in RAD51.

RAD51 phosphomimetic mutants efficiently overcome the kinetic barrier to the nucleoprotein filament formation presented by the ssDNA binding protein RPA.

The RPA protein plays multiple important roles in HR and in recombination dependent DNA repair. While processing DNA double strand breaks, the ssDNA produced by end resection machinery is immediately bound by RPA preventing secondary structure formation and further stimulates end resection (Symington, 2014). The bound RPA is eventually replaced by RAD51 nucleoprotein filament in a reaction facilitated by the recombination mediator BRCA2 (Jensen et al., 2010; Liu et al., 2010b; Prakash et al., 2015). The *in-vitro* reconstituted DNA strand exchange reaction requires RPA (Baumann et al., 1996; Sung, 1994), which removes secondary structures in the ssDNA (Sigurdsson et al., 2001; Sugiyama et al., 1997) and as the reaction progresses, sequesters the displaced ssDNA thereby preventing non-productive RAD51 nucleoprotein complexes. When added first, however, RPA kinetically impedes the RAD51 nucleation (New et al., 1998). Thus, the order of RAD51 and RPA addition while performing the DNA strand exchange reaction has a very significant impact on the reaction outcome. As expected, under our experimental conditions, the efficiency of nicked circular product formation was significantly higher when RAD51-nucleoproten filaments were formed prior to the addition of RPA protein. However, while the appearance of the joint molecules and nicked circular products was delayed and diminished in the "RPA-first" reactions, it was not completely abolished. Not only did the Y54 phosphomimetic mutants display higher DNA strand exchange activity in the "RAD51-first" reactions, $RAD51^{Y54pCMF}$ displayed no difference in the efficiency between the "RAD51-first" and "RPA-first" reactions (Figure 3.5 & 3.6). These experiments suggest that phosphorylated RAD51 may efficiently displace RPA from the ssDNA in the absence of recombination mediators. As with the "RAD51-first" reaction, $RAD51^{Y315pCMF}$ was indistinguishable from the wild type protein (Figure 3.5). A slight decrease in the reaction yield was due to the effect of the storage buffer and was recapitulated in the wild type protein reaction adjusted to match the buffer. $RAD51^{Y315D}$ and $RAD51^{Y315E}$ were more robust DNA strand-exchange proteins than the wild type, $RAD51^{Y315pCMF}$

and RAD51^{Y315F} yet again suggesting that the glutamate and aspartate mutations inaccurately represent tyrosine phosphorylation.

Tyrosine phosphorylation limits RAD51 dsDNA binding.

RAD51 can form nucleoprotein filament on both ssDNA and dsDNA substrates. However, only filaments formed on ssDNA substrates are productive in homologous pairing and DNA strand exchange (Sung and Robberson, 1995). The recombination mediator BRCA2 is known to promote a bias in binding affinity favoring the ssDNA (Jensen et al., 2010; Liu et al., 2010b; Thorslund et al., 2010b). We investigated the effect of RAD51 phosphorylation on the RAD51-dsDNA interaction by incubating the RAD51 phosphomimetic mutants with ϕX174 dsDNA and analyzing the electrophoretic mobility shift. The RAD51^{Y54pCMF}, RAD51^{Y54D} and RAD51^{Y54E} showed a significant decrease in the dsDNA binding (Figure 3.7) whereas only a slight decrease was observed in RAD51^{Y315D} and RAD51^{Y315E}. As expected, the Y/F substitution had no effect on the RAD51-dsDNA interaction. The decrease in dsDNA binding in all RAD51 mutants directly correlated with the stimulation of the DNA strand exchange activity. The only deviation from this trend was RAD51^{Y315pCMF}, which upon incubation with the supercoiled ϕX174 dsDNA formed large aggregates that failed to enter the gel. Similar experiments were performed on shorter dsDNA substrates with results consistent with those from the ϕX174 dsDNA substrate (data not shown).

The altered ssDNA binding behavior of phosphomimetic mutants

To elucidate the regulation of RAD51 nucleoprotein filament formation on ssDNA, we investigated the RAD51-ssDNA interaction using Forster Resonance Energy Transfer (FRET) based assays (see methods). These experiments monitored binding of RAD51 to a Poly dT_{60} substrate which was labelled internally with Cy3 (FRET donor) and Cy5 (FRET acceptor) fluorescent dyes separated by 25 nucleotides. Free ssDNA yields a high FRET value due to the proximity of the donor and acceptor dyes. Addition of RAD51 into the reaction mixture and ensuing nucleoprotein filament formation extends the ssDNA, which can be measured as a decrease in FRET as the two fluorophores move away from each other (Subramanyam et al., 2013). RAD51 protein binds and extends ssDNA~1.5 fold beyond its contour length with each RAD51 monomer occluding approximately three nucleotides as evident by the FRET of approximately 0.55, compared to 0.19 expected for the fully extended B-form ssDNA (Figure 3.8). ATP bound

RAD51 forms stable nucleoprotein filaments (Namsaraev and Berg, 1998) while hydrolysis of ATP to ADP leads to the protein turnover. Using an NADH coupled ATPase assay (see methods), we measured ATP hydrolysis activity for all phosphomimetic mutants. All mutants could hydrolyze ATP similar to the wild type protein. (Figure 3.9D).

The characteristic ssDNA binding and extension was observed in buffers which permitted ATP hydrolysis (Mg^{2+}/ $NaNH_4PO_4$) as well as under conditions which inhibited ATP hydrolysis (Ca^{2+}/Mg^{2+}/KCl) (Figure 3.8). Under both sets of experimental conditions we observed a stoichiometric binding of the RAD51 protein to 600 nM (nucleotides) ssDNA. Under these conditions, $RAD51^{Y54pCMF}$ binding to and extension of the ssDNA was indistinguishable from the wild type protein, while $RAD51^{Y54D}$ and $RAD51^{Y54E}$ showed weaker binding to the ssDNA substrate (Figure 3.9A & B). Under conditions permitting ATP hydrolysis, binding of the $RAD51^{Y54pCMF}$ to the ssDNA was no longer stoichiometric as higher concentrations of $RAD51^{Y54pCMF}$ were required to saturate the ssDNA substrate compared to wild type. This binding deficiency was even more pronounced in the $RAD51^{Y54D}$ were $RAD51^{Y54E}$ proteins, which respectively required approximately 6- and 160-fold more protein to saturate the ssDNA substrate (Figure 3.8). No difference between the wild type RAD51 and $RAD51^{Y54F}$ was observed. $RAD51^{Y315E}$ and $RAD51^{Y315pCMF}$ mutants displayed ssDNA binding only slightly weaker than the wild type under ATP hydrolysis conditions while formed stable nucleoprotein filaments similar to wild type when hydrolysis was inhibited (Figure 3.9C & D). All other RAD51 Y315 mutants showed binding characteristics similar to the wild type protein under both, permissive and inhibitory conditions for ATP hydrolysis. The inverse correlation between the ssDNA binding and the capacity to carry out the DNA strand exchange activity was somewhat unexpected. This is especially true for the $RAD51^{Y54D}$ which in contrast to other mutants does not show ssDNA nucleoprotein filament formation even at the tens of µM concentrations exceeding those of the DNA strand exchange reactions. This apparent discrepancy can be explained from the previous studies (Carreira et al., 2009) which showed that RAD51 forms less stable complexes with short oligonucleotides: while RAD51 dissociates from the shorter oligonucleotides normally, reassembly of the nucleoprotein filament is blocked under the steady-state conditions. RAD51 forms stable nucleoprotein filaments on longer ssDNA much more efficiently. The same is likely applies to the phosphomimetic mutants.

Equilibrium single-molecule measurements quantify RAD51-ssDNA binding

To further understand the effect of the RAD51 Y54 phosphorylation, we employed single-molecule Total Internal Reflection Fluorescence Microscopy (smTIRFM). A FRET-based analysis (Joo et al., 2006) was used to follow binding of RAD51 to the individual surface-tethered partial DNA duplex molecules, where the Cy3 and Cy5 dyes were separated by 21 nucleotides and the overall ssDNA region was 60 nucleotides (Figure 3.8 & Figure 3.10). Similar to the bulk ssDNA-binding experiments described above, the Cy3 dye is excited with the green (532 nm) laser and the Cy5 dye is excited via FRET. Measurements collected for 3839 RAD51-free DNA molecules from fifteen second movies yielded a normally distributed FRET peak with an average FRET value of 0.5. A second smaller peak distributed around 0 FRET value represents the fraction of molecules that underwent acceptor photo-bleaching. Upon binding by RAD51, the donor and acceptor dyes move apart as the ssDNA molecule is stretched. This manifests in the change from high to low FRET. The DNA molecules fully extended by the bound RAD51 yielded a FRET distribution with an average FRET of approximately 0.1. Using this system, we compared the binding properties of the wild type RAD51 with RAD51^{Y54pCMF}. The concentration-dependence of these reactions recapitulates the data observed in bulk, except binding of the RAD51 protein here is not stoichiometric and therefore can be directly compared to that of RAD51^{Y54pCMF}. A clear difference was observed wherein, 250nM of RAD51 was required to fully extend the ssDNA substrates whereas 2.5μM of the Y54pCMF mutant was required to achieve similar results (Figure 3.8). This is approximately a 10-fold increase in protein required to form stable nucleoprotein filaments. Notably, these measurements predict that both proteins will bind and extend the ssDNA at concentrations used in the DNA strand exchange reaction.

RAD51 nucleates ssDNA as a dimer

To visualize the nucleation of the RAD51 filament on ssDNA in real time, we carried out smTIRFM flow experiments. In these experiments RAD51 was introduced into a reaction chamber containing tethered ssDNA molecules as described earlier, but 10 seconds into the recording. Movies were recorded for three minutes at a time resolution of 100 milliseconds. The fluorescence trajectories (time-based changes in Cy3 and Cy5 fluorescence originating from distinct spots on the slide) were collected. The trajectories that contained binding events were then corrected for donor leakage and trimmed to contain all events prior to donor or acceptor photo-bleaching (Figure

3.11A & B show representative trajectories). Only events containing anti-correlated trajectories for donor and acceptor intensities were selected for further FRET analysis. Selected trajectories were then analyzed collectively using ebFRET (van de Meent et al., 2014) to determine the number of distinct states in the trajectories. This program uses an Empirical Bayesian method to generate Hidden Markov Models (HMM) to analyze complex FRET trajectories. In our experiments, on addition of 250nM of the wild type RAD51, filament formation was observed to proceed in a step wise manner. The trajectories were fit to models for two to eight distinct FRET states. The model that best fitted the trajectories was determined by calculating the statistical mean lower bound per series output from the ebFRET program. The model that showed the best fit was then visually compared with the raw data (Figure 3.11A & B). All data points in the fit dataset were also plotted as frequency histograms which indicate the value of the individual FRET states as a Gaussian distribution. Models that under-fit or over-fit the dataset can be distinguished using this method and eliminated. Four distinct states were observed between FRET values of 0.5 and 0.1 (Figure 3.11C & D) indicating three steps. Since the donor and acceptor dyes are separated by 21 bases and each RAD51 monomer binds ssDNA as a triplet (Lee et al., 2015), we conclude that the basic unit for binding and extension of RAD51 nucleoprotein filaments is a dimer. At similar concentrations, no binding events were observed with RAD51^{Y54pCMF}. However, at 2.5μM RAD51^{Y54pCMF}, multiple trajectories were observed with assembly and dissociation events for both the wild type protein and for the RAD51^{Y54pCMF} mutant. Both proteins were found to bind and extend the nucleating nucleoprotein filament as a dimer. To determine the dynamics of the nucleus assembly, the transitions between multiple FRET states were plotted as transition density plots (TDP) (McKinney et al., 2006) (Figure 3.12). This plot represents the FRET state of a molecule prior to and after the transition on the horizontal and vertical axes respectively. Interestingly, at a lower concentration of 250nM, RAD51 nucleation includes transitions in the forward (S1→S2→S3→S4) and the reverse direction (S4→S3→S2→S1) with similar frequencies indicating a very dynamic process whereby RAD51 dimers rapidly bind and dissociate at each step of the filament formation. At higher concentrations of 2.5μM, a bias was observed towards transitions between the final two states, i.e. S3→S4 transitions are the most prevalent with lower occupancy in the density of S4→S3 transitions. This shows that the nucleoprotein filament nucleation proceeds in the forward direction. Analysis of the RAD51^{Y54pCMF} showed no binding events at 250nM protein concentration. However, the RAD51^{Y54pCMF} filament nucleation was

highly efficient compared to wild type RAD51 at higher concentrations of RAD51^{Y54pCMF}, i.e. 2.5μM. Experiments were also performed at an intermediate concentration of 750nM of protein where RAD51^{Y54pCMF} was able to form filaments with a stronger forward directionality compared to the wild type protein.

It is important to note that the studies reported above are pre-equilibrium measurements that follow the initial steps of the filament nucleation process. When the reactions were allowed to reach an equilibrium (i.e. 5 – 10 min after the addition of RAD51 or RAD51^{Y54pCMF} to the surface-tethered DNA), we observed a stable low FRET signal indicative of the fully-formed nucleoprotein filaments.

To determine whether the differences between the nucleation mechanisms of the wild type RAD51 and RAD51^{Y54pCMF} are due to differences in their respective capacity to form dimers required for the filament nucleation we analyzed the oligomeric state of the phosphomimetic proteins by native PAGE (Figure 3.13). This analysis was carried out at 7.5 μM of each protein, which is the concentration used in our DNA strand exchange reactions. Prior to loading on the gel, the proteins were incubated in the strand exchange buffer. While all assayed proteins showed a tendency to form large oligomeric complexes, slightly lower molecular weight and more discrete bands were observed for the RAD51^{Y54pCMF}, RAD51^{Y54D} and RAD51^{Y54E} mutants.

Discussion

Here, we addressed the biochemical consequences of RAD51 phosphorylation by c-Abl/BCR-ABL using phosphomimetic mutants that accurately mimic the effect of tyrosine phosphorylation (Honda et al., 2009; Xie et al., 2007). First, we showed that RAD51^{Y54pCMF} protein has an enhanced DNA strand exchange activity on the plasmid-length substrates (Figure 3.3A). Our findings contrast previous biochemical studies where phosphorylation of Y54 on RAD51 was thought to inhibit the DNA strand exchange (Yuan et al., 1998), but is in an excellent agreement with cellular studies where RAD51 nucleoprotein formation was enhanced after c-Abl phosphorylation (Yuan et al., 2003). These earlier biochemical studies were performed using *sce*Rad51 phosphorylated *in vitro* by human c-Abl. While the tyrosine residue analogous to Y54 is conserved in *sce*Rad51, yeast cells lack tyrosine phosphorylation (Lim and Pawson, 2010) resulting in different

mechanisms of regulation for the yeast and human proteins. This makes our studies the first to mechanistically address and to parse out the effects of the two tyrosine phosphorylation events on RAD51 activities. We observed that the Y54 phosphorylation enhances the recombinase activity of RAD51 in at least three important ways. Firstly, RAD51^{Y54pCMF} appeared to efficiently promote nicked circular DNA formation, the DNA strand exchange product, while the joint molecule intermediates accumulated to a much greater extent in the reactions utilizing the unmodified RAD51 protein (Figure 3.3). Human RAD51 readily forms the DNA joint-molecules, but, in contrast to its bacterial functional homolog RecA, is inefficient at their branch migration (Baumann et al., 1996). Y54 phosphorylation appears to alleviate this deficiency allowing for a rapid clearing of the intermediates and formation of the DNA strand exchange products. Second and most unexpectedly, RAD51^{Y54pCMF} was able to enhance the so-called "RPA-first" reaction allowing it to reach the extent of the "RAD51-first" reaction (Figure 3.5A). The ssDNA binding protein RPA plays both inhibitory and stimulatory roles at different HR steps. RPA stimulates the DNA strand exchange by removing regions of the secondary structure in the ssDNA, which allows RAD51 to form a contiguous nucleoprotein filament (Sugiyama et al., 1997), and also sequesters the displaced DNA strand thereby assisting in progression of the reaction from the joint molecule intermediates to the nicked circular products by competing with RAD51 for free ssDNA (Sigurdsson et al., 2001). If added to the ssDNA first ("RPA-first" reaction), RPA delays the RAD51 nucleoprotein filament formation by kinetically inhibiting the nucleation step (New et al., 1998). This inhibition is overcome by the recombination mediator activity of BRCA2 (Jensen et al., 2010; Liu et al., 2010b). Robust DNA strand exchange reaction promoted by RAD51^{Y54pCMF} suggests that RAD51 protein phosphorylated by the c-Abl kinase may nucleate on the RPA-coated ssDNA and form contiguous active nucleoprotein filaments even in the absence of the recombination mediator. However, a scenario where the DNA damage response leads to a coordinated activation of several simultaneous pathways that ensure faithful and efficient recombination mediated DNA repair is more likely. BRCA2 has been shown to diffuse in the nucleus together with RAD51 delivering it to the sites of DNA damage (Reuter et al., 2014) with each BRCA2 binding 4-5 RAD51 monomers and promoting the filament nucleation (Jensen et al., 2010; Shahid et al., 2014). Phosphorylation of RAD51 may then ensure the filament growth on the RPA-coated ssDNA, as the RAD51 focus formation is reduced in c-Abl$^{-/-}$ cells (Yuan et al., 2003). Finally, the selectivity of the RAD51^{Y54pCMF} towards binding the ssDNA over dsDNA (Figures 3.7

and 3.8) is another means of enhancing the RAD51-mediated DNA strand exchange reaction by favoring the formation of active nucleoprotein filament, and is also akin to the function of the recombination mediator (Jensen et al., 2010; Sung and Robberson, 1995).

In contrast to RAD51^{Y54pCMF}, RAD51^{Y315pCMF} displayed the biochemical activities essentially similar to those of the unmodified protein. These data are consistent with a sequential phosphorylation of RAD51 proposed earlier (Colicelli, 2010; Popova et al., 2011). According this this model, c-Abl phosphorylates RAD51 on Y315 by binding the consensus PXXP motif through its SH3 domain. Phosphorylation of Y315 residue then forms a pYXXP motif which is a consensus binding motif for the SH2 domain of c-Abl. This in turn leads to phosphorylation of the Y54 residue which upregulates DNA repair activity through stimulation of the RAD51-mediated DNA strand exchange.

The ostensible discrepancy between the enhanced DNA strand exchange activity of RAD51^{Y54pCMF} and its reduced capacity to bind ssDNA was unexpected. A ten-fold difference in the concentration of the unmodified RAD51 and RAD51^{Y54pCMF} required to bind and extend the ssDNA under conditions identical to those used in the DNA strand exchange reactions (Figure 3.8B) may be explained as reduced affinity for ssDNA, change in the binding cooperativity, or in the altered RAD51 oligomerization. RAD51 nucleoprotein filaments were shown to grow from heterogeneous nuclei ranging size from dimers and even monomers to large oligomers (Candelli et al., 2014). Using single-molecule TIRFM we were able to follow the formation kinetics of these nuclei at a single monomer resolution. We observed no RAD51 monomers binding to ssDNA, but showed that both RAD51 and RAD51^{Y54pCMF} interact with ssDNA as a dimer. The nucleus then grows by dynamic addition and dissociation of the RAD51 dimers. At intermediate and high protein concentrations, growth of the RAD51^{Y54pCMF} nucleus is more directional than that of the unmodified RAD51, which manifests in fewer transitions in the reverse direction (Figure 3.12B). The lack of the DNA binding at 250 nM RAD51^{Y54pCMF} may be due to the decreased oligomerization of the phosphomimetic protein brought about by a possibly slightly distorted monomer-monomer interface. Y54 is located in the N-terminal domain of RAD51 near the monomer-monomer interface. In the unmodified protein, it participates in the stacking interaction with F195 of the adjacent monomer (Subramanyam et al., 2013). Introduction of the negative charge upon Y45 phosphorylation or due to the presence of phosphomimetic residue may affect this interface so that the RAD51 oligomerization is delayed until higher protein concentrations.

Although RAD51 nucleates the nucleoprotein filaments at much lower protein concentrations, these complexes are dynamic, unstable, and are likely non-productive with respect to formation of the recombination proficient nucleoprotein filaments. On the other hand, RAD51^{Y54pCMF} while unable to form these early unstable filaments at lower protein concentrations forms more stable, recombination proficient nucleoprotein filaments at higher protein concentrations. The same properties that allow RAD51^{Y54pCMF} dimers to directionally and rapidly assemble into the contiguous nuclei, are also likely underlie its ability to displace RPA and to efficiently carry out the "RPA-first" DNA strand exchange reactions.

Materials and Methods

Protein Expression & Purification

To achieve a robust expression system compatible with both, the expression of the soluble RAD51 protein and efficient pCMF incorporation, the *RAD51* open reading frame was codon optimized for *E. coli* expression and cloned into a pCOLADuet (Novagen) expression vector which also encoded for the *E. coli* GroE Operon. The resulting plasmid is referred herein as pCH1-RAD51. Protein incorporating pCMF was expressed in *E. coli* Acella™ Cells (F^- *ompT* *hsd*S_B(r_B^- m_B^-) *gal dcm* (DE3) Δ*endA* Δ*recA*). These cells were transformed with the pCH1 vector encoding *E. coli* codon optimized RAD51 Y54TAG and Y54315TAG sequences, along with the pSUPT-UaaRS vector encoding the amber tRNA and pCMF-RS expressed under *proK* and *araBAD* promoters respectively. *E. coli* Acella™/pCH1-RAD51 cells were taken from a glycerol stock and inoculated in 10 ml of LB Broth containing 40µg/ml Kanamycin and incubated at 37°C overnight. 8ml of the overnight culture was used to seed a 1L LB culture containing 40µg/ml Kanamycin. The culture was grown at 37°C until an OD_{600} of ~0.6 was reached. RAD51 expression was then induced for 4 hours at 37°C with 0.1 mM IPTG. The cells were pelleted and stored at -80°C for protein purification. RAD51 protein was then purified using the protocol described previously (Subramanyam et al., 2013). Briefly, cells from the RAD51 expression were thawed on ice for two hours and resuspended in lysis buffer containing protease inhibitors and lysozyme. The cells were then sonicated and the soluble fraction was dialyzed against spermidine acetate buffer. Spermidine precipitates RAD51 which is selectively resuspended in buffers containing increasing amounts of sodium chloride. The fractions containing RAD51 are then purified through a Blue Agarose

column followed by a Heparin column. Finally, the protein collected from this step is concentrated using a MonoQ anion exchange column. RAD51 pCMF mutants were expressed similarly to the wild type protein except that, *E. coli* Acella™/pCH1-RAD51/pSUPT-UaaRS starter cultures were seeded to 2L culture volumes and grown in the presence of 34µg/ml chloramphenicol and 1% arabinose in addition to 40µg/ml Kanamycin. These cultures were then induced as described above with IPTG along with 300µg/ml of p-carboxymethyl-L-phenylalanine (custom synthesized by AsisChem). The wild type protein concentration was determined using molar extinction coefficient of 14900 $M^{-1}cm^{-1}$ at A_{280}. The concentrations for all mutants were measured using molar extinction coefficient of 13410 $M^{-1}cm^{-1}$ at A_{280}. Human RPA was purified as described earlier (Henricksen et al., 1994).

DNA Strand Exchange Assay

RAD51 DNA Strand Exchange Assay was performed similarly as described earlier (Sigurdsson et al., 2001). The assay was performed using the φX174 virion ssDNA and the φX174 RFI dsDNA as substrates for the reaction. The φX174 RFI dsDNA was linearized using ApaLI to generate 4 base pair, 5' overhangs, and purified using phenol-chloroform-isoamyl alcohol extraction. The "RAD51-first" DNA strand exchange reactions were initiated by incubating 7.5µM RAD51 with 30µM (Nucleotide) φX174 ssDNA and 2.5mM ATP in the Reaction Buffer (20mM HEPES-NaOH pH 7.5, 10% Glycerol, 1mM $MgCl_2$ and 1mM DTT) for 5 minutes at 37°C. To this, 150mM of $NaNH_4PO_4$ was added along with 2µM RPA unless indicated otherwise, and incubated for another 5 minutes. This was followed by addition of 15µM (base pair) of ApaLI-digested Linear φX174 dsDNA and further incubation at 37°C over for 30, 60, 90 or 120 min. All concentrations were determined for a 90µl Reaction volume. The samples were then deproteinized by addition of 0.8% SDS and Proteinase K (800µg/ml) followed by incubation at 37°C for 30 minutes and resolved on a 0.8% Agarose gel (UltraPure Agarose) at 3.5V/cm (60V) at 25°C for 16 hours. The gel was then stained for 10 minutes using SYBR Gold and then destained for 20 minutes in MilliQ H_2O. In the "RPA-first" reactions the order of addition of RAD51 and RPA was reversed. The appearance of the nicked circular products were quantitated using ImageJ software normalized to the amount of initial linear dsDNA substrate. The intensities obtained for the mutants were then statistically compared to the wild type strand exchange reaction using two way ANOVA and verified for significance. All data were plot using GraphPad Prism 6.

ATPase Assay

A coupled NADH assay was performed as described earlier (Kreuzer and Jongeneel, 1983). A reaction mixture was assembled containing 20mM HEPES-NaOH pH7.5, 5mM MgCl2, 150mM $NaNH_4PO_4$, 1mM DTT, 0.2mg/ml NADH, 7.5mM Phosphoenol Pyruvate, 2.5mM ATP, 20U of Lactate Dehydrogenase and Pyruvate Kinase and 7.5μM RAD51. The reaction was incubated at 37°C and A_{340} measured using an Agilent 8453 UV-Vis Spectrophotometer. 80μM nucleotides of poly dT_{100} was added to start the DNA dependent ATP hydrolysis. The slope of the decrease of NADH absorbance at 340nm was used to calculate the rate of ATP hydrolysis using the following conversion: rate of A_{340} decrease (s^{-1}) × 9880 = rate of ATP hydrolysis (μM/min). Data were plotted using GraphPad Prism 6.

RAD51 DNA Binding & Extension Assay

The FRET based DNA binding assay for RAD51 was performed as described in (Subramanyam et al., 2013). Six hundred nM (nucleotides) of dT_{60} oligo labeled with the Cy3 and Cy5 dyes separated by 25 nt was titrated with RAD51 protein in FRET Reaction Buffer (20mM HEPES pH 7.5, 5mM $CaCl_2$, 5mM $MgCl_2$, 150mM KCl, 1mM ATP, 1mM DTT or 20mM HEPES pH 7.5, 150mM $NaNH_4PO_4$, 2mM $MgCl_2$, 1mM ATP, 1mM DTT) at 37˚C. Cy3 and Cy5 fluorescence was recorded using Cary Eclipse Fluorimeter. FRET was calculated as a fraction of acceptor intensity relative to the total donor and acceptor intensity adjusted by correction factors –

$$\frac{4.2*Icy5}{(4.2*Icy5) + (1.7*Icy3)}.$$

FRET value was plotted against the concentration to observe RAD51 DNA Binding and Extension characteristics. Data were plotted using GraphPad Prism 6.

dsDNA Binding Assay

RAD51 dsDNA binding assay was performed largely as described in (Takizawa et al., 2004). The reaction was initiated by incubating 2.5µM, 5µM and 7.5µM RAD51 with 15µM (base pair) φX174 dsDNA in 4X Reaction Buffer (80mM HEPES-NaOH pH7.5, 40% Glycerol, 4mM $MgCl_2$, 4mM Dithiothreitol and 0.4mg/ml BSA) and 2.5mM ATP and 150mM of indicated salt for 20 minutes at 37°C. The samples were then mixed with 6X DNA Loading Dye and resolved on a 0.8% Agarose gel at 3.5V/cm (50V) at 25°C for ~3 hours. The gel was then stained for 10 minutes using SYBR Gold and then de-stained for 10 minutes in MilliQ H_2O.

Acquisition of Single Molecule Data

A diode-pumped solid state (DPSS) 532nm green laser was used to excite Cy3 molecules in an evanescent field produced in a reaction chamber. Scattered light was removed using Cy3/Cy5 dual band-pass filter (Semrock, FF01-577/690) in the emission optical pathway Images were chromatically separated into Cy3 image and Cy5 image using 630-nm dichroic mirror inside the dual view system (DV2; Photometrics). The data acquisition was carried out through an Andor IXON CCD camera, using software written in Visual C++. The movies obtained with the CCD were analyzed first using IDL and the intensities of the fluorophores and the time traces were visualized using customized MATLAB programs. FRET was calculated using the equation (Ha, 2001) –

$$E = \frac{1}{\left(1 + \gamma \frac{I_D}{I_A}\right)^{-1}}$$

The where I_D and I_A are the sensitized emission intensity of the donor and acceptor, respectively. The measured raw intensities of the donor and acceptor channels were then corrected by measuring the leakage intensities from the donor channel β to the acceptor channel. The donor leakage was subtracted from the acceptor channel intensities and added it back to the donor intensity. For our optical setup, the donor leakage correction was measured to be 7%. Applying this correction, we get the equation for FRET given below –

$$E = \frac{(I_A^* - \beta)}{(I_A^* - \beta) + \gamma(I_D^* + \beta)}$$

Where, I_A^* and I_D^* are the background corrected acceptor and donor intensity. $\gamma = 1$, is calculated as the ratio of change in the acceptor intensity, ΔI_A to change in the donor intensity, ΔI_D upon acceptor photobleaching ($\gamma = \Delta I_A / \Delta I_D$) (Joo and Ha, 2012b; Roy et al., 2008).

Single-Molecule RAD51 DNA Binding Assay (TIRFM)

An oligo made of 18 random bases and was labelled with a 5' Cy5 dye and a 3' biotin tag. This oligo was annealed to a 78 base oligo containing the complementary 18 bases to the biotinylated oligo followed by a T_{60} sequence and an internal Cy3 dye spaced 21 bases apart from the Cy5 dye on the tethered oligo. This arrangement ensures that all oligos tethered to the slide have a complete FRET pair. PEGylated slides and sample chamber were prepared as described earlier (Joo and Ha, 2012a). 100µl of 0.2mg/ml Neutravidin was passed through the reaction chamber and incubated for 3 minutes. The chamber was washed with 300µl T50 buffer (Tris pH 8.0, 50mM NaCl). Fifty pM Biotinylated DNA substrate was then added to the Imaging Buffer (Trolox, 20mM HEPES pH 7.5, 2mM $MgCl_2$, 0.8% Glucose, 150mM $NaNH_4PO_4$, 1mM ATP, 1mM DTT, 0.1mg/ml BSA, 0.04 mg/ml Catalase and 1 mg/ml Glucose Oxidase) and flowed into the chamber. The indicated concentrations of RAD51 were then added into the Imaging Buffer and flowed into the imaging chamber. Forty movies of 15 second duration (100 ms time resolution, 400 background value and 1600 data scalar) were recorded using gain of 230 at 45.6 mW green laser intensity. Data were analyzed using customized programs for IDL & MATLAB. The DNA substrates used for the experiment are 5Cy5-Bot18-3Bio (GCCTCGCTGCCGTCGCCA) and Top18-(T21)Cy3(T39) (TGGCGACGGCAGCGAGGCTTTTTTTTTTTTTTTTTTTT/iCy3/TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT).

For experiments performed under flow, similar conditions were used to set up the experiment except 75 pM of the DNA substrate was used instead of 50 pM to account for uniform concentration of molecules in a larger flow chamber. Once the DNA substrate was bound to the PEGylated slide, the imaging chamber was washed with the RAD51 Imaging Buffer. Using the same values for laser power and gain, three minute movies were recorded at the same time

resolution with the indicated concentration of RAD51 being injected into the imaging chamber at ~10 seconds into the movie recording.

Analysis of Single Molecule binding Data

To generate FRET state histograms, five frames from each movie were used to calculate the FRET values for each frame for each molecule in the movie and plotted as a FRET histogram after the data were normalized in MATLAB. Each histogram was fit to a Gaussian distribution or to a sum of Gaussians using GraphPad Prizm. The area under the histogram defined the percentage of molecules in each FRET state.

The time dependent trajectories obtained from flow experiments were analyzed using Hidden Markov Modeling. Empirical Bayesian methods were used to determine transitions indicating RAD51 filament formation. The leakage corrected trajectories were trimmed to include all data points until donor/acceptor photo-bleaching events. ebFRET analysis suite (van de Meent et al., 2014) was used to statistically determine the FRET states observed in the course of the RAD51 nucleoprotein filament formation. Mean lower bound was used to determine the evidence for the model that best described the number of states for each series. This was confirmed by visually comparing the idealized trajectories from the model with the raw data to confirm the fit.

Figures

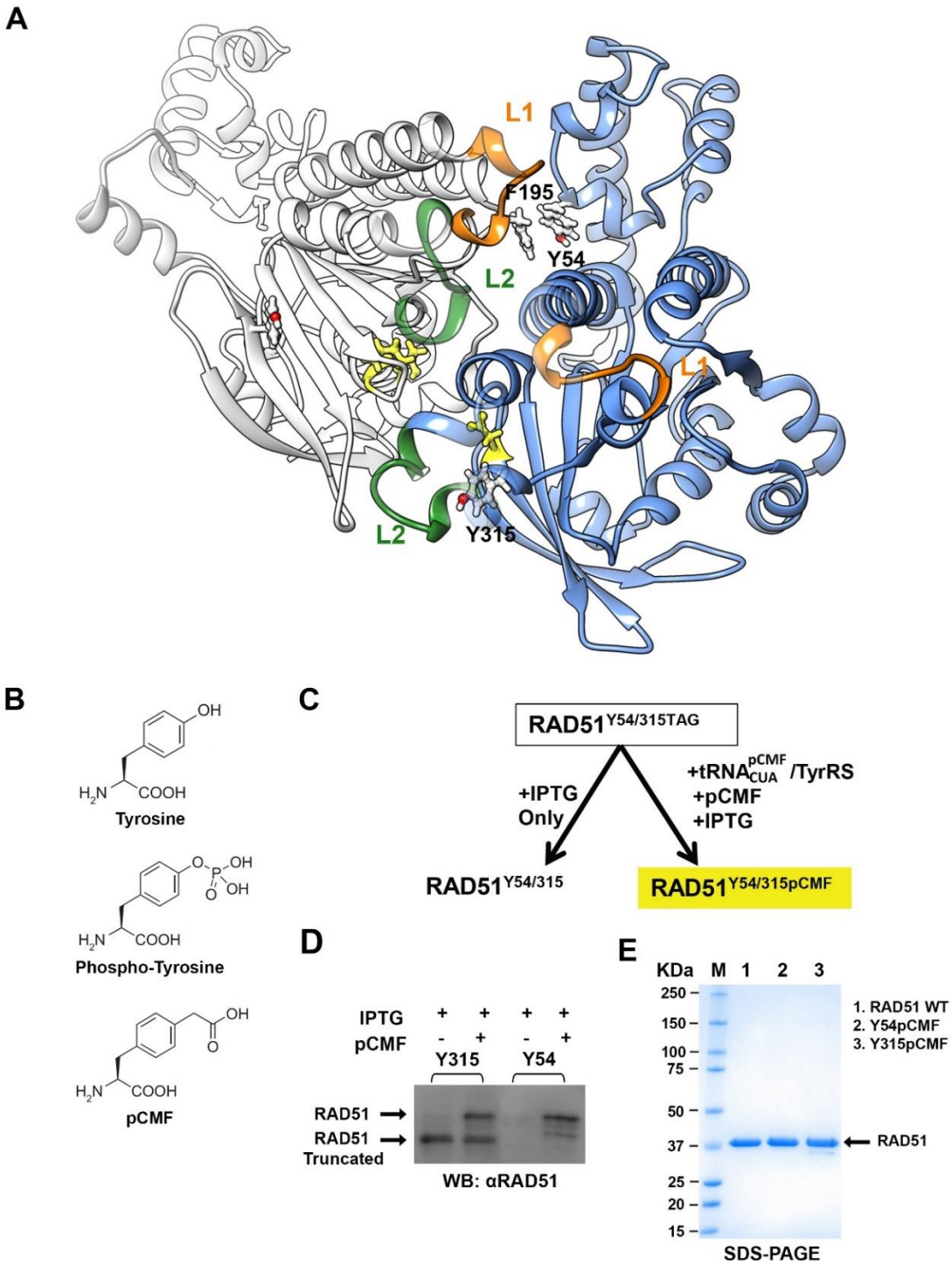


Figure 3.1. Incorporation of pCMF into RAD51

Figure 3.1. Incorporation of pCMF into RAD51 **(A)** Homology model of two adjacent RAD51 monomers. RAD51 DNA binding loops L1 & L2 are shown in orange and green respectively. Residues Y54, Y315, and F195 are shown in ball and stick rendering. **(B)** The structures of tyrosine, phosphorylated tyrosine and the non-natural amino acid pCMF (p-Carboxymethyl-L-phenylalnine). **(C)** Schematic showing expression system for the pCMF incorporation using the amber suppressor system. RAD51 expression in the absence of pCMF leads to translation of a truncated product, whereas full length protein is produced on addition of the pCMF amino acid into the expression media. **(D)** Western blot showing expression of full length $RAD51^{Y54pCMF}$ and $RAD51^{Y315pCMF}$ proteins using mouse anti-RAD51 (3C10) antibody. Truncated products are observed in the absence of pCMF (Y54 truncations are too small to be verified by standard gel-electrophoresis). **(E)** SDS-PAGE gel showing purified RAD51, $RAD51^{Y54pCMF}$ and $RAD51^{Y315pCMF}$. The $RAD51^{Y315pCMF}$ have a very small amount of truncated product (<10%) that could not be separated using our purification scheme (Figure S1B).

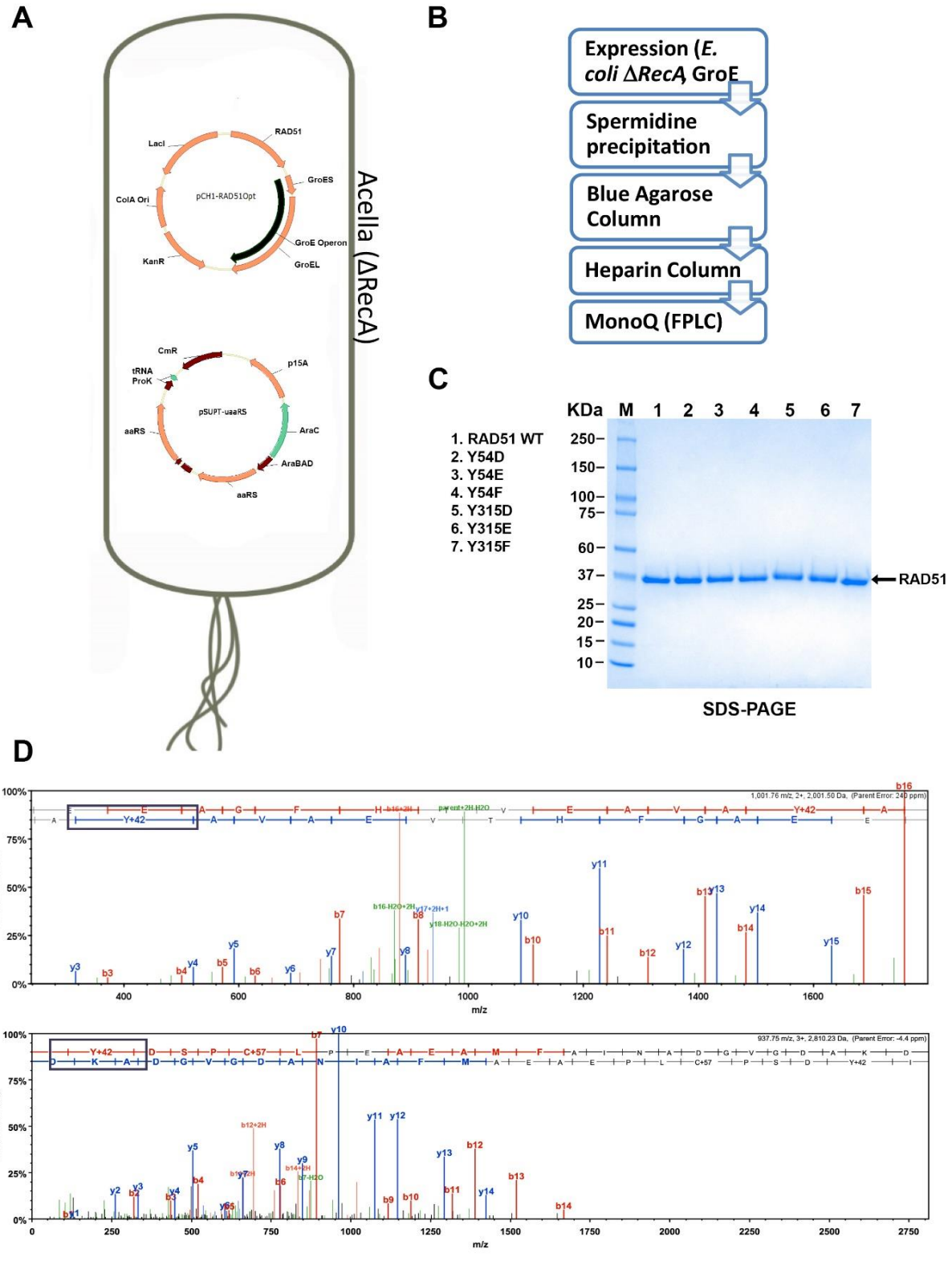


Figure 3.2. Incorporation of non-natural amino acid into RAD51 protein.

Figure 3.2. Incorporation of non-natural amino acid into RAD51 protein. **(A)** Expression system established for incorporation of the pCMF amino acid consists of *E. coli* Acella™ cells transformed with two plasmids. The first plasmid, pCH1-RAD51opt contains an*RAD51* ORF optimized for *E. coli* expression as well as the GroE operon from *E. coli*. The second plasmid contains components for the amber suppressor system. **(B)** Purification scheme for obtaining homogeneous RAD51 protein. **(C)** SDS-PAGE gel showing purified RAD51 and phosphomimetic mutants. **(D)** Data from MALDITOF-MS experiments confirming incorporation of pCMF amino acid at the Y54 (top) and Y315 (bottom) residues respectively.

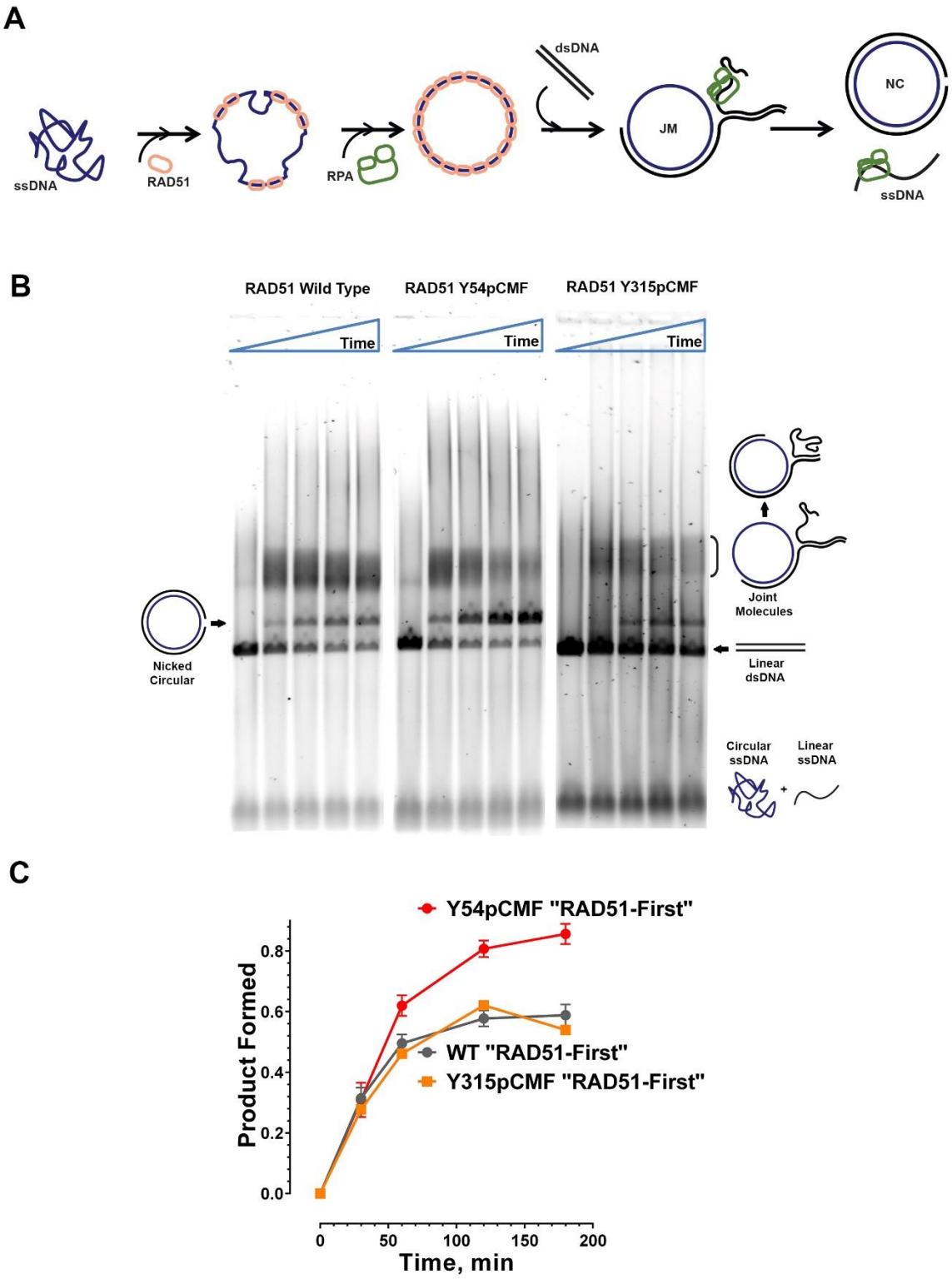


Figure 3.3. Y54 phosphorylation stimulates DNA strand exchange activity of RAD51

Figure 3.3. Y54 phosphorylation stimulates DNA strand exchange activity of RAD51 (A) Schematic showing the biochemically reconstituted DNA strand exchange reaction. Circular $\phi X174$ ssDNA is incubated with RAD51 in reaction conditions permitting ATP hydrolysis (see methods) followed by addition of RPA which helps remove secondary structures in the ssDNA allowing stable nucleoprotein filament formation. The RAD51 nucleoprotein then invades $\phi X174$ linear dsDNA to form nicked circular dsDNA products through several joint-molecule intermediates showing various stages of branch migration. **(B)** Strand exchange reactions carried out by RAD51 (left), RAD51^{Y54pCMF} (middle) and RAD51^{Y315pCMF} (right). For each protein, the respective panel shows reactions stopped at 0, 30, 60, 120 and 180 min. All substrates, joint-molecules and nicked circular products are observed on a 0.8% TAE agarose gel stained with SYBR Gold. **(C)** Quantitative analysis of formed nicked circular products. The RAD51Y54pCMF mutant (red) is able to convert ~80% of linear dsDNA substrate into nicked circular products compared to ~58% products formed for RAD51 wild type (grey) and RAD51^{Y315pCMF} mutant (orange) over a 180 minute time course. All reactions were performed in triplicate with data represented as mean ± SEM. The results for both mutants were compared to the wild type protein using two way ANOVA to verify significance of results.

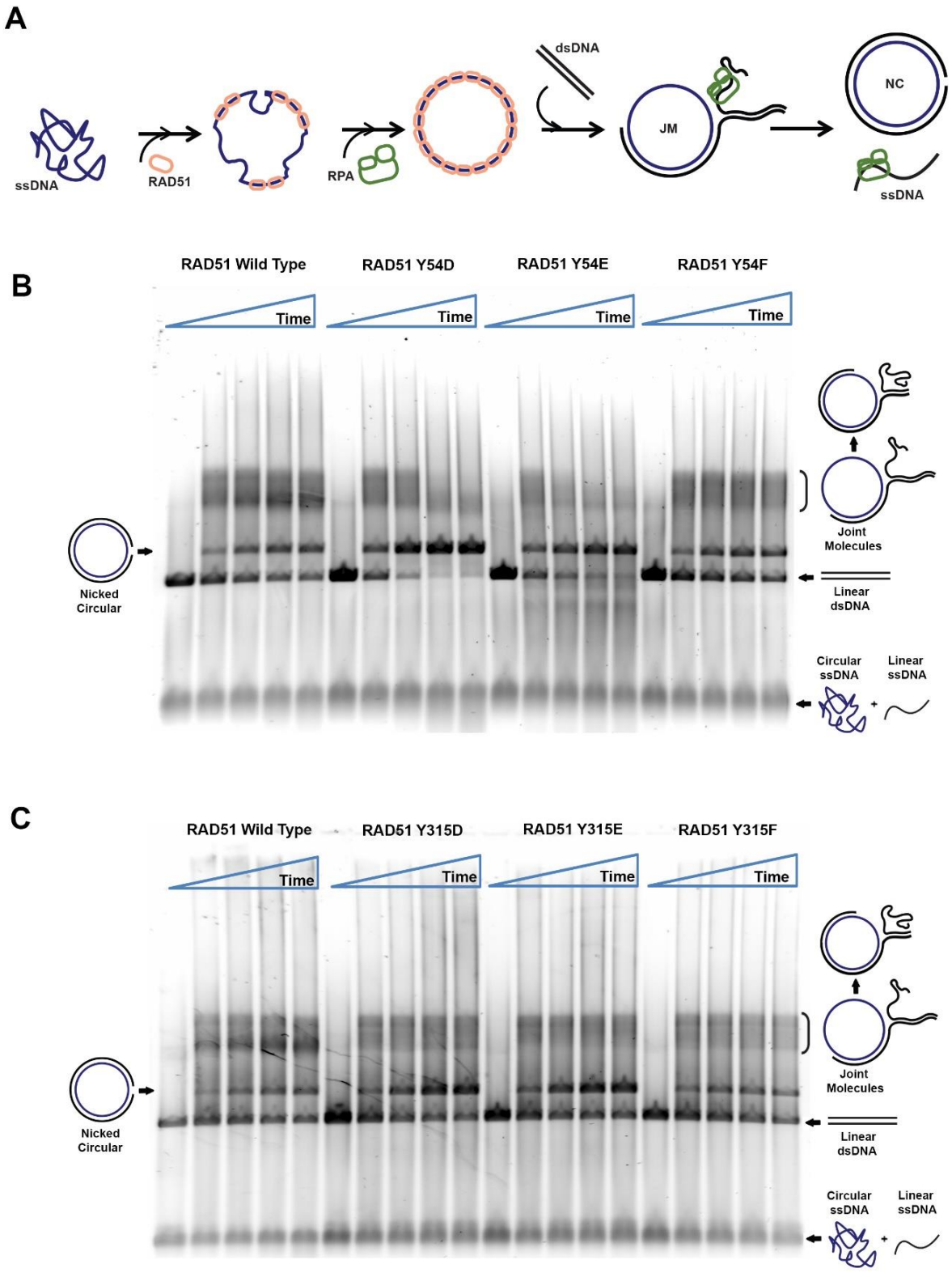


Figure 3.4. Strand exchange reactions with RAD51 Y54 & Y315 mutants

Figure 3.4. Strand exchange reactions with RAD51 Y54 & Y315 mutants **(A)** Schematic showing the *in vitro* reconstituted "RAD51-first" DNA strand exchange reaction (Figure 2). **(B)** Strand exchange reactions comparing RAD51 wild type to $RAD51^{Y54D}$, $RAD51^{Y54E}$ and $RAD51^{Y54F}$ mutants respectively. **(C)** Strand exchange reactions comparing RAD51 wild type to $RAD51^{Y315D}$, $RAD51^{Y315E}$ and $RAD51^{Y315F}$ mutants respectively.

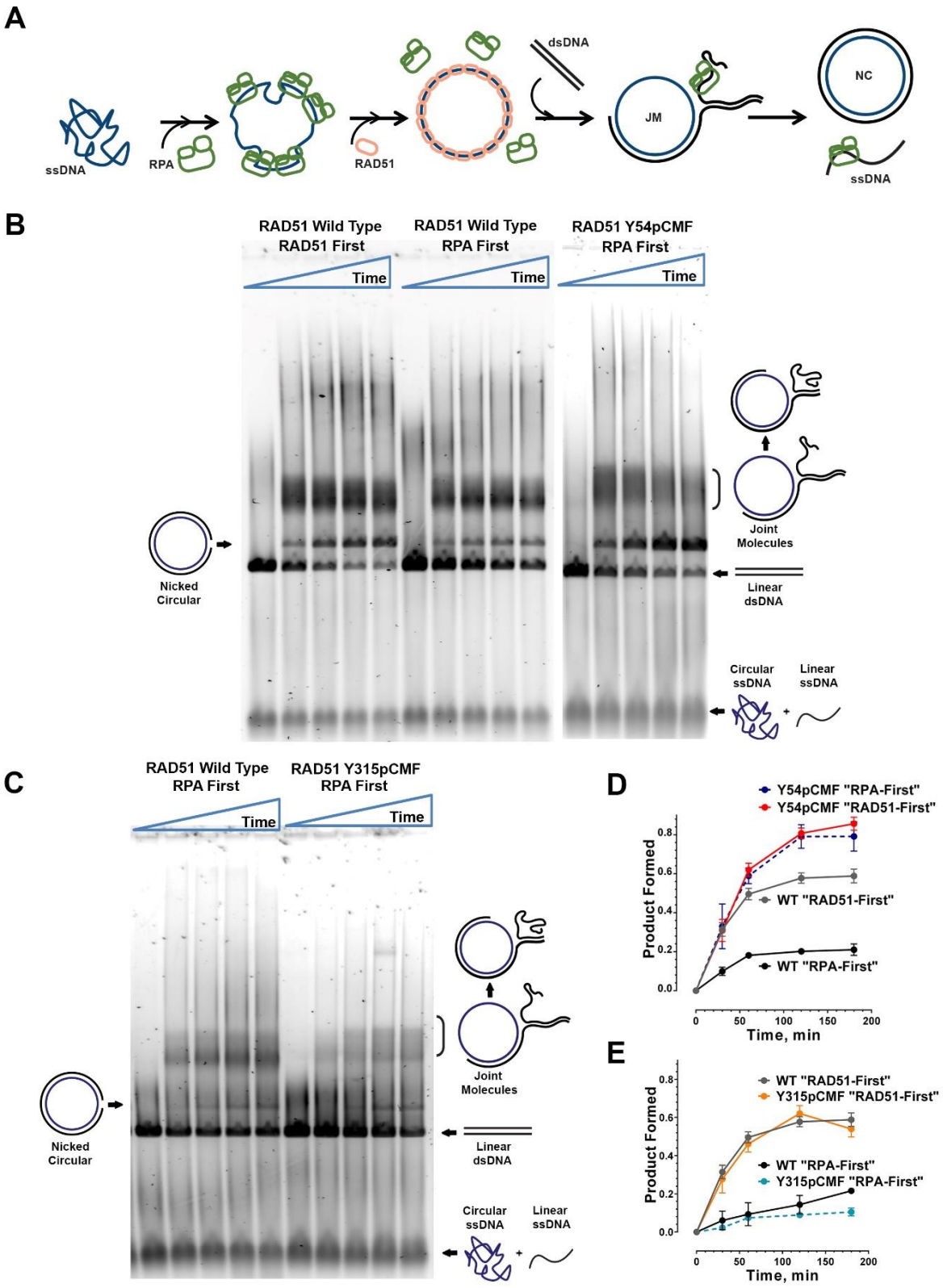


Figure 3.5. RAD51^{Y54pCMF} performs efficient "RPA-first" DNA strand exchange

Figure 3.5. RAD51^{Y54pCMF} performs efficient "RPA-first" DNA strand exchange (A) Schematic explaining the "RPA-first" reactions. Circular $\phi X174$ ssDNA is incubated with RPA protein. This is followed by RAD51 protein which displaces the bound RPA to form nucleoprotein filaments which can then invade $\phi X174$ linear dsDNA to form nicked circular products through a series of joint-molecule intermediates. (**B**) "RPA-first" DNA strand exchange reactions carried out by RAD51 (left), RAD51^{Y54pCMF} (middle) and RAD51^{Y315pCMF} (right). For each protein, the respective panel shows reactions stopped at 0, 30, 60, 120 and 180 min. **(C)** "RPA-first" DNA strand exchange reactions carried out by RAD51 (left) and RAD51^{Y315pCMF} (right). **(D&E)** Quantitation of product formed in the DNA strand exchange reactions. The RAD51^{Y54pCMF} mutant (blue) converts ~80% of linear dsDNA substrate into nicked circular products in the presence of RPA compared to ~20% products formed for RAD51 wild type (black) while RAD51^{Y315pCMF} mutant (cyan) is similar to wild type protein in both "RAD51-first" and "RPA-first" reactions. All reactions were performed in triplicate with data represented as mean ± SEM. The results for both mutants were compared to the wild type protein using two way ANOVA to verify significance of results.

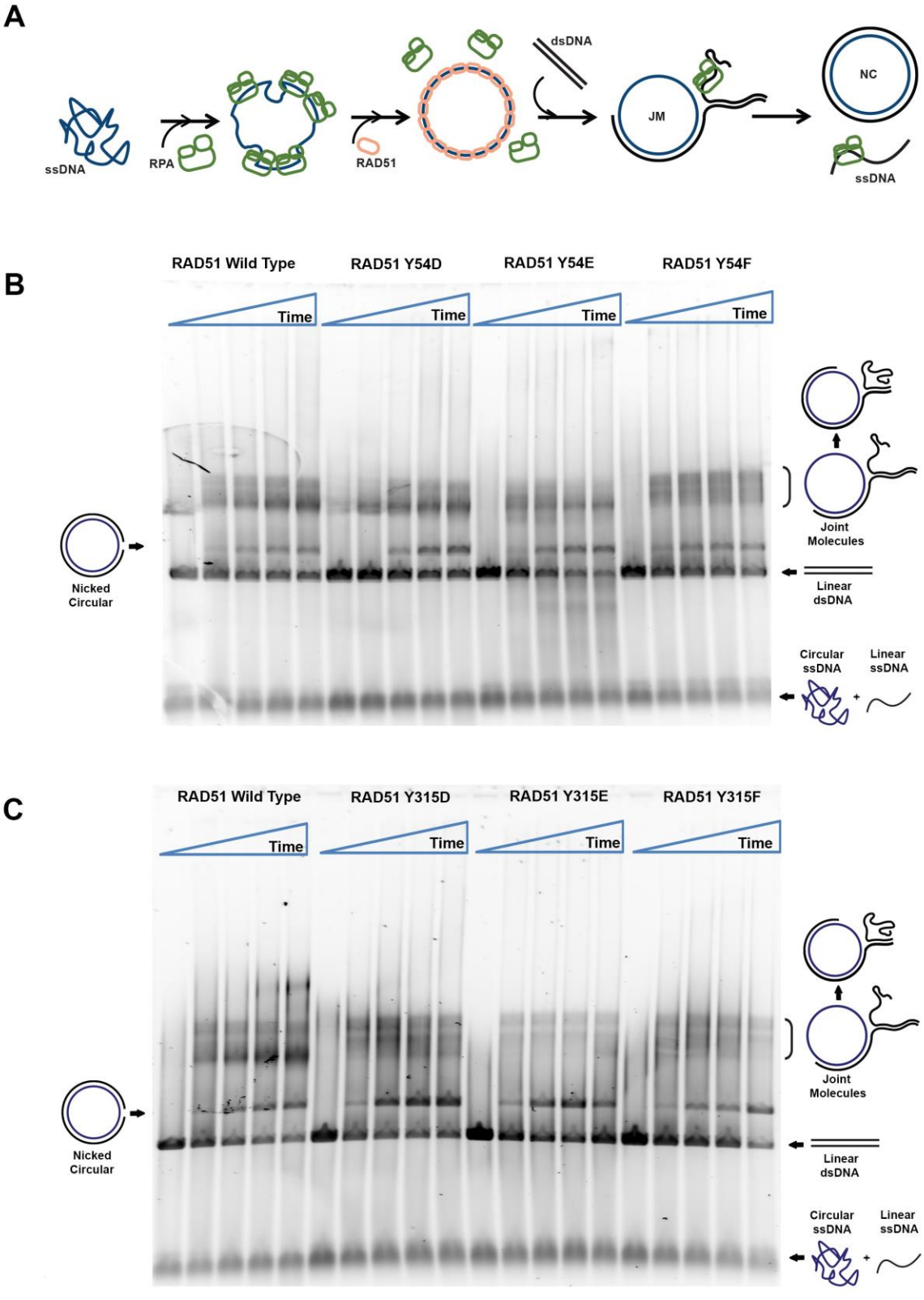


Figure 3.6. "RPA-First" Strand exchange reactions with RAD51 Y54 & Y315 mutants

Figure 3.6. "RPA-First" Strand exchange reactions with RAD51 Y54 & Y315 mutants (A) Schematic showing the *in vitro* reconstituted "RPA-first" DNA strand exchange reaction (Figure 3). **(B)** Strand exchange reactions comparing RAD51 wild type to $RAD51^{Y54D}$, $RAD51^{Y54E}$ and $RAD51^{Y54F}$ mutants respectively. **(C)** Strand exchange reactions comparing RAD51 wild type to $RAD51^{Y315D}$, $RAD51^{Y315E}$ and $RAD51^{Y315F}$ mutants respectively.

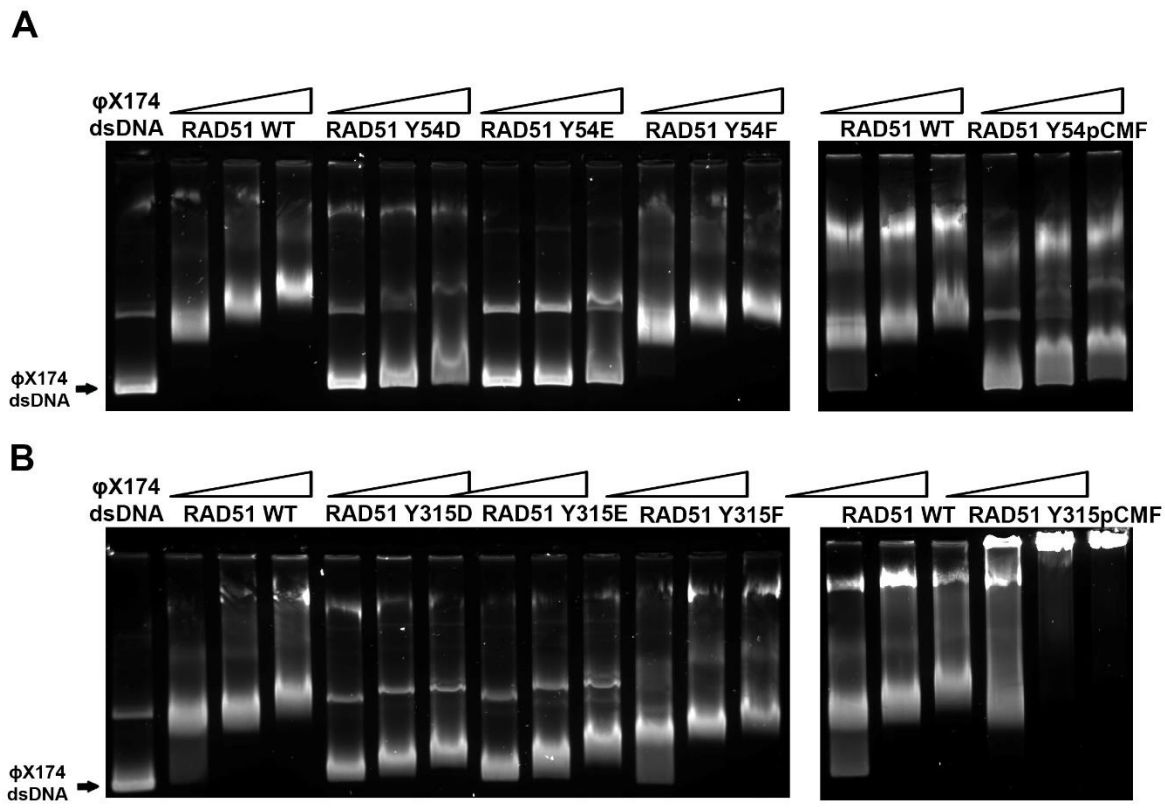


Figure 3.7. RAD51 phosphorylation limits dsDNA binding. **(A)** Electrophoretic mobility shift assay (EMSA) RAD51 Y54 mutants binding supercoiled *ϕX*174 dsDNA. 25μM (base-pair) dsDNA was incubated with increasing concentrations of 2.5μM, 5.0μM or 7.5μM RAD51 and loaded on a 0.8% TAE agarose gel and stained with SYBR gold. RAD51^{Y54pCMF}, RAD51^{Y54D}, RAD51^{Y54E} showed decrease in dsDNA binding compared to wild type protein and RAD51^{Y54F}. **(B)** In a similar assay RAD51^{Y315pCMF} formed aggregates that failed to enter into the gel whereas RAD51^{Y315D}, RAD51^{Y315E} mutants showed slightly reduced dsDNA binding compared to RAD51 and RAD51^{Y315F}.

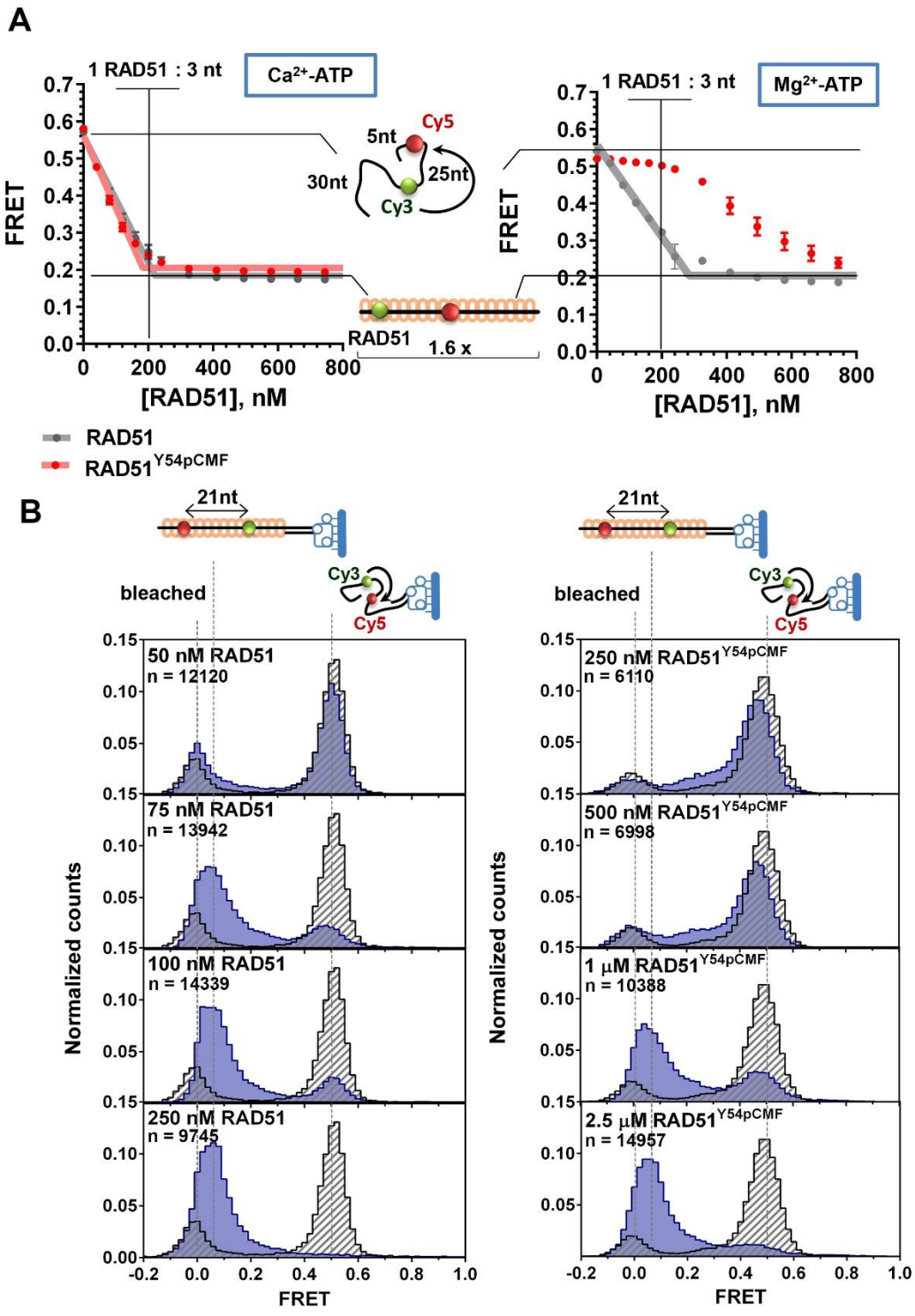


Figure 3.8. Effects of Y54 phosphorylation on ssDNA binding activity

Figure 3.8. Effects of Y54 phosphorylation on ssDNA binding activity **(A)** RAD51 binding to ssDNA was observed by following the extension of the 60-mer oligonucleotide poly(dT)-60 containing Cy3 (FRET donor) and Cy5 (FRET acceptor) fluorophores separated by 25 nucleotides. Binding of RAD51 to 600nM (Nucleotide) ssDNA moves the two dyes apart which can be seen as a change from high FRET (0.55) to low FRET (0.19). Under conditions preventing ATP hydrolysis (Ca^{2+}) the $RAD51^{Y54pCMF}$ binds and extends the ssDNA substrate similar to wild type RAD51 with a 1:3 RAD51:ssDNA ratio, whereas under reactions conditions permitting ATP hydrolysis (Mg^{2+}) $RAD51^{Y54pCMF}$ binding in non-stoichiometric. **(B)** Analysis of the equilibrium ssDNA binding using single-molecule TIRFM. FRET histograms indicating binding of RAD51 to immobilized ssDNA substrate (see methods, Figure S5). Unbound ssDNA yeilds a histogram centered on FRET value of (~0.5), while fully extended RAD51 nucleoprtoein filament yields a histogram centered on ~0.1 FRET value. Grey and blue histograms represent FRET values in the absence and presence of RAD51 respectively. Concentrations of RAD51 or $RAD51^{Y54pCMF}$, as well the number of molecules used to build each histogram are indicated in each panel.

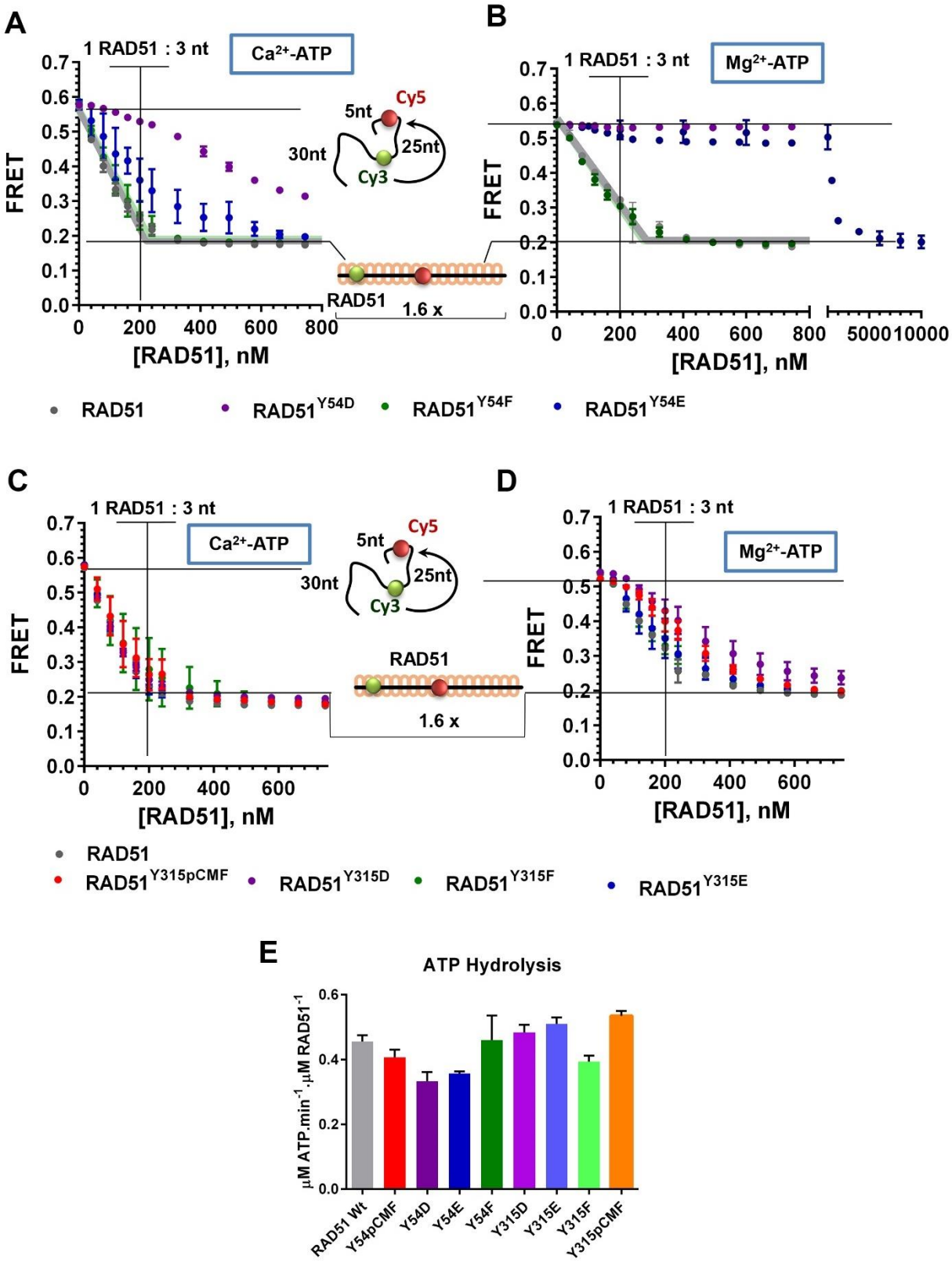


Figure 3.9. Altered ssDNA binding of RAD51 Y54 and Y315 mutants

Figure 3.9. Altered ssDNA binding of RAD51 Y54 and Y315 mutants **(A)** FRET based assay to measure binding for RAD51 Y54 mutants to ssDNA. Under conditions inhibiting ATP hydrolysis (Ca2+) RAD51Y54D & RAD51Y54E mutants deviate from ideal RAD51 binding behavior. **(B)** This effect is significantly exaggerated under conditions permitting ATP hydrolysis (Mg2+) **(C & D)** Similar FRET based experiments were conducted for Y315 mutants which were observed to have similar ssDNA binding characteristics under both, prohibitive conditions for ATP hydrolysis (Ca2+) as well as permissive conditions (Mg2+). Data is represented as mean ± SEM with n = 3. **(E)** Rates of ssDNA dependent ATP hydrolysis of all RAD51 mutants measured using the coupled NADH assay. Data is represented as mean ± SEM with n = 3.

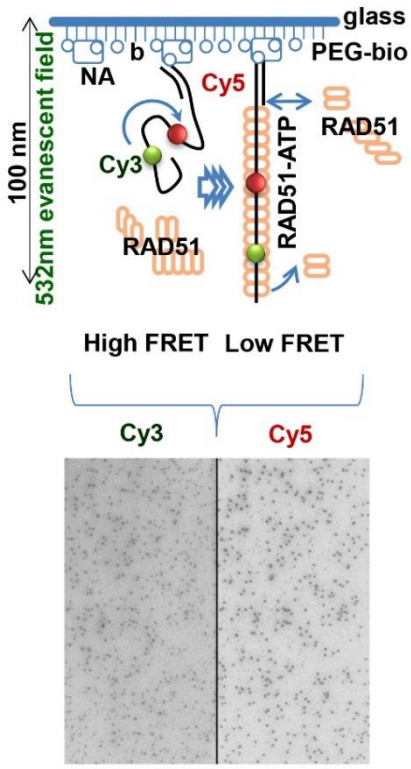


Figure 3.10. Single-Molecule TIRF ssDNA binding Assay (A) A Poly dT(60) ssDNA substrate is immobilized to a biotinylated quartz slide using the biotin-Neutravidin interaction (see methods). Evanescent wave produced by a prism-based total internal reflection (TIR) illumination and a 532nm laser is used to excite the Cy3 dye on the surface-tethered DNA substrate. In the absence of protein, the acceptor Cy5 dye on the partial duplex DNA is excited via FRET. Upon binding, RAD51 extends the ssDNA substrate leading to a decrease in energy transfer and reduction in Cy5 emission with a corresponding increase in Cy3 intensity. Both Cy3 and Cy5 emission can be tracked simultaneously using a dual view system. The change in Cy3 and Cy5 intensities on RAD51 binding to ssDNA substrate are tracked over time by recording movies over the course of the experiment.

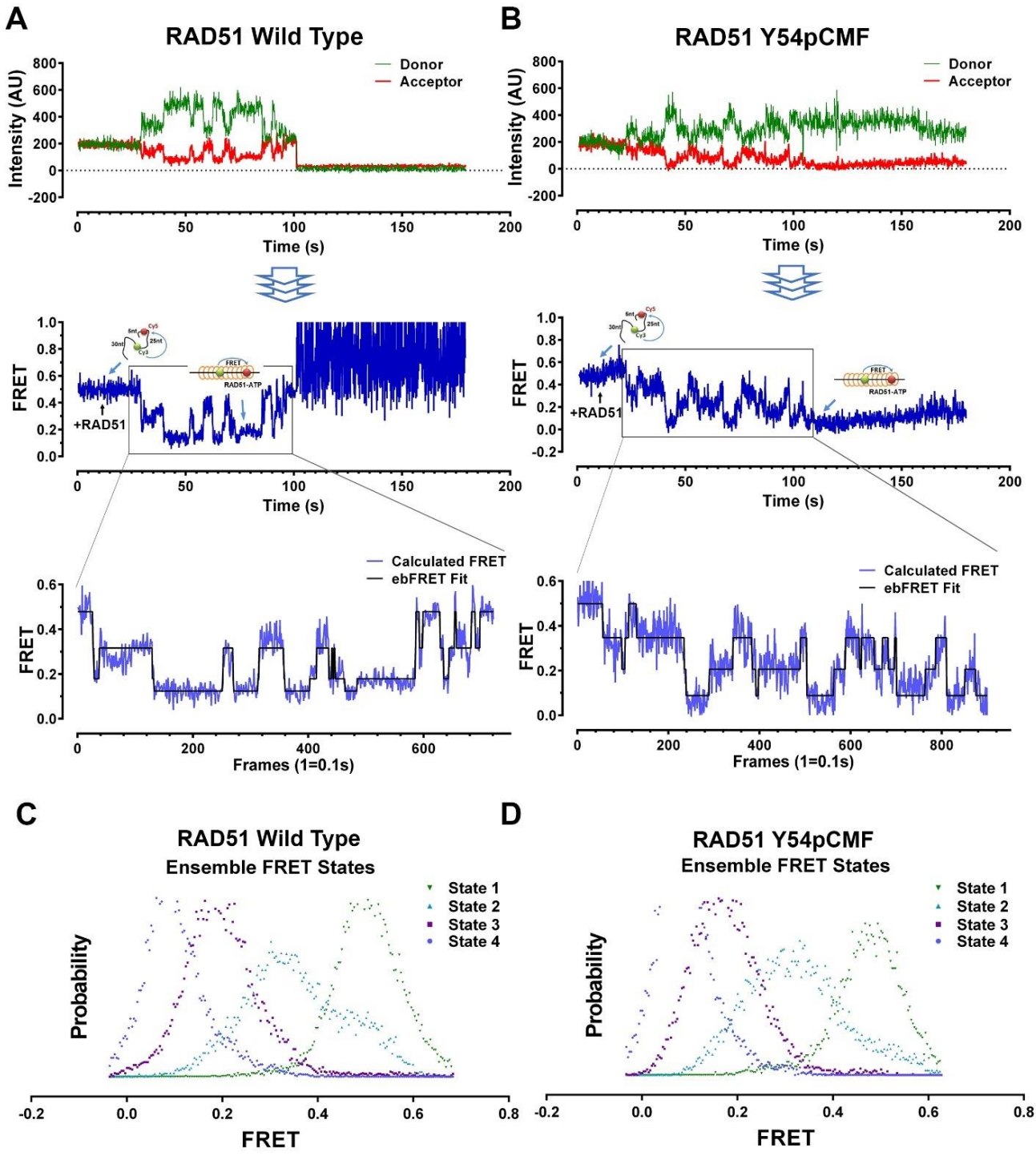


Figure 3.11. Hidden Markov Analysis shows that RAD51 and RAD51^{Y54pCMF} nucleate on ssDNA as dimers

Figure 3.11. Hidden Markov Analysis shows that RAD51 and RAD51^{Y54pCMF} nucleate on ssDNA as dimers **(A & B)** Representative Cy3 (donor) and Cy5 (acceptor) trajectories from pre-equilibrium experiments visualizing RAD51 **(A)** and RAD51^{Y54pCMF} **(B)** nucleation onto ssDNA in real time. RAD51 protein was introduced into the reaction chamber at ~10 seconds into a 3 minute recording. Trajectories that showed anti-correlated behavior were selected and corrected for background intensities as well as donor leakage followed by calculation of FRET trajectories. Regions containing transitions were trimmed and used for Hidden Markov analysis (ebFRET). The idealized trajectories are overlaid on the raw FRET trajectories showing the fit. . **(C &D)** Histograms showing distribution of observed FRET states from 281 and 179 molecules for RAD51 **(C)** and RAD51^{Y54pCMF} **(D)**. Four distinct FRET states were observed with FRET values between 0.5 and 0.1. Since the Cy3-Cy5 FRET pair is 21 nucleotides apart, this shows the RAD51 molecules bound ssDNA in three steps indicating each step as a dimer.

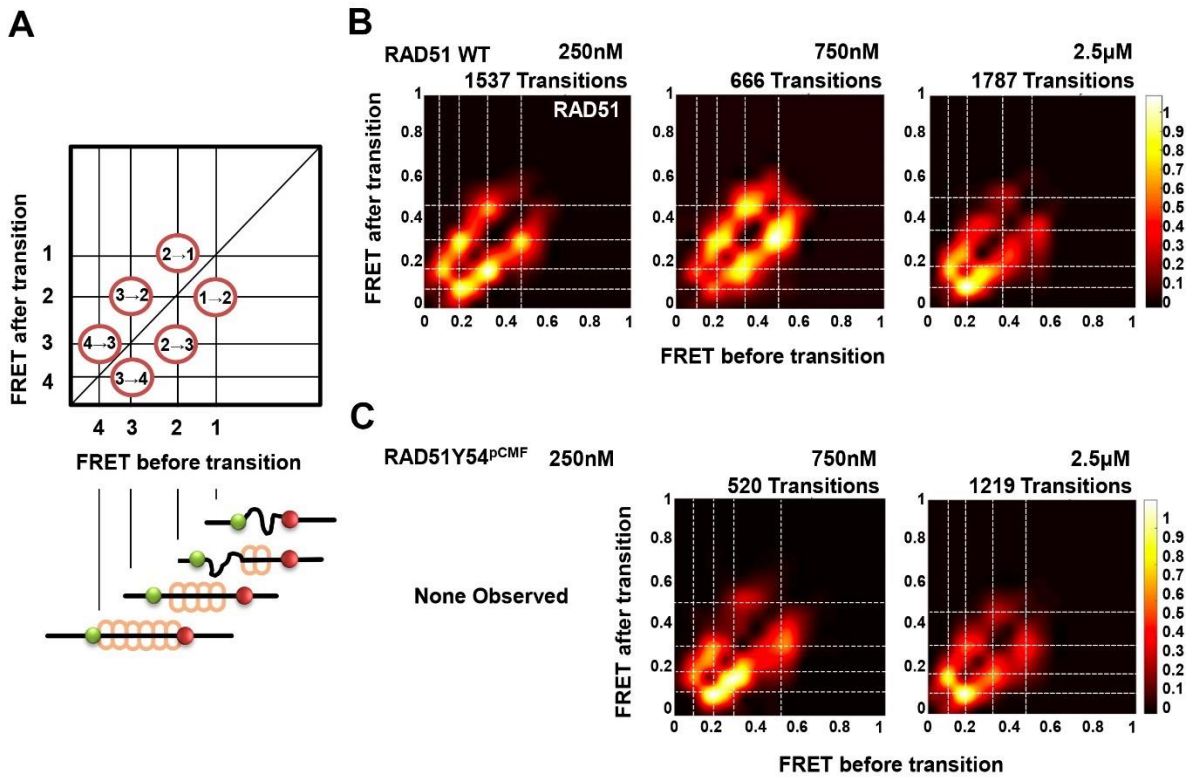


Figure 3.12. Formation of the dynamic RAD51 nuclei vs directional nuclei formation by RAD51^{Y54pCMF}. (A) Schematic explaining the 3D Transition Density Plots (TDP). Four distinct FRET states 1, 2, 3 and 4 correspond to respective states in the RAD51 nucleus formation on ssDNA with three steps between these states required to fully extend the ssDNA substrate. Transitions 1→2→3→4 correspond to the filament formation, while filament disassembly is reflected in transitions in the reverse direction. **(B)** TDPs showing the transition densities corresponding to each state for the wild type RAD51 protein at increasing concentrations. The transitions were calculated from 85, 50 and 146 molecules for 250nM, 750nM and 2.5μM concentrations respectively. Brighter colors represent more frequent transitions. The frequency scale is shown on the right of the graphs. **(C)** TDP plots showing transition densities for the RAD51^{Y54pCMF} mutant. No transitions were observed at 250nM concentrations. Transitions were calculated from 63 and 116 molecules for 750nM and 2.5μM concentrations, respectively.

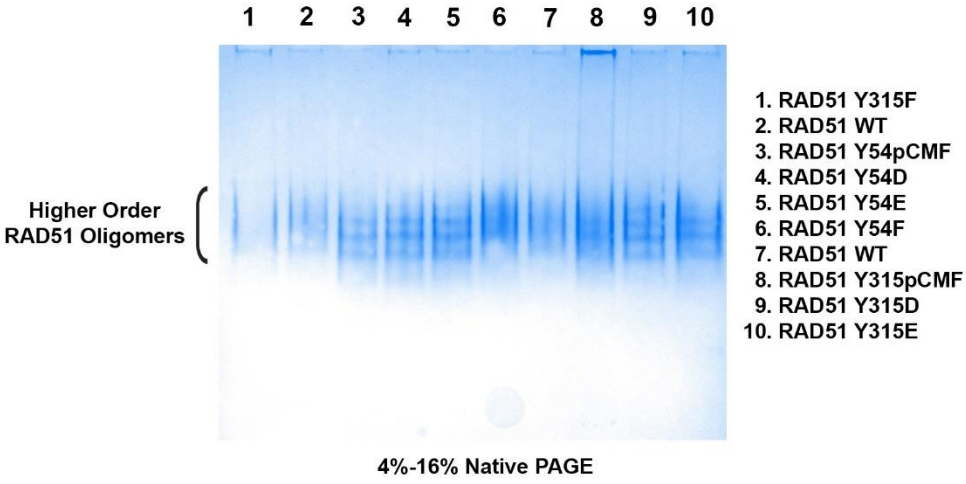


Figure 3.13. Effect of phosphomimetics on the RAD51 oligomerization Native-PAGE gel of all RAD51 mutants loaded at 7.5μM concentration similar to that used in strand exchange assays. All RAD51 Y54 mutants except RAD51^{Y54F} display increase in the electrophoretic mobility where the higher order oligomeric species are shifted to a lower molecular weight with more discrete bands.RAD51^{Y315D} & RAD51^{Y315E} mutants migrate higher compared to the Y54 mutants. A fraction of RAD51^{Y315pCMF} formed higher molecular weight aggregates that were slow to enter the gel.

CONCLUSIONS & FUTURE DIRECTION

In the work described in this dissertation, I have answered some fundamental questions regarding aspects of regulation of DNA repair by Homologous Recombination. My studies have focused mainly around the human RAD51 protein and its regulation in DNA repair. Using previously known structural information from the yeast Rad51 protein as well as information from a partial Human RAD51 crystal structure, I helped build and validate a homology model that explains the structural features of the complete RAD51 recombinase. Using this model, I was also able to describe new interactions between RAD51 protein and the BRCA2 tumor suppressor protein. These studies make significant contributions to the role of the BRCA2 mediator in RAD51 mediated DNA repair.

Using information from the homology model, I decided to probe the mechanism of regulation of the RAD51 protein by the c-Abl/BCR-ABL kinase protein. Biochemical information on this topic is very sparse and conflicts with observation made in cellular studies. Using a novel method to incorporate and artificial amino acid pCMF that is a more relevant replacement for the a phosphotyrosine residue, I was able to reconcile results with cellular data and for the first time, provide a complete understanding of how c-Abl phosphorylation affects RAD51 mechanistically using high resolution single molecule methods and traditional biochemical and biophysical approaches.

The results of my studies and methods developed by me, will be used to further understand the regulation of the RAD51 protein as well as study other mediators and antirecombinases involved in homologous recombination which will generate insights into one of the most critical steps in DNA repair and to find an ‘Achilles Heel’ in RAD51 function and exploit it to develop novel anticancer therapies.

BIBLIOGRAPHY

Aguilera, A., and Gaillard, H. (2014). Transcription and Recombination: When RNA Meets DNA. Cold Spring Harbor Perspectives in Biology *6*.

Amunugama, R., Fishel, R., and Paul, W.D. (2012a). Homologous Recombination in Eukaryotes. In Progress in Molecular Biology and Translational Science, P.W. Doetsch, ed. (The Boulevard, Langford Lane, Kidlington, Oxford, OX5 1GB, UK: Academic Press), pp. 155-206.

Amunugama, R., Groden, J., and Fishel, R. (2013). The HsRAD51B–HsRAD51C stabilizes the HsRAD51 nucleoprotein filament. DNA repair *12*, 723-732.

Amunugama, R., He, Y., Willcox, S., Forties, R.A., Shim, K.-S., Bundschuh, R., Luo, Y., Griffith, J., and Fishel, R. (2012b). RAD51 Protein ATP Cap Regulates Nucleoprotein Filament Stability. Journal of Biological Chemistry *287*, 8724-8736.

Aylon, Y., Liefshitz, B., and Kupiec, M. (2004). The CDK regulates repair of double-strand breaks by homologous recombination during the cell cycle. The EMBO Journal *23*, 4868-4875.

Baskaran, R., Wood, L.D., Whitaker, L.L., Canman, C.E., Morgan, S.E., Xu, Y., Barlow, C., Baltimore, D., Wynshaw-Boris, A., Kastan, M.B., *et al.* (1997). Ataxia telangiectasia mutant protein activates c-Abl tyrosine kinase in response to ionizing radiation. Nature *387*, 516-519.

Baumann, P., Benson, F.E., Hajibagheri, N., and West, S.C. (1997). Purification of human Rad51 protein by selective spermidine precipitation. Mutat Res *384*, 65-72.

Baumann, P., Benson, F.E., and West, S.C. (1996). Human Rad51 protein promotes ATP-dependent homologous pairing and strand transfer reactions in vitro. Cell *87*, 757-766.

Bignell, G., Micklem, G., Stratton, M.R., Ashworth, A., and Wooster, R. (1997). The BRC Repeats are Conserved in Mammalian BRCA2 Proteins. Human Molecular Genetics *6*, 53-58.

Bizard, A.H., and Hickson, I.D. (2014). The Dissolution of Double Holliday Junctions. Cold Spring Harbor Perspectives in Biology *6*.

Bork, P., Blomberg, N., and Nilges, M. (1996). Internal repeats in the BRCA2 protein sequence. Nat Genet *13*, 22-23.

Brown, S.P., and Muchmore, S.W. (2009). Large-scale application of high-throughput molecular mechanics with Poisson-Boltzmann surface area for routine physics-based scoring of protein-ligand complexes. J Med Chem *52*, 3159-3165.

Budke, B., Logan, H.L., Kalin, J.H., Zelivianskaia, A.S., Cameron McGuire, W., Miller, L.L., Stark, J.M., Kozikowski, A.P., Bishop, D.K., and Connell, P.P. (2012). RI-1: a chemical inhibitor of RAD51 that disrupts homologous recombination in human cells. Nucleic Acids Res.

Bugreev, D.V., and Mazin, A.V. (2004). Ca2+ activates human homologous recombination protein Rad51 by modulating its ATPase activity. Proceedings of the National Academy of Sciences of the United States of America *101*, 9988-9993.

Candelli, A., Holthausen, J.T., Depken, M., Brouwer, I., Franker, M.A., Marchetti, M., Heller, I., Bernard, S., Garcin, E.B., Modesti, M., *et al.* (2014). Visualization and quantification of nascent RAD51 filament formation at single-monomer resolution. Proceedings of the National Academy of Sciences of the United States of America *111*, 15090-15095.

Carreira, A., Hilario, J., Amitani, I., Baskin, R.J., K, M.K., Venkitaraman, A.R., and Kowalczykowski, S.C. (2010). The BRC Repeats of BRCA2 Modulate the DNA Binding Selectivity of RAD51. Cancer Cell *136*, 1032-1043.

Carreira, A., Hilario, J., Amitani, I., Baskin, R.J., Shivji, M.K., Venkitaraman, A.R., and Kowalczykowski, S.C. (2009). The BRC repeats of BRCA2 modulate the DNA-binding selectivity of RAD51. Cell *136*, 1032-1043.

Carreira, A., and Kowalczykowski, S.C. (2011). Two classes of BRC repeats in BRCA2 promote RAD51 nucleoprotein filament function by distinct mechanisms. Proceedings of the National Academy of Sciences of the United States of America *108*, 10448-10453.

Chen, J.-J., Silver, D., Cantor, S., Livingston, D.M., and Scully, R. (1999). BRCA1, BRCA2, and Rad51 Operate in a Common DNA Damage Response Pathway. Cancer Research *59*, 1752s-1756s.

Chen, P.L., Chen, C.F., Chen, Y., Xiao, J., Sharp, Z.D., and Lee, W.H. (1998). The BRC repeats in BRCA2 are critical for RAD51 binding and resistance to methyl methanesulfonate treatment. Proceedings of the National Academy of Sciences of the United States of America *95*, 5287-5292.

Chen, Z., Yang, H., and Pavletich, N.P. (2008). Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. Nature *453*, 489-484.

Chiruvella, K.K., Liang, Z., and Wilson, T.E. (2013). Repair of Double-Strand Breaks by End Joining. Cold Spring Harbor Perspectives in Biology *5*, a012757.

Ciccia, A., and Elledge, S.J. (2010). The DNA Damage Response: Making It Safe to Play with Knives. Molecular cell *40*, 179-204.

Colicelli, J. (2010). ABL tyrosine kinases: evolution of function, regulation, and specificity. Science signaling *3*, re6.

Conway, A.B., Lynch, T.W., Zhang, Y., Fortin, G.S., Fung, C.W., Symington, L.S., and Rice, P.A. (2004). Crystal structure of a Rad51 filament. Nature structural & molecular biology *11*, 791-796.

Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., and Kollman, P.A. (1995). A 2nd Generation Force-Field for the Simulation of Proteins, Nucleic-Acids, and Organic-Molecules. J Am Chem Soc *117*, 5179-5197.

Costanzo, V. (2011). Brca2, Rad51 and Mre11: performing balancing acts on replication forks. DNA repair *10*, 1060-1065.

Couedel, C., Mills, K.D., Barchi, M., Shen, L., Olshen, A., Johnson, R.D., Nussenzweig, A., Essers, J., Kanaar, R., Li, G.C., *et al.* (2004). Collaboration of homologous recombination and nonhomologous end-joining factors for the survival and integrity of mice and cells. Genes & development *18*, 1293-1304.

Daley, J.M., Gaines, W.A., Kwon, Y., and Sung, P. (2014a). Regulation of DNA Pairing in Homologous Recombination. Cold Spring Harbor Perspectives in Biology *6*, a017954.

Daley, J.M., Gaines, W.A., Kwon, Y., and Sung, P. (2014b). Regulation of DNA Pairing in Homologous Recombination. Cold Spring Harbor perspectives in biology.

Darden, T., York, D., and Pedersen, L. (1993). Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. J Chem Phys *98*, 10089-10092.

Doksani, Y., and de Lange, T. (2014). The Role of Double-Strand Break Repair Pathways at Functional and Dysfunctional Telomeres. Cold Spring Harbor Perspectives in Biology *6*.

Essers, J., van Steeg, H., de Wit, J., Swagemakers, S.M.A., Vermeij, M., Hoeijmakers, J.H.J., and Kanaar, R. (2000). Homologous and non-homologous recombination differentially affect DNA damage repair in mice. The EMBO Journal *19*, 1703-1710.

Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H., and Pedersen, L.G. (1995). A Smooth Particle Mesh Ewald Method. J Chem Phys *103*, 8577-8593.

Forget, A.L., and Kowalczykowski, S.C. (2012). Single-molecule imaging of DNA pairing by RecA reveals a three-dimensional homology search. Nature *482*, 423-427.

Franchitto, A., and Pichierri, P. (2002). Bloom's syndrome protein is required for correct relocalization of RAD50/MRE11/NBS1 complex after replication fork arrest. The Journal of cell biology *157*, 19-30.

Galkin, V.E., Esashi, F., Yu, X., Yang, S., West, S.C., and Egelman, E.H. (2005). BRCA2 BRC motifs bind RAD51-DNA filaments. Proceedings of the National Academy of Sciences of the United States of America *102*, 8537-8542.

Game, J.C. (1983). Radiation-Sensitive Mutants and Repair in Yeast. In Yeast Genetics: Fundamental and Applied Aspects, J.F.T. Spencer, D.M. Spencer, and A.R.W. Smith, eds. (New York, NY: Springer New York), pp. 109-137.

Grimme, J.M., Honda, M., Wright, R., Okuno, Y., Rothenberg, E., Mazin, A.V., Ha, T., and Spies, M. (2010). Human Rad52 binds and wraps single-stranded DNA and mediates annealing via two hRad52-ssDNA complexes. Nucleic Acids Research *38*, 2917-2930.

Grimme, J.M., and Spies, M. (2011). FRET-based assays to monitor DNA binding and annealing by Rad52 recombination mediator protein. Methods Mol Biol *745*, 463-483.

Ha, T. (2001). Single-Molecule Fluorescence Resonance Energy Transfer. Methods *25*, 78-86.

Head, M.S. (2010). Docking: a domesday report. In: Kenneth M. Merz, Jr et al. (eds.) Drug Design. pp. 98-119. (Cambridge University Press).

Henricksen, L.A., Umbricht, C.B., and Wold, M.S. (1994). Recombinant replication protein A: expression, complex formation, and functional characterization. The Journal of biological chemistry *269*, 11121-11132.

Heyer, W.-D. (2015). Regulation of Recombination and Genomic Maintenance. Cold Spring Harbor Perspectives in Biology *7*.

Heyer, W.-D., Ehmsen, K.T., and Liu, J. (2010). Regulation of Homologous Recombination in Eukaryotes. Annual review of genetics *44*, 113-139.

Hoeijmakers , J.H.J. (2009). DNA Damage, Aging, and Cancer. New England Journal of Medicine *361*, 1475-1485.

Holloman, W.K. (2011). Unraveling the mechanism of BRCA2 in homologous recombination. Nature structural & molecular biology *18*, 748-754.

Holthausen, J.T., Wyman, C., and Kanaar, R. (2011). Regulation of DNA strand exchange in homologous recombination. DNA repair *9*, 1264-1272.

Honda, M., Okuno, Y., Hengel, S.R., Martín-López, J.V., Cook, C.P., Amunugama, R., Soukup, R.J., Subramanyam, S., Fishel, R., and Spies, M. (2014). Mismatch repair protein hMSH2–hMSH6 recognizes mismatches and forms sliding clamps within a D-loop recombination intermediate. Proceedings of the National Academy of Sciences *111*, E316-325.

Honda, M., Okuno, Y., Yoo, J., Ha, T., and Spies, M. (2011). Tyrosine phosphorylation enhances RAD52-mediated annealing by modulating its DNA binding. EMBO J *30*, 3368-3382.

Honda, M., Park, J., Pugh, R.A., Ha, T., and Spies, M. (2009). Single-molecule analysis reveals differential effect of ssDNA-binding proteins on DNA translocation by XPD helicase. Molecular cell *35*, 694-703.

Hooft, R.W., Vriend, G., Sander, C., and Abola, E.E. (1996). Errors in protein structures. Nature *381*, 272.

Huang, F., Mazina, O.M., Zentner, I.J., Cocklin, S., and Mazin, A.V. (2012). Inhibition of Homologous Recombination in Human Cells by Targeting RAD51 Recombinase. Journal of Medicinal Chemistry *55*, 3011-3020.

Huang, F., Motlekar, N.A., Burgwin, C.M., Napper, A.D., Diamond, S.L., and Mazin, A.V. (2011). Identification of Specific Inhibitors of Human RAD51 Recombinase Using High-Throughput Screening. ACS Chem Biol.

Ira, G., Pellicioli, A., Balijja, A., Wang, X., Fiorani, S., Carotenuto, W., Liberi, G., Bressan, D., Wan, L., Hollingsworth, N.M., *et al.* (2004). DNA end resection, homologous recombination and DNA damage checkpoint activation require CDK1. Nature *431*, 1011-1017.

Islam, M.N., Paquet, N., Fox, D., Dray, E., Zheng, X.-F., Klein, H., Sung, P., and Wang, W. (2012). A Variant of the Breast Cancer Type 2 Susceptibility Protein (BRC) Repeat Is Essential for the RECQL5 Helicase to Interact with RAD51 Recombinase for Genome Stabilization. Journal of Biological Chemistry *287*, 23808-23818.

Iyama, T., and Wilson Iii, D.M. (2013). DNA repair mechanisms in dividing and non-dividing cells. DNA repair *12*, 620-636.

Jakalian, A., Jack, D.B., and Bayly, C.I. (2002). Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. J Comput Chem *23*, 1623-1641.

Jasin, M., and Rothstein, R. (2013a). Repair of Strand Breaks by Homologous Recombination. Cold Spring Harbor Perspectives in Biology *5*.

Jasin, M., and Rothstein, R. (2013b). Repair of strand breaks by homologous recombination. Cold Spring Harbor perspectives in biology *5*, a012740.

Jayathilaka, K., Sheridan, S.D., Bold, T.D., Bochenska, K., Logan, H.L., Weichselbaum, R.R., Bishop, D.K., and Connell, P.P. (2008). A chemical compound that stimulates the human homologous recombination protein RAD51. Proceedings of the National Academy of Sciences of the United States of America *105*, 15848-15853.

Jensen, R.B., Carreira, A., and Kowalczykowski, S.C. (2010). Purified human BRCA2 stimulates RAD51-mediated recombination. Nature *467*, 678-683.

Joo, C., and Ha, T. (2012a). Preparing Sample Chambers for Single-Molecule FRET. Cold Spring Harbor Protocols *2012*, pdb.prot071530.

Joo, C., and Ha, T. (2012b). Single-Molecule FRET with Total Internal Reflection Microscopy. Cold Spring Harbor Protocols *2012*, pdb.top072058.

Joo, C., McKinney, S.A., Nakamura, M., Rasnik, I., Myong, S., and Ha, T. (2006). Real-Time Observation of RecA Filament Dynamics with Single Monomer Resolution. Cell *126*, 515-527.

Juffer, A.H., Botta, E.F.F., Vankeulen, B.A.M., Vanderploeg, A., and Berendsen, H.J.C. (1991). The Electric-Potential of a Macromolecule in a Solvent - a Fundamental Approach. Journal of Computational Physics *97*, 144-171.

Kanaar, R., Hoeijmakers, J.H., and van Gent, D.C. (1998). Molecular mechanisms of DNA double strand break repair. Trends in cell biology *8*, 483-489.

Kass, E.M., and Jasin, M. (2010). Collaboration and competition between DNA double-strand break repair pathways. FEBS letters *584*, 3703-3708.

Keeney, S. (2008). Spo11 and the Formation of DNA Double-Strand Breaks in Meiosis. In Recombination and Meiosis: Crossing-Over and Disjunction, R. Egel, and D.-H. Lankenau, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 81-123.

Kisker, C., Kuper, J., and Van Houten, B. (2013). Prokaryotic Nucleotide Excision Repair. Cold Spring Harbor Perspectives in Biology *5*.

Kolodner, R. (1996). Biochemistry and genetics of eukaryotic mismatch repair. Genes & development *10*, 1433-1442.

Konagurthu, A.S., Whisstock, J.C., Stuckey, P.J., and Lesk, A.M. (2006). MUSTANG: a multiple structural alignment algorithm. Proteins *64*, 559-574.

Kowalczykowski, S.C., and Krupp, R.A. (1987). Effects of Escherichia-Coli Ssb Protein on the Single-Stranded DNA-Dependent Atpase Activity of Escherichia-Coli Reca Protein - Evidence That Ssb Protein Facilitates the Binding of Reca Protein to Regions of Secondary Structure within Single-Stranded-DNA. Journal of molecular biology *193*, 97-113.

Kreuzer, K.N., and Jongeneel, C.V. (1983). Escherichia coli phage T4 topoisomerase. In Methods in enzymology, L.G.K.M. Ray Wu, ed. (Academic Press), pp. 144-160.

Krieger, E. YASARA Biosciences GmbH, Wagramer Strasse 25/3/45, 1220 Vienna, Austria / Europe, www.yasara.org. In YASARA Biosciences GmbH, Wagramer Strasse 25/3/45, 1220 Vienna, Austria / Europe, wwwyasaraorg (YASARA Biosciences GmbH, Wagramer Strasse 25/3/45, 1220 Vienna, Austria / Europe, www.yasara.org).

Krieger, E., Joo, K., Lee, J., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D., and Karplus, K. (2009). Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. Proteins: Structure, Function, and Bioinformatics *77*, 114-122.

Krieger, E., Nielsen, J.E., Spronk, C.A., and Vriend, G. (2006). Fast empirical pKa prediction by Ewald summation. J Mol Graph Model *25*, 481-486.

Krokan, H.E., and Bjørås, M. (2013). Base Excision Repair. Cold Spring Harbor Perspectives in Biology *5*.

Labute, P. (2010). LowModeMD—Implicit Low-Mode Velocity Filtering Applied to Conformational Search of Macrocycles and Protein Loops. Journal of Chemical Information and Modeling *50*, 792-800.

Lam, I., and Keeney, S. (2015). Mechanism and Regulation of Meiotic Recombination Initiation. Cold Spring Harbor Perspectives in Biology *7*.

Lee, J.H., and Paull, T.T. (2007). Activation and regulation of ATM kinase activity in response to DNA double-strand breaks. Oncogene *26*, 7741-7748.

Lee, J.Y., Terakawa, T., Qi, Z., Steinfeld, J.B., Redding, S., Kwon, Y., Gaines, W.A., Zhao, W., Sung, P., and Greene, E.C. (2015). Base triplet stepping by the Rad51/RecA family of recombinases. Science *349*, 977-981.

Li, X., and Heyer, W.D. (2008a). Homologous recombination in DNA repair and DNA damage tolerance. Cell Res *18*, 99-113.

Li, X., and Heyer, W.D. (2008b). Homologous recombination in DNA repair and DNA damage tolerance. Cell research *18*, 99-113.

Li, X., and Heyer, W.D. (2009). RAD54 controls access to the invading 3'-OH end after RAD51-mediated DNA strand invasion in homologous recombination in Saccharomyces cerevisiae. Nucleic acids research *37*, 638-646.

Li, X., Zhang, X.-P., Solinger, J.A., Kiianitsa, K., Yu, X., Egelman, E.H., and Heyer, W.-D. (2007). Rad51 and Rad54 ATPase activities are both required to modulate Rad51-dsDNA filament dynamics. Nucleic acids research *35*, 4124-4140.

LiCata, V.J., and Wowor, A.J. (2008). Applications of fluorescence anisotropy to the study of protein-DNA interactions. Methods in cell biology *84*, 243-262.

Lim, D.S., and Hasty, P. (1996). A mutation in mouse rad51 results in an early embryonic lethal that is suppressed by a mutation in p53. Mol Cell Biol *16*, 7133-7143.

Lim, W.A., and Pawson, T. (2010). Phosphotyrosine Signaling: Evolving a New Cellular Communication System. Cell *142*, 661-667.

Lin, Z., Kong, H., Nei, M., and Ma, H. (2006). Origins and evolution of the recA/RAD51 gene family: Evidence for ancient gene duplication and endosymbiotic gene transfer. Proceedings of the National Academy of Sciences *103*, 10328-10333.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. Nature *362*, 709-715.

Liu, J., Doty, T., Gibson, B., and Heyer, W.-D. (2010a). Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA. Nature structural & molecular biology *17*, 1260-1262.

Liu, J., Doty, T., Gibson, B., and Heyer, W.D. (2010b). Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA. Nature structural & molecular biology *17*, 1260-1262.

Maréchal, A., and Zou, L. (2013). DNA Damage Sensing by the ATM and ATR Kinases. Cold Spring Harbor Perspectives in Biology *5*.

Martinez, S.F., Renodon-Corniere, A., Nomme, J., Eveillard, D., Fleury, F., Takahashi, M., and Weigel, P. (2011). Targeting human Rad51 by specific DNA aptamers induces inhibition of homologous recombination. Biochimie *92*, 1832-1838.

Masuda-Ozawa, T., Hoang, T., Seo, Y.-S., Chen, L.-F., and Spies, M. (2013). Single-molecule sorting reveals how ubiquitylation affects substrate recognition and activities of FBH1 helicase. Nucleic acids research.

Masuda, T., Ito, Y., Terada, T., Shibata, T., and Mikawa, T. (2009). A Non-canonical DNA Structure Enables Homologous Recombination in Various Genetic Systems. Journal of Biological Chemistry *284*, 30230-30239.

Mazin, A.V., Mazina, O.M., Bugreev, D.V., and Rossi, M.J. (2010). Rad54, the motor of homologous recombination. DNA repair *9*, 286-302.

McKinney, S.A., Joo, C., and Ha, T. (2006). Analysis of Single-Molecule FRET Trajectories Using Hidden Markov Modeling. Biophysical Journal *91*, 1941-1951.

Mehta, A., and Haber, J.E. (2014). Sources of DNA Double-Strand Breaks and Models of Recombinational DNA Repair. Cold Spring Harbor Perspectives in Biology *6*, a016428.

Molecular Operating Environment (MOE) 2010.09, C.C.G.I., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7 (2012).

Morrical, S.W. (2015). DNA-Pairing and Annealing Processes in Homologous Recombination and Homology-Directed Repair. Cold Spring Harbor Perspectives in Biology *7*, a016444.

Moynahan, M.E., and Jasin, M. (2010). Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. Nat Rev Mol Cell Biol *11*, 196-207.

Murakami, K., and Trakselis, M.A. (2014). Nucleic Acid Polymerases, Vol 1, 1 edn (Springer-Verlag Berlin Heidelberg: Springer-Verlag Berlin Heidelberg).

Namsaraev, E.A., and Berg, P. (1998). Interaction of Rad51 with ATP and Mg2+ induces a conformational change in Rad51. Biochemistry *37*, 11932-11939.

Nathanson, K.L., Wooster, R., and Weber, B.L. (2001). Breast cancer genetics: what we know and what we need. Nature medicine *7*, 552-556.

New, J.H., Sugiyama, T., Zaitseva, E., and Kowalczykowski, S.C. (1998). Rad52 protein stimulates DNA strand exchange by Rad51 and replication protein A. Nature *391*, 407-410.

Nimonkar, A.V., Genschel, J., Kinoshita, E., Polaczek, P., Campbell, J.L., Wyman, C., Modrich, P., and Kowalczykowski, S.C. (2011). BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. Genes & development *25*, 350-362.

Nishinaka, T., Shinohara, A., Ito, Y., Yokoyama, S., and Shibata, T. (1998). Base pair switching by interconversion of sugar puckers in DNA extended by proteins of RecA-family: a model for homology search in homologous genetic recombination. Proceedings of the National Academy of Sciences of the United States of America *95*, 11071-11076.

Nomme, J., Renodon-Corniere, A., Asanomi, Y., Sakaguchi, K., Stasiak, A.Z., Stasiak, A., Norden, B., Tran, V., and Takahashi, M. (2011). Design of potent inhibitors of human RAD51 recombinase based on BRC motifs of BRCA2 protein: modeling and experimental validation of a chimera peptide. J Med Chem *53*, 5782-5791.

Oganesian, L., and Karlseder, J. (2011). Mammalian 5' C-Rich Telomeric Overhangs Are a Mark of Recombination-Dependent Telomere Maintenance. Molecular cell *42*, 224-236.

Ogawa, T., Yu, X., Shinohara, A., and Egelman, E.H. (1993). Similarity of the yeast RAD51 filament to the bacterial RecA filament. Science *259*, 1896-1899.

Pal, J., Bertheau, R., Buon, L., Qazi, A., Batchu, R.B., Bandyopadhyay, S., Ali-Fehmi, R., Beer, D.G., Weaver, D.W., Reis, R.S., *et al.* (2011). Genomic evolution in Barrett's adenocarcinoma

cells: critical roles of elevated hsRAD51, homologous recombination and Alu sequences in the genome. Oncogene.

Pellegrini, L., Yu, D.S., Lo, T., Anand, S., Lee, M., Blundell, T.L., and Venkitaraman, A.R. (2002). Insights into DNA recombination from the structure of a RAD51-BRCA2 complex. Nature *420*, 287-293.

Petermann, E., and Helleday, T. (2010). Pathways of mammalian replication fork restart. Nature reviews Molecular cell biology *11*, 683-687.

Petermann, E., Orta, M.L., Issaeva, N., Schultz, N., and Helleday, T. (2010). Hydroxyurea-stalled replication forks become progressively inactivated and require two different RAD51-mediated pathways for restart and repair. Molecular cell *37*, 492-502.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—A visualization system for exploratory research and analysis. Journal of Computational Chemistry *25*, 1605-1612.

Popova, M., Henry, S., and Fleury, F. (2011). Posttranslational Modifications of Rad51 Protein and Its Direct Partners: Role and Effect on Homologous Recombination – Mediated DNA Repair. In DNA repair (InTech).

Popova, M., Shimizu, H., Yamamoto, K., Lebechec, M., Takahashi, M., and Fleury, F. (2009). Detection of c-Abl kinase-promoted phosphorylation of Rad51 by specific antibodies reveals that Y54 phosphorylation is dependent on that of Y315. FEBS letters *583*, 1867-1872.

Prakash, R., Satory, D., Dray, E., Papusha, A., Scheller, J., Kramer, W., Krejci, L., Klein, H., Haber, J.E., Sung, P., *et al.* (2009). Yeast Mph1 helicase dissociates Rad51-made D-loops: implications for crossover control in mitotic recombination. Genes & development *23*, 67-79.

Prakash, R., Zhang, Y., Feng, W., and Jasin, M. (2015). Homologous Recombination and Human Health: The Roles of BRCA1, BRCA2, and Associated Proteins. Cold Spring Harbor Perspectives in Biology *7*.

Qi, Z., Redding, S., Lee, Ja Y., Gibb, B., Kwon, Y., Niu, H., Gaines, William A., Sung, P., and Greene, Eric C. (2015). DNA Sequence Alignment by Microhomology Sampling during Homologous Recombination. Cell *160*, 856-869.

Qiao, G.B., Wu, Y.L., Yang, X.N., Zhong, W.Z., Xie, D., Guan, X.Y., Fischer, D., Kolberg, H.C., Kruger, S., and Stuerzbecher, H.W. (2005). High-level expression of Rad51 is an independent prognostic marker of survival in non-small-cell lung cancer patients. Br J Cancer *93*, 137-143.

Ravindranathan, K., Tirado-Rives, J., Jorgensen, W.L., and Guimarães, C.R.W. (2011). Improving MM-GB/SA Scoring through the Application of the Variable Dielectric Model. Journal of Chemical Theory and Computation *7*, 3859-3865.

Reuter, M., Zelensky, A., Smal, I., Meijering, E., van Cappellen, W.A., de Gruiter, H.M., van Belle, G.J., van Royen, M.E., Houtsmuller, A.B., Essers, J., *et al.* (2014). BRCA2 diffuses as oligomeric clusters with RAD51 and changes mobility after DNA damage in live cells. The Journal of cell biology *207*, 599-613.

Ristic, D., Modesti, M., van der Heijden, T., van Noort, J., Dekker, C., Kanaar, R., and Wyman, C. (2005). Human Rad51 filaments on double- and single-stranded DNA: correlating regular and irregular forms with recombination function. Nucleic acids research *33*, 3292-3302.

Roy, R., Hohng, S., and Ha, T. (2008). A practical guide to single-molecule FRET. Nat Meth *5*, 507-516.

Schärer, O.D. (2013). Nucleotide Excision Repair in Eukaryotes. Cold Spring Harbor Perspectives in Biology *5*.

Schipler, A., and Iliakis, G. (2013). DNA double-strand–break complexity levels and their possible contributions to the probability for error-prone processing and repair pathway choice. Nucleic acids research *41*, 7589-7605.

Schutz, C.N., and Warshel, A. (2001). What are the dielectric "constants" of proteins and how to validate electrostatic models? Proteins: Structure, Function, and Bioinformatics *44*, 400-417.

Scott, D.E., Ehebauer, M.T., Pukala, T., Marsh, M., Blundell, T.L., Venkitaraman, A.R., Abell, C., and Hyvönen, M. (2013). Using a Fragment-Based Approach To Target Protein–Protein Interactions. ChemBioChem *14*, 332-342.

Shafman, T., Khanna, K.K., Kedar, P., Spring, K., Kozlov, S., Yen, T., Hobson, K., Gatei, M., Zhang, N., Watters, D., *et al.* (1997). Interaction between ATM protein and c-Abl in response to DNA damage. Nature *387*, 520-523.

Shahid, T., Soroka, J., Kong, E.H., Malivert, L., McIlwraith, M.J., Pape, T., West, S.C., and Zhang, X. (2014). Structure and mechanism of action of the BRCA2 breast cancer tumor suppressor. Nature structural & molecular biology *21*, 962-968.

Shammas, M.A., Shmookler Reis, R.J., Koley, H., Batchu, R.B., Li, C., and Munshi, N.C. (2009). Dysfunctional homologous recombination mediates genomic instability and progression in myeloma. Blood *113*, 2290-2297.

Shibata, T., Nishinaka, T., Mikawa, T., Aihara, H., Kurumizaka, H., Yokoyama, S., and Ito, Y. (2001). Homologous genetic recombination as an intrinsic dynamic property of a DNA structure induced by RecA/Rad51-family proteins: a possible advantage of DNA over RNA as genomic material. Proceedings of the National Academy of Sciences of the United States of America *98*, 8425-8432.

Shim, K.S., Schmutte, C., Yoder, K., and Fishel, R. (2006). Defining the salt effect on human RAD51 activities. DNA repair *5*, 718-730.

Shimizu, H., Popova, M., Fleury, F., Kobayashi, M., Hayashi, N., Sakane, I., Kurumizaka, H., Venkitaraman, A.R., Takahashi, M., and Yamamoto, K. (2009). c-ABL tyrosine kinase stabilizes RAD51 chromatin association. Biochemical and biophysical research communications *382*, 286-291.

Shinohara, A., Ogawa, H., Matsuda, Y., Ushio, N., Ikeo, K., and Ogawa, T. (1993). Cloning of human, mouse and fission yeast recombination genes homologous to RAD51 and recA. Nat Genet *4*, 239-243.

Shirts, M.R., Mobley, D.L., and Brown, S.P. (2010). Drug design : structure- and ligand-based approaches

(Cambridge [U.K.] ; New York: Cambridge University Press).

Short, S.C., Giampieri, S., Worku, M., Alcaide-German, M., Sioftanos, G., Bourne, S., Lio, K.I., Shaked-Rabi, M., and Martindale, C. (2011). Rad51 inhibition is an effective means of targeting DNA repair in glioma models and CD133+ tumor derived cells. Neuro Oncol.

Sigurdsson, S., Trujillo, K., Song, B., Stratton, S., and Sung, P. (2001). Basis for avid homologous DNA strand exchange by human Rad51 and RPA. The Journal of biological chemistry *276*, 8798-8806.

Simandlova, J., Zagelbaum, J., Payne, M.J., Chu, W.K., Shevelev, I., Hanada, K., Chatterjee, S., Reid, D.A., Liu, Y., Janscak, P., *et al.* (2013). FBH1 disrupts RAD51 filaments in vitro and modulates homologous recombination in mammalian cells. Journal of Biological Chemistry.

Slupianek, A., Dasgupta, Y., Ren, S.Y., Gurdek, E., Donlin, M., Nieborowska-Skorska, M., Fleury, F., and Skorski, T. (2011). Targeting RAD51 phosphotyrosine-315 to prevent unfaithful recombination repair in BCR-ABL1 leukemia. Blood *118*, 1062-1068.

Slupianek, A., Schmutte, C., Tombline, G., Nieborowska-Skorska, M., Hoser, G., Nowicki, M.O., Pierce, A.J., Fishel, R., and Skorski, T. (2001). BCR/ABL regulates mammalian RecA homologs, resulting in drug resistance. Molecular cell *8*, 795-806.

Sorensen, C.S., Hansen, L.T., Dziegielewski, J., Syljuasen, R.G., Lundin, C., Bartek, J., and Helleday, T. (2005). The cell-cycle checkpoint kinase Chk1 is required for mammalian homologous recombination repair. Nat Cell Biol *7*, 195-201.

Sotriffer, C., and Matter, H. (2011). The Challenge of Affinity Prediction: Scoring Functions for Structure-Based Virtual Screening. In Virtual Screening: Principles, Challenges, and Practical Guidelines, C. Sotriffer, ed. (Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA), pp. 177-221.

Soulas-Sprauel, P., Rivera-Munoz, P., Malivert, L., Le Guyader, G., Abramowski, V., Revy, P., and de Villartay, J.P. (2007). V(D)J and immunoglobulin class switch recombinations: a paradigm to study the regulation of DNA end-joining. Oncogene *26*, 7780-7791.

Spies, M., and Fishel, R. (2015). Mismatch Repair during Homologous and Homeologous Recombination. Cold Spring Harbor Perspectives in Biology *7*.

Spies, M., and Kowalczykowski, S.C. (2006). The RecA Binding Locus of RecBCD Is a General Domain for Recruitment of DNA Strand Exchange Proteins. Molecular Cell *21*, 573-580.

Steinbrecher, T., and Labahn, A. (2010). Towards accurate free energy calculations in ligand protein-binding studies. Curr Med Chem *17*, 767-785.

Subramanyam, S., Jones, W.T., Spies, M., and Spies, M.A. (2013). Contributions of the RAD51 N-terminal domain to BRCA2-RAD51 interaction. Nucleic acids research *41*, 9020-9032.

Sugiyama, T., Zaitseva, E.M., and Kowalczykowski, S.C. (1997). A single-stranded DNA-binding protein is needed for efficient presynaptic complex formation by the Saccharomyces cerevisiae Rad51 protein. The Journal of biological chemistry *272*, 7940-7945.

Sung, P. (1994). Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast RAD51 protein. Science *265*, 1241-1243.

Sung, P., and Robberson, D.L. (1995). DNA strand exchange mediated by a RAD51-ssDNA nucleoprotein filament with polarity opposite to that of RecA. Cell *82*, 453-461.

Surade, S., and Blundell, TomÂ L. (2012). Structural Biology and Drug Discovery of Difficult Targets: The Limits of Ligandability. Chemistry & Biology *19*, 42-50.

Suwaki, N., Klare, K., and Tarsounas, M. (2011). RAD51 paralogs: Roles in DNA damage signalling, recombinational repair and tumorigenesis. Seminars in Cell & Developmental Biology *22*, 898-905.

Symington, L.S. (2014). End Resection at Double-Strand Breaks: Mechanism and Regulation. Cold Spring Harbor Perspectives in Biology *6*.

Takizawa, Y., Kinebuchi, T., Kagawa, W., Yokoyama, S., Shibata, T., and Kurumizaka, H. (2004). Mutational analyses of the human Rad51-Tyr315 residue, a site for phosphorylation in leukaemia cells. Genes to cells : devoted to molecular & cellular mechanisms *9*, 781-790.

Tavtigian, S.V., Simard, J., Rommens, J., Couch, F., Shattuck-Eidens, D., Neuhausen, S., Merajver, S., Thorlacius, S., Offit, K., Stoppa-Lyonnet, D., *et al.* (1996). The complete BRCA2 gene and mutations in chromosome 13q-linked kindreds. Nat Genet *12*, 333-337.

Tham, K.-C., Hermans, N., Winterwerp, Herrie H.K., Cox, Michael M., Wyman, C., Kanaar, R., and Lebbink, Joyce H.G. (2013). Mismatch Repair Inhibits Homeologous Recombination via Coordinated Directional Unwinding of Trapped DNA Structures. Molecular cell *51*, 326-337.

Thorslund, T., McIlwraith, M.J., Compton, S.A., Lekomtsev, S., Petronczki, M., Griffith, J.D., and West, S.C. (2010a). The breast cancer tumor suppressor BRCA2 promotes the specific targeting of RAD51 to single-stranded DNA. Nature structural & molecular biology *17*, 1263-1265.

Thorslund, T., McIlwraith, M.J., Compton, S.A., Lekomtsev, S., Petronczki, M., Griffith, J.D., and West, S.C. (2010b). The breast cancer tumor suppressor BRCA2 promotes the specific targeting of RAD51 to single-stranded DNA. Nature structural & molecular biology *17*, 1263-1265.

Tombline, G., and Fishel, R. (2002). Biochemical characterization of the human RAD51 protein. I. ATP hydrolysis. The Journal of biological chemistry *277*, 14417-14425.

Tombline, G., Shim, K.S., and Fishel, R. (2002). Biochemical characterization of the human RAD51 protein. II. Adenosine nucleotide binding and competition. The Journal of biological chemistry *277*, 14426-14433.

Tsuzuki, T., Fujii, Y., Sakumi, K., Tominaga, Y., Nakao, K., Sekiguchi, M., Matsushiro, A., Yoshimura, Y., and MoritaT (1996). Targeted disruption of the Rad51 gene leads to lethality in embryonic mice. Proceedings of the National Academy of Sciences of the United States of America *93*, 6236-6240.

Turner, N., Tutt, A., and Ashworth, A. (2004). Hallmarks of 'BRCAness' in sporadic cancers. Nature reviews Cancer *4*, 814-819.

van de Meent, J.-W., Bronson, Jonathan E., Wiggins, Chris H., and Gonzalez Jr, Ruben L. (2014). Empirical Bayes Methods Enable Advanced Population-Level Analyses of Single-Molecule FRET Experiments. Biophysical Journal *106*, 1327-1337.

Vannier, J.-B., Pavicic-Kaltenbrunner, V., Petalcorin, Mark I.R., Ding, H., and Boulton, Simon J. (2012). RTEL1 Dismantles T Loops and Counteracts Telomeric G4-DNA to Maintain Telomere Integrity. Cell *149*, 795-806.

Wallace, S.S. (2014). Base excision repair: A critical player in many games. DNA repair *19*, 14-26.

Whitmore, L., and Wallace, B.A. (2004). DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. Nucleic acids research *32*, W668-W673.

Wiese, C., Dray, E., Groesser, T., San Filippo, J., Shi, I., Collins, D.W., Tsai, M.-S., Williams, G.J., Rydberg, B., Sung, P., *et al.* (2007). Promotion of Homologous Recombination and Genomic Stability by RAD51AP1 via RAD51 Recombinase Enhancement. Molecular cell *28*, 482-490.

Wyatt, H.D.M., and West, S.C. (2014). Holliday Junction Resolvases. Cold Spring Harbor Perspectives in Biology *6*, a023192.

Xie, J., Supekova, L., and Schultz, P.G. (2007). A genetically encoded metabolically stable analogue of phosphotyrosine in Escherichia coli. ACS Chem Biol *2*, 474-478.

Xing, X., and Bell, C.E. (2004). Crystal structures of Escherichia coli RecA in complex with MgADP and MnAMP-PNP. Biochemistry *43*, 16142-16152.

Yoshimura, Y., Morita, T., Yamamoto, A., and Matsushiro, A. (1993). Cloning and sequence of the human RecA-like gene cDNA. Nucleic acids research *21*, 1665.

Yuan, S.-S.F., Chang, H.-L., and Lee, E.Y.H.P. (2003). Ionizing radiation-induced Rad51 nuclear focus formation is cell cycle-regulated and defective in both ATM−/− and c-Abl−/− cells. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis *525*, 85-92.

Yuan, Z.-M., Huang, Y., Ishiko, T., Nakada, S., Utsugisawa, T., Kharbanda, S., Wang, R., Sung, P., Shinohara, A., Weichselbaum, R., *et al.* (1998). Regulation of Rad51 Function by c-Abl in Response to DNA Damage. Journal of Biological Chemistry *273*, 3799-3802.

Zauhar, R.J., Morgan, R.S., and Shaw, P.B. (1985). A Polarization Charge Model for the Computation of Macromolecular Electric-Fields. Biophysical Journal *47*, A21-A21.

Zelensky, A., Kanaar, R., and Wyman, C. (2014a). Mediators of Homologous DNA Pairing. Cold Spring Harbor Perspectives in Biology *6*, a016451.

Zelensky, A., Kanaar, R., and Wyman, C. (2014b). Mediators of Homologous DNA Pairing. Cold Spring Harbor perspectives in biology *6*.

APPENDIX: IMPORTANT DNA SEQUENCES RELATED TO RAD51 STUDIES

>Homo sapiens RAD51 homolog (RecA homolog, E. coli) (S. cerevisiae) (RAD51), transcript variant 1, CDS (SPIES LAB VERSION)

ATGGCAATGCAGATGCAGCTTGAAGCAAATGCAGATACTTCAGTGGAAGAAGAAAGCTTTGGCCCACAACCCATTTC
ACGGTTAGAGCAGTGTGGCATAAATGCCAACGATGTGAAGAAATTGGAAGAAGCTGGATTCCATACTGTGGAGGCTG
TTGCC**TAT**GCGCCAAAGAAGGAGCTAATAAATATTAAGGGAATTAGTGAAGCCAAAGCTGATAAAATTCTGGCTGAG
GCAGCTAAATTAGTTCCAATGGGTTTCACCACTGCAACTGAATTCCACCAAAGGCGGTCAGAGATCATACAGATTAC
TACTGGCTCCAAAGAGCTTGACAAACTACTTCAAGGTGGAATTGAGACTGGATCTATCACAGAAATGTTTGGAGAAT
TCCGAACTGGGAAGACCCAGATCTGTCATACGCTAGCTGTCACCTGCCAGCTTCCCATTGACCGGGGTGGAGGTGAA
GGAAAGGCCATGTACATTGACACTGAGGGTACCTTTAGGCCAGAACGGCTGCTGGCAGTGGCTGAGAGGTATGGTCT
CTCTGGCAGTGATGTCCTGGATAATGTAGCATATGCTCGAGCGTTCAACACAGACCACCAGACCCAGCTCCTTTATC
AAGCATCAGCCATGATGGTAGAATCTAGGTATGCACTGCTTATTGTAGACAGTGCCACCGCCCTTTACAGAACAGAC
TACTCGGGTCGAGGTGAGCTTTCAGCCAGGCAGATGCACTTGGCCAGGTTTCTGCGGATGCTTCTGCGACTCGCTGA
TGAGTTTGGTGTAGCAGTGGTAATCACTAATCAGGTGGTAGCTCAAGTGGATGGAGCAGCGATGTTTGCTGCTGATC
CCAAAAAACCTATTGGAGGAAATATCATCGCCCATGCATCAACAACCAGATTGTATCTGAGGAAAGGAAGAGGGGAA
ACCAGAATCTGCAAAATCTACGACTCTCCCTGTCTTCCTGAAGCTGAAGCTATGTTCGCCATTAATGCAGATGGAGT
GGGAGATGCCAAAGAC**TGA**

>gi|49168602|emb|CAG38796.1| RAD51 [Homo sapiens]

MAMQMQLEANADTSVEEESFGPQPISRLEQCGINANDVKKLEEAGFHTVEAVAYAPKKELINIKGISEAKADKILAE
AAKLVPMGFTTATEFHQRRSEIIQITTGSKELDKLLQGGIETGSITEMF**GEFRTGK**TQICHTLAVTCQLPIDRGGGE
GKAMYIDTEGTFRPERLLAVAERYGLSGSDVLDNVAYARAFNTDHQTQLLYQASAMMVES**RYALLIVDSATALYR**TD
YSGRGELSARQMHLARFLRMLLRLADEFGVAVVITNQVVAQ**VDGAAMFAADPK**KPIGGNIIAHASTTRLYLRKGRGE
TRICKIYDSPCLPEAEAMFAINADGVGDAKD*

GEFRTGK – Walker A Domain

RYALLIVDSATALYR – Walker B Domain

TDYSGRG – DNA Binding Loop 1

VDGAAMFAADPK – DNA Binding Loop 2

PKKP – PXXP Consensus Sequence for c-Abl SH3 Binding

YDSP – (p)YXXP Consensus Sequence for c-Abl SH2 Binding

PCLP – PXXP Consensus Sequence for c-Abl SH3 Binding

>RAD51 (OPT) Sequence Codon Optimzed for E. coli expression

ATGGCAATGCAGATGCAGCTGGAAGCGAACGCAGATACCTCAGTGGAAGAAGAAAGTTTTGGTCCGCAACCGATTAG
CCGCCTGGAACAGTGCGGTATCAACGCAAATGATGTGAAAAAGCTGGAAGAAGCTGGCTTTCATACCGTTGAAGCAG
TCGCTTATGCGCCGAAAAAGGAACTGATCAACATCAAGGGTATCAGCGAAGCCAAAGCAGACAAGATCCTGGCCGAA
GCGGCCAAACTGGTTCCGATGGGCTTTACCACGGCAACCGAATTTCATCAGCGTCGCAGCGAAATTATCCAAATTAC
CACGGGTTCTAAAGAACTGGATAAGCTGCTGCAGGGCGGTATTGAAACCGGCTCTATCACGGAAATGTTTGGTGAAT
TTCGCACCGGCAAAACGCAGATTTGCCACACCCTGGCGGTGACGTGTCAACTGCCGATCGATCGTGGCGGTGGCGAG
GGTAAAGCCATGTACATTGACACCGAAGGCACGTTTCGTCCGGAACGCCTGCTGGCTGTTGCGGAACGCTATGGTCT
GAGCGGCTCTGATGTGCTGGACAACGTTGCCTACGCACGTGCTTTCAATACCGATCATCAGACGCAACTGCTGTATC
AGGCGAGTGCCATGATGGTCGAATCCCGCTACGCGCTGCTGATTGTGGATAGCGCAACCGCTCTGTATCGTACGGAC
TACTCAGGTCGCGGCGAACTGTCGGCACGTCAAATGCACCTGGCTCGTTTTCTGCGCATGCTGCTGCGTCTGGCGGA
CGAATTTGGTGTTGCCGTGGTTATCACCAACCAGGTCGTGGCGCAAGTCGATGGCGCAGCTATGTTCGCGGCCGACC
CGAAAAAGCCGATTGGTGGCAATATTATCGCGCACGCCAGTACCACGCGCCTGTATCTGCGTAAAGGTCGCGGCGAA
ACCCGTATTTGCAAAATCTATGATTCCCCGTGTCTGCCGGAAGCCGAAGCCATGTTTGCCATCAATGCCGACGGTGT
TGGTGACGCTAAGGAT**TAA**

NcoI-RAD51o FP – AATAGCCC**ATG**GCAATGCAGATGCAGCTGG

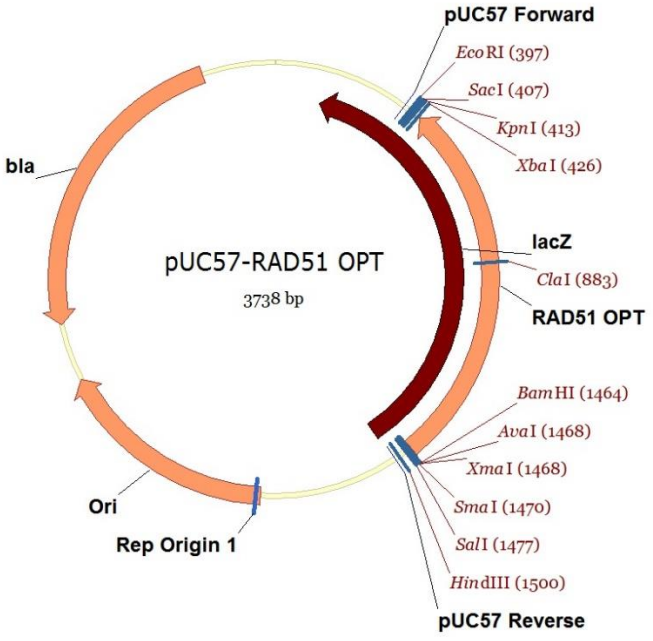
NotI-RAD51o RP – GCTATTGCGGCCGC**TTA**ATCCTTAGCGTCACCAACACCGTC

NcoI-RAD51o FPN - TAAATAGCCC**ATG**GCAATGCAGATGCAGCTGG

NotI-RAD51o RPN – ATGCTATTGCGGCCGC**TTA**ATCCTTAGCGTCACCAACACCGTC

RAD51o F1 - TCGCACCGGCAAAACGCAG

RAD51o R1 – TCAGAGCCGCTCAGACCATAGC



>pCOLADuet-1/GroE/RAD51_OPT (pCH1-RAD51Opt)

ggggaattgtgagcggataacaattcccctgtagaaataattttgtttaactttaataaggagatataCC**ATG**GCAA
TGCAGATGCAGCTGGAAGCGAACGCAGATACCTCAGTGGAAGAAGAAAGTTTTGGTCCGCAACCGATTAGCCGCCTG
GAACAGTGCGGTATCAACGCAAATGATGTGAAAAAGCTGGAAGAAGCTGGCTTTCATACCGTTGAAGCAGTCGCTTA
TGCGCCGAAAAAGGAACTGATCAACATCAAGGGTATCAGCGAAGCCAAAGCAGACAAGATCCTGGCCGAAGCGGCCA
AACTGGTTCCGATGGGCTTTACCACGGCAACCGAATTTCATCAGCGTCGCAGCGAAATTATCCAAATTACCACGGGT
TCTAAAGAACTGGATAAGCTGCTGCAGGGCGGTATTGAAACCGGCTCTATCACGGAAATGTTTGGTGAATTTCGCAC
CGGCAAAACGCAGATTTGCCACACCCTGGCGGTGACGTGTCAACTGCCGATCGATCGTGGCGGTGGCGAGGGTAAAG
CCATGTACATTGACACCGAAGGCACGTTTCGTCCGGAACGCCTGCTGGCTGTTGCGGAACGCTATGGTCTGAGCGGC
TCTGATGTGCTGGACAACGTTGCCTACGCACGTGCTTTCAATACCGATCATCAGACGCAACTGCTGTATCAGGCGAG
TGCCATGATGGTCGAATCCCGCTACGCGCTGCTGATTGTGGATAGCGCAACCGCTCTGTATCGTACGGACTACTCAG
GTCGCGGCGAACTGTCGGCACGTCAAATGCACCTGGCTCGTTTTCTGCGCATGCTGCTGCGTCTGGCGGACGAATTT
GGTGTTGCCGTGGTTATCACCAACCAGGTCGTGGCGCAAGTCGATGGCGCAGCTATGTTCGCGGCCGACCCGAAAAA
GCCGATTGGTGGCAATATTATCGCGCACGCCAGTACCACGCGCCTGTATCTGCGTAAAGGTCGCGGCGAAACCCGTA
TTTGCAAAATCTATGATTCCCCGTGTCTGCCGGAAGCCGAAGCCATGTTTGCCATCAATGCCGACGGTGTTGGTGAC
GCTAAGGAT**TAA**GCGGCCGCataatgcttaagtcgaacagaaagtaatcgtattgtacacggccgcataatcgaaat
taatacgactcactataggggaattgtgagcggataacaattccccatcttagtatattagttaagtataagaaggga
gatataCAT**ATG**AATATTCGTCCATTGCATGATCGCGTGATCGTCAAGCGTAAAGAAGTTGAAACTAAATCTGCTGG
CGGCATCGTTCTGACCGGCTCTGCAGCGGCTAAATCCACCCGCGGCGAAGTGCTGGCTGTCGGCAATGGCCGTATCC
TTGAAAATGGCGAAGTGAAGCCGCTGGATGTGAAAGTTGGCGACATCGTTATTTTCAACGATGGCTACGGTGTGAAA
TCTGAGAAGATCGACAATGAAGAAGTGTTGATCATGTCCGAAAGCGACATTCTGGCAATTGTTGAAGCG**TAA**TCCGC
GCACGACACTGAACATACGAATTTAAGGAATAAAGATA**ATG**GCAGCTAAAGACGTAAAATTCGGTAACGACGCTCGT
GTGAAAATGCTGCGCGGCGTAAACGTACTGGCAGATGCAGTGAAAGTTACCCTCGGTCCAAAAGGCCGTAACGTAGT
TCTGGATAAATCTTTCGGTGCACCGACCATCACCAAAGATGGTGTTTCCGTTGCTCGTGAAATCGAACTGGAAGACA
AGTTCGAAAATATGGGTGCGCAGATGGTGAAAGAAGTTGCCTCTAAAGCAAACGACGCTGCAGGCGACGGTACCACC
ACTGCAACCGTACTGGCTCAGGCTATCATCACTGAAGGTCTGAAAGCTGTTGCTGCGGGCATGAACCCGATGGACCT
GAAACGTGGTATCGACAAAGCGGTTACCGTTGCAGTTGAAGAACTGAAAGCGCTGTCCGTACCATGCTCTGACTCTA
AAGCGATTGCTCAGGTTGGTACCATCTCCGCTAACTCCGACGAAACCGTAGGTAAACTGATCGCTGAAGCGATGGAC
AAAGTCGGTAAAGAAGGCGTTATCACCGTTGAAGACGGTACCGGTCTGCAGGACGAACTGGACGTGGTTGAAGGTAT
GCAGTTCGACCGTGGCTACCTGTCTCCTTACTTCATCAACAAGCCGGAAACTGGCGCAGTAGAACTGGAAAGCCCGT
TCATCCTGCTGGCTGACAAGAAAATCTCCAACATCCGCGAAATGCTGCCGGTTCTGGAAGCTGTTGCCAAAGCAGGC
AAACCGCTGCTGATCATCGCTGAAGATGTAGAAGGCGAAGCGCTGGCAACTCTGGTTGTTAACACCATGCGTGGCAT
CGTGAAAGTCGCTGCGGTTAAAGCACCGGGCTTCGGCGATCGTCGTAAAGCTATGCTGCAGGATATCGCAACCCTGA
CTGGCGGTACCGTGATCTCTGAAGAGATCGGTATGGAGCTGGAAAAAGCAACCCTGGAAGACCTGGGTCAGGCTAAA
CGTGTTGTGATCAACAAAGACACCACCACTATCATCGATGGCGTGGGTGAAGAAGCTGCAATCCAGGGCCGTGTTGC
TCAGATCCGTCAGCAGATTGAAGAAGCAACTTCTGACTACGACCGTGAAAAACTGCAGGAACGCGTAGCGAAACTGG
CAGGCGGCGTTGCAGTTATCAAAGTGGGTGCTGCTACCGAAGTTGAAATGAAAGAGAAAAAAGCACGCGTTGAAGAT
GCCCTGCACGCGACCCGTGCTGCGGTAGAAGAAGGCGTGGTTGCTGGTGGTGGTGTTGCGCTGATCCGCGTAGCGTC

TAAACTGGCTGACCTGCGTGGTCAGAACGAAGACCAGAACGTGGGTATCAAAGTTGCACTGCGTGCAATGGAAGCTC
CGCTGCGTCAGATCGTATTGAACTGCGGCGAAGAACCGTCTGTTGTTGCTAACACCGTTAAAGGCGGCGACGGCAAC
TACGGTTACAACGCAGCAACCGAAGAATACGGCAACATGATCGACATGGGTATCCTGGATCCAACCAAAGTAACTCG
TTCTGCTCTGCAGTACGCAGCTTCTGTGGCTGGCCTGATGATCACCACCGAATGCATGGTTACCGACCTGCCGAAAA
ACGATGCAGCTGACTTAGGCGCTGCTGGCGGTATGGGCGGCATGGGTGGCATGGGCGGCATGATG**TAA**AGATCTcaa
ttggatatcggccggccacgcgatcgctgacgtcggtaccctcgagtctggtaaagaaaccgctgctgcgaaatttg
aacgccagcacatggactcgtctactagcgcagcttaattaacctaggctgctgccaccgctgagcaataactagca
taacccccttggggcctctaaacgggtcttgaggggtttttgctgaaacctcaggcatttgagaagcacacggtcac
actgcttccggtagtcaataaaccggtaaaccagcaatagacataagcggctatttaacgaccctgccctgaaccga
cgacaagctgacgaccgggtctccgcaagtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttt
ctaaatacattcaaatatgtatccgctcatgaattaattcttagaaaaactcatcgagcatcaaatgaaactgcaat
ttattcatatcaggattatcaataccatatttttgaaaaagccgtttctgtaatgaaggagaaaactcaccgaggca
gttccataggatggcaagatcctggtatcggtctgcgattccgactcgtccaacatcaatacaacctattaaGCCGC
tttcccctcgtcaaaaataaggttatcaagtgagaaatcaccatgagtgacgactgaatccggtgagaatggcaaaa
gtttatgcatttctttccagacttgttcaacaggccagccattacgctcgtcatcaaaatcactcgcatcaaccaaa
ccgttattcattcgtgattgcgcctgagcgagacgaaatacgcggtcgctgttaaaaggacaattacaaacaggaat
cgaatgcaaccggcgcaggaacactgccagcgcatcaacaatattttcacctgaatcaggatattcttctaatacct
ggaatgctgttttcccggggatcgcagtggtgagtaaccatgcatcatcaggagtacggataaaatgcttgatggtc
ggaagaggcataaattccgtcagccagtttagtctgaccatctcatctgtaacatcattggcaacgctacctttgcc
atgtttcagaaacaactctggcgcatcgggcttcccatacaatcgatagattgtcgcacctgattgcccgacattat
cgcgagcccatttatacccatataaatcagcatccatgttggaatttaatcgcggcctagagcaagacgtttcccgt
tgaatatggctcatactcttccttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacat
atttgaatgtatttagaaaaataaacaaataggcatgctagcgcagaaacgtcctagaagatgccaggaggatactt
agcagagagacaataaggccggagcgaagccgtttttccataggctccgccccccctgacgaacatcacgaaatctga
cgctcaaatcagtggtggcgaaacccgacaggactataaagataccaggcgtttccccctgatggctccctcttgcg
ctctcctgttcccgtcctgcggcgtccgtgttgtggtggaggctttacccaaatcaccacgtcccgttccgtgtaga
cagttcgctccaagctgggctgtgtgcaagaacccccgttcagcccgactgctgcgccttatccggtaactatcat
cttgagtccaacccggaaagacacgacaaaacgccactggcagcagccattggtaactgagaattagtggatttaga
tatcgagagtcttgaagtggtggcctaacagaggctacactgaaaggacagtatttggtatctgcgctccactaaag
ccagttaccaggttaagcagttccccaactgacttaaccttcgatcaaaccgcctccccaggcggtttttttcgttta
cagagcaggagattacgacgatcgtaaaaggatctcaagaagatcctttacggattcccgacaccatcactctagat
ttcagtgcaatttatctcttcaaatgtagcacctgaagtcagccccatacgatataagttgtaattctcatgttagt
catgccccgcgcccaccggaaggagctgactgggttgaaggctctcaagggcatcggtcgagatcccggtgcctaat
gagtgagctaacttacattaattgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcat
taatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgccagggtggtttttcttttcaccagtgaga
cgggcaacagctgattgcccttcaccgcctggccctgagagagttgcagcaagcggtccacgctggtttgccccagc
aggcgaaaatcctgtttgatggtggttaacggcgggatataacatgagctgtcttcggtatcgtcgtatcccactac
cgagatgtccgcaccaacgcgcagcccggactcggtaatggcgcgcattgcgcccagcgccatctgatcgttggcaa

ccagcatcgcagtgggaacgatgccctcattcagcatttgcatggtttgttgaaaaccggacatggcactccagtcg
ccttcccgttccgctatcggctgaatttgattgcgagtgagatatttatgccagccagccagacgcagacgcgccga
gacagaacttaatgggcccgctaacagcgcgatttgctggtgacccaatgcgaccagatgctccacgcccagtcgcg
taccgtcttcatgggagaaaataatactgttgatgggtgtctggtcagagacatcaagaaataacgccggaacatta
gtgcaggcagcttccacagcaatggcatcctggtcatccagcggatagttaatgatcagcccactgacgcgttgcgc
gagaagattgtgcaccgccgctttacaggcttcgacgccgcttcgttctaccatcgacaccaccacgctggcaccca
gttgatcggcgcgagatttaatcgccgcgacaatttgcgacggcgcgtgcagggccagactggaggtggcaacgcca
atcagcaacgactgtttgcccgccagttgttgtgccacgcggttgggaatgtaattcagctccgccatcgccgcttc
cactttttcccgcgttttcgcagaaacgtggctggcctggttcaccacgcgggaaacggtctgataagagacaccgg
catactctgcgacatcgtataacgttactggtttcacattcaccaccctgaattgactctcttccgggcgctatcat
gccataccgcgaaaggttttgcgccattcgatggtgtccgggatctcgacgctctcccttatgcgactcctgcatta
ggaaattaatacgactcactata

>MCS1-RAD51 OPT

*aatagcCC***ATG**GCAATGCAGATGCAGCTGGAAGCGAACGCAGATACCTCAGTGGAAGAAGAAAGTTTTGGTCCGCAA
CCGATTAGCCGCCTGGAACAGTGCGGTATCAACGCAAATGATGTGAAAAAGCTGGAAGAAGCTGGCTTTCATACCGT
TGAAGCAGTCGCTTATGCGCCGAAAAAGGAACTGATCAACATCAAGGGTATCAGCGAAGCCAAAGCAGACAAGATCC
TGGCCGAAGCGGCCAAACTGGTTCCGATGGGCTTTACCACGGCAACCGAATTTCATCAGCGTCGCAGCGAAATTATC
CAAATTACCACGGGTTCTAAAGAACTGGATAAGCTGCTGCAGGGCGGTATTGAAACCGGCTCTATCACGGAAATGTT
TGGTGAATTTCGCACCGGCAAAACGCAGATTTGCCACACCCTGGCGGTGACGTGTCAACTGCCGATCGATCGTGGCG
GTGGCGAGGGTAAAGCCATGTACATTGACACCGAAGGCACGTTTCGTCCGGAACGCCTGCTGGCTGTTGCGGAACGC
TATGGTCTGAGCGGCTCTGATGTGCTGGACAACGTTGCCTACGCACGTGCTTTCAATACCGATCATCAGACGCAACT
GCTGTATCAGGCGAGTGCCATGATGGTCGAATCCCGCTACGCGCTGCTGATTGTGGATAGCGCAACCGCTCTGTATC
GTACGGACTACTCAGGTCGCGGCGAACTGTCGGCACGTCAAATGCACCTGGCTCGTTTTCTGCGCATGCTGCTGCGT
CTGGCGGACGAATTTGGTGTTGCCGTGGTTATCACCAACCAGGTCGTGGCGCAAGTCGATGGCGCAGCTATGTTCGC
GGCCGACCCGAAAAAGCCGATTGGTGGCAATATTATCGCGCACGCCAGTACCACGCGCCTGTATCTGCGTAAAGGTC
GCGGCGAAACCCGTATTTGCAAAATCTATGATTCCCCGTGTCTGCCGGAAGCCGAAGCCATGTTTGCCATCAATGCC
GACGGTGTTGGTGACGCTAAGGAT**TAA***GCGGCCGC**aatagcat*

Cloning and sequencing primers

NcoI-RAD51o FP – AATAGCCC**ATG**GCAATGCAGATGCAGCTGG

NotI-RAD51o RP – GCTATTGCGGCCGC**TTA**ATCCTTAGCGTCACCAACACCGTC

NcoI-RAD51o FPN - TAAATAGCCC**ATG**GCAATGCAGATGCAGCTGG

NotI-RAD51o RPN – ATGCTATTGCGGCCGC**TTA**ATCCTTAGCGTCACCAACACCGTC

RAD51o F1 - TCGCACCGGCAAAACGCAG

RAD51o R1 – TCAGAGCCGCTCAGACCATAGC

>MCS2-GroE Operon

*aatagc*CAT**ATG**AATATTCGTCCATTGCATGATCGCGTGATCGTCAAGCGTAAAGAAGTTGAAACTAAATCTGCTGG
CGGCATCGTTCTGACCGGCTCTGCAGCGGCTAAATCCACCCGCGGCGAAGTGCTGGCTGTCGGCAATGGCCGTATCC
TTGAAAATGGCGAAGTGAAGCCGCTGGATGTGAAAGTTGGCGACATCGTTATTTTCAACGATGGCTACGGTGTGAAA
TCTGAGAAGATCGACAATGAAGAAGTGTTGATCATGTCCGAAAGCGACATTCTGGCAATTGTTGAAGCG**TAA**TCCGC
GCACGACACTGAACATACGAATTTAAGGAATAAAGATA**ATG**GCAGCTAAAGACGTAAAATTCGGTAACGACGCTCGT
GTGAAAATGCTGCGCGGCGTAAACGTACTGGCAGATGCAGTGAAAGTTACCCTCGGTCCAAAAGGCCGTAACGTAGT
TCTGGATAAATCTTTCGGTGCACCGACCATCACCAAAGATGGTGTTTCCGTTGCTCGTGAAATCGAACTGGAAGACA
AGTTCGAAAATATGGGTGCGCAGATGGTGAAAGAAGTTGCCTCTAAAGCAAACGACGCTGCAGGCGACGGTACCACC
ACTGCAACCGTACTGGCTCAGGCTATCATCACTGAAGGTCTGAAAGCTGTTGCTGCGGGCATGAACCCGATGGACCT
GAAACGTGGTATCGACAAAGCGGTTACCGTTGCAGTTGAAGAACTGAAAGCGCTGTCCGTACCATGCTCTGACTCTA
AAGCGATTGCTCAGGTTGGTACCATCTCCGCTAACTCCGACGAAACCGTAGGTAAACTGATCGCTGAAGCGATGGAC
AAAGTCGGTAAAGAAGGCGTTATCACCGTTGAAGACGGTACCGGTCTGCAGGACGAACTGGACGTGGTTGAAGGTAT
GCAGTTCGACCGTGGCTACCTGTCTCCTTACTTCATCAACAAGCCGGAAACTGGCGCAGTAGAACTGGAAAGCCCGT
TCATCCTGCTGGCTGACAAGAAAATCTCCAACATCCGCGAAATGCTGCCGGTTCTGGAAGCTGTTGCCAAAGCAGGC
AAACCGCTGCTGATCATCGCTGAAGATGTAGAAGGCGAAGCGCTGGCAACTCTGGTTGTTAACACCATGCGTGGCAT
CGTGAAAGTCGCTGCGGTTAAAGCACCGGGCTTCGGCGATCGTCGTAAAGCTATGCTGCAGGATATCGCAACCCTGA
CTGGCGGTACCGTGATCTCTGAAGAGATCGGTATGGAGCTGGAAAAAGCAACCCTGGAAGACCTGGGTCAGGCTAAA
CGTGTTGTGATCAACAAAGACACCACCACTATCATCGATGGCGTGGGTGAAGAAGCTGCAATCCAGGGCCGTGTTGC
TCAGATCCGTCAGCAGATTGAAGAAGCAACTTCTGACTACGACCGTGAAAAACTGCAGGAACGCGTAGCGAAACTGG
CAGGCGGCGTTGCAGTTATCAAAGTGGGTGCTGCTACCGAAGTTGAAATGAAAGAGAAAAAAGCACGCGTTGAAGAT
GCCCTGCACGCGACCCGTGCTGCGGTAGAAGAAGGCGTGGTTGCTGGTGGTGGTGTTGCGCTGATCCGCGTAGCGTC
TAAACTGGCTGACCTGCGTGGTCAGAACGAAGACCAGAACGTGGGTATCAAAGTTGCACTGCGTGCAATGGAAGCTC
CGCTGCGTCAGATCGTATTGAACTGCGGCGAAGAACCGTCTGTTGTTGCTAACACCGTTAAAGGCGGCGACGGCAAC
TACGGTTACAACGCAGCAACCGAAGAATACGGCAACATGATCGACATGGGTATCCTGGATCCAACCAAAGTAACTCG
TTCTGCTCTGCAGTACGCAGCTTCTGTGGCTGGCCTGATGATCACCACCGAATGCATGGTTACCGACCTGCCGAAAA
ACGATGCAGCTGACTTAGGCGCTGCTGGCGGTATGGGCGGCATGGGTGGCATGGGCGGCATGATG**TAA**AGATCT*aat
agc*

Cloning and Sequencing Primers

NdeI-GroE FP – AATAGCCAT**ATG**AATATTCGTCCATTGCATGATCG

BglII-GroE RP – GCTATTAGATCTTTACATCATGCCGCCCATGCCAC

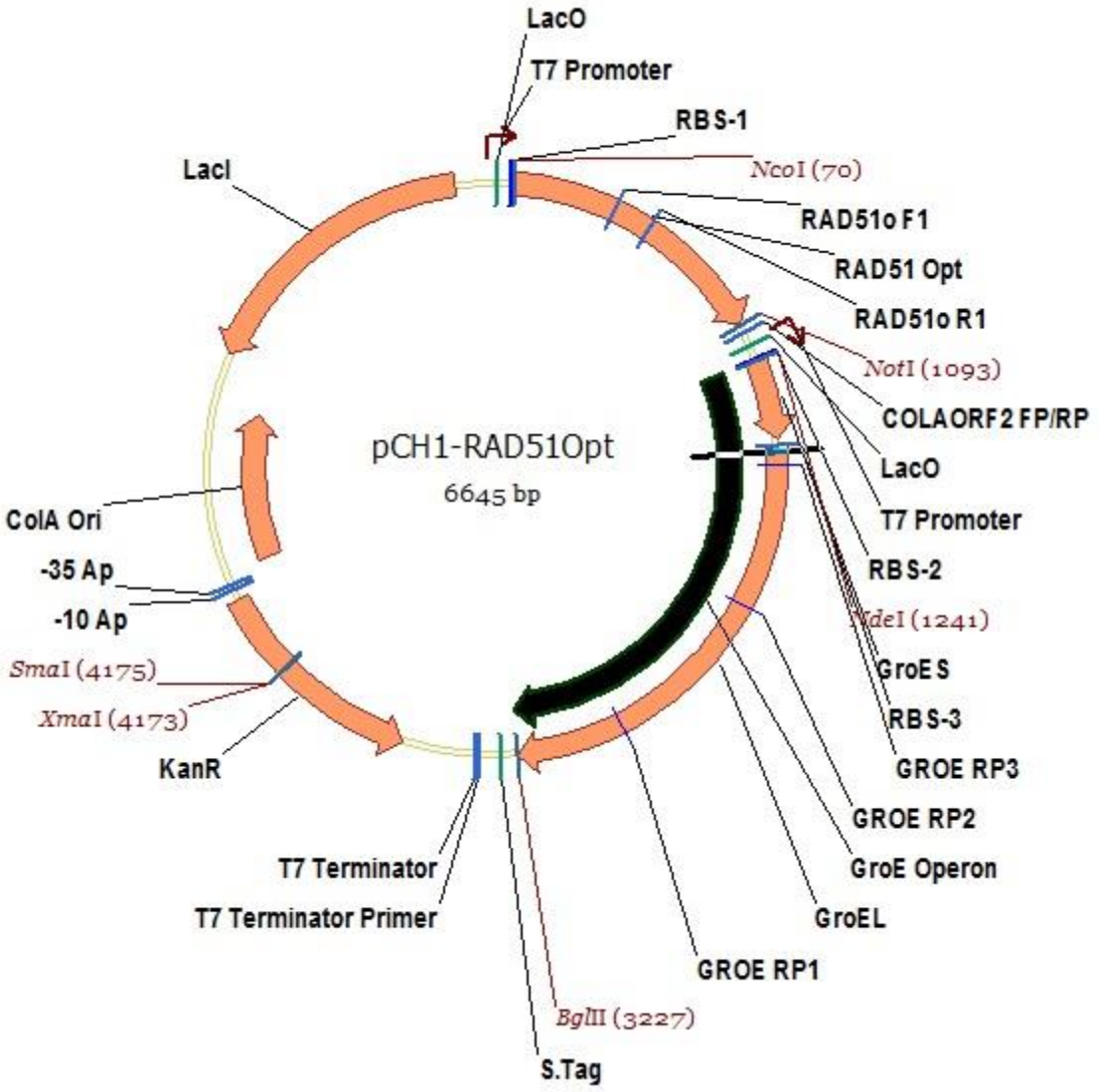
GROE RP1 - ACGCGGATCAGCGCAACACCAC

GROE RP2 – TCCGGCTTGTTGATGAAGTAAGGAG

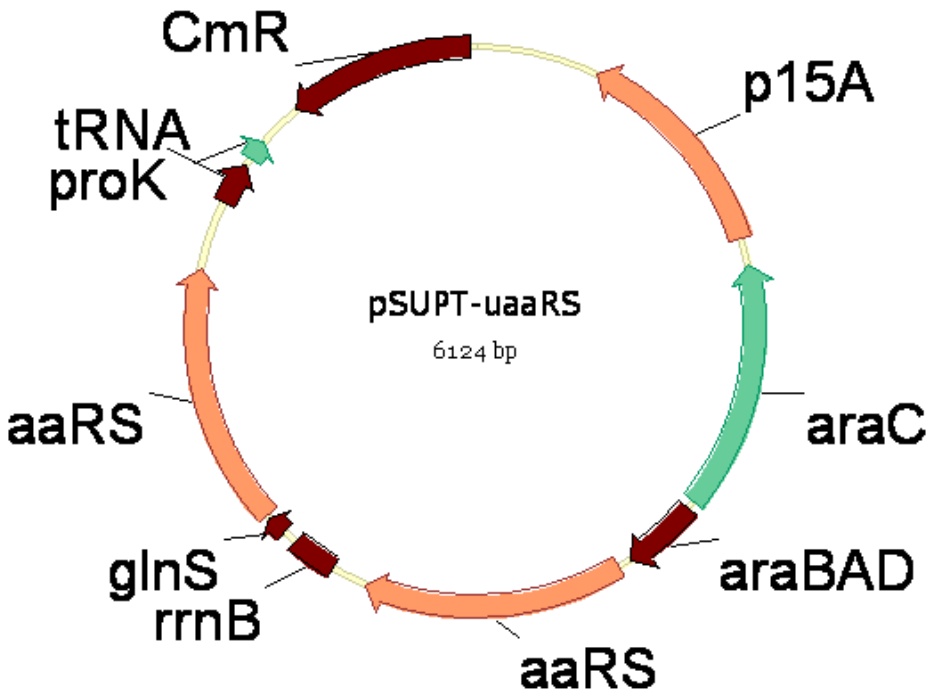
GROE RP3 – ACTGCATCTGCCAGTACGTTTACG

COLAORF2 FP – CGTATTGTACACGGCCGCATAATCG

COLAORF2 RP - CGATTATGCGGCCGTGTACAATACG



Vector Map of the pCH1-RAD51Opt Plasmid



Vector Map of the pSUPT-uaaRS Plasmid