The Role of Eukaryotic Recombinase Loop L1 During Homologous Recombination

Justin Benjamin Steinfeld

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2018

ABSTRACT

The Role of Eukaryotic Recombinase Loop L1 During Homologous Recombination

Justin Benjamin Steinfeld

Within the life of an organism, its deoxyribonucleic acid (DNA) is constantly bombarded with damaging agents from exogenous and endogenous sources. One of the most deleterious types of damage is the double-stranded break (DSB) in which a continuous strand of DNA is broken in two. As a result, the information stored in their connection is lost. If improperly repaired, a cell will either not survive or transform into a neoplasm. Homologous recombination (HR) is a mechanism by which the cell processes these broken ends and uses proteins called recombinases to search for an undamaged homologous DNA template for repairing the break, the homology search. Generally for eukaryotes, the recombinase, Rad51, performs the homology search. Without it, cells cannot repair spontaneous DSBs by recombination and instead, must use alternative, less efficacious pathways. This type of reparative homologous recombination generally occurs during mitosis and is thus called mitotic recombination.

In addition to its role in repair, HR is employed by eukaryotes during the first stage of meiosis to create crossover events, or chiasmata, between DNA homologs. The formation of these chiasmata is necessary for proper segregation of the chromosomes, preventing aneuploidy in the haploid cells destined for sexual reproduction. These crossover events have an added evolutionary benefit of mixing genes between the parental chromosomes, creating allelic diversity in the haploid cells. Eukaryotes have evolved a subset of meiotically-expressed proteins to mediate this process. Dmc1 is a meiosis-specific, second recombinase

that eukaryotes require to properly form these crossover events between homologs. It is not entirely understood why most eukaryotes require a second recombinase specifically designed for meiotic HR. A potential reason for this second recombinase may lie in the preferred templates for recombination that Rad51 and Dmc1 seek. Rad51 is employed mitotically to repair spontaneous DSBs and thus searches for the perfect undamaged copy, the sister chromatid, to prevent the loss of genetic information. Conversely, Dmc1 is employ meiotically to purposely form crossover events between homologs, which carry single-nucleotide polymorphisms (SNPs) between parental chromosomes. Thus, Dmc1 must be able to anneal DNA strands that aren't perfectly the same.

This work uses the single-molecule technique of DNA curtains to understand the factors that effect Rad51 and Dmc1 homologous DNA-capture stability. The first part of Chapter 1 is a historical exploration of homologous recombination research and a review of the current understanding of the pathway. The second part of Chapter 1 discusses human diseases that are associated with the failure to properly repair double-strand breaks. Chapter 2 will explain the single-molecule DNA curtain technique used throughout this work. Chapter 3 will show that Dmc1 is more tolerant of mismatches in captured DNA than Rad51. Chapter 4 will test the limits of Dmc1's tolerance to imperfect DNA and attempts understand how it accomplishes this tolerance. Chapter 5 will demonstrate that this tolerance of mismatches is mediated by a specific structural element in recombinases, loop L1, and a chimeric Rad51 with a Dmc1-like L1 can tolerate mismatches *in vitro* and *in vivo*. Chapter 6 will explore how recombinase mediators such as BARD1 and BRCA1 enhance RAD51's ability to capture DNA during the homology search.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would first like to thank, Dr. Eric C. Greene, for mentoring me throughout my PhD and inspiring me to continue my pursuit of a research career through his inexhaustible passion and excitment for science. I would like to also thank all the members of my thesis commitee: Drs. Rodney Rothstein (Chair), Lorraine Symington, Ruben Gonzalez, and Alberto Ciccia. I also thank current and previous members of the Greene lab for willingly providing their respective expertise and making life in the lab truly an enjoyable experience. Most importantly, I would like to thank my parents, Bart and Rebecca Steinfeld, and my sisters, Carrie Schulmeister and Elizabeth Charette, for providing the love, support, and encouragement that enabled me to get to where I am today.

Dedication

To Emil Botti

musician, teacher, and friend

Chapter 1

Introduction

As a museum that holds the great works of humanity, each living being holds the magnum opus of evolution in the form of its DNA. For humans, this work represents an almost four-billion year process of natural selection to produce over six billion base pairs of sequence information stored in every cell in the body. Although variations in DNA account for the diversity of life across species and the human population, alterations in certain genetic regions can contribute to the morbidity and mortality that plague all life. Thus, there is a balance between the preservation of these alterations and the restoration of the original work to realize each organism's Darwinian potential.

Deleterious alterations can come in many forms and from various sources. Each form poses a formidable challenge to genomic integrity. Thus, organisms have evolved many proteins designed specifically for each type of damage. One such form is a discontinuation in the series of phosphodiester linkages that hold a strand of DNA together. Normally, this single-stranded nick can be seamlessly repaired through ligation; however, when there is also a second break on the opposite strand of DNA, the information stored in their connection can be forever lost. This type of alteration is referred to as a double-strand break (DSB), one of the most deleterious forms of DNA damage, making proper repair of DSBs critical for longterm survival.

1.1 Double-Strand Break Repair

1.1.1 Formation of Double-Strand Breaks

The causes of double-strand breaks (DSBs) can be categorized into spontaneous and programmed events. Spontaneous DSBs can originate from exogenous sources, such as ionizing radiation (IR) or chemotherapeutic agents, and endogenous sources, such as free radicals from metabolism or DNA stress during replication [1, 2]. The frequency of DSBs has been estimated to be about one per 10^8 bp per day—a rate consistent across *E. coli* [3], yeast [1], chicken DT40 cells [4], and normal humans cells [3]. For humans, this represents about 50 DSBs per cell per day or more than 50 quintillion breaks in the entire body over an average human lifetime. DSBs during DNA replication, either by ssDNA breaks or stalled replication forks, appear to represent a majority of spontaneous events [1, 5]. However, at certain times, cells intentionally create DSBs, such as during mating-type switching in yeast, during meiosis, or for adaptive immunity (e.g. V(D)J recombination, class switching, and somatic hypermutation); and recent studies have even suggested DSBs contribute to normal physiological neural function [1, 2, 6]. These programmed events are highly regulated and involve additional proteins outside the canonical double-strand break repair (DSBR) pathway [1]. Since most of this work is related to *Saccharomyces cerevisiae* proteins, this model organism will be the primary focus of this part of the introduction and often referred to as just yeast. *E. coli*, archaea, and human homologs will be noted when appropriate.

1.1.2 Recognition of the Double-Strand Break

As with finding any solution, recognition of the problem is often the most important step. In yeast, the sentinel of DSBs is the MRX (MRN in humans) complex composed of Mre11, Rad50, and Xrs2 (Nbs1 in humans) and appears to be one of the first complexes to recognize DSB damage [7, 8]. The MRX complex appears to be at least partially responsible for physically tethering the two DSB ends through the Rad50 Zn-hook structure, preventing the broken ends from diffusing away from one another [8–10]. Mre11 and Rad50 are conserved in bacteria and archaea, but Xrs2 appears to be unique to eukaryotes, ensuring Mre11 nuclear localization [11]. MRX complex bound to DSB ends recruits the serine/threonine kinase, Tel1 (ATM in humans), through direct interaction with Xrs2, activating Tel1 for phosphorylation of histone H2A (H2AX in humans) at serine 129 and cell-cycle arrest [8, 12, 13]. Once the DSB is recognized, the cell must now act to repair the damage. There are two primary DSBR pathways: non-homologous end joining (NHEJ) and homologous recombination (HR) and the predominant pathway choice is dependent on where in the cell cycle the DSB has occurred and whether initial DNA processing has occurred (Fig. 1.1) [5, 14, 15].

1.1.3 Non-homologous End Joining (NHEJ)

Although NHEJ occurs throughout the cell cycle, it is the predominant mode of repair during the G1 phase of haploid yeast cells [8, 15]. The pathway choice between HR and NHEJ is determined by whether initial resection of the dsDNA ends has occurred, as inhibition of the resection pathway is necessary for NHEJ to proceed [8, 15]. This inhibition is achieved

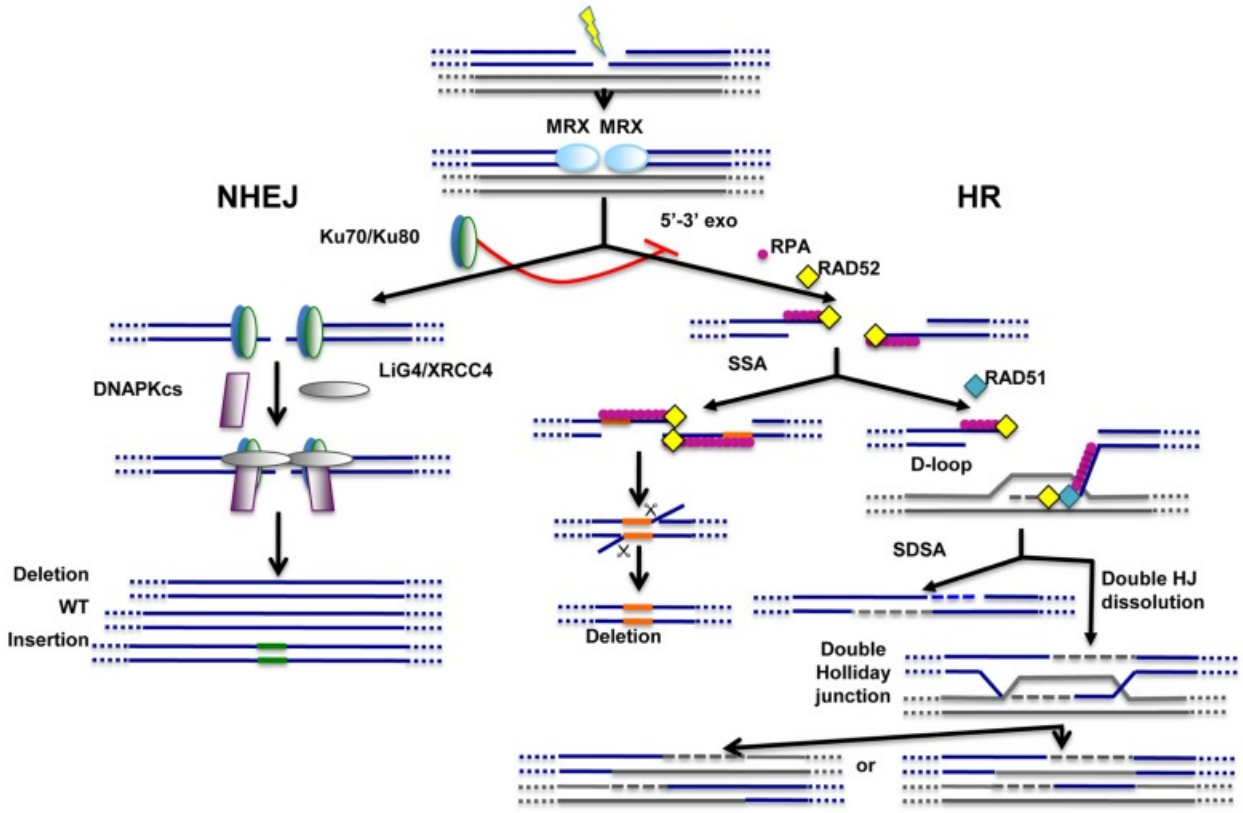

Figure 1.1: **Main pathways of DNA repair.** Non-homologous end-joining (NHEJ) and homologous recombination (HR) pathways act competitively to repair DNA double-strand breaks (DSBs). Key players of NHEJ and HR are depicted. The MRE11/RAD50/XRS2 (MRX) complex is recruited very early at DNA ends and appears to play important roles for both NHEJ and HR. Ku70/Ku80 heterodimer is required for NHEJ and, through inhibition of DNA end resection (5'→3'exo), acts as a repressor of HR. Fidelity of NHEJ-dependent DSB repair is low and, most of the time, associated with nucleotide deletions and/or insertions at repair junctions. The common early step of HR-dependent mechanisms is the formation of ssDNA which is then coated by replication protein A (RPA). Single-strand annealing (SSA) mechanism requires the presence of direct repeats (shown in orange) on both sides of the break. SSA does not imply any strand invasion process and is therefore not dependent on RAD51 protein. Strand invasion and D-loop formation are however common steps of synthesis-dependent strand annealing (SDSA) and double Holliday junction (HJ) dissolution mechanisms. In the latter case, double Holliday junctions are resolved with or without crossing-over. This figure was reproduced with permission from Decottignies, A. Alternative end-joining mechanisms: a historical perspective. *Front Genet* (2013).

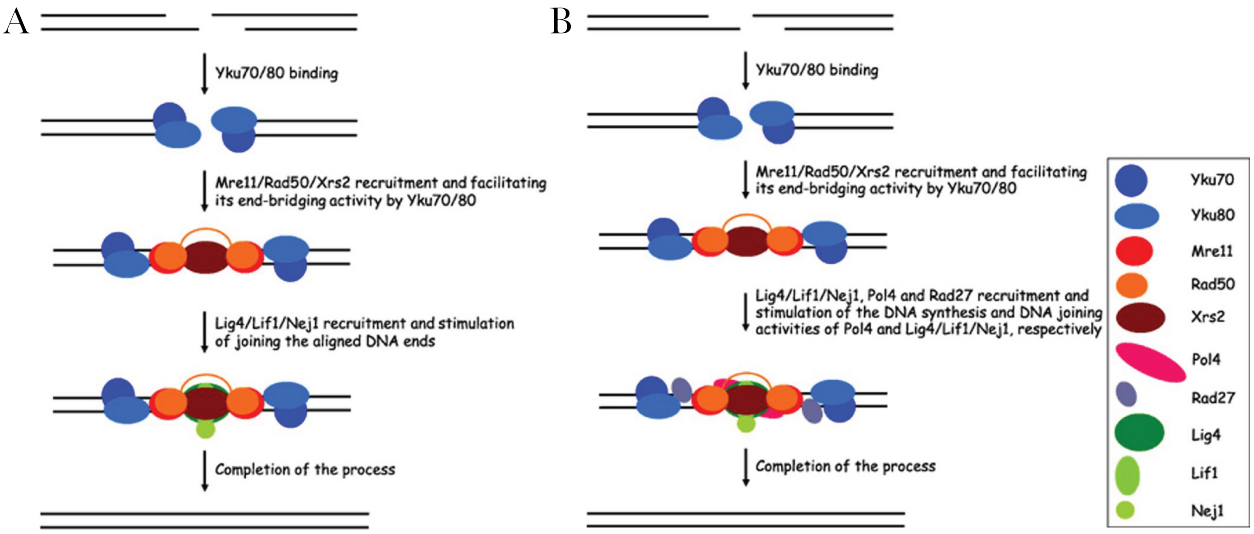

Figure 1.2: **Current model of the *Saccharomyces cerevisiae* NHEJ.** (**A**) Simple religation. In a similar manner to the process of mammalian NHEJ, Yku70/80 heterodimer is the first NHEJ factor to bind to broken DNA ends. Subsequently, DNA-bound Yku70/80 recruits the Mre11/Rad50/Xrs2 complex and facilitates its end-bridging activity. This activity enables Lig4/Lif1/Nej1 complex to join Yku70/80-bound DNA ends. (**B**) DNA-ends-processing-dependent NHEJ. If broken DNA ends are non-complementary, DNA end-processing reactions, generating ligatable structure and occurring prior simple ligation, are required. These reactions include nucleolytic end-processing and gap-filling mediated by Rad27 and Pol4, respectively. This figure was reproduced with permission from Dudasova, Z., A. Dudas, and M. Chovanec, Non-homologous end-joining factors of Saccharomyces cerevisiae. *FEMS Microbiol Rev* (2004).

by the MRX-independent recruitment of the Ku (Yku70–Yku80) complex to DSB ends, protecting the ends from degradation and directly inhibiting the resection activity of the MRX complex [8, 15, 16]. The Ku complex is then the scaffold to which the remainder of NHEJ proteins are recruited and allowed to complete repair (Fig. 1.2) [9, 16–19].

The core NHEJ machinery includes the proteins DNA Ligase IV (Dnl4/Lig4 in humans), Lif1 (XRCC4 in humans), and Nej1 (XLF in humans) (Fig. 1.2A) [16, 17]. Lif1/XRCC4 has been shown to be necessary for stabilizing DNA Ligase IV *in vivo* and for recruiting Dnl4 to DSBs [20, 21]. The role of Nej1 is not entirely understood; however, it has been shown to directly interact with Lif1, and when deleted, to be necessary for efficient NHEJ [9, 22, 23]. Nej1 in concert with Lif1 appears to be critical for stable binding of

the Ku complex to DSB ends [23, 24]. More recently, Nej1 may play an independent and additive role in recruiting Rad27 and Pol4, which are responsible for end-processing and gap-filling, respectively, and are occasionally required to ligate DSB ends by NHEJ (Fig. 1.2B) [17, 23]. Ultimately, the critical step in NHEJ is Dnl4 ligation, which due to its footprint, presumably occurs one strand at a time, even if two Dnl4 molecules are present [9]. Overall, NHEJ is error-prone with genetic loss due to end processing and the potential of chromosomal translocations events [8, 15, 25, 26]. Thus, alternative methods of repair are preferred when available.

1.1.4 Homologous Recombination

Similar to the restoration of a painting, homologous recombination (HR) proteins can use homologous copies of the damaged DNA as a reference for repairing with greater fidelity than NHEJ. During S phase and G2 phases, DNA replication produces homologous templates for repair, making HR a viable pathway [8, 27, 28]. Interestingly, diploid cells always prefer HR for repairing breaks even if there is only a homologous template [29]. As discussed earlier, the occurrence of 5'→3' nucleolytic resection commits the cell to HR. Resection proceeds in two steps: i) MRX (Mre11)/Sae2-mediated; ii) Exo1/Dna2-Sgs1-mediated [8]. The initial MRX and Sae2 resection is more robust in its ability to remove irregular end structures, such as hairpins and bulky adducts; however, it is much less processive, leaving behind small 3'-ssDNA overhangs of a couple hundred nucleotides (nt) [8, 30]. This initial resection step is facilitated by phosphorylation of Ku70 during S phase, enabling the MRX complex nuclease activity [31] and activation of Sae2 by direct phosphorylation by Cdk1, a cell-cycle-regulated

kinase involved in the G1/S phase transition [32]. In addition, these ssDNA overhangs, in turn, inhibit binding of Ku to the DSB ends [8]. Efficient HR requires longer tracts of up to 1-kb ssDNA overhangs mediated by the second step in resection [30]. The Sgs1 helicase may be critical for paving the way for processive 5'→3' nucleolytic resection by the Dna2 nuclease with Exo1 nuclease playing a less critical but non-redundant role [30, 33, 34].

The second-step of resection causes dissociation of MRX, Sae2, and Tel1 and binding of replication protein A (RPA) to 3'-ssDNA overhangs, serving to remove secondary structure, prevent degradation, and recruit additional checkpoint signaling proteins (Fig. 1.3) [8, 35]. With the aid of various mediators, such as Rad52 (BRCA2 in humans), the recombinase, Rad51 (RAD51 in humans) replaces RPA on the ssDNA, forming a filamentous nucleoprotein structure known as the presynaptic filament (Fig. 1.3) [36–39]. From this point, the presynaptic filament engages in the homology search, scouring the genome for homologous DNA to use as a template for repair; additional details describing this homology search mechanism will be provided below [27, 28]. Once a homologous dsDNA template is captured, the presynaptic filament invades the captured dsDNA, displacing the non-complementary strand, forming a D-loop structure and postsynaptic structure (Fig. 1.3) [27, 28].

Repair and resolution of the break can take two paths: i) synthesis-dependent strand-annealing (SDSA) or ii) classical double-strand break repair (DSBR), forming double Holliday junctions (dHJ) (Fig. 1.3) [27, 28]. During SDSA, polymerases, Pol δ and Pol ϵ, are thought to fill in resected DNA using the complementary strand as a template and second-end capture is prevented by anti-recombinase helicases, Srs2, Irc20, and Mph1 (FANCM in humans) (Fig. 1.3) [1, 40, 41]. Thus, SDSA obligatorily produces noncrossover events (NCO), limiting loss of heterozygosity (LOH). During DSBR, the non-complementary strand

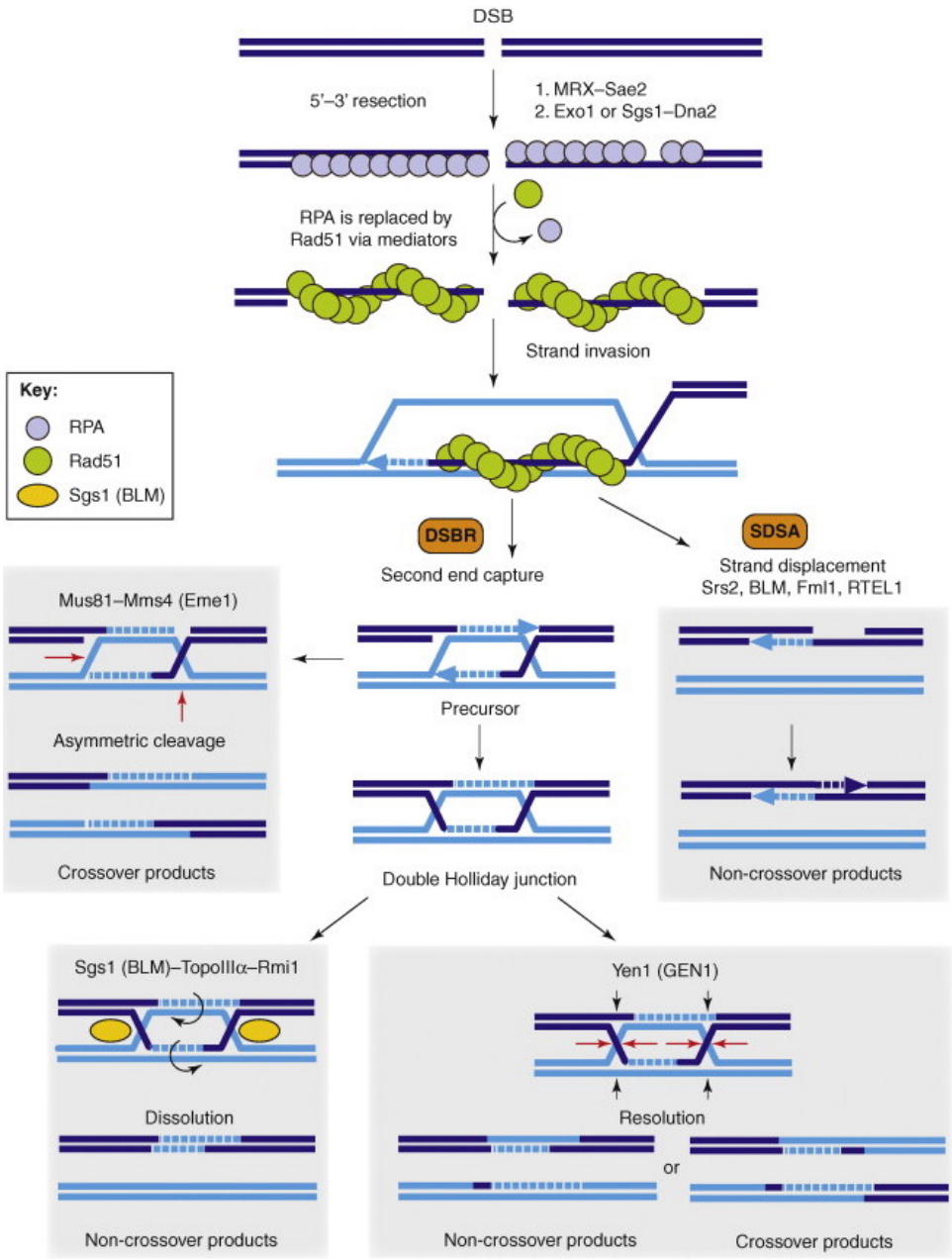

Figure 1.3: **Models and outcomes of the homologous recombination pathway.** 5'→3' resection of the broken ends creates 3'-ssDNA tails that are rapidly coated by RPA (light blue). RPA is replaced by Rad51 (green) to form the nucleoprotein filament, which can initiate pairing and strand invasion with the homologous duplex DNA. The 3'-end of the invading strand is extended by DNA synthesis using the donor duplex as a template. In the SDSA model, the invading strand is displaced and pairs with the other 3'-ssDNA tail, allowing DNA synthesis to complete repair. In the DSBR model, second end capture forms an early strand exchange intermediate. Processing of this precursor by Mus81–Mms4 (Eme1) generates crossover products, and ligation of this precursor creates a dHJ. Dissolution of the dHJ (via Sgs1 [BLM]–TopoIII -Rmi1) gives rise to noncrossover products, whereas resolution (via Yen1 [GEN1]) can lead to either crossover or noncrossover products. This figure was reproduced with permission from Mimitou, E.P. and L.S. Symington, Nucleases and helicases take center stage in homologous recombination. *Trends Biochem Sci* (2009).

of the captured dsDNA is bound and used as a template for repairing the second end of the ssDNA overhang, also known as second-end capture (Fig 1.3) [1, 27, 28]. Once second-end captured has occurred, the joint-molecule (JM) can either resolve into NCO or crossover (CO) events (Fig. 1.3) [1, 27, 28]. When the JM evolves into a dHJ structure, the helicase Sgs1 (BLM in humans), topoisomerase Top3 (Topo III in humans), accessory factor Rmi1 (RMI1 in humans), and potentially helicase Mph1 are responsible for dissolution of the dHJ by convergent branch migration into NCO events [34, 42–44]. Endonucleases Yen1 (GEN1), Mms81-Mms4, and Slx1-Slx4 appear to play a large role in the formation of crossover events, but it is not entirely understood how all of these resolvases act independently to achieve crossover events (Fig. 1.3) [34, 42, 45, 46]. In addition to these resolvases, recent work has pointed to Mlh1-Mlh3 and Sgs1 as the primary actors in the formation of crossover events during yeast meiosis [44]. As discussed below, the pathway chosen for DSB repair is dependent upon whether the DSB was formed during mitosis or meiosis [47, 48].

1.1.5 Mitotic Recombination

In the late 1960s, a series of experiments looking at yeast sensitivity to γ-radiation led to the discovery of a confluence of mutant genes displaying inherited spontaneous mutability [49, 50]. It was soon found that a series of proteins in the *RAD50* to *RAD57* loci were also critical for γ-radiation sensitivity [51–53]. Understanding that ionizing radiation (IR) caused double-stranded breaks, Michael Resnick and Patricia Martin discovered the first protein, Rad52, to be directly responsible for repairing these breaks [54, 55]. Eventually, additional loci were discovered to be responsible for the repair of double-strand breaks and

the *RAD52* epitasis group was formed which includes, *RAD50*, *RAD51*, *RAD52*, *RAD54*, *RAD55*, *RAD57*, *RAD59*, *RDH54*, *MRE11*, and *XRS2* [56]. Although the MRX (Rad50-Mre11- Xrs2) complex was mentioned earlier as part of DNA-end resection, the focus of the section will be to cover the recombination roles of members of the *RAD52* epistasis group post-resection and required recombination protein, RPA. The reader is encouraged to seek a recent review that explores DNA-end resection more extensively [57].

1.1.5.1 Replication Protein A

As discussed earlier, after DNA-end resection and formation of 3'-ssDNA overhangs, replication protein A (RPA) binds to the ssDNA to remove secondary structure, prevent degradation, and recruit checkpoint and other HR proteins (Fig. 1.3) [8, 35]. Interestingly, unlike many of the other HR proteins, the discovery of RPA has its origins in bulk biochemistry rather than genetics. In an attempt to understand the host proteins required for simian virus 40 (SV40) replication, HeLa cell extracts were systematically fractionated with one fraction containing the first eukaryotic protein with DNA binding activity highly specific for ssDNA over dsDNA and resistant to high salt, reminiscent of previous studies of *E. coli* SSB [58]. Due to its involvement in replication, this protein was soon purified, characterized as a heterotrimer of 70, 32, and 14 kDa subunits with ssDNA-binding activity, and given the name replication protein A (RPA) [59]. Soon after, Steven Brill and Bruce Stillman unsuccessfully tried to reproduce *in vitro* SV40 replication with yeast extracts. However, they identified a yeast homolog of human RPA with similar structural characteristics (heterotrimer with 70, 34, and 11kDa subunits) and ssDNA-binding activity [60]. In a separate attempt to isolate the *S. cerevisiae* proteins that catalyze strand exchange as seen previously

in *E. coli* [61], a single 132 kDa polypeptide was identified to be heavily overrepresented in the fractions [62]. The 34 kDa subunit was purified and shown to be able to mediate *E. coli* RecA strand-exchange in a similar manner to *E. coli* SSB [63]. Eventually, all of these yeast protein complexes were realized to be one in the same, yeast RPA, expressed by three genes *RFA1*, *RFA2*, and *RFA3*, and with roles in both replication and strand exchange [64–66]. Deletion of any of these three genes is lethal, making every subunit essential for yeast survival [67]. Additionally, genetic studies have successfully parsed RPA's diverse roles in DNA replication and repair with the discovery of the *rfa1-t11* (L45E) mutant, which appears to hinder DNA repair but not DNA replication [68, 69]. This section will focus on RPA's role in recombination.

Biochemically, RPA is a highly conserved ssDNA-binding protein with subnanomolar affinities ($K_d \sim 10^{-8} - 10^{-10}$ M), low cooperativity, and 50-fold greater affinity for pyrimidines over purines with low sequence specificity overall [70–73]. This exceptional ssDNA-binding affinity serves two primary purposes: i) protection of ssDNA from degradation and ii) secondary structure removal. *In vitro*, RPA has been shown to protect 3'-ssDNA overhangs from degradation while simultaneously promoting end-resection and processing by MRX and Exo1 [74]. Similar results were found *in vivo* with the prediction that RPA protects ssDNA degradation from endonuclease Dna2 [35, 75]. RPA's duplex melting properties in low salt conditions was one of the first properties discovered and appeared to not require ATP [76–78]. More recently, RPA has been show to be necessary to prevent DNA-end hairpin formation at palindromic sequences during DNA processing [35]. Additionally, hRPA has been show to melt G-quadruplexes in a 5'→3' direction and this directionality was shown kinetically in duplex DNA melting [79, 80]. Finally, this DNA melting has been

shown visually in single-molecule experiments with long tracts of M13mp18 ssDNA extended with the aid of hydrodynamic flow [81].

RPA's exceptional affinity for ssDNA poses a problem for recombination since RPA must eventually exchange with Rad51 and mediators to form the presynaptic filament for the homology search to begin (Fig. 1.3) [8]. Single-molecule experiments with ssDNA curtains have shown that RPA is, in fact, quite dynamic in its ssDNA binding activity [38, 70]. Although fluorescently-labeled RPA can remain stably bound to ssDNA for hours without free RPA in solution, once unlabeled RPA is introduced into the system, the unlabeled RPA quickly exchanges with labeled RPA and visa versa [38, 70]. Furthermore, without the need of mediators and as long as free RPA is not in the system, Rad51 can freely exchange with RPA-ssDNA complexes [38, 72]. This phenomenon led to the proposal of microscopic dissociation events by RPA allowing facilitated exchange with other RPA or Rad51 molecules without mediators [38, 72]. RPA's ability to dynamically bind ssDNA lies in its unique structure.

Structurally, RPA is a heterotrimer composed of 70, 32, and 14 kDa subunits, Rpa1, Rpa2, and Rpa3, respectively [73]. In total, RPA has six oligonucleotide-binding (OB) folds that mediate its DNA-binding properties and a winged helix (Wh) domain that mediates its protein partner interactions [82]. Rpa1 contains four of these OB folds, DNA-binding domains (DBD) F, A, B, and C with flexible linkers in between. Rpa2 has one fold, DBD-D, and the Wh domain, while Rpa3 has DBD-E. *In vitro*, these six DBD's coordinate to produce three distinct salt-concentration-dependent DNA-binding modes [83, 84]. The first 8-10 nt binding mode is mediated by Rpa1's DBD-A and DBD-B domains followed by an intermediate 12-20 nt mode thought to be mediated by the addition of DBD-C [83]. The

final 30 nt mode is mediated by the addition of Rpa2's DBD-D [83]. It is hypothesized that the coordination of these OB folds and their dynamic binding modes are what allow RPA to have both high affinity for ssDNA and microscopic dissociation events that enable facilitated exchange [38, 70, 72, 85].

1.1.5.2 Rad52

Although the Rad51 presynaptic filament can form on RPA-bound ssDNA in the absence of free RPA, free RPA is presumably always present in the cell. *In vitro*, when free RPA is present, it competitively inhibits Rad51's exchange and thus mediators are required to allow Rad51 nucleation on the ssDNA, the rate-limiting step for Rad51 filament formation [38, 72, 86]. Rad52 is one of these critical mediators. Although, non-essential, *rad52* yeast strains are deficient in DSBR, thus, connoting its critical role in the pathway [15, 28, 87]. As discussed earlier, Rad52 was the first DSBR protein found and thus the titular member of the *RAD52* epistasis group [51–55]. Rad52 quickly became of great interest to the scientific community with the gene being cloned, sequenced, and primary structure determined soon after [88, 89]. However, due to its role in DSBR, Rad52 remained rather elusive until the characterization of recombinases Rad51 and Dmc1 (discussed later). Genetically, it was shown that *K. lactis* Rad52 could partially complement a *S. cerevisiae rad52* background and a dominant negative effect in a wild-type background, which could be overcome by Rad51 overexpression, suggesting a competitive interaction with Rad51 between a non-productive *K. lactis* Rad52 and wild-type *S. cerevisiae* Rad52 [90]. This study further demonstrated a Rad51-Rad52 interaction with a yeast two-hybrid experiment [90].

One of the first *in vitro* biochemical properties found for Rad52 was its ability to

specifically bind Rad51 in a study focused on the identification of a eukaryotic RecA-like protein [91]. A Rad52-focused study showed that recombinantly expressed Rad52 could bind both single- and double-stranded DNA with a slight preference for ssDNA [92]. This study also demonstrated that Rad52 could anneal a partially melted "Y" DNA structure in a homology-dependent manner with second-order kinetics [90]. Rad52's binding to the recombinase Rad51 in addition to its strand annealing property, led to the hypothesis that it participates in strand annealing after Rad51 strand exchange [90]. The significant reduction in the formation of Holliday junctions *in vivo* in a *rad52* background, but not for strains with deletions of other members of the *RAD52* epistasis group, further implicated the purpose of Rad52's strand-annealing property [93]. The importance of Rad52's strand-annealing property for second-end capture was then shown *in vitro*, establishing Rad52's critical role in the DSBR pathway (Fig. 1.4) [94].

Rad52 was also found to have a mediator function for Rad51 presynaptic filament formation. It was shown that Rad52 could help Rad51 overcome free RPA inhibition of Rad51 strand-exchange activity, elucidating its mediator role in recombination [86, 95]. Simultaneously, a yeast two-hybrid study showed that Rad52 interacts directly with RPA *in vivo* [96]. All of these biochemical properties and genetic findings coalesced into a coherent Rad52 mechanism of action when an *in vitro* study demonstrated that Rad51 nucleation on RPA-ssDNA filaments is the rate-limiting step for presynaptic filament formation, Rad52 stimulates Rad51 displacement of RPA on ssDNA, and Rad52 binds to RPA-bound ssDNA [97]. A new model was proposed: RPA first removes secondary structure from ssDNA and recruits Rad52 to this RPA-ssDNA complex to allow nucleation and presynaptic filament formation by Rad51 [97].

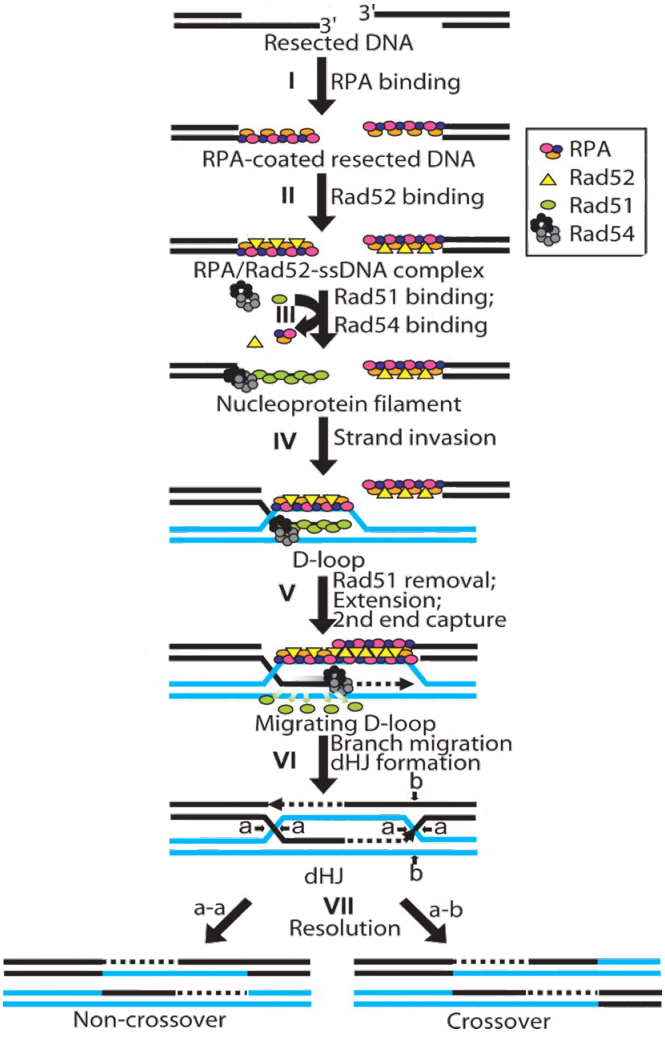

Figure 1.4: **Rad52-mediated second-end capture as part of DSBR.** Rad52-mediated second-end capture as part of DSBR. (Step I) RPA binds to the 3'-terminated ssDNA tails of the resected DSB. (Step II) Rad52 binds to the RPA–ssDNA complex. (Step III) Rad52 mediated displacement of RPA by Rad51 to form a Rad51 nucleoprotein filament (Left end); Rad54 binds and stabilizes the Rad51 nucleoprotein filament. (Step IV) The Rad54–Rad51–DNA nucleoprotein filament catalyzes DNA strand invasion to form a D-loop, displacing the resident strand in the target dsDNA; the displaced strand is bound by RPA and Rad52. (Step V) Rad54 clears Rad51; the invading 3'-end primes DNA synthesis; extension causes the displaced strand to advance until an ssDNA region complementary to second resected DNA end is exposed; Rad52 mediates annealing of the RPA-complexed displaced strand to second end (Right end). (Step VI) Additional DNA synthesis, branch migration, and ligation generate dHJs. (Step VII) Resolution to produce cross-overs or noncrossovers. This figure was reproduced with permission from Nimonkar, A.V., R.A. Sica, and S.C. Kowalczykowski, Rad52 promotes second-end DNA capture in double-stranded break repair to form complement-stabilized joint molecules. *PNAS* (2009).

Recent single-molecule work on ssDNA curtains has confirmed all of these results with interesting additions to the proposed model in Figure 1.4 [38]. When Rad52 foci were bound to the labeled-RPA-ssDNA complex and unlabeled RPA was flowed into the system, most RPA turned over by the facilitated exchange mechanism mentioned earlier; however, those RPA clusters with Rad52 did not turnover suggesting that Rad52 prevents the proposed microscopic dissociations by RPA [38]. As expected, these Rad52 foci nucleated Rad51 presynaptic assembly, but many RPA-Rad52 clusters were left behind and were shown to

be able to then nucleate Rad52 filament formation from these clusters [38]. This Rad52-RPA-Rad51-ssDNA presynaptic complex is proposed to be mechanistically important for stabilizing the noncomplementary strand after dsDNA capture and further downstream HR processes [38]. The biological relevance of this complex is further corroborated by *in vivo* evidence of RPA and Rad52 present at DSB's before, during, and after Rad51 presynaptic filament formation [38].

Structurally, there is little know about yeast Rad52 other than its ability to form multimeric rings by electron microscopy (EM) [98], with Rad52's diversity in function delineated by protein truncations and point mutations. Rad52's RPA, Rad51, and DNA-interacting domains are thought to be located in its C-terminus [87]. Rad52's N-terminus is responsible for oligomerization of Rad52, DNA-binding and ssDNA-annealing activity for second-end capture, and binding of Rad59 (discussed later) [99]. When this N-terminus is mutated, *rad52-R70A* strains are capable of promoting Rad51 loading as well as wild-type, mediating strand invasion; however, these strains are unable to complete recombination, confirming the importance of Rad52's DNA-binding and annealing activity for second-end capture [99]. Overall, there remains significant work in understanding the structure-function relationship of Rad52 and Rad52's exact mechanism during multiple steps of DSBR.

1.1.5.3 Rad55 and Rad57

Rad55 and Rad57 were first identified as part the original *RAD52* epistasis during an IR-sensitivity screen [51–53]. The role of Rad55/57 in regulating Rad51 filament stability was first shown by the effects of temperature on IR sensitivity [100, 101]. While *rad55* and *rad57* yeast strains show IR sensitivity similar to *rad51* at lower temperatures (20 °C), rad55

IR sensitivity was significantly suppressed at 30 °C [100, 101]. Through a yeast two-hybrid screen with Rad51, Rad55, and Rad57, Rad55 and Rad57 interact with each other and with Rad51 through an interaction with Rad55, suggesting Rad55/57's role as a heterodimer [102]. Temporal analysis of Rad55/57 *in vivo* shows Rad55/Rad57 foci form at the break after Rad51 foci formation, indicating that the complex is unlikely to serve a role in Rad51 nucleation in addition to its stabilization activity [103, 104].

In vitro bulk biochemical assays demonstrated Rad55 and Rad57 form a heterodimer with a $K_d < 2.0 \times 10^{-10}$ M [105]. Simple structural analysis shows that Rad55 and Rad57 are paralogs of Rad51, sharing 20-30% sequence identity, containing putative nucleotide-binding Walker A and B motifs [27]. Rad51 paralogs represent an important, but poorly understood class of proteins that promote the stability and/or assembly of the presynaptic complex [27]. Consistent with their similarity to Rad51, both Rad55 and Rad57 show ssDNA-binding and ATPase activity; however, this activity is not ssDNA- or dsDNA-dependent and these proteins are not independently capable of strand-exchange activity [105]. In this same study, it was shown that Rad55/57 increased Rad51 strand-exchange activity by aiding in Rad51 nucleation in the presence of RPA; however, this conflicts with previous temporal *in vivo* studies and potentially an artifact of the Rad55/57-stabilization behavior [105]. A combination of genetics and bulk biochemistry further progressed the Rad55/57-stabilization hypothesis by uncovering a series of Rad51 mutations that partially suppressed ionizing radiation in a *rad55 rad57* background [106]. One of these mutants, Rad51-I345T, was shown to be a dominant mutation *in vivo* and to have a higher binding affinity for both ss- and dsDNA and to overcome RPA competitive inhibition for DNA binding *in vitro* [106]. Rad55/57 role in Rad51 stabilization was further elucidated by gel electrophoresis assays

involving high salt conditions or the presence of the antirecombination helicase, Srs2, that showed Rad51 presynaptic instability can be overcome by the presence of Rad55/57 [107]. Although there are potentially more functions of the Rad55/57 heterodimer, it appears that its main function in HR is as an essential accessory factor for Rad51.

1.1.5.4 Rad54

As with other members of the *RAD52* epistasis group, Rad54 was first identified during the original IR-sensitivity screen in the 1970s [51–53]. Interestingly, during the same temperature-sensitive study for Rad55 mentioned earlier, Rad54 was found to have the opposite effect—the *rad54* strain IR sensitivity was suppressed by colder temperatures [100]. Similarly to Rad52, Rad54 was then shown to be involved in DSBR and the *rad54* repair defect showed similar temperature-sensitive effects to the earlier IR sensitivity study [108]. Further evidence of Rad54's involvement in the IR damage response was shown by Rad54's upregulation after γ-radiation exposure [109]. It was later demonstrated that Rad51 directly interacts with Rad54 *in vivo* by yeast two-hybrid studies [110, 111] and through the N-terminus of Rad54 by *in vitro* far-western and pull down assays [111]. Additional *in vivo* evidence of the potential antirecombinase activity Rad54 was demonstrated by the synthetic lethality of *rad54 srs2* [112].

The first biochemical characterizations of yeast Rad54 came over twenty years after its identification in the original IR-sensitivity screen and presented results seemingly in conflict with the antirecombination hypothesis from *in vivo* studies [113]. Rad54 was shown to have DNA-dependent ATPase activity with double the magnitude in the presence of dsDNA versus ssDNA and three-orders of magnitude greater than that of Rad51 [113]. Although it was

previously shown that Rad54 carried the putative helicase motifs of the Swi2/Snf2 translocase family [114], it did not show helicase activity *in vitro* [113]. It was shown that Rad54 promoted Rad51 homologous pairing and D-loop formation but did not have pairing activity by itself [113]. Soon after, a study showed that this homologous-pairing-mediation activity was dependent on Rad54's ATPase activity and ATPase dead mutants were sensitive to MMS *in vivo* [115, 116]. This study also demonstrated that Rad54 forms a dimer/oligomer in a DNA-dependent manner and this oligomer state could negatively supercoil relaxed DNA plasmid in an ATP-hydrolysis-dependent manner, lending credence to the earlier finding of Rad54 containing helicase motifs [116]. This helicase activity is hypothesized to be responsible for Rad51's homologous pairing enhancement *in vitro* and important for the homology search *in vivo* [116]. Later *in vitro* bulk biochemistry experiments established an optimal order and stoichiometry for Rad54's role in mediating Rad51 homologous pairing. Specifically, Rad54 is introduced after the formation of Rad51-ssDNA presynaptic filaments but before homologous dsDNA in a 1:1 ratio [117]. This work presented a new model of Rad54 binding as a co-complex along presynaptic filaments in preparation for the homology search [117]. This stoichiometry was later confirmed by immobilizing ssDNA-Rad51 on polystyrene beads by a biotin-streptavidin interaction, showing Rad54 co-complex stoichiometrically and Rad54 stabilization of the presynaptic filament [118].

At this point, the exact purpose of Rad54's helicase activity and its connection to homologous pairing mediation still needs to be established. Rad54 was shown *in vitro* to produce both negative and positive supercoils, which was simulated by Rad51 [119]. Additionally, this remodeling allowed relaxed DNA to be contorted such that ssDNA was exposed, allowing access for Rad51 presynaptic filaments and a new hypothesis for Rad54's

homologous pairing mediating activity [119]. This remodeling behavior was then shown to directly act on chromatin, mediating bidirectional movement of nucleosomes across DNA and this activity was stimulated by Rad51-ssDNA presynaptic filaments [120]. This ATP-dependent chromatin remodeling behavior was then demonstrated *in vivo* during HR at the MAT locus [121]. Eventually, it was directly shown that Rad54's chromatin remodeling behavior mediated the joint formation on chromatin *in vitro*, further evidence of the Rad54's remodeling behavior in mediating Rad51 pairing [122]. In addition to applying supercoiling and chromatin-remodeling, Rad54 was shown to disassemble partially Rad51-bound dsDNA dead-end complexes in an ATP-dependent manner, thus demonstrating antirecombinase activity and drawing the connection to the *rad54 srs2* synthetic mutant discussed earlier [112, 123, 124]. Thus, Rad54's ATPase activity serves a multitude of functions towards mediating Rad51 homologous pairing.

An *in vivo* study demonstrated that Rad54's helicase activity also acts postsynaptically, showing that gene conversion to heterologous DNA templates decreased with overexpression of Rad54 and increased with expression of ATPase-dead mutants [125]. This potential postsynaptic behavior was then shown *in vitro* with human RAD54 preferentially binding to branched DNA molecules and could branch migrate synthetically generated Holliday Junction-like DNA molecules [126].

Single-molecule experiments have elucidated some of Rad54's biophysical properties, including a dsDNA translocation rate of ∼301 bp/s, processivity of ∼11.5 kbp, and ATP concentration-dependent velocity with a K_m ∼97 μM [127]. More recently, it has been shown by single-molecule DNA curtains that Rad54 can bind to ssDNA only when Rad51-ssDNA filaments have formed, confirming earlier bulk biochemistry results [128]. Additionally, Rad54

remained bound upon Rad51 dissolution by ATP depletion [128]. Together, these results combined with known Rad54 DNA-binding activity suggest that Rad54 does not stabilize Rad51 filaments and binds in clusters at the gaps between filaments that are too small for RPA-binding [128]. This finding suggests that Rad54 does not form co-filaments with Rad51-ssDNA and that the 1:1 optimal stoichiometry shown in earlier bulk experiments is not due to homogenous binding of Rad54 but clustering between filaments [128].

Structurally, Rad54 has a primary and secondary Rad51 binding site at the N-terminus and is not necessary for remodeling [129]. The C-terminus, chromatin-modeling domain has been crystalized, revealing interesting relationships between structure and function (Fig. 1.5) [129]. Of note, there are two superfamily 2 (SF2) helicase-like lobes that act almost as a palm along the presumed dsDNA-binding axis with two α-helical domains (HD1 and 2) which are the Swi2/Snf2-related domains that act as the fingers onto the chromatin for remodeling (Fig 1.5) [129].

1.1.6 Meiotic Recombination

Since mitotic and meiotic recombination both abide by the guiding principles of the DSBR model, there is significant overlap in the major proteins involved. However, as mentioned earlier and to be discussed further later, the initiation and end products of meiosis and mitosis are different. Thus, the roles of mitotic factors may change during meiotic recombination, and additional HR-related proteins are expressed exclusively during meiosis.

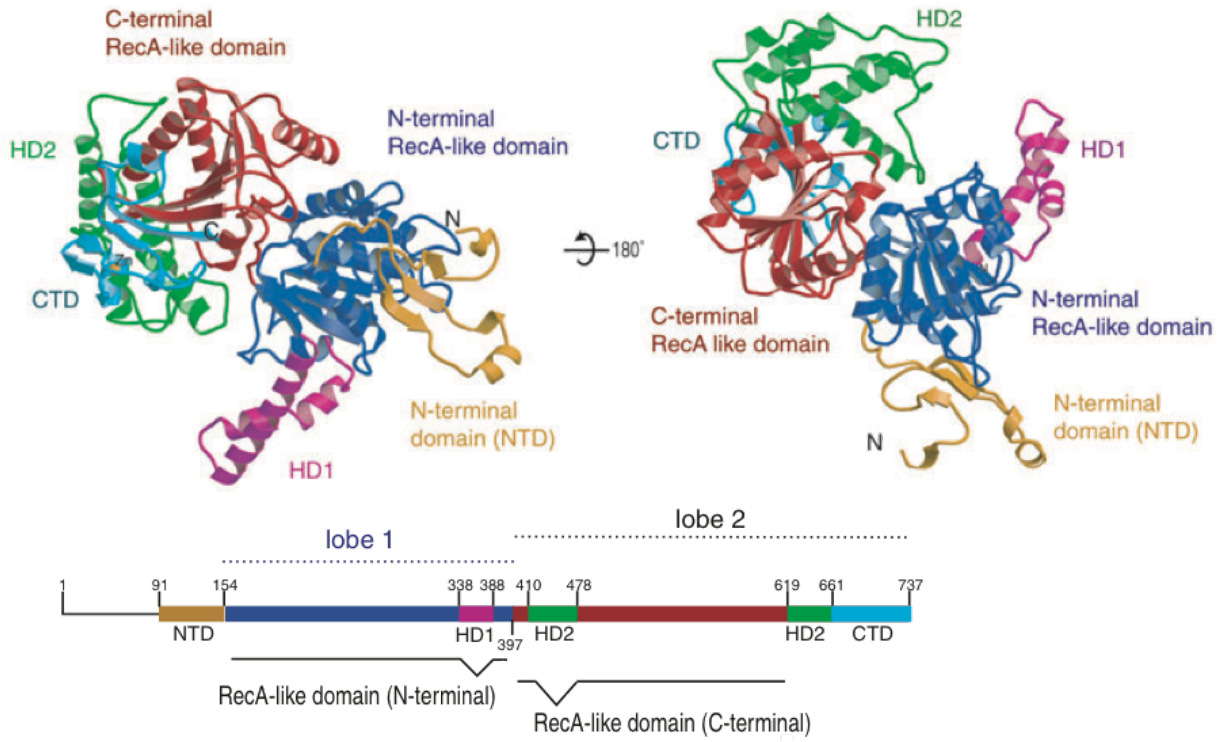

Figure 1.5: **Overview of the structure of zebrafish Rad54 (dnRad54ΔN).** The structure of the core of Rad54. SWI2/SNF2 specific elements are depicted in magenta (HD1) and green (HD2), the RecA-like helicase domains are blue (lobe 1) and red (lobe 2), and NTD and CTD are yellow and cyan, respectively. The locations of the individual domains in the sequence are shown below the structure. This figure was reproduced with permission from Thoma, N.H., et al., Structure of the SWI2/SNF2 chromatin-remodeling domain of eukaryotic Rad54. *NSMB* (2005).

1.1.6.1 Spo11

Since DSBs need to be present for recombination to occur, arguably one of the most critical meiotic proteins is Spo11, which produces the programmed break. *SPO11* was first gene discovered in a series of genetic experiments looking for mutants that were deficient in sporulation [130–133]. Twenty years after this finding, a genetic study showed that DSB formation was necessary for meiotic recombination initiation by monitoring changes in a yeast plasmid containing a meiotic initiation site and strategically placed restriction sites [134]. Spo11 became a candidate for producing this DSB when examination of a meiotic recombination hot spot at *HIS4-LEU2* failed to show a break in *spo11* or *rad50* backgrounds

[135]. Further explorations demonstrated that X-ray-induced breaks could suppress the non-sporulating phenotype of *spo11* strains, implicating Spo11's role in initiation of meiotic recombination [136]. Eventually, two separate groups, in quick succession, demonstrated Spo11 was responsible for initiation DSBs during meiotic recombination [137, 138]. Both groups used the findings of a previous study, which showed that *rad50S* (a non-null Rad50 mutation) caused the accumulation of DSBs without further processing and an unknown covalently-bound protein attached to the 5' of the DNA fragments hypothesized to be a type II topoisomerase [139]. One of these groups used primary structure alignment to demonstrate that Spo11 had homology with subunit A of a known *Sulfolubus shibatae* type II topoisomerase and predicted that Tyr135 was the catalytic amino acid [137]. Furthermore, they demonstrated that a Y135F mutation prevented the formation of DSBs during mitosis [137]. More definitively, another group used trypsin-digestion and mass spectrometry to show that Spo11 was the covalently-bound protein on the 5'-end of unprocessed DSBs [138].

Structurally, Spo11 is considered a type IIB topoisomerase, an archaeal relative of topoisomerase VI, which initiates the programmed DSB during leptotene to zygotene stage of meiotic prophase I by cutting both strands of the DNA through the formation of covalent protein-5'-end linkages in a topoisomerase-like mechanism [140, 141]. Specifically, a Spo11 dimer binds to the dsDNA and each Spo11, with a divalent ion cofactor, performs a trans-esterification reaction by attacking the phosphorus in the phosphodiester linkage with the nucleophilic hydroxyl-group on Try135, forming a tyrosyl phosphodiester linkage (Fig 1.6) [140, 141]. Spo11 must be removed to allow 5'→3' resection and the subsequent HR steps. Evidence points to Mre11 endonuclease of the MRX complex along with Sae2 cleaving the DNA behind each Spo11 [142, 143]. It is unclear whether Spo11 can turnover its covalently

bound oligonucleotide and is hypothesized to be on a "suicide" mission that self-regulates the number of DSBs that can form [140, 141, 143]. It would not be efficacious for meiotic recombination to happen across totally randomly placed DSBs across the genome, and thus DSB hotspots have been hypothesized [144]. Recently, by combining deep sequencing with microarray hybridization, a set of general principles of Spo11 preference has been generated with >88% of DSBs occurring at promoters [144].

1.1.6.2 Mei5-Sae2

After resection, RPA binds to 3'-ssDNA overhangs allowing formation of Rad51 and meiotic-specific recombinase, Dmc1, nucleofilaments through the aid of mediators. One of these mediators, the Mei5-Sae3 heterodimer, is meiosis specific (only in budding yeast), and appears to serve a comparable function to Rad52 for Rad51 in mitosis [87]. Sae3 was the first of the two to be determined necessary for the proper progression of meiosis [145]. In this study, transcription analysis by Northern blot demonstrated that Sae3 was a meiotically-expressed protein, and physical analysis of recombination by Southern blot showed *sae3* strains had hyper-resected DSBs [145]. Additionally, cytological analysis showed cells were stalled in the third stage of meiotic prophase, pachytene [145]. Mei5 was first shown to be critical in meiosis a few years later in a systematic genetic screen of over 301 ORFs with monitoring of meiosis by fluorescence microscopy and flow cytometry [146]. Strains with a *mei5* background were able to complete premeiotic S phase but did not initiate nuclear division nor sporulate [146]. By fluorescence microscopy, two studies showed that Mei5, Sae3, and Dmc1 spatially and temporally overlap during meiosis, and this localization is mutually dependent on both Mei5 and Sae3. Additionally, Mei5 and Sae3 promoted Dmc1

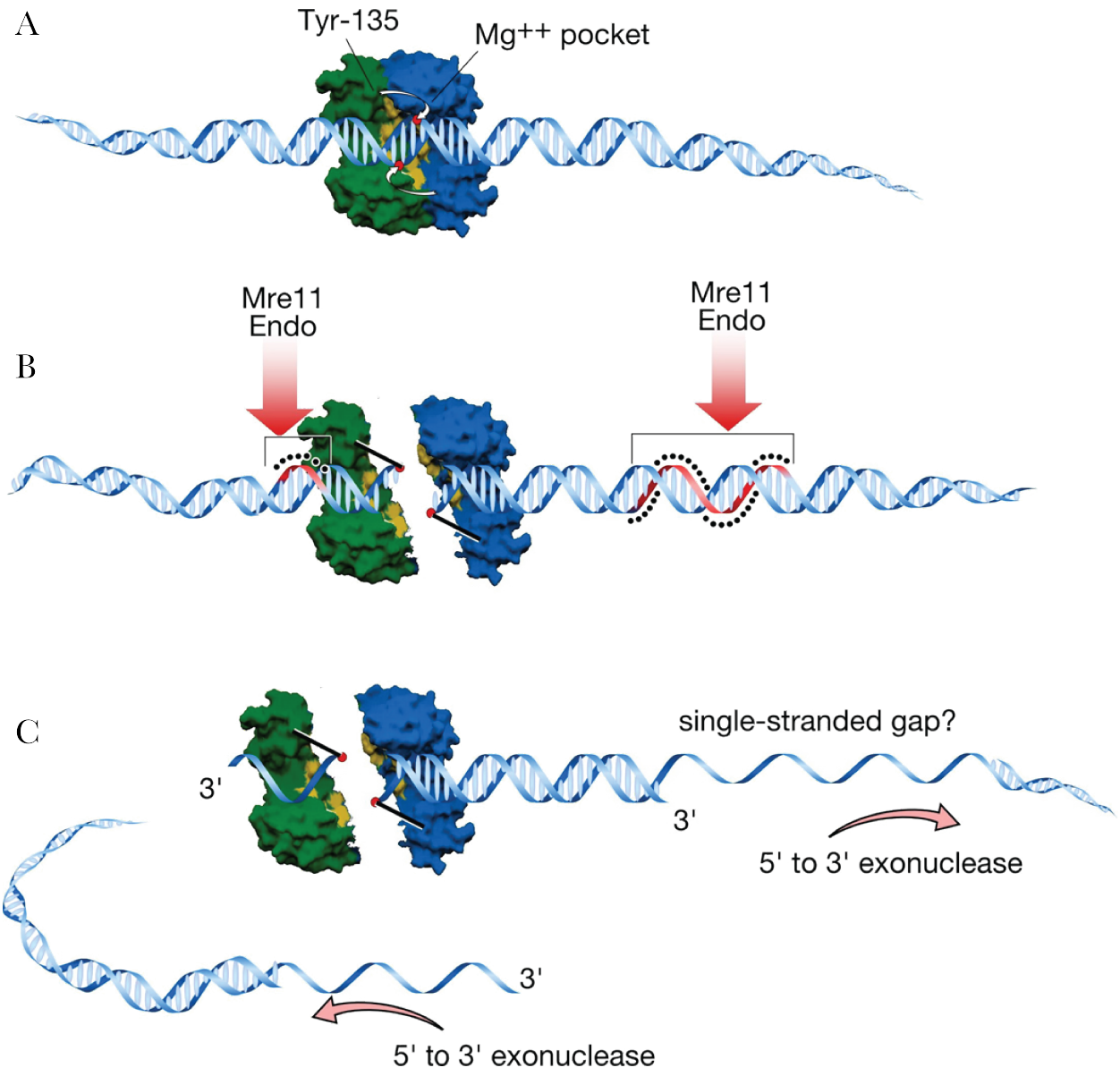

Figure 1.6: **Structural model for Spo11-induced DSB formation.** (**A**) A Spo11 dimer is modeled using a surface rendering of the structure of Methanococcus jannaschi TopoVIA, with one monomer in green and the other in blue. Approximate positions of the catalytic tyrosine of one monomer and the metal-binding pocket of the other monomer are indicated. B-form DNA (drawn approximately to scale) is docked onto the dimer. The metal binding pockets of the two monomers are appropriately spaced to place them in proximity to the scissile phosphates which would give the two-nucleotide 5'overhang known to be generated by Spo11 and TopoVI. The tyrosines in the TopoVIA structure are far from these phosphates; however, so it is assumed that a conformational change in the protein moves the tyrosines into position to cleave (small white arrows). (**B**) Cleavage of the DNA backbone leaves Spo11 covalently attached to the DSB ends. For clarity, the Spo11 dimer interface has been interrupted to show how the two monomers are attached. Analogous separation of the DNA ends is what opens a gate to allow strand passage for TopoVI; whether the Spo11 dimer interface might be disrupted in this manner is not known. After DNA cleavage, Spo11 is released from DSB ends by endonucleolytic single-strand cleavage on either side of the break, most likely mediated by Mre11 endonuclease activity. The strands where nicking occurs are

Figure 1.6: (cont.) colored red and highlighted with the black dots. Asymmetric nick spacing is shown, but the disposition of nicks around individual DSBs is not yet known. (**C**) The DSB ends are further processed by 5'→3' single-strand resection. The shorter Spo11- bound oligonucleotides may be small enough to be readily dissociated, but the longer oligonucleotides would need to be actively unwound. This feature may result in capping of one of the DSB ends, which may in turn influence the behavior of this end during subsequent recombination reactions. This figure was reproduced with permission from Keeney, S., Spo11 and the Formation of DNA Double-Strand Breaks in Meiosis. *Genome Dyn Stab* (2008).

assembly due to the reduced number of Dmc1 foci compared to Rad51 foci in *mei5* and *sae3* backgrounds [147, 148]. Furthermore, one of these studies showed direct interaction of Mei5 and Sae3 *in vivo* by co-immunoprecipitation [147].

The first biochemical work was performed only in the last 10 years with Mei5-Sae3 shown to directly bind RPA and DNA with a preference for ssDNA and prevent inhibition of Dmc1 assembly by free RPA, analogously to Rad52 [149]. More recently, Mei5-Sae3 was shown to preferentially bind DNA fork structures over ss- and dsDNA and to interact with Rad51 through the N-terminus of Mei5. However, Mei5-Sae3 lacked Rad52's homologous-pairing-mediator and ssDNA-annealing activity [150]. This Rad51-Mei5-Sae3 interaction was later shown to enhance Dmc1 D-loop formation [151]. Together, these biochemical studies demonstrate that the Mei5-Sae3 heterodimer is important for Dmc1 assembly, stabilization, and homologous pairing; however, further work is needed to understand the heterodimer's exact mechanism [152].

1.1.6.3 Hed1

During a screen for suppressor mutants of non-sporulating *red1* (a meiotic protein discussed later) strains, a new ORF was found, *HED1* (high-copy suppressor of *red1*), previously undetected due to a one-nucleotide error in the Saccharomyces Genome Database [153]. By

sequencing upstream of the new ORF, they found that *HED1* had a previously identified meiotic-specific gene expression sequence [153, 154]. In this same study, it was shown that *hed1* could suppress *dmc1*'s reduced meiotic DSB repair, Hed1 co-localized with Rad51 at meiotic DSB foci by immunostaining, and Hed1 interacted with Rad51, *in vivo*, by yeast two-hybrid analysis [153]. Hypothesizing that Hed1 might inhibit Rad51 during vegetative growth, Hed1 was expressed through a galactose-inducible promoter, resuilting in a strain more sensitive to MMS than wild-type [153]. Another interesting finding from this study that will be discussed in more detail later was that although *hed1* could suppress *dmc1*'s lack of meiotic DSB repair, the level of crossover events were far less than wild-type, suggesting that Rad51 is capable of repairing meiotic DSB's but not able to form crossover events to the same extent [153]. A subsequent study, discussed in more detail later, expanded on this finding, showing that *hed1 dmc1* lost the meiotic bias towards homologs [155].

The first biochemical characterization of Hed1 came soon after this original genetic study, showing that Hed1 binds ssDNA and directly interacts with Rad51 without interfering with presynaptic filament assembly but blocks its interaction with Rad54 and the subsequent Rad54 mediation behavior discussed earlier [156]. They then showed that this Rad54 blocking occurs *in vivo* by expressing Hed1 mitotically during galactose-inducible DSB formation and demonstrating a lack of Rad54 recruitment to the DSB by ChIP [156].

Recent single-molecule work has shed more light on Hed1's Rad54-blocking mechanism [128]. Fluorescently-labeled Hed1 bound tightly and uniformly to Rad51-ssDNA presynaptic filaments, suggesting a Hed1-Rad51-ssDNA co-filament [128]. Interestingly, this complex did not increase presynaptic filament stability, and Hed1 appeared to dissociate with Rad51 once ATP was depleted and hydrolyzed [128]. This co-filament also prevented the binding

of Rad54, leading to the hypothesis that Hed1 fills in the gaps between Rad51 filaments that Rad54 usually binds to [128]. It was then shown that Hed1 and Rad54 competitively inhibit each other when added in unison on naked Rad51-ssDNA filaments and neither could remove the other once bound to the filament [128]. These results together led to a new meiotic kinetic model where the timing of Rad51, Rad54, Dmc1, and Hed1 expression is critical for determining the presynaptic filament character in early and late meiosis (Fig. 1.7) [128].

1.1.6.4 Hop2-Mnd1

HOP2 was first identified in a screen for mutants that were defective in meiotic gene conversion [157]. It was shown to have a similar transcription profile to another meiotic-specific gene discussed later, *RED1*, and to localize to meiotic chromosomes prior to and during synapsis [157]. The *hop2* mutants arrested in prophase, failed to repair DSBs, and showed high densities of Dmc1 remaining on the chromosomes, suggesting that recombination was not completed [157]. Curiously, *hop2* mutants seemed to show synapsis between nonhomologous chromosomes, and thus, it was hypothesized that its primary role is to ensure homologous pairing during meiosis [157]. *MND1* was discovered genetically and found to be critical for meiosis in the same genome-wide screen used to find *MEI5* discussed earlier [146]. By genetic deletion strains, *MND1* was then shown to be important in early prophase I, chromosome synapsis, and the interhomolog bias during meiosis, suggesting a role with Dmc1 (discussed further later) [158]. Eventually, interdependent colocalization to meiotic chromosomes and coprecipitation assays established the Hop2-Mnd1 complex *in vivo* [159].

Using these genetics studies, a group used bulk biochemistry to show that Hop2-Mnd1

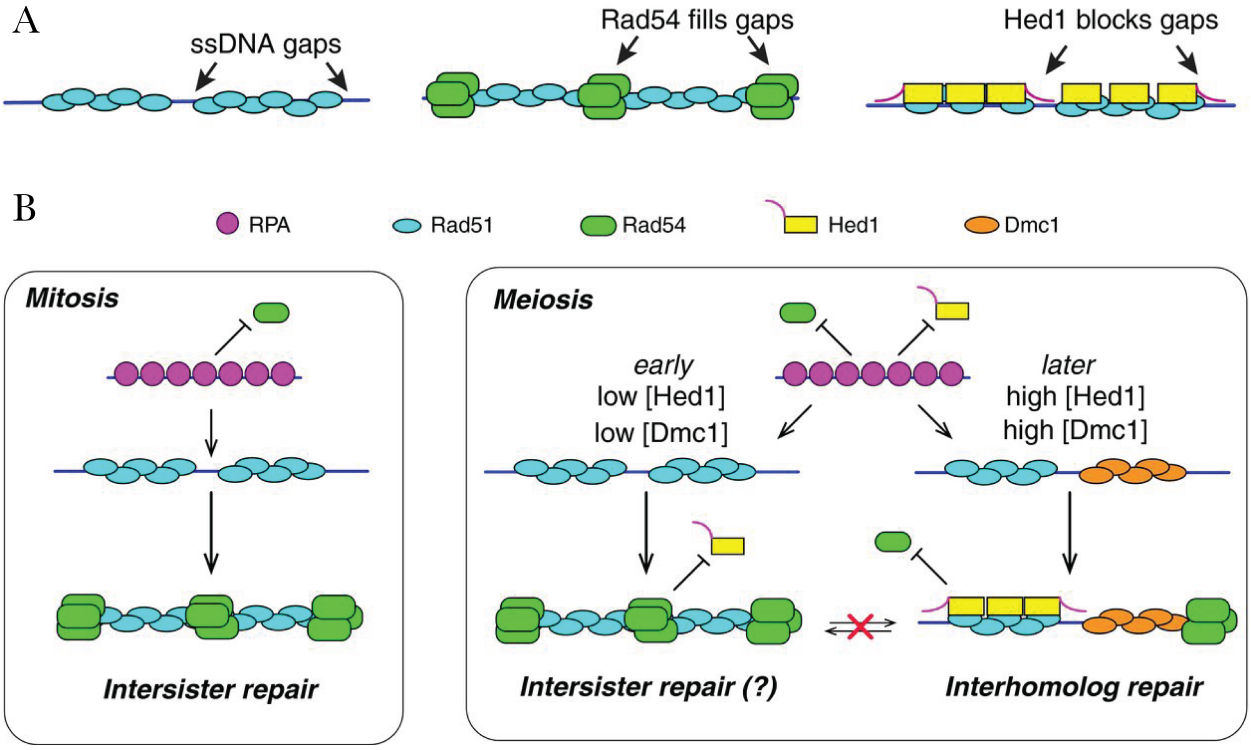

Figure 1.7: **Model for kinetic completion between Rad54 and Hed1 as a mechanism for defining presynaptic filament identity.** (**A**) Cartoon diagram depicting proposed models for Rad54 or Hed1 binding to the Rad51 presynaptic complex. We propose that Rad54 binds to small regions of exposed ssDNA present between Rad51 filaments. This model postulates that while initial recruitment to the filament is dependent on Rad51, stable association of Rad54 with the presynaptic complex is mediated primarily through Rad54 interactions with the ssDNA. In contrast, Hed1 associates directly with Rad51 and does not require stable interactions with the underlying ssDNA for binding stability. Hed1 prevents Rad54 association by blocking access to the small ssDNA gaps present at the ends of the Rad51 filaments, and this inhibition requires the ssDNA-binding amino acids within Hed1. (**B**) Cartoon diagram showing the differences between mitotic and meiotic presynaptic complexes and also depicting that transition from mitotic presynaptic filaments to meiotic presynaptic filaments is controlled by a kinetic competition between Rad54 and Hed1 that is ultimately related to the relative abundance of each protein. During meiosis, Rad54 will bind to the Rad51 presynaptic complexes when Hed1 abundance is low, whereas Rad54 will be blocked from interacting with Rad51 presynaptic complexes when Hed1 abundance is high. Once assembled, transitions between the two types of complexes will only take place if the proteins are actively removed from the ssDNA. Additional details are presented in the main text. This figure was reproduced with permission from Crickard, J.B., et al., Regulation of Hed1 and Rad54 binding during maturation of the meiosis-specific presynaptic complex. *EMBO* (2018).

formed a heterodimer in solution which could bind dsDNA and stimulate Dmc1 D-loop formation by more than 30-fold in a concentration-dependent manner [160]. A later *in vitro* study showed that Hop2-Mnd1 may mediate this stimulation in two ways: i) by stabilizing the Dmc1 presynaptic filament from disassembly and ii) by enhancing presynaptic filament dsDNA capture [161].

A recent crystal structure of full length *G. lamblia* Hop2-Mnd1 heterodimer is a remarkable example of the structure-function relationship (Fig. 1.8) [162]. The dsDNA-binding C-teriminus of each protein features a wing-helical domain, which, as a dimer, was predicted to distort dsDNA base pairs using MD simulations [162]. The N-terminus features three consecutive leucine zipper domains that were modeled to perfectly fit between the grooves of a known RecA presynaptic crystal structure (Fig. 1.8) [162]. Using this structural information, a combined model using known genetic, biochemical, and structure was developed and represented in Figure 1.8 [162].

1.1.6.5 The Synaptonemal Complex (SC)

At the formation of presynaptic filaments, it is important to understand the extreme complexity of the meiotic homology search problem. Not only must a 1kb length of ssDNA find its homologous target within the genome, but it must additionally avoid the sister chromatid and recombine with one of the two homologs (discussed further later) [141]. Thus, architectural organization in the form of the synaptonemal complex (SC) is critical for proper meiotic recombination (Fig 1.9) [163].

Due to its impressive size and organization, the synaptonemal complex has been observed during meiosis for as least one hundred years [164], although work by Moses [165] and

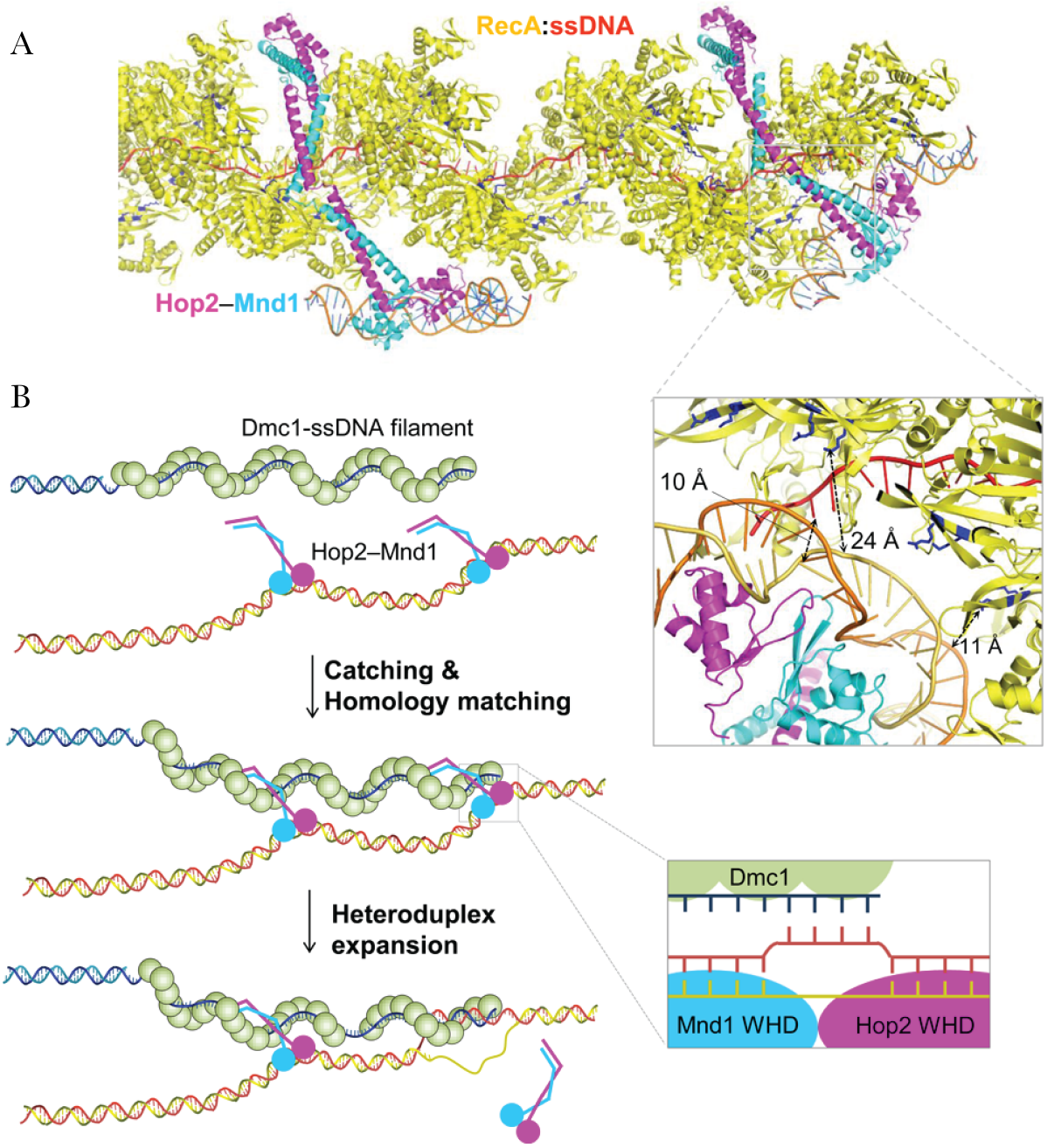

Figure 1.8: **A model for Hop2–Mnd1-assisted strand invasion.** (**A**) A model of Hop2–Mnd1 binding to RecA nucleofilament. The structural model obtained from the MD simulation was fitted into the groove of the crystal structure of RecA (yellow) bound to ssDNA (red) (PDB entry: 3CMU). The Site II residues are shown in sticks (blue), and representative distances are indicated. (**B**) A model for Hop2–Mnd1-assisted strand invasion. The Hop2–Mnd1 molecules catch the Dmc1 nucleofilament, which is then closely juxtaposed to (highly distorted) dsDNA bound to the WHD pair. If sequence matches most preferably at the end of the filament, DNA joint molecule is formed and strand invasion proceeds subsequently. This would require concomitant with Hop2–Mnd1 detachment from both dsDNA and the Dmc1 filament for the propagation of the heteroduplex. This figure was reproduced with permission from Costa, Y., et al., Two novel proteins recruited by synaptonemal complex protein 1 (SYCP1) are at the centre of meiosis. *JCS* (2005).

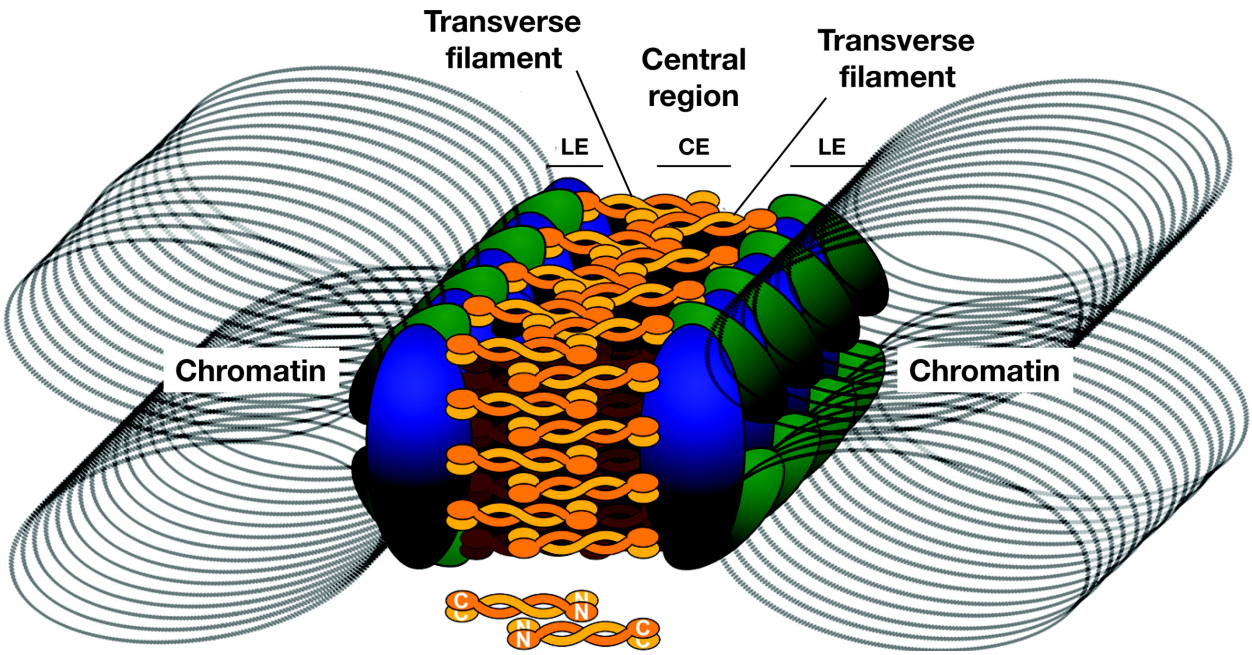

Figure 1.9: **Model of SC structure.** Shown is a cross section of a segment of the SC with lateral elements (LE), transverse filaments, central element (CE), and central region. The arrangement of transverse filament proteins, as determined experimentally for Zip1p and SCP1, is shown at bottom. Also shown is a hypothetical arrangement of cohesins/condensins (blue ovals) and other LE proteins (green ovals) along the LEs. This figure was reproduced with permission from Page, S.L. and R.S. Hawley, The genetics and molecular biology of the synaptonemal complex. *Annu Rev Cell Dev Biol* (2004).

Fawcett [166] in 1956 is often credited for the original description in crayfish and various vertebrate spermatocytes. The wealth of meticulous cytological descriptions of the observed phenomena by light and electron microscope merited hundreds of papers and many reviews well before the widespread adoption of molecular genetics as a tool for understanding the basis of such structures [167, 168]. In order to grasp the significance of this structure and its meiotic arrival, a review of meiosis is essential. Briefly, meiosis involves two rounds of chromosome segregation after DNA replication: i) meiosis I, which separates homologs, and ii) meiosis II, which separate sister chromatids to form haploid gametic cells. These rounds of segregation can be further sub-divided cytologically into prophase (chromosomes condense and recombination occurs), metaphase (chromosomes align across equator of cell), anaphase

(chromosomes begin separating towards opposite poles of cell), telophase/cytokinesis (cytoplasm division and formation of the cleavage furrow). The prophase stage during meiosis I, prophase I, can be further subdivided cytologically into five substages: leptotene, zygotene, pachytene, diplotene, and diakinesis. During leptotene, the first appearance of SC elements occur with axial elements (orange elements in Fig. 1.10) assembling along the length of the each homolog (Fig. 1.10) [169]. By zyogotene, transverse elements (blue and white chains in Fig. 1.10) begin assembling across the homologs in a zipper-like fashion (Fig. 1.10) [169]. At pachytene, the homologs are fully synapsed. This is when crossover events are thought to occur, leading to visible chiasmata by diplotene (Fig 1.10) [169, 170]. The formation of these chiasmata is critical for proper segregation of the homologs [170, 171]. Although complex SC formation is not required for homologous pairing in yeast, the lack of SC formation greatly reduces these pairing interactions [172].

One of the first SC-associated genes identified in yeast was *HOP1* by showing a lack of SC formation by EM in a homozygous *hop1* background [173]. It was later shown that *HOP1* was transcriptionally expressed during meiosis and primary structure analysis showed the existence of a zinc-finger motif that when mutated from a cysteine to a serine, led to a similar phenotype to *hop1 null* [173]. Simultaneously, another gene, *RED1*, was shown to lead to chromosome nondisjunction during meiosis I and significant aneuploidy in the few spores that did form [174]. A follow-up study showed that *red1* chromosomes did not form the SC either [175]. In addition, Red1 is implicated in reducing the interhomolog template bias seen during meiosis, suggesting the importance of the SC in template bias [176].

Eventually, it was shown using fluorescently-labeled antibodies that Red1 and Hed1 localize at meiotic chromosomes, which combined with their early arrival during prophase,

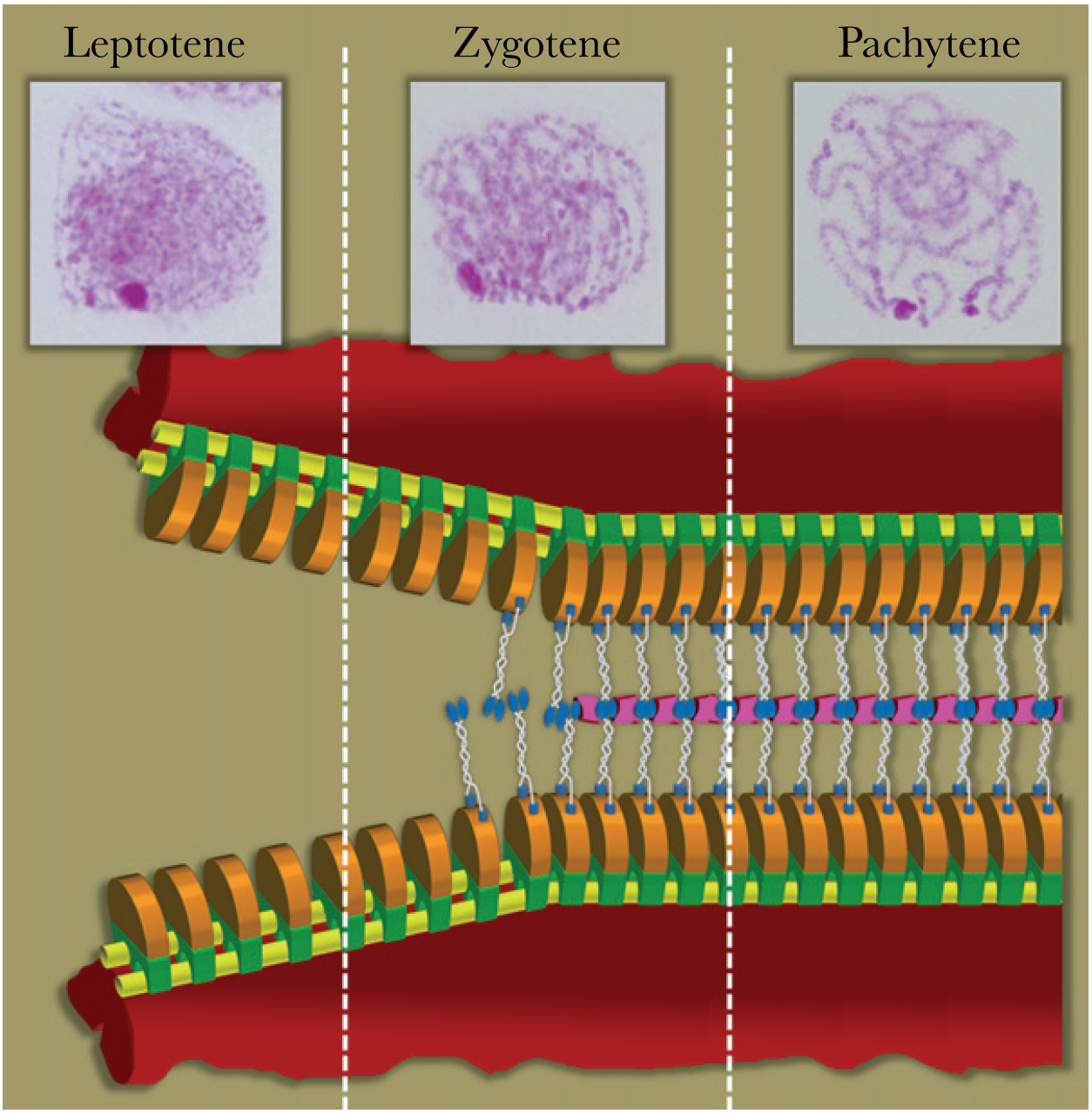

Figure 1.10: **Diagrammatic representation of the development and structure of the synaptonemal complex (SC) throughout the prophase I stages.** The two chromatids of each homologue (red) are associated with the chromosome scaffold (yellow axes) that underlies the position of the cohesin axes (green pieces). The axial/lateral elements (orange pieces) of the SC are assembled at each homologue during leptotene. By zygotene, the homologous chromosomes pair and align provoking that their lateral elements run parallel. At this stage, synapsis initiates as the transverse filaments (white and blue chains) interconnect the lateral elements in a zipper-like structure, which is completed by the arrangement of the central element (purple pieces). At pachytene, the SC extends the entire length of homologues, completing synapsis. Meiotic stages and structures are indicated in the figure. This figure was reproduced with permission from Maloy, S., K. Hughes, and ebrary Inc., Brenner's encyclopedia of genetics. *Elsevier* (2013).

suggested that they formed the lateral elements of the SC (blue elements in Fig. 1.9, orange elements in Fig. 1.10)[177, 178]. A third gene, *ZIP1*, was identified to not properly form SC's during this same period in the late 80s and early 90s [170]. Due to its URS1-like meiotic gene promoter and a clever Zip1-β-galactosidase fusion assay, it was determined that Zip1 is expressed exclusively during meiosis [170]. After primary structure analysis of Zip1 showed an α-helical coiled-coil motif and fluorescent antibodies to Zip1 localized to meiotic chromosomes, Zip1 was hypothesized to be a structural element of the SC (orange elements in Fig. 1.9, transverse elements in Fig. 1.10, represented by blue and white chains) [170]. Beyond homolog proximity, there is additional evidence that the SC can transduce signals through structural changes due to DSBs in the attached chromosomes such that additional DSBs are not made nearby, a concept referred to as interference [171]. Of course, there are many more proteins involved in the assembly of the SC, but these original three provided the molecular genetic evidence for the importance of the SC in homologous pairing, crossover events, template bias and meiotic recombination more generally.

1.1.6.6 Meiosis in Dmc1-less Eukaryotes

Neither *C. elegans* nor *D. melanogaster* have Dmc1 and thus use Rad51 as the recombinase for meiotic recombination [171]. Unlike yeast, these species are able to enagage in homologous pairing without DSB formation. *C. elegans* rapidly aligns homologs during meiosis without the initiation of recombination or synapsis-related proteins [171]. *D. melanogaster* seems to have DSB-independent somatic pairing associations [171]. Together, it appears that homologous pairing is recombination independent in these species [171]. If there are eukaryotic species that can thrive with one recombinase, why did eukaryotes evolve

a separate recombinase for meiosis?

1.1.7 Mitotic versus Meiotic HR

For the most part, mitotic and meiotic recombination differ in their purpose and thus, their respective outcomes (Fig 1.11) [47, 48, 179]. Generally, mitotic recombination is employed to repair spontaneous DSBs with minimal LOH, whereas meiotic recombination is employed to produce chiasmata between homologs from programmed breaks to ensure proper segregation [47, 48, 179]. Consequently, mitotic and meiotic recombination differ in two distinct ways: i) production of noncrossover (NCO) versus crossover (CO) events; ii) homologous template choice for repair [47, 48, 179]. Since the goal of mitotic recombination is to repair spontaneously formed DSBs with minimal LOH, NCO events are heavily biased [47, 48, 179]. In order to ensure NCO events, SDSA is the primary mode of repair for mitotic HR [47, 48, 179]. Meiotic recombination, on the other hand, requires CO events to ensure physical linkages between homologs, and thus, DSBR is the primary pathway [47, 48, 179]. However, DSBR can result in CO and NCO events so, as mentioned earlier, Mlh1-Mlh3 and Sgs1 appear to be employed during meiosis to ensure that CO events occur [44]. In addition, evidence points to the Mms4-Mms81 endonuclease complex as an additional player in CO formation with the central regulator of meiosis I, Cdc5, phosphorylating and activating the complex for resolution during meiosis [45].

Mitotic and meiotic recombination differ in their choice of templates for repair (Fig. 1.11) [179]. During mitosis in diploids and meiosis, a broken chromosome has the choice of three templates: its exact copy sister chromatid and two homologs containing single

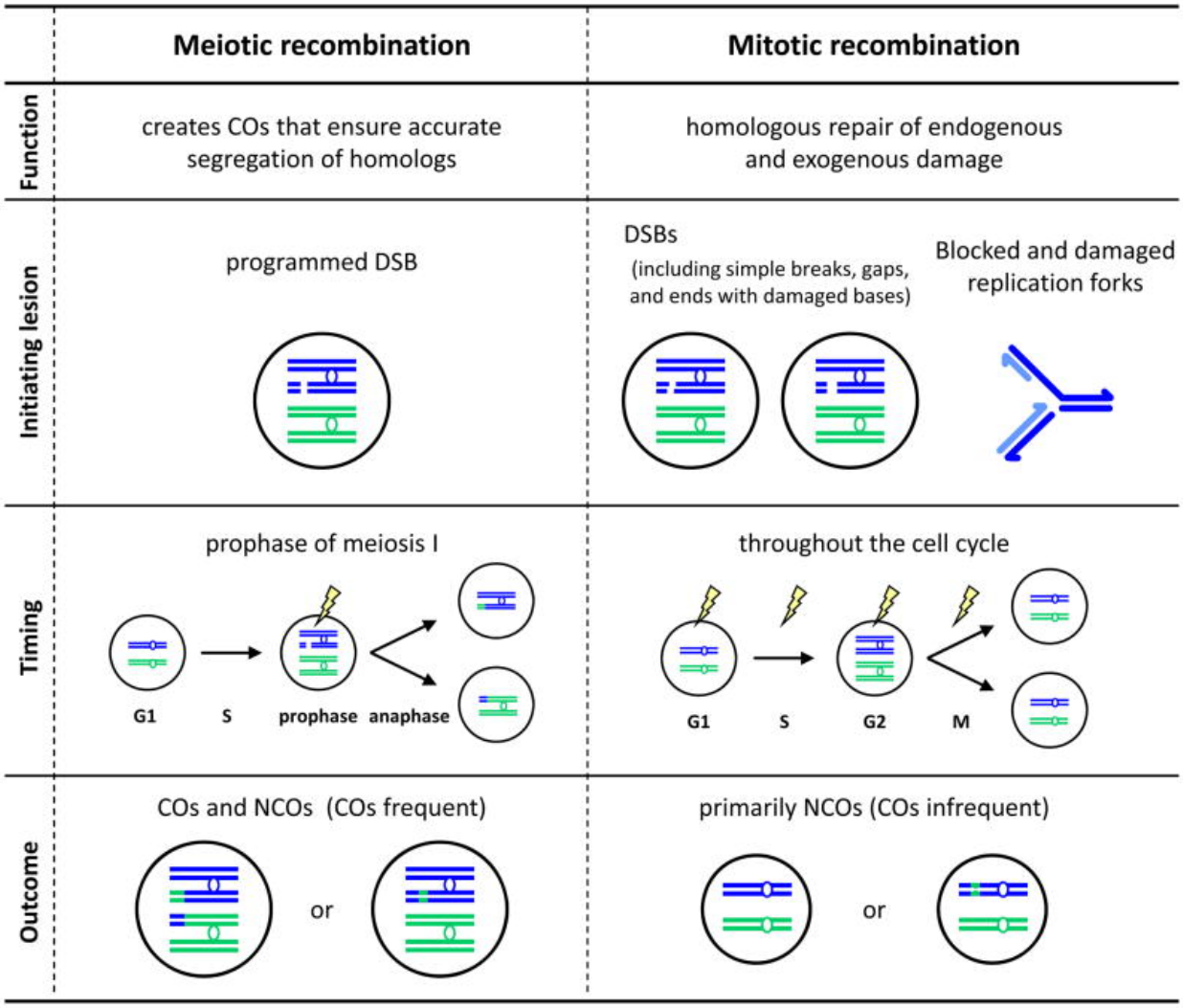

Figure 1.11: **Meiotic versus mitotic recombination.** Meiotic and mitotic recombination are fundamentally different in several aspects, including function, initiating lesions, timing during the cell cycle, and outcome. In this figure, a single pair of homologous chromosomes is shown in each nucleus. Both DNA strands of each chromatid are shown. This figure was reproduced with permission from Andersen, S.L. and J. Sekelsky, Meiotic versus mitotic recombination: two different routes for double-strand break repair: the different functions of meiotic versus mitotic DSB repair are reflected in different pathway usage and different outcomes. *Bioessays* (2010).

nucleotide polymorphisms (SNPs) [179]. In order to prevent LOH, mitotic recombination prefers the sister chromatid as a template with an inter-sister (IS) to inter-homolog (IH) ratio of 4:1, despite the fact that there are 1:2 sister chromatid to homolog substrates for repair [179]. To ensure formation of chiasmata and proper chromosomal segregation, meiotic recombination biases the homolog with a 5:1 IH:IS ratio [179]. Although the factors that contribute to differences in template bias remain poorly understood, these biases appear to be in part mediated by the recombinases involved in meiosis and mitosis [179].

During mitosis, Rad51 is the only recombinase expressed for repair of spontaneous breaks [179]. During meiosis, Rad51 is expressed together with the Rad51 paralog Dmc1 (DMC1 in humans) [179]. Normally, during meiosis, Rad51 and Dmc1 form mixed presynaptic filaments with recent evidence pointing to self-segregation of Dmc1 and Rad51 *in vitro* [180, 181]. However, the meiotically-expressed Rad51 regulatory protein, Hed1, is expressed and inhibits the activity of Rad51 during meiosis [179, 181]. In fact, *hed1* strains appear to decrease the IH:IS bias from 5:1 to 4:1, implicating Rad51 in a IS bias (Fig. 1.11) [179]. Furthermore, *hed1 dmc1* strains show a complete reversal of bias with a 1:9 IH:IS bias during meiotic recombination [179]. In addition, when a dsDNA-binding-deficient rad51-II3A mutant is introduced, the IH:IS bias is seemingly unaffected, pointing to Dmc1's primary role in meiotic recombination [151, 179]. Thus, this work, in aggregate, suggests that Rad51 and Dmc1 have differing properties to facilitate their roles in mitotic and meiotic recombination template bias.

1.1.8 Eukaryotic Recombinases Rad51 and Dmc1

In the 1960s, *E. coli RecA* was the first recombinase identified, playing a crucial role in recombination by UV-sensitivity experiments [182, 183]. It wasn't for another ten years that RecA was purified and shown to have ATP-dependent homologous pairing activity *in vitro* [61]. Thirteen years later, Rad51 [91] and Dmc1 [184] were determined to be the yeast recombinases, by comparing the amino acid sequences with RecA's, demonstrating the accumulation of DSBs in *rad51* and *dmc1* strains, and additionally, showing that Rad51 could bind to mediator, Rad52, and DNA [91, 184]. Two years later, yeast Rad51 was purified and shown to have ATP-dependent homologous-pairing and strand-exchange activity [185]. Yeast Dmc1 was finally shown to have homologous pairing activity seven years later but much weaker than Rad51 and RecA due to incomplete filament formation [186]. It was later determined that Ca^{2+} was necessary to promote longer Dmc1 filaments *in vitro* and comparable activity could be recovered when this divalent ion was included into the reaction [187]. Overall, biochemically, yeast Rad51 and Dmc1 share similar ATP-dependent presynaptic filament assembly, DNA-dependent ATPase activity [185, 187], and strand-exchange activity.

Structurally, both eukaryotic recombinases Rad51 and Dmc1 are remarkably similar [37, 39, 188, 189]. Yeast Rad51 and Dmc1 share 45% sequence identity (55% for humans) [190], both carry Walker A and B ATPase motifs [191]. By transmission electron microscopy (TEM), yeast Rad51 and Dmc1 appeared to form similar right-handed helical filaments across ssDNA with persistence lengths of 543 $\pm$ 45.3 nm and 507 $\pm$ 45.2 nm, respectively, helical pitches of 11.9 $\pm$ 0.8 nm and 11.2 $\pm$ 0.7 nm, respectively, and helical widths of 10.2

$\pm$ 0.8 nm and 10.4 $\pm$ 0.8, respectively [192]. Both recombinases have been shown to stretch DNA by $\sim$1.5 times [192]. Crystal structures of human RAD51 and *E. coli* RecA in pre- and postsynaptic forms have elucidated a rather unique asymmetry in this stretching with each recombinase binding 3-nt of DNA with B-form-like properties and a $\sim$8 Å between each triplet [36, 39]. Presumably, yeast Rad51 and Dmc1 share these properties as well.

1.1.9 Recombinase Loops L1 and L2

The Rad51/RecA family of DNA-binding proteins interact with DNA substrates through two highly conserved loop regions, referred to as loop 1 (L1) and loop 2 (L2). The recombinase loops L1 and L2 were first identified by Thomas Steitz in his *E. coli* RecA crystal structure in 1992 [193]. Although the electron density did not appear due to presumed disorder, the loops were defined as between residues G157-M164 and I195-T209, respectively [193]. When aligned with the available low-resolution EM-structure of the RecA filament, the loops appeared to lie closest to the ssDNA in presynaptic form [193]. Later crystal structures of *E.coli* RecA [36, 194, 195], *S. solfataricus* RadA (L1: R217-R229; L2: M258-H275) [196], yeast Rad51(L1: F290-L296; L2: V328-N348) [37], human DMC1 (L1: R230-R242; L2: N268–R300) [188, 189], and human RAD51 (L1: Y232-L238; L2: V270-I287)[39] confirmed that these loops are conserved across the RecA/Rad51 family of recombinases. Interestingly, although the loops in RecA were unstructured when RecA is unbound to DNA, portions of the loops become ordered once RecA forms a presynaptic filament [36]. In the postsynaptic structure, a few residues in L1 become more ordered and appear to make contacts with the captured DNA [36]. Unfortunately, the RecA L1 significantly diverges from

eurkaryotic L1's so directly comparing them is difficult. Yeast Rad51's loops are disordered in the filamentous form of the protein [37] and hRAD51's loops are disordered in the pre- and postsynaptic forms of the protein [39]. Conversely, hDMC1's L1 is fully structured and L2 partially structured (268-272, 285-300) although hDMC1 is crystalized unbound to DNA and as an octomeric ring [188, 189]. This may imply that the differing amino acids in hDMC1's loop from those in yeast and human Rad51's loops confer a more stable state, contributing to how it functions in meiotic recombination.

One function of these loops is ssDNA binding. Crosslinking [197, 198] and mutational fluorescence spectroscopic analysis [199] confirmed that *E. coli* RecA loop's L1 and L2 intimately interact with ssDNA. One group purified a 20 amino acid peptide from L2 (193-212) with a F203W mutation and used linear dichroism to determine that this peptide alone was capable of stiffening and restricting the motion of ssDNA in a similar manner to full-length RecA [200]. Using the previously discussed postsynaptic RecA crystal structure, another group decided to mutate D161 to an alanine due to its close proximity to complementary DNA strand and implied electrostatic repulsion with the DNA backbone [201]. They found that the D161A mutant lost the wild-type preference for ssDNA over dsDNA and that D161N mutation had the same effect, suggesting that D161 is important for RecA's ssDNA preference and charge is one of the primary reasons for this effect [201]. During a study with yeast Rad51 looking for suppression of *rad57* IR sensitivity, three L2 mutations V328A, P339S and I345T were discovered and shown to have much greater affinity for ssDNA by EMSA and salt titration [106, 202]. In addition, these mutants showed no significant sensitivity to IR in an otherwise wild-type background but a greater sensitivity in an *srs2* background than *srs2* and *RAD51*, further demonstrating these mutants increase presynaptic filament

stability and L2's role in ssDNA binding [106, 202]. A study looking at human RAD51 found that an L2 mutation, F279W, could bind and hydrolyze ATP and bind dsDNA but was significantly defective in ssDNA binding [203]. A dominant mutation in mouse DMC1 L2, A272P, causing male sterility and decreased fertility in females, was found to have a 3-fold reduction in ssDNA binding and no strand-exchange activity[1] [204].

In addition, the loops appear to be involved in recombinase ATPase activity. One study systematically made all 380 single mutations in *E. coli* RecA L2 (195-209) and used lambda plaque, UV sensitivity, and mitomycin C sensitivity assays, finding that any substitution of R196 ablated ATPase activity [205]. Combined with the previously known DNA-binding function of L2, this study led to the hypothesis that L2 is responsible for linking RecA's DNA-binding to its ATPase activity [205]. Similarly, yeast Rad51 F290A (L1) was able to bind ssDNA and dsDNA in an ATP-dependent manner but was ATPase dead, suggesting this amino acid is also important for communicating between the ATP hydrolysis and DNA binding [206]. A third study showed that mutation, K342E, in yeast Rad51 L2 led to DNA-independent ATPase activity, again suggesting L2's role in linking ATP-binding, DNA-binding, and ATP hydrolysis [207, 208]. Mutations in *Methanococcus vitae* RadA's L2, H276N and H280N, reduced and ablated ATPase activity [209].

Recombinase loops have also been found to contribute to dsDNA-capture and strand-exchange activity. The yeast Rad51 L2, K342E, mention earlier, caused a defect in dsDNA capture, implicating L2's role in dsDNA capture [207, 208]. Interestingly, this mutation led to filament formation in the absence of DNA by direct observation on EM [207, 208]. In addition, yeast Rad51 L2 mutants, V328, P339S and I345T, all showed greater *in vitro*

[1] ATPase activity was not tested.

strand-exchange activity than WT [106, 202]. A mutation in human RAD51 L1, R235E, could bind ssDNA but not bind to dsDNA by gel shift or to dsDNA single-molecule curtains [210]. In this same study, hRAD51 L2 K284A did not affect assembly or binding to dsDNA curtains but was deficient in a bulk strand-exchange assay [210]. During a mutational fluorescence spectroscopic analysis, hRAD51 L1 mutant, D231W, had normal ssDNA- and dsDNA-binding activity but had no strand-exchange activity [203]. Additionally, in this study, L1 mutant, Y232W, had normal ssDNA binding but had a dsDNA-binding defect and had no strand-exchange activity [203]. A series of studies using site-specific linear dichroism and modeling found hRAD51 Y232 is important for intercalating stretched DNA triplets in presynaptic ssDNA, and R235 is near the phosphate backbone of captured strand. [211–213].

Overall, structural and biochemical studies have demonstrated that recombinase loops L1 and L2 are intimately involved in the critical functions of ATP, ssDNA, and dsDNA binding, ATP hydrolysis, and strand-exchange activity. Many independent studies have shown that L1 and L2 are proximal to ssDNA and captured DNA in pre- and postsynaptic structures. Thus, any potential differences between Rad51 and Dmc1 at these loops could translate into differences in how Rad51 and Dmc1 engage captured dsDNA and suggest why eukaryotes evolved an additional meiotic recombinase.

1.1.10 Conclusion

The hypothesis-driven scientific method requires independent discoveries across many fields that converge on common truths of nature. This is no different in the case of homologous recombination for which decades of work from initial studies using molecular genetics

laid the foundation for biochemical, structural biological, and single-molecule fields to uncover the vastly complex coordination of the many proteins involved in the HR pathway. Yet, questions still remain. The focus of this work is to understand why eukaryotes evolved two recombinases for mitotic and meiotic recombination. Developments in single-molecule biophysics have allowed high-throughput analysis of recombinase on ssDNA curtains (Chapter 2), revealing subtle differences between Rad51 and Dmc1 in response to mismatches in captured dsDNA that were otherwise lost in bulk biochemical assays (Chapter 3). By altering additional properties of the captured dsDNA, we test the limits of Dmc1's tolerance towards various sequence imperfections in order to understand the potential mechanism of this tolerance (Chapter 4). Using primary and crystal structure alignment, we predict that recombinase loop L1 is responsible for this differential behavior through single-molecule and *in vivo* studies (Chapter 5). Finally, through the combination of bulk biochemistry and single-molecule techniques, a novel function for BARD1-BRCA1 in facilitating human RAD51 dsDNA capture will be shown (Chapter 6).

1.2 DSBR and Disease

1.2.1 Introduction

Unrepaired double-strand breaks pose an incredible hazard to human health as seen in the plethora of diseases related to mutated DNA-repair proteins. These diseases not only reveal the importance of studying DNA repair, but their common disease signs, symptoms, and imaging findings suggest human cell-types where DSBR plays a disproportionately large

role. Hypomorphic autosomal recessive mutations in repair proteins often support normal overall health but a higher propensity for developing cancer later in life if the patient is heterozygous for the gene. However, in the cases of homozygosity or dominant mutations, the prognosis is much worse, leading to an almost inevitable death by cancer early in life with additional immunodeficiency, infertility, and neurological abnormalities. Since most of the syndromes discussed have overlapping pathophysiology, diseases will be organized based on their "primary" pathophysiology with additional discussion of their mechanistic cellular pathways, diagnosis, prognosis, and treatment. When applicable, footnotes will note yeast homologs, which can be referenced in the previous section for more detailed mechanistic details. Footnotes will also contain medical definitions when appropriate.

1.2.2 DSBR and Oncology

1.2.2.1 BRCA1/2 and Cancer Risk

Due to its role in maintaining genetic integrity, DSBR is critical for preventing the development of malignant neoplasms. When individuals are heterozygous for recessive mutations causing hypomorphic DSBR proteins, the only resulting phenotype may be a higher rate of malignancy several decades after birth. Two of the most well-known tumor suppressor genes within the DSBR pathway are *BRCA1* and *BRCA2*. Amongst all women, the prevalence of *BRCA1* mutations is about 1 in 800-1400, whereas *BRCA2* has a prevalence of 1 in 450-800 [214]. This prevalence is as high as 2.5% amongst Ashkenazi Jews [215]. For both BRCA1/2, the lifetime risk of breast cancer is 45-80% [214]. For BRCA1, the lifetime

risk for ovarian cancer is 45- 60% whereas for BRCA2, this risk is slightly lower at 11-35% [2] [214]. In fact, between the high risk and prevalence, *BRCA1/2* mutations account for 2-6% of breast cancers and 10-15% of epithelial ovarian cancers [214]. Most clinically relevant *BRCA1/2* alterations result in an inactivated protein due to missense, frameshift, nonsense, or splice mutations [214].

Cells without BRCA1 or BRCA2[3] are sensitive to DNA-damage, and these proteins have been shown to play a critical role in repair [216]. BRCA1 has been implicated in many cell functions including transcription regulation, cell-cycle checkpoint activation, and DNA repair, resulting in critical tumor suppressor activity [216]. During HR, BRCA1 is thought to aid in 5'→3' resection of DSBs to generate 3'-ssDNA overhangs directly through interactions with CtIP[4] and the MRN[5] complex and indirectly through exclusion of anti-resection protein, 53BP1 [216]. BRCA1 has also been shown to aid in RAD51 loading by recruiting BRCA2 to DSBs [217]. Most recently, in collaboration with the Patrick Sung Lab at Yale University, this work will later (Chapter 6) show how, in complex with BARD1, BRCA1 promotes RAD51 homologous dsDNA capture without affecting the stability of the dsDNA on the postsynaptic filament [218]. BRCA2 is modeled to facilitate RAD51 nucleation on ssDNA and stabilization of the RAD51 presynaptic—essentially the human analog to yeast Rad52 [219].

Since the primary phenotype of *BRCA1/2* mutations is malignancy several decades

[2]Lifetime risk of breast cancer and ovarian cancer amongst general female population is 1 in 8 and 1 in 70, respectively [214]

[3]Yeast Rad52 is considered BRCA2's functional equivalent although not homologous.

[4]Yeast Sae2 homolog.

[5]Yeast MRX homolog.

into life, early detection is critical to long-term survival of patients. Reasons for genetically testing individuals can be a history of breast or ovarian cancer in the family or Ashkenazi Jewish ancestry [214]. For those individuals who carry a mutation, prophylactic[6] bilateral mastectomy[7] and bilateral salpingo-oophorectomy[8] (BSO) can be an extremely effective preventative measure [214, 220]. In one study, BSO was shown to have a 96% reduction in *BRCA*-related gynecological cancer [220]. Pre-menopausal patients who receive a BSO will undergo early-onset menopause and can receive hormone replacement therapy[9] (HRT) to ease symptoms [214]. Many patients may instead opt for surveillance, which can be moderately effective for breast cancer through mammography and physical breast exams; however, no effective screening strategy has been found for ovarian cancer [214]. Once ovarian cancer is found, the classic standard of care is surgical removal by BSO, total abdominal hysterectomy or other procedures depending on cancer stage and then postoperative chemotherapy with combination platinum (e.g. cisplatin) and taxane (e.g. paclitaxel) compounds [220]. More targeted therapies, such as bevacizumab[10], pazopanib, nintedanib, and PARPi, in combination with surgery and/or chemotherapy have been shown to improve outcomes [221]. The standard of care for breast cancer is similar to ovarian with additional potential targeted therapies including olaparib, veliparib, talazoparib, and rucaparib [222].

A still lingering question in the field of *BRCA1/2* research is if these proteins are

[6]A medical procedure or treatment to prevent probable disease

[7]Partial or complete removal of breast tissue.

[8] Concatenation of salpingectomy, removal of fallopian tubes, and oophorectomy, removal of ovary

[9]Typically a mixture of estrogen and progesterone.

[10]New immuno/targeted therapies are often referenced as biologics. Suffix -mab, means it is a monoclonal antibody, -zumab: humanized antibody; -umab: human; -ximab: chimeric; -omab: mouse; -amab: rat. Suffixes with –ib typically indicate a small-molecule inhibitor.

important for DNA-repair why do the malignancies tend to be localized to gender-specific and hormone-regulated tissues? Little is known about BRCA2's role in this gender bias; however, studies trying to connect these dots have shown that BRCA1 serves an inhibitory role in activation and transcription of the estrogen receptor, providing negative regulation to the gene expression and cell proliferation from estrogen [223]. In addition, BRCA1 appears to reduce transcription of progesterone receptors, PRA and PRB, and has a direct interaction with the progesterone receptors, leading to polyubiquitination and protein degradation [223]. BRCA1 also appears to be directly involved in estrogen biosynthesis regulation by decreasing aromatase expression through the ovary-specific promoter PII, although the exact mechanism is not entirely understood [223]. This finding was confirmed by a study showing higher serum estradiol and progesterone levels in postmenopausal women with *BRCA1/2* mutations [224]. This increase in steroid hormone likely promotes breast epithelial cells and, combined with deficient HR repair, could bias malignancies toward these hormone-regulated tissues [225]. Together, these results begin to tackle the connection between *BRCA1/2* and female hormone-regulated tissues.

BRCA1/2 have also been implicated in fertility. *BRCA1* mutations have been shown to reduce fertility in male mice but not females [226]. BRCA1 appears necessary for proper formation of X-pericentric heterochromatin and implicated in chromatin changes to produce a functional XY body [226]. As with RAD51, BRCA2 has been shown to promote DMC1's strand-exchange activity [227]. Together, these results suggest that females with *BRCA1/2* mutations would have reduced fertility, but retrospective studies have suggested the opposite. Women with *BRCA1/2* mutations bear more children, have shorter intervals between births, and reproduce later in life than their *BRCA1/2 WT* counterparts [228]. One potential

explanation is that *BRCA1/2* mutations have been shown to prevent normal telomerase shortening with age [228, 229]. A separate study looking at telomere length and *in vitro* fertilization (IVF) showed that eggs with longer telomeres were significantly more likely to conceive than those with shorter telomeres [228]. Another study looking at 50-year-old women showed a positive correlation between telomere length and reproductive lifespan [228]. Although the evidence is still weak, the hypothesis of the intersection of *BRCA1/2*, telomeres, and fertility seems quite compelling.

1.2.2.2 RAD51 and Cancer Risk

Rad51 null mutations in mice results in embryonic lethality [230], thus, loss of function mutations in *RAD51* rarely result in viability. Interestingly, there have been a few cases of hypomorphic mutations where heterozygous *RAD51* phenotype only increases cancer risk. One of these mutations is RAD51 R150Q, which was found in two patients with a strong family history and bilateral breast cancer [231]. This mutant protein was later purified and shown to have WT-like ATPase activity and similar ssDNA- and dsDNA-binding preference in low (45mM) and high (200mM) NaCl conditions. However, the RAD51 R150Q mutant showed faster migration on ssDNA-binding gel-shift assay as compared with WT [232]. The subtly of the defect might explain the minor phenotype as compared with the A293T mutation. This study also identified two more mutations, D149N and G141D, associated with tumorigenesis, which along with R150Q, mapped to the Schellman loop of RAD51 [231, 233].

A study looking at all three mutations showed that R150Q, contrary to the previous group, and G141D had reduced ATPase activity, whereas D149N has similar activity to WT [233]. Interestingly, their strand-exchange activity was similar to WT [233]. The only

perceivable difference was a noticeable structural defect of G151D and R150Q filaments on ssDNA by EM, showing less stiff filaments and pronounced differences in the migration of mixed presynaptic filaments with WT [233]. Together these results point to the potential of these mutants to negatively impact the structure and stiffness of the presynaptic filaments, a property hypothesized to be important for yeast Rad51 presynaptic filament mobility [234].

In addition to mutations within the *RAD51* ORF, a 135G>C polymorphism in the 5'-untranslated region of *RAD51* has been shown to reduce mRNA transcription and cause an increase in risk for breast cancer [235, 236]. Beyond germline mutations, there is also evidence of two mutations in RAD51, Q268P and Q272L, which occur during tumorigenesis, further increasing genomic instability [237]. These mutations were found to have reduced ATPase, DNA-binding, and strand-exchange activity along with a dominant negative effect on strand-exchange activity when WT and mutant mixed filaments were formed [227]

1.2.2.3 Fanconi Anemia

Individuals with a dominant negative *RAD51* mutation or homozygous for *BRCA1/2* develop a condition called Fanconi Anemia (FA). Patients with FA present with developmental abnormalities, progressive bone marrow failure, and a strong disposition to cancer [238]. FA is generally autosomal recessive and quite genetically heterogeneous, resulting from mutations in one of sixteen different *FANC*[11] genes (*FANCA*-*FANCT*), all part of the DNA-damage repair pathway [239, 240]. Mutations in *FANCA*, *C*, and *G* (*XRCC9*) represent 85% of FA cases and mutations in *FANCB*[12], *D1* (*BRCA2*), *D2*, *E*, and *F* represent 13%

[11]Many of these *FANC* genes have more common names (e.g. *FANCD1* = *BRCA2*) and will be noted in parenthesis when appropriate.

[12]Only X-linked, non-autosomal FA gene [240].

[240]. Many of these genes express proteins within the FANC "core complex" involved in DNA-damage signaling by ubiquitination or independently signal and process DNA-damage upstream of many of the main HR and NHEJ proteins discussed earlier [240].

Due to FA's genetic complexity, proteins within the FANC family involved primarily in HR will be the focus in this section, including FANCD1 (BRCA2), FANCU (XRCC2), FANCO (RAD51C), and FANCR (RAD51). The most common FA protein of this group is FANCD1 (BRCA2), the mechanism of which was discussed earlier. A review of 30 reported biallelic mutations showed that the most severe and fast majority of cases resulted in a frameshift or truncation [239]. Within this group, there were recurrent congenital birth defects, including microencephaly (Fig. 1.12A), café-au-lait spots (Fig. 1.12B) , short stature, abnormal thumbs, and renal anomalies [239]. The most severe case was a patient homozygous for the 1548del4 mutation that had VATER[13] features, a Wilm's tumor[14], and a neuroblastoma by age 3 [239]. The prognosis can vary greatly depending on the extent of protein dysfunction. In terms of treatment, chemotherapy options are limited due to the progressive bone marrow dysfunction and sensitivity to chemotherapeutic agents, such as diepoxybutane and mitomycin C [240].

Due to the infrequency of carriers, the more rare forms of FA are often the result of consanguinity [241]. One of these rare forms is due to mutations in *FANCO* (*RAD51C*), a gene expressing a RAD51 paralog part of the BCDX2 and CX3 complexes [241, 242]. A subcomplex of BCDX2, the BC (RAD51B-RAD51C) complex, has been shown to be important in RAD51 loading when put in competition with RPA for ssDNA binding [243]. The CX3

[13]VATER features: vertebral defects, anal atresia, cardiac defects, tracheo-esophageal fistula, renal anomalies, and limb abnormalities.

[14]Primary renal cancer commonly found in children

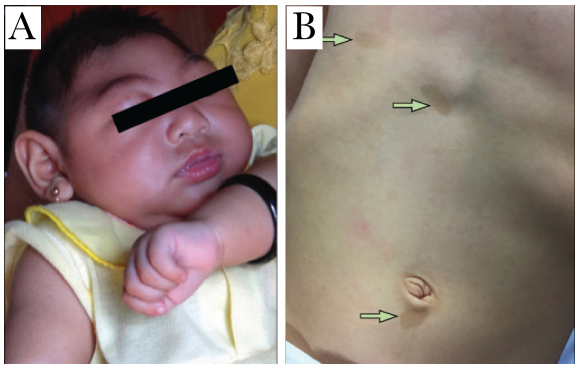

Figure 1.12: **Common features of Fanconi Anemia.** (**A**) Microencephaly of 4-mo child. (**B**) Café-au-lait spots. A was adapted with permission from Moi, M.L., et al., Zika virus infection and microcephaly in Vietnam. *Lancet Infect Dis* (2017). B was adapted with permission from Klein, C., E. Haraux, and R. Gouron, Pathological tibia fracture in an 18-month-old child. *The Lancet Child & Adolescent Health* (2017).

(RAD51C-XRCC3) complex has been shown to be involved in Holliday Junction (HJ) processing with mutations in either *XRCC3* or *RAD51C*, causing decreased HJresolvase activity [242]. Four children from first cousins who both carried a FANCO R258H mutation had very severe phenotypes [241]. Both the second and third child showed sensitivity to mitomycin C and died 2 days and 2 months after birth, respectively, with extensive congenital abnormalities including absent and vestigial thumbs, congenital heart defects, imperforate anuses, and hydronephrosis[15] [241]. The fourth pregnancy resulted in miscarriage after 11 weeks [241]. At the time of the study, the fifth child was still alive at age 10 but had sensitivity to diepoxybutane and mitomycin C with congenital abnormalities including short stature, bilateral radial hypoplasia[16], anal atresia[17], bilateral cryptorchidism[18], small genitalia, bilateral cystic dysplasia of the kidneys[19], and chronic renal failure [241].

FANCU (XRCC2) is another RAD51 paralog and part of the BCDX2 (RAD51B-RAD51C-RAD51D-XRCC2) complex and DX2 (RAD51D-XRCC2) subcomplex. A study looking at RAD51 foci post-gamma radiation demonstrated that the BCDX2 acts upstream

[15]Kidney swelling due to urine accumulation generally because of a blockage downstream.

[16]Underdevelopment of the radius bone.

[17]Imperforate anus.

[18]Absence of one or both testicles.

[19]Formation of cysts in the kidney.

of presynaptic filament formation, where as the CX3 complex acts downstream. Although, its exact role in HR is still not entirely understood [244]. Similarly to RAD51, the DX2 subcomplex has been shown to form a multimeric ring structure in solution, assemble into a filamentous structure on ssDNA, and have strand-exchange activity *in vitro* [245]. When two healthy first cousins with a *FANCU* truncating nonsense mutation conceived a homozygous child, he was born with left facial nerve palsy (paralysis), microcephaly (Fig. 1.12), bilaterally absent thumbs, and an ectopic left kidney [246]. He was later found to have severe growth deficiency, continued microcephaly, and sensitivity to diepoxybutane [246].

The first discovered autosomal dominant mutation causing FA was RAD51 A293T [238]. The *de novo* mutation caused growth retardation, microcephaly (Fig. 1.12), hydrocephalus, skeletal abnormalities, an imperforate anus, and a left testicular malformation when examined at 2.5 years of age [238]. The researchers then recombinantly expressed the protein in *E. coli* and showed that the protein had a lower ATPase activity that was not stimulated by DNA, formed short and distorted filaments by scanning force microscopy (SFM), and inhibited RAD51 WT from binding ssDNA and dsDNA, explaining the dominant negative effect [238]. More recently, it was found that this A293T mutation prevents replication fork protection from MRE11-mediated degradation [247].

1.2.2.4 Bloom, Werner, AND Rothmund-Thomson Syndromes

The last set of rare autosomal recessive diseases discussed in DSBR and oncology derive from mutations in the *RECQ* family of SF2 helicases, *BLM*, *WRN*[20], and *RECQ4*, which

[20]Yeast Sgs1 homolog.

cause Bloom, Werner, and Rothmund-Thomson[21] syndromes, respectively [248]. Common features of RECQ helicases are their binding to ssDNA, movement in the 3'→5' direction, and ATP- and ssDNA-dependent unwinding of dsDNA [248]. However, these helicases differ in the types of DNA structures that they bind. For example, BLM, in complex with topoisomerase IIIα, RMI1 and RM2 (BTR complex), has been shown to promote dissolution of double HJs, suppressing crossover events between sister chromatids [249]. Patients homozygous for mutations in BLM have pronounced dwarfism due to pre- and postnatal growth retardation, congenital telangiectatic erythema (Fig. 1.13A), immunodeficiency, male infertility, and an extremely high risk of cancer [250]. Generally, patients do not survive past childhood, with most succumbing to cancer [248].

The WRN helicase has 3'→5' exonuclease activity at the N-terminus and appears to localize at stalled replication forks in response to DNA-damage agents, such as, camptothecin (CPT), etoposide, 4-nitroquinolin-N-oxide and bleomycin, facilitating repair by HR [251]. Patients with Werner syndrome typically present with hyperkeratosis (Fig.1.13B), "bird-like" facial features (Fig 1.13C), alopecia (baldness), juvenile bilateral cataracts, diffuse arteriosclerosis, telangiectasia (Fig. 1.13D), diabetes mellitus, and hypogonadism [252]. Patients with Werner syndrome have a far better outcome than those with Bloom syndrome with a median age survival of 54 years. Atherosclerotic cardiovascular disease and cancer are the most common causes for death [253].

RECQ4, also known as RECQL4, is markedly different from BLM and WRN in that it is the only RECQ helicase in mitochondria and appears to be expressd disproportionately

[21]RECQ4 mutations are also the cause of RAPADILINO Syndrome and Baller-Gerold-Syndrome (BGS) and have similar clinical findings to Rothmund-Thomson syndrome.

Figure 1.13: **Common features of RecQ-related syndromes.** (**A**) Telangiectatic erythema of forehead. (**B**) Palmer hyperkeratosis. (**C**) "Bird-like" facial features of teenager with Seckel Syndrome. (**D**) Ocular telangiectasia. A was adapted with permission from Martinez, C.A., et al., Adenocarcinoma of the Right Colon in a Patient with Bloom Syndrome. *Case Rep Surg* (2016). B adapted reproduced with permission from Sehgal, V.N., et al., Hand dermatitis/eczema: current management strategy. *J Dermatol* (2010). C was adapted with permission from Arora, S., B. Ghai, and V. Rattan, Anesthetic management of a child with Seckel syndrome for multiple extractions and restoration of teeth. *J Anaesthesiol Clin Pharmacol* (2012). D was adapted with permission from Alakloby, O.M., et al., Ataxia telangiectasia with abnormal cellular immunity. *International Journal of Case Reports and Images* (2012).

in thymus and testis cells during the S phase [248, 254]. The N-terminus has been shown to be responsible for specific binding to RNA, HJ, and G-quadruplex DNA [254]. At birth, individuals with Rothmund-Thomson syndrome (RTS) have typically unremarkable findings but the RTS rash begins to manifest within three to six months of birth, eventually developing into poikiloderma[22] [255]. Eventually, the clinical findings evolve into small stature, alopecia, premature aging, cataracts, skeletal abnormalities and a high risk for neoplasm, especially osteosarcoma [255]. Patients with RTS have the best outcomes of the *RECQ*-related diseases with an unaffected life span assuming cancer does not develop [256].

1.2.3 DSBR and Immunology

1.2.3.1 V(D)J Recombination

Adaptive immunity is critical for humans to combat short-lived, rapidly evolving pathogens. With the ability to create a combinatorial library of 108 unique receptors, V(D)J recombination is the means by which progenitor T and B cells can recognize various antigens produced by these evasive pathogens within a single lifetime of a human [257]. There are six total loci involved in this process: three immunoglobulin (Ig) loci, one heavy chain and two light chains, κ and λ, and three TCR loci, α/δ[23], β, and γ [257]. Each locus has multiple V (variable) and J (junction) segments and sometimes D (diversity) segments with the final recombination products containing one of each, creating the combinatorial diversity.[24] Immunoglobulins (Ig) have two heavy and two light chains with the heavy chain containing

[22]Skin condition consisting of hypopigmentation, hyperpigmentation, telangiectasias and atrophy.

[23]TCRδ is located within the TCRα locus [258].

[24]Constant regions are part of immunoglobulin and TCR development, however, processed after transcription and polyadenylation by RNA splicing [258].

V, D, and J segments while light chains contain only V and J [257]. Similarly, TCRβ and δ have all three segments while TCRα and γ contain only V and J [259]. This process of removing and joining segments is initiated by RAG1/2 which introduces DSBs at the recombination-signal sequences (RSS) between segments and recruits Ku70/80, initiating NHEJ [257].[25]

1.2.3.2 Severe Combined Immunodeficiency (SCID)

When RAG1/2 or downstream proteins in NHEJ are mutated, the resulting syndrome is called T-/B-/NK+ Severe Combined Immunodeficiency (SCID) due to the absence of developed B and T cells but presence of natural killer cells. SCID affects 1 in 58,000 people in the U.S with a 1 in 2000 incidence amongst certain Native American populations and is generally autosomal recessive with some dominant and X-linked exceptions [260]. At least a quarter of SCID cases are of the T-/B-/NK+ SCID subtype and can be further subdivided into non-radiosensitive, *RAG1/2*, and radiosensitive, downstream NHEJ proteins [260, 261]. Depending on the severity of mutations, the T-/B-/NK+ SCID course can vary greatly, however, the general immunodeficiency findings are similar. For the most part these patients present with life-threatening bacterial, viral, and fungal infections early in life, chronic diarrhea, and failure to thrive [261, 262]. Interestingly, findings can be delayed by several months after birth due to residual maternal antibodies [261]. The only cure for the disease is hematopoietic cell transplantation (HCT) with 50-88% longterm[26] survival and RAG1/2 deficiency having the best outcomes for patients [261]. Radiosensitive forms of

[25]NHEJ is also required for class switch recombination (CSR) in B-cells (i.e. from IgG to IgM) but follows a separate DSB initiation mechanism from RAG1/2 [257].

[26]Patients were followed for up to 25 years after birth.

T-/B-/NK+ SCID have additional symptoms that lead to their poorer prognosis.

1.2.3.3 Artemis Deficiency

The most common radiosensitive T-/B-/NK+ SCID is Artemis deficiency [263]. Artemis is a protein involved in dsDNA processing after the formation of DSBs, eventually allowing ligation of dsDNA ends. Alone, Artemis has 5'→3' exonuclease activity, however, in complex with DNA-PKcs, can open hairpins, such as those formed by RAG1/2, and endonucleolytically cleave 5'- and 3'-ssDNA overhangs [264]. Based on the clinical findings of the disease, it would appear that Artemis's function is most critical in response to V(D)J recombination and ionizing-radiation repair since immunodeficiency and radiosensitivity are its defining clinical features. Malignancies are not common in these patients although often patients either are cured or die within one year of birth[27] [261]. There has also been one reported case of a DNA-PKcs missense mutation (L3062R), which only seemed to affect its interaction with Artemis and had similar clinical features [265]. In terms of cures beyond HCT, recent advances in gene therapy have demonstrated that Artemis deficiency can be corrected using the lentivirus and potentially a transformative technique for other SCID's [266].

1.2.3.4 LIG4 Syndrome

Two very rare forms of radiosensitive SCID are DNA Ligase IV deficiency or LIG4 syndrome with only 28 known global cases and XLF/Cernunnos deficiency [263, 267]. DNA

[27]One exception is the hypomorphic forms of Artemis cause a less severe immunodeficiency and lack of radiosensitivity, however, partially functional B-cells allow a predisposition to Epstein-Barr virus (EBV)-associated lymphomas [263].

ligase IV in complex with XRCC4 and XLF is responsible for ligation of DNA ends during NHEJ [268, 269]. Thus, due to XLF and DNA ligase IV's widespread function, the clinical features of LIG4 syndrome can be more extensive than those of radiosensitive SCID depending on the severity of the mutation with some individuals showing no phenotype except for radiosensitivity and susceptibility to malignancy [267]. The more severe forms can present with serious Seckel-like musculoskeletal abnormalities (Fig. 1.13C), hypogonadism, and microcephaly [267]. Treatment for these severe forms is limited and supportive, although there have been isolated cases of successful HCT [267, 270]. HCT's lack of efficacy for these deficiencies may be due in part to the frequent occurrence of acute Graft-Versus-Host Disease[28] (GVHD) [263].

1.2.4 DSBR and Neurology

1.2.4.1 DSBR and Neuroscience

The neurological findings in patients with mutated DSBR genes, in many ways, are the most fascinating since unlike oncology, immunology, and fertility, it is not immediately obvious why DSBR plays such a critical role in the nervous system. Due to the wide range of neurological diseases with faulty DSBR, recent work has uncovered a critical role of HR and NHEJ in normal neurological function and development. When examining mouse embryos, there was unexpected spatiotemporal demarcation of NHEJ and HR during neurological development [271]. Specifically, Xrcc2-/- mice, eliminating the ability of mice to repair with HR, showed significant apoptosis in early proliferating neural stem cells in the ventricular

[28] As apposed to the more common transplant rejection by host, GVHD is the graft rejecting the host, leading to wide spread organ failure.

zone, whereas Lig4-/- mice, eliminating NHEJ, showed greater apoptosis in the later stage, postmitotic subventricular zone [271].[29] Although it is not entirely surprising that HR is more important than NHEJ in rapidly dividing cells, these findings hone in on how smaller deficits in DSBR could lead to subtle disruptions in neural development and manifest into neurological findings at birth and later in life.

More recently, several studies have begun to draw direct links between DSBR and physiological brain activity [6, 272]. While examining the effects of Alzheimer's-related human amyloid precursor protein (hAPP) on DSB formation in transgenic mice, researchers found an unexpected result [272]. By looking at phosphorylation of the histone protein H2A variant X at Ser139 (γH2A.X), which in focal accumulations connotes DSB formation, researchers expectedly found that hAPP mice had 2-3.5 times more γH2A.X positive cells than WT in several different brain regions which was further confirmed by finding a three-fold increase in γH2A.X after culturing primary neurons from WT mice with hAPP [272]. Serendipitously, while moving WT mice from their home cages to a new environment, they found a several fold increase in γH2A.X in many parts of the brain, most pronounced in the dentate gyrus, a brain region critical for learning and memory [272]. Fascinatingly, this effect went almost completely away after 24 hours in the environment [272]. To show that neural activity and not stress hormones from moving to a novel environment was sufficient for increases in γH2A.X, they exposed anesthetized mice to a visual stimulus and showed increases of γH2A.X in the primary visual cortex, V1 [272]. In addition, activation of the striatum by optogenetics showed greater levels of γH2A.X [272]. With these findings, the

[29]These mutations were embryonically lethal, however, could be rescued with an addition of p53-/-, further confirming p53's importance in cell cycle regulation and apoptosis—though these mice died from tumors quickly after birth.

group hypothesized that DSBs might serve as an adaptive function to aid in the significant chromatin remodeling and gene expression necessary for physiological neural activity [272]. Although seemingly farfetched, this hypothesis was further demonstrated with a more recent study examining the relationship between DSBs and the expression of early-response genes in neurons[6].

In order to understand the connection between DSBR and neural activity, groups out of MIT incubated cultured primary neurons with etoposide, a chemotherapy drug known to bind topoisomerase II (Topo II) and DNA and then induce double-stranded breaks [6]. By using next-generation RNA sequencing (RNA-seq), the groups noticed an enrichment of Fos, FosB, Npas4, and Egr1, genes associated with neural activity or early-response genes [6]. To demonstrate that this enrichment was not a product of the DNA-damage signaling pathway, the group repeated this experiment with an inhibitor to ATM (ATMi), leading to a reduction in γH2A.X but no effect on enrichment [6]. Interestingly, other DSB-inducing agents, such as bleomycin and neocarzinostatin, did not lead to enrichment, suggesting that enrichment is a Topo II-mediated process [6].

Knowing that neural activity can induce DSBs from previous work [272], the groups wanted to see if there were particular areas of the genome with a greater likelihood of breaks after neural activation. By incubating NMDA, a neural agonist, with primary neurons and using γH2A.X ChIP-seq, they found enrichment of 21 loci including Fos, FosB, Npas4, Egr1, Nr4a1, and Nr4a3. The groups then sought to understand the connection by performing Topo IIβ ChIP-qPCR under basal conditions, finding enrichment of Topo II at the Fos and Npas4 promoters. Since Topo II is known to relieve torsional stress on chromatin [273], they showed that this precipitated Topo IIβ could relax supercoiled plasmid DNA

in vitro [6]. They hypothesized at this point that Topo IIβ , relieves torsional stress to allow greater transcription. To ensure that transcription itself does not create torsional stress leading to breaks, they used an RNA Polymerase II inhibitor, 5,6-Dichloro-1-b-D-ribofuranosylbenzimidazole (DRB), prior to NMDA treatment and showed no change in the enrichment of DSBs [6]. They also pointed out that, although there was a greater amount of Topo IIβ binding after NMDA treatment, the DSBs only formed at loci where Topo IIβ had been previously bound during basal conditions [6].

To show that DSB formation at Fos and Npas4 promoters leads to greater expression of the genes that these transcription factors act on, the groups created primary neurons with Fos and Npas4-regulated luciferase and, as a control, Bdnf-regulated luciferase, and showed increases in fluorescence of the former and not the latter [6]. Using Topo IIβ ChIP-qPCR at different time points after NMDA treatment, they showed that activity-induced DSBs were repaired within two hours [6]. Finally, the group shows that NHEJ is critical for repairing these DSBs by preincubating primary neurons with a specific inhibitor of DNA-PK (NU7026) and showed a lack of repair two hours after NDMA treatment [6]. All together, these results demonstrate that there is a direct linkage between DSB and physiological neural activity, potentially explaining the neurological findings when DSBR is disrupted.

1.2.4.2 Ataxia Telangiectasia

One DSBR-related neurodegenerative disease is Ataxia Telangiectasia[30] (A-T), an autosomal recessive disorder resulting from the truncation or inactivation of the ATM protein—

[30]Hypomorphic mutations in MRE11A is A-T like disorder (ATLD) with milder symptoms and a later onset of symptoms [274].

typically via missense mutations [274]. The incidence of A-T is about 1 in 20,000, although about 1.2-2% of Caucasians in the U.S. are carriers [275]. The disease presents with the classical clinical triad of ataxia (loss of voluntary coordination), cerebellar degermation (atrophy) (Fig. 1.14A), and ocular telangiectasia (Fig. 1.13D) [276] with additional features including radiosensitivity, thymic degeneration, immunodeficiency, recurrent sinopulmonary infections, premature aging, gonadal dysgenesis and significant predisposition to lymphoma and leukemia [277].

Mechanistically, ATM[31] is a phosphoinositol-3 kinase-related kinase (PIKK) in the family of Ser/Thr protein kinases and part of the DNA-damage signaling pathway, arresting cells at the G1/S, intra-S, or G2/M checkpoints [278]. After the MRN complex binds at the site of a DSB, it recruits and activates ATM, converting it from its dimer and oligomer forms to a monomer [278]. ATM then phosphorylates a number of checkpoint proteins including p53, BRCA1, and CHK2, arresting the cell cycle. Although cells without ATM show a defect in the activation of the cellular response to DSBs, the DSBR response is attenuated not abolished, suggesting redundancy in protein kinases [274].

Clinically, children often show no signs of illness for the first year of life and begin walking at a normal age, however, the earliest symptom of ataxia presents soon after [275]. The neurodegeneration progresses slowly, leading to further problems with gait and balance and eventually deficits in speech [275]. The prognosis is poor with the median age of survival being 25 years and most succumbing to cancer or interstitial lung disease (Fig. 1.14B) secondary to repeated pulmonary infections [275]. Treatment options for A-T remain limited and supportive management is the consensus course of action [275].

[31] Yeast Tel1 homolog.

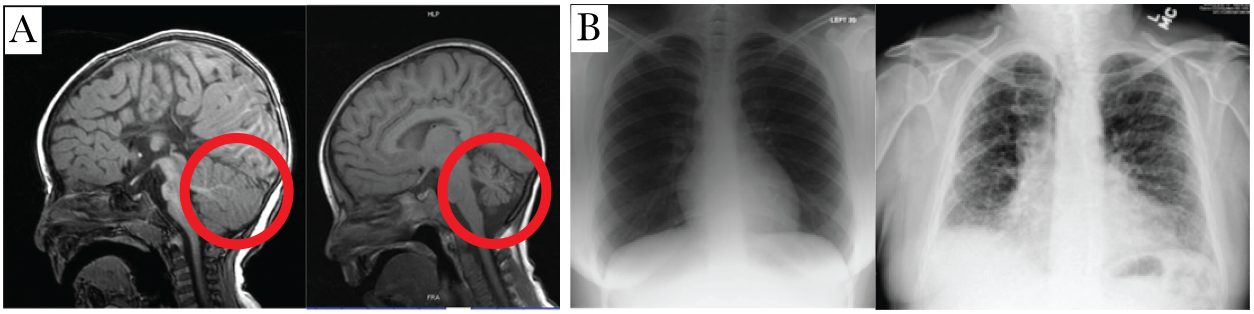

Figure 1.14: **Common features of Ataxia Telangiectasia (A-T)** (**A**) Cerebellar degermation (atrophy). Left image features sagittal MRI of normal cerebellum in a child (but not corpus collusum) circled in red. Right image features sagittal MRI of atrophied cerebellum in a child circled in red. (**B**) Interstitial lung disease by chest x-ray. Left, normal; right, clinical finding. A (left/normal) was adapted with permission from Siffredi, V., et al., Neuropsychological profile of agenesis of the corpus callosum: a systematic review. *Dev Neuropsychol* (2013). A (right/abnormal) was adapted with permission from Alakloby, O.M., et al., Ataxia telangiectasia with abnormal cellular immunity. *International Journal of Case Reports and Images* (2012). B (left/normal) was adapted with permission from Joarde, R. and N. Crundwell, The Normal Chest X-ray: An Approach to Interpretation, in Chest X-Ray in Clinical Practice, J. R. and C. N., Editors. *Springer* (2009). B (right/abnormal) was adapated with permission from Chapman, J.T. Interstitial Lung Disease. *Center for Continuing Education* (2010)

1.2.4.3 Seckel Syndrome

Seckel syndrome is an autosomal recessive disease characterized by short stature, microcephaly, and "bird-like" facial features (Fig. 1.13C) most frequently resulting from mutations in the ATR (ATM and Rad3-related) protein [279, 280]. Seckel syndrome was once an over-diagnosed syndrome and, in fact, only four families have been reported to have this ATR-related disease [280]. Like ATM, ATR is also a phosphoinositol-3 kinase-related kinase (PIKK) that is part of the DNA-damage signaling pathway [279]. In complex with ATR interacting protein (ATRIP), ATR/ATRIP phosphorylates H2AX, 53BP1, p53, NBS1 and CHK1, arresting cells at the intra-S and G2/M checkpoints [279]. However unlike ATM, ATR is activated by RPA-bound ssDNA resulting from stalled replication forks and is essential for embryonic development and somatic cell growth [279]. Since it is essential, reported mutations result in hypomorphic ATR that reduce its activity and do not completely abolish

it [279]. Due to limited cases of this disease, prognostic information is limited but anecdotal reports suggest these individuals can live past age 50, assuming no malignancies. No treatments are available and symptom management is the goal of care [280].

1.2.4.4 Nijmengen Breakage Syndrome

Alternatively known as Ataxia Telangiectasia Variant 1, Nijmegen Breakage Syndrome (NBS) is a recessive disorder presenting with microcephaly at birth and later with progressive microcephaly (Fig. 1.12A), dysmorphic "bird-like" facial features (Fig. 1.13C), growth deficits, intellectual disability, and, in females, hypogonadism [281, 282]. NBS is relatively rare and isolated to Eastern European countries with an estimated incidence of 1 in 95,000 in the former Czechoslovakia and 1 in 154-190 carrying the mutation in three Slavic countries [283]. The mutant protein responsible for NBS, NBS1[32], was first discovered by haplotype mapping of genomic DNA extracted from the blood samples of 51 patients with related symptoms [284, 285].

In the cell, NBS1[33] is apart of the MRN complex and thus essential for proper DNA repair, although its exact function is not fully understood as it does not have DNA-binding nor kinase activity [282, 286]. NBS1 may also have critical functions outside of DSBR, including interstrand crosslink repair[34], ICL, potentially explaining the wider range of malignancies presented in heterozygous carriers of mutant NBS1 [287]. Although NBS does not cause neurodegeneration, progressive microcephaly results in mild to moderate intellectual

[32]Yeast Xrs2 homolog.

[33]Rad50 deficiency causes an NBS-like disorder [285].

[34]NBS1 has been shown to bind RAD18 and initiate Polη-dependent translesion synthesis (TLS) by PCNA ubiquitination [287].

deficits by age 7 to 10 with particularly reduced growth in the frontal lobes and corpus callosum [283]. Interestingly, the cerebellum appears relatively intact by MRI, potentially explaining the lack of ataxia or abnormal gait, as seen in A-T [283]. NBS also presents with café-au-lait spots (Fig. 1.12) and vitiligo spots in about 50-70% of patients [282]. As with the other DSBR syndromes, the prognosis is poor with most patients dying of cancer[35] and infections and the median age of survival being seven years [283]. There is no cure for NBS and treatment is limited to symptom management and infection prophylaxis [283].

1.2.5 Conclusion

The purpose of MD/PhD programs is to bridge the gap between scientists and medical professionals. Often, the mechanistic findings of molecular genetics, biochemistry, and biophysics are cited in aiding our understanding of disease. However, less often cited is the reverse relationship—the manifestations of disease can tell us something about the pathways that might otherwise be overlooked. For DSBR, the common neurological findings in many patients with mutations in the pathway suggest that proper DSB repair is critical for physiological neural function. Recent studies have begun exploring this connection although significant questions still remain. Thus, physician scientists have the unique privilege to not only use their scientific pursuits to aid in the curing of disease but may use the clinical manifestations as a way of understanding the mechanisms behind disease.

[35] 40% develop cancer before 20 years of age with most being lymphomas although there are cases of glioma, rhabdomyosarcoma, and medulloblastoma [257].

Chapter 2

Single-Stranded DNA Curtains for Studying Homologous Recombination

This chapter is adapted from work originally published as: "Single-Stranded DNA Curtains for Studying Homologous Recombination," Chu J. Ma*, **Justin B. Steinfeld***, and Eric C. Greene. Methods in Enzymology (2017). *C.J.M. and I contributed equally to researching and writing this methods paper.

2.1 Abstract

Homologous recombination is an important pathway involved in the repair of double-stranded DNA breaks. Genetic studies form the foundation of our knowledge on homologous recombination. Significant progress has also been made toward understanding the biochemical and biophysical properties of the proteins, complexes, and reaction intermediates involved in this essential DNA repair pathway. However, heterogeneous or transient recombination intermediates remain extremely difficult to assess through traditional ensemble methods, leaving an incomplete mechanistic picture of many steps that take place during homologous recombination. To help overcome some of these limitations, we have established DNA curtain methodologies as an experimental platform for studying homologous DNA re-

combination in real-time at the single- molecule level. Here, we present a detailed overview describing the preparation and use of single-stranded DNA curtains in applications related to the study of homologous DNA recombination with emphasis on recent work related to the study of the eukaryotic recombinase Rad51.

2.2 Introduction

2.2.1 Homologous Recombination

Homologous recombination (HR) is a highly conserved pathway that enables the exchange of genetic information between identical or closely related DNA molecules, and is an important driving force in genome evolution. HR plays crucial roles in the repair of double-stranded DNA breaks (DSBs), the rescue of stalled or collapsed replication forks, chromosomal rearrangements, horizontal gene transfer, and meiosis in sexually reproducing organisms [28, 288–290].

Much of our knowledge of HR comes from the study of DSB repair in the budding yeast Saccharomyces cerevisiae [291–293]. Here, we briefly highlight some of the key steps and proteins involved during the early stages of DSB repair in S. cerevisiae (Fig. 2.1); for a more in-depth discussion of HR, we refer the reader to several excellent reviews [28, 289, 293]. Upon formation of a DSB the newly liberated DNA ends are processed to yield long 30 single-stranded DNA (ssDNA) overhangs. Replication protein A (RPA) binds to these overhang to remove any potential secondary structure and also protect the ssDNA from degradation by nucleases. RPA is then replaced, with the help of mediators such as Rad52, by the

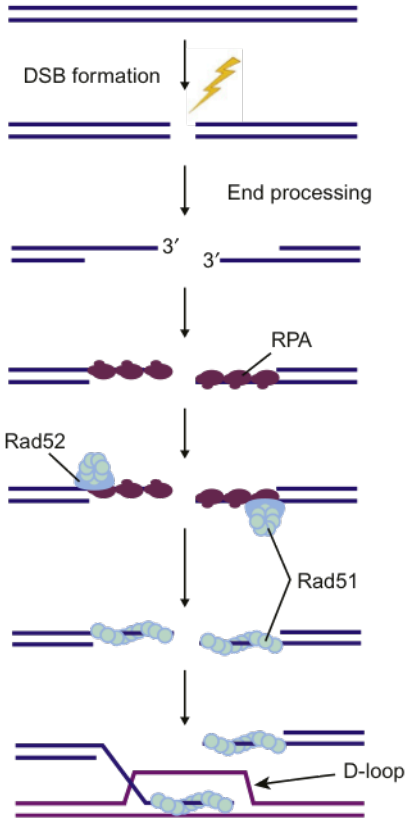

Figure 2.1: **Early stages of eukaryotic homologous recombination.** DSBs are resected to yield long 3' ssDNA overhangs that are first bound by RPA. Rad52 then binds and assists loading of the Rad51, which forms long filaments on the ssDNA and these Rad51–ssDNA filaments are referred to as the presynaptic complex. The presynaptic complex then searches for a homologous DNA and pairs the processed ssDNA overhang with its homologous partner to generate a D-loop intermediate.

recombinase Rad51, an ATP-dependent DNA-binding protein that forms an extended right-handed helical filament on the ssDNA overhang [294, 295]. This nucleoprotein filament, referred to as the presynaptic complex, is responsible for aligning and pairing the ssDNA overhang with a homologous double-stranded DNA (dsDNA) sequence present elsewhere in the genome. The 3' end of the presynaptic ssDNA can prime replication using homologous dsDNA as a template and the resulting intermediates can be channeled through a number of distinct pathways that will restore the originally broken DNA molecule [28, 289, 293].

The highly simplified view of HR presented earlier belies the fact that recombination requires the coordinated action of a complex repertoire of proteins. Highly organized macromolecular assemblies are responsible for sensing DNA damage, recruiting essential factors to the damaged sites, and repairing the damaged DNA. Many of these proteins belong to the

RAD52 epistasis group of genes, which were initially identified in S. cerevisiae as mutants that exhibited extreme sensitivity to ionizing radiation [51], and many other HR proteins have been identified in subsequent years. In total, at least 45 different proteins are known to be directly involved DSB repair in S. cerevisiae [293].

Genetic experiments have provided the basis for most of the current body of knowledge in HR and continue to yield important new insights [292, 293]. Cell biology studies have also yielded an enormous amount of information regarding the protein and DNA components involved in eukaryotic HR [8, 296, 297]. Biochemical and biophysical studies have been employed to study several aspects of the reaction, and these experiments have revealed some crucial insights into HR mechanisms [294, 295]. More recently, structural approaches have unveiled atomic-level details of some proteins and protein–DNA complexes that are essential to HR [36, 37]. However, many questions related to the HR mechanisms cannot be addressed through any of these types of approaches because the underlying intermediates are either transient or heterogeneous. Single molecule fluorescence-based methods offer the potential for direct visual analysis of individual HR reaction components or complexes, which in turn can enable the direct detection of subpopulations within an otherwise heterogeneous mixture and can also capture rare or transient intermediates along a reaction trajectory.

2.2.2 Single-Molecule Biology

Single-molecule approaches are technically demanding, and it is relevant to ask whether it is worth the trouble to study complex biological systems at the scale of individual components. As an analogy, take the study of salmon swimming upstream to spawn. One could

tag 1000 salmon at a river inlet, and then wait upstream to determine, for instance, how many fish reached their spawning grounds, how long it took them to get there, and how many fish actually spawned. Such a study would provide valuable information about the salmon population and basic insight into their life cycle. However, the study as presented could not answer why some salmon survived and some did not, what factors dictate precise arrival time, or why some salmon successfully spawn whereas others do not. To address these types of questions, the study must be redesigned to follow individual fish within the population as they swim toward the spawning grounds. Through this type of study, one might find that some salmon were harvested by anglers or bears, others may be unable to navigate fish ladders, and still others may have experienced too much stress during their travels to spawn successfully. Suddenly, the picture of salmon reproduction becomes much more detailed. Similarly, ensemble biochemical or genetic studies typically can look only at some well-defined intermediate state or the final output from a process, and may overlook heterogeneous intermediates or transient states that are important for understanding the overall nature a particular reaction. Experimental approaches capable of interrogating individual macromolecules or complexes over the course of a biochemical reaction trajectory now offer new possibilities for understanding many types of biological problems in greater depth than previously possibly.

2.2.3 Overview of DNA curtains

Two challenges users of single-molecule methods face are the difficulties associated with collecting statistically relevant information and the problem of nonspecific surface absorp-

tion, which arises because most single-molecule bases methods require that the biomolecules under investigation be anchored to a solid supporting surface without compromising biological activity. To help overcome these problems we have developed "DNA curtains," in which ds- or ssDNA molecules can be organized into defined patterns on the surface of a microfluidic sample chamber (Fig. 2.2) [298–301]. In brief, DNA curtains are prepared by first depositing metal barriers and anchors on the surface of a fused silica microscope slide by electron beam (e-beam) lithography. The slide is then coated with a fluid lipid bilayer, which prevents nonspecific surface adsorption and provides a mobile platform for anchoring DNA molecules through a biotin–streptavidin linkage. Buffer flow is then used to push the DNA molecules into the barriers where they all align with one another [298, 301]. If desired, the second end of the DNA can be attached to a downstream anchor point [299, 300]. This approach allows for the direct observation of hundreds of individual DNA molecules within the typical field of view of an optical microscope, providing a flexible experimental platform that can be used to study different types of protein–DNA interactions [299, 300, 302–310]. In subsequent sections, we describe how to prepare ssDNA curtains [81] and provide brief examples of how this technique has been applied to the study of HR with emphasis on recent experiments using the eukaryotic DNA recombinase Rad51 [305, 306].

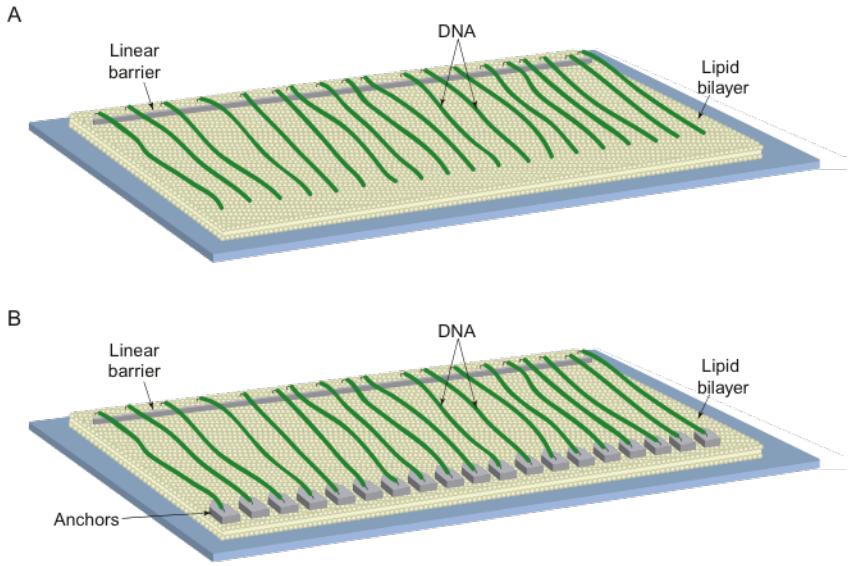

Figure 2.2: **Different types of DNA curtains.** (**A**) Schematic illustration of a single-tethered DNA curtain made with a linear barrier. (**B**) Double-tethered DNA curtain where the downstream ends of the DNA are tethered to the exposed anchor points that project above the bilayer. Both formats are compatible with either dsDNA or ssDNA. Adapted with permission from Silverstein, T. D., Gibb, B., & Greene, E. C. (2014). Visualizing protein movement on DNA at the single-molecule level using DNA curtains. DNA Repair (Amst), 20, 94–109.

2.3 Methods

2.3.1 Total Internal Reflection and Instrumentation

We use total internal reflection fluorescence microscopy (TIRFM) for visualizing DNA curtains. TIRFM uses spatially selective laser excitation to reduce background signal by several orders of magnitude relative to conventional wide-field illumination techniques [311]. Detailed descriptions of TIRF microscopes are widely available, and below provide a component list describing our most recent instruments, which uses inverted Nikon Eclipse microscopes equipped with custom laser illumination systems for dual-color prism-type TIRFM illumination (Fig. 2.3 and Table 2.1).

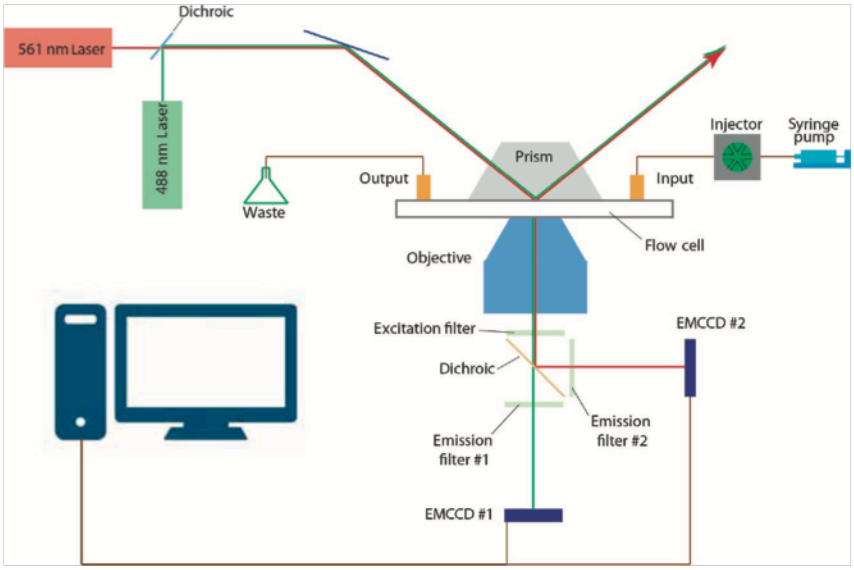

Figure 2.3: **Total internal reflection fluorescence microscopy.** This schematic highlights important details of the TIRFM systems used to visualize DNA curtains. Details of the schematic and system components are presented in the main text.

Table 2.1: TIRF Microscopy Instrumentation.

Microscope Body	Nikon Eclipse Ti-E with Perfect Focus System (PFS)
Camera	Two Andor iXon X3 EMCCDs (Model: DU-897E-C50-#BV)
Objective	Nikon CFI PLAN APO 60X WI
Filter cube	Chroma TE2000/Ti filter cube (Part No. 91020) Chroma ET525/50 m band pass filter Chroma ET575lp long pass filter Chroma ZT561rdc dichroic mirror
Additional Filter	Chroma ZT488rde-UF1 dichroic
Illumination	Coherent Sapphire LP (488 nm, 200 mW) Coherent Sapphire LP (561 nm, 200 mW)
TIRF prism	Thor Labs, custom-made, uncoated, fused silica
Shutter	Vincent Uniblitz VCM-D1 Single Channel Uni-Stable Shutter
Objective heater	Bioptech Objective Heater (150819-19)/Controller (150803)
Slide heater	Custom Al casing, Omega Mini Benchtop Controller (CSC32J)

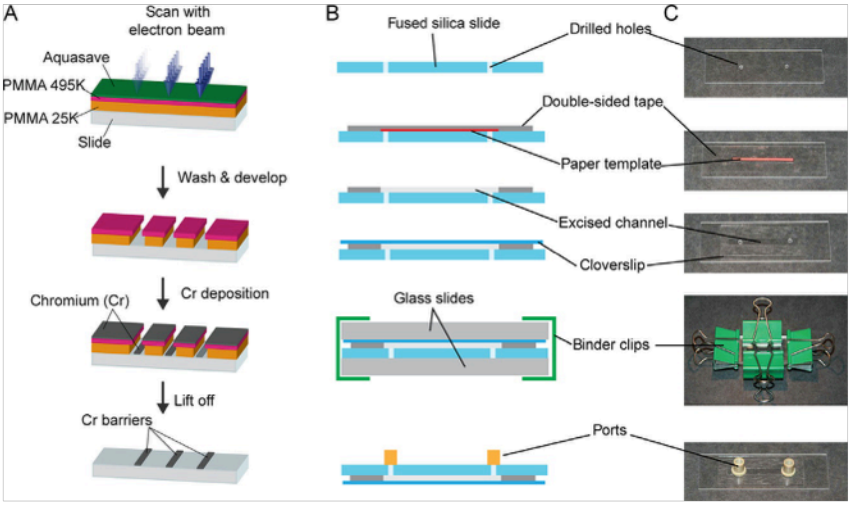

Figure 2.4: **Flow cell fabrication.** (**A**) Slides are first coated with two layers of PMMA, a layer of AquaSAVE, and an electron beam is then used to etch through these layers. Chromium is deposited on the surface and the remaining PMMA is removed, leaving behind the nanofabricated barriers. Schematic illustrations (**B**) and photographs (**C**) depicting the different stages of flow cell assembly. Adapted with permission from Greene, E. C., Wind, S., Fazio, T., Gorman, J., & Visnapuu, M. L. (2010). DNA curtains for high-throughput single-molecule optical imaging. Methods in Enzymology, 472, 293–315.

2.3.2 Flow Cell Fabrication

Our DNA curtain experiments are performed using flow cells that are machined and assembled in-house (Fig. **??**). In brief, each flow cell is made from a fused silica microscope slide. Inlet and outlet holes are bored into each slide using a diamond-coated drill bit. Metallic patterns are deposited on the surface of the slide by e-beam lithography, and the patterns can be designed for either single- or double-tethered formats. These modified slides are then cleaned and assembled into the flow cells that are used for making DNA curtains. Later, we summarize each of these procedures, and detailed explanations of each step have also been previously published [312].

2.3.2.1 Slide Preparation

First, two holes are cut through each fused silica slide to allow sample delivery in the completed flow cell (Fig. 2.4). The holes are drilled using a 1.4-mm diamond-coated drill bit (Shor International, Cat. No. DIB-211.00). While drilling, the slides are submerged in a water bath to cool the bit and prevent inhalation of fused silica dust [312]. The drilled slides are then cleaned by submersion in piranha solution (3:1 mixture of concentrated sulfuric acid (97%) and 30% hydrogen peroxide). For cleaning, up to 10 slides are placed in a glass slide rack and the rack is placed within rectangular glass staining dish (e.g., Electron Microscopy Sciences, Cat. No. 70312-20). The dish is then filled with 150 mL of sulfuric acid followed by the addition of 50 mL cold hydrogen peroxide. The slides are incubated in the piranha solution for 30 min. Extreme care must be taken when handling piranha solution to avoid contact with exposed skin. Following the 30-min submersion in piranha solution, the slides should be rinsed with copious amounts of deionized water.

The cleaned and drilled slides must now be prepped for lithography by deposition of a positive photoresist and conductive polymer layer (Fig. 2.4A). For this, the slides are spin-coated with a layer of 3% (w/v) polymethylmethacrylate (PMMA; molecular weight 25 kDa), dissolved in anisole (MicroChem, Newton, MA), followed by a layer of 1.5% (w/v) PMMA (495 kDa), also dissolved in anisole. The PMMA layers are then topped off with a final layer of AquaSAVE conducting polymer (Mitsubishi Rayon, Tokyo, Japan). Each of these layers is spun at 4000 rpm for 45 s using a ramp rate of 300 rpm/s using a Laurell Technologies Corp spincoater (WS-650MZ-23NPP). The slides are now ready for e-beam lithography. We use an FEI scanning transmission electron microscope equipped with nano pattern generator

system software (Nabity, Inc.), which controls the e-beam as it writes patterns on coated fused silica slides. Dosage tests can be used to determine the current and writing time that yield the best combination of time and resolution for the desired patterns. The slides are then developed to remove the PMMA from the areas exposed to the e-beam. This step is performed by placing the slide within a 50-mL falcon tube containing a developing solution comprised of a 3:1 mixture of methyl isobutyl ketone:isopropanol cooled to –20°C. The tube is sonicated in an icy water bath sonicator (Branson 1800, power LOW, 60 s processing time), and rinsed off with isopropyl alcohol (IPA). A thin layer of chrome (∼20-nm) is then deposited onto the patterned surfaces using a Semicore E-beam Evaporation System (SC 2000LT). The remaining PMMA is then removed by first washing the slides with acetone from a squirt bottle. The slides are then submerged in acetone and sonicated for 5–10 min. Finally, the patterned slides are rinsed with clean acetone from a wash bottle. To prevent formation of deposits from the acetone drying, acetone is cleaned away with IPA and dried with a stream of nitrogen gas.

2.3.2.2 Flow Cell Assembly and Disassembly

Prior to use, the patterned slides must be assembled into flow cells that can be connected to a sample deliver system (Fig. 2.4B,C). The flow cells are made by using double-sided tape to create a sample chamber between the patterned slide and a glass coverslip. Later we provide a step-by-step description of the flow cell assembly procedure, as well as information describing how the flow cells can be dismantled and reused.

1. Center a rectangular paper template (35 x 5 mm) over a piece of double- sided tape

(19 mm width) and tape over the chrome pattern on the slide.

2. Use the paper template as a guide to excise a channel in the double-sided tape.
3. Place a coverslip (Fisher Scientific, Fisher Finest Premium cover glass, Cat. No. 12-548-5E) on top of the tape and apply pressure to seal the coverslip to the tape.
4. The assembled flow cell is sandwiched between glass slides (Fisher Scientific, Frosted microscope slides 12-550-343) and held by binder clips on all four sides to distribute pressure evenly.
5. Bake the assembly under vacuum for 45 min at 140°C to seal the tape.
6. Remove from the oven, release the binder clips and the glass slides.
7. Glue Nanoports (Idex, Cat.No.N-333) over the drilled port holes with a hot glue gun. The flow cell is now complete and can be stored under vacuum at room temperature until use.
8. Patterned slides can be reused after each experiment. Submerge the slide in ethanol for 48 h, and then remove the ports, coverslip, and double-sided tape. The slides are then cleaned by submersion in the following solutions with constant stirring: 2% Hellmanex solution for 48 h; rinse with Milli-Q water; 1 M NaOH for 40 min, rinse with Milli-Q water; 100% ethanol for 30 min. The cleaned slides are then ready for reuse.

2.3.3 ssDNA Curtains

The assembled flow cells can now be used for the preparation of the ssDNA curtains. First, a lipid bilayer is used to passivate the flow cell surface. The ssDNA substrate is then

attached to the lipid bilayer through a biotin– streptavidin–biotin linkage, and buffer flow is used to push the anchored ssDNA molecules into position along the chrome barriers. Finally, the ssDNA is labeled and extended by injecting fluorescently tagged RPA into the sample chamber. The following sections provide step-by-step details necessary to complete each of these procedures.

2.3.3.1 Liposome Preparation

1. Lipid stocks are prepared by dissolving the following components in 10 mL of chloroform: 1 g DOPC (1,2-dioleoyl-sn- glycero-3-phosphocholine), 100mg PEG-2000 DOPE (18:1 PEG- 2000: 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine-N-[methoxy (polyethylene glycol)-2000] (ammonium salt)), 5 mg biotinylated DOPE. The dissolved lipid mixtures can then be stored at –20°C. All lipids are purchased from Avanti Polar Lipids, Inc.
2. The lipid stock solutions are used to make liposomes. Liposomes are typically prepared in 2 mL batches, using the following step-by-step procedure, and can stored at 4°C for 4–8 weeks.
    a) Clean an organic-solvent compatible syringe with chloroform and transfer 200 μL (1/10th of the final desired volume) of the lipid stock to a new 2-mL glass vial (National Scientific, Cat. No. C4015).
    b) Using a very low pressure stream of nitrogen gas, evaporate the chloroform from the lipid stock slowly over several minutes. During this time, the lipid stock will form a solid residue on the side of the vial. After all the chloroform is evaporated, slightly increase the pressure and continue blowing nitrogen until all

traces of liquid are removed.

c) Place the uncapped glass vial under vacuum overnight.

d) Add 2 mL of lipid buffer (10 mM Tris–HCl [pH 8.0], 100 mM NaCl) to the dried lipid stock and cap the vial. Incubate at room temperature for an hour and then vortex until all the lipid stock has dissolved into solution.

e) Transfer the mixture to a 5-mL polypropylene culture tube(Falcon, Cat. No. 35-2058) and sonicate in an ice bath using a microtip sonicator (Misonix S-4000) until the solution becomes clear.

f) The solution is then filtered through a $0.22-\mu$m nylon syringe filer (Fisherbrand, Cat. No. 09-720-3) and the resulting liposomes are ready for use.

2.3.3.2 Preparation of ssDNA

Our ssDNA curtain experiments make use of relative long ssDNA substrates (40,000 nucleotides) that are made by rolling circle replication with a bio- tinylated oligonucleotide primer and a circular ssDNA template [81], as described later.

1. The biotinylated primer is first annealed to a circular M13 DNA tem- plate in a $100-\mu$L reaction containing: 40 mM Tris–HCl [pH 8.0], 50 mM NaCl, 10 mM MgCl2, 10 g (89.4 nM) of M13mp18 (New England Biolabs, Cat. No. N4040S), and 45nM primer (5'-BIO- TEG-TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT GTA AAA CGA CGG CCA GT). The sample is placed in near boiling water (95°C) in a 1-L beaker for 5 min, and the beaker is then transferred to the benchtop and allowed to cool slowly to room temperature. The annealing reactions are then in buffer containing

10 mM Tris–HCl [pH 8.0], 50 mM NaCl, and 5 mM MgCl2 to a total volume of 300 μL and stored at –20°C until use.

2. A fresh preparation of ssDNA is made for each ssDNA curtain experi- ment. Rolling circle reactions (50 μL) are prepared containing: 10 μL of 5 reaction buffer (250 mM Tris–HCl [pH 7.5]), 20 mM DTT, 50 mM ammonium sulfate, and 50 mM MgCl2), 1 μL annealed M13 template (see earlier), 1 μL of 10 mM dNTP mix, 1 μL purified ϕ29 polymerase (10 μM stock) [81], and 37 μL water. Mix by pipetting, do not vortex. Incubate at 30°C for 25 min, and use immediately after preparation.

2.3.3.3 Lipid Bilayer Deposition and ssDNA Attachment

The following section describes how the bilayer is deposited onto the flow cell surface and how the ssDNA substrate is attached to the bilayer. These steps are all performed manually at the benchtop using hand held syringes. Extreme care should be taken to prevent any air from entering the flow cell once the bilayer has been deposited, and all tubing and syringe attachments should be made using drop-to-drop connections. If air bubbles pass through the sample chambers, they will destroy the lipid bilayer.

1. Fill two 3 mL syringes with Milli-Q water. Connect one of the syringes to the inlet port and push 1 mL of water through the flow cell. Connect the second syringe to the outlet port and push–pull the water between the inlet and outlet syringes to remove any air from the flow cell. Very small air bubbles may appear near the edge of the tape, but these are generally not problematic.
2. Mix 40 μL of liposome solution with 1 mL of lipid buffer (10 mM Tris– HCl [pH 8.0],

100 mM NaCl). Push 250 μL of the mixture into the flow cell approximately every 5 min until all of the liposome mixture is used. After the final injection, allow the lipids to form a cohesive bilayer by incubating at room temperature for 30 min.

3. Further passivate the surface with 1 mL of BSA buffer (40 mM Tris–HCl [pH 8.0], 2 mM MgCl2, 1 mM DTT, and 0.2 mg/mL BSA). Incubate 5 min. The rolling circle replication reaction can be started at this time (see Section 2.4.2).
4. Mix 10 μL of streptavidin (1 mg/mL, Invitrogen, Cat. No. S888) with 790 μL of BSA buffer and push the solution through the sample chamber in two 500 μL steps with a 5 min incubation between steps.
5. Rinse the sample chamber with 3 mL of BSA buffer to remove free streptavidin.
6. Dilute the freshly prepared rolling circle reaction with 450 μL of BSA buffer and slowly push the ssDNA solution through the sample chamber over a 10 min period. Mount the flow cell on the microscope stage and adjust the focus as necessary.
7. After mounting the flow cell on the microscope stage, the input and the output ports are connected to a sample injection system comprised of a syringe pump (KD Scientific, KDS-201) and a high-pressure switch valve (IDEX Health & Science, MXP9900-000). Again, it is essential that all connections be made using drop-to-drop connections to avoid inadvertently injecting air bubbles through the sample chamber.

2.3.3.4 Using RPA–eGFP to Visualize ssDNA

DNA curtains made with dsDNA can be visualized using an intercalating dye such as YOYO1. However, ssDNA is not readily labeled with intercalating dyes and the reactive oxygen species generated when these dyes are illuminated by laser light can rapidly nick the

ssDNA, which is problematic because a single nick will result in release of the ssDNA from the flow cell surface. In addition, ssDNA forms extensive secondary structure, which must be removed in order to visualize extended molecules along their full contour lengths. To overcome these problems, we use a GFP-tagged version of the eukaryotic ssDNA-binding protein RPA both to label the ssDNA and to remove secondary structure so that the molecules can be easily extended by buffer flow (Fig. 2.5)[81]. RPA offers the additional benefit that RPA-coated ssDNA is the physiological substrate for assembly of the eukaryotic presynaptic complex[73, 294]. The following steps outline a typical procedure for labeling and extending the ssDNA curtains with RPA–eGFP.

1. Begin by diluting an appropriate amount of S. cerevisiae RPA–eGFP into 20 mL of BSA buffer. RPA has a very high affinity for ssDNA [73], so working concentrations of just 0.1 nM RPA–eGFP are sufficient to label and extend the ssDNA.
2. Flush the RPA-containing buffer through the sample chamber at a rate of 1.0 mL/min for approximately 15 min. The ssDNA will immediately be visible by TIRFM upon injection of RPA–eGFP. The RPA–eGFP– ssDNA complexes initially appear as short molecules that slowly extend with time as secondary structure is removed (Fig. 2.5).
3. After 2 min of RPA buffer flow, a 500 μL pulse of 7 M urea is flushed through the sample chamber at 1 mL/min to help remove any residual ssDNA secondary structure, ϕ29 DNA polymerase, or M13 circular ssDNA template.
4. For double-tethering, a row of pentagon-shaped pedestals is lithographed downstream of the barriers (Fig.2.2). These pedestals serve as anchor points for the nonspecific adsorption of the RPA–ssDNA, which stick to the exposed chromium surfaces. Nonspe-

cific interactions between the downstream end of the RPA–ssDNA and the chromium pedestals allow the ssDNA to remain extended and visualized by TIRFM even in the absence of buffer flow.

2.3.3.5 Presynaptic Complex Assembly

The RPA–eGFP-coated ssDNA can serve as the starting point for assembly of the presynaptic complex, and we have used this as a substrate for the assembly of presynaptic complexes made from a variety of Rad51/RecA recombinases, including Escherichia coli RecA, S. cerevisiae Rad51, human Rad51, as well as the meiosis-specific recombinases S. cerevisiae Dmc1 and human Dmc1 [304, 306]. We use unlabeled Rad51/RecA recombinases for our experiments so successful assembly of the presynaptic complex is revealed by the displacement of the fluorescent RPA–eGFP (Fig. 2.6). Here, we briefly describe the procedure used for making an S. cerevisiae Rad51 presynaptic complex, but this protocol can be readily adapted for other Rad51/RecA recombinases.

1. Wash the RPA–eGFP-bound ssDNA curtains with HR buffer (30 mM Tris-Acetate [pH 7.5], 20 mM Mg-Acetate, 50 mM KCl, 1 mM DTT, 0.2 mg/mL BSA) plus 2.5 mM ATP for 2 min at a flow rate of 1 mL/ min to remove any free RPA and equilibrate the sample chamber in HR buffer.
2. Inject a $50-\mu$L sample of S. cerevisiae Rad51 (2 M) in HR buffer plus 2.5 mM ATP. Terminate buffer flow once Rad51 enters the sample chamber and incubate the sample in the absence of buffer flow for 15 min at 30°C.
3. Confirm assembly of the presynaptic complex by visual inspection of the ssDNA before,

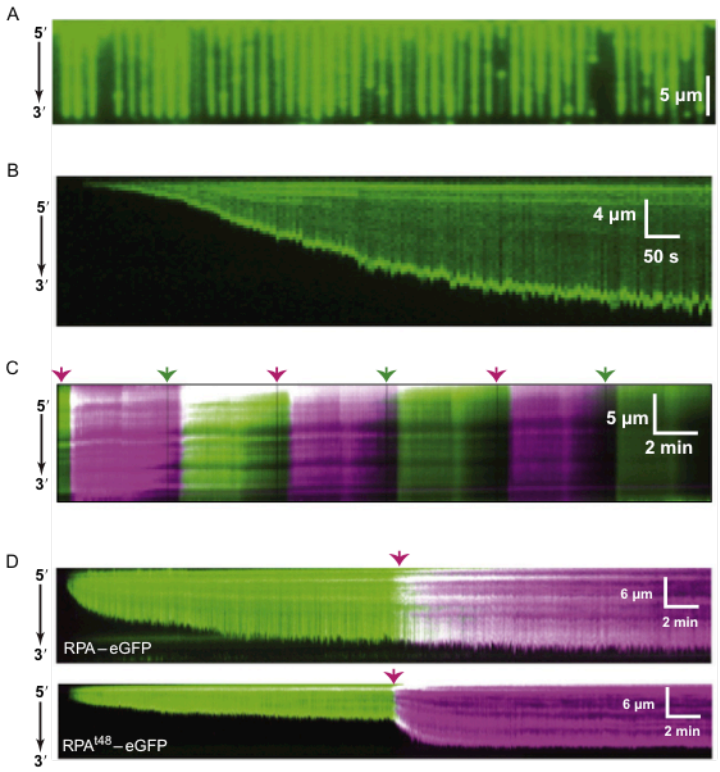

Figure 2.5: **RPA-coated ssDNA curtains.** (**A**) Wide-field TIRFM image of a double-tethered ssDNA curtain bound by RPA–eGFP. (**B**) Kymograph showing what takes place when single-tethered ssDNA molecules are labeled with RPA–eGFP. The ssDNA slowly becomes longer as RPA–eGFP binds and disrupts existing secondary structure. (**C**) Kymograph showing facilitated dissociation of RPA from the ssDNA when free RPA is injected into the sample chamber. RPA–eGFP is shown in green, and RPA– mCherry is shown in magenta, and the color-coded arrowheads indicate successive injections of each protein. (**D**) Kymographs of a single-tethered ssDNA showing that the exchange of wild-type RPA–eGFP with wild-type RPA–mCherry does not alter ssDNA length (upper panel), whereas exchange of the RPAt48 mutant, which is defective for ssDNA binding, with wild-type RPA–mCherry coincides with an increase in ssDNA length. Adapted with permission from Deng, S. K., Gibb, B., de Almeida, M. J., Greene, E. C., & Symington, L. S. (2014). RPA antagonizes 405–412; Gibb, B., Silverstein, T. D., Finkelstein, I. J., & Greene, E. C. (2012). Single- stranded DNA curtains for real-time single-molecule visualization of protein-nucleic acid interactions, Analytical Chemistry, 84, 7607–7612; Gibb, B., Ye, L. F., Gergoudis, S. C., Kwon, Y., Niu, H., Sung, P., et al. (2014). Concentration-dependent exchange of replication protein A on single-stranded DNA revealed by single-molecule imaging. PloS One, 9, e87922; Gibb, B., Ye, L. F., Kwon, Y., Niu, H., Sung, P., & Greene, E. C. (2014). Protein dynamics during presynaptic-complex assembly on individual single-stranded DNA molecules. Nature Structural and Molecular Biology; and Qi, Z., Redding, S., Lee, J. Y., Gibb, B., Kwon, Y., Niu, H., et al. (2015). DNA sequence alignment by microhomology sampling during homologous recombination. Cell, 160, 856–869.

during, and after the Rad51 injection. Successful assembly of the presynaptic complex results in dissociation of RPA–eGFP from the ssDNA (Fig. 2.6). Once RPA–eGFP has been displaced from the ssDNA flush the sample chamber with additional HR buffer to remove any unbound Rad51.

4. The resulting Rad51 presynaptic complexes remain stable for at least 2 h if ATP is maintained in the buffer, and the stability of the complexes can be assessed using RPA–eGFP [306]. RPA–eGFP will only bind to the ssDNA after dissociation of Rad51, so the integrity of the presynaptic complex can be readily confirmed by injecting HR buffer containing 0.1 nM RPA–eGFP (Fig. 2.6).

2.4 Applications

Here, we briefly summarize some of the studies we have conducted using ssDNA curtains, which include analysis of RPA-binding dynamics, Rad51 presynaptic complex assembly, the association of protein cofactors with the Rad51 presynaptic complex, and interactions between the presynaptic complex and fluorescently tagged dsDNA fragments.

2.4.1 RPA-Binding Dynamics

RPA is a heterotrimeric complex composed of Rfa1, Rfa2, and Rfa3 sub- units, and participates in all aspects of nucleic acid metabolism involving an ssDNA intermediate [73, 313]. RPA–ssDNA complexes are stable for over 2 h, without any detectable dissociation when free RPA is removed from solution [38, 70]. However, the bound RPA can undergo rapid exchange when free RPA is present in solution, which can be visualized as a change in fluo-

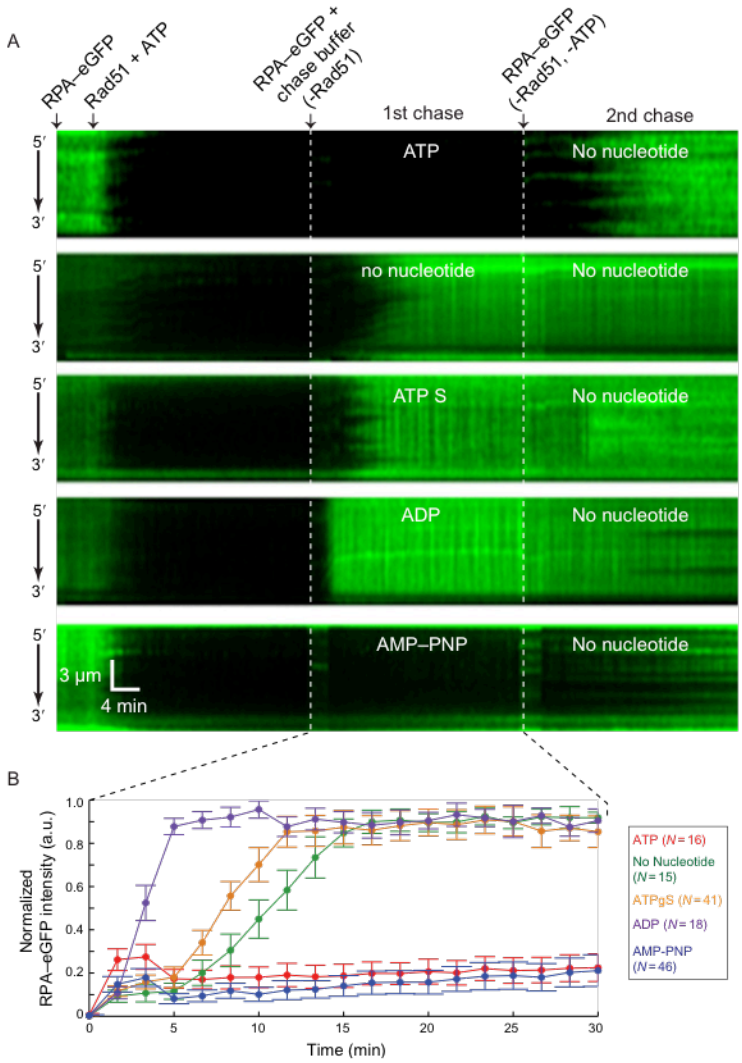

Figure 2.6: **Kymographs showing presynaptic complex assembly reactions.** (**A**) RPA–eGFP– ssDNA curtains were first incubated with 1 μM S. cerevisiae wild-type (unlabeled) Rad51 and 2.5 mM ATP. Binding of Rad51 to the ssDNA is revealed as a rapid loss of RPA–eGFP fluorescence signal. The sample chambers were then flushed (1st chase) with buffer containing 1.0 nM RPA–eGFP and either no nucleotide or 2.5 mM of the indicated nucleotide cofactor, followed by a 30-min incubation. Disassembly of the Rad51–ssDNA presynaptic filaments is revealed by the binding of RPA–eGFP to the exposed ssDNA. The sample chambers were then flushed (2nd chase) with additional buffer containing 0.1 nM RPA–eGFP and no nucleotide cofactor, and incubated for an additional 30-min. (**B**) Quantitation of the Rad51 filament stability in the presence of various nucleotide cofactors, as indicated. Error bars represent s.d. Adapted with per- mission from Qi, Z., Redding, S., Lee, J. Y., Gibb, B., Kwon, Y., Niu, H., et al. (2015). DNA sequence alignment by microhomology sampling during homologous recombination. Cell, 160, 856–869.

rescence color of the ssDNA when switching between RPA–eGFP and RPA–mCherry (Fig. 2.5C,D) [38, 70, 314]. This unusual behavior suggests that ssDNA-bound RPA undergoes constant microscopic dissociation under all conditions, but these microscopic dissociation events only result in macroscopically detectable dissociation into free solution when other ssDNA-binding proteins are present to compete with the transiently unbound species for exposed patches of ssDNA [38, 70]. This concentration-dependent disso- ciation mechanism has been referred as facilitated dissociation [315], and it has now been reported for several different proteins [316–320], suggesting that facilitated dissociation may have a widespread impact on the turnover of nucleic acid- binding proteins.

2.4.2 Protein Cofactor Association with the Presynaptic Complex

ssDNA curtains can be used to monitor the assembly and disassembly of pre- synaptic complexes with either single- or double-tethered ssDNA molecules, and can also be used to determine how other HR proteins bind to the presynaptic complexes. For instance, Rad52 is a mediator protein that promotes assembly of the Rad51 presynaptic complex during the early stages of HR [321, 322], and is also required for the second strand capture and strand annealing reactions that take place during the later stages of recombination [323–325]. We have used ssDNA curtains to study the spatial and temporal progression of RPA and Rad52 association with Rad51 during presynaptic complex assembly [38, 70]. These studies revealed that Rad52 can suppress RPA turnover, highlighting an unanticipated influence on protein dynamics, and also showed that both Rad52 and RPA can remain associated with the Rad51

presynaptic complex. As indicate earlier, there are ∼45 different proteins that participate in DSB repair in S. cerevisiae, and in many instances we have only a cursory understanding of how these proteins function. Future studies using ssDNA curtains help reveal how these proteins interact with and influence the Rad51 presynaptic complex.

2.4.3 Duplex DNA Binding by the Presynaptic Complex

We have used ssDNA curtains along with short (70-bp) fluorescently tagged dsDNA molecules containing short tracts of sequence microhomology complementary to the presynaptic ssDNA to study processes that take place as S. cerevisiae Rad51, human Rad51, S. cerevisiae Dmc1, human Dmc1, and E. coli RecA presynaptic complexes are attempting to align and pair homologous DNA sequences [304, 306]. This work revealed that dsDNA sequences bearing fewer than 8-nucleotides (nts) of micro- homology are rapidly sampled and rejected within seconds through a mechanism that gives rise to characteristic power law kinetics [306]. However, dsDNA molecules bearing 8-nts of microhomology are more tightly bound, and the resulting intermediates exhibit single-exponential kinetics with lifetimes corresponding to tens of minutes (Fig. 2.7). Increasing the length of microhomology from 8- to 15-nt reveals changes in lifetimes that take place in 3-nt increments, suggesting that strand exchange takes place in 3-nt steps, with each step exhibiting a characteristic energetic signature that appears to be broadly conserved among the Rad51/RecA family members (Fig. 2.7) [305, 306].

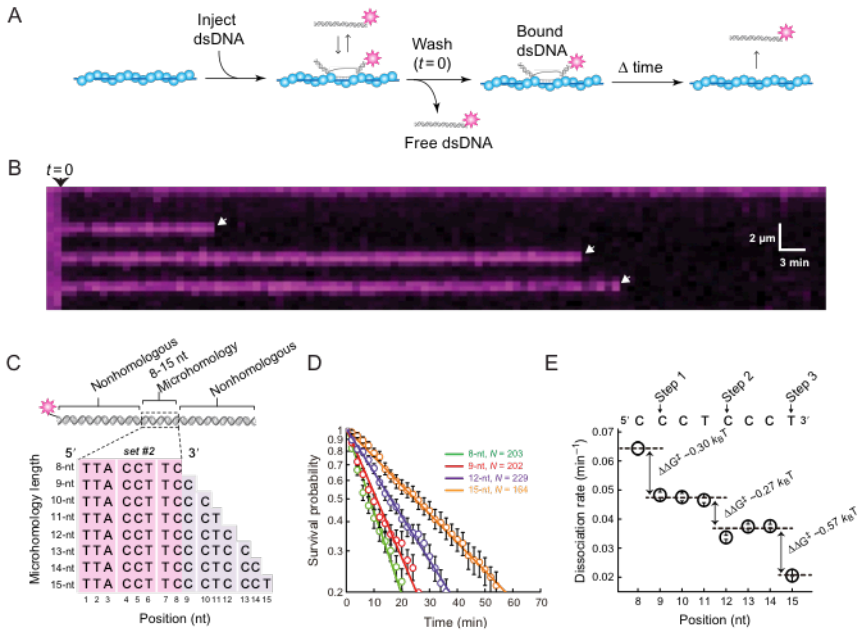

Figure 2.7: **Duplex DNA binding by the S. cerevisiae Rad51 presynaptic complex.** (**A**) Experimental schematic for measuring the survival probability of fluorescently tagged dsDNA oligonucleotides bound the presynaptic complex. (**B**) Example of a kymograph showing the binding of single Atto565–dsDNA molecules to a ScRad51 pre- synaptic complex. White arrowheads highlight individual dsDNA dissociation events. (**C**) Schematic of the 70-bp dsDNA substrates. All substrates contain an internal 8- to 15-nt tract of microhomology (as indicated) flanked by nonhomologous sequence. (**D**) Survival probability data for substrates with 8- to 15-nt of microhomology, as indi- cated; survival probability curves for the 10- and 11-nt substrates superimpose with the 9-nt and are omitted for clarity, and the 13- and 14-nt substrates superimpose 12-nt data sets and are also omitted for clarity. Data were analyzed by measuring the amount of time (dwell time) that each Atto565–DNA molecule remained bound to the Rad51– ssDNA presynaptic complexes after flushing unbound DNA from the sample chamber. (**E**) Atto565–dsDNA dissociation rates (mean ± s.d.) for reactions with S. cerevisiae Rad51 in the presence of AMP–PNP. Each data point was calculated from an average of ∼150 molecules (N = 70–250). Arrows indicate stepwise reductions in dissociation rates coincident with recognition of the 3rd base of each triplet, dashed lines report the mean rate for each step, and the free energy changes ($\Delta\Delta G^{\ddagger}$) associated with each triplet step are indicated. Adapted with permission from Lee, J. Y., Terakawa, T., Qi, Z., Steinfeld, J. B., Redding, S., Kwon, Y., et al. (2015). DNA recombination. Base triplet stepping by the Rad51/RecA family of recombinases. Science, 349, 977–981.

2.5 Data Collection and Analysis

In the following sections, we describe how to collect and analyze data from ssDNA curtain experiments using measurements of dsDNA binding to an unlabeled S. cerevisiae Rad51–ssDNA presynaptic complex as an example. Similar procedures can be used to study dsDNA-binding properties of other Rad51/RecA family members, or adapted to study interactions between the unlabeled presynaptic complexes and any fluorescently tagged protein or DNA component that binds to the presynaptic complex.

2.5.1 dsDNA Binding by the Rad51 Presynaptic Complex

To measure stable dsDNA binding, Rad51 presynaptic complexes are first prepared as described earlier (see Section 2.3.3.5). Atto565-tagged dsDNA oligonucleotides (10 nM) 70-bp in length are injected into the sample chamber in HR buffer (30 mM Tris-Acetate [pH 7.5], 20 mM Mg-Acetate, 50 mM KCl, 1 mM DTT, 0.2 mg/mL BSA) plus 2.5 mM ATP. Reactions are incubated for 10 min at 30°C in the absence of buffer flow and without laser illumination. The free dsDNA is removed by quickly flushing the sample chamber with HR buffer for 40 s at a flow rate of 1 mL/min. The flow rate is then reduced to 0.2 mL/min to allow continuous replenishment of ATP and removal of any dissociated dsDNA fragments. Data are acquired by capturing single 100 ms frames at either 20-, 30-, 40-, or 60-s intervals and the laser remains shuttered between each acquired image to minimize photobleaching. These measurements probe intermediates with lifetimes spanning a minute to tens of minutes and the image acquisition frequency and overall duration of the experiments are adjusted to accommodate the lifetime of each particular dsDNA substrate. Kymographs are then

generated from the resulting images as described later.

2.5.2 Generating Kymographs

Images are acquired through Nikon NIS Elements software and exported as individual tiff files for each exposure in the experiment. Using Fiji (ImageJ 1.48b, Wayne Rasband, National Institutes of Health, USA), the tiff files are stacked to create a tiff stack (i.e., movie) of the entire field of view. Kymographs representing individual presynaptic complexes are then generated from the resulting tiff stack using the "Reslice" function in Fiji. A straight line is superimposed on an ssDNA molecule and the corresponding image information for every image within the tiff stack is compiled as a new image. Each resulting kymograph represents a two-dimension projection of the events relating to a single presynaptic complex over the course of a reaction trajectory—the y-component reflects the position information and x-component represents time (Fig. 2.5-7). The resulting kymographs can then be used to assess the dsDNA-binding distributions, as previously described [306], as well as the lifetimes of fluorescent molecules bound to the presynaptic complex (see later) [305, 306].

2.5.3 Survival Probability

Dissociation kinetics are analyzed by measuring the amount of time (dwell time) that each molecule of Atto565-labeled dsDNA remained bound to the Rad51–ssDNA presynaptic complexes after flushing the unbound dsDNA from the sample chamber (Fig. 2.7A-C). Survival probability analysis allows one to extract lifetime information for the labeled dsDNA fragments bound to the unlabeled presynaptic complex. Multiple kymographs are

first analyzed manually to determine the lifetime of each individual binding event that is observed within the kymographs (Fig. 2.7B). These lifetimes are then plotted from shortest to longest and the resulting curves are analyzed to extract dissociation parameters. For a simple dissociation process, the distribution of lifetimes is expected to resemble a single-exponential decay. The probability of survival at time t is then defined as the proportion of molecules that remain bound to the ssDNA at time t, which can be defined as the number of molecules whose lifetime exceeds a particular time point, divided by the total number of observed binding events (Fig. 2.7D). Bootstrapping is a standard statistic method used to estimate confidence intervals of a population mean by randomly resampling a subset of data from within a larger data set [326]. Error bars for the survival probability measurements are then expressed as 70% confidence intervals obtained through this bootstrap analysis. Use of 70% confidence intervals for the bootstrapped data provides a close approximation to expectations for one standard deviation from the mean because for any normally distributed data set 68.27% of the values lie within one standard deviation of the mean.

2.5.4 Free Energy Changes During Base Triplet Stepping

Survival curves generated from experiments using the 70-bp dsDNA substrates bearing 8–15-nts of microhomology (Fig. 2.7C) can be fitted to a simple exponential decay function of the form e^{-kt}, where k is the experimentally observed dissociation rate constant and t is time (Fig. 2.7D). This observed dissociation rate constant is comprised of two components: the actual dsDNA dissociation rate constant and the Atto565 photobleaching rate constant. Subtracting photobleaching rate, which must be determined separately [306], from the ex-

perimentally observed dissociation rate k allows one to determine the actual dissociation constant for the dsDNA substrates (k_d). The Arrhenius equation can then be used to relate the experimentally determined k_d to the free energy barrier ($\Delta G^{\ddagger}$) for dsDNA binding as follows:

$$k_d = Ae^{-\frac{\Delta G^{\ddagger}}{k_b T}} \tag{2.1}$$

where A is the jump frequency, k_b is the Boltzmann constant, and T is temperature [306]. The difference in the barrier heights ($\Delta\Delta G^{\ddagger}$) between two different dsDNA substrates can be compared using the following relation:

$$\Delta\Delta G^{\ddagger} = \Delta G_2^{\ddagger} - \Delta G_1^{\ddagger} = k_b T ln \frac{k_d^1}{k_d^2} \tag{2.2}$$

These $\Delta\Delta G^{\ddagger}$ values can be normalized such that $\Delta G^{\ddagger}$ for a dsDNA substrate containing a single 8-nt of microhomology is set to zero, and all other $\Delta\Delta G^{\ddagger}$ values for dsDNA fragments harboring longer lengths of microhomology are expressed relative to the substrate bearing 8-nts of microhomology (Fig. 2.7E) [305].

2.5.5 Real-Time Binding Measurements

The experiments described earlier pertain to relatively stable reaction inter- mediates with lifetimes on the order of tens of minutes. However, less stable intermediates can also be readily detected by increasing the data acquisition frequency. For instance, transient dsDNA sampling by the Rad51–ssDNA presynaptic complex can be detected over much shorter time regimes [306]. This is accomplished by injecting the Atto565-tagged dsDNA substrate (10

nM) into a sample chamber containing assembled pre- synaptic complexes. Buffer flow is then terminated without flushing away the free dsDNA and images are acquired with 60 ms exposure time using continuous laser illumination in the absence of shuttering. The resulting data for the transient binding intermediates can then be analyzed as described earlier based on kymographs generated from the resulting tiff stacks. The dwell times for each binding event are then defined as the difference between the first frame and the last frame in which a particular molecule of Atto565- labeled dsDNA is observed bound to the Rad51 presynaptic complex. The data describing all of the transient dsDNA-binding events are then synchronized such that the initial appearance of each bound Atto565-DNA was defined as time zero, and then the probability that a bound molecule survived up to a particular time point (t) was determined as the fraction of Atto565–DNA molecules that remained bound at time t. Survival probability graphs can then constructed from the resulting data to analyze dwell times, as described earlier for the more stable dsDNA-binding intermediates.

2.6 Conclusion

ssDNA curtains provide a powerful experimental platform, enabling new avenues of investigation into the biochemical and biophysical proper- ties of Rad51/RecA–ssDNA presynaptic complexes. These studies offer the potential for new insights into the assembly, stability, and regulation of this crucial HR intermediate, and the procedures described here can be adapted to study many different questions related to HR. Future ssDNA curtains studies may help provide additional insights into the DNA transactions that take place during HR, and may also provide important new clues into the dozens of other proteins that are neces-

sary for HR to take place within living cells. Of particular interest will be work looking at how nucleosomes and chromatin impact the interactions of the Rad51–ssDNA presynaptic complex with dsDNA, and how these interactions are modulated by nucleosome-remodeling proteins and posttranslational histone modifications. In addition, these ssDNA curtain methods can be adapted for studies involving other types of ssDNA-binding proteins, and with additional development it may even be possible to extend these research tools to study of single-stranded RNA substrates.

Chapter 3

Base Triplet Stepping by the Rad51/RecA Family of Recombinases

This chapter is adapted from work originally published as: "Base triplet stepping by the Rad51/RecA family of recombinases," Ja Yil Lee, Tsuyoshi Terakawa, Zhi Qi, **Justin B. Steinfeld**, Sy Redding, YoungHo Kwon, William A. Gaines, Weixing Zhao, Patrick Sung, and Eric C. Greene. Science (2015). I was responsible for all the single-molecule stepping data represented in Figure 3.5B.

3.1 Abstract

DNA strand exchange plays a central role in genetic recombination across all kingdoms of life, but the physical basis for these reactions remains poorly defined. Using single-molecule imaging, we found that bacterial RecA and eukaryotic Rad51 and Dmc1 all stabilize strand exchange intermediates in precise three-nucleotides steps. Each step coincides with an energetic signature ($0.3k_BT$) that is conserved from bacteria to humans. Triplet recognition is strictly dependent on correct Watson-Crick pairing. Rad51, RecA, and Dmc1 can all step over mismatches, but only Dmc1 can stabilize mismatched triplets. This finding provides insight into why eukaryotes have evolved a meiosis-specific recombinase. We propose that

canonical Watson-Crick base triplets serve as the fundamental unit of pairing interactions during DNA recombination.

3.2 Introduction

Homologous recombination enables the exchange of genetic information between DNA molecules and is a driving force in evolution. During homologous recombination, a presynaptic single-stranded DNA (ssDNA) is paired with the complementary strand of a homologous double-stranded DNA (dsDNA), resulting in displacement of the noncomplementary strand [294]. This strand exchange reaction plays essential roles in double-strand DNA break (DSB) repair [293, 294], the rescue of stalled or collapsed replication forks [293, 327], chromosomal rearrangements [293], horizontal gene transfer [**1603**, 328], and meiosis [141, 152]. These reactions are promoted by the Rad51/RecA family of DNA recombinases, which are adenosine triphosphate (ATP)–dependent proteins that form helical filaments on DNA [192, 294]. Crystal structures of RecA-ssDNA presynaptic complexes and RecA-dsDNA postsynaptic complexes reveal that the DNA is organized into near B-form base triplets separated by ∼7.1 to 8.4 between adjacent triplets (Fig. 3.1A)[36, 329]; for brevity, we refer to these nucleic acids as RS-DNA (Rad51/RecA-stretched DNA).

We previously developed single-molecule methods to study DNA recombination by total internal reflection fluorescence microscopy (TIRFM) [306]. Using this approach, we showed that presynaptic complexes search for homology by sampling dsDNA for ≥8-nucleotide (nt) tracts of microhomology and rapidly reject sequences with ≤7 nt of microhomology [306]. We have also reported that strand exchange by Saccharomyces cerevisiae

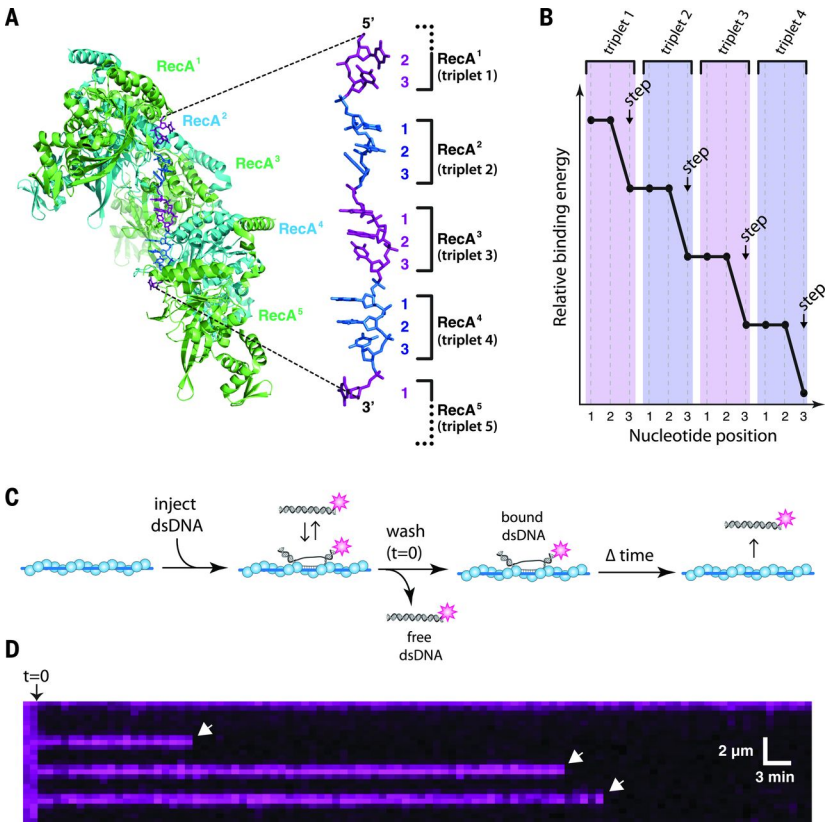

Figure 3.1: **Structure of the presynaptic complex and experimental overview.** (**A**) Structure of the RecA-ssDNA filament highlighting the base triplet organization of the presynaptic RS-ssDNA [36]. (**B**) Cartoon illustration of base triplet stepping for S. cerevisiae Rad51. Quantized reductions in binding energy are proposed to coincide with the third base of each triplet. (**C**) Experimental outline for measuring the survival probability of fluorescently tagged dsDNA oligonucleotides bound to the presynaptic complex. (**D**)Example of a kymograph showing the binding of single Atto565-dsDNA molecules to a ScRad51 presynaptic complex. White arrowheads highlight individual dsDNA dissociation events.

Rad51 (ScRad51) occurs in 3-nt steps, presumably reflecting the base triplet organization of RS-DNA (Fig. 3.1A, B) [306], and 3-nt stepping has also been proposed for bacterial RecA [330]. Here, we sought to determine the underlying principles that contribute to base triplet stepping and establish how these principles influence the mechanism and fidelity of DNA recombination. To address these questions, we assembled presynaptic complexes on ssDNA curtains (Fig. 3.1C and 3.2) using one of four different recombinases: Escherichia coli RecA (EcRecA)[294]; ScRad51 or human Rad51 (hRAD51), eukaryotic orthologs of RecA [293]; or S. cerevisiae Dmc1 (ScDmc1), which is specialized for meiotic recombination (Fig. 3.3) [152, 184].

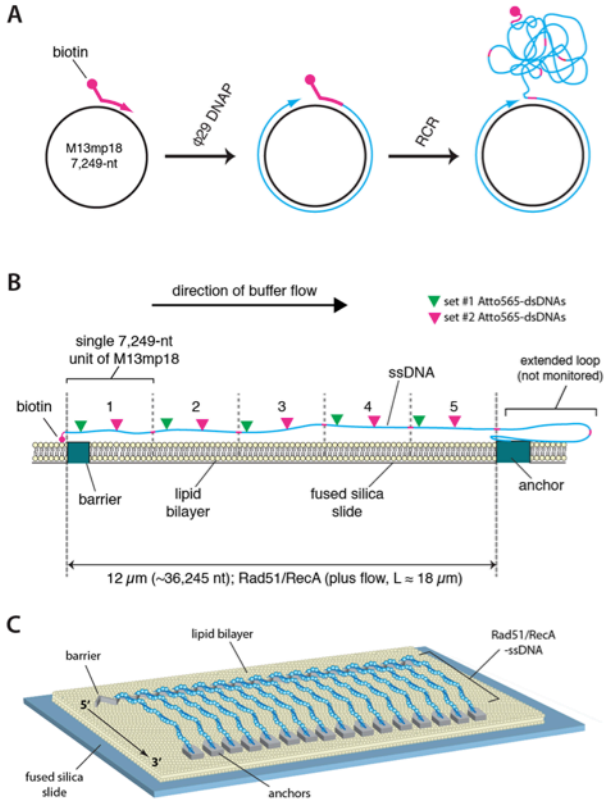

Figure 3.2: **Preparation of ssDNA curtains by rolling circle replication.** (**A**) Schematic showing the production of the presynaptic ssDNA by rolling circle replication (RCR) using an M13mp18 circular ssDNA template. Each substrate is biotinylated at the 5'end and is comprised of multiple units of M13mp18. (**B**) Schematic showing the ssDNA anchored to the surface of the microfluidic sample chamber. Each anchored substrate reflects 5 units of M13mp18 (numbered 1 through 5); any ssDNA longer than the 5 units extends beyond the anchor and is not monitored in the experiments (as indicated). Green and magenta arrowheads indicate the relative locations of the two sets of dsDNA oligonucleotides used in the binding experiments (see Tables 3.1, 3.2, **??**. (**C**) Schematic of an ssDNA curtain bound by Rad51.

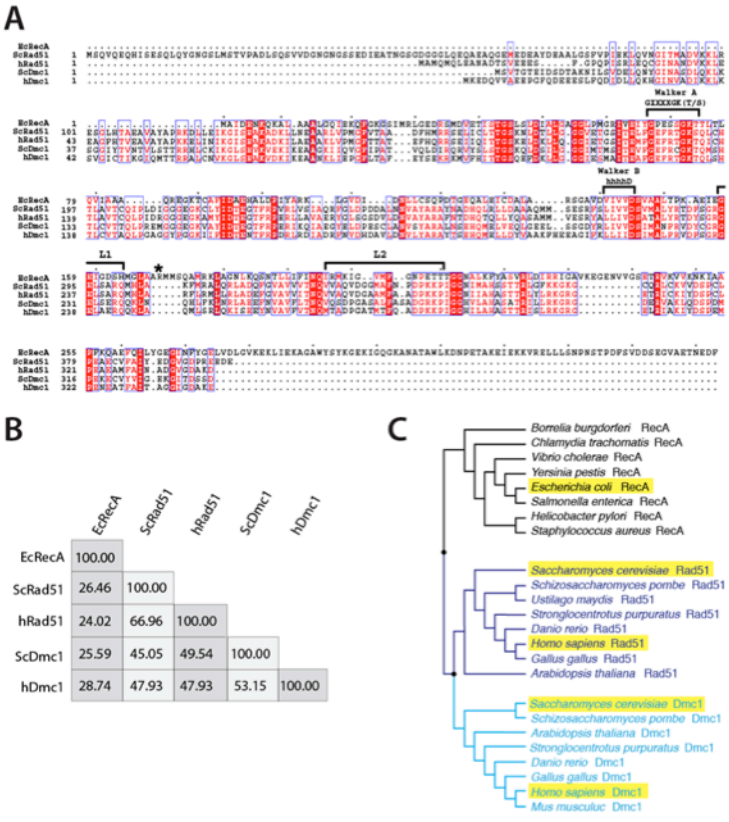

Figure 3.3: **Relationship among Rad51/RecA proteins used in this study.** (**A**) Sequence alignment of EcRecA, ScRad51, hRAD51, ScDmc1 and hDMC1. Identical amino acids are highlighted by a red background, similar amino acids are highlighted in red text and the locations of the Walker A and B motifs are indicated, as are the L1 and L2 loops, which contact the presynaptic ssDNA. Arginine 169 in EcRecA, which has been proposed to sense the minor groove geometry of the postsynaptic dsDNA (9), is marked with an asterisk. The numbering scheme for EcRecA begins at residue 2 because the N–terminal methionine is removed in vivo. (**B**) Percent identity between EcRecA, ScRad51, hRAD51, ScDmc1, and hDMC1. (**C**) Phylogenetic relationship of RecA, Rad51 and Dmc1; the three different lineages are color–coded, and the five recombinases used in this study are highlighted in yellow.

3.3 Results

For visualization of strand exchange intermediates, 70–base pair (bp) Atto565-dsDNA substrates were briefly incubated with the presynaptic complexes and unbound dsDNA was flushed away. Complexes were visualized by TIRFM (Fig. 3.1C, D) and dissociation rates were obtained from the survival probabilities of the bound Atto565-dsDNA (Fig. 3.4). The dsDNA substrates bore 8- to 15-nt tracts of microhomology targeted to two different regions of the presynaptic ssDNA (Fig. 3.2B, 3.5A and 3.6A and Table 3.1) [306]. The dissociation rates for both sets of substrates scaled with microhomology length for each of the recombinases. In each instance, pronounced changes in dissociation rates coincided with recognition of the 9th, 12th, and 15th nucleotides, and similar results were observed in reactions with the nonhydrolyzable ATP analogs adenylyl imidodiphosphate (AMP-PNP) and adenosine 5'-O-(3-thiotriphosphate) (ATPγS) (Fig. 3.5 and 3.6). These binding patterns demonstrate that 3-nt stepping is a broadly conserved feature of the Rad51/RecA family of DNA recombinases.

Comparison of reactions with EcRecA, ScRad51, hRAD51, and ScDmc1 revealed that the free energy change ($\Delta\Delta G^{\ddagger}$) associated with the binding of each base triplet was similar for all recombinases (Fig. 3.5 and 3.6), corresponding to 0.30 $\pm$ 0.14 k_BT (mean $\pm$ SD) for completion of a single triplet step (Fig. 3.7). This result supports the conclusion that the free energy changes associated with triplet steps during DNA recombination are broadly conserved. In addition, reactions with AMP-PNP or ATPγS revealed no appreciable shift in the free energy change associated with each triplet step (Fig. 3.5 and 3.6). Thus, the physical determinants governing the energetics of strand exchange have been retained during

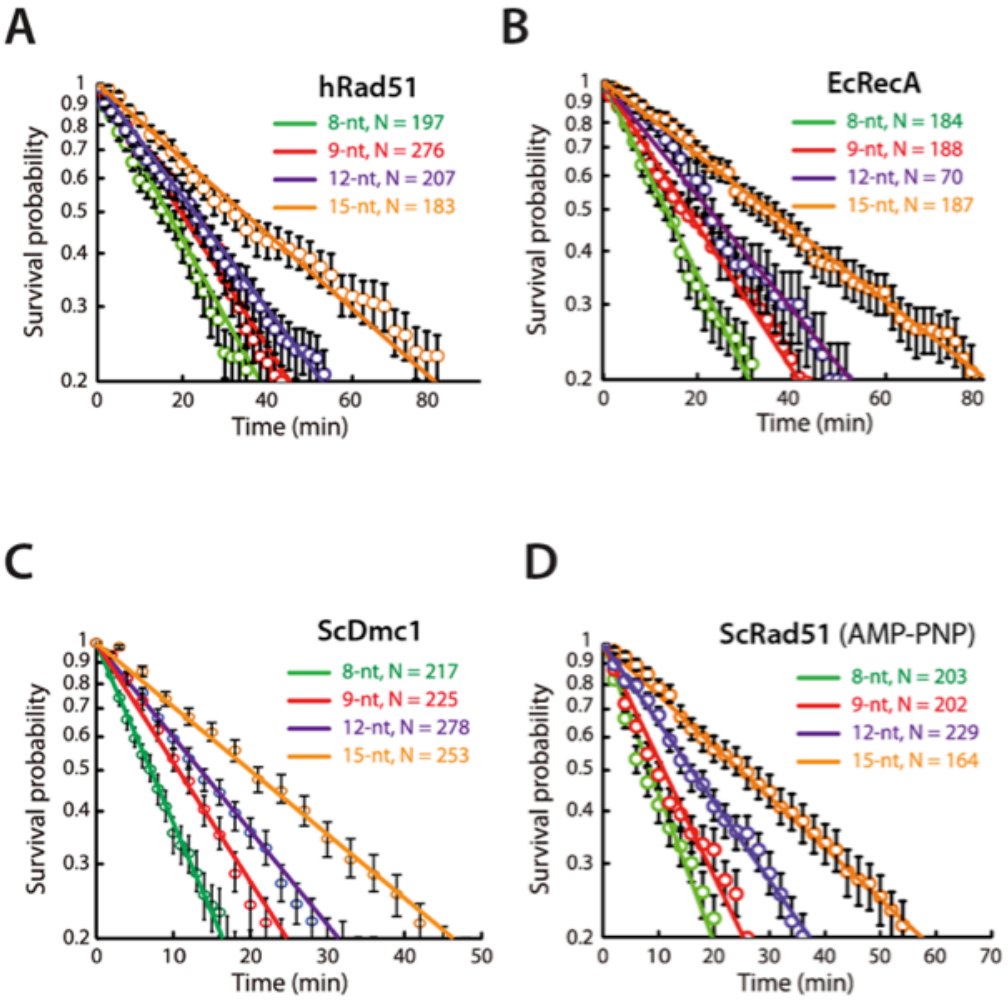

Figure 3.4: **Representative survival probability curves.** (**A**) hRAD51, (**B**) EcRecA, (**C**) ScDmc1, and (**D**) ScRad51 (plus AMP–PNP) using set #1 substrates with 8– to 15–nt of microhomology, as indicated; survival probability curves for the 10–, 11–, 13–, and 14–nt substrates superimpose with the 9–nt and 12–nt data sets and are omitted for clarity. Data were analyzed by measuring the amount of time (dwell time) that each Atto565–DNA molecule remained bound to the Rad51–ssDNA presynaptic complexes after flushing unbound DNA from the sample chamber. The probability that a bound molecule survived up to time point (t) was determined as the fraction of Atto565–DNA molecules that remained bound at time t, and survival probability graphs for the different substrates were constructed from the resulting data. N corresponds to the total number of Atto565–DNA molecules measured. Error bars for survival probability plots represent 70% confidence intervals obtained through bootstrap analysis; our choice of 70% confidence intervals for the bootstrapped data provides a close approximation to expectations for one standard deviation from the mean.

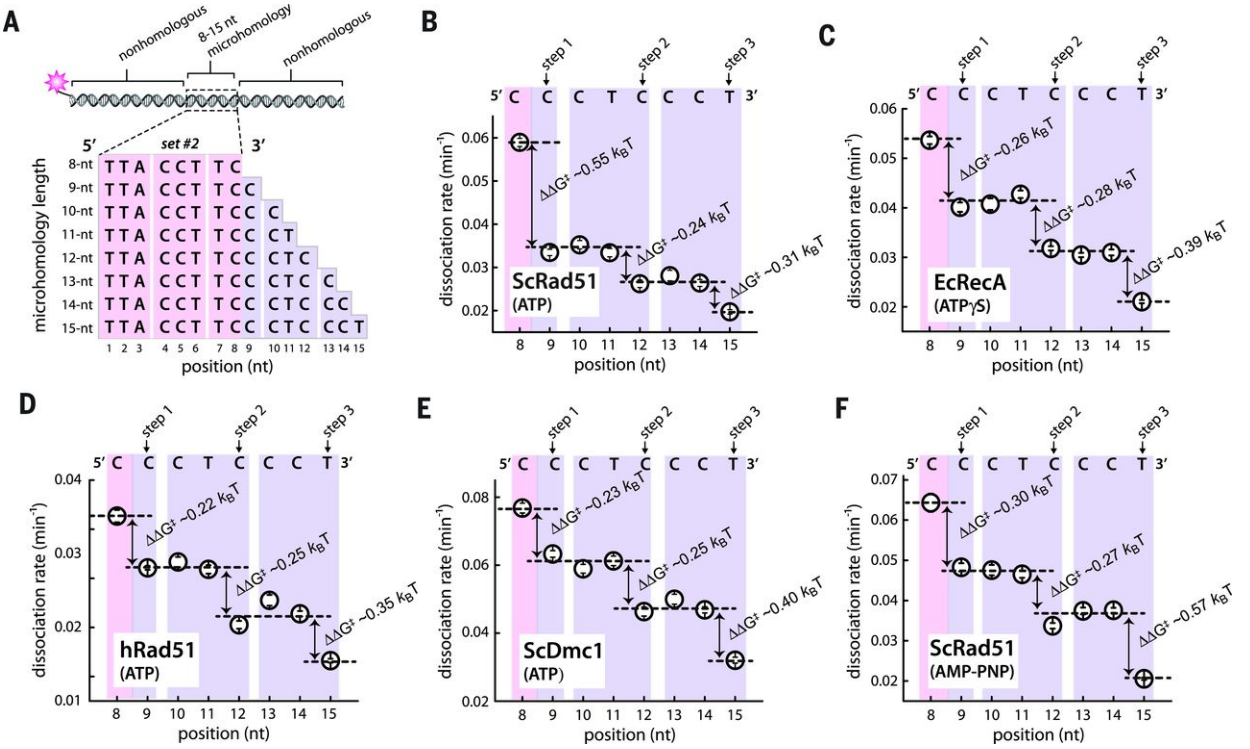

Figure 3.5: **Conservation of base triplet stepping among Rad51/RecA family members.** (**A**) Schematic of the 70-bp dsDNA substrates (set #2; see Table 3.1). All substrates contain an internal 8- to 15-nt tract of microhomology (as indicated) flanked by nonhomologous sequence. (**B**) Atto565-dsDNA dissociation rates (mean ± SD) for ScRad51 plus ATP. (**C**) EcRecA plus ATPγS. (**D**) hRAD51 plus ATP. (**E**) ScDmc1 plus ATP. (**F**) ScRad51 plus AMP-PNP. In (B) to (F), each data point was calculated from an average of ∼150 molecules (N = 70 to 250); the color-coded shading highlights each base triplet; magenta shading indicates the minimum 8 nt necessary for binding; purple shading corresponds to additional homologous nucleotides; arrows indicate stepwise reductions in dissociation rates coincident with recognition of the third base of each triplet; dashed lines report the mean rate for each step; and the free energy changes ($\Delta\Delta G^{\ddagger}$) associated with each triplet step are indicated.

the evolution of the Rad51/RecA gene family.

Whereas base stacking dominates the stability of B-DNA [331], stacking interactions are disrupted in RS-DNA, which suggests that the structure of RS-DNA may enhance the fidelity of recombination by relying more on correct Watson-Crick pairing. Therefore, we next asked whether noncomplementary bases affect individual strand exchange steps (Fig. 3.8A). These experiments demonstrated that a single mismatch anywhere within a base triplet completely abolishes recognition of the entire triplet (Fig. 3.8B and 3.9). All noncomple-

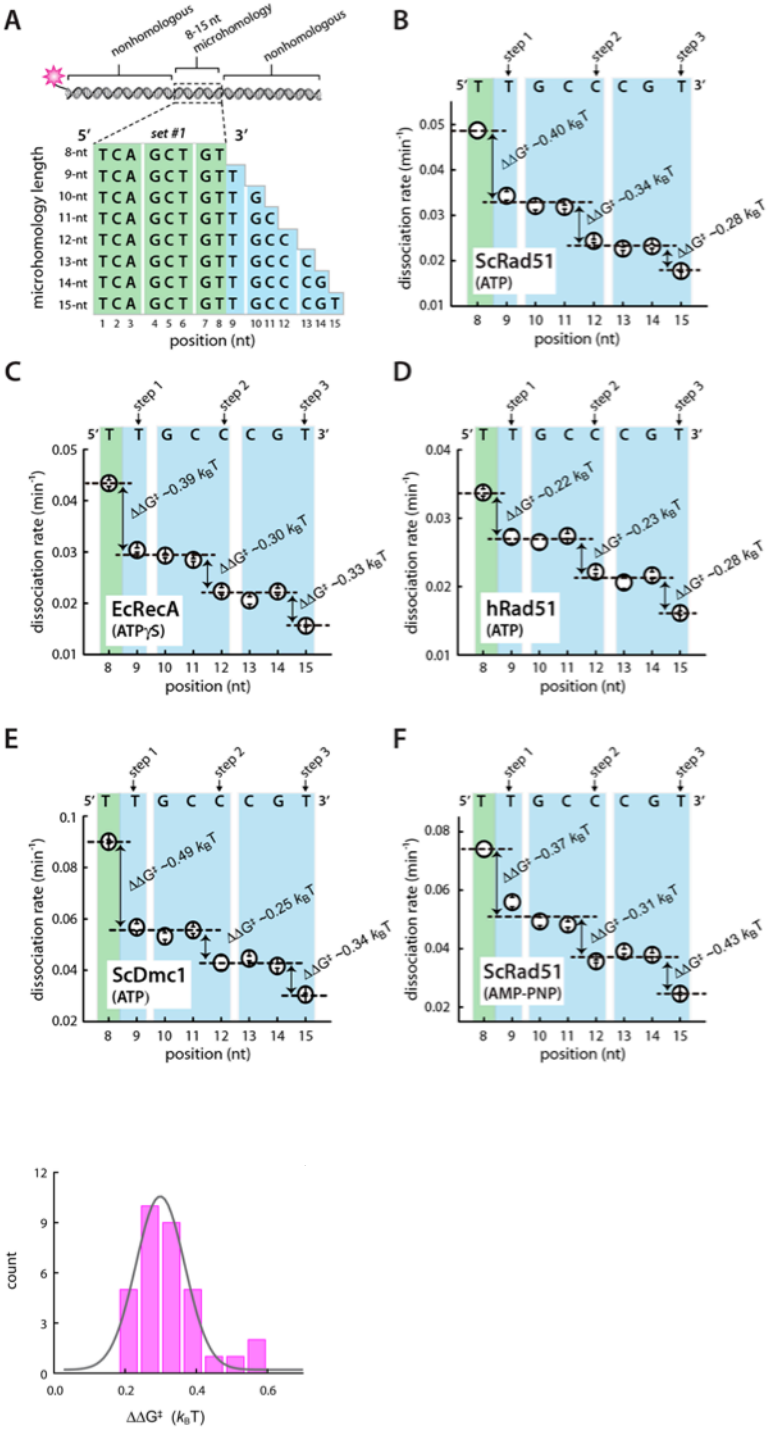

Figure 3.6: **Conservation of base triplet stepping.** (**A**) Schematic of the 70–bp dsDNA substrates (set #1; see Table 2.1). All substrates contain an internal 8– to 15–nt tract of microhomology (as indicated) flanked by nonhomologous sequence. (**B**) Atto565– dsDNA dissociation rates (mean ± s.d.) for ScRad51 plus ATP. (**C**)EcRecA plus ATPγS. (**D**) hRAD51 plus ATP. (**E**) ScDmc1 plus ATP. (**F**) ScRad51 plus AMP–PNP. In (B-F), the color–coded shading highlights each base triplet, arrows indicate stepwise reductions in dissociation rates coincident with recognition of the 3rd base of each triplet, and the free energy changes ($\Delta\Delta G^{\ddagger}$) associated with each step are indicated.

Figure 3.7: **Shared energetic signature for triplet stepping.** Histogram showing the free energy changes associated with base triplet recognition for all individual steps (N=33 total) measured for EcRecA, ScRad51, hRad51, ScDmc1 and hDmc1 (see Fig. 3.5, 3.6, 3.11). A Gaussian fit to the distribution yields a value of 0.3 ± 0.14 kBT (mean ± s.d.) for the free energy change that takes place during 3–nt strand exchange steps.

Set	Microhomology length	Top strand sequence†
#1	8–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GT***A GGC CTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	9–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT*** AGC CTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	10–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT G***AC CTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	11–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GC***G GTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	12–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GCC*** GTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	13–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GCC C***AT AGC TAG CTA GCT AGC TAG CTA GCT
#1	14–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GCC CG***G AGC TAG CTA GCT AGC TAG CTA GCT
#1	15–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GCC CGT*** GGC TAG CTA GCT AGC TAG CTA GCT
#2	8–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TC***A GGC CTT AGC TAG CTA GCT AGC TAG CTA GCT
#2	9–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC*** AGC CTT AGC TAG CTA GCT AGC TAG CTA GCT
#2	10–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC C***GC CTT AGC TAG CTA GCT AGC TAG CTA GCT
#2	11–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CT***A ATT AGC TAG CTA GCT AGC TAG CTA GCT
#2	12–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CTC*** ACT AGC TAG CTA GCT AGC TAG CTA GCT
#2	13–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CTC C***TT AGC TAG CTA GCT AGC TAG CTA GCT
#2	14–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CTC CC***A AGC TAG CTA GCT AGC TAG CTA GCT
#2	15–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CTC CCT*** AGC TAG CTA GCT AGC TAG CTA GCT

Table 3.1: **Sequences of 70–bp oligonucleotides with 8–nt to 15–nt tracts of microhomology.** †Only the Atto565 labeled top strands of each oligonucleotide are shown; the tracts of microhomology complementary to the M13mp18 ssDNA substrate are designated in underlined bold italics. The corresponding bottom strands are complementary to the top strand sequences, but lack a fluorescent label (not shown). Microhomology analysis confirmed that the sequences were targeted to only a single region on M13mp18, corresponding the complement of the underlined sequences.

mentary nucleotides abolished triplet recognition regardless of mismatch identity, and this high level of discrimination was conserved across the Rad51/RecA family (Fig. 3.8B and 3.9). In addition, reactions with ATPγS or AMP-PNP revealed that ATP hydrolysis played no discernible role in mismatch discrimination at the level of a single triplet step (Fig. 3.8B and 3.9C,F).

We next sought to determine whether strand exchange could progress beyond a mis-

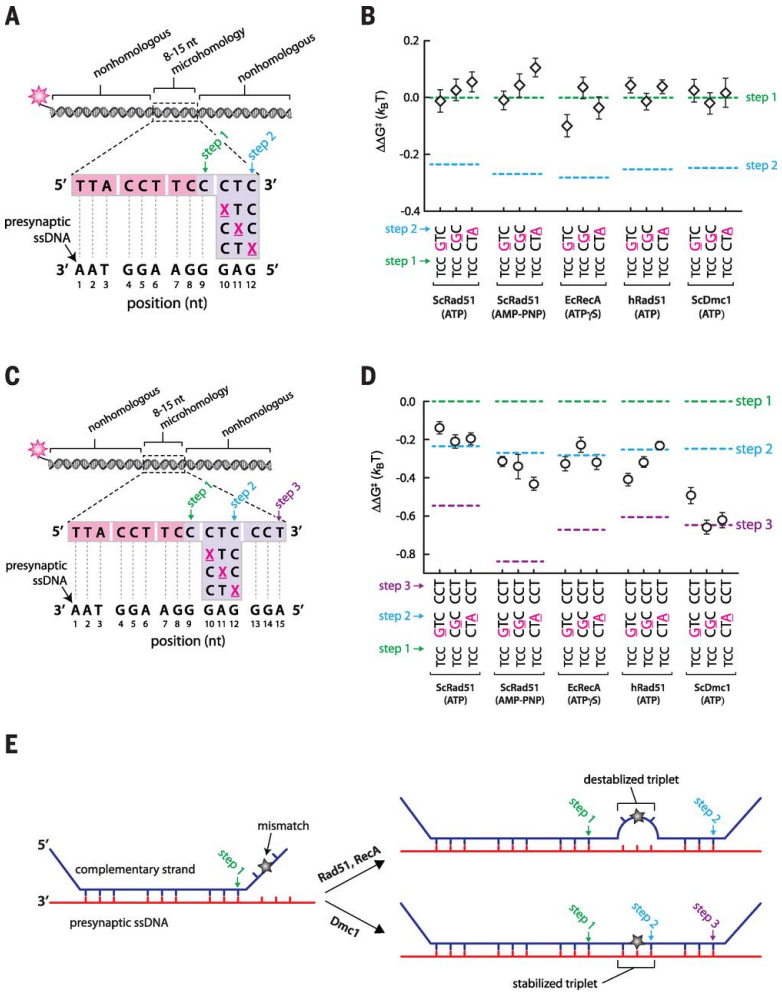

Figure 3.8: **Effects of mismatches on base triplet recognition.** (**A**) Schematic highlighting the design of dsDNA substrates (set #2; see Table 3.2) bearing mismatched bases within the fourth nucleotide triplet, corresponding to the second strand exchange step; each mismatch position is highlighted as an underlined magenta X. (**B**)Mismatch substrate binding for ScRad51, EcRecA, hRAD51, and ScDmc1, as indicated. All indicated $\Delta\Delta G^{\ddagger}$ values are relative to the step 1 binding data for the substrate bearing 9 nt of microhomology (see Fig. 2.5); each data point was calculated from an average of ∼150 molecules. The green and blue dashed lines correspond to data obtained for non-mismatched substrates with each of the four different recombinases (see Fig. 3.5B-F). (**C**) Substrate design (set #2; see Table 3.3) for testing whether Rad51/RecA is capable of stepping over a mismatched triplet. (**D**) Triplet binding data illustrating how base mismatches within the fourth triplet affect recognition of the fifth triplet. (**E**) Schematic highlighting the different models predicted for mismatched substrates in reactions with RecA and Rad51 versus Dmc1.

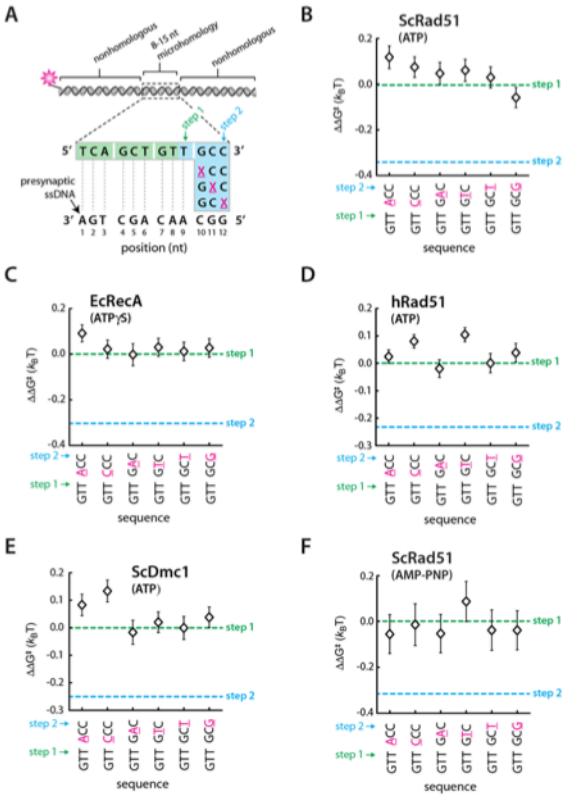

Figure 3.9: **Effects of mismatches on base triplet recognition (cont.).** (**A**) Schematic illustration of mismatched set #1 dsDNA substrates. The location of each mismatch within the fourth base triplet is indicated as an underlined, magenta "X". The corresponding sequence of the M13mp18 presynaptic ssDNA is shown to highlight the identity of the mismatched pair (Also see Table 3.2). (**B**) Changes in binding free energy for mismatch substrates in reactions with ScRad51, EcRecA, hRad51, and ScDmc1, as indicated. The green and light blue dashed lines correspond to step 1 and step 2, respectively, obtained for non– mismatched substrates with each of the different recombinases (see Fig. 3.6)

match (Fig. 3.8C and 3.10A). Extending the length of homology by just a single triplet allowed each recombinase to step past the mismatches, as evidenced by the corresponding reduction in binding free energy (Fig. 3.8D and 3.10B). When EcRecA, ScRad51, or hRAD51 stepped over a mismatched triplet, the mismatched triplets did not contribute to the binding free energy of the resulting intermediates. This conclusion is based on the observation that the corresponding reduction in binding free energy was comparable to products reflecting a second rather than a third strand exchange step (Fig. 3.8D and 3.10B). This result supports a model in which these internal mismatch-bearing triplets remain destabilized (Fig. 3.8E). In contrast, when ScDmc1 stepped over a mismatched triplet, the stability of the resulting intermediates was comparable to expectations for completion of both the second and the third strand exchange steps (Fig. 3.8D and 3.10B). The simplest interpretation of this result is that in reactions with ScDmc1, the internal mismatched triplet was

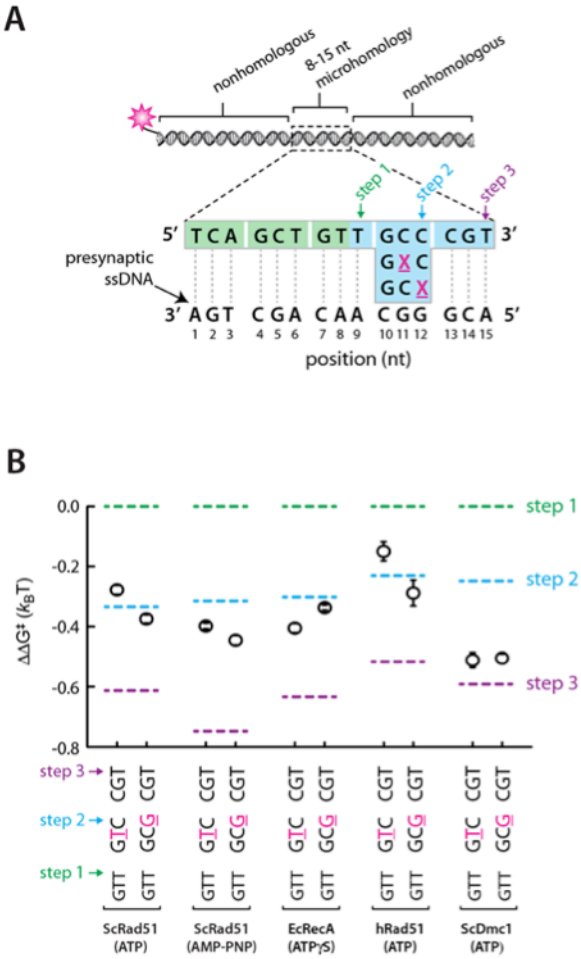

Figure 3.10: **RecA, Rad51 and Dmc1 can step over mismatched triplets.** (**A**) Substrate design based on oligonucleotide set #1 (Table **??**) for testing whether Rad51/RecA is capable of stepping over a mismatched triplet; note that position 10 could not be altered without introducing new tracts of microhomology, so mutations at this position were not tested. (**B**) Triplet binding data illustrating how base mismatches within the 4th triplet affect recognition of the 5^{th} triplet.

either partially or fully paired with presynaptic ssDNA (Fig. 3.8E). To determine whether the ability to stabilize mismatched triplets is a conserved feature of the Dmc1 lineage, we next performed experiments using human DMC1 (hDMC1). Like the other recombinases, hDMC1 strand exchange intermediates were stabilized in 3-nt steps, triplet recognition coincided with a change in free energy of $\sim 0.3 k_B T$, and mismatches abolished triplet recognition (Fig. 3.11A-D). Like ScDmc1, hDmc1 is able to step over mismatches, and also stabilizes the mismatched triplets (Fig. 3.11E,F). Although Rad51, RecA, and Dmc1 can all step over mismatches, the ability to stabilize mismatched triplets embedded within longer tracts of homology is only conserved within the Dmc1 lineage.

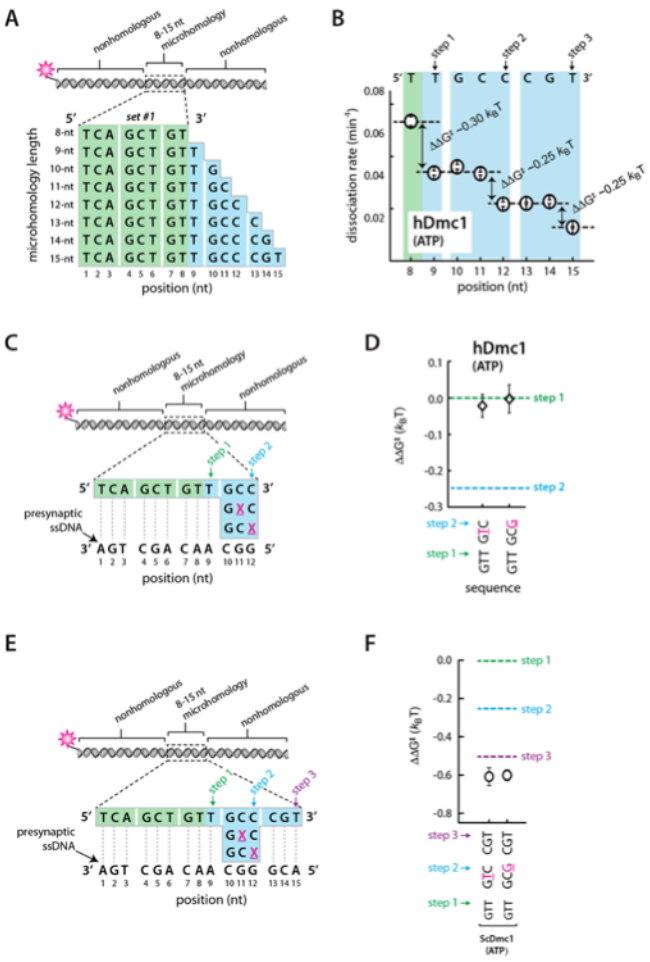

Figure 3.11: **Strand exchange characteristics of human DMC1.** (**A**) Schematic of the 70–bp dsDNA substrates (set #1). All substrates contain an internal 8– to 15–nt tract of microhomology (as indicated) flanked by nonhomologous sequences. (**B**) Atto565– dsDNA dissociation rates (mean /*pm* s.d.) for hDMC1 plus ATP. (**C**) Schematic of mismatched substrates. (**D**) Mismatch substrate binding for hDMC1. All indicated $\Delta\Delta G^{\ddagger}$ values are relative to the step 1 binding data for the substrate bearing 9–nt of microhomology. (**E**) Substrate design for testing mismatch bypass. (**F**) hDMC1 binding data illustrating how base mismatches within the 4^{th} triplet affect recognition of the 5^{th} triplet.

It is not known why eukaryotes have evolved two recombinases. Rad51 is expressed ubiquitously, but Dmc1 is present only during meiosis[152, 184]. A fundamental difference between mitotic and meiotic recombination is the choice of template used to direct DSB repair [141, 152, 332]. Mitotic recombination is biased toward use of the identical sister chromatid to ensure accurate repair. Meiotic recombination favors use of the homolog, yielding crossovers and increasing genome diversity through reshuffling of parental alleles and meiotic gene conversion [141, 152, 332]. The mechanisms that direct template choice remain poorly defined [141, 152, 332]. Our findings reveal that Dmc1 can stabilize mismatches, which may play a role in guiding template choice. For example, the inability

of Rad51 to stabilize mismatches could help bias mitotic recombination toward use of the sister chromatid by disfavoring interhomolog recombination. This bias may be enforced through mismatch repair (MMR) anti-recombination activity [333, 334], which could selectively disrupt mismatched intermediates arising from any attempts by Rad51 to promote interhomolog recombination. Conversely, the ability of Dmc1 to stabilize mismatches may allow recombination between polymorphic maternal and paternal alleles during meiosis by masking mismatched intermediates from premature dissolution by the MMR machinery. Dissociation of Dmc1 upon completion of strand invasion would then allow efficient gene conversion through MMR-mediated repair of the mismatched heteroduplex.

We next sought a quantitative explanation for why strand exchange takes place in 3-nt steps. Given our findings, it is conceivable that base-pairing transitions within the interior of the presynaptic complexes are governed primarily by the thermodynamic characteristics of RS-DNA. Concordant with this view is the finding that the complementary strand within the RecA-dsDNA postsynaptic complex is held in place mainly by Watson-Crick hydrogen bonds [36]. However, the thermodynamic properties of RS-DNA cannot be experimentally accessed outside the context of Rad51/RecA nucleoprotein filaments. As an alternative, we used Monte Carlo and molecular dynamics simulations to explore how extension of every third phosphodiester bond may alter the melting and annealing transitions of RS-DNA. These simulations used a coarse-grained model that recapitulates structural, thermodynamic, and mechanical characteristics of DNA [335]. Although more complex in silico models can be envisaged, such models would require further assumptions that could only serve to confuse the general argument that the physical architecture of RS-DNA itself may be sufficient to define some mechanistic attributes of strand exchange. Monte Carlo simula-

tions reveal that RS-DNA annealing is highly unfavorable because of the energetic penalty associated with extension of the complementary ssDNA strand (Fig. 3.12A). However, annealing becomes favorable when the incoming homologous ssDNA strand is constrained into base triplets to mimic the effect of the recombinases (Fig. 3.12A). This result agrees with the expectation that Rad51/RecA must locally stretch the incoming homologous DNA to more closely conform to the extended configuration of the presynaptic RS-ssDNA [36, 329, 336]. Moreover, the free energy profiles for RS-DNA reveal pronounced energetic barriers with 3-nt periodicity (Fig. 3.12A). Molecular dynamics simulations confirm that B-DNA annealing and melting transitions occur in 1-bp increments; consistent with a 1-bp zippering mechanism [335], the lifetimes of pairing intermediates are largely independent of nucleotide position (Fig. 3.12B,D and Fig. 3.13A-C). In contrast, the annealing and melting transitions for RS-DNA occur in 3-nt steps (Fig. 3.12C,D and Fig. 3.13D-F), and partially paired RS-DNA triplets are highly transient relative to fully paired triplets (Fig. 3.12D and Fig. 3.13E). These simulations reveal that partially paired RS-DNA triplets are not stable, which recapitulates a key feature of our experimental data—base triplet stepping—even though the model intentionally omits any detailed contributions of the proteins or amino acid side chains other than to stretch the DNA strands into an RS-DNA configuration. Together, the simulations and experimental work suggest that canonical Watson-Crick base triplets act as the fundamental pairing unit within RS-DNA.

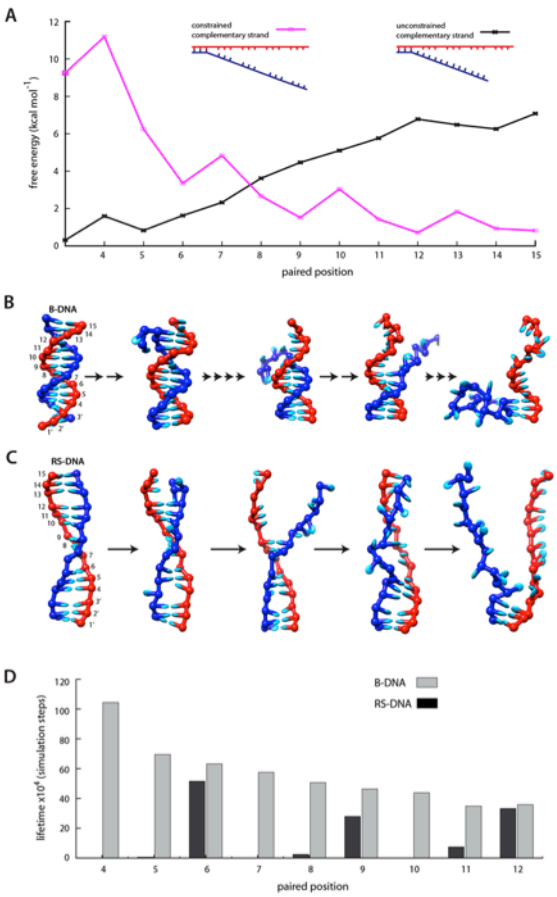

Figure 3.12: **RS–DNA melting occurs in 3–nt steps.** (**A**) Free energy profile of RS–DNA annealing generated by Monte Carlo simulation. (**A**) Molecular dynamics snapshots of B– DNA melting intermediates. (**C**) Molecular dynamics snapshots of RS–DNA melting intermediates. (**D**) Lifetimes of B–DNA and RS–DNA intermediates calculated from 50 separate simulation runs (5×10^8 total simulation steps). In (B) and (D), the number of arrowheads between each image corresponds to the number of intermediates necessary to reach the depicted state. The data in (D) do not include the terminal triplets.

3.4 Discussion

RecA, Rad51, and Dmc1 differ in many structural, functional, and mechanistic details; they each interact with different accessory factors; and they are adapted for diverse biological functions [152, 293, 294]. Our work reveals that despite these differences, unifying mechanistic principles underlie the process of strand exchange: Rad51/RecA family members all stabilize strand exchange intermediates in 3-nt steps; each step coincides with a broadly conserved energetic signature; and a single mismatch can abolish recognition of an entire base triplet. We also show that RecA, Rad51, and Dmc1 can all step over mismatches, but only Dmc1 can stabilize mismatched triplets; this distinction likely reflects the role of Dmc1 in promoting strand exchange between polymorphic alleles during meiosis.

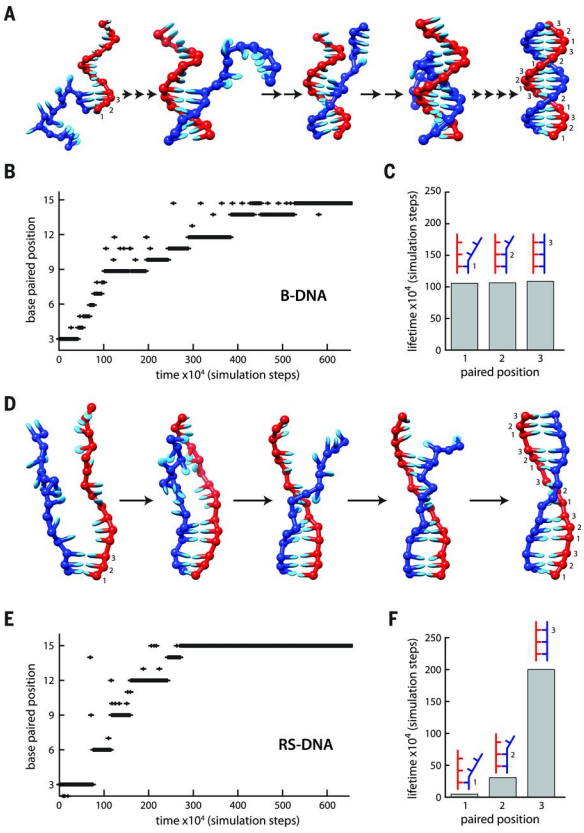

Figure 3.13: **RS-DNA governs triplet stepping during strand exchange.** (**A**) Molecular dynamics snapshots of B-DNA annealing intermediates. Paired bases are numbered 1 to 3 for comparison to RS-DNA triplets. (**B**) Sample annealing trajectory for B-DNA. (**C**) Lifetimes of B-DNA intermediates calculated from 50 separate simulation runs (5×10^8 simulation steps total). (**D**) Molecular dynamics snapshots of RS-DNA. (**E**) Sample annealing trajectory for RS-DNA. (**F**) Lifetimes of RS-DNA intermediates calculated from 50 separate simulation runs (5×10^8 simulation steps total). In (A) and (D), the number of arrows between each image corresponds to the number of steps necessary to reach the depicted state. The y axes in (B) and (E) denote the position of the most distal paired base relative to the pre-annealed end of the DNA. The data in (C) and (F) do not include the terminal triplets.

Our findings support a model in which a primary role of the Rad51/RecA family members during strand exchange is to establish the structure of RS-DNA, whereas the mechanism, energetics, and fidelity of the base-pairing transitions that take place during recombination are all governed by the physical architecture of RS-DNA itself. We propose that RS-DNA was selected at least ∼2 billion years ago by an ancestral recombinase as the most energetically favorable solution for allowing efficient exchange of genetic information between related DNA molecules.

3.5 Materials and Methods.

3.5.1 Fluorescent dsDNA substrates

Oligonucleotides were obtained from Integrated DNA Technologies (IDT) and purified by HLPC reverse phase chromatography. Stock solutions were quantified by UV absorbance and diluted to 100 μM oligonucleotide using 10 mM Tris–HCl [pH 8.5]. Fluorescently–tagged and untagged complementary ssDNA oligonucleotides were mixed in a 1:1.2 ratio and annealed in buffer containing 40 mM Tris–HCl [pH 8.0], 50 mM NaCl, and 10 mM $MgCl_2$. Reaction mixes were heated 5 minutes at 90 °C, then cooled slowly to room temperature over a period of 4 hours. Annealed oligonucleotides were purified using a QIAquick Gel Extraction Kit (#28704), and the purity of the annealed substrates was verified by electrophoresis on 12% DNA polyacrylamide gels. Fluorescent substrates were protected from light during preparation and storage.

All dsDNA sequences (Table 3.1) were analyzed for microhomology to ensure that they were targeted to a single unique location on the M13mp18 presynaptic ssDNA using a MatLab algorithm (MathWorks, Inc., Natick, MA), as described [306]. This analysis was essential to ensure that the survival probability data reflected the dissociation of Atto565–dsDNAs bound to identical locations on the presynaptic ssDNA. In brief, the dsDNA strands (top and bottom stands) were scanned in 3– to 10–nt increments along the M13mp18 ssDNA sequence to identify all corresponding tracts of microhomology. This process was also repeated for all mismatched substrates (Tables 3.2 and 3.3) to ensure that introduction of the mismatched bases did not also result in new $\geq$8–nt tracts of microhomology. If a desired mismatch resulted in a new tract of microhomology, then we first attempted to modify

Set	4th Triplet‡	Top strand sequence†
#1	GCC	5'–Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GCC*** GTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	XCC	5'–Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT ACC*** CTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	XCC	5'–Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT CCC*** GTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	GXC	5'–Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GAC*** CTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	GXC	5'–Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GTC*** CTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	GCX	5'–Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GCT*** GTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	GCX	5'–Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GCG*** GTT AGC TAG CTA GCT AGC TAG CTA GCT
#2	CTC	5'–Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CTC*** ACT AGC TAG CTA GCT AGC TAG CTA GCT
#2	XTC	5'–Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC GTC*** TCT AGC TAG CTA GCT AGC TAG CTA GCT
#2	CXC	5'–Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CGC*** ACT AGC TAG CTA GCT AGC TAG CTA GCT
#2	CTX	5'–Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CTA*** ACT AGC TAG CTA GCT AGC TAG CTA GCT

Table 3.2: **DNA sequences for 70–bp oligonucleotides with 12–nt microhomology tracts bearing mismatched triplets.** †Only the Atto565 labeled top strands of each oligonucleotide are shown; the tracts of microhomology complementary to M13mp18 ssDNA designated in underlined bold italics. The corresponding bottom strands are complementary to top strand sequences, but lack fluorescent label. ‡"X" corresponds to the mismatched position, and the identity and location of the mismatched bases in the DNA sequences are highlighted in magenta. All oligonucleotide sequences were analyzed for microhomology to ensure that the introduction of the mismatched bases did not result in new regions of microhomology; mismatches that resulted in new regions of microhomology were excluded from this study.

nucleotides in the flanking sequences to remove the undesired microhomology. If it was not possible to ensure that a particular mismatched dsDNA substrate was targeted to a single region on M13mp18, then that substrate was omitted from the study.

3.5.2 DNA binding experiments and data analysis

Flowcells, presynaptic ssDNA substrates, and ssDNA curtains were prepared essentially as described [81, 312]. Rad51/RecA presynaptic complexes were assembled as previously described [306], using reaction buffers adapted from prior studies of EcRecA [337–339], ScRad51 [340], hRAD51 [341], and ScDmc1 [342], as summarized in Table 3.4. Atto565–

Set	4th Triplet‡	Top strand sequence†
#1	GCC	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GCC CGT*** AGC TAG CTA GCT AGC TAG CTA GCT
#1	GXC	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GTC CGT*** GGC TAG CTA GCT AGC TAG CTA GCT
#1	GCX	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GCG CGT*** GGC TAG CTA GCT AGC TAG CTA GCT
#2	CTC	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CTC CCT*** AGC TAG CTA GCT AGC TAG CTA GCT
#2	XTC	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC GTC CCT*** AGC TAG CTA GCT AGC TAG CTA GCT
#2	CXC	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CGC CCT*** CGC TAG CTA GCT AGC TAG CTA GCT
#2	CTX	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CTA CCT*** AGC TAG CTA GCT AGC TAG CTA GCT

Table 3.3: **DNA sequences for 70–bp oligonucleotides with 15–nt microhomology tracts bearing mismatched triplets.** †Only the Atto565 labeled top strands of each oligonucleotide are shown; the tracts of microhomology complementary to M13mp18 ssDNA designated in underlined bold italics. The corresponding bottom strands are complementary to top strand sequences, but lack fluorescent label. ‡corresponds to the mismatched position, and the identity and location of the mismatched bases in the DNA sequences are highlighted in magenta. All oligonucleotide sequences were analyzed for microhomology to ensure that the introduction of the mismatched bases did not result in new regions of microhomology; mismatches that resulted in new regions of microhomology were excluded from this study.

tagged dsDNA oligonucleotides (2–10 nM) were then injected into the sample chamber in the same buffers, and reactions were incubated for a period of 10 minutes at 30 °C in the absence of buffer flow. Flowcells were quickly flushed (40 sec at 1.0 ml min^{-1}) with fresh buffer to remove unbound Atto565–dsDNA, and the flow rate was then reduced (0.2 ml min^{-1}) to ensure removal of dissociated dsDNA and replenishment of free nucleotide cofactor while images were being collected. Data were obtained by acquiring single 100–msec frames at 30– or 60–second intervals. The data collection intervals were optimized relative to the overall lifetime of each dsDNA substrate and the laser was shuttered between acquired images to minimize photobleaching. Kymographs were then generated from the resulting movies using Fiji. Survival probabilities were determined from analysis of the

Protein	Reaction buffer
EcRecA	25 mM Tris–Acetate [pH 7.5], 4 mM Mg–Acetate, 10 mM Na–Acetate, 1 mM ATPγS, 1 mM DTT, and 0.2 mg ml^{-1} BSA
hRad51	30 mM Tris–Acetate [pH 7.5], 1 mM $MgCl_2$, 5 mM $CaCl_2$, 100 mM KCl, 1 mM ATP, 1 mM DTT, and 0.2 mg ml^{-1} BSA
ScRad51	30 mM Tris–Acetate [pH 7.5], 20 mM Mg–Acetate, 50 mM KCl, 1 mM DTT, 2.5 mM ATP or 1 mM AMP–PNP (as indicated), and 0.2 mg ml^{-1} BSA
ScDmc1 and hDmc1	40 mM Tris–HCl [pH 7.5], 2 mM $MgCl_2$, 1.5 mM $CaCl_2$, 100 mM KCl, 2.5 mM ATP, 1 mM DTT, and 0.2 mg ml^{-1} BSA

Table 3.4: **Reaction conditions required for each of the recombinase.**

resulting kymographs by measuring the time (dwell time) that each molecule of Atto565–DNA remained bound to the presynaptic complexes after flushing the unbound DNA from the sample chamber. The probability that a bound molecule survived to a particular time point (t) was determined as the fraction of Atto565–dsDNA molecules that remained bound at time t, and survival probability graphs were constructed from the resulting data. All reported data points were calculated from an average of 150 different molecules (N=70–250). In each instance, the dissociation kinetics for the different dsDNA substrates were well described by single exponential fits to the survival probability data (Fig. 3.4) [306]. Error bars for the survival probability measurements and binding distributions represent 70% confidence intervals obtained through bootstrap analysis, providing a close approximation of expectations for one standard deviation from the mean [306]. All reported $\Delta\Delta G^{\ddagger}$ values were calculated from the dissociation rate data for the Atto565–dsDNA substrates, as previously described [306], and the shared energetic signature (mean ± s.d.) reported in the main text for EcRecA, ScRad51, hRad51, and ScDmc1 reflects the combined average of the first, second and third triplet steps observed for all four recombinases using both sets of oligonucleotide substrates (Fig. 3.5 and 3.6).

3.5.3 Simulations

Simulations were performed using the Academic Center for Computing and Media Studies (ACCMS) Cray XE6 supercomputer at Kyoto University. Monte Carlo and molecular dynamics simulations were performed using a course–grained oxDNA model [335, 343]. The RS–DNA model was prepared by mapping the center of mass (COM) of the sugar atoms in each base in the crystal structure of RecA–dsDNA postsynaptic filament (PDB ID: 3CMX) [36] to the COM of each base in the coarse–grained representation of a DNA molecule 15 base pairs in length. We performed a short low temperature (0 °C) simulation with the COM of each base restrained in space to relax positions of base and backbone beads. The structure of B–DNA was prepared using generate–sa.py provided by the oxDNA software package.

In all of the simulations, one of the two DNA strands was constrained in space to mimic the protein–bound presynaptic ssDNA. We refer to this constrained strand as the presynaptic strand, which is colored red in Fig. 3.12 and 3.13. We refer to the other strand as the complementary strand, which is colored blue in Fig. 3.12 and **??**. For both the B– DNA and RS–DNA models the first three base pairs were also constrained in space by imposing a harmonic potential (Stiffness parameter k = 34.3 kJ mol^{-1-2}) on the COM of each of these three base pairs. This constraint was applied so that the initial structural models within the simulations represented a partially paired recombination intermediate. Two additional constraints were applied for the RS–DNA simulations. First, distances between adjacent base triplets the complementary strand were restrained by a harmonic potential. Second, the end–to–end distance of the complementary strand was also restrained by a harmonic poten-

tial (Stiffness parameter k = 3.43 kJ mol^{-1-2}). These constraints were intended to emulate the expectation that Rad51/RecA would need to locally stretch the incoming complementary strand to more closely match the extended configuration of the presynaptic ssDNA [36, 329, 336, 344–346]. Monte Carlo simulations (see below) confirmed that RS–DNA duplex formation was strongly disfavored in the absence of these constraints on the complementary strand (Fig. 3.12A). Using these initial structures and constraints, we performed virtual move Monte Carlo (MC) simulations to obtain canonical NVT ensemble at 27 °C [347], and this ensemble was used to elucidate the free energy profiles of duplex formation for RS–DNA (Fig. 3.12A). Umbrella sampling, where the number of base pairs with a negative hydrogen–bonding energy was used as an order parameter, was used to enhance conformational sampling [348]. Weight for each biasing potential was calibrated manually by iterating relatively short MC simulations (108 steps). Each simulation was then performed for 109 steps and the conformations were collected every 10^4 steps. The free energy profile was calculated from the canonical conformational ensemble using the equation, $-k_BTlog(P)$, where k_B is the Boltzmann constant, T is temperature, and P is the probability of each conformational state.

Molecular dynamics (MD) simulations were performed by integrating a Langevin equation using a diffusion coefficient of 12 ^{2} $picosecond^{-1}$ and an integral step size of 15 femtoseconds. The initial structure for the annealing simulations was prepared by performing a short (106 steps) high temperature (100 °C) simulation to melt the duplex. The RS–DNA and B–DNA annealing simulations were conducted at 27 °C, RS–DNA melting simulations were performed at 80 °C, and B–DNA melting simulations were performed at 100 °C. The temperature was kept constant using an Anderson–like algorithm. Each simulation was

performed for 107 steps, and conformations were collected every 104 steps.

The melting and annealing trajectories of B–DNA and RS–DNA were then analyzed to determine which bases were paired (Fig. 3.13B,E) as well as the duration (lifetime) of these pairing interactions (Fig. 3.12B and 3.13C,F). Within the simulations, a base was considered paired if the beads representing the two bases on complementary strands were within 8.5 of one another, and the bases were considered unpaired if this distance exceeded 8.5 . We define the "n bases paired state" as when the nth base pair forms (where n = 1 to 15) and the base pair does not form (where k = 1 to 15 – n); examples of these states are shown as snapshots in Figure 3.13A,D and graphical representation of the data are presented in 3.13B,D where the y–axes reflect the position of the most distal base pair relative to the pre–annealed triplet. We then calculated the lifetime (in simulation steps) of each paired state. For annealing simulations, we report the lifetimes of each bases paired state (Fig. 3.13C,F). For the melting simulations, we report the lifetimes as the average value for each nucleotide position within the base triplets (Fig. 3.12D). In both cases, lifetime data for the terminal base triplets are excluded from the reported values.

Chapter 4

Sequence Imperfections and Base Triplet Recognition by the Rad51/RecA Family of Recombinases

This chapter is adapted from work originally published as: "Sequence imperfections and base triplet recognition by the Rad51/RecA family of recombinases," Ja Yil Lee, **Justin B. Steinfeld**, Zhi Qi, YoungHo Kwon, Patrick Sung, and Eric C. Greene. J. Biol. Chem. (2017). I was responsible for some of the single molecule experiments along with experimental design and data analysis.

4.1 Abstract

Homologous recombination plays key roles in double-strand break repair, rescue, and repair of stalled replication forks and meiosis. The broadly conserved Rad51/RecA family of recombinases catalyzes the DNA strand invasion reaction that takes place during homologous recombination. We have established single-stranded (ss)DNA curtain assays for measuring individual base triplet steps during the early stages of strand invasion. Here, we examined how base triplet stepping by RecA, Rad51, and Dmc1 is affected by DNA sequence imperfections, such as single and multiple mismatches, abasic sites, and single nucleotide insertions. Our work reveals features of base triplet stepping that are conserved among these three

phylogenetic lineages of the Rad51/RecA family and also reveals lineage-specific behaviors reflecting properties that are unique to each recombinase. These findings suggest that Dmc1 is tolerant of single mismatches, multiple mismatches, and even abasic sites, whereas RecA and Rad51 are not. Interestingly, the presence of single nucleotide insertion abolishes recognition of an adjacent base triplet by all three recombinases. On the basis of these findings, we describe models for how sequence imperfections may affect base triplet recognition by Rad51/RecA family members, and we discuss how these models and our results may relate to the different biological roles of RecA, Rad51, and Dmc1.

4.2 Introduction

Homologous recombination allows for the regulated exchange of genetic information between two different DNA molecules of identical or nearly identical sequence composition, and it is a driving force in evolution [349, 350]. Homologous recombination contributes to double-strand DNA break (DSB)[1] repair [292], the rescue of stalled or collapsed replication forks [293, 327, 351], chromosomal rearrangements [1, 352, 353], horizontal gene transfer [354], and meiosis [141, 152, 355]. Defects in recombination compromise genome integrity and lead to the gross chromosomal rearrangements that are a hallmark of cancer [1, 352, 353]. During recombination, single-stranded DNA (ssDNA), derived from the nucleolytic processing of a DSB or collapsed replication fork, is paired with the complementary strand of a homologous double-stranded DNA (dsDNA), resulting in the displacement of the non-

[1]The abbreviations used are: DSB, double-strand DNA break; ssDNA, single- stranded DNA; nt, nucleotide; eGFP, enhanced GFP; ATPγS, adenosine 5-O-(thiotriphosphate); RS-DNA, recombinase stretched-DNA; RCR, rolling circle replication; MD, molecular dynamics.

complementary strand from the duplex to generate a D-loop [295, 356]. This reaction is referred to as strand invasion, and the resulting intermediate can be channeled through a number of alternative pathways, any of which can allow for the repair of the originally broken DNA molecule using information derived from the template [292, 293].

The homology search and strand invasion reactions are catalyzed by the Rad51/RecA family DNA recombinases [294, 295, 356]. These recombinases are among the most highly conserved of all DNA repair proteins [357], and prominent family members include bacterial RecA, the archaeal protein RadA, and the eukaryotic recombinase Rad51 [294, 358]. In addition to Rad51, most eukaryotes also have the meiosis-specific recombinase Dmc1, although the reason most eukaryotes require a second recombinase for meiosis remains an enduring mystery [152, 184].

RecA is the archetypal recombinase originally identified in genetic screens for Escherichia coli mutants defective in recombination [182], and much of our current understanding of recombination mechanisms can be attributed to studies of this recombinase [295, 359]. Rad51/RecA family members are all ATP-dependent DNA-binding proteins that form extended helical filaments on DNA [294, 295]. The bound DNA is extended by ∼50% relative to the contour length of B-form DNA [185, 192, 360–362], and crystal structures of RecA-ssDNA and RecA-dsDNA pre- and post-synaptic complexes reveal that the bound DNA is organized into near B-form base triplets separated by ∼ 8 between adjacent triplets (Fig. 4.1A) [36]. We have referred to this unique DNA architecture as recombinase stretched-DNA (RS-DNA) to distinguish it from other forms of mechanically stretched DNA and as a reflection of its unique structural and mechanistic properties [363]. To help better understand the mechanisms of genetic recombination, we have established single-molecule

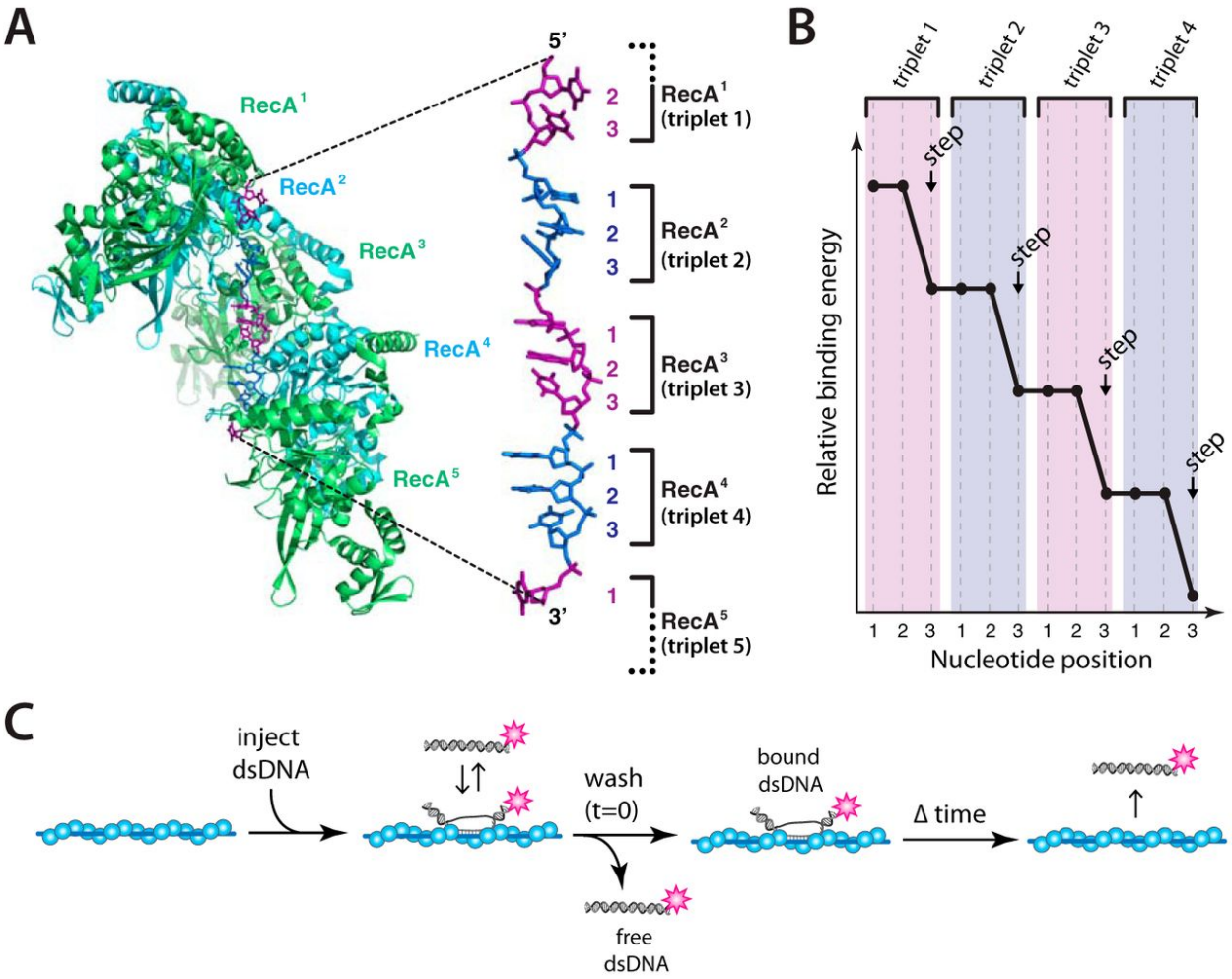

Figure 4.1: **Base triplet stepping by the Rad51/RecA family of recombinases.** (**A**) Crystal structure of the RecA-ssDNA pre-synaptic complex highlighting the base triplet organization of the bound pre-synaptic ssDNA. (**B**) Schematic illustration of base triplet stepping, in which each step coincides with complete pairing of all three bases within an RS-DNA triplet. (**C**) Experimental assay for quantitating the binding of Atto565-labeled dsDNA fragments to an unlabeled pre-synaptic complex. These panels were adapted with permission from Lee, J.Y. et al. (2015) DNA recombination. Base triplet stepping by the Rad51/RecA family of recombinases. Science 349, 977–981.

assays for studying how pre-synaptic complexes engage dsDNA during the early stages of homologous recombination [306, 363, 364]. These experiments reveal that strand invasion products are stabilized in 3-nt increments (Fig. 4.1B), with each 3-nt step exhibiting a characteristic energetic signature that is conserved among the Rad51/RecA family members [306, 363, 364].

Here, we examine how sequence imperfections affect individual base triplet steps by *E. coli* RecA, *Saccharomyces cerevisiae* Rad51, and *S. cerevisiae* Dmc1. We show that RecA, Rad51, and Dmc1 all require perfect Watson-Crick pairing interactions to stabilize base triplet pairing interactions at either the 5'or 3'termini of paired intermediates. Single nucleotide insertions completely disrupt base triplet recognition by all three recombinases, underscoring the importance of maintaining proper sequence register during strand exchange. RecA and Rad51 can step over base triplets harboring single mismatches or abasic sites, and Rad51 can step over tandem mismatches, but neither of these recombinases can stabilize the imperfectly paired triplets. In contrast, the meiosis-specific recombinase Dmc1 can step over and stabilize base triplets harboring single mismatches, multiple mismatches, and even abasic sites. Our results are consistent with a model where Dmc1 makes compensatory protein-DNA contacts with lesion-bearing or otherwise mismatched base triplets allowing stabilization of the resulting intermediates.

4.3 Results

4.3.1 Assay for base triplet recognition with ssDNA curtains

We used ssDNA curtains and total internal reflection fluorescence microscopy to visualize RecA, Rad51, or Dmc1 pre-synaptic complexes. The ssDNA was generated using M13 as a template for rolling circle replication and then tethered to a lipid bilayer through a biotin-streptavidin linkage and aligned along chromium barriers, as described previously [81, 306, 363, 365]. The ssDNA unravels when incubated with RPA-eGFP, and the downstream

ends of the RPA-ssDNA are anchored to exposed chromium pedestals through non-specific surface adsorption. Addition of RecA, Rad51, or Dmc1, along with the required nucleotide cofactor, results in the displacement of the fluorescent RPA-eGFP from the ssDNA [306, 363]. Once assembled, unbound proteins are flushed away, and dsDNA-binding activity of the pre-synaptic complexes is probed by the addition of short (70-bp) dsDNA oligonucleotides bearing a single tract of sequence microhomology 8–15 nucleotides (nt) in length that is complementary to a unique location on the M13 ssDNA [306, 363]. Following a short incubation, the unbound dsDNA is flushed away, and the stability of the remaining dsDNA molecules is measured by survival probability analysis (Fig. 4.1C) [306, 363, 365].

4.3.2 Microhomology and base triplet recognition

We have previously reported assays in which the internal tract of microhomology was incrementally extended from 8- to 15-nt in the 5'→ 3'direction relative to the complementary strand within the fluorescent dsDNA [306, 363]. These assays, together with accompanying molecular dynamics simulations, revealed that strand invasion intermediates were stabilized in 3-nt increments [306, 363]. Here, we sought to determine whether similar results would be obtained if we instead extended the tract of microhomology in the 3'→ 5'direction (Fig. 4.2A). To address this question, we used a series of 70-bp oligonucleotides labeled at one end with a single Atto565 dye and bearing an internal tract of microhomology targeted to a specific region on the pre-synaptic ssDNA. The internal tracts of microhomology were flanked by non-homologous sequences that lacked any microhomology exceeding 7-nt in length (Table 4.1 and 4.2). *E. coli* RecA, *S. cerevisiae* Rad51, or *S. cerevisiae* Dmc1 pre-synaptic complexes

were assembled onto the ssDNA curtains. The fluorescent dsDNA substrates were then injected into the sample chamber and incubated for 10 min at 30 °C, and the unbound DNA was then flushed away while collecting data. Dissociation rates were then obtained from survival probability analysis of the bound dsDNA fragments (Fig. 4.3) [306, 363]. These experiments revealed that extension of the internal tract of microhomology in the 3'→ 5'direction yielded changes in dissociation rates for RecA, Rad51, and Dmc1, similar to our previously reported results for 5'→ 3'extension (Fig. 4.2B-D) [306, 363]. In addition, each step coincided with a 30% change in relative dissociation rates, corresponding to change in free energy ($\Delta\Delta G^{\ddagger}$) of $\sim 0.3 k_B T$ (Fig. 4.2B-D). From these results, we concluded that 3'→ 5'extension of an internal tract of microhomology gave rise to periodic changes in dsDNA dissociation rates for RecA, Rad51, and Dmc1. These findings are consistent with our previous experimental observations using substrates in which we incrementally extended the microhomology length in the 5'→ 3', as anticipated [306, 363]. We were careful to note that although we obtained similar results for these two sets of substrates, the findings themselves do directly report upon the polarity of strand exchange because we are only assessing the stability of the reaction intermediates after they have already formed.

4.3.3 Mismatch discrimination at the 5 terminus of microhomology

We next asked whether RecA, Rad51, and Dmc1 could discriminate against mismatched bases located within the 5'-terminal triplet of an embedded tract of microhomology. For these assays, a single mismatch was introduced at each of three possible positions at the

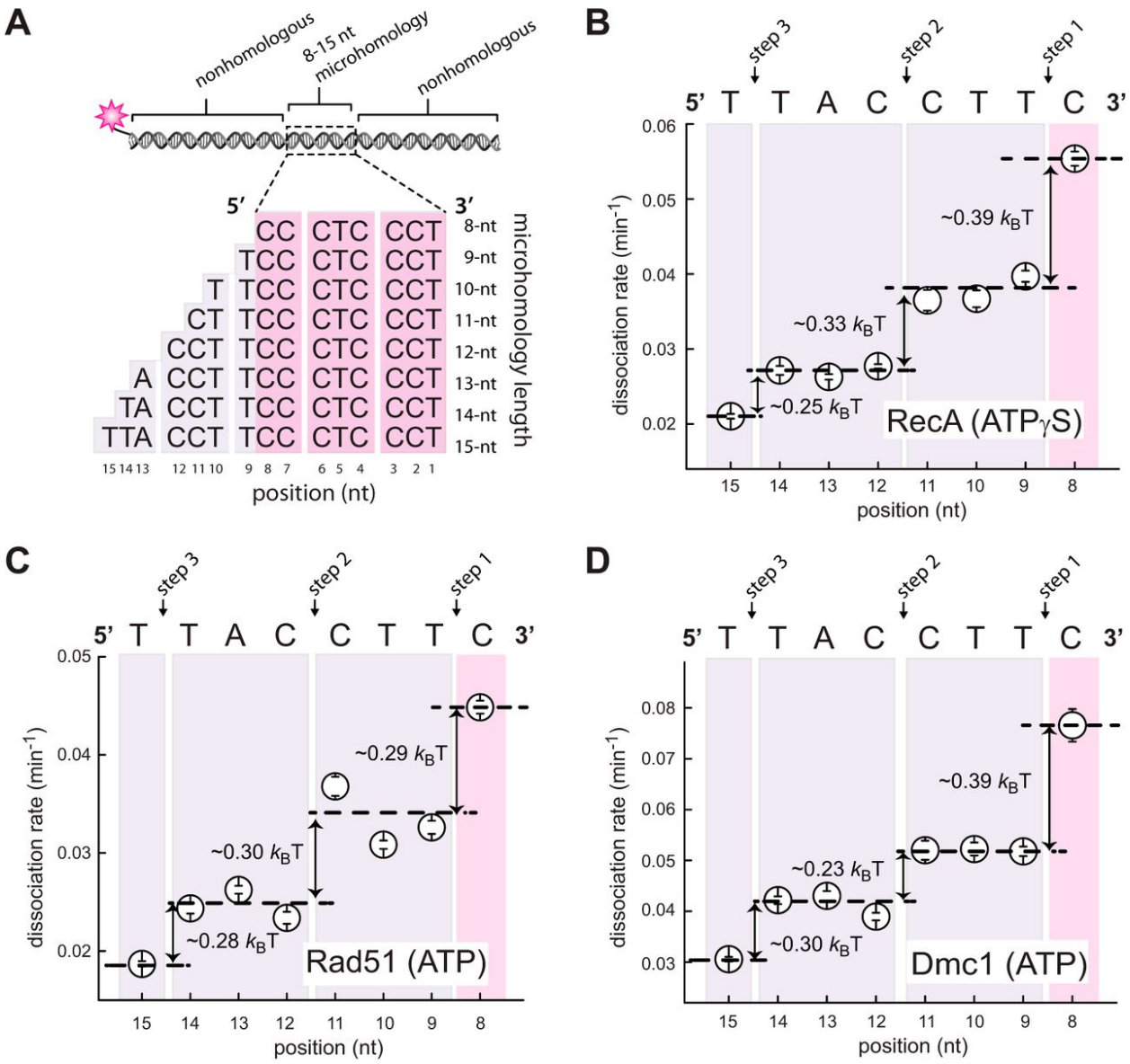

Figure 4.2: **Base triplet stepping observed with incremental extension of the 5'-microhomology.** (**A**) Schematic illustration of the 70-bp dsDNA substrates with incrementally lengthened 5'-microhomology. (**B**) Atto565-dsDNA dissociation rates (mean ± S.D.) for RecA in the presence of ATP S; (**C**) *S. cerevisiae* Rad51 plus ATP; and (**D**) *S. cerevisiae* Dmc1 plus ATP. All data points were calculated from an average of 200 molecules (n = 150 –250); the color-coded shading highlights each base triplet; magenta coloring indicates the minimum 8 nt necessary for stable binding, and purple shading corresponds to addition of homologous nucleotides; arrows indicate stepwise reductions in dissociation rates coincident with the third base of each triplet; dashed lines report the mean rate for each step; and the free energy ($\Delta\Delta G^{\ddagger}$) calculated for each triplet step are indicated.

	Designation	Sequence†
1	15-nt "wild-type"	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GCC CGT*** GGC TAG CTA GCT AGC TAG CTA GCT*–3'*
2	Double mismatch #1	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT CAC CGT*** GGC TAG CTA GCT AGC TAG CTA GCT*–3'*
3	Double mismatch #2	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT ACA CGT*** GGC TAG CTA GCT AGC TAG CTA GCT*–3'*
4	Triple mismatch	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT AAA CGT*** GGC TAG CTA GCT AGC TAG CTA GCT*–3'*
5	Single-base insertion	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GCC G CGT*** GGC TAG CTA GCT AGC TAG CTA GC*-3'*
6	Abasic #1	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT BCC CGT*** GGC TAG CTA GCT AGC TAG CTA GC*-3'*
7	Abasic #2	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GBC CGT*** GGC TAG CTA GCT AGC TAG CTA GC*-3'*
8	Abasic #3	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GCB CGT*** GGC TAG CTA GCT AGC TAG CTA GC*-3'*

Table 4.1: **Oligonucleotide sequences.** Lesions near the 3'end of the microhomology.

	Microhomology length	Sequence†
9	8-nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG TTA CCT G***CC CTC CCT*** AGC TAG CTA GCT AGC TAG CTA GCT*-3'*
10	9-nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG TTA CCA ***TCC CTC CCT*** AGC TAG CTA GCT AGC TAG CTA GCT*-3'*
11	10-nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG TTA CA***T TCC CTC CCT*** AGC TAG CTA GCT AGC TAG CTA GCT*-3'*
12	11-nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG TTA A***CT TCC CTC CCT*** AGC TAG CTA GCT AGC TAG CTA GCT*-3'*
13	12-nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG TAT ***CCT TCC CTC CCT*** AGC TAG CTA GCT AGC TAG CTA GCT*-3'*
14	13-nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG AC***A CCT TCC CTC CCT*** AGC TAG CTA GCT AGC TAG CTA GCT*-3'*
15	14-nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG G***TA CCT TCC CTC CCT*** AGC TAG CTA GCT AGC TAG CTA GCT*-3'*
16	15-nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CTC CCT*** AGC TAG CTA GCT AGC TAG CTA GCT*-3'*

Table 4.2: **Oligonucleotide sequences (cont.).** 5'extension of the microhomology.

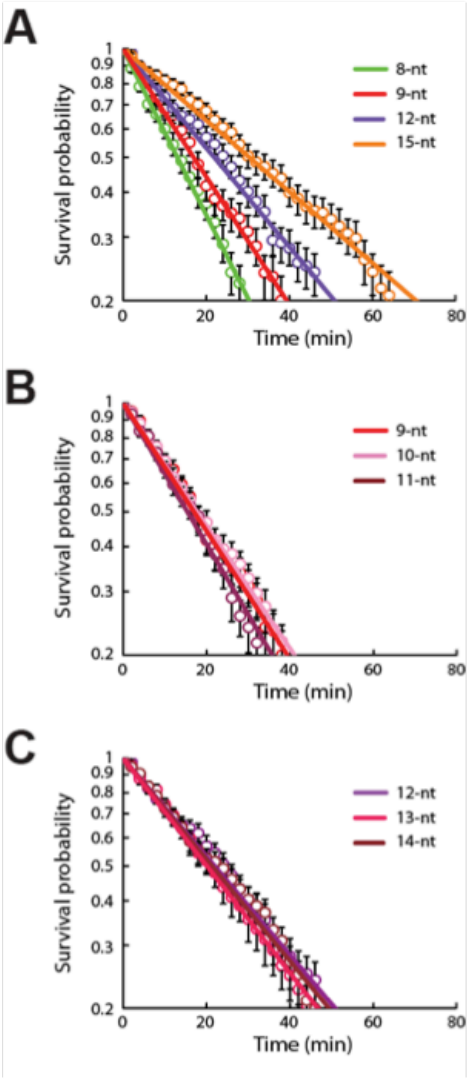

Figure 4.3: **Survival probability plots.** (**A**) Examples of survival probability data for 70-bp oligonucleotide substrates with 8-, 9-, 12-, or 15-nt of microhomology, as indicated. (**B**) Survival probability data for the 9-, 10-, and 11-nt substrates and (**C**) 12-, 13-, and 14-nt substrates are presented separately, for clarity. All plots are for reactions with S. cerevisiae Rad51 in the presence of ATP. Error bars represent 70% confidence intervals obtained through bootstrap analysis, providing a close approximation to expectations for one standard deviation from the mean, as previously described from Qi, Z. et al. (2015) DNA sequence alignment by microhomology sampling during homologous recombination. Cell 160, 856-869. Lee, J.Y. et al. (2015) DNA recombination. Base triplet stepping by the Rad51/RecA family of recombinases. Science 349, 977–981.

5'end of a 12-nt tract of microhomology (Fig. 4.4A). These results revealed that a single mismatch at any position within the 5'-terminal triplet abolishes detectable binding of the mismatched triplet by the RecA, Rad51, and Dmc1 (Fig. 4.4B-D). These findings are similar to our previous observations for mismatches located at the 3'end of the internal tract of microhomology [363]. Taken together, our results support the conclusion that RecA, Rad51, and Dmc1 are all intolerant of single nucleotide mismatches within terminal base triplets located at either the 5'or 3'end of an internal tract of sequence homology.

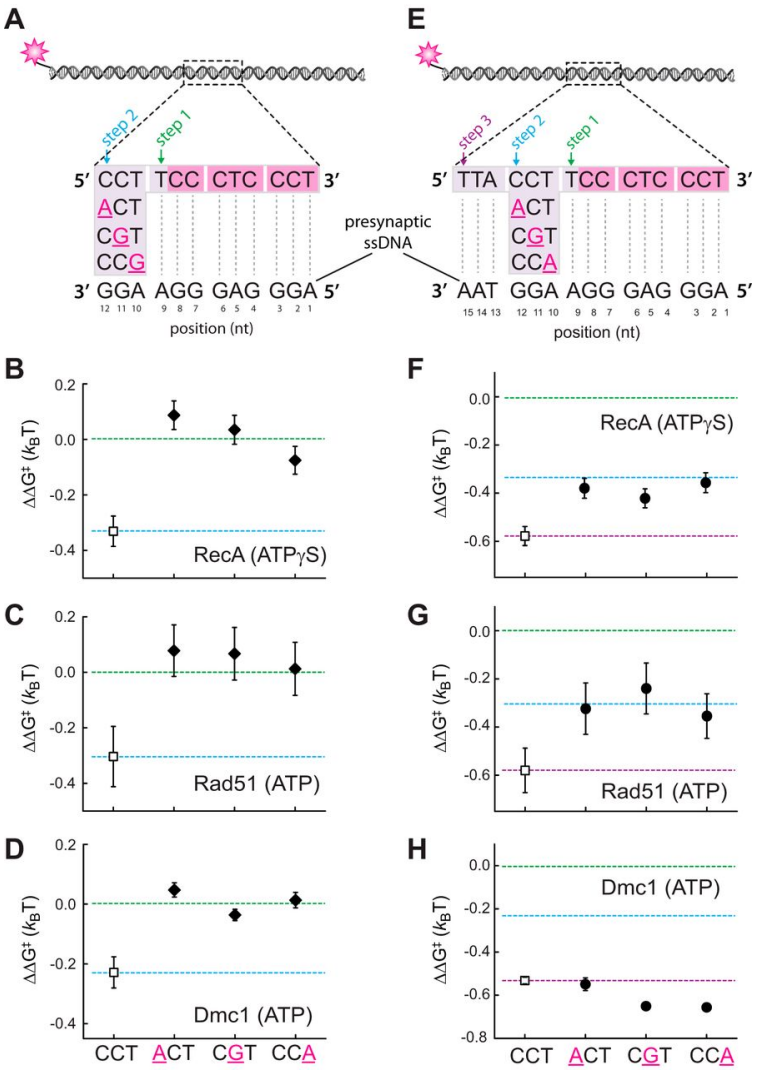

Figure 4.4: **Effect of single nucleotide mismatches on 5'-base triplet recognition.** (**A**) Schematic illustration of dsDNA substrates bearing single nucleotide mismatches within the terminal triplet at the 5'end of a 12-nt tract of microhomology. Each mutated nucleotide is highlighted in magenta and underlined, and the corresponding sequence of the pre-synaptic ssDNA is indicated. Mismatch substrate-binding data for RecA (**B**), Rad51 (**C**), and Dmc1 (**D**) are as indicated. All indicated ($\Delta\Delta G^{\ddagger}$) values are relative to the step 1 binding data for the substrate bearing 9 nt of microhomology (see Fig. 4.2). The green and blue dashed lines correspond to the average binding stability for data obtained from the non-mismatched 9 –11-nt substrates (reflecting stable association equivalent to three complete triplets) and the 12–14-nt substrates (reflecting stable association equivalent to four complete triplets), respectively. (**E**) Schematic illustration of dsDNA substrates bearing single nucleotide mismatches within the penultimate triplet at the 5 end of a 15-nt tract of microhomology. Corresponding binding data for RecA (**F**), Rad51 (**G**), and Dmc1 (**H**)are as indicated. Green, blue, and purple dashed lines correspond to the averaged binding stability for data obtained from the non-mismatched 9 –11-nt substrates (reflecting stable association equivalent to three complete triplets); the 12–14-nt substrates (reflecting stable association equivalent to four complete triplets); and 15-nt substrates (reflecting stable association of 5 complete triplets), respectively, reflecting steps 1–3 in A.

4.3.4 Internal triplets are destabilized within RecA and Rad51 complexes

RecA and Rad51 can step over mismatches located within the penultimate base triplet near the 3'end of an internal tract of microhomology, but the reduced stability of the resulting intermediates suggests that the mismatched triplet may remain unpaired [363]. Dmc1 can also step over internal triplets near the 3'end of an internal tract of microhomology, but these mismatched triplets still contribute to binding free energy of the resulting intermediate [363]. As an extension of these previous findings, we next asked whether RecA, Rad51, and Dmc1 could step over mismatches positioned within the penultimate base triplet near the 5'end of an internal tract of microhomology (Fig. 4.4E). Here, a single mismatch was introduced at each of three possible positions within the penultimate base triplet within a 15-nt tract of microhomology (Fig. 4.4E). For both RecA and Rad51, the binding stability of the resulting intermediate was most consistent with destabilization of the mismatch-bearing base triplet (Fig. 4.4F,G). In contrast, the stability of the mismatched intermediates generated by Dmc1 was comparable with the fully paired intermediate (Fig. 4.4H). We conclude that Dmc1 can stabilize single mismatches located in the penultimate base triplet near either the 5'or 3'end of an embedded tract of microhomology. In contrast, RecA and Rad51 were unable to stabilize these mismatched triplets, but both were able to step over the mismatch-bearing triplets to stabilize an adjacent base triplet that is fully homologous to the pre-synaptic ssDNA.

4.3.5 Dmc1 can stabilize triplets bearing multiple mismatches

To explore the potential mechanism of mismatch stabilization by Dmc1 we next sought to determine whether Dmc1 could stabilize base triplets bearing more than one mismatch. For these assays, we utilized a 70-bp dsDNA substrate bearing a 15-nt tract of microhomology in which the penultimate 3'-triplet harbored either two or three nucleotide mismatches with the pre-synaptic ssDNA (Fig. 4.5A and 4.6). Consistent with the observation that RecA and Rad51 are intolerant of even a single mismatch, neither protein was capable of stabilizing substrates bearing either two or three mismatches (Fig. 4.5B,C and 4.6). In contrast, Dmc1 was able to stabilize each mismatched substrate such that the relative binding energy of the mismatch-bearing substrates was comparable with the fully paired substrate (Fig. 4.5B).

Interestingly, RecA and Rad51 can both step over triplets bearing single mismatches, but only Rad51 appears to be capable of stepping over triplets bearing either two or three mismatches (Fig. 4.5B,C and Fig. 4.6). Based upon the observed changes in binding free energy, RecA appears to be incapable of stabilizing a homologous triplet that lies beyond a triplet bearing either two or three mismatches in reactions using ATPγS. This differential response to substrates harboring multiple mismatches is the only substantive difference we have found between RecA and Rad51 in response to base triplet imperfections. Bulk biochemical assays have revealed that RecA is unable to bypass longer ($\geq$50 bp) heterologous insertions during strand invasion when ATPγS is used as the nucleotide cofactor, but it can bypass these insertions in reactions with ATP [366, 367]. Therefore, we sought to determine whether RecA might be able to stabilize the substrates with multiple mismatches in the presence of ATP. These experiments revealed that in the presence of ATP, RecA is capa-

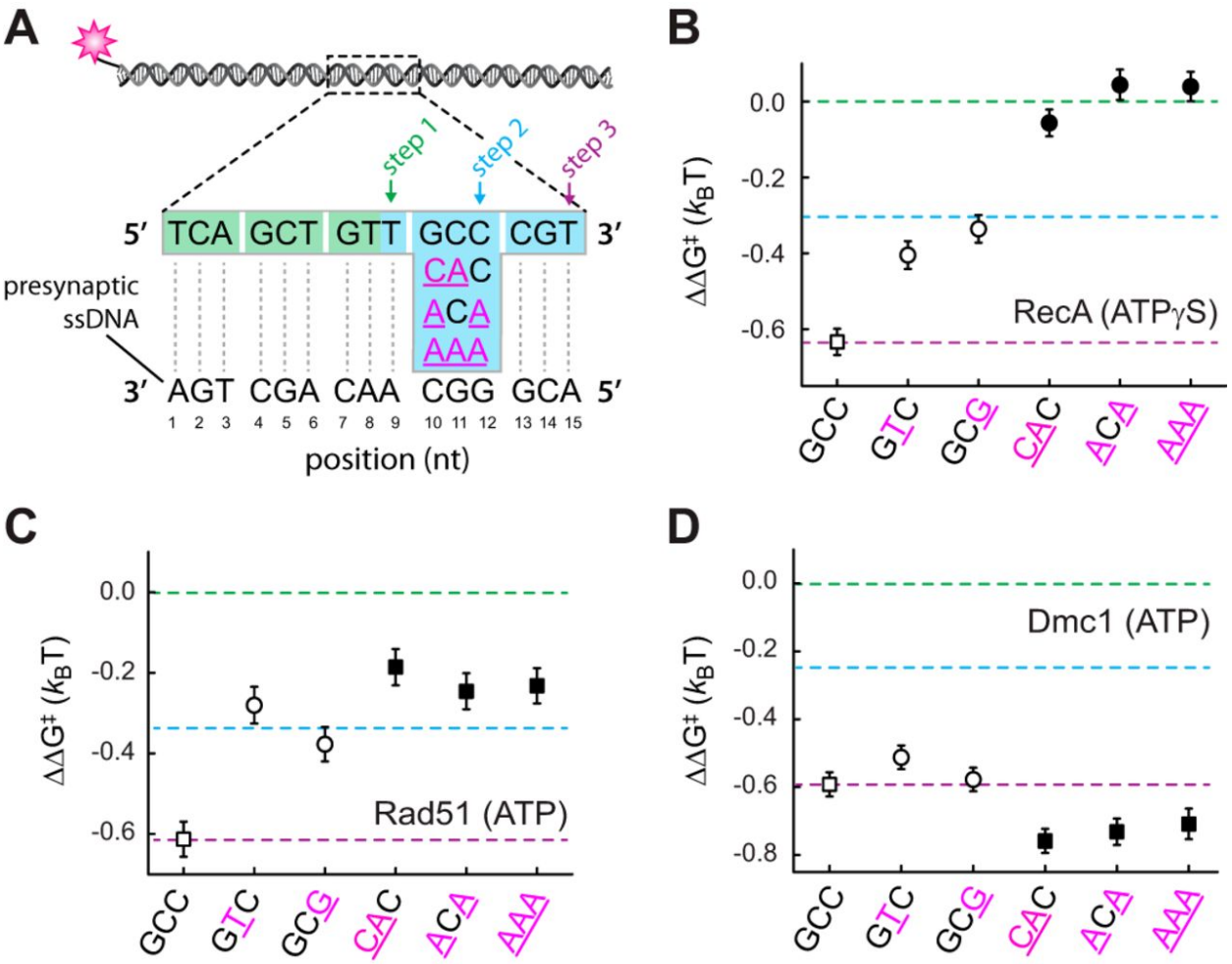

Figure 4.5: **Impact of multiple mismatches on base triplet recognition.** (**A**) Schematic illustration of dsDNA substrates bearing either two or three nucleotide mismatches within the penultimate triplet at the 3'end of a 15-nt tract of microhomology. Corresponding binding data for RecA (**B**), Rad51 (**C**), and Dmc1 (**D**) are as indicated. Green, blue, and purple dashed lines correspond to the averaged binding stability for data obtained from the non-mismatched 9–11-nt substrates; the 12–14-nt substrates; and 15-nt substrates, respectively, reflecting steps 1–3 in A. Previously reported data for substrates bearing single mismatches are shown (open circles) for comparison see Lee, J.Y. et al. (2015) DNA recombination. Base triplet stepping by the Rad51/RecA family of recombinases. Science 349, 977–981.

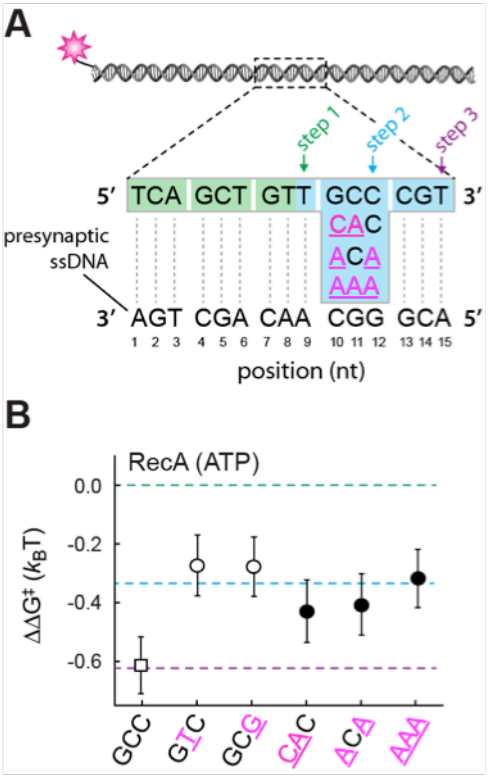

Figure 4.6: **RecA plus ATP with substrates bearing multiple mismatches.** (**A**) Schematic illustration of the substrates bearing multiple mismatches in the penultimate triplet near the 3'end of the 15-nt tract of microhomology. The position and identity of the mismatches are highlighted in magenta underlined text. (**B**) Corresponding data for each of the different DNA substrates. Data for the complementary triplet is shown as a reference (open square), previously reported data for substrates bearing single mismatches are also shown (open circles) for comparison see Lee, J.Y. et al. (2015) DNA recombination. Base triplet stepping by the Rad51/RecA family of recombinases. Science 349, 977–981, along with the data for the multiple mismatch substrates (closed circles).

ble of stabilizing a homologous triplet that lies beyond an internal triplet bearing multiple mismatches (Fig. 4.7). We conclude that in the presence of ATP, RecA and Rad51 respond similarly to internal base triplets bearing multiple mismatches.

4.3.6 Dmc1 can stabilize triplets bearing abasic sites

We next tested substrates bearing a single abasic site at each of the three possible positions within the penultimate base triplet near the 3'or 5'end of the microhomology (Fig. 4.8). Neither RecA nor Rad51 was able to stabilize internal triplets bearing an abasic site, consistent with their intolerance for single base mismatches (Fig. 4.8B,C,F,G). In contrast, Dmc1 was able to stabilize these substrates regardless of the location of the abasic site within the triplet, and similar findings were obtained for abasic sites located near the

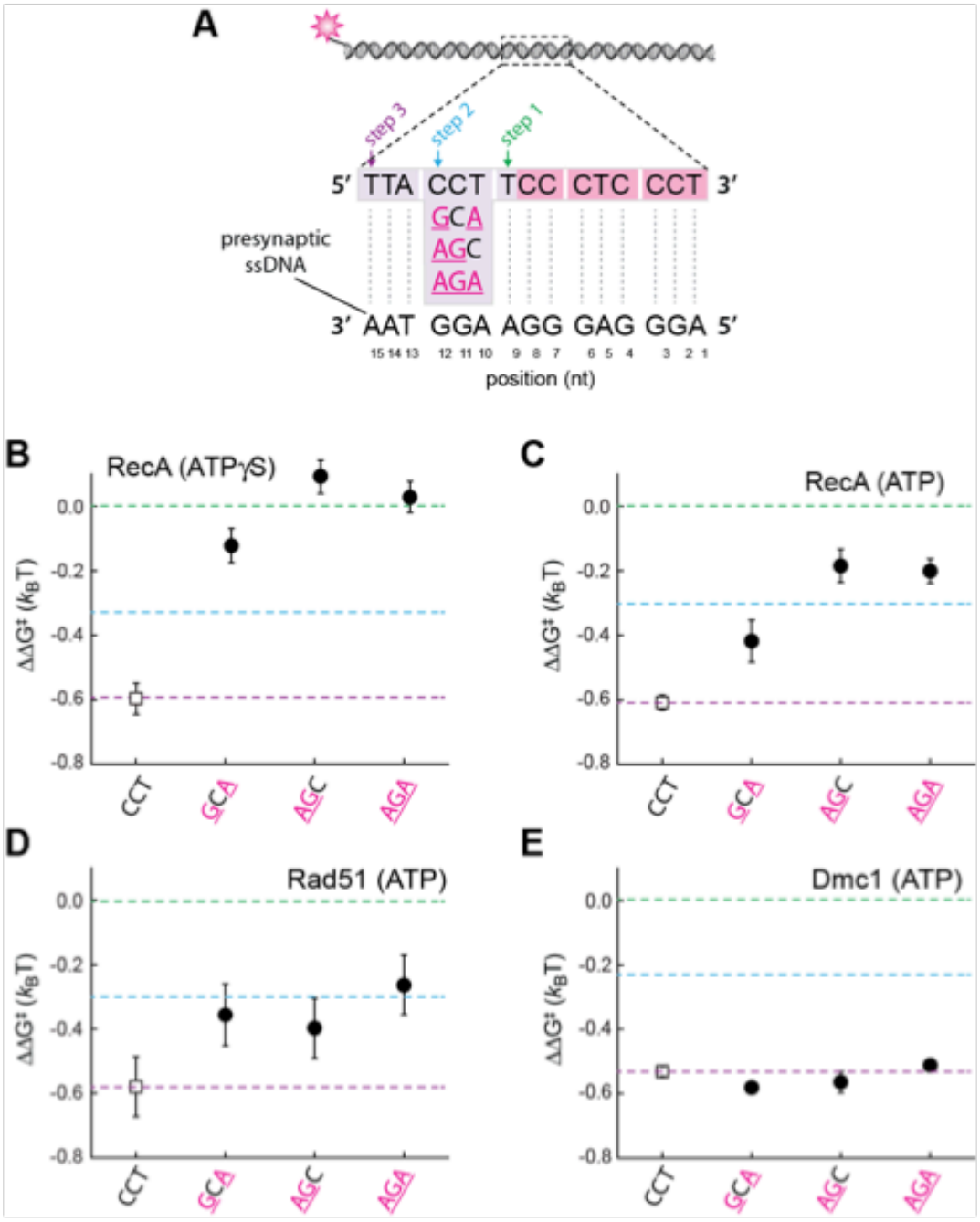

Figure 4.7: **Reactions with multiple mismatches near the 5'end of the microhomology.** (**A**) Schematic illustration of the substrates bearing multiple mismatches in the penultimate triplet near the 5'end of the 15-nt tract of microhomology. The position and identity of the mismatches are highlighted in magenta underlined text. Corresponding data for each of the different DNA substrates with (**B**) RecA and ATPγS, (**C**) RecA and ATP, (**D**) *S. cerevisiae* Rad51 and ATP, and (**E**) *S. cerevisiae* Dmc1 and ATP. Data for the complementary triplet is shown as a reference (open square), along with the data for the multiple mismatch substrates (closed circles).

3'or 5'ends of the internal tract of microhomology (Fig. 4.8D,H). These findings provide further support for the hypothesis that Dmc1 does not stabilize mismatched intermediates through a mechanism involving protein-enhanced pairing of non-Watson-Crick interactions, because abasic sites cannot form any type of non-canonical pairing interaction. The finding that Dmc1 can stabilize substrates bearing abasic sites also suggests that the stabilization mechanism likely does not involve direct protein contacts with the nucleotide bases of the suboptimal triplet, but instead it may involve contacts with the ribose-phosphate backbone of the DNA.

4.3.7 Non-bridging oxygen modifications do not prevent mismatch stabilization by Dmc1

We next tested whether Dmc1-mediated mismatch stabilization might involve recombinase contacts with the phosphate backbone. To address this possibility, we conducted assays with Dmc1 using mismatched substrates harboring either methylphosphonate or phosphorothioate substitutions within the phosphate backbone of the mismatched triplet (Fig. 4.9). Surprisingly, the methylphosphonate substitutions had no appreciable impact upon the ability of Dmc1 to stabilize the mismatched base triplets. These findings suggest that if Dmc1 contacts the phosphate backbone of incoming complementary strand during mismatch stabilization, then these protein-DNA contacts may not involve the non-bridging oxygen atoms, but it would more likely involve other contacts with the ribose-phosphate backbone. Future work will be essential to determine the structural basis for mismatch stabilization by Dmc1.

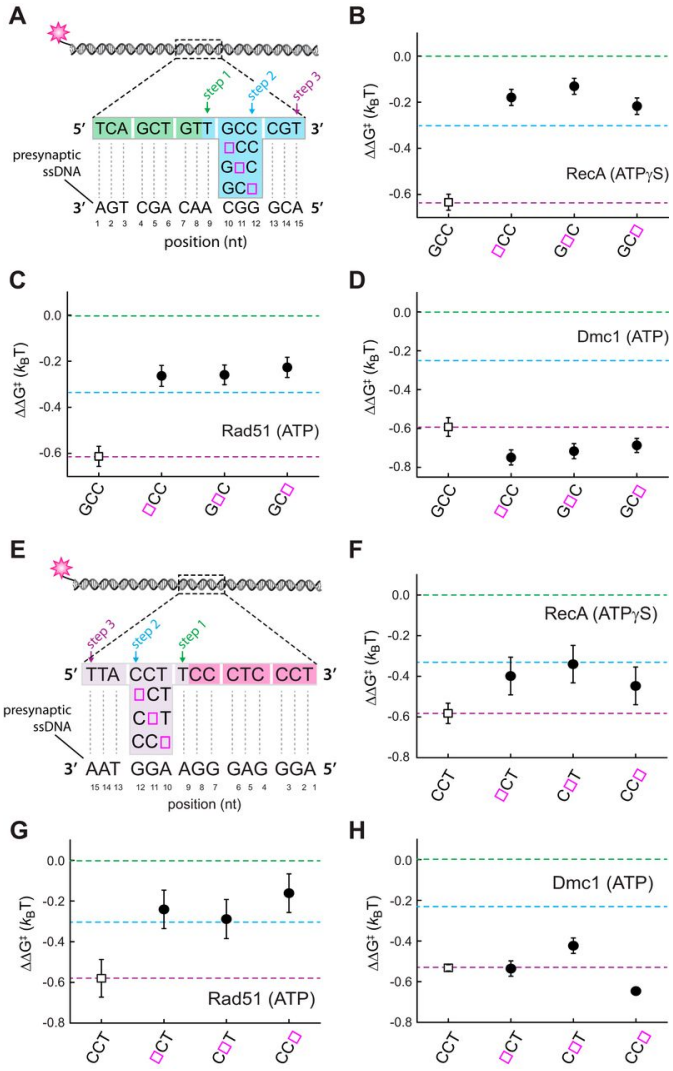

Figure 4.8: **Dmc1 can stabilize triplets bearing abasic sites.** (**A**) Schematic illustration of dsDNA substrates bearing single abasic sites within the penultimate triplet at the 3'end of a 15-nt tract of microhomology. Open magenta boxes signify the locations of the abasic sites. Corresponding to binding data for RecA (**B**), Rad51 (**C**), and Dmc1 (**D**) are as indicated. Green, blue, and purple dashed lines correspond to the averaged binding stability for data obtained from the non-mismatched 9-nt substrates; the 12–14-nt substrates; and 15-nt substrates, respectively, reflecting steps 1–3 in A. (**E**) Schematic illustration of dsDNA substrates bearing single abasic sites within the penultimate triplet at the 3'end of a 15-nt tract of microhomology. Open magenta boxes signify the locations of the abasic sites. Corresponding binding data for RecA (**F**), Rad51 (**G**), and Dmc1 (**H**) are as indicated. Green, blue, and purple dashed lines correspond the averaged binding stability for data obtained from the non-mismatched 9 –11-nt substrates; the 12–14-nt substrates; and 15-nt substrates, respectively, reflecting steps 1–3 in A.

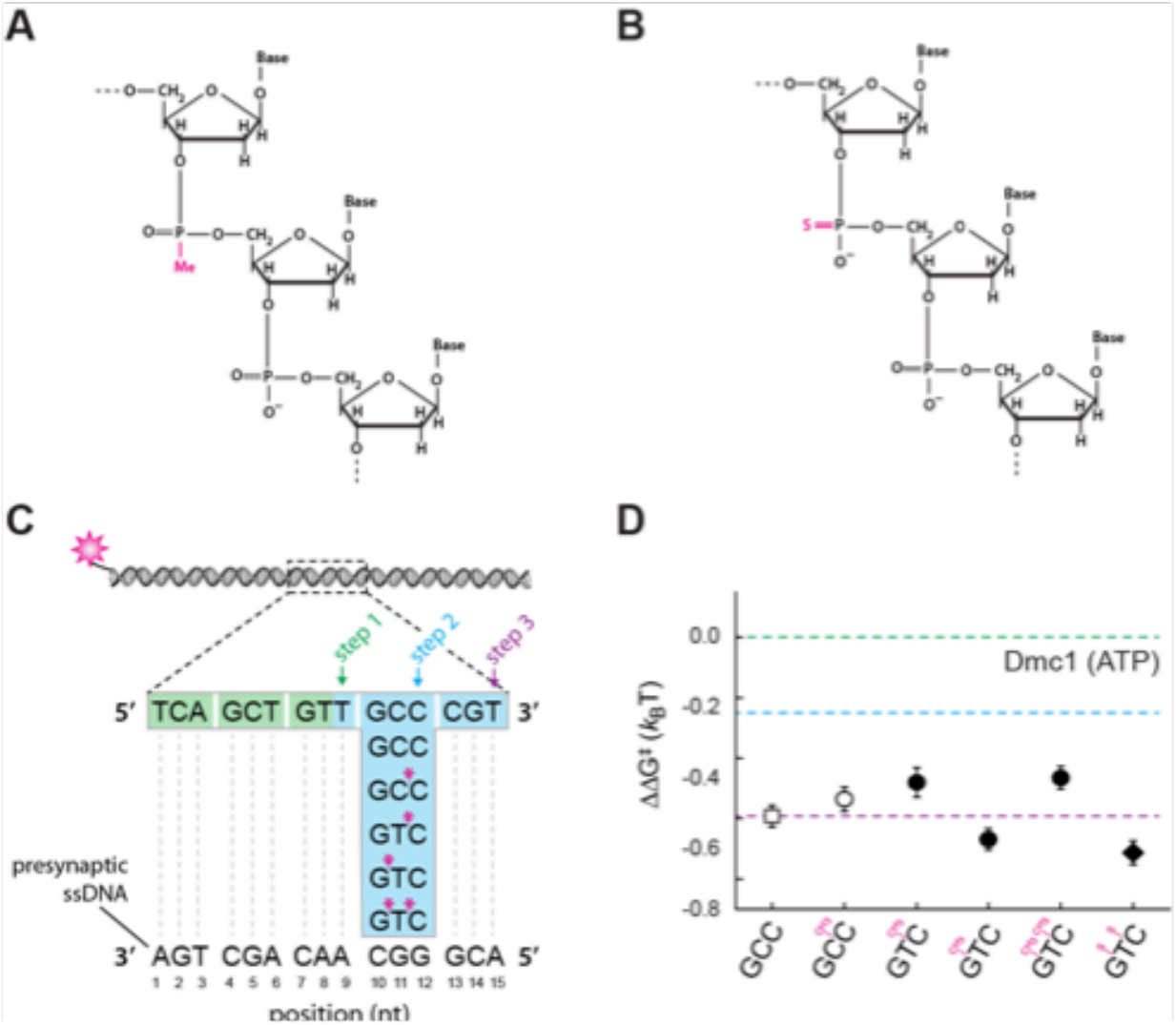

Figure 4.9: **Substitutions of the non-bridging oxygens do not affect mismatch stabilization by Dmc1.** Illustrations of (**A**) methylphosphonate and (**B**) phosphothioate backbone substitutions. (**C**) Schematic of dsDNA substrates bearing either methylphosphonate or phosphorothionate substitutions within the penultimate triplet at the 3'end of a 15-nt tract of microhomology. The locations of the backbone modifications are indicated with a magenta asterisk. (**D**) Binding data for the methyl phosphonate and phosphorothionate substituted substrates (as indicated) in reactions with Dmc1.

4.3.8 Single nucleotide insertions disrupt triplet stabilization

We next sought to establish whether RecA, Rad51, or Dmc1 could tolerate single nucleotide insertions. For this purpose, we designed a dsDNA substrate harboring a single base insertion between the penultimate and terminal triplet at either the 3'or 5'end of an internal tract of microhomology, which is anticipated to position the terminal base triplets just 1-nucleotide out of register with the pre-synaptic ssDNA (Fig. 4.10A,C). Surprisingly, the results revealed that a single base insertion results in a change in binding free energy consistent with the loss of stabilizing interactions involving the terminal base triplets for each of the three recombinases (Fig. 4.10B,D). We conclude that RecA, Rad51, and Dmc1 cannot stabilize a homologous 3'- or 5'-terminal base triplet that is moved out of register by the presence of a single nucleotide insertion.

We have established that the eukaryotic meiosis-specific recombinase Dmc1 is more tolerant of imperfect base triplet-pairing interactions relative to Rad51 and RecA. As discussed below, our findings suggest a model in which Dmc1 stabilizes imperfectly paired triplets through a mechanism involving protein-mediated contacts that compensate for the binding energy that would otherwise be lost due to the unpaired triplet.

4.3.9 Strand exchange takes place in 3-nt steps

Our work suggests that incrementally increasing the length of homology in either the 5'$\rightarrow\rightarrow$3'or 3'$\rightarrow$5'around a central 8-nt tract of microhomology gives rise to stepwise changes in binding free energy that follow a highly predictable base triplet pattern. In addition, the free energy changes associated with base triplet stabilization appear to be indistinguishable

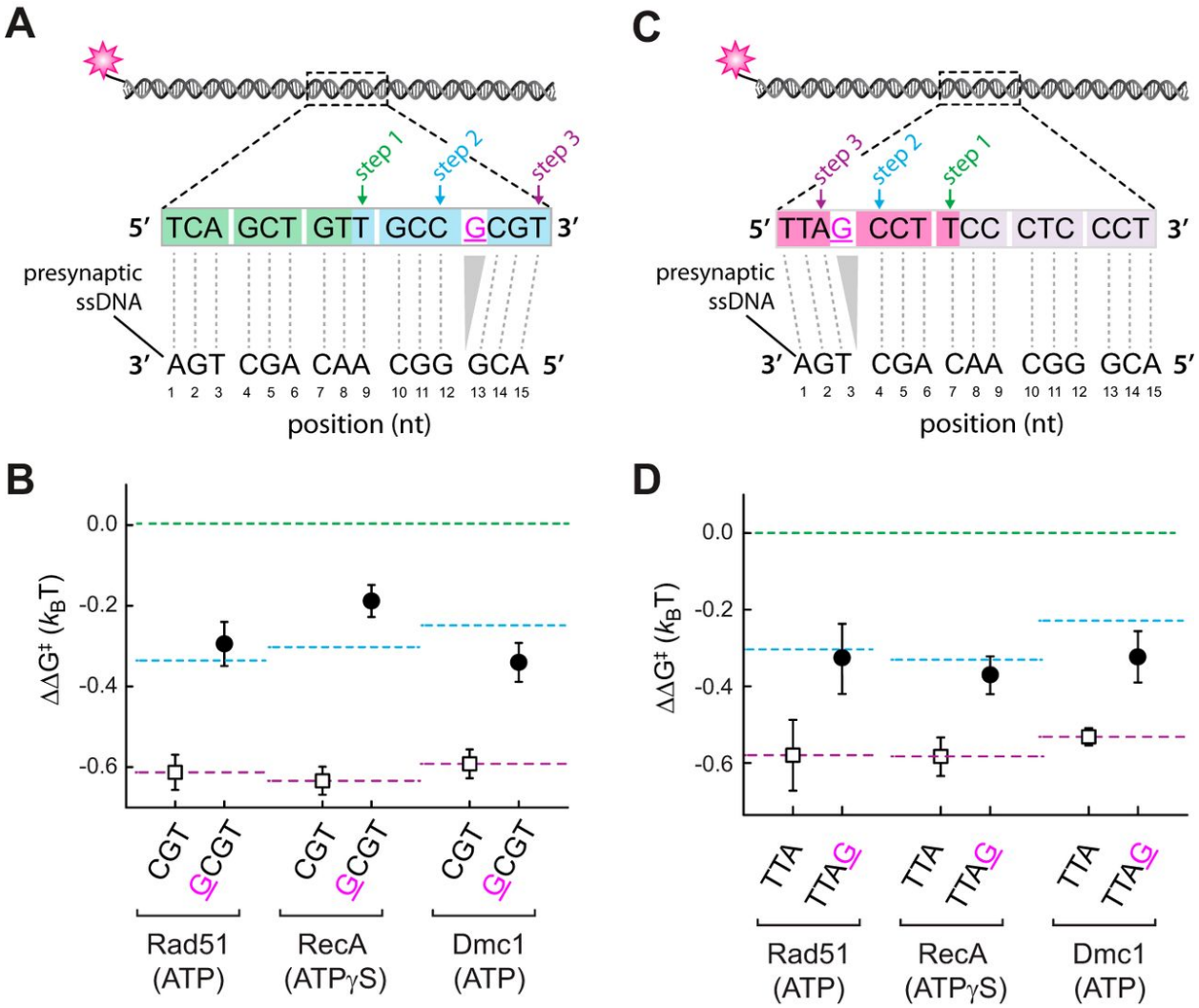

Figure 4.10: **Single base insertion can disrupt triplet stabilization.** (**A**) Schematic illustration of dsDNA substrates bearing a single nucleotide insertion adjacent to the penultimate triplet at the 3'end of a 15-nt tract of microhomology. (**B**) Corresponding binding data for Rad51, RecA, and Dmc1 for dsDNA substrates with (closed circles) and without (open squares) the nucleotide insertion. (**C**) Schematic illustration of dsDNA substrates bearing a single nucleotide insertion adjacent to the penultimate triplet at the 5'end of a 15-nt tract of microhomology. (**D**) Corresponding binding data for Rad51, RecA, and Dmc1 for dsDNA substrates with (closed circles) and without (open squares) the nucleotide insertion.

for base triplets located at either end of the internal tract of microhomology. These observations, together with the finding that the recombinases can all step over mismatches or abasic sites positioned within the penultimate triplet at either the 5'or 3'ends of the internal tract of microhomology, support a model in which strand exchange might proceed in either the 5'→ 3'or 3'→ 5'directions, and they suggest that bidirectional strand exchange may be conserved across the Rad51/RecA family of recombinases (Fig. 4.11A). Several studies have suggested that strand exchange can take place with a defined polarity [368–371], whereas others have reported that strand exchange can occur in either direction [372, 373]. We note that our assays do not directly observe strand exchange steps as they are taking place, but rather assess the stability of these early recombination intermediates only after they have already formed. Therefore, we do not rule out the possibility that there may exist polarity-dependent differences in the rates of strand exchange that cannot be assessed in our experimental system.

4.3.10 Responses of RecA, Rad51, and Dmc1 to sequence imperfections

Single mismatches within the terminal base triplet at either the 5'or 3'end of the microhomology prevent stabilization of these terminal triplets by RecA, Rad51, and Dmc1, suggesting that this may be a universal response of Rad51/RecA family members to imperfectly paired terminal base triplets (Fig. 4.11A). This result is surprising for Dmc1 because this recombinase can stabilize all other tested imperfections (i.e.single, double, and triple mismatches, and abasic sites), with the exception of single nucleotide insertions, so long as

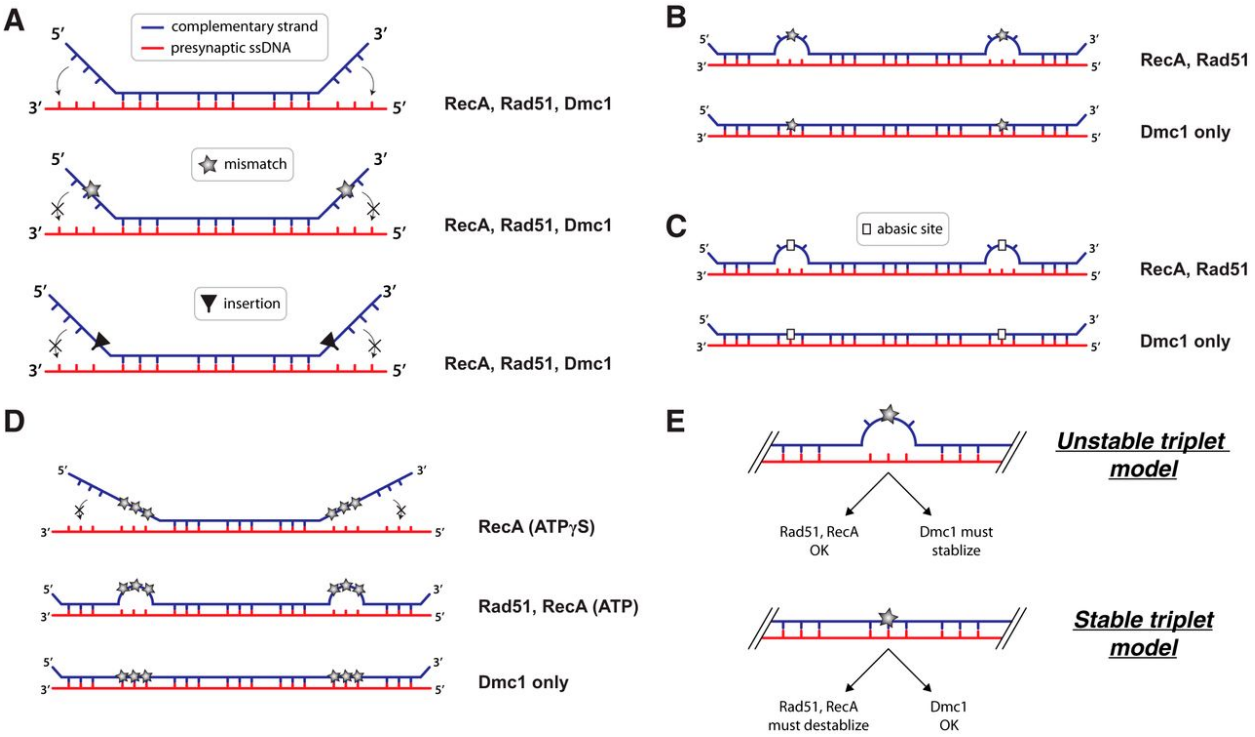

Figure 4.11: **Effects of lesions on base triplet recognition by RecA, Rad51, and Dmc1.** Details of these models are presented under "Discussion."

the imperfections are flanked by at least one fully homologous base triplet (Fig. 4.11B). These findings raise the question of whether the inability of Rad51/RecA recombinases to stabilize imperfect terminal triplets provides some specific biological advantage for homologous recombination or whether it might be a reflection of the thermodynamic properties of RS-DNA or both. One possibility is that imperfect terminal triplets play an important role in ensuring accurate recombination, perhaps by providing a kinetic barrier to slow the progression of stand invasion, and providing a decision point for the pre-synaptic complex (and associated factors) to either continue strand invasion or dissociate to sample other regions of the genome for a more perfectly matched donor sequence.

4.3.11 Triplet pairing requires perfect alignment

Single base insertions abolish recognition of the adjacent triplet by all three recombinases, suggesting that this may be a conserved feature of base triplet recognition by the Rad51/RecA family of DNA recombinases (Fig. 4.11A). This finding highlights intolerance of the pre-synaptic complex to what might have been considered a relatively minor structural defect. For instance, one might have assumed that a single base insertion would simply have been accommodated between the two homologous triplets, which would in principle eliminate the need to fully extend the phosphodiester linkage between these two triplets. Instead, the inserted base appears to abolish interactions with the downstream triplet, suggesting that amino acids within the pre-synaptic complex may be positioned to prevent accommodation of extra nucleotides. Indeed, inspection of the RecA crystal structure reveals that the L2 DNA-binding loop lies between adjacent base triplets, suggesting that steric hindrance by L2 may be responsible for preventing accommodation of base insertions [36]. Our experiments cannot yet address whether the presence of additional homologous base triplets beyond the nucleotide insertion would eventually allow stabilization, although we favor the idea that they would, given that homologous recombination can take place between longer nucleic acid substrates bearing either nucleotide insertions or deletions. Future work will be necessary to determine how much sequence homology is necessary on either side of a base insertion to allow triplet pairing interactions with both flanking sequences.

Interestingly, RecA and Rad51 can both step over triplets bearing single mismatches at either the 5'or 3'ends of the microhomology, but neither Rad51 nor RecA can stabilize the resulting mismatched triplets (Fig. 4.11B). This finding contrasts with results for Dmc1,

which can both step over and stabilize single mismatches at either the 5'or 3'ends of the microhomology (Fig. 4.11B). RecA and Rad51 can also step over abasic sites, but cannot stabilize the resulting triplets bearing the abasic residues. This finding is again in striking contrast with Dmc1, which can step over and stabilize base triplets bearing abasic sites regardless of the location of the abasic site within the base triplet (Fig. 4.7C). These results have important implications for understanding the mechanistic basis by which Dmc1 stabilizes substrates bearing imperfect triplets (see below).

4.3.12 Effects of multiple mismatches

Experiments looking at multiple mismatches yielded several important findings. First, RecA can step over triplets bearing a single mismatched nucleotide but not two or three mismatches in reactions with ATPγS (Fig 4.11D). This contrasts with the findings for Rad51, which is able to step over base triplets bearing two or even three mismatches (Fig. 4.11D). We speculate that in the presence of ATPγS, RecA might be capable of stepping over these tandem mismatches if there were additional sequence homology beyond the tandem mismatches, perhaps by initiating an independent strand invasion reaction downstream of the mismatches. Interestingly, RecA can step over multiple mismatches in reactions with ATP (Fig. 4.11D), which is consistent with prior biochemical assays demonstrating that RecA can bypass larger sequence heterologies only in the presence of ATP [366, 367]. This requirement for ATP has led to the hypothesis that RecA may have an ATP-dependent motor activity that is coupled to heterology bypass [374]. An alternative possibility is that ATP hydrolysis-dependent protein turnover may be necessary to reduce the stiffness of the

RecA filament, allowing a separate segment of the pre-synaptic complex to more readily re-engage the dsDNA substrate beyond tandem mismatches.

These experiments also revealed that Dmc1 can stabilize base triplets bearing either one, two, or three mismatches, so long as those triplets are flanked by at least one homologous triplet (Fig. 4.11D). This finding highlights the marked tolerance of Dmc1 to sequence imperfections at the level of a single base triplet, and it also has crucial implications for understanding how Dmc1 stabilizes imperfectly matched RS-DNA triplets (see below).

4.3.13 Models for base triplet stability

Our results can be considered within the context of two contrasting models describing the physical stability of mismatched RS-DNA base triplets. The first model is that imperfectly matched RS-DNA base triplets are themselves inherently unstable (Fig. 4.11E). This unstable triplet model would account for all of our observations with RecA and Rad51, indicating that the inability of these two proteins to stabilize imperfect triplets stems from the lesion-bearing triplets themselves being unstable relative to a perfectly paired triplet. The unstable triplet model is also supported by the previous reports that RecA makes very limited contact with the incoming complementary DNA strand, which is instead held in place primarily by Watson-Crick hydrogen bond pairing interactions with the pre-synaptic ssDNA [36]. However, if the unstable triplet model is correct, then Dmc1 must somehow stabilize the imperfect triplet, or otherwise compensate for the loss of binding free energy that would be expected if the imperfect triplets were not paired within the Dmc1-dsDNA postsynaptic complex.

An alternative model, which we do not favor, is that imperfect RS-DNA base triplets are themselves inherently stable (Fig. 4.11E). This model could explain all of our results with Dmc1, but it would then require that RecA and Rad51 have built-in mechanisms for sensing and actively destabilizing imperfect base triplets. Moreover, it is difficult to envision how an RS-DNA base triplet would tolerate single, double, and triple mismatches and even abasic sites without any loss of triplet pairing stability. Indeed, molecular dynamics (MD) simulations of either RS-DNA or isolated base triplets reveal that RS-DNA is highly dynamic relative to B-DNA, and single base mismatches cause rapid destabilization of triplet pairing interactions.[2] Recent reports suggest that non-complementary sequences are stable within RecA filaments during a 10-ns MD simulation [375], so future work will be required to more fully understand how mismatched sequences behave during recombination. However, our experimental results seem to argue against a model where mismatched triplets are inherently stable and must be actively destabilized by RecA and Rad51. Specifically, Dmc1 can stabilize base triplets bearing single abasic sites at any of the three possible positions, and it is again difficult to envision that a base triplet harboring an abasic site could be inherently stable within an RS-DNA triplet. Together, we consider our findings to be most consistent with the unstable triplet model, suggesting that RecA and Rad51 need not actively destabilize imperfect RS-DNA base triplets, but instead Dmc1 must somehow compensate for the loss of Watson-Crick binding free energy due to mismatched base triplets.

[2]T. Terawaka and E. C. Greene, unpublished data.

4.3.14 Possible implications for homologous recombination

Our results suggest that RecA and Rad51 require perfect Watson-Crick pairing to allow stabilization of RS-DNA base triplets, whereas Dmc1 can stabilize non-complementary triplets embedded with longer tracts of homology. We emphasize that we do not yet know the biological implications of these findings. One possibility is that this biophysical difference may reflect the biological specialization of the meiosis-specific recombinase Dmc1, namely the requirement for Dmc1 to promote recombination between polymorphic alleles of different parental origins (see below). Future work will therefore be essential to determine whether or not the differences we observe between Rad51 and Dmc1 in response to mismatches and abasic sites reflect some broader biological difference between these two proteins.

Importantly, our work only reflects the effects of mutations within individual base triplets. The overall pairing interactions are highly stable, with half-lives still exceeding tens of minutes, even when mismatches or basic sites are present. Furthermore, we anticipate that the relative differences in stability for longer strand-exchange products will be very small. A single unpaired triplet within a paired intermediate that is tens or perhaps hundreds of base triplets in length will not in and of itself destabilize pairing interactions. Indeed, experimental and theoretical studies of RecA have recently shown that after initial recognition of an 8-nt homologous sequence, the presence of mismatches outside this region has little impact on overall stability [375]. The previous work with RecA by Prentiss and co-workers [375] is in good general agreement with our findings. At this stage, we can only infer that the differential stabilization of imperfectly matched RS-DNA triplets might reflect some broader difference in the overall stability of a complete strand-exchange product, which in turn may

manifest as a recombinase- and mismatch-dependent difference in recombination efficiency. Future work will be essential to continue testing for potential relationships between triplet pairing, or mispairing, and recombination efficiencies.

Finally, the results reported here reflect the basal activities of *E. coli* RecA, *S. cerevisiae* Rad51, and *S. cerevisiae* Dmc1. These recombinases acts in concert with many other proteins; for example, a total of ∼20 proteins or protein complex participate in recombination in *E. coli* [356, 376, 377], and 45 distinct proteins or protein complexes participate in recombination in *S. cerevisiae* [27, 293]. Interestingly, recent in vivo studies have shown that Rad51-mediated recombination is unexpectedly tolerant of mismatches [378]. Therefore, the crucial next step will be to start understanding how recombination accessory proteins augment the basal base triplet recognition activities of the recombinases during strand invasion.

4.4 Experimental Procedures

4.4.1 DNA curtains

Single-stranded DNA substrates were generated using M13mp18 (7,249-nt; New England Biolabs) as a template for rolling circle replication (RCR), as described [81, 306]. In brief, a biotinylated primer was annealed to the M13mp18 template (∼30 nM) at a 1:1.2 molar ratio (primer, M13mp18) in buffer containing 40 mM Tris-HCl (pH 8.0), 50 mM NaCl, and 10 mM $MgCl_2$. The annealing mixture was heated for 5 min at ∼90 °C and then cooled slowly to room temperature. Annealed products were diluted to a final concentration of 15 nM and stored at 4 °C until use. Replication reactions were prepared in RCR buffer (50 mM

Tris-HCl (pH 7.5), 4 mM DTT, 10 mM ammonium sulfate, and 10 mM $MgCl_2$), containing annealed primer/M13mp18 mixture (350 pM), 150 nM ϕ29 DNA polymerase, and dNTP mixture (200 μM each). Reactions were incubated at 30 °C for 25 min and then diluted 10-fold into buffer containing 40 mM Tris-HCl (pH 7.5), 2 mM $MgCl_2$, 50 mM NaCl, 1mM DTT, and 0.2 mg/ml BSA prior to use.

Flow cells were made by depositing chromium barriers onto the surface of a fused silica slide by electron beam lithography and then assembled into a sample chamber using double-sided tape and a borosilicate cover slide, as described [312]. Bilayers were prepared with 91.5% 1,2-dioleoyl-sn-glycero-3-phosphocholine, 0.5% biotinylated 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine-N-(biotinyl), and 8% mPEG 2000–1,2-dioleoyl-sn-glycero-3-phosphoethanolamine, and the biotinylated ssDNA molecules were attached to the bilayer through a biotin-streptavidin linkage. Anchored ssDNA molecules were then aligned at the barriers by application of flow in buffer containing 40 mM Tris-HCl (pH 7.5), 2 mM $MgCl_2$, 50 mM NaCl, 1 mM DTT, 0.2 mg/ml BSA, and 0.1 nM RPA-eGFP for 15 min at 1 ml/min. All experiments were conducted with a custom-built prism-type total internal reflection fluorescence microscope (Nikon) equipped with a 488-nm blue laser (Coherent Sapphine, 200 milliwatt) and a 561-nm yellow laser (Coherent Sapphine, 200 milliwatt).

4.4.2 Reaction conditions and data analysis

E. coli RecA, *S. cerevisiae* Rad51, and *S. cerevisiae* Dmc1 were prepared as described [306, 363]. Rad51/RecA pre-synaptic complexes were assembled as described previously [306, 363], using the following reaction buffers: E. coli RecA buffer, 25 mM Tris acetate

(pH 7.5), 4 mM magnesium acetate, 10 mM sodium acetate, 1 mM ATPγS, or 2.5 mM ATP plus an ATP-regenerating system (20 mM creatine and 0.04 mg/ml creatine kinase), 1 mM DTT, and 0.2 mg ml^{-1} BSA; S. cerevisiae Rad51 buffer, 30 mM Tris acetate (pH 7.5), 20 mM magnesium acetate, 50 mM KCl, 1 mM DTT, 2.5 mM ATP, and 0.2 mg ml^{-1} BSA; *S. cerevisiae* Dmc1, 40 mM Tris-HCl (pH 7.5), 2 mM $MgCl_2$, 1.5 mM $CaCl_2$, 100 mM KCl, 2.5 mM ATP, 1 mM DTT, and 0.2 mg ml^{-1} BSA. Note that the identity of the nucleotide cofactor used in each different experiment is specified in each of the corresponding figure panels. All data were collected at 30 °C. All dsDNA binding data were collected and analyzed as described previously [306, 363]. In brief, Atto565-tagged dsDNA oligonucleotides (2–10 nM; Table 4.1 and 4.2) were injected into the sample chamber in the same buffers (see above), and reactions were incubated for 10 min at 30 °C in the absence of buffer flow. Flow cells were quickly flushed (40 s at 1.0 ml min^{-1}) with fresh buffer to remove unbound Atto565-dsDNA, and the flow rate was then reduced (0.2 ml min^{-1}) to remove dissociated dsDNA and to replenish free nucleotide cofactor. Data were obtained by acquiring 100-ms frames at 30- or 60-s intervals; collection intervals were optimized relative to the overall lifetime of each dsDNA substrate, and the laser was shuttered between acquired images to minimize photobleaching. Kymographs were generated using Fiji and were analyzed by measuring the time that each Atto565-dsDNA molecule remained bound to the pre-synaptic complexes after flushing unbound DNA from the sample chamber. The probability that a bound molecule survived to a particular time point (t) was determined as the fraction of Atto565-dsDNA molecules that remained bound at time t, and survival probability graphs were constructed from the resulting data. All data points were calculated from an average of ∼200 molecules (n = 150–250). The dissociation kinetics for all dsDNA substrates were

well described by single exponential fits to the survival probability data (Fig. 4.2). Free energy calculations were performed as described previously [306, 363]. In brief, for dsDNA substrates harboring ≥8-nt of microhomology, the free energy barrier, ΔG‡, for escape from a potential well can be related to the rate as shown in Equation 4.1,

$$k_d = Ae^{-\frac{\Delta G^{\ddagger}}{k_b T}} \tag{4.1}$$

where A is the jump frequency, k_b is the Boltzmann constant, and T is temperature [306]. The difference in the barrier heights between two different dsDNA substrates can be compared , leading to Equation 4.2,

$$\Delta\Delta G^{\ddagger} = \Delta G_2^{\ddagger} - \Delta G_1^{\ddagger} = k_b T ln \frac{k_d^1}{k_d^2} \tag{4.2}$$

All reported $\Delta\Delta G^{\ddagger}$ values were normalized such that $\Delta G^{\ddagger}$ for the dsDNA containing a single 8-nt tract of microhomology is zero, and the experimentally measured data used to calculate $\Delta\Delta G^{\ddagger}$ were the k_d values for each different substrate obtained from survival probability data.

Chapter 5

Rad51 and Dmc1 lineage–specific amino acids influence the fidelity of genetic recombination

This chapter is adapted from a manuscript in submission as: "Rad51 and Dmc1 lineage-specific amino acids regulate the fidelity of genetic recombination," **Justin B. Steinfeld**, Ondrej Belan, YoungHo Kwon, Tsuyoshi Terakawa, Amr Al-Zain, Michael J. Smith, Zhi Qi, Weixing Zhao, Rodney Rothstein, Lorraine S. Symington, Patrick Sung, Simon J. Boulton, Eric C. Greene. Manuscript in submission to *Molecular Cell* (2018). I did everything in this paper except single molecule experiments for CeRad51-WT by ZQ, Rad54 foci experiments by MJS, MD simulations by TT, bulk biochemistry assays for yeast and human proteins by YK and for *C. elegans* by OB, and zeocin and MMS plating by AA.

5.1 Summary

The vast majority of eukaryotes possess two DNA recombinases, Rad51, which is ubiquitously expressed, and the meiosis–specific recombinase Dmc1. The evolutionary origins and potential benefits of this two–recombinase system remain poorly understood. Interestingly, Dmc1 can stabilize heteroduplex recombination intermediates bearing mismatches, whereas Rad51 cannot. Here, we demonstrate that this difference can be attributed to three

amino acids conserved within the Dmc1 lineage of the Rad51/RecA family, but absent from Rad51. Chimeric Rad51 mutants harboring Dmc1 lineage–specific amino acids gain the ability to stabilize mismatches *in vitro* and exhibit higher mismatch tolerance *in vivo*. We also show that RAD–51 from *Caenorhabditis elegans*, an organism that lacks Dmc1, has acquired "Dmc1–like" characteristics. We propose that the ability to stabilize imperfectly paired recombination intermediates reflects a fundamental distinction between Dmc1 and Rad51 that may have been established early in the evolutionary history of the Rad51/RecA family of recombinases.

5.2 Introduction

Homologous recombination (HR) enables the exchange of genetic information between DNA molecules and is a major driving force in evolution. HR plays essential roles in double–strand DNA break (DSB) repair [379], the rescue of stalled or collapsed replication forks [327, 379], and meiosis [141, 380]. During HR, a presynaptic single–stranded DNA (ssDNA) is paired with the complementary strand of a homologous double stranded DNA (dsDNA), resulting in displacement of the non–complementary strand [295, 356], and the resulting D–loop intermediates can then be channeled through several mechanistically distinct pathways to complete repair [292, 379]. The DNA pairing reactions that take place during HR are promoted by the Rad51/RecA family of DNA recombinases, which are ATP–dependent proteins that form extended helical filaments on DNA, referred to as presynaptic complexes [295, 356, 359]. Crystal structures of RecA–ssDNA presynaptic and RecA–dsDNA postsynaptic complexes reveal that the DNA is organized into near B–form base triplets separated

by 8 Å between adjacent triplets [36, 359]. This structural organization likely underpins homology recognition mechanisms and the ability of the Rad51/RecA family of recombinases to promote DNA strand invasion in 3–nt steps [306, 330, 359, 363, 381].

In mitotic cells, HR is primarily used for the repair of spontaneous DNA breaks, such as those associated with DNA replication errors [292, 327, 379]. Mitotic HR is strongly biased towards inter–sister recombination, and the resulting intermediates are preferentially channeled through the synthesis–dependent strand annealing pathway (SDSA), and double Holliday junction (dHJ) dissolution pathways, which yield only noncrossover recombinants and thus help avoid chromosome arm translocations that may stem from crossover recombination [379]. Meiotic HR is used to repair programmed DSBs generated by the Spo11 complex and is biased towards inter–homolog recombination to allow for the formation of crossovers necessary for accurate chromosome segregation in the first meiotic division [141, 355, 380].

Rad51 is the only recombinase in mitotic cells, whereas Rad51 and Dmc1 are both expressed during meiosis in the vast majority of eukaryotes [141, 355, 380]. Dmc1 is responsible for catalyzing inter–homolog recombination during meiosis, while Rad51 promotes Dmc1 presynaptic filament assembly and as participates in inter–sister repair that give rise to noncrossover outcomes[141, 155, 380]. Although Rad51 and Dmc1 interact with different subsets of accessory factors, there are few other biochemical differences between them that might explain the evolutionary advantages of using different recombinases during mitosis and meiosis [141, 380]. Biophysical studies have shown that Rad51–ssDNA filaments can bind to dsDNA fragments containing short tracts of sequence microhomology, but the introduction of a single nucleotide mismatch causes a reduction in binding lifetime commensurate with

the loss of one base triplet pairing interaction [363, 381]. In contrast, Dmc1–ssDNA can tolerate base triplets bearing single, double, or triple mismatches and even abasic sites with no change in the binding lifetimes of the heteroduplex DNA intermediates [363, 381]. These findings suggest that Dmc1 can stabilize heteroduplex DNA joints containing mismatched base triplets, whereas Rad51 cannot [363, 381, 382]. We have hypothesized that the ability of Dmc1 to stabilize imperfectly paired recombination intermediates might reflect an intrinsic difference in fidelity between the two eukaryotic recombinases [363, 381]. However, the molecular basis for these biophysical differences, and their biological implications remained unexplored.

Here, we use structural analysis and bioinformatics to identify Rad51 lineage–specific amino acid residues and Dmc1 lineage–specific amino acid residues that might contribute to their unique responses to mismatched base triplets. Based upon these analyses, we swapped Rad51 lineage–specific amino acid residues present at the putative DNA–binding interfaces with their lineage–specific counterparts from Dmc1, and vice versa. Single molecule biophysical analysis of these chimeric recombinases reveals that the differential responses of *S. cerevisiae* and human Rad51 and Dmc1 to mismatches can be attributed to three lineage–specific amino acid residues within DNA–binding loop L1. MAT switching analysis provides genetic evidence that these L1 residues affect recombination between divergent sequences *in vivo*. Remarkably, *C. elegans* RAD–51 L1 amino acid residues more closely resemble Dmc1. Accordingly, wild–type *C. elegans* RAD–51 stabilizes mismatched substrates, as is observed for yeast and human Dmc1, whereas mutation of *C. elegans* L1 residues to their "canonical" Rad51 counterparts abolishes mismatch stabilization. Together, our results show that lineage–specific amino acid residues in the L1 DNA–binding domain are responsible for the

differential response of Rad51 and Dmc1 to mismatched sequences and suggest that the functions of these amino acid residues in regulating recombination fidelity may be broadly conserved among eukaryotes.

5.3 Results

5.3.1 Identification of Rad51 and Dmc1 lineage-specific amino acids

The Rad51 and Dmc1 lineages within the Rad51/RecA family of recombinases arose early in the evolutionary history of eukaryotes [358, 380, 383, 384]. These proteins remain closely related, for instance, *S. cerevisiae* Rad51 (ScRad51) and Dmc1 (ScDmc1) share 56% sequence similarity and 45% sequence identity [141, 380]; for brevity, we use the nomenclature Rad51 and Dmc1 as general designations, and we use ScRad51, ScDmc1, hRAD51 (human RAD51) and hDMC1 (human DMC1) when referring to specific recombinases. Rad51 and Dmc1 form similar filaments on ssDNA, and both promote DNA strand invasion [141, 380]. However, Dmc1 can stabilize imperfectly paired base triplets, whereas Rad51 cannot [363, 381]. We speculated that Dmc1–specific amino acid residues might be responsible for this differential response to mismatches. Furthermore, we presumed that residues responsible for this effect might fulfill three criteria: *(i)* they should be conserved within the Dmc1 lineage of the recombinase family; *(ii)* they should be absent from the Rad51 lineage; and *(iii)* they would likely be within one of the two known DNA binding motifs, DNA–binding loop 1 (L1) or DNA–binding loop 2 (L2), which are present in all recombinase family members

[36, 384]. Given these criteria, we sought to determine whether Rad51 and Dmc1 harbor lineage–specific residues within the L1 and L2 DNA–binding loops.

L1 and L2 were originally identified from examination of the crystal structure of *E. coli* RecA [384]. Therefore, we utilized *E. coli* RecA as a model to verify the boundaries of the L1 and L2 motifs (Fig. 5.1A) [36]. *E. coli* RecA shares 26.5 and 25.6% sequence identity with ScRad51 and ScDmc1, respectively, and the core domain of *E. coli* RecA co–aligns with the core domain of ScRad51 with root mean square deviation of 1.5 Å [37]. We then mapped these regions onto a primary structure alignment of Rad51 and Dmc1 from *S. cerevisiae*, *H. sapiens*, *P. carinii*, *O. sativa*, *M. musculus*, *E. histolytica*, *T. gondii*, and *S. scrofa* (Fig. 5.2). From this initial comparison, we identified four amino acid residues within L1 and five amino acid residues within L2 that are conserved within either the Rad51 lineage, or the Dmc1 lineage, but are divergent between the two recombinases (Fig. 5.2A). For ScRad51, these lineage–specific amino acid residues correspond to L1 residues T288, A298, M301, and H302; and L2 residues V328V, Q330, V331, D332, and N348. The ScDmc1 lineage–specific amino acids include L1 residues V224, E234, Q237, and K238; and L2 residues Q264, D266, P267, G268, and H285. We validated this initial assignment by analysis of 600 Rad51 protein sequences and 270 Dmc1 sequences (Fig. 5.2B,C and Table 5.1 and 5.2).

5.3.2 Biochemical characterization of Rad51 and Dmc1 chimeras

If Dmc1 lineage–specific residues present within L1, L2, or both are responsible for mismatch tolerance, then mutation of these residues to those present in Rad51 might abolish this property. Conversely, mutation of the Rad51 lineage–specific residues to their Dmc1

	ScRad51 Loop 1					ScDmc1 Loop 1			
aa	T288	A298	M301	H302		V224	E234	Q237	K238
A	2	454	6	5		0	3	0	6
C	0	0	2	0		6	0	0	0
D	0	7	0	0		0	81	0	0
E	0	12	0	9		0	185	0	0
F	0	0	0	0		0	0	0	0
G	0	0	36	0		0	0	1	0
H	0	0	1	554		0	0	0	1
I	0	1	15	0		0	0	0	1
K	0	0	0	8		0	1	0	236
L	0	1	11	4		0	0	0	1
M	0	1	359	3		0	0	0	1
N	0	35	38	3		1	0	0	0
P	0	4	0	0		0	0	0	0
Q	0	1	26	4		12	0	267	0
R	0	0	0	0		0	0	0	13
S	55	71	26	10		4	0	1	2
T	533	11	66	0		22	0	0	3
V	8	2	14	0		225	0	1	6
W	0	0	0	0		0	0	0	0
Y	0	0	0	0		0	0	0	0
total	600	600	600	600		270	270	270	270

Table 5.1: **Rad51 and Dmc1 L1 amino acid conservation.**

	ScRad51 Loop 2					ScDmc1 Loop 2			
aa	V328	Q330	V331	D332	N348	D266	P267	G268	H285
A	0	1	2	1	0	0	0	5	0
C	0	0	1	0	0	1	0	0	0
D	0	5	0	589	0	262	0	6	0
E	0	2	0	3	0	2	0	0	0
F	0	0	0	0	0	0	0	0	0
G	0	1	0	4	0	0	0	246	0
H	0	0	0	0	96	0	0	0	265
I	5	0	0	0	0	0	0	0	0
K	0	29	0	0	0	0	0	0	0
L	5	0	15	0	0	0	1	0	0
M	17	0	0	0	0	0	0	0	0
N	0	49	0	2	502	1	0	0	5
P	0	0	29	0	1	0	268	0	0
Q	4	471	0	0	0	1	0	0	0
R	0	3	0	0	0	0	0	0	0
S	3	28	0	0	0	3	0	13	0
T	28	10	0	1	0	0	0	0	0
V	538	1	553	0	0	0	1	0	0
W	0	0	0	0	0	0	0	0	0
Y	0	0	0	0	0	0	0	0	0
total	600	600	600	600	599	270	270	270	270

Table 5.2: **Rad51 and Dmc1 L2 amino acid conservation.**

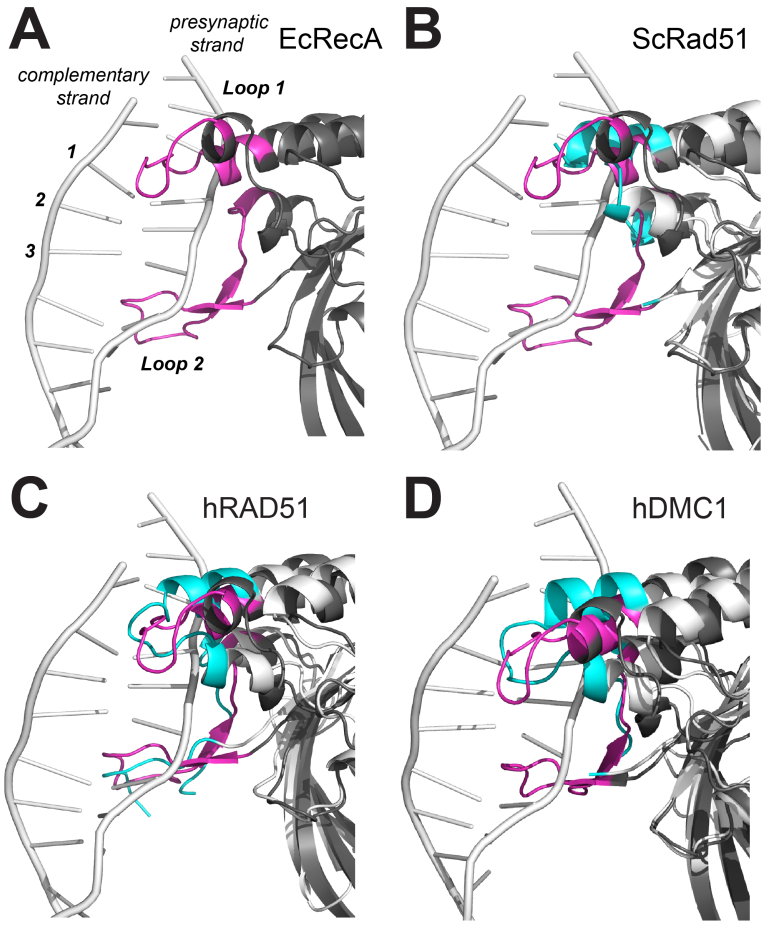

Figure 5.1: **Structural alignments of L1 and L2 DNA binding loops for Rad51/RecA family members.** (**A**) Crystal structure of *E. coli* RecA (PDB: 1CMX) highlighting the L1 and L2 DNA binding loops. L1 and L2 are shown in magenta, the remainder of the protein is in dark gray, the presynaptic DNA strand and its complementary strand are labeled, and the number designations highlight one of the base triplets. Structure of *E. coli* RecA aligned with (**B**) ScRad51 (PDB: 1SZP), (**C**) hRAD51 (PDB: 5HB1), and (**D**) hDMC1 (PDB: 4HYY). In (B-D), the RecA color coding is as shown in panel (A), the eukaryotic protein ribbon diagrams are light gray, and the eukaryotic L1 and L2 domains are shown in cyan.

counterparts might enable Rad51 to stabilize mismatched recombination intermediates. To test these hypotheses, we designed chimeric recombinases by swapping the entire L1 and L2 motifs (Table 5.3). For brevity, we assigned names to the mutants based on the identity of the altered residues (e.g. ScDmc1–RL1 refers to *S. cerevisiae* Dmc1 harboring the ScRad51 L1; hRAD51–DL12 refers to human RAD51 with the amino acid residues from hDMC1 L1 and L2; etc.; Table 5.3). All mutants behaved like their wild–type (wt) counterparts during expression and purification, with the exception of the hDMC1 chimeras, which were significantly less soluble than the wt protein *(not shown)*; therefore, we were unable to analyze these hDMC1 mutants.

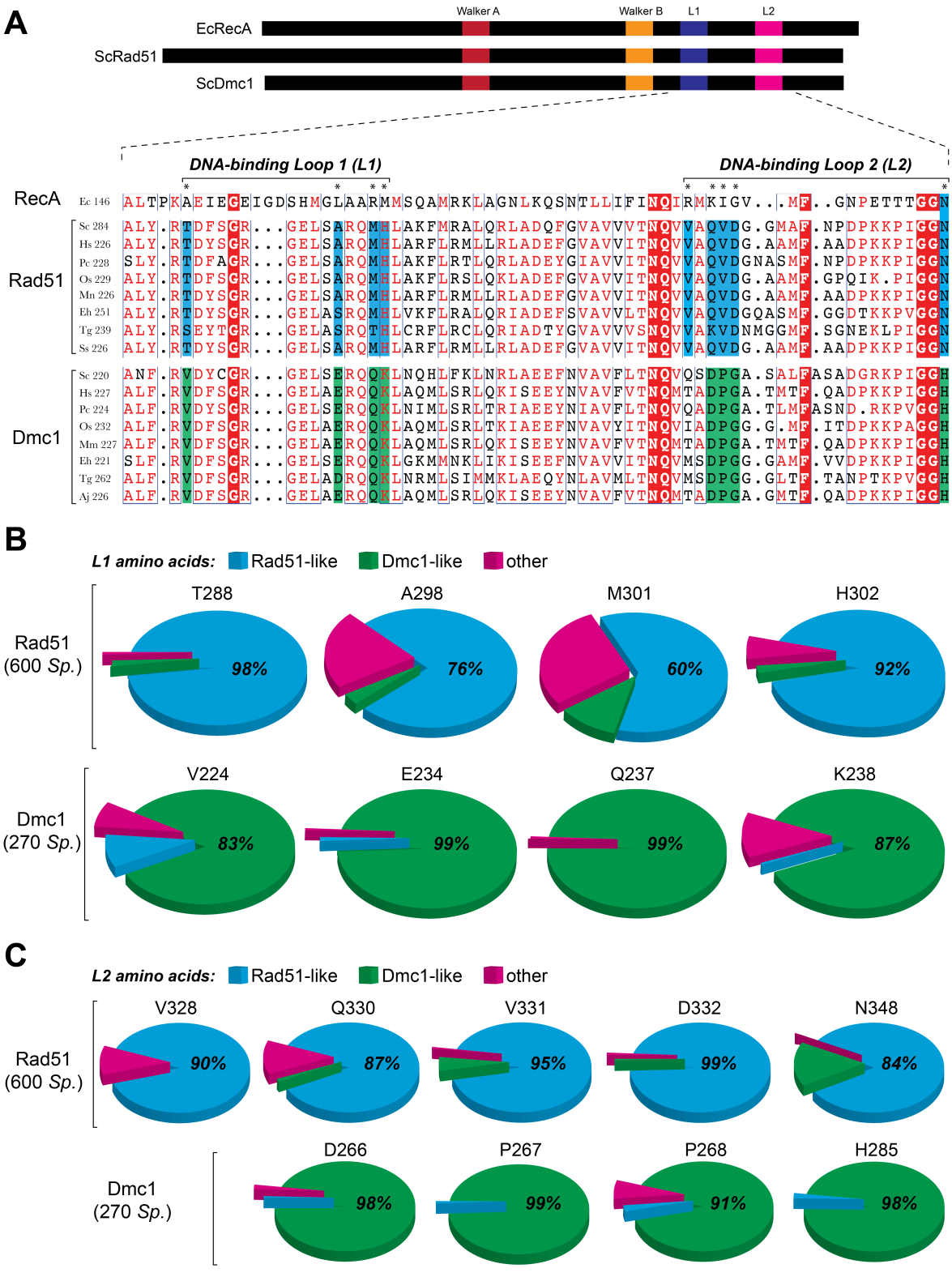

Figure 5.2: **Identification of Rad51 and Dmc1 L1 and L2 lineage-specific amino acids.** (**A**) Location and sequences of the L1 and L2 DNA-binding loops from RecA, Rad51, and Dmc1. Amino acids conserved in all three lineages are highlighted in red, Rad51– lineage specific amino

Figure 5.2: (cont.) acids are highlight in blue, and Dmc1-lineage specific amino acids are highlighted in green. Included in the alignments are recombinases from *S. cerevisiae*, *Homo sapiens*, *Pneumocystis carinii*, *Oryza sativa*, *Mus musculus*, *Entamoebe histolytica*, *Toxoplasma gondii*, *Sus scrofa* and *Anguilla japonica*. (**B**) Conservation and identity of L1 amino acids based upon analysis of 600 Rad51 and 270 Dmc1 sequences. The analyzed positions correspond to ScRad51 amino acids T288, A298, M301 and H302, and ScDmc1 amino acids V224, E234, Q237 and K238. Color-coding indicates Rad51-like, Dmc1-like and other amino acids (see also Tables 5.1 and 5.2). (**C**) Conservation and identity of L2 amino acids.

Protein	**L1 aa sequence**	**L2 aa sequence**
ScRad51	TDFSGRGELSARQMH	VAQVDGGMAFNPDPKKPIGGN
ScRad51-DL1	**V**D**YC**GRGELS**E**RQ**QK**	VAQVDGGMAFNPDPKKPIGGN
ScRad51-DL2	TDFSGRGELSARQMH	**QSDPGASALFASA**D**GR**KPIGGH
ScRad51-D12	**V**D**YC**GRGELS**E**RQ**QK**	**QSDPGASALFASA**D**GR**KPIGGH
ScDmc1	VDYCGRGELSERQQK	QSDPGASALFASADGRKPIGGH
ScDmc1-RL1	**T**D**FS**GRGELS**A**RQ**MH**	QSDPGASALFASADGRKPIGGH
ScDmc1-RL2	VDYCGRGELSERQQK	**VAQVDGGMAFNP**D**PK**KPIGG**N**
ScRad51-RL12	**T**D**FS**GRGELS**A**RQ**MH**	**VAQVDGGMAFNP**D**PK**KPIGG**N**
hRAD51	TDYSGRGELSARQMH	VAQVDGAAMFAADPKKPIGGN
hRAD51-DL1	**V**DYSGRGELS**E**RQ**QK**	VAQVDGAAMFAADPKKPIGGN
hRAD51-DL2	TDYSGRGELSARQMH	**T**A**DPGATMTFQ**ADPKKPIGG**H**
hRAD51-DL12	**V**DYSGRGELS**E**RQ**QK**	**T**A**DPGATMTFQ**ADPKKPIGG**H**
hDMC1	VDYSGRGELSERQQK	TADPGATMTFQADPKKPIGGH
hDMC1-RL1	**T**DYSGRGELS**A**RQ**MH**	TADPGATMTFQADPKKPIGGH
hDMC1-RL2	VDYSGRGELSERQQK	**VAQVDGAAMFA**ADPKKPIGG**N**
hDMC1-RL12	**T**DYSGRGELS**A**RQ**MH**	**VAQVDGAAMFA**ADPKKPIGG**N**

Table 5.3: **Chimeric recombination design at L1 and L2.**

Each chimeric protein was tested for ATP hydrolysis and DNA strand exchange activity (Figure 5.3A). Most of the chimeras retained DNA–dependent ATPase activity, albeit typically at a lower level relative to the wt recombinases, and the single loop swaps exhibited greater ATP hydrolysis activity than the double loop swaps (Figure 5.3A). DNA strand exchange assays revealed that the mutant proteins with a single chimeric loop swap exhibited activity comparable to that of their wt counterparts (Figure 5.3B,C). However, chimeric

recombinases in which both loops were swapped, were deficient in strand exchange activity, in particular, ScDmc1–RL12, was significantly compromised for both strand exchange and ATP hydrolysis (Figure 5.3). Interestingly, ScDmc1–RL1 exhibited 3–fold more strand exchange activity compared to wt Dmc1 (Figure 5.3B,C). These findings indicate that the identity of the lineage–specific amino acid residues contribute to recombinase interactions with DNA substrates, and also show that it is possible to swap the Rad51 and Dmc1 L1 or L2 regions without abolishing basic protein activities *in vitro*.

5.3.3 Assembly of presynaptic filaments with chimeric recombinases

The chimeric recombinases were tested for the ability to assemble into stable presynaptic complexes using ssDNA curtain assays (Fig. 5.4). As reported, the addition of wt (unlabeled) Rad51 or Dmc1 results in ATP–dependent displacement of RPA–GFP from the ssDNA, reflecting the assembly of the presynaptic complexes (Fig. 5.4)[385]. The RPA–GFP reappears when ATP, or both ATP and Ca^{2+} in the case of ScDmc1 and hRAD51, is flushed from the sample chamber, reflecting presynaptic complex disassembly (Fig. 5.4)[385]. Most of the chimeras assembled into presynaptic filaments, which remained stable for $\geq$30 minutes so long as ATP (and Ca^{2+} when appropriate) was maintained in the reaction buffer (Fig. 5.4 and Table 5.4). One exception was ScDmc1–RL12, which failed to assemble into stable filaments *(not shown)* and was not characterized further.

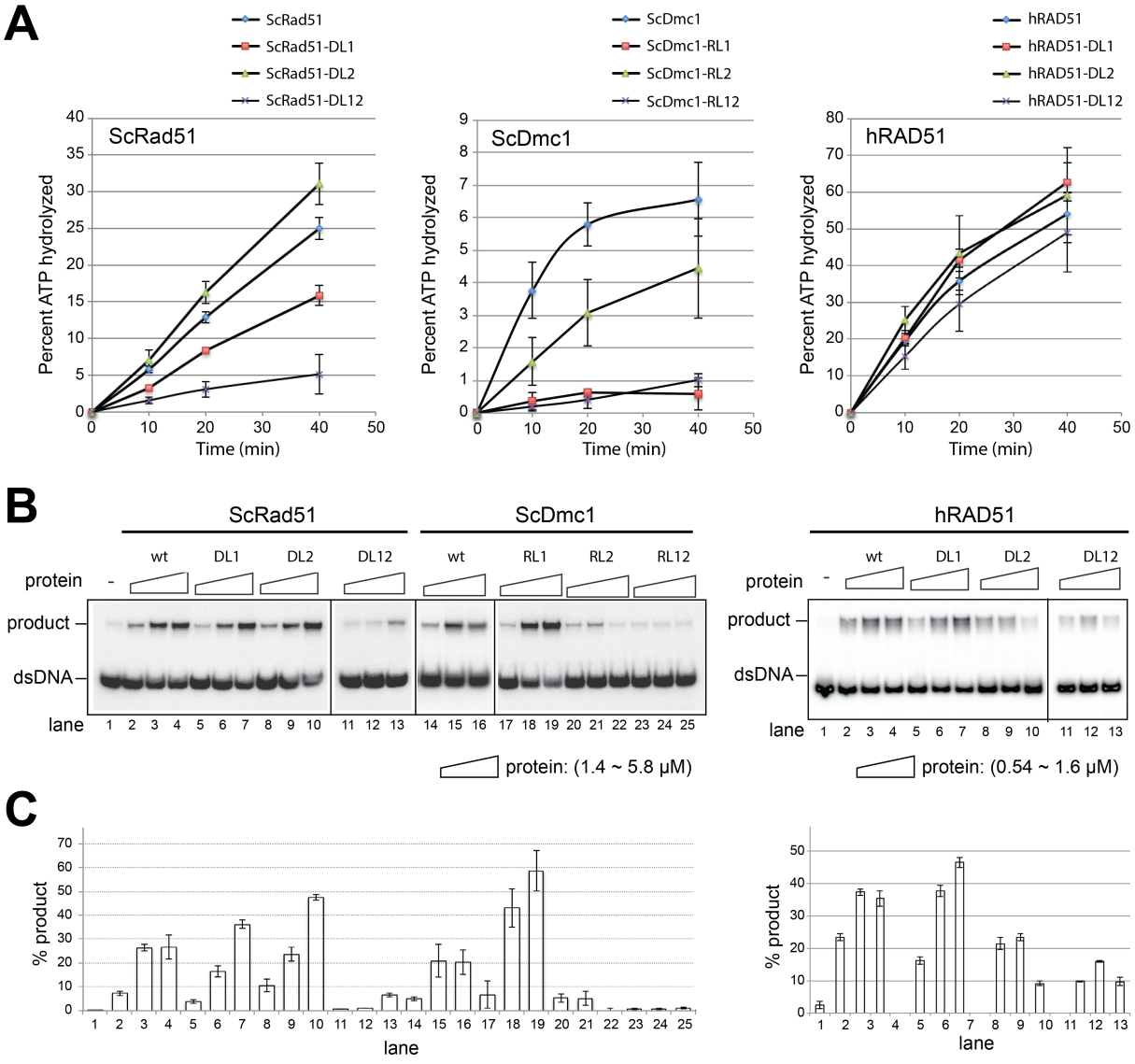

Figure 5.3: **Biochemical characterization of human and yeast recombinases.** (**A**) ATP hydrolysis assays, (**B**) D-loop formation assays, and (**C**) quantitation of D-loop formation for each of the yeast and human chimeric recombinases. Lane designations in (B) correspond to the quantitation in panel (C). Error bars in (A) and (C) represent the mean ± s.d. from three separate experiments.

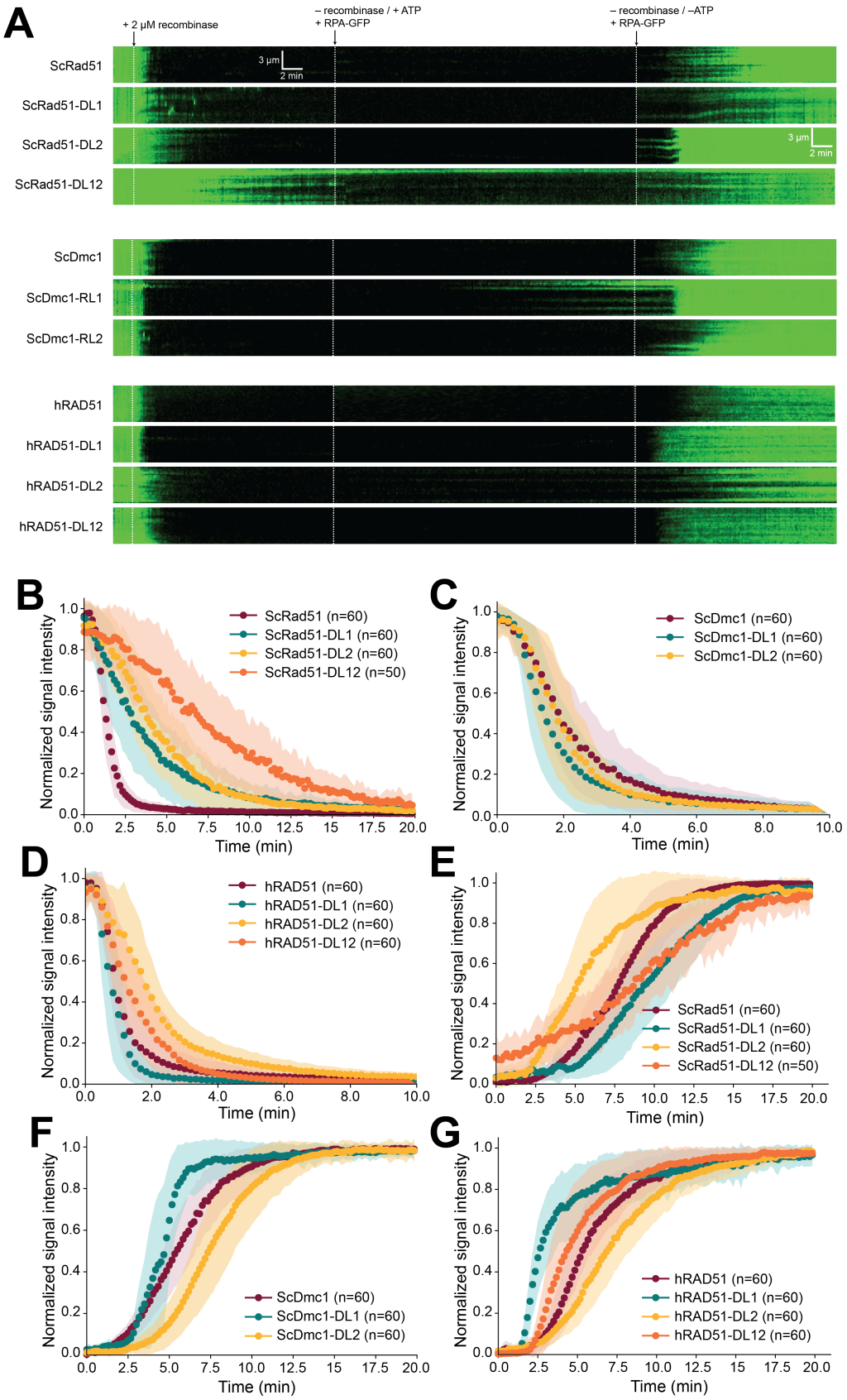

Figure 5.4: **Characterization of presynaptic complexes.** These ssDNA curtain assays use

Figure 5.4: (cont.) fluorescent GFP-tagged RPA (replication protein A) as a proxy for Rad51 or Dmc1 binding. In brief, ssDNA was generated using M13mp18 (7,249-nt) as a template for rolling circle replication, anchored to a lipid bilayer within a microfluidic chamber through a biotin-streptavidin linkage, and aligned along chromium (Cr) barriers, as described. The ssDNA unravels when incubated with RPA-GFP, and the downstream ends of the RPA-ssDNA are anchored to Cr pedestals, and these nucleoproteins complexes can be visualized by total internal reflection fluorescence microscopy (TIRFM). (**A**) Kymographs showing typical examples of presynaptic complex assembly, stability and disassembly assays for the yeast and human chimeric recombinases. These assays use GFP-RPA (green) as a read out for protein occupancy on the ssDNA. The ssDNA is bound by GFP-RPA at the outset of the measurements, and filament assembly is initiated by injection of the appropriate recombinase and ATP at the indicated time point (first dashed white line). Filament assembly is revealed as the loss of GFP-RPA signal. Unbound recombinase is then flushed from the sample chamber (second dashed white line) and the filaments are observed for $\geq$30 minutes to verify that they remain intact. Filament disassembly is then triggered by flushing free ATP from the sample chamber (third dashed white line) while monitoring the rebinding of GFP-RPA. (**B-D**) Filament assembly kinetics for each of the indictated recombinases. (**E-G**) Filament disassembly kinetics for each indicated recombinase. Error bars in (B-G) represent mean $\pm$ s.d. Assembly and disassembly lifetimes are presented in Table 5.4.

5.3.4 Chimeric recombinases exhibit base triplet stepping

We have developed a ssDNA curtain assay for visualizing DNA strand exchange intermediates at the single–molecule level [306, 363, 381]. In brief, a series of Atto565–labeled dsDNA substrates (70–bp), harboring 8– to 15–nt tracts of microhomology targeted towards unique sequences in the M13 ssDNA, are incubated with the presynaptic complexes, and unbound dsDNA is flushed away (Fig. 5.5A,B). The resulting intermediates are visualized by total internal reflection fluorescence microscopy (TIRFM), and dsDNA dissociation rates are obtained from the survival probabilities of the bound dsDNA fragments [306, 363]. Using this assay, we have shown that RecA, Rad51 and Dmc1 stabilize paired heteroduplex intermediates in 3–nt increments, each base triplet "step" coincides with an energetic signature ($\Delta\Delta G^{\ddagger}$) of $\sim 0.3 k_B T$, corresponding to an $\sim$30% change in the dissociation rates [306, 363, 381]. Importantly, presynaptic complexes prepared with the chimeric recombinases could bind the Atto565–labeled dsDNA, the resulting dsDNA dissociation rates were

Protein	Assembly $t_{1/2}$ (min)	± S.D.	Disassembly $t_{1/2}$ (min)	± S.D.
ScRad51	1.35	± 0.013	7.72	± 0.037
ScRad51-DL1	3.15	± 0.038	9.63	± 0.036
ScRad51-DL2	3.98	± 0.033	5.81	± 0.048
ScRad51-DL12	6.90	± 0.053	8.59	± 0.044
ScDmc1	2.11	± 0.027	5.59	± 0.024
ScDmc1-RL1	1.59	± 0.027	4.51	± 0.031
ScDmc1-RL2	1.90	± 0.025	7.56	± 0.032
ScDmc1-RL12	NA	NA	NA	NA
hRAD51	1.06	± 0.018	5.80	± 0.028
hRAD51-DL1	0.81	± 0.019	3.32	± 0.031
hRAD51-DL2	1.90	± 0.022	7.26	± 0.026
hRAD51-DL12	1.34	± 0.020	4.82	± 0.037
CeRad51	1.24	± 0.022	30.4	± 0.5
CeRad51-TM	1.50	± 0.036	22.5	± 0.2
ScRad51-T288V	1.07	± 0.016	6.28	± 0.054
ScRad51-A298E	1.49	± 0.015	4.35	± 0.024
ScRad51-M301Q	1.10	± 0.017	5.62	± 0.052
ScRad51-H302K	1.38	± 0.027	6.97	± 0.036
ScRad51-A298E Q302K	2.96	± 0.046	6.91	± 0.026

Table 5.4: **Presynaptic complex assembly and disassembly kinetics.**

comparable to those measured for the wt proteins, and the dissociation rates also varied in 3–nt increments (Fig. 5.5C-J, 5.6, and 5.7). We conclude that the chimeric Rad51 and Dmc1 recombinases possess dsDNA–binding and base triplet stepping attributes similar to those determined for their wt counterparts.

5.3.5 Dmc1 L1 lineage–specific amino acid residues regulate mismatch stabilization

We next asked how the chimeric recombinases responded to DNA mismatches. Rad51, RecA, and Dmc1 require perfect Watson–Crick pairing interactions to stabilize base triplets

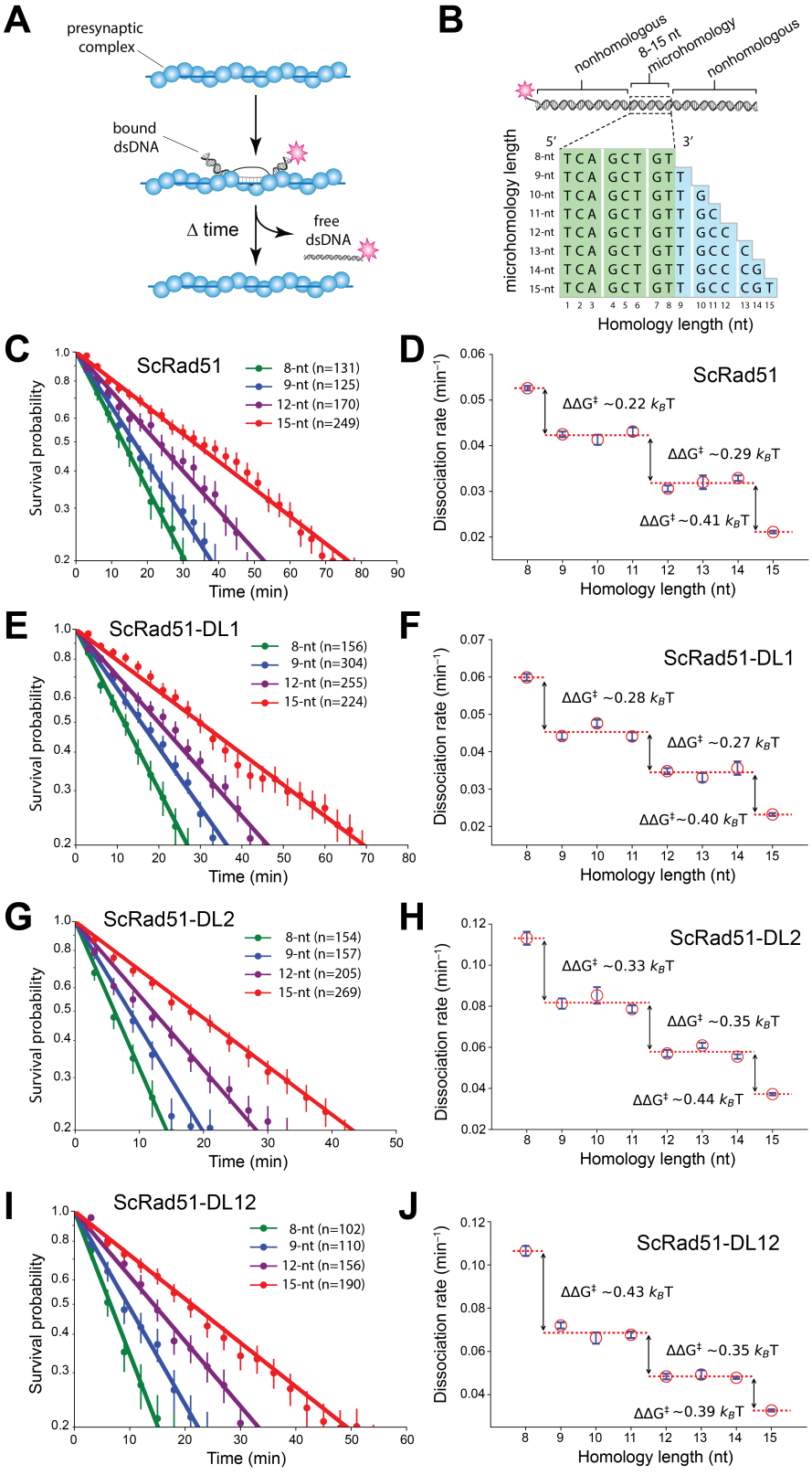

Figure 5.5: **Chimeric recombinases undergo base triplet stepping.** (**A**) Schematic of dsDNA capture assay. (**B**) Illustration of dsDNA substrates used for testing base triplet stepping.

Figure 5.5: (cont.) The 8-nt tract of microhomology highlighted in green is required for efficient binding, and the nucleotides highlighted in blue represent incremental increases in the microhomology length. (**C**) Survival probabilities and (**D**) dissociation rate data obtained from the survival probability plots for wt ScRad51. (**E**) Survival probabilities and (**F**) dissociation rate data for ScRad51-DL1. (**G**) Survival probabilities and (**H**) dissociation rate data for ScRad51-DL2. (**I**) Survival probabilities and (**J**) dissociation rate data for ScRad51-DL12. For all dissociation rate graphs, arrows indicate stepwise reductions in dissociation rates coincident with recognition of the third base of each triplet, dashed lines report the mean rate for each step, and the free energy changes ($\Delta\Delta G^{\ddagger}$) associated with each triplet step are indicated. Here and throughout, error bars for survival probability plots represent 70% confidence intervals obtained through bootstrap analysis, error bars for the dissociation rate data represent mean ± s.d., the number of events used to calculate these values are shown the survival probability panels.

located at the terminal positions of tracts of microhomology tracts embedded within the dsDNA (Fig. 5.8Ai)[363, 381]. Rad51 and RecA also require perfect Watson–Crick base pairing interactions to stabilize base triplets located at internal positions (Fig. 5.8Aii)[363, 381]. In contrast, Dmc1 can stabilize mismatches at internal positions (Fig. 5.8Aiii)[363, 381]. Indeed, Dmc1 can stabilize single, double and triple mismatches and even abasic sites, so long as these imperfect triplets are flanked by homologous sequences [381].

Consistent with previous results, none of the recombinases was capable of stabilizing a base triplet located at the terminal position of a 12–nt tract of microhomology (Fig. 5.8B), and the resulting substrates exhibited dissociation rates similar to those measured for a substrate with only 9–nts of microhomology (Fig. 5.8D–F). We next tested the chimeric recombinases with mismatched triplets that were juxtaposed to a single perfectly paired triplet with a 15–nt tract of microhomology (Fig. 5.8C). As previously shown, ScRad51 and hRAD51 could step over internal mismatches, but could not stabilize the internal mismatched triplet, instead yielding dissociation rates comparable to a substrate bearing only 12–nts of microhomology (Fig. 5.8H,I)[363, 381]. In contrast, ScDmc1 and hDMC1 yielded dissociation rates comparable to the corresponding substrate bearing 15–nts of perfect mi-

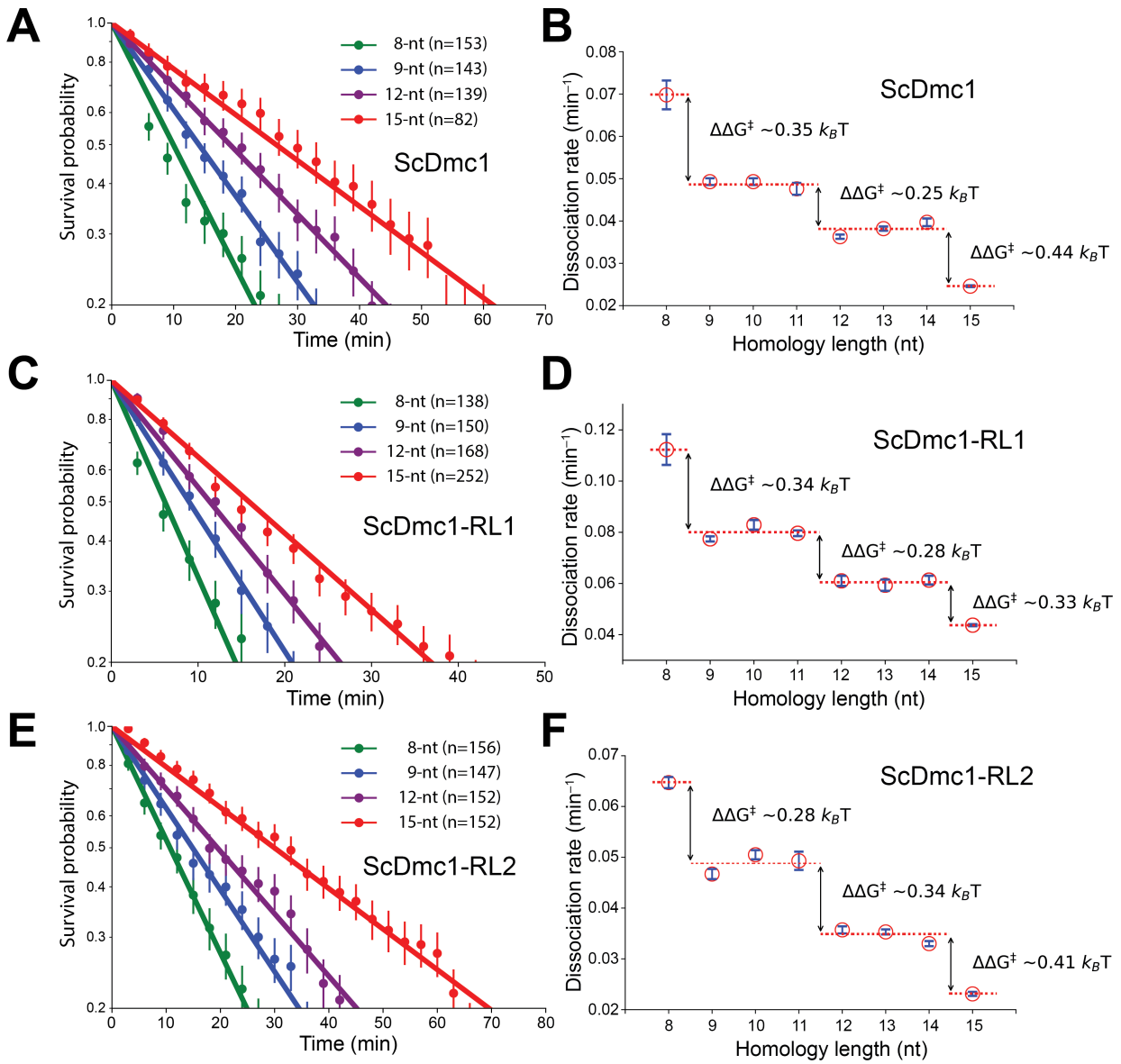

Figure 5.6: **dsDNA binding characteristics of presynaptic complexes prepared with ScDmc1 chimeric proteins.** (**A**) Survival probabilities and (**B**) dissociation rate data for wt ScDmc1. (**C**) Survival probabilities and (**D**) dissociation rate data for ScDmc1-RL1. (**E**) Survival probabilities and (**F**) dissociation rate data for ScDmc1-RL2. As in the main text figures, error bars for survival probability plots represent 70% confidence intervals calculated by bootstrap analysis, error bars for the dissociation rate data represent mean ± s.d., the number of events used to calculate these values are shown the survival probability panels.

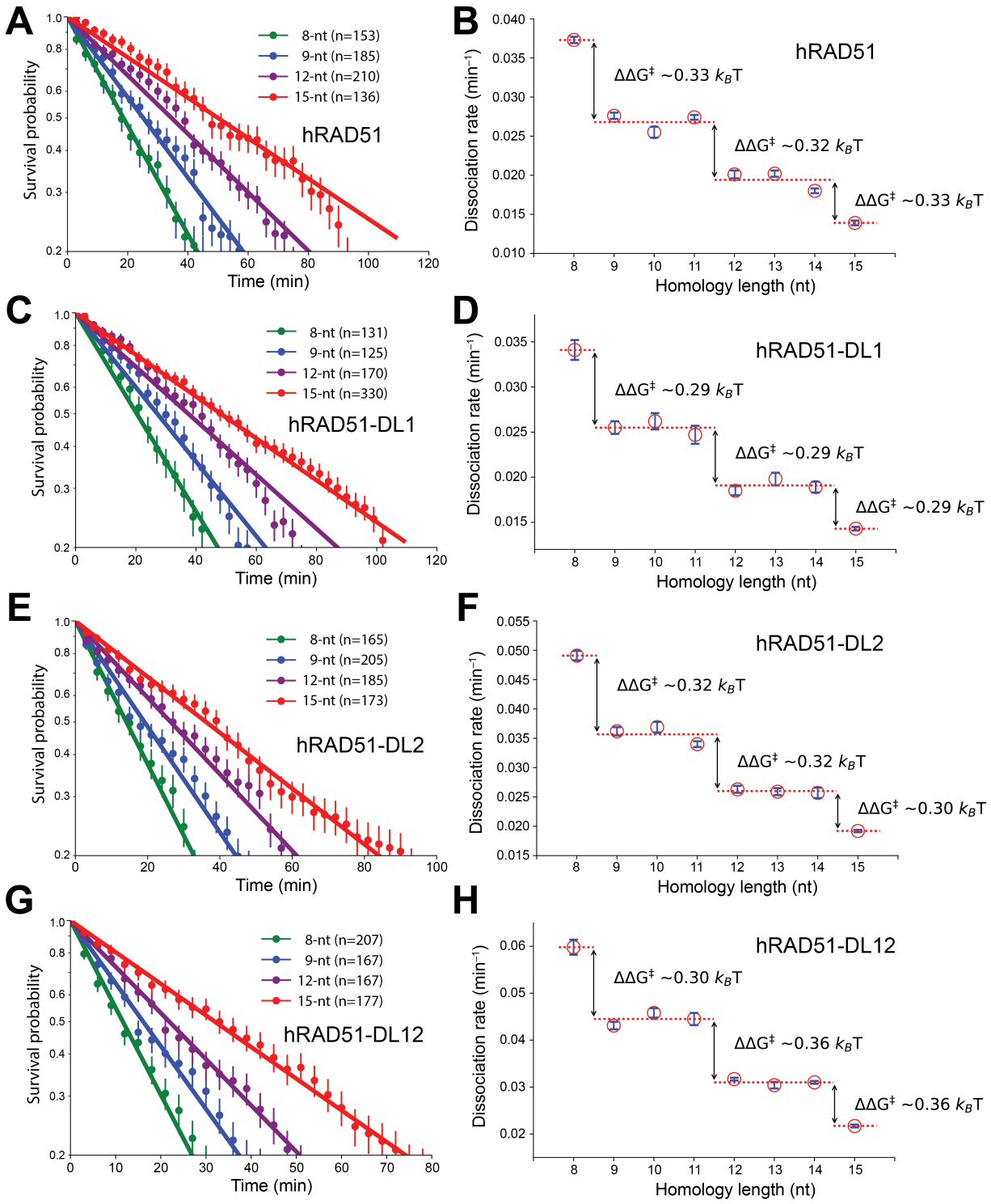

Figure 5.7: **dsDNA binding characteristics of presynaptic complexes prepared with hRAD51 chimeric proteins.** (**A**) Survival probabilities and (**B**) dissociation rate data for wt hRAD51. (**C**) Survival probabilities and (**D**) dissociation rate data for hRAD51-RL1. (**E**) Survival probabilities and (**F**) dissociation rate data for hRAD51-RL2. (**G**) Survival probabilities and (**H**) dissociation rate data for hRAD51-RL12.

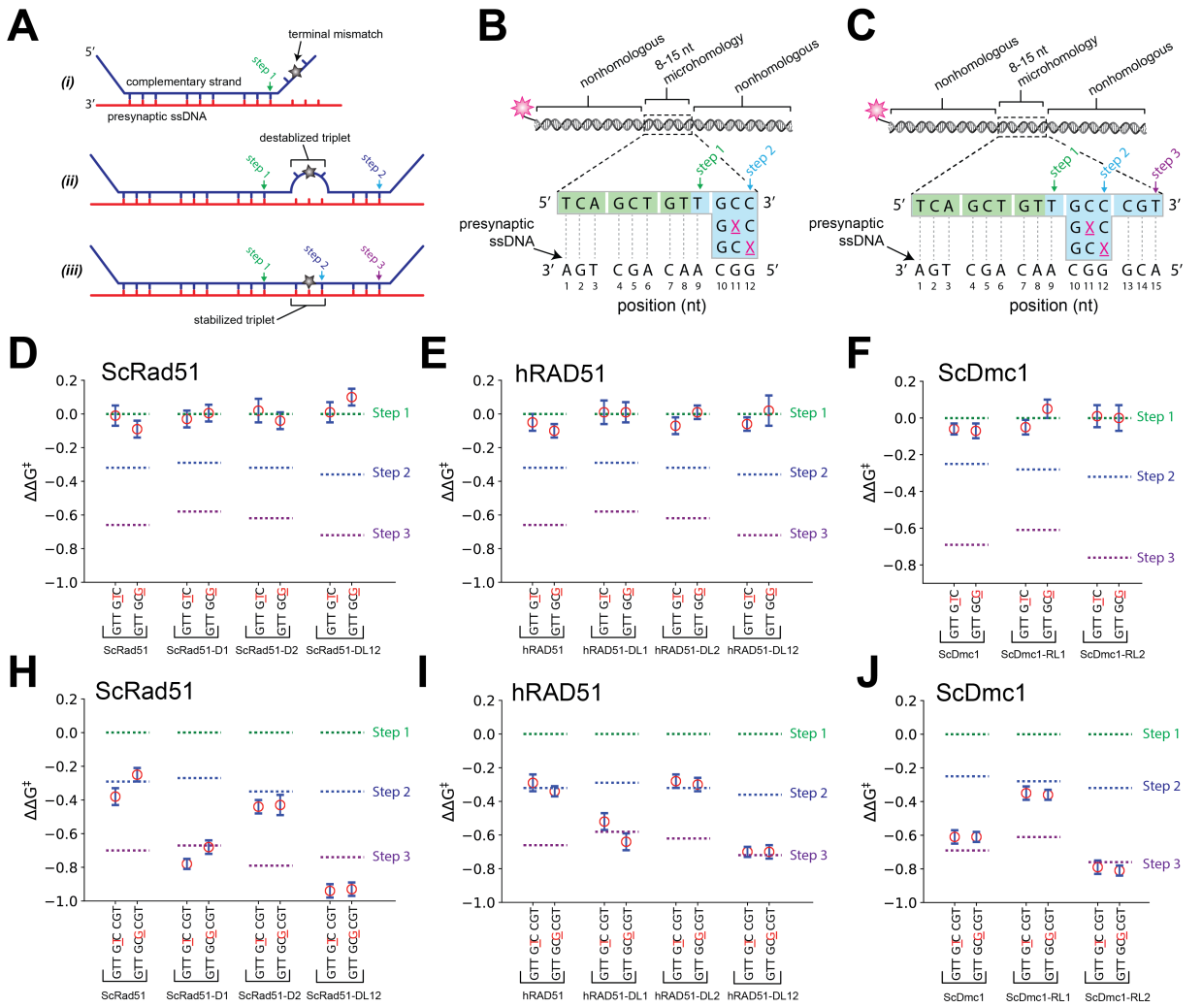

Figure 5.8: **The Dmc1 L1 DNA-binding loops allows for mismatch stabilization.** (**A**) Schematic illustration of a dsDNA substrate (shown in blue; non-complementary strand is omitted for clarity) bound to a presynaptic ssDNA (in red). A single mismatch is highlighted and is positioned either in the (i) terminal base triplet, (ii) an internal base triplet that is not stabilized (as with Rad51), or (iii) an internal triplet that is stabilized (as with Dmc1). The binding steps relative to triplet length are highlighted, where step 1 corresponds to the initial binding interaction, whereas steps 2 and 3 reflect the changes in dissociation rates that occur for each 3-nt increase in length. (**B**) Schematic of dsDNA substrates with mismatches positioned within the terminal base triplet. The location of the mismatches are highlighted as an underlined, magenta X. (**C**) Schematic of dsDNA substrates with mismatches positioned within the internal triplet. Assays with terminal mismatch substrates for wild type and chimeric versions of (**D**) ScRad51, (**E**) hRAD51, and ScDmc1. Assays with internal mismatch substrates for wild type and chimeric versions of (**D**) ScRad51, (**E**) hRAD51, and ScDmc1. In (**D-J**), the location and identity of the mismatched nucleotide is highlighted in red and underlined, and the error bars represent mean ± s.d.

crohomology (Figure 3J)[363]. Remarkably, ScRad51–DL1, ScRad51–DL12, hRAD51–DL1, and hRAD51–DL12 could stabilize the mismatched substrates similar to wt Dmc1 (Fig. 5.8H,I). In contrast, ScRad51–DL2 and hRAD51–DL2 were unable to stabilize internal mismatches, and instead exhibited behaviors more comparable to wt Rad51 (Fig. 5.8H,I). Moreover, ScDmc1–RL1 was unable to stabilize the mismatched substrates, and instead exhibited behavior similar to ScRad51 (Fig. 5.8J). Finally, ScRad51–DL2, ScDmc1–DL2, and hRAD51–DL2 all exhibited responses to the mismatched triplets comparable to their wt counterparts (e.g. ScRad51 and ScRad51–DL2 behaved similarly, ScDmc1 and ScDmc1–RL2 behaved similarly, as did hRAD51 and hRAD51–DL2; Fig. 5.8H–J). These findings demonstrate that Rad51 chimeras harboring L1 amino acid residues from Dmc1 attain the ability to stabilize mismatched base triplets, whereas Dmc1 chimeras harboring L1 residues from Rad51 lose the ability to stabilize mismatched triplets.

5.3.6 *C. elegans* RAD–51 behaves like "canonical" Dmc1

Some eukaryotes, such as *Caenorhabditis sp.*, have lost the DMC1 gene, although the reasons for this loss remain uncertain [380]. Inspection of RAD–51 from *Caenorhabditis sp.* revealed that the lineage–specific residues present in L1 were not the same as "canonical" Rad51 (we use the term "canonical" to identify Rad51 from species that have both recombinases), but instead more closely resembled Dmc1 (Fig. 5.9A). If our hypothesis regarding the role of L1 in mismatch stabilization is correct, then *C. elegans* RAD–51 (CeRAD–51) may stabilize mismatches. Whereas, a RAD–51 mutant, in which the "Dmc1–like" amino acids were converted to the Rad51 lineage–specific residues, might lose the ability to stabilize

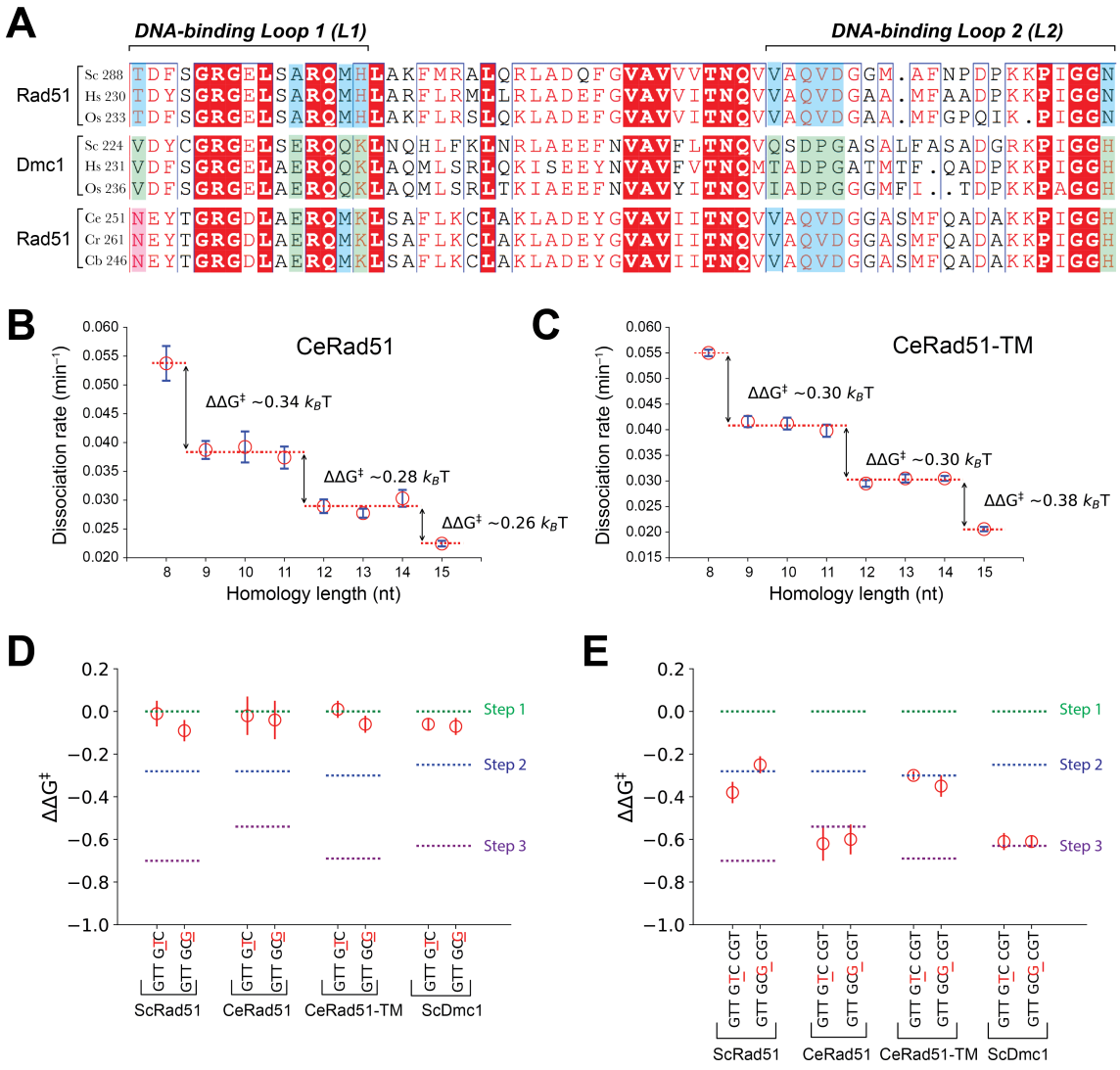

Figure 5.9: ***C. elegans* Rad51 has Dmc1-like amino acids and dsDNA-binding properties.** (**A**) Comparison of Caenorhabditis sp. Rad51 sequences (*C. elegans*, *C. remanei* and *C. brenneri*) with Rad51 and Dmc1 sequences from organisms with both recombinases. Color coding is the same as shown in Figure 5.2A. Base triplet stepping data for (**B**) wt CeRad51 and (**C**) the CeRad51 triple mutant. (**D**) Terminal mismatch triplet assays for CeRad51 and CeRad51-TM. Data for wt ScRad51 and ScDmc1 are shown for comparison and are reproduced from Figures 5.8D and F. (**E**) Internal mismatch triplet assays for CeRad51 and CeRad51-TM. Data for wt ScRad51 and ScDmc1 are shown for comparison and are reproduced from Figures 5.8H and J.

mismatches.

To test this hypothesis, we made a CeRAD–51 N251T, E261A, K265H triple mutant protein, which we refer to as CeRAD–51–TM, for brevity. Both CeRAD–51 [386] and

CeRAD–51–TM form presynaptic complexes in the ssDNA curtains with similar assembly and disassembly kinetics (Fig. 5.10C,D and Table 5.2). Assays with Atto565–dsDNA fragments confirmed that both CeRAD–51 proteins exhibited base triplet stepping (Figure 5.9B,C and 5.10E,F), and as with the other recombinases, wt and triple mutant RAD–51 were unable to stabilize mismatches present at the end of an embedded tract of microhomology (Fig. 5.9D). However, in contrast to ScRad51 and hRAD51, wt CeRAD–51 could stabilize mismatches located at an internal position within the embedded tract of microhomology (Fig. 5.9E). This finding demonstrates that CeRAD–51 does not behave like "canonical" Rad51 when presented with a mismatched substrate, rather it responds similarly to Dmc1. However, the ability to stabilize mismatched base triplets was abolished for the CeRAD–51–TM (Fig. 5.9E). Remarkably, CeRAD–51 could also promote D–loop formation with mismatched substrates (32% sequence divergence), albeit at low efficiency, whereas CeRAD–51–TM lacks this activity (Fig. 5.10G,H). Together, these findings provide additional support for the premise that lineage–specific L1 residues in Dmc1 confer the ability to stabilize mismatched base triplets during DNA strand exchange.

5.3.7 Three Dmc1 L1 amino acid residues contribute to DNA mismatch stabilization

ScRad51–DL1 has a total of six amino acid residues from ScDmc1 (Tables 5.1, 5.2, and 5.3), and we sought to determine whether a smaller subset of these residues could confer the ability to stabilize mismatches. We focused on the ScRad51 chimeric point mutants T288V, A298E, M301Q, and H302K because these residues were most conserved across the Rad51

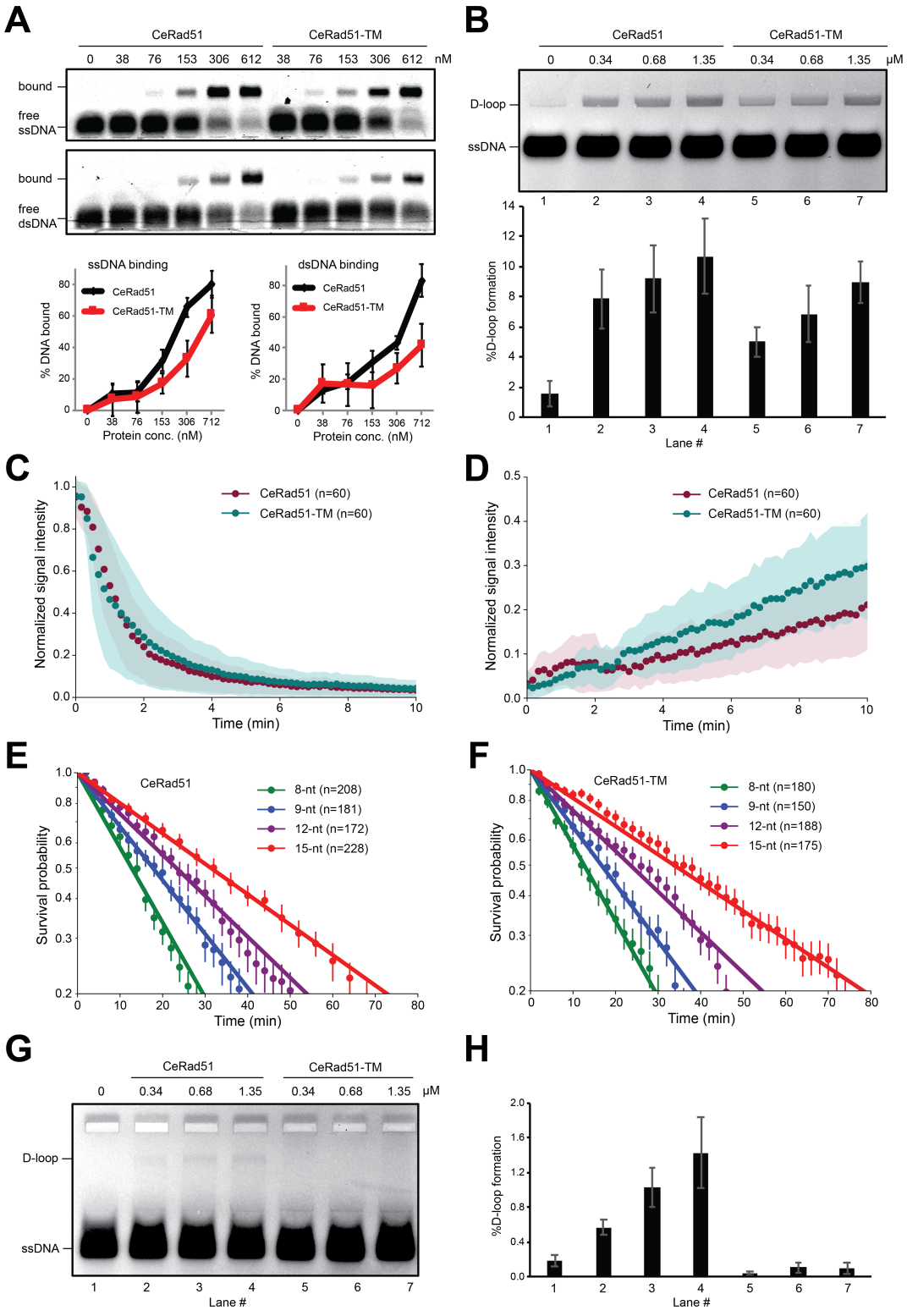

Figure 5.10: **Biochemical and single molecule analysis of CeRad51 proteins.** (**A**) ssDNA and dsDNA binding assays an quantitation for wt CeRad51 and CeRad51-TM; error bars represent s.d. from three separate measurements. (**B**) D-loop formation assays and quantitation for wt CeRad51 and CeRad51-TM; error bars represent s.d. from three separate measurements. (**C**)

Figure 5.10: (cont.) Assembly and (**D**) disassembly kinetics for CeRad51 and CeRad51-TM filaments obtained from ssDNA curtain assays; error bars represent s.d. cacluated from the indicated number of single ssDNA molecules. Corresponding rate data are presented in Table 5.3. Survival probability plots from base triplet stepping assays for (**E**) wt CeRad51 and (**F**) CeRad51-TM. (**G**) D-loop formation assay and (**H**) corresponding quantitation for a 90-nt D-loop substrate harboring 29 mutations, corresponding to 32% sequence divergence with the dsDNA plasmid substrate. Error bars in (H) represent s.d. from three separate measurements.

or Dmc1 lineages (Fig. 5.2A). The resulting mutants displayed filament assembly and disassembly kinetics that closely resembled the wt protein (Table 5.4). Each mutant exhibited dissociation rates for the Atto565–dsDNA fragments harboring the 9–nt, 12–nt, and 15–nt tracts of microhomology, which were essentially indistinguishable from wt ScRad51 (*not shown*). As with all recombinases, the point mutants were unable to stabilize mismatches present at the terminal position of an embedded tract of microhomology (Fig. 5.11A–C). ScRad51 T288V behaved like wt ScRad51 as it was able to step–over the internal mismatches but did not stabilize the mismatch (Fig. 5.11B,D). However, ScRad51 M301Q stabilized mismatches, regardless of the relative position of the mismatch within the base triplet, exhibiting behavior comparable to wt ScDmc1 (Fig. 5.11D). Interestingly, both ScRad51 A298E and ScRad51 H302K mutants were able to stabilize an internal mismatch located at the edge of a base triplet (nt position 12 in Fig. 5.11B), but neither mutant could stabilize an internal mismatch located at the center of the triplet (nt position 11 in Fig. 5.11B; Fig. 5.11D); these properties were confirmed using an alternative set of dsDNA substrates targeted to a different region of the presynaptic ssDNA (*not shown*)[363]. However, a ScRad51 A298 H302K double mutant (Table 5.4) could stabilize internal mismatched base triplets regardless of whether the mismatch was located at the center or edge of the triplet (Fig. 5.11F). These results suggest that ScRad51 L1 amino acids E298, Q301, H302, when mutated to the

corresponding Dmc1 L1 residues, can all contribute to mismatch stabilization.

5.3.8 Genetic characteristics of ScRad51 chimeras

A major challenge in understanding why eukaryotes have two recombinases is that Rad51 and Dmc1 cannot simply be replaced for one another because of the many mitotic– and meiotic–specific cofactors necessary for each of their *in vivo* functions, respectively [141, 380]. However, the chimeric proteins described here offer the opportunity to examine potential benefits of the dual recombinase system within the context of mutants that have a well–defined biochemical characteristic, namely, the ability or inability to stabilize mismatched recombination intermediates.

For genetic testing of the chimeric ScRad51 proteins, we constructed *S. cerevisiae* strains wherein the chromosomal RAD51 gene was replaced with *rad51* mutants containing Dmc1 lineage–specific amino acids. Western blot analysis confirmed that all proteins were expressed at similar levels (*not shown*). In addition, wt ScRad51, ScRad51–DL1, ScRad51 A298E and ScRad51 M301Q all supported YFP–Rad54 DNA repair focus formation upon exposure to ionizing radiation (Fig. 5.12A,B). Since the assembly of DNA repair–specific Rad54 foci is known to be dependent upon Rad51–ssDNA filaments [103], these results provide evidence that the mutant Rad51 proteins could form presynaptic filaments *in vivo*. In contrast, ScRad51 H302K and the ScRad51 A298E H302K double mutant were unable to support YFP–Rad54 focus formation, suggesting that these mutants were defective in presynaptic filament assembly in cells (Fig. 5.12B). We next asked whether ScRad51 chimeras could support cell growth on media containing the DNA–damaging agents methyl methane-

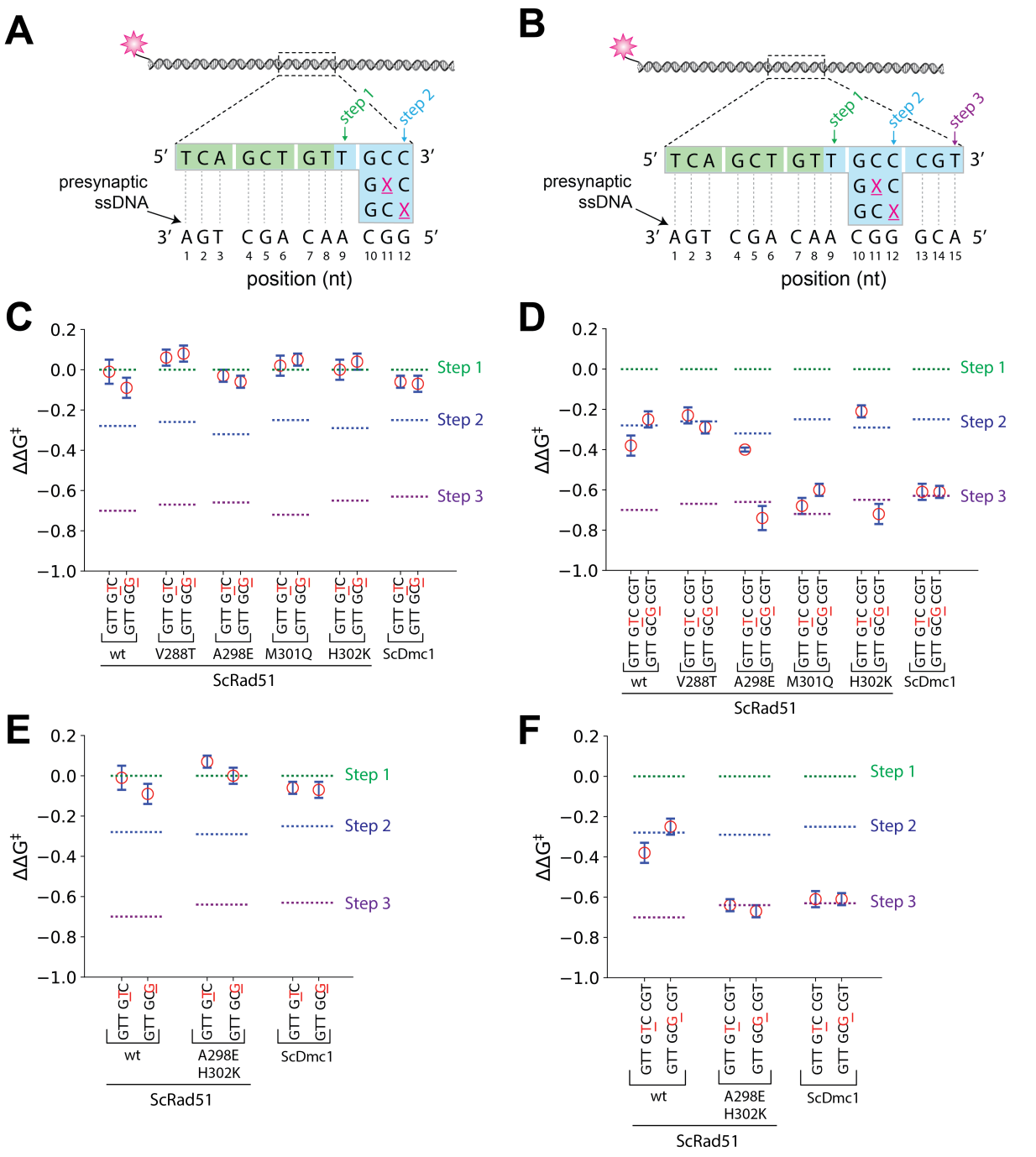

Figure 5.11: **Identification of Dmc1 lineage-specific amino acids involved in stabilizing mismatches.** Schematic illustrations of dsDNA sequences used to analyze the effects of (**A**) terminal and (**B**) internal mismatched triplets. (**C**)Terminal mismatch assays for the ScRad51 point mutants (V288T, A298E, M301Q and H302K). (**D**) Internal mismatch assays for the ScRad51 point mutants. (**E**) Terminal mismatch assays and (**F**) internal mismatch assays for the ScRad51 double mutant A298E H302K. In (C-D), data for wt ScRad51 and ScDmc1 are reproduced from Figures 5.8D and F, and in (E-F) Data for wt ScRad51 and ScDmc1 are shown for comparison and are reproduced from Figures 5.8H and J, for comparison.

sulfonate (MMS) or zeocin (Fig. 5.12C). Importantly, strains expressing either ScRad51 A298E or ScRad51 M301Q exhibited near wild–type levels of cell growth in the presence of MMS or zeocin (Fig. 5.12B), demonstrating that these point mutants retain function *in vivo*. As expected, ScRad51 H302K and ScRad51 A298E H302K were compromised for growth on plates with MMS or zeocin (Fig. 5.12C). Interestingly, ScRad51–DL1 was also compromised for growth on MMS or zeocin plates, even though this chimera was functional *in vitro* and supported Rad54 focus formation (Fig. 5.12B). We speculate that swapping the entire Rad51 L1 motif may hinder some downstream step in the HR pathway.

5.3.9 Lineage–specific amino acid residues contribute to recombination fidelity

We used a modified version of the mating¬ type (MAT) switching assay to determine whether lineage–specific L1 residues might contribute to recombination fidelity. In this assay, the mating type locus (*MAT*) is cleaved by the HO endonuclease, and Rad51–mediated recombination takes place between the cleaved mating type locus and either the *HMRa* or *HML*α donor locus [291]. Strand invasion initiates from the Z box within the *MAT* locus, which is homologous to sequences in the donor loci (Fig. 5.13A)[291]. The cleaved intermediates can be repaired by HR or non–homologous end–joining (NHEJ), and the identity of the resulting products can be defined by Southern Blot analysis (Fig.5.13A,B). To examine the effects of mismatches on recombination, we introduced point mutations at every 8th, 7th, 6th, 5th or 4th position within the Z box, corresponding to 12.5, 14.2, 16.7, 20 or 25% sequence divergence between donor and acceptor loci, respectively (Fig. 5.13A and Table

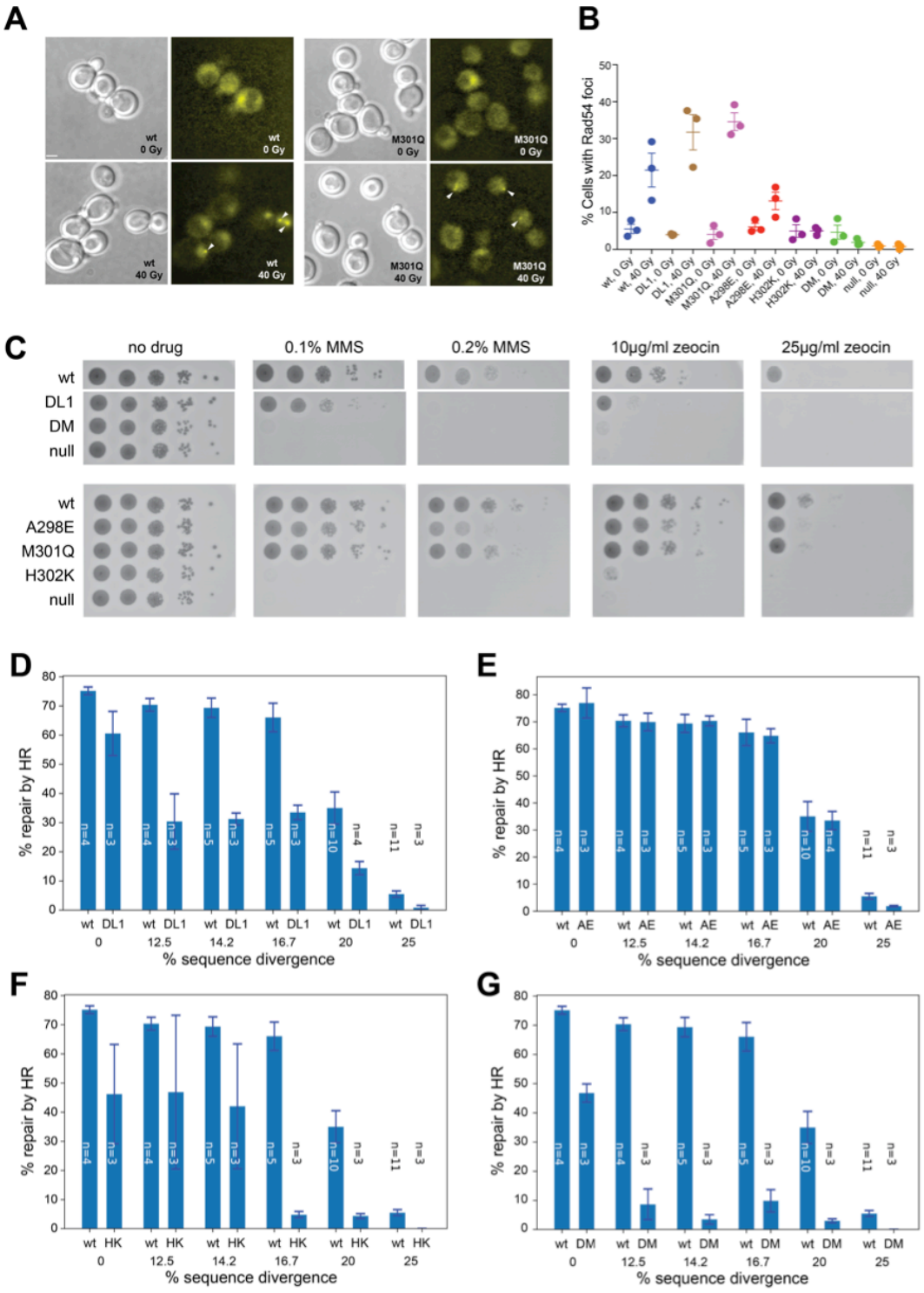

Figure 5.12: ***In vivo* characterization of chimeric *S. cerevisiae* Rad51 mutants.** (**A**) Rad54–YFP focus formation assays; white arrowheads highlight examples of Rad54–YFP; and (**B**) focus quantitation, (**C**) MMS and zeocin resistance assays, and MAT switching assays for (**D**) ScRad51–DL1, (**E**) ScRad51 A298E, (**F**) ScRad51 H302K and (**G**) the ScRad51 A298E H302K double mutant (abbreviated as DM), as indicated. Each bar in panels D–G represents the mean ± s.d. from n independent experiments, as indicated.

5.5).

For the strain expressing wt ScRad51, MAT switching was remarkably tolerant of mismatches (Fig. 5.13C). For instance, with the wt templates (0% divergence) 75.1 ± 1.33% of the repair products could be attributed to HR–mediated repair, while 66.1 ± 4.75% of the repair products could be attributed to HR for templates with 16.7% sequence divergence (Fig. 5.13C). However, HR efficiency drops markedly for templates with mismatches at every 5th (20% divergence) and 4th position (25% divergence), yielding values of 35.0 ± 4.75% and 5.52 ± 1.46%, respectively (Fig. 5.13C). These observations are in good agreement with previous reports indicating that wt ScRad51 supports efficient break–induced replication (BIR) for templates with similar levels of sequence divergence [378].

Consistent with the MMS and zeocin resistance assays, ScRad51–DL1, ScRad51 H302K, and the ScRad51 A298E H302K double mutant were all compromised for MAT switching even at 0% divergence, and these defects were exacerbated at higher levels of sequence divergence (Fig. 5.12D,F,G). In contrast, the ScRad51 M301Q point mutant supported levels of HR–mediated repair comparable to wt ScRad51 for the substrates with up to 16.7% divergence. Remarkably, at 20% sequence divergence, representing the midpoint of the curve for wt ScRad51 (Fig. 5.13B), there was a 26.7 ± 5.00% (Student t–test, $p<0.001$) increase in HR–mediated repair by ScRad51 M301Q relative to wt ScRad51, and there was a 22.2 ± 10.2% (Student t–test, $p<0.05$) increase in HR for the template with 25% sequence divergence (Fig. 5.13B,C). These findings, together with our biophysical data, demonstrate that ScRad51 M301Q, which can stabilize mismatched HR intermediates *in vitro*, also supports a higher recombination frequency for mismatched substrates *in vivo*. Interestingly, although ScRad51 A298E was functional for MAT switching, this mutant showed no dif-

ferences in HR efficiency compared to wt ScRad51 for templates with increasing sequence divergence (Fig. 5.12E). Given this outcome, it is notable that ScRad51 A298E could only stabilize mismatches located at the edge of a base triplet (Fig. 5.11B,D), and the divergent Z box sequences will always have mismatches at both the center and edge positions of the mismatched base triplets (irrespective of the frame of reference) (Table 5.5). The inability of ScRad51 A298E to stabilize mismatches located in the center of a base triplet could explain why this particular mutant does not behave like ScRad51 M301Q in the MAT switching assays.

5.4 Discussion

We have defined the structural elements that are responsible for the differential response of Rad51 and Dmc1 to mismatched base triplets. We propose that these structural and functional differences represent a fundamental distinction between the Rad51 and Dmc1 lineages of the Rad51/RecA family of DNA recombinases. Here, we discuss possible implications of these findings with respect to recombinase structures, recombination fidelity and mechanisms, and the evolution of Rad51/RecA family members.

5.4.1 Lineage–specific amino acid residues help determine recombinase behaviors

L1 and L2 DNA–binding loops of the Rad51/RecA family of DNA recombinases contain amino acids that are specifically conserved within either the Rad51 lineage, or the Dmc1 lineage, but not both. Our data demonstrate that three amino acid residues within the L1

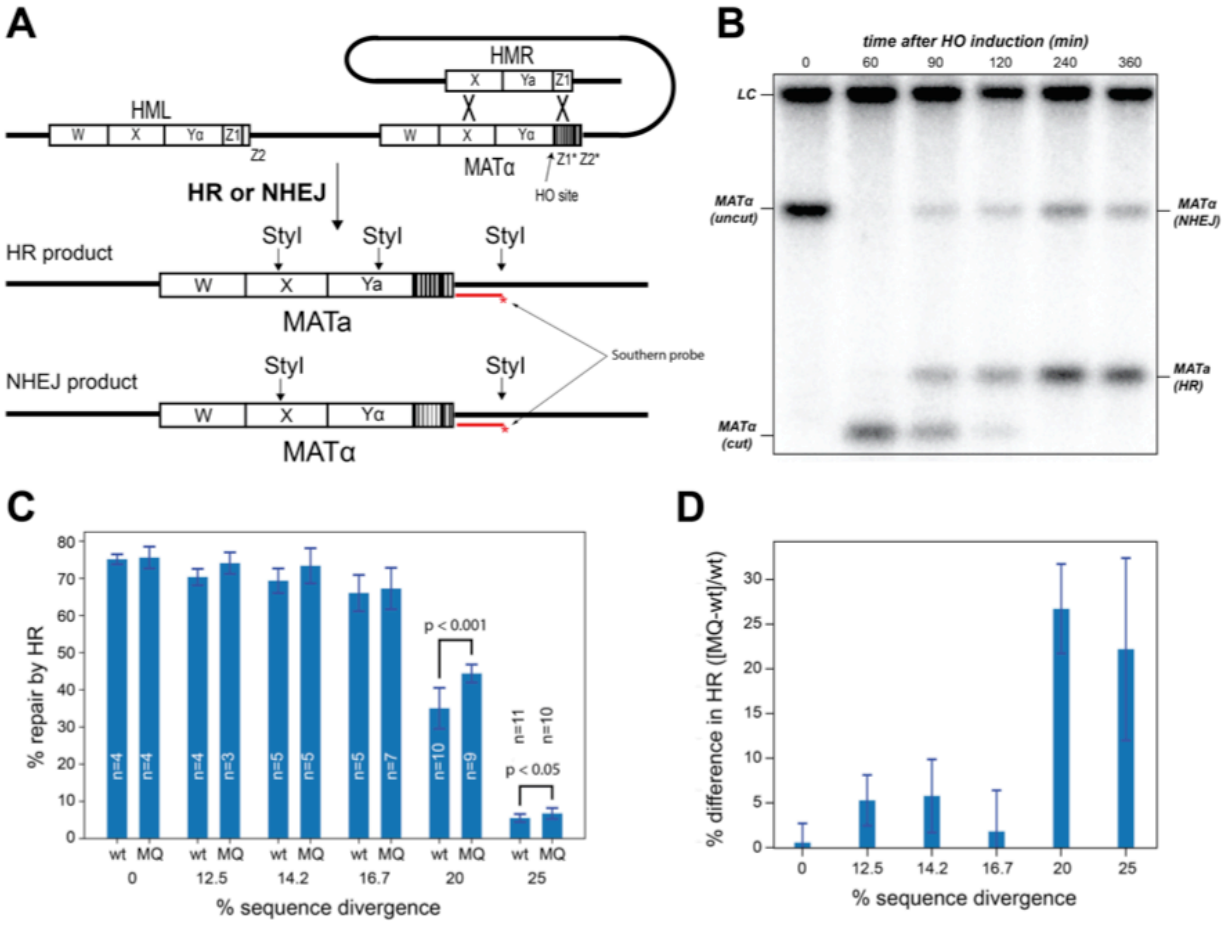

Figure 5.13: **L1 amino acid identity influences recombination between divergent DNA sequences.** (**A**) Schematic illustration of the MAT switching assay highlighting the products that arise from NHEJ and HR. (**B**) Example of a southern blot analysis of *StyI*–digested DNA after galactose induction of HO endonuclease. (**C**) Bar graph showing the fraction of repair products attributed to HR for strains expressing either wt Rad51 or Rad51 M301Q for templates with varying levels of sequence divergence. Each bar represents the mean ± s.d. from n independent experiments, as indicated. The statistical significance for the observed differences between wt and mutant for the 20 and 25% sequence divergence data sets are indicted (p values from an unpaired two–tailed Student's t test are indicated). (**D**) Bar graph illustration the percent difference for HR–mediated MAT switching between Rad51 M301Q and wt Rad51. Each bar represents the mean ± s.e.m. calculated from data in panel C.

wild-type *(0% divergence):*
ATT TTA TAA ACC CTG GTT TTG GTT TTG TAG AGT GGT TGA CGA ATA ATT ATG CTG AAG TAC GTG GTG ACG
GAT ATT GGG AAG ATG TGT TTG TAC ATT TGG CCT TAT AGA GTG TGG TCG TGG CGG AGG TTG TTT ATC TTT
CGA GTA CTG AAT GTT GTC AGT ATA GCT ATC CTA TTT GAA ACT CCC CAT CGT CTT GCT CTT GTT CCC AAT
GTT TGT TTA TAC ACT CAT ATG GCT ATA CCC TTA TCT ACT TGC CTC TTT TGT TTA TGT CTA TGT ATT TGT
ATA AAA TAT GAT ATT ACT CAG ACT CAA GCA AAC AAT CAA

1 in 8 *(12.5% divergence):*
ATT TTA TTA ACC CTG CTT TTG GTA TTG TAG ACT GGT TGA GGA ATA ATA ATG CTG ATG TAC GTG CTG ACG
GAA ATT GGG ATG ATG TGA TTG TAC AAT TGG CCT AAT AGA GTC TGG TCG TCG CGG AGG ATG TTT ATG TTT
CGA GAA CTG AAT CTT GTC AGA ATA GCT TTC CTA TTA GAA ACT CGC CAT CGT GTT GCT CTA GTT CCC ATT
GTT TGT ATA TAC ACA CAT ATG GGT ATA CCC ATA TCT ACA TGC CTC TAT TGT TTA AGT CTA TGA ATT TGT
AAA AAA TAT CAT ATT ACA CAG ACT CTA GCA AAC TAT CAA

1 in 7 *(14.2% divergence):*
ATT TTA AAA ACC CAG GTT TTC GTT TTG AAG AGT GCT TGA CGT ATA ATT TTG CTG ATG TAC GTC GTG ACG
CAT ATT GCG AAG ATC TGT TTG AAC ATT TCG CCT TAA AGA GTG AGG TCG TCG CGG AGC TTG TTT TTC TTT
CCA GTA CTC AAT GTT CTC AGT AAA GCT ATG CTA TTT CAA ACT CGC CAT CGA CTT GCT GTT GTT CGC AAT
GTA TGT TTA AAC ACT CTT ATG GCA ATA CCC ATA TCT AGT TGC CTG TTT TGT ATA TGT CAA TGT ATA TGT
ATA TAA TAT GTT ATT ACA CAG ACT GAA GCA ATC AAT CAT

1 in 6 *(16.7% divergence):*
ATT TTG TAA ACT CTG GTC TTG GTA TTG TAT AGT GGC TGA CGT ATA ATA ATG CTT AAG TAT GTG GTA ACG
GAG ATT GGT AAG ATT TGT TTC TAC ATG TGG CCA TAT AGC GTG TGC TCG TGC CGG AGC TTG TTG ATC TTC
CGA GTG CTG AAA GTT GTT AGT ATT GCT ATT CTA TTA GAA ACG CCC CAA CGT CTA GCT CTA GTT CCT AAT
GTA TGT TTG TAC ACA CAT ATA GCT ATT CCC TTC TCT ACC TGC CTA TTT TGC TTA TGA CTA TGG ATT TGC
ATA AAC TAT GAA ATT ACG CAG ACA CAA GCC AAC AAA CAA

1 in 5 *(20% divergence):*
ATT TAA TAA TCC CTC GTT TAG GTT ATG TAC AGT GCT TGA GGA ATT ATT AAG CTG TAG TAG GTG GAG ACG
CAT ATA GGG ATG ATG AGT TTC TAC AAT TGG GCT TAA AGA GAG TGG ACG TGC CGG ACG TTG ATT ATG TTT
CCA GTA GTG AAA GTT GAC AGT TTA GCA ATC CAA TTT CAA ACA CCC CTT CGT GTT GCA CTT GAT CCC TAT
GTA TGT TAA TAC TCT CAA ATG GGT ATA GCC TTT TCT AGT TGC GTC TTA TGT TAA TGT GTA TGA ATT TCT
ATA TAA TAA GAT AAT ACT GAG ACA CAA GGA AAC TAT CAT

1 in 4 *(25% divergence):*
ATT ATA TTA ACG CTG CTT TAG GTA TTG AAG ACT GGA TGA GGA AAA ATA ATG CAG AAC TAC CTG GAG
ACC GAT TTT GCG AAC ATG AGT TAG TAG ATT AGG CGT TAA AGA CTG TCG TCC TGG AGG ACG TTC TTT TTC
TAT CGT GTA GTG ATT GTA GTC TGT AAA GCA ATC GTA TAT GAT ACT GCC CTT CGA CTT CCT CAT GTA CCC
TAT GAT TGA TTA AAC AGT CAA ATG CCT AAA CCG TTA ACT AGT TGG CTC ATT TCT TTT TGT GTA TCT ATA
TGT TTA ATA TAA GAT TTT AGT CAC ACT GAA GGA AAG AAT GAA

Table 5.5: **MAT switching Z box templates.**

DNA–binding loop regulate the response of Rad51 and Dmc1 to mismatched base triplets. Mutations that swap the identity of these residues alter the response of the respective recombinase to match that of its paralog, and this behavior appears to be conserved between recombinases from both *S. cerevisiae* and humans. We speculate that within the Rad51 lineage, these residues may help enhance recombination fidelity, whereas identity of these residues within the Dmc1 lineage may contribute meiosis, perhaps by allowing for more favorable recombination between polymorphic parental alleles. Indeed, a single point mutation in Rad51 (M301Q) is sufficient to recapitulate Dmc1–like mismatch stabilization *in vitro* and also enhances recombination between divergent sequence *in vivo*. However, we also note that the precise spatial geometry of the L1 DNA–binding loop is likely to be influenced by its native context, in particular its connectivity to rest of the recombinase core domain. Thus, it is possible, that the mismatch tolerance observed for Rad51 M301Q may not exactly match that of Dmc1. Indeed, Dmc1 itself may prove to be even more mismatch–tolerant in vivo than the Rad51 chimeras harboring Dmc1 residues. Although the natural prevalence of sequence polymorphisms between parental alleles would not approach the high levels of sequence divergence in our MAT switching assays, we anticipate that more subtle differences in recombination efficiency involving templates with lower divergence may be important when considered over long evolutionary time scales. Moreover, MAT switching is a highly efficient HR–dependent process that is greatly facilitated by local chromosomal architecture [291], which may make MAT switching inherently more tolerant to mismatches. It is possible that less efficient HR–mediated repair events could exhibit even more pronounced effects at lower levels of sequence divergence.

Interestingly, ScRad51–DL2 retains basic biochemical functions (Fig. 5.3, 5.4, 5.5, and

5.8 and Table 5.4), but does not support Rad54 focus formation or allow for growth on MMS or zeocin plates (*not shown*). The disparity between the *in vivo* and *in vitro* activities of this chimera suggest that lineage–specific amino acids in L2 may be important for assembly of the native presynaptic complex. Finally, in addition to the L1 and L2 amino acids described in this study, we have also identified ∼19 lineage–specific amino acid residues present in other regions of recombinases (*not shown*). We anticipate that some of these residues may mediate Rad51– and Dmc1–specific protein–protein contacts. Exploring the roles of these lineage–specific amino acids may yield further insights into the differential properties of the two eukaryotic recombinases.

5.4.2 Recombination between divergent sequences

Rad51/RecA family recombinases require at least consecutive 8 nucleotides of microhomology for efficient recognition of short dsDNA substrates *in vitro* [306, 330, 375, 387]. However, the efficiency of BIR [378] and MAT switching (this study) indicate that the *in vivo* requirements for donor DNA recognition and strand invasion are significantly less stringent. Two factors may account for these differences. First, *in vivo* recombination requires many other proteins, a number of which may influence the activities of Rad51 [356, 379]. Second, long substrates likely allow for multiple, simultaneous interactions, the cumulative effects of which may circumvent the kinetic benefits observed *in vitro* for shorter substrates bearing a single contiguous tract of homology. For example, the Z box with mismatches at every 8th position has 40 adjacent 7–nt tracts of microhomology separated from one another by single mismatches, and the Z box with mismatches at every 6th position has 52 adjacent 5–nt tracts

of microhomology (Table 5.5). At present, the length and complex sequence composition of these substrates preclude detailed biophysical analysis in ssDNA curtain assays, nor can we recapitulate the protein composition of a native presynaptic complex *in vitro*. Future work will be necessary to understand how these parameters influence the efficiency and kinetics of HR reaction mechanisms.

Our data reveal that the chimeric ScRad51 M301Q mutant, which stabilizes mismatches in biophysical assays, can also support more efficient HR–mediated repair during MAT switching for a Z box bearing ≥20% sequence divergence. It is interesting to note that for wt ScRad51, we observe the largest decline in HR efficiency only for templates with ≥20% divergence (Fig. 5.13C). Inspection of these DNA sequences reveals an interesting feature that coincides with this dramatic change in HR efficiency. Namely, for templates ranging up to 16.7% sequence divergence, the mismatched base triplets are never adjacent to one another, but this is not true for the templates with ≥20% sequence divergence, which instead are comprised of repeating patterns where there are always two or three mismatched triplets adjacent to one another (Table 5.5). One possible explanation for marked decline in HR efficiency at ≥20% sequence divergence is that Rad51 may have particular difficulties in promoting recombination when two or more adjacent base triplets contain mismatches, whereas Dmc1, or ScRad51 harboring Dmc1 L1 amino acids, may be more tolerant of these substrates.

5.4.3 Potential mechanisms of DNA mismatch stabilization

Three lineage–specific residues clustered together within L1 DNA–binding loop contribute mismatch stabilization by Dmc1. We proposed that these amino acids may allow Dmc1 to make compensatory DNA contacts that are independent of Watson–Crick pairing interactions (Fig. 5.14A). Consistent with this hypothesis, molecular dynamics (MD) simulations of wt hRAD51 and an hRAD51 chimeric mutant bearing three hDMC1 residues in the L1 domain, corresponding to the mutations A241E, M244K and H245K (numbering based on hRAD51; Fig. 5.2A) reveal that the introduction of these Dmc1–specific residues results in more intimate contacts between L1 and the phosphate backbone of the complementary strand of the postsynaptic complex (Fig. 5.14B,C). The notion that Dmc1 residues contact the phosphate backbone of the nascent DNA joint is also consistent with the observation that Dmc1 can stabilize base triplets bearing abasic sites [381]. At present, these simulations must be interpreted with caution, given that they are based upon structural data that lacks sufficient information to fully describe the relevant protein–nucleic acid interfaces. In the case of hDMC1, the crystal structures lack DNA [189]. For hRAD51, existing cryo–EM structures lack sufficient resolution to accurately define the L1 protein–DNA interface [39, 388]. Similarly, although there is a crystal structure of ScRad51 bound to DNA, the L1 and L2 contacts are not visible [37]. Finally, there is as yet no structural information regarding how recombinases interact with a DNA joint that harbors mismatches, nor do our simulations take into account the potential influence of mismatches on recombinase–DNA interactions. Regardless, the MD simulations, together the observation that Dmc1 can stabilize abasic sites [381], are consistent with the general notion that the lineage–specific amino

acid residues responsible for the differential responses of Rad51 and Dmc1 to mismatches are positioned to interact with the phosphate backbone of the complementary DNA strand within the nascent DNA joint.

5.4.4 What is the mechanistic impact of mismatches on recombination?

We can envision at least two general mechanisms by which DNA mismatches could affect recombination: by altering the intrinsic stability of the heteroduplex DNA intermediates; or by making these intermediates more susceptible to disruption by regulatory enzymes. Importantly, although a single mismatch impacts the binding lifetimes of dsDNA fragments in our biophysical assays, on the whole, these are all still relatively long–lived intermediates (e.g. for ScRad51, a 70–bp dsDNA fragment with 12–nt of microhomolgy has a lifetime of ∼33 minutes, increasing the microhomology length to 15–nts yields a lifetime of ∼47 minutes, and introduction of a single mismatch within the 15–nt tract of microhomology reduces the lifetime to ∼35 minutes). This conclusion is also consistent with studies demonstrating that bacterial RecA is surprisingly tolerant of mismatches *in vitro* [375]. We cannot yet predict how our *in vitro* observations will scale for the longer *in vivo* substrates, nor can we measure the equivalent biophysical parameters *in vitro* with longer substrates. Nevertheless, the available biophysical data imply that mismatched HR intermediates are not intrinsically unstable. Thus, we favor the hypothesis that mismatches may render HR intermediates more susceptible to disruption by regulatory enzymes. Several proteins are known to dissociate HR intermediates, including the *S. cerevisiae* helicases Srs2 and Sgs1 (BLM in humans), and

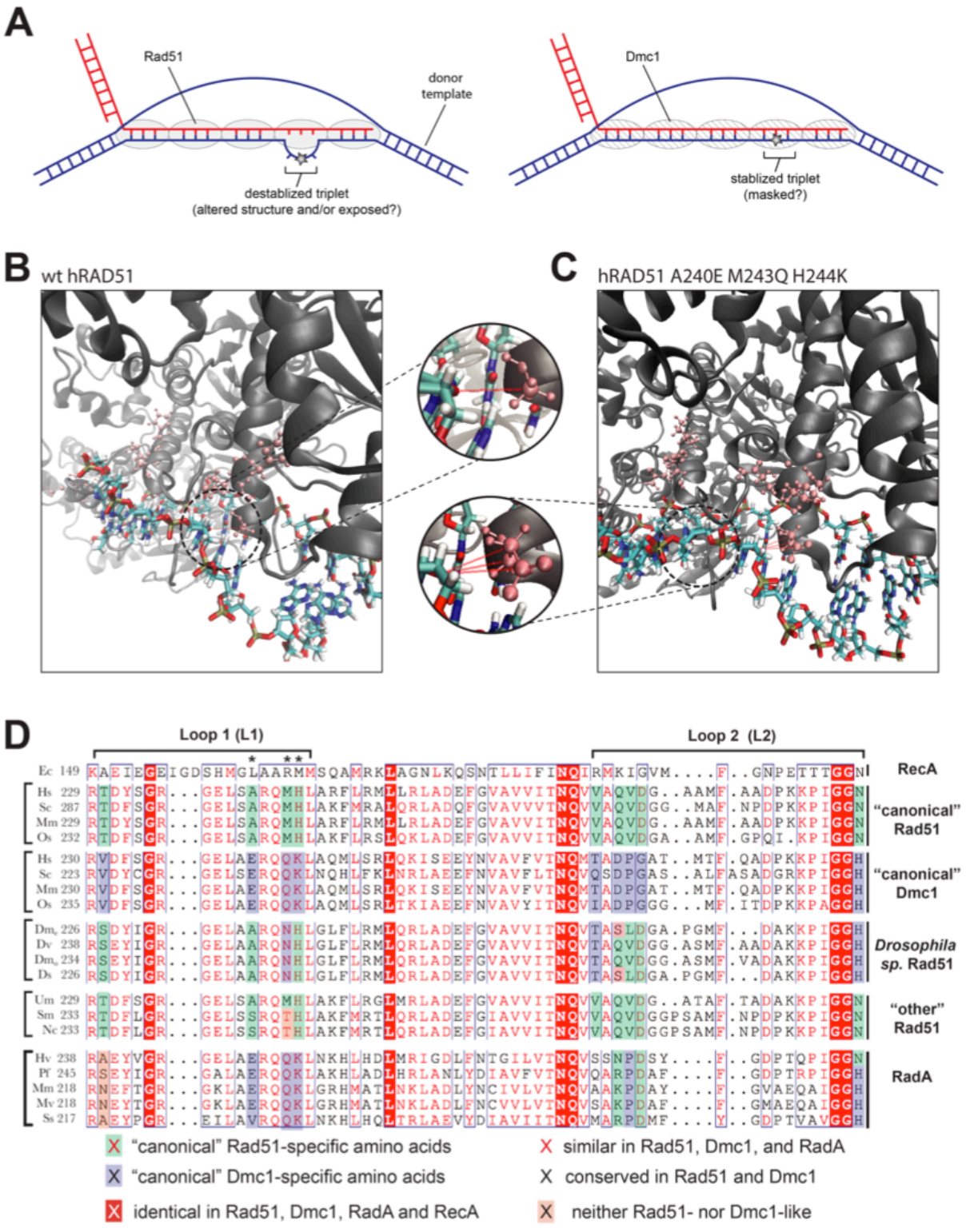

Figure 5.14: **Potential mechanism of mismatch stabilization and L1 conservation among different Rad51/RecA family members.** (**A**) Model for differences between Rad51 and Dmc1 interactions involving imperfectly paired HR intermediates. (**B**) Snapshots taken from MD

Figure 5.14: (cont.) simulations of hRAD51 and (**C**) hRAD51 harboring three Dmc1 lineage–specific amino acids substitutions (A240E M243Q H244K), suggesting that the Dmc1 amino acids are better positioned to contact the incoming complementary DNA strand. Insets highlight potential protein contacts (red dashed lines) with the ribose ring of the phosphate backbone. (**D**) Comparison of L1 and L2 sequences from: E. coli RecA; "canonical" Rad51 and Dmc1, from organisms harboring both recombinases; Rad51 from four *Drosophilia sp.* (*D. melanogaster*, *D. virilis*, *D. mojavenis* and *D. simulans*); Rad51 sequences designated as "other" from *U. maydis*, *S. macrospora*, and *N. crassa*; and RadA from *H. volcanii*, *Pyrococcus furiousus*, *Methanococcus maripaludis*, *Methanococcus voltae*, and *sulfolobus solfataricus*. Asterisks denote the amino acids that contribute to mismatch stabilization for Dmc1.

the post–replicative mismatch repair (MMR) machinery also plays a role in minimizing HR between divergent sequences [389–393]. One possibility is that these enzymes may recognize some distinct mismatch–dependent structural feature that enables them to more readily disrupt Rad51–bound intermediates, whereas Dmc1 may shield mismatched intermediates from these enzymes (Fig. 5.14A).

5.4.5 *C. elegans* RAD–51 and other "non–canonical" recombinases

Not all eukaryotes follow the typical two recombinases paradigm. *Ecdysozoans*, such as *Caenorhabditis sp.* and *Drosophila sp.*, possess RAD51 but have lost DMC1 as well as genes encoding core meiotic proteins necessary for Dmc1 activity (e.g. Hop2/Mnd1) [355, 358, 383, 394, 395]. It remains unknown why these organisms have lost DMC1, and the evolutionary implications of this loss remain unexplored. However, *Caenorhabditis sp.* and *Drosophila sp.*, have among the highest rates of evolution for the RAD51 genes [358]. Interestingly, the lineage–specific amino acid residues within the L1 DNA–binding loop of RAD–51 from *Caenorhabditis sp.* have adapted to more closely match Dmc1, and our

findings show that CeRAD–51 can stabilize mismatched base triplets *in vitro*. Thus, the loss of DMC1 may have caused, or allowed for, adaptation of CeRAD–51 L1 to become more "Dmc1–like". Our findings suggest that CeRAD–51 may have inherently lower fidelity than "canonical" Rad51 proteins from other organisms. Consistent with this idea, recent studies have shown that recombination can occur in *C. elegans* meiosis between sequences with very limited sequence homology [396]. Interestingly, *Drosophila sp.* Rad51 L1 amino acids closely match those found in canonical Rad51 (Fig. 5.14C). However, N301 (numbering based on *S. cerevisiae* Rad51 for comparison) in *Drosophila* Rad51 does not match the methionine found in canonical Rad51 but is instead more similar to the glutamine from Dmc1 (Fig. 5.14C). Thus, *Drosophila* Rad51 may also behave like Dmc1 with respect to mismatches.

Among organisms that have Dmc1, chromosome pairing is initiated through recombination–dependent mechanisms [141, 355, 380]. In contrast, both *C. elegans* and *D. melanogaster* have evolved alternative, recombination–independent mechanisms for initiating chromosome pairing during meiosis [395, 397]. However, use of alternative pairing mechanisms is not necessarily the defining attribute of organisms lacking Dmc1. For instance, *Ustilago maydis*, *Sodaria macrospora* and *Neurospora crassa* have all lost the *DMC1* gene, but require recombination to initiate chromosome pairing during meiosis [398]. Interestingly, *U. maydis*, *S. macrospora* and *N. crassa* all possess canonical Rad51 lineage–specific amino acids in the L1 DNA–binding loop (Fig. 5.14C). Finally, it should be noted that the L1 (and L2) amino acids from bacterial RecA are highly divergent from the eukaryotic recombinases (Fig. 5.2A and 5.14C). Therefore, the mechanisms by which RecA interacts with the DNA intermediates may differ from those that define the action of the eukaryotic recombinases.

5.4.6 Potential origins of the dual–recombinase paradigm

Rad51 and Dmc1 arose early in eukaryotic evolution from a gene duplication event involving an ancestral Archaeal RadA recombinase, and this gene duplication event may have coincided with or allowed for the emergence of meiosis and sexual reproduction [358, 383, 394]. Some extant Archaea, such as *Haloferax volcanii*, undergo a conjugation process involving the exchange of highly divergent genetic information via HR, and this conjugation processes bears some resemblance to the eukaryotic meiotic program [399–404]. Intriguingly, RadA L1 DNA–binding loop amino acid residues located at the positions involved in mismatch stabilization are identical to those found in the Dmc1 lineage of the Rad51/RecA family (Fig. 5.14E). One speculative possibility is that a "lower–fidelity" recombinase (i.e. more "Dmc1–like") may have predated the "higher–fidelity" Rad51 present in modern eukaryotic lineages, and the emergence of Rad51 may have allowed eukaryotes to take better advantage of HR as a high–fidelity DNA repair pathway.

5.5 Experimental Procedures

5.5.1 Sequence and crystal structure alignments

Crystal structures of E. coli RecA (1CMX) [36], ScRad51 (1SZP) [37], hDMC1 [189] and hRAD51 (5HB1)[39] were obtained from the RCSB PDB and aligned in MacPyMOL [405]. All protein sequences were obtained from the NCBI database and sequence alignments were performed using NIH COBALT [406]. Most of the sequenced RAD51 and DMC1 genes have not been experimentally validated, therefore to help ensure the alignments reflected

information from bonafide Rad51 and Dmc1 proteins, we restricted our analysis to proteins that were 200 amino acids in length and also contained the highly conserved L1 motif GRGEL (or GRGDL; corresponding to amino acids 294-298 in ScRad51). Sequences that did not fulfill these two criteria were excluded from the analysis. Aligned sequences were further analyzed for common features and annotated using ESPript 3.0 [407].

5.5.2 Single molecule dsDNA binding assays

All single molecule dsDNA binding experiments were performed as previously described [306, 363, 381]. In brief, presynaptic filaments were assembled, as described above, by injecting 2 µM of the indicated recombinase in the presence of 2 mM ATP, followed by a 20-minute incubation at 30 °C. Free protein was then flushed from the sample chamber, followed by an injection of HR buffer (as indicated above for each recombinase) containing 10 nM to 10 µM dsDNA oligonucleotide substrate (as indicated), and the reactions were incubated for an additional 10 minutes. Unbound dsDNA was then quickly flushed from the sample chamber using a 30 second wash at 1 ml/min, flow was then reduced to 0.2 ml/min and images (90 millisecond integration) were collected at 30- to 160-second intervals for 1 to 2 hours, as previously described [306, 380, 381]. The data collection intervals were optimized relative to the overall lifetime of each dsDNA substrate and the laser was shuttered between acquired images to minimize photo-bleaching. Kymographs were then generated from the resulting movies using Fiji. Survival probabilities were determined from analysis of the resulting kymographs by measuring the time (dwell time) that each molecule of Atto565–DNA remained bound to the presynaptic complexes after flushing the unbound

DNA from the sample chamber. Error bars for the survival probability measurements and binding distributions represent 70% confidence intervals obtained through bootstrap analysis, providing a close approximation of expectations for one standard deviation from the mean [306, 363]. All reported $\Delta\Delta G^{\ddagger}$ values were calculated from the dissociation rate data for the Atto565–dsDNA substrates, as described [306, 363].

5.5.3 ScRad51, ScDmc1, hRAD51 and hDMC1 purification and characterization

All recombinant yeast and human Rad51 proteins were expressed in E. coli and purified as previously described [342, 408, 409]. For ATPase assays, hRAD51 was assayed as previously described [410] with some minor modifications. Reactions (10 µl) were performed with 4 µM hRAD51 incubated with 50 µM ATP supplemented with 0.05 μCi/μl [γ-32P]-ATP in buffer containing 35 mM Tris-HCl [pH 7.5], 1 mM DTT, 1 mM $MgCl_2$, 100 ng/µl BSA, 100 mM KCl, and ϕX174 viral DNA [45 μM nucleotides] at 37 $^\circ$C. Aliquots (2 μl) were removed at the indicated time points, and terminated by the addition of 2 μl 500 mM EDTA. Reaction products were spotted onto TLC cellulose PEI plates (Select Scientific) and developed with 0.5 M LiCl and 0.5 M formic acid. ScRad51 and ScDmc1 were assayed for ATP hydrolysis as described [342] with minor modifications. Reactions (10 μl) contained 3 μM recombinase and were performed in buffer containing 125 μM ATP supplemented with 0.05 μCi/μl [γ-32P]-ATP in 35 mM Tris-HCl [pH 7.5], 1 mM DTT, 1.5 mM $MgCl_2$, 100 ng/µl BSA, 100 mM KCl, and ϕX174 viral DNA [45 μM nucleotides] at 37 $^\circ$C. Reactions were terminated using 500 mM EDTA and products were resolved by TLC chromatography,

as above. Strand exchange assays with hRAD51 assays were performed as described [411]. Reactions contained 10 nM 150-mer ssDNA, 10 nM [α-32P]-labeled 40- bp dsDNA, and either 0.43, 0.86, or 1.3 μM hRAD51, as indicated. ScRad51 and ScDmc1 strand exchange assays were also performed as previously described [342]. Reactions contained 40nM 150-mer ssDNA, 40-nM [γ-32P]-labeled 40-bp dsDNA; and either 1.4, 2.9, or 5.8 μM recombinase, as indicted.

5.5.4 CeRad51 expression and purification

Wt CeRad51 and CeRad51-TM were expressed and purified essentially as described previously [386]. Briefly, CeRad51 was expressed using the Champion pET-SUMO system (Life Technologies) in BL21(DE3) One Shot *E. coli* in LB supplemented with 50 μg/ml kanamycin at 37 $^\circ$C, before induction for 4 h with 1 mM IPTG at 30 $^\circ$C. Pellets were resuspended in 400 ml ice cold Lysis Buffer (50 mM potassium phosphate [pH 7.8], 1 M KCl, 10% glycerol) supplemented with cOmplete, EDTA-free protease inhibitor cocktail tablets (Roche). All further steps were carried out at 4 $^\circ$C. Triton X-100 was added to 0.1% and cells were sonicated. Lysate was cleared using a Ti45 rotor (Beckman Coulter) at 40,000 rpm for 60 min. Imidazole was added to the supernatant to a final concentration of 25 mM and applied to Ni-NTA agarose affinity gel (Qiagen, Cat No. 30210) which had been pre-washed with Binding Buffer (50 mM potassium phosphate [pH 7.8], 1 M KCl, 10% glycerol, 25 mM imidazole [pH 7.5]). The protein was bound to the beads by rotating for 2 h. Beads were washed with Binding Buffer and Binding Buffer containing 50 mM imidazole. The protein was eluted with Binding Buffer containing 200 mM imidazole and dialyzed against

Dialysis Buffer (20 mM Tris- HCl [pH 8.0], 300 mM KCl, 10% glycerol) overnight using 10 kDa MWCO SnakeSkin dialysis tubing (Thermo Scientific). The His-SUMO tag was cleaved to yield native CeRad51 by addition of His-tagged Ulp1 SUMO protease for 45 min. The protein was centrifuged and the soluble fraction collected and bound Ni-NTA agarose affinity gel to remove the SUMO protease and His-SUMO tag. The flowthrough containing native CeRad51 was collected and diluted at 1:1 ratio with Dilution Buffer (20 mM Tris-HCl [pH 8.0], 10% glycerol, 2 mM EDTA, 1 mM DTT) to reduce salt concentration to 150 mM KCl. The protein was bound to a 1 ml Mono Q 5/50 GL column (GE Healthcare) at washed with R buffer supplemented with 150 mM KCl. The protein was eluted with a gradient of 150-640 mM KCl in R buffer. The peak fractions were pooled and concentrated and frozen directly in the elution buffer.

5.5.5 CeRad51 D-loop formation assay

Proteins were diluted in Dilution Buffer (25 mM Tris-HCl [pH 7.5], 10% (v/v) glycerol, 0.5 mM EDTA [pH 7.5], 100 mM KCl, 1 mM DTT, 0.01% NP40). To start the reaction, proteins were incubated with a master mix (35 mM Tris, pH 7.5, 50 mM KCl, 1 mM DTT, 2 mM MgCl2, 2 mM ATP) and 30 nM of either 5´-FITC labelled 90mer (AAA TCA ATC TAA AGT ATA TAT GAG TAA ACT TGG TCT GAC AGT TAC CAA TGC TTA ATC AGT GAG GCA CCT ATC TCA GCG ATC TGT CTA TTT) or mismatched 5'-FITC labelled 90mer (TTT TCT TTC TTT TGT TTT TTT GTG TTT TCT TGG TCT GTC TGT TTC CTT TGC TTT TTC TGT GTG GCT CCT TTC TCT GCG TTC TGT CTT TTT; mismatches underlined) oligonucleotide at 25 ° C for 10 min; pBluescript SK(–) (540

ng in 2 l) was then added to bring the final reaction volume to 10 l and incubated for further 10 minutes. The samples were deproteinized with 0.1% SDS and 10 g proteinase K for 10 min at 37 ° C and resolved in 0.8% agarose gels in 1X TAE (90 V, 35 min). Gels were imaged on a Typhoon FLA 9500 scanner (GE Healthcare) and quantified using ImageJ software.

5.5.6 CeRad51 ectrophoretic mobility shift assay

Proteins were diluted in Dilution Buffer (25 mM Tris-HCl pH 7.5, 10% (v/v) glycerol, 0.5 mM EDTA, pH 7.5, 100 mM KCl, 1 mM DTT, 0.01% NP40). To start the reaction, proteins were incubated with M buffer: 20 mM Tris, pH 7.5, 10% (v/v) glycerol, 50 mM sodium acetate, 1 mM DTT, 2 mM MgCl2, 2 mM ATP and 25 nM 5´ FITC labelled 49mer ssDNA (AGC TAC CAT GCC TGC ACG AAT TAA GCA ATT CGT AAT CAT GGT CAT AGC T) or 5' FITC labelled 49mer dsDNA (AGC TAC CAT GCC TGC ACG AAT TAA GCA ATT CGT AAT CAT GGT CAT AGC T, prepared by annealing of complementary DNA strand) at 25 ° C for 10 min. Reactions were terminated by crosslinking with 0.8% glutaraldehyde for 10 min at 25 ° C and resolved in 0.8% TAE agarose gels (50 V, 50 min, 4 ° C). Gels were imaged on a Typhoon FLA 9500 scanner (GE Healthcare) and quantified using ImageJ software. The proportion of bound DNA was calculated from free ssDNA relative to "no protein" control.

5.5.7 Presynaptic complex assembly and disassembly assays

Single molecule ssDNA curtains were prepared and imaged by total internal reflection fluorescence microscopy, as previously described [306, 385]. All single-molecule presynaptic kinetic assays were performed essentially as previously described [38, 72, 306, 363, 385, 386]. In brief, ssDNA curtains were first prepared using RPA-GFP, and presynaptic complex assembly was initiated by injecting buffer solutions containing 2 µM of the indicated recombinase and 2 mM ATP into the sample chamber at a flow rate of 1 ml/min, buffer flow was then terminated and reactions were incubated without flow while capturing images (90 millisecond integration) at 10 second intervals for the duration of the assembly reactions. The resulting data were analyzed by integrating the RPA-GFP signal intensity over entire ssDNA molecules, and then plotting the normalized signal intensity versus time, as previously described [306]. Assembly rates were then extracted from the resulting graphs by fitting the data to single exponential curves. The presynaptic complex assembly conditions for each different recombinase were as follows: hRAD51(30 mM Tris-Acetate [pH 7.5], 1 mM $MgCl_2$, 5 mM $CaCl_2$, 100 mM KCl, 1 mM ATP, 1 mM DTT, and 0.2 mg/ml BSA; 37 $^\circ$C); hDMC1 (40 mM Tris-HCl [pH 7.5], 2 mM $MgCl_2$, 1.5 mM $CaCl_2$, 100 mM KCl, 2.5 mM ATP, 1 mM DTT, and 0.2 mg/ml BSA; 37 $^\circ$C); ScRad51 (30 mM Tris-Acetate [pH 7.5], 20 mM Mg-Acetate, 50 mM KCl, 1 mM DTT, 2.5 mM ATP, and 0.2 mg/ml BSA; 30 $^\circ$C); ScDmc1 (40 mM Tris-HCl [pH 7.5], 2 mM $MgCl_2$, 1.5 mM $CaCl_2$, 100 mM KCl, 2.5 mM ATP, 1 mM DTT, and 0.2 mg/ml BSA; 30 $^\circ$C); CeRad51 (50 mM Tris-HCl [pH 7.5], 5 mM $MgCl_2$, 50 mM NaCl, 2 mM ATP, 1 mM DTT, and 0.2 mg/ml BSA). For all recombinase, the stability of the presynaptic complexes was confirmed by flushing free protein from the

sample chambers while retaining ATP in the reaction buffers (as indicated above) while monitoring the samples at 30 ˚ C (to emulate the conditions of the dsDNA binding assays, see below). Filament disassembly kinetics were measured by flushing the sample chambers with buffers (as indicated above) lacking ATP (or in the case of hRAD51, lacking both ATP and Ca^{2+}), while collecting images (90 millisecond integration) at 10 second intervals for the duration of the assembly reactions, and the resulting data were analyzed similarly to the assembly reactions.

5.5.8 Yeast Strain construction

A W303 strain, LS1009-1 (MAT α ade3::GAL-HO leu2-3,112 trp1-1 ura3-1 can1-100; his3-11,15 ade2-1), was genetically modified at the RAD51 and MAT loci. First, the endogenous Z box was replaced with the commercially produced mutated Z box segments using a previously described method [412]. Briefly, two fragments were created by PCR: Fragment 1 consists of 100 nt upstream of Z box, the mutated 315 nt Z box, 100 nt downstream of the Z box, 300 nt upstream of K. lactis URA3 start codon, and 550 nt downstream from the K. lactis URA3 start codon; Fragment 2 contains100 nt downstream of the K. lactis URA3 start codon, 200 nt downstream K. lactis URA3 stop codon, 100 nt upstream of Z box, the mutated 315 nt Z box, and 100 nt downstream of the Z box. The two fragments were transformed into LS1009-1 and transformants were selected on SC-ura plates after growth at 30 ˚ C for two days. Colonies were patched onto SC-ura plates and grown at 30 ˚ C overnight. Colony PCR was performed checking for the presence of K. lactis URA3 at the MAT locus. Positive patches were transferred to YPD and grown for one day and then replica plated

to 5-FOA plates to select for colonies that lost the K. lactis URA3 gene. Colonies were patched onto YPD, grown overnight, and colony PCR of the MAT locus was performed. PCR fragments were then digested with the following restriction enzymes: MfeI (1 in 8 and 1 in 5), BglII (1 in 7), AflII (1 in 6), SpeI (1 in 4). Clones that yielded PCR fragments with the modified Z boxes were confirmed by DNA sequencing. The rad51 mutations were made in the Z box variant strains using a standard "pop-in/pop-out" method "pop-in/pop-out" method [413]. Briefly, rad51 mutants were inserted into pRS406 (URA3) plasmid using XhoI and BamHI cut sites. Each rad51 mutant was additionally altered with silent mutations to create unique restriction enzyme cut sites: HindIII (Rad51-DL1, DM, DL2), AfeI (M301Q, H302K), and PstI (A298E). Plasmids were then cut with EcoRI, transformed into the Z box variant strains, plated on SC-ura plates, and grown for 2 days at 30 ° C. Colonies were patched onto SC-ura plates and grown overnight. Patches were tested for plasmid integration by colony PCR. Positive patches were then transferred to YPD and grown overnight. Patches were then struck onto 5-FOA plates and grown for 2 days at 30 °C. Colonies were patched onto YPD, grown overnight, and colony PCR of the RAD51 locus was performed. Clones that yielded PCR fragments with the modified Z boxes were identified by restriction digestion and confirmed by DNA sequencing.

5.5.9 Plating assays

Cells were cultured overnight in YPD, and ten-fold serial dilutions of the cultures were spotted on YPD media containing either MMS or Zeocin at the indicated concentrations.

5.5.10 Rad54 focus formation

Cells were grown overnight in SC with supplemental adenine at 23 ˚C before being pelleted, resuspended at higher densities, and immobilized on a microscope slide by mixing with 1.4% agarose. Images were acquired on a Leica DM5500B upright microscope (Leica Microsystems) illuminated with a 100W mercury arc lamp. A Chroma 41028 high-efficiency filter cube was used for Rad54-YFP imaging. Images were captured with a Hamamatsu Orca AG cooled digital CCD (charged-coupled device), and analysis of image data was performed with Volocity software (Perkin-Elmer). We acquired 20 z-stacks spaced by 300 nm. Exposure times were as follows: DIC images (30 milliseconds), YFP fluorescence images (5 seconds).

5.5.11 MAT switching assays

This assay was modified from previously published protocols [291]. For each strain, colonies were grown in 2 ml of YPD overnight. Cultures were spun down at 3000 x g for 2 min, washed with 2 ml of YPL, and resuspended in 5 ml of YPL and grown for 8-10 hours. 240 ml of YPL media was inoculated at the end of the day such that the OD600 was around 0.02 and grown overnight at 30 ˚C. Cultures were allowed to grow until OD600 0.4 to 0.6 and 60 ml of culture was removed for DNA extraction (time point 0). 20 ml of 20% galactose was then added to the media and cells were allowed to grow for one hour. An additional 60 ml of culture was removed for DNA extraction. The rest of the culture was spun down, washed with 10 ml of sterile deionized water, and then resuspended in 160 ml of YPL with 2% glucose. 50, 45, 35, and 25 ml of cultures were then taken at 30, 60, 180, 300 min after YPL-glu resuspension, respectively. DNA was extracted as described previously [414]. For

each sample, 10 μg of DNA was digested with StyI for 12 hours at 37 $^\circ$C and then incubated for 20 min at 65 $^\circ$C to inactivate StyI. Samples were loaded into a 1% agarose gel and run in 1x TBE buffer for 2.5 hours at 160 V. DNA was fixed to GE Amersham Hybond-N+ (Cat. No. RPN303B) membrane by capillary transfer for 5 hours and UV-crosslinked (120 mJ/cm^2) Membranes were equilibrated with 25ml of GE Amersham Rapid-hyb buffer (Cat. No. RPN1636) at 65 $^\circ$C for 30 min. 20 ng of SAE2 and MAT probes were labeled using Thermo Scientific RadPrime DNA Labeling System (Cat. No. 18428011) and CTP-[-32P] (PerkinElmer Cat. No. BLU508H250UC) in a 50 µl reaction. Labeled probes were denatured at 95 $^\circ$C for 5 min and immediately put on ice and diluted with 100 μl of ice cold H2O. 75 μl of diluted probe was added to the hybridization buffer and membranes were incubated for 4 hours at 65 $^\circ$C. Membranes were then washed with 2x SSC with 0.1% SDS for 15min at 65 $^\circ$C and then two more washes with 1x SSC with 0.1% SDS for 20 min each at 65 $^\circ$C. A phosphor screen was exposed to the membrane for 12 hours and imaged on GE Typhoon FLA 9000 Imager and analyzed using ImageJ.

5.5.12 Molecular dynamics simulations

To gain insight into the mechanism by which the three lineage-specific amino acid residues in Dmc1 might stabilize mismatched base pairs, we sought to determine locate the position of these residues relative to the presynaptic ssDNA. A structure of the Dmc1 filament bound with the presynaptic ssDNA in not yet available. Therefore, as an initial alternative we replaced the corresponding residues of hRAD51 filament with the corresponding amino acids from Dmc1, yielding an initial structure for the hRAD51 with the following three

mutations: A240E, M243Q, and H244K. An initial structure wherein the side chains of these amino acids were randomly oriented was generated using Pymol (https://pymol.org), and the resulting mutant hRAD51 structure was then relaxed using all-atom molecular dynamics (MD) simulations, which were performed using Gromacs 5.0.6 (http://www.gromacs.org). In the simulation, the mutated structure was solvated with 27,267 water molecules in the 13.8 nm x 13.8 nm x 13.8 nm periodic boundary box. We added 186 Na+ and 158 Cl- ions to neutralize the system and to simulate a 100 mM salt concentration. The AMBER99SB-ILDN force fields were used for the protein and DNA molecules, while the TIP3P force field was used for water molecules. In all the simulations, we used the particle mesh Ewald method to calculate the electrostatic energy using a grid spacing of 0.16 Å and four-order interpolation. The cut-off length of electrostatic and Van der Waals interaction calculation was set to 1.0 Å. The bond lengths that include hydrogen atoms were constrained by p-LINCS for the protein and DNA molecules and by SETTLE for the water molecules. The initial structures were equilibrated before the start of the production simulation by 5×10^4 steps of the steepest descent energy minimization and by 1 x 105 steps of MD simulations with position restraints of the protein and DNA molecules. The first half of the equilibration MD simulation was conducted in the NVT ensemble at T = 300 [K] and second half in the NPT ensemble at T = 300 [K] and P = 1 [atm]. The 5×10^6 steps (10-ns) production simulations were performed in the NPT ensemble at T = 300 [K] and at P = 1 [atm] by the leap-frog integrator with 2-fs time step. The snapshots presented in Figure **??**B and C were obtained from the last frame of the simulations.

From the production simulations, we found that E240 in the hRAD51 A240E M243Q H244K triple mutant makes contacts with the complementary DNA strand, which are absent

in the simulation performed for wt hRAD51. We calculated the distances between the three atoms in E240 (Cβ, Cδ, Cγ) and the O4' atom in the backbone sugar, yielding distances of 5.0 ± 0.5 (Cβ), 6.8 ± 0.7 Å (Cδ), and 5.8 ± 0.7 (Cγ). This result suggests that these three atoms may make contact with the complementary DNA strand. For wt hRAD51, the corresponding residue is A240. The distance between the Cβ atom in A240 and the O4' atom in the backbone sugar is 7.9 in the structure obtained by the electron microscopy [39]. Thus, more atoms of E240 are indicated to have contacts with the complementary DNA strand compared to A240. This more intimate interaction is consistent with a model where Dmc1 may stabilizes mismatches by making nonspecific contacts with the phosphate backbone of the complementary DNA strand.

Chapter 6

BRCA1–BARD1 Promotes RAD51-Mediated Homologous DNA Pairing

This chapter is adapted from work originally published as: "Sequence imperfections and base triplet recognition by the Rad51/RecA family of recombinases," Weixing Zhao, **Justin B. Steinfeld**, Fengshan Liang, Xiaoyong Chen, David G. Maranon, Chu Jian Ma, YoungHo Kwon, Timsi Rao, Weibin Wang, Chen Sheng, Xuemei Song, Yonhong Deng, Judit Jimenez-Sainz, Lucy Lu, Ryan B. Jensen, Yong Xiong, Greg M. Kupfer, Claudia Wiese, Eric C. Greene, Patrick Sung. Nature. (2017). I was responsible for all of the single molecule experiments along with associated experimental design and data analysis.

6.1 Abstract

The tumor suppressor complex BRCA1–BARD1 functions in the repair of DNA double-stranded breaks by homologous recombination. During this process, BRCA1–BARD1 facilitates the nucleolytic resection of DNA ends to generate a single-stranded template for the recruitment of another tumor suppressor complex, BRCA2–PALB2, and the recombinase RAD51. Here, by examining purified wild-type and mutant BRCA1–BARD1, we show that both BRCA1 and BARD1 bind DNA and interact with RAD51, and that BRCA1–BARD1

enhances the recombinase activity of RAD51. Mechanistically, BRCA1–BARD1 promotes the assembly of the synaptic complex, an essential intermediate in RAD51-mediated DNA joint formation. We provide evidence that BRCA1 and BARD1 are indispensable for RAD51 stimulation. Notably, BRCA1–BARD1 mutants with weakened RAD51 interactions show compromised DNA joint formation and impaired mediation of homologous recombination and DNA repair in cells. Our results identify a late role of BRCA1–BARD1 in homologous recombination, an attribute of the tumor suppressor complex that could be targeted in cancer therapy.

6.2 Introduction

Mutations in *BRCA1* (breast cancer susceptibility gene 1) are linked to familial breast and ovarian cancers, and also to Fanconi anaemia [216, 415–418]. Since its discovery over twenty years ago [419–421], BRCA1 has been implicated in various biological processes, including mRNA splicing and microRNA biogenesis [422–426], DNA damage signaling, cell cycle checkpoints [416, 427], the avoidance of replication–transcription conflicts [428, 429], and DNA double-strand break (DSB) repair by homologous recombination [415, 418, 430–432]. The role of BRCA1 in these processes has remained mostly undefined, largely because of difficulties in obtaining high-quality protein preparations for biochemical analyses. BRCA1 (1,863 residues) forms a stable complex with BARD1 (BRCA1-associated RING domain protein 1; 777 residues) [433, 434]. Depletion of BARD1 engenders DNA damage sensitivity, homologous recombination deficiency, and genome destabilization [433, 435–438]. The ablation of BARD1 in mice leads to cancer susceptibility [437], and probable disease-causing

mutations are found in patients with cancer [439–443].

We have investigated the multifaceted role of BRCA1–BARD1 in homologous recombination-mediated DSB repair. During the repair process, the DSB ends are resected to yield 3 single-stranded DNA tails [444]. These DNA tails become coated with replication protein A (RPA), which is subsequently displaced by the recombinase protein RAD51 to form a nucleoprotein complex termed the presynaptic filament. The presynaptic filament searches for, engages, and then invades a homologous duplex target to form a nascent heteroduplex DNA joint, the displacement loop or D-loop. This is followed by DNA synthesis and resolution of DNA intermediates to complete repair [28]. There is evidence that BRCA1 promotes DNA end resection by acting as an antagonist of 53BP1 and regulating the MRE11–RAD50–NBS1–CtIP resection nuclease complex, and also participates in RAD51-mediated presynaptic filament formation with the tumor suppressors BRCA2 and PALB2 [216, 445]. To gain mechanistic insights into how BRCA1–BARD1 promotes homologous recombination, we have developed a robust system for co-expressing BRCA1 and BARD1 in insect cells and a protocol to obtain BRCA1–BARD1 for biochemical testing. Our results reveal novel attributes of BRCA1–BARD1 and a role for this protein complex in the DNA strand invasion step of homologous recombination-mediated chromosome damage repair.

6.3 Results

6.3.1 DNA Binding by BRCA1 and BARD1

BRCA1–BARD1 was expressed in insect cells and purified to near homogeneity (Fig. 6.1A-D). We used the DNA electrophoretic mobility shift assay (EMSA) to test the ability of BRCA1–BARD1 to bind radiolabelled single-stranded DNA (ssDNA), double-stranded DNA (dsDNA), the replication fork, the D-loop and DNA bubble (a short section of unwound DNA that forms during biological processes such as transcription). We also performed competition experiments in which the nucleoprotein complex of BRCA1–BARD1 and radiolabelled D-loop was challenged with an unlabeled DNA species. The results revealed that BRCA1–BARD1 has the highest affinity for the D-loop and DNA bubble, followed by the replication fork, dsDNA and ssDNA (Fig. 6.1E-H, Fig. 6.2A,B and Fig. 6.3A,B).

BRCA1 is known to bind DNA [446, 447]. Using the southwestern assay, we found that both BRCA1 and BARD1 bind the D-loop, with BARD1 showing an apparently higher affinity for the substrate (Fig. 6.2C). Consistent with this, BRCA1–BARD1$^{1–142}$ (full-length BRCA1 in complex with only the RING domain of BARD1) exhibited a lower affinity for various DNA substrates (Fig. 6.1I,J). Together, our results show that both BRCA1 and BARD1 contribute to the DNA-binding capability of the BRCA1–BARD1 complex. The DNA-binding domain of BRCA1 was previously found to reside within the protein's middle region [446, 447] and our mapping effort has led to the isolation of the BARD1 DNA-binding domain (Fig. 6.3C-E). Notably, the BARD1 domain exhibited similar DNA-binding properties to the complex (Fig. 6.3F-I). Thus, BARD1 is a structure-specific DNA-binding protein with the highest affinity for the D-loop and DNA bubble.

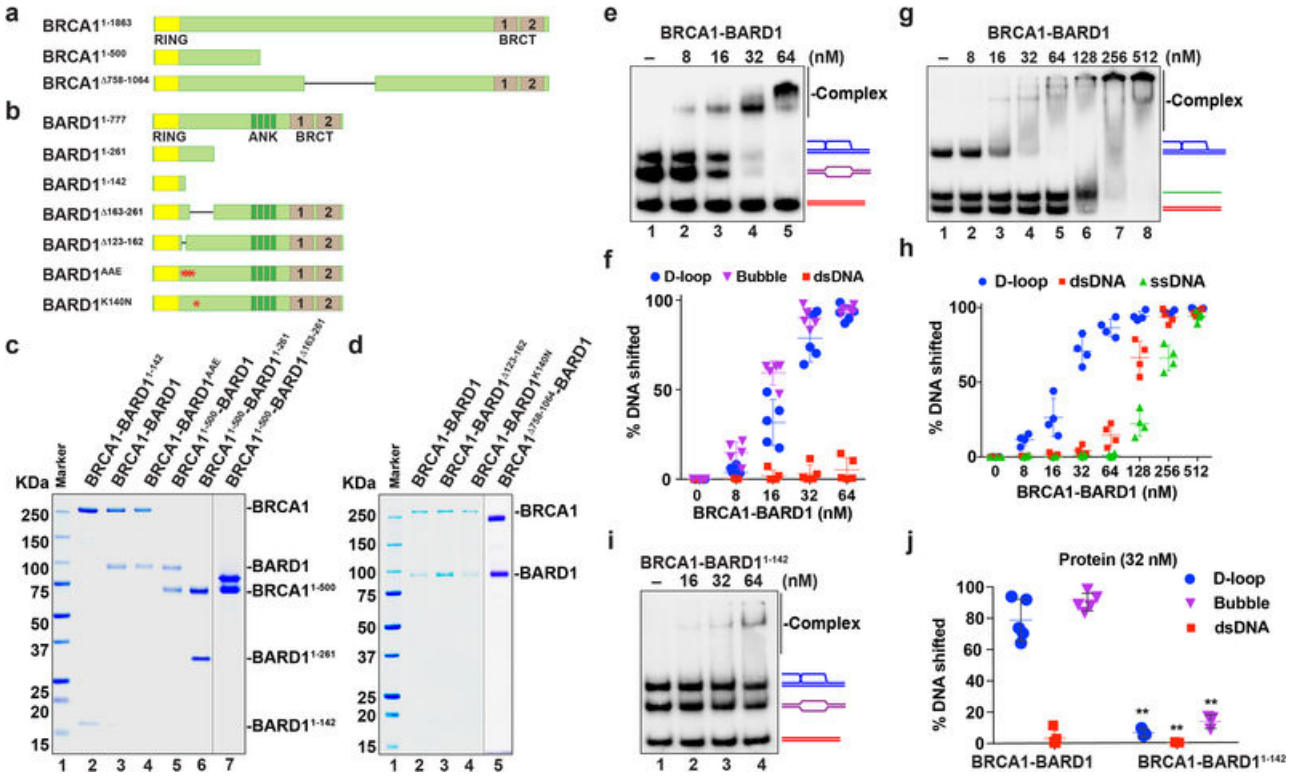

Figure 6.1: **Purification of BARD1-BRCA1 and mutant variants, and DNA-binding properties of BRCA1-BARD1 and BRCA1-BARD1$^{1-142}$** Schematics of BRCA1 (**A**) and BARD1 (**B**) and mutant variants of these proteins tested in this study. (**C**) SDS–PAGE of purified BRCA1–BARD1$^{1-142}$ (lane 2), BRCA1–BARD1 (lane 3), BRCA1–BARD1AAE (lane 4), BRCA1$^{1-500}$–BARD1 (lane 5), BRCA1$^{1-500}$–BARD1$^{1-261}$ (lane 6) and BRCA1$^{1-500}$–BARD1$^{\Delta 163-261}$ (lane 7). Size markers were run in lane 1. (**D**) SDS–PAGE of purified BRCA1–BARD1 (lane 2), BRCA1–BARD1$^{\Delta 123-162}$ (lane 3), BRCA1–BARD1K140N (lane 4) and BRCA1 758–1064–BARD1 (lane 5). Size markers were run in lane 1. (**E**)DNA binding test of BRCA1–BARD1 with a mixture of D-loop, DNA bubble and dsDNA. (F) Quantification of data from experiments in E. Data are means ± s.d., n = 5. (G) DNA binding test of BRCA1–BARD1 with a mixture of D-loop, dsDNA and ssDNA. (H) Quantification of data from experiments in G. Data are means ± s.d., n = 4. (I) DNA binding test of BRCA1–BARD1$^{1-142}$ with a mixture of D-loop, DNA bubble and dsDNA. (J) Quantification of the results obtained with 32 nM of protein complexes in E and I. Data are means ± s.d., n = 3 (BRCA1–BARD1$^{1-142}$) or 5 (BRCA1–BARD1). **P < 0.01.

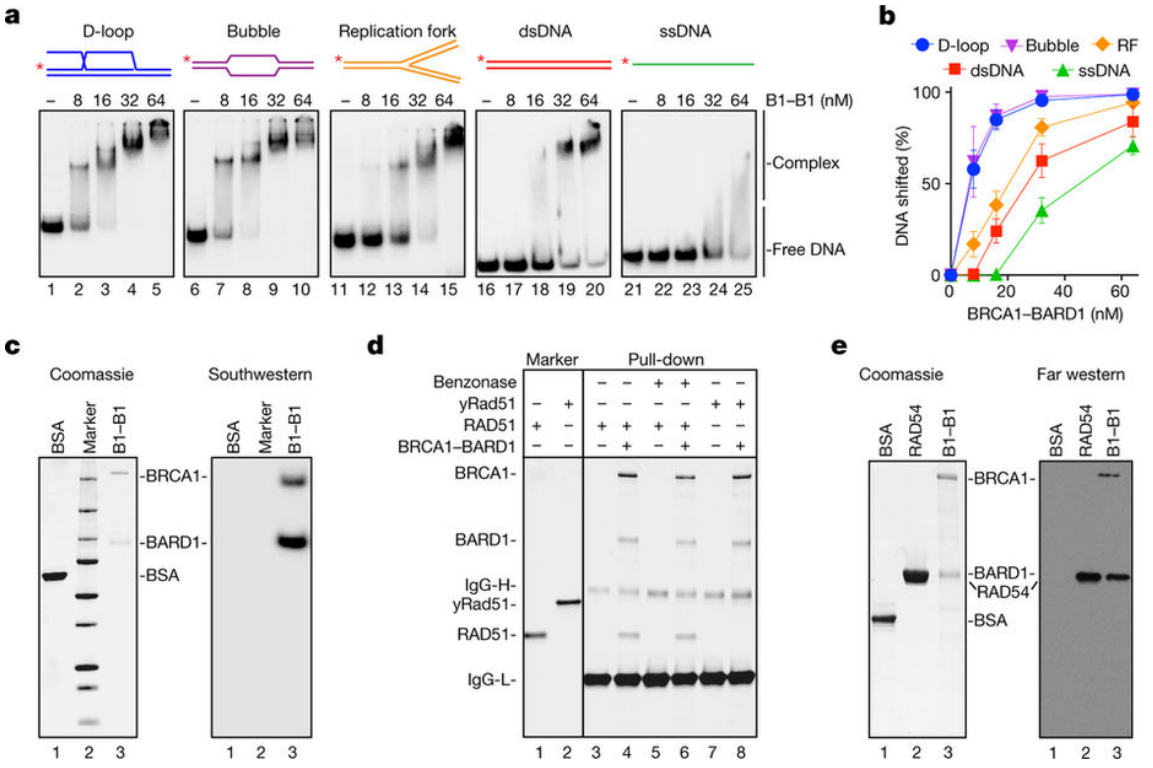

Figure 6.2: **DNA-binding and RAD51-interaction attributes of BRCA1-BARD1.** (**A**) Binding of D-loop, DNA bubble (Bubble), replication fork (RF), dsDNA and ssDNA. (**B**) Quantification of interaction shown in A. Data are means ± s.d., n = 3 (bubble, RF, dsDNA and ssDNA) or 5 (D-loop). (**C**) Southwestern analysis to test D-loop binding. Bovine serum albumin (BSA) was the negative control. (**D**) Pull-down analysis for interaction of RAD51 or yRad51 with BRCA1–BARD1. (**E**) Far western analysis for interaction of BRCA1 and BARD1 with RAD51. B1–B1, BRCA1–BARD1. BSA and RAD54 were the negative and positive controls, respectively.

6.3.2 RAD51 Interaction with BRCA1-BARD1

BRCA1 has been shown to co-immunoprecipitate with RAD51 from cell extracts[432], but it has remained unclear whether it associates with RAD51 directly. By affinity pull-down, we found that BRCA1–BARD1 interacts with human RAD51 but has little or no affinity for yeast Rad51 (yRad51) (Fig. 6.2D) or Escherichia coli RecA (Fig. 6.4A). We also determined that four or five RAD51 molecules are bound by BRCA1–BARD1 (Fig. 6.4B,C). Notably, formation of the BRCA1–BARD1–RAD51 complex was not affected by benzonase or ethidium bromide (Fig. 6.2D and 6.4D), indicating that the association is not bridged

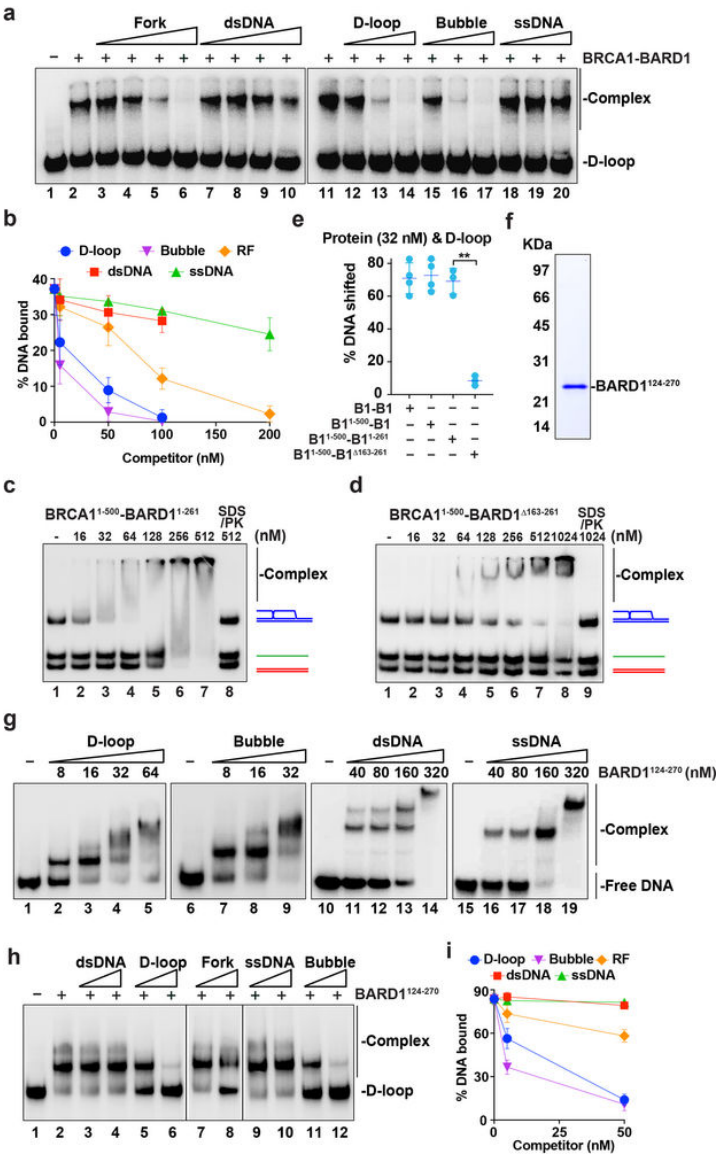

Figure 6.3: **DNA Binding by BARD1.** (**A**) BRCA1–BARD1 (5 nM) was incubated with radiolabelled D-loop (10 nM) and then the nucleoprotein complex was presented with an increasing concentration of unlabelled ssDNA, dsDNA, fork, bubble or D-loop as indicated. (**B**) Quantification of data from experiments in A. Data are means ± s.d., n = 2 (ssDNA) or 3 (all other substrates). (**C**) DNA binding test of BRCA1$^{1-500}$–BARD1$^{1-261}$ with a mixture of D-loop, dsDNA and ssDNA. (**D**) DNA binding test of BRCA1$^{1-500}$–BARD1$^{\Delta 163-261}$ with a mixture of D-loop, dsDNA and ssDNA. (**E**) Comparison of results obtained using 32 nM of BRCA1–BARD1 (from Fig. 6.1G) BRCA1$^{1-500}$–BARD1 (from Fig. 6.15A), BRCA11–500–BARD11–261 (from C) and BRCA11–500–BARD1 163–261 (from D). Data are means ± s.d., n = 3 (BRCA11–500–BARD11–261 and BRCA11–500–BARD1 163–261) or 4 (BRCA1–BARD1 and BRCA1$^{1-500}$–BARD1). **P < 0.01. (**F**) SDS–PAGE of purified BARD1$^{124-270}$. (**G**) EMSA to test BARD1$^{124-270}$ for binding to the D-loop, DNA bubble (Bubble), double-stranded DNA (dsDNA) and single-stranded DNA (ssDNA). (**H**) Nucleoprotein complex consisting of BARD1$^{124-270}$ (16 nM) and radiolabelled D-loop (10 nM) was challenged with an increasing concentration of unlabelled ssDNA, dsDNA, fork, DNA bubble or D-loop as indicated. (**I**) Quantification of data from experiments in H. Data are means ± s.d., n = 3 (D-loop and ssDNA) or 4 (Bubble, RF and dsDNA).

by nucleic acid. Surprisingly, both BRCA1 and BARD1 retained RAD51 in the far western assay, with BARD1 showing a more robust signal (Fig. 6.2E), whereas the homologous recombination factors RAD51D–XRCC2 and DSS1 did not bind RAD51 under the same conditions (Fig. 6.4E). These results helped to establish that BRCA1–BARD1 associates with RAD51 in a species-specific manner, and that both proteins in the complex participate in this interaction.

6.3.3 BRCA1-BARD1 enhances homologous DNA pairing

Given that BRCA1–BARD1 binds DNA and interacts with RAD51 (Fig. 6.2), we hypothesized that it would enhance either the assembly of the presynaptic filament or the potential of the presynaptic filament to mediate DNA strand invasion, or both. We used a DNA strand exchange assay [448, 449] (Fig. 6.5A) to test whether BRCA1–BARD1 could facilitate RAD51 presynaptic filament assembly. Although the BRCA2–DSS1 complex promotes RAD51 presynaptic filament assembly on RPA-coated ssDNA [449, 450], BRCA1–BARD1 does not (Fig. 6.5B,C). Moreover, unlike BRCA2–DSS1 [448, 450], BRCA1–BARD1 cannot target RAD51 to ssDNA when dsDNA is present (Fig. 6.5D-F).

Next, we conducted a D-loop assay to test whether BRCA1–BARD1 could promote DNA strand invasion (Fig. 6.6A). Notably, BRCA1–BARD1, in amounts substoichiometric to RAD51, strongly enhanced the reaction, regardless of whether ATP (Fig. 6.5G-I) or the non-hydrolysable analogue AMP-PNP (Fig. 6.6A-C) was used as the nucleotide cofactor, while BRCA2–DSS1 did not stimulate strand invasion (Fig. 6.6A-C). By contrast, BRCA1–BARD1 did not enhance the activity of yRad51 (Fig. 6.5J,K). BRCA2–DSS1 did

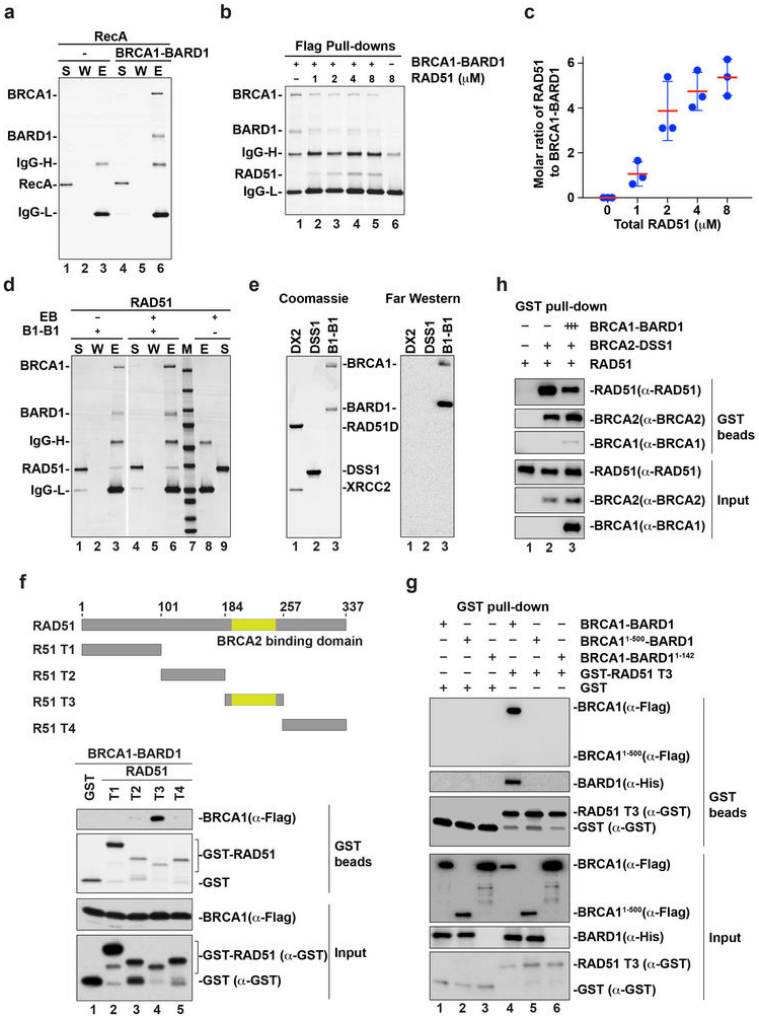

Figure 6.4: **The RAD51 interaction attributes of BRCA1-BARD1.** (**A**) Affinity pull-down to test for the interaction of RecA with BRCA1–BARD1 (B1–B1) via the Flag tag on BRCA1. The supernatant (S), wash (W) and eluate (E) fractions were analysed by SDS–PAGE and Coomassie blue staining. (**B**) Affinity pull-down with Flag-tagged BRCA1–BARD1 (66 nM) and an increasing concentration of RAD51 (1, 2, 4 and 8 μM). The eluates from the pull-down experiment were analysed by SDS–PAGE with Coomassie blue staining. (**C**) The amount of BRCA1–BARD1 and RAD51 in lanes 2–5 of B was quantified against known quantities of these protein species, run and stained in the same SDS polyacrylamide gel. Data are means $\pm$ s.d., n = 3. (**C**) Affinity pull-down to test for the interaction of RAD51 with BRCA1–BARD1 with or without ethidium bromide (EB) being present. (**E**) Far western analysis to examine RAD51D–XRCC2 (DX2), GST–DSS1 (DSS1) and BRCA1–BARD1 for RAD51 interaction. (**F**) Schematic of the GST-tagged RAD51 fragments examined (top). Results from the pull-down experiment to test for interaction of BRCA1-BARD1 with the RAD51 fragments via the GST tag on the latter (bottom). RAD51 fragments and BRCA1 were revealed by immunoblot analysis using anti-GST or anti-Flag antibodies, respectively. (**G**) GST pull-down assay to test for the interaction of the RAD51-T3 fragment with BRCA1–BARD1, BRCA1$^{1-500}$–BARD1 and BRCA1–BARD1$^{1-142}$ The RAD51 fragment, GST, BRCA1 and BARD1 were revealed by immunoblot analysis using anti-GST, anti-Flag or anti-His antibodies, respectively. (**H**) GST pull-down assay to test for competition between BRCA1–BARD1 (198 nM) and BRCA2–DSS1 (66 nM) for RAD51 (1 μM); DSS1 was GST-tagged. RAD51, BRCA1 and BRCA2 were revealed by immunoblot analysis using antibodies specific for them.

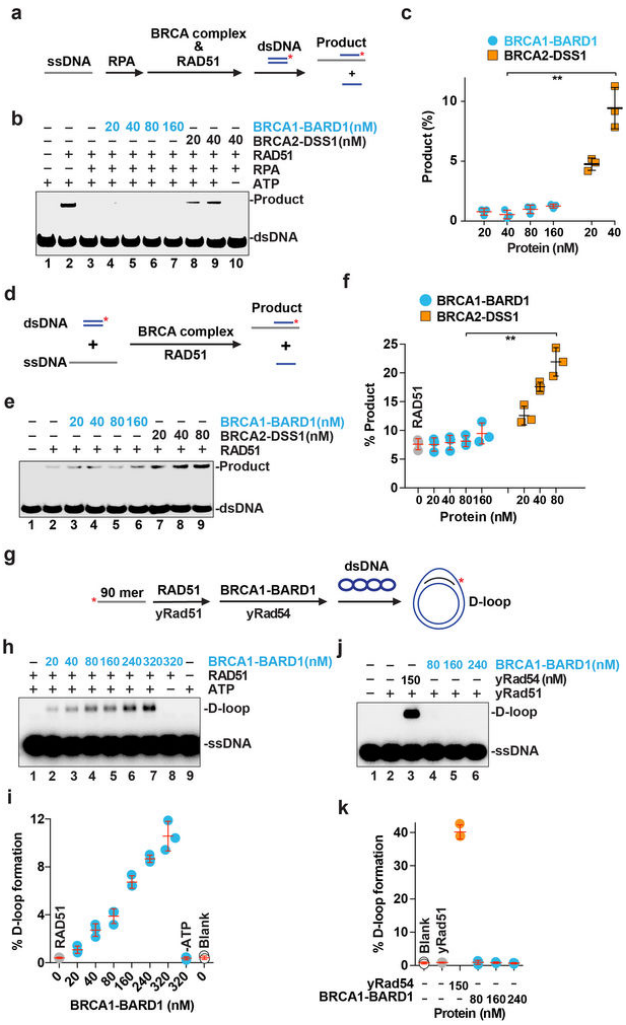

Figure 6.5: **Lack of recombination mediator activity in BRCA1-BARD1 and species-specific enhancement of RAD51 recombinase by BRCA1-BARD1.** (**A**) Schematic of the test for mediator activity of BRCA complex (BRCA1–BARD1 and BRCA2–DSS1). (**B**) BRCA1–BARD1 and BRCA2–DSS1 were tested for recombination mediator activity with RPA-coated ssDNA as substrate. (**C**) Quantification of data from experiments in B. Data are means ± s.d., n = 3. (**D**) Schematic of the test for ssDNA targeting activity of BRCA complex (BRCA1–BARD1 and BRCA2–DSS1). (**E**) BRCA1–BARD1 was tested alongside BRCA2–DSS1 for the ability to target RAD51 to ssDNA. (**F**) Quantification of data from experiments in E. Data are means ± s.d., n = 3. (**G**) Schematic of the D-loop assay. (**H**) D-loop reactions were carried out with the indicated concentration of BRCA1–BARD1 and ATP as the nucleotide co-factor. (**I**) Quantification of data from experiments in H. Data are means ± s.d., n = 3. (J) BRCA1–BARD1 and Saccharomyces cerevisiae Rad54 (yRad54) were tested for their influence on D-loop formation catalysed by S. cerevisiae Rad51 (yRad51). (**K**) Quantification of data from experiments in J. Data are means ± s.d., n = 3.

not stimulate D-loop formation by BRCA1–BARD1–RAD51 when the ssDNA substrate was pre-incubated with RAD51 or when ssDNA and plasmid DNA were premixed (Fig. 6.7A,B). However, with RPA-coated ssDNA, D-loop formation became more robust in the presence of BRCA1–BARD1 and BRCA2–DSS1 than with either complex alone (Fig. 6.7C,D). Together, these results reveal an unexpected role of BRCA1–BARD1 in promoting DNA joint formation that is catalyzed by RAD51 (Fig. 6.6D).

In homologous DNA pairing, the presynaptic filament captures the duplex partner and then assembles the synaptic complex in which the recombining DNA molecules are aligned in homologous registry and base switching has occurred [28]. By monitoring the protection

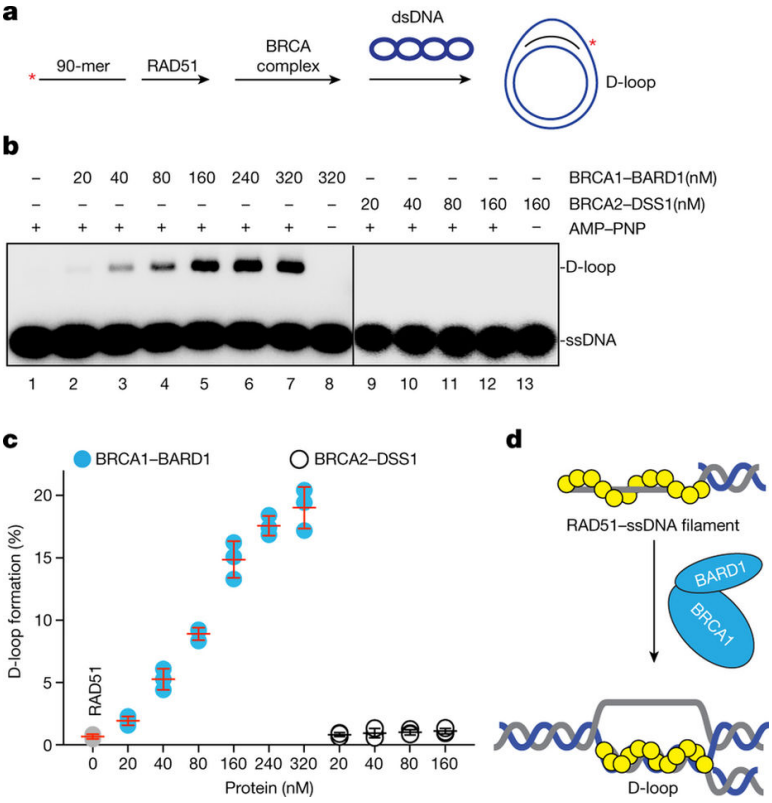

Figure 6.6: **Enhancement of RAD51-mediated D-loop formation by BRCA1-BARD1.** (**A**) Assay schematic of D-loop reactions. (**B**) D-loop reactions performed with the BRCA1–BARD1 or BRCA2–DSS1 complex. (**C**) Quantification of the reactions in B. Data are means ± s.d., n = 3. (**D**) Schematic depicting the role of BRCA1–BARD1 in DNA strand invasion during homologous recombination.

of dsDNA against restriction enzyme digestion (Fig. 6.8A), we found that BRCA1–BARD1 stimulates synaptic complex formation (Fig. 6.8B,C). Next, we used our DNA curtain assay [306, 363] to examine the pairing of homologous DNA sequences in real time (Fig. 6.8D). As previously reported [306, 363], the RAD51 presynaptic filament was able to engage a 70-base pair dsDNA fragment harboring 9-nucleotide homology (Fig. 6.8E). Importantly, the results showed enhancement of DNA engagement by BRCA1–BARD1 (Fig. 6.8E). However, we found no evidence that BRCA1–BARD1 has any effect on the binding site distributions, the pairwise distance distributions, or the resident time (k_{off}) of the bound dsDNA (Fig. 6.7E and Fig. 6.8F,G). Since BRCA1–BARD1 does not affect the k_{off} of the aligned dsDNA, we speculate that it acts by increasing the kon of dsDNA engagement. We note that BRCA1–BARD1 mutants impaired for BARD1–RAD51 interaction or lacking the RAD51-interaction domain of BRCA1 cannot promote pairing with the duplex target (see below). We also

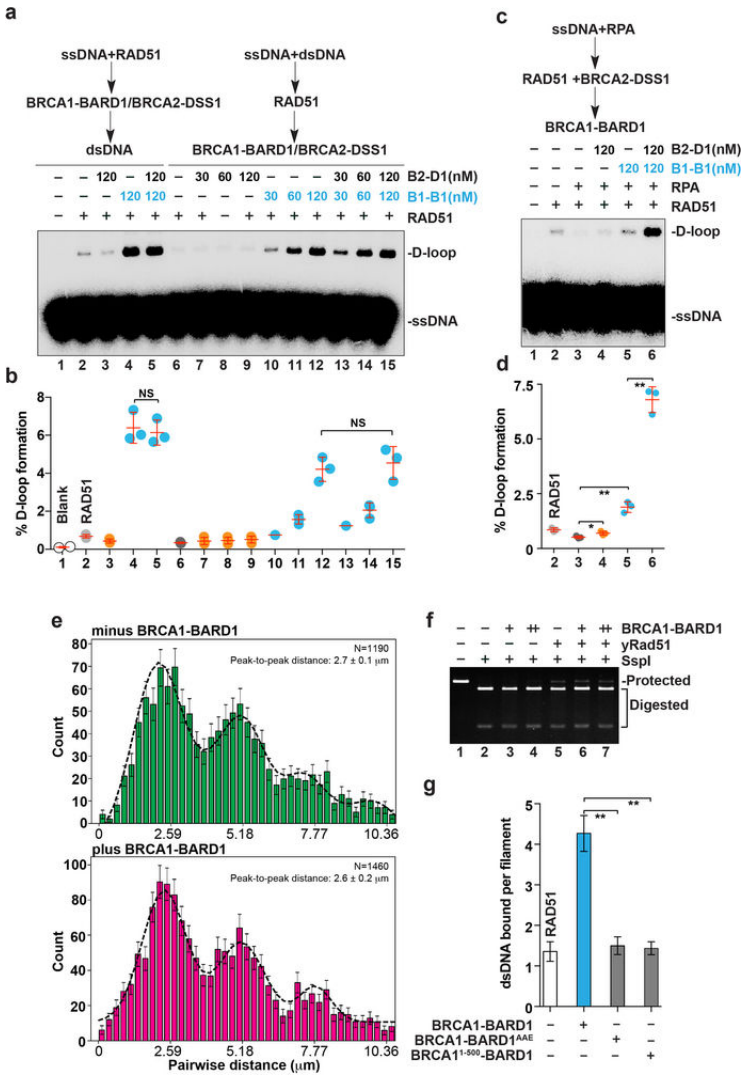

Figure 6.7: **Interplay between BRCA2-DSS1 and BRCA1-BARD1.** (**A**) D-loop reactions performed with the indicated concentration of BRCA1–BARD1 (B1–B1), BRCA2–DSS1 (B2–D1), and order of addition of reaction components. (**B**) Quantification of data from experiments in A. Data are means ± s.d., n = 3. NS, non-significant. (**C**) D-loop reactions performed with the indicated concentration of BRCA1–BARD1, BRCA2–DSS1, and order of addition of reaction components. (**D**) Quantification of data from experiments in C. Data are means ± s.d., n = 3. *P < 0.05; **P < 0.01. (**E**) Pairwise distance distributions for Atto565-dsDNA bound to the RAD51–ssDNA filaments with or without BRCA1–BARD1. Data are means ± errors (determined by bootstrapping). (**F**) BRCA1–BARD1 (100 and 200 nM) was tested with filaments of yRad51–ssDNA in synaptic complex assembly as assayed by protection against restriction digest. (**G**) Number of dsDNA oligonucleotides bound by the RAD51–ssDNA filament without (n = 49) and with BRCA1–BARD1 (n = 54), BRCA1–BARD1^{AAE} (n = 50) or BRCA1$^{1-500}$–BARD1 (n = 50). Data are means ± 95% confidence intervals. **P < 0.01.

verified that BRCA1–BARD1 does not affect the ability of presynaptic filaments harboring yRad51 to engage dsDNA (Fig. 6.7F and Fig. 6.8).

6.3.4 Functional relevance of BARD1-RAD51 interaction

We sought to isolate RAD51-binding defective mutants of BRCA1–BARD1 for biochemical and genetic testing. First, we co-expressed RAD51 with various BRCA1 fragments in insect cells and conducted co-immunoprecipitation. Consistent with a previous study [432], BRCA1$^{1–1527}$ could interact with RAD51, whereas BRCA1$^{1–1000}$ and BRCA1$^{1–500}$ were impaired in this regard (Fig. 6.9A-C). Notably, BRCA1–BARD1 could co-precipitate much more RAD51 than BRCA1 alone. This result, together with the far western data (Fig. 6.2E), indicates that BARD1 harbors a major RAD51-interaction domain (Fig. 6.10A). Deletion analysis showed that the region between residues 123 and 162 of BARD1 is indispensable for RAD51 interaction (Fig. 6.9D-G and Fig. 6.10A,B). Moreover, a GST-tagged BARD1 fragment harboring these residues could efficiently associate with RAD51 (Fig. 6.9H), indicating that it encompasses the RAD51 interaction domain. We also discovered that the core domain of RAD51 (referred to as T3), which has been implicated in BRCA2 binding via the BRC4 repeat of BRCA2 [451], can interact with BRCA1–BARD1 (Fig. 6.4F), but not with BRCA1–BARD1$^{1–142}$ or BRCA1$^{1–500}$–BARD1 (Fig. 6.4G). Interestingly, BRCA1–BARD1 could compete with BRCA2–DSS1 for RAD51 association (Fig. 6.4H).

We expressed and purified the mutant BRCA1–BARD1$^{\Delta 123–162}$ complex, in which the RAD51-interaction domain of BARD1 has been deleted. BRCA1–BARD1$^{\Delta 123–162}$ retained normal DNA-binding activity (Fig. 6.11A,B) but was defective in RAD51 interaction (Fig.

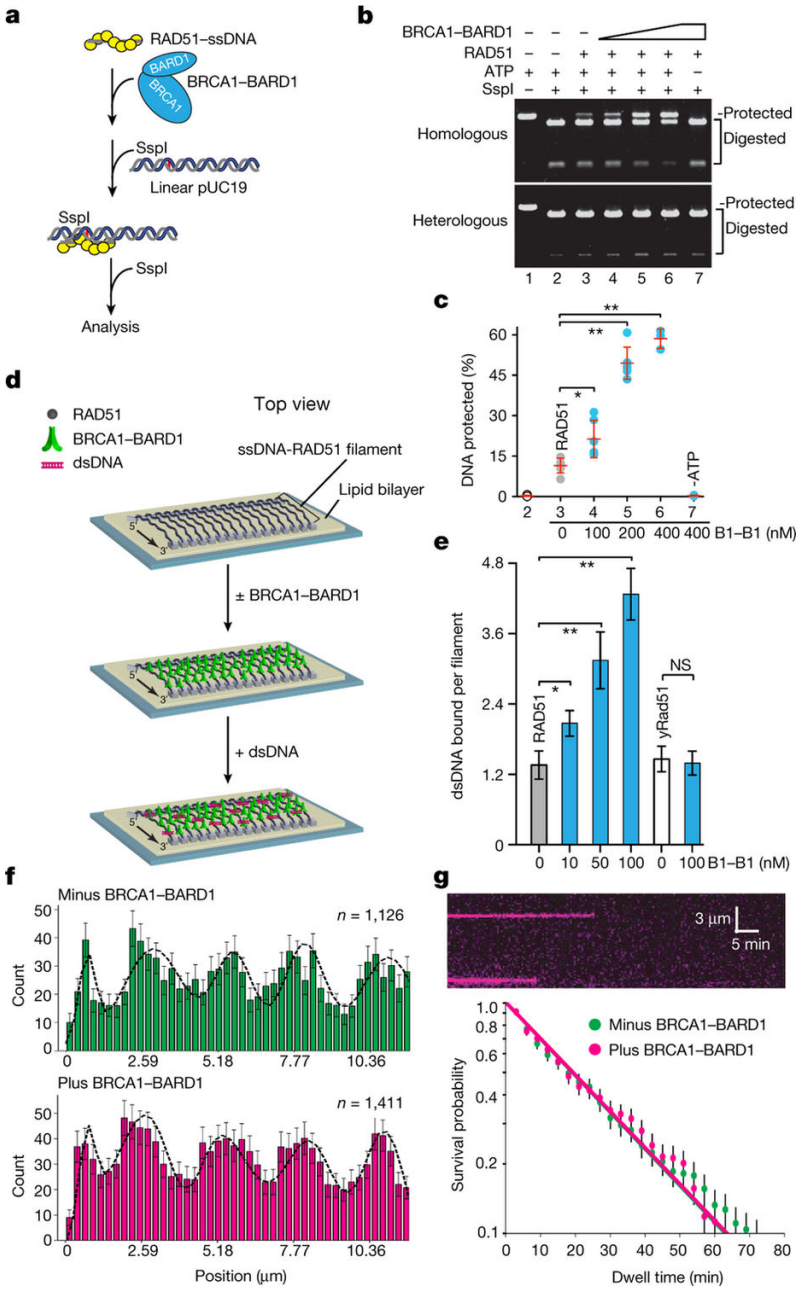

Figure 6.8: **Promotion of synaptic complex formation by BRCA1-BARD1.** (**A**) Schematic of the synaptic complex assay. (**B**) Synaptic complex formation by the RAD51–ssDNA filament and BRCA1–BARD1. (**C**) Quantification of synaptic complex formation. Data are means ± s.d., n = 3 (for lanes 6 and 7) or 6 (for lanes 2–5). (**D**) Schematic of DNA curtain assay. (**E**) Number of dsDNA oligonucleotides bound by each RAD51–ssDNA or yRad51–ssDNA filament as a function of BRCA1–BARD1 concentration. Data are means ± 95% confidence intervals, n = 49 (RAD51), 50 (RAD51 + 10 nM BRCA1–BARD1), 38 (RAD51 + 50 nM BRCA1–BARD1), 54 (RAD51 + 100 nM BRCA1–BARD1), 51 (yRad51) or 53 (yRad51 + 100 nM BRCA1–BARD1). (**F**) Binding distribution for Atto565–dsDNA with or without BRCA1–BARD1. (**G**) Semi-log survival plot of the synaptic complex with and without 100 nM BRCA1–BARD1. *P < 0.05; **P < 0.01. NS, non-significant. In F and G, data are means ± errors (determined by bootstrapping). The multiGaussian in F and the lines in G were fitted with least squares analysis.

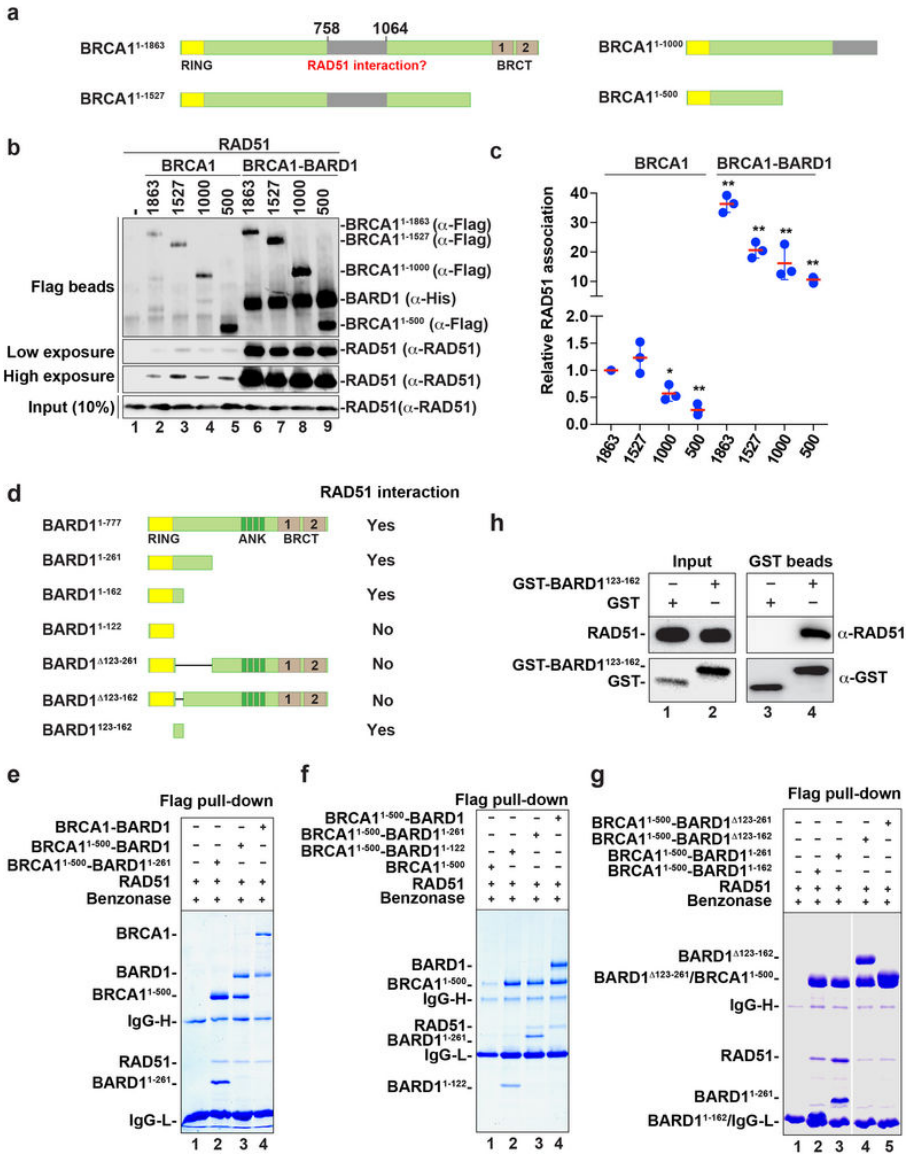

Figure 6.9: **Identification of the RAD51 interaction domain in BRCA1-BARD1.** (**A**) Schematic of the BRCA1 deletion variants37 examined in this study. (**B**) Testing BRCA1 deletion variants alone or in complex with BARD1 for the ability to co-immunoprecipitate RAD51 from insect cell extracts using anti-Flag resin with Benzonase treatment. The immunoprecipitates were analyzed by western blotting with antibodies against the Flag epitope (for BRCA1), the His6 epitope (for BARD1), or RAD51, as indicated. The cell extracts (10% of total) were probed for their RAD51 content. (**C**) Quantification of data from experiments in B. Data are means ± s.d., $n = 3$. *$P < 0.05$; **$P < 0.01$. (**D**) Summary of the RAD51 interaction ability of BARD1 truncation mutants, based on the pull-down analyses in (**E**) (for BRCA1–BARD1, BRCA1$^{1–500}$–BARD1 and BRCA1$^{1–500}$–BARD1$^{1–261}$), (**F**) (for BRCA1$^{1–500}$–BARD1, BRCA1$^{1–500}$–BARD1$^{1–261}$ and BRCA1$^{1–500}$–BARD1$^{1–122}$), (**G**) (for BRCA1$^{1–500}$–BARD1 123–261, BRCA1$^{1–500}$–BARD1$^{\Delta 123–162}$, BRCA1$^{1–500}$–BARD1$^{1–261}$ and BRCA1$^{1–500}$–BARD1$^{1–162}$) and (**H**) (for BARD1$^{123–162}$). In E, F and G, the eluates from the affinity resin were analyzed by SDS–PAGE and Coomassie blue staining. In H, the interaction between RAD51 and GST–BARD1$^{123–162}$ was tested by pull-down using glutathione resin. The input and eluate fractions were analyzed by western blotting with antibodies against GST or RAD51, as indicated.

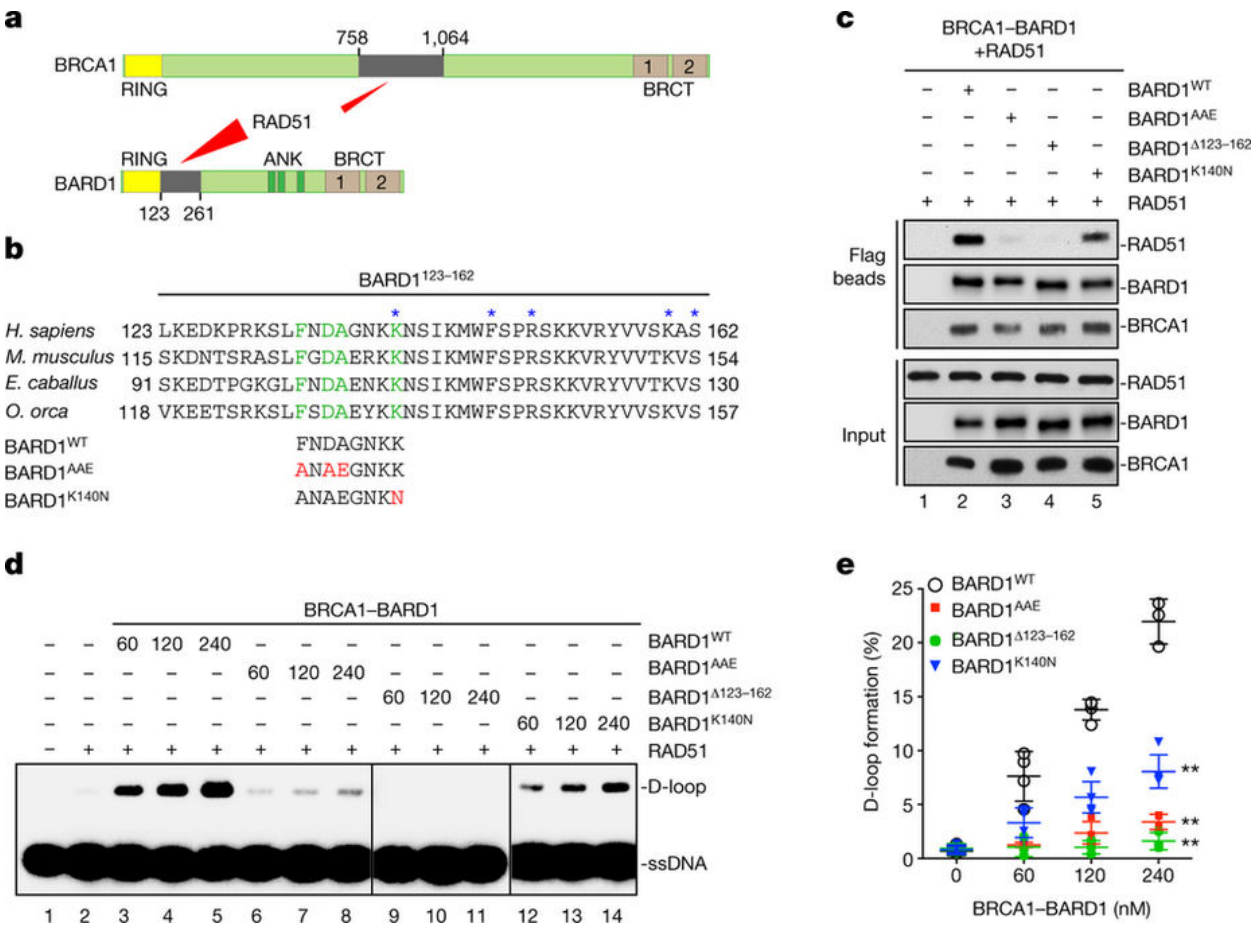

Figure 6.10: **Relevance of the BARD1-RAD51 complex in DNA strand invasion.** (**A**) Domains in BRCA1–BARD1. (**B**) Alignment of the RAD51 interaction domain in BARD1 orthologues. The highlighted residues (green) were changed to AAE or N (red). The asterisks denote BARD1 mutations found in human cancers (cBioPortal for Cancer Genomic). (**C**) Testing of RAD51 interaction with wild-type or mutant BRCA1–BARD1. (**D**) Examination of BRCA1–BARD1 mutants in the D-loop reaction. (**E**) Quantification of data from experiment in D. Data are means ± s.d., n = 3 (BARD1$^{\Delta 123–162}$), 4 (BARD1WT and BARD1^{AAE}) or 5 (BARD1^{K140N}). P values were calculated using two-way ANOVA and multiple comparisons were corrected by the Bonferroni method. **P < 0.01.

6.10C) and, accordingly, failed to enhance D-loop formation (Fig. 6.10D,E) or synaptic complex assembly (Fig. 6.11C,D). Sequence alignment of the RAD51-interaction domain in BARD1 orthologues revealed a number of conserved amino acid residues (including the FXDA motif; Fig. 6.10B). On the basis of this information, we generated a compound mutant that changes the conserved residues F133 and D135 to alanine and A136 to glutamic acid (the AAE mutant); F133 was included because other RAD51-interaction motifs [451–453], such as BRC4 in BRCA2 [451], also harbor a functionally indispensable phenylalanine residue. We expressed and purified the mutant BRCA1–BARD1^{AAE} complex. Biochemical testing revealed that, even though the mutant complex binds DNA normally (Fig. 6.11A,B), it is impaired not only for RAD51 association (Fig. 6.10C), but also for the ability to stimulate D-loop formation and synaptic complex assembly (Fig. 6.8G, 6.10D,E, and Fig. 6.11C,D). Together, these results provided evidence that the BRCA1–BARD1–RAD51 complex is indispensable for the enhancement of RAD51-mediated DNA strand invasion.

Cancer-associated mutations have been identified within the RAD51-interaction domain of BARD1. One such mutation (K140N), found in two patients with colorectal adenocarcinoma or uterine corpus endometrial carcinoma, alters the conserved residue K140 (cBioPortal for Cancer Genomic) [454, 455] next to the FXDA motif (Fig. 6.10B). To determine the relevance of this mutation, we expressed and purified the BRCA1–BARD1K140N mutant complex for testing. Although the mutation had no impact on DNA binding (Fig. 6.11A,B), it attenuated the affinity of BRCA1–BARD1 for RAD51 (Fig. 6.10C) and also compromised the ability of the tumor suppressor complex to enhance D-loop formation and synaptic complex assembly (Fig. 6.10D, E and Fig. 6.11C,D).

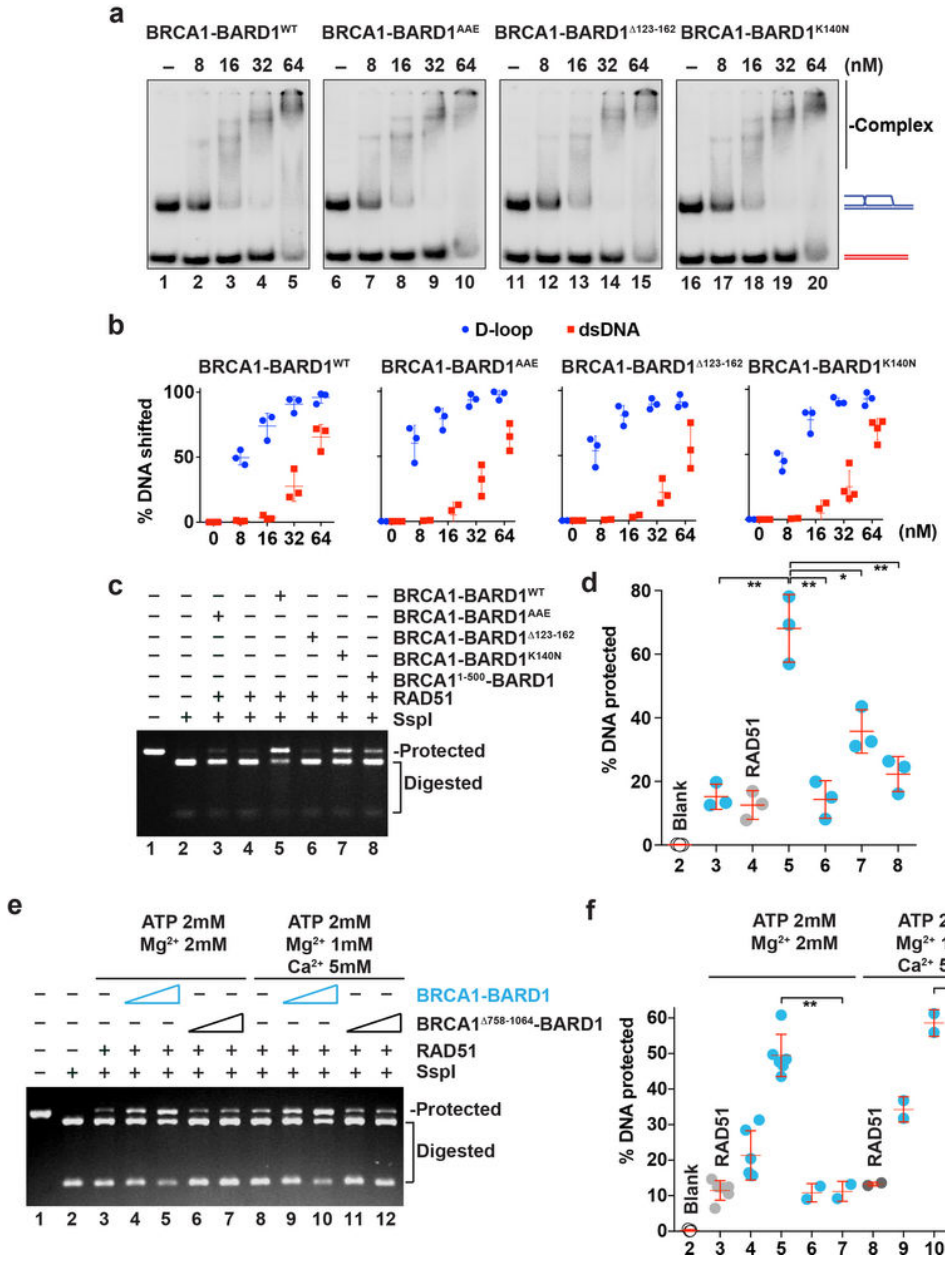

Figure 6.11: **Characterization of BRCA1-BARD1 mutants.** (**A**) BRCA1–BARD1 (n = 3), BRCA1–BARD1^{AAE} (n = 3), BRCA1–BARD1$^{\Delta 123–162}$ (n = 3), and BRCA1–BARD1^{K140N} (n = 4) were tested for their DNA binding activity using a mixture of radiolabelled D-loop and dsDNA as substrates. (**B**) Quantification of data from experiments in A. Data are means ± s.d. (**C**) Wild-type and mutant variants of BRCA1–BARD1 (300 nM each) were tested for the ability to promote synaptic complex formation. (**D**) Quantification of data from experiments in C. Data are means ± s.d., n = 3. (E) Synaptic complex formation by RAD51–ssDNA filament with BRCA1–BARD1 (100 and 200 nM) and BRCA1 758–1064–BARD1 (100 and 200 nM). (**F**) Quantification of data from experiments in E. Data are means ± s.d., n = 6 (BRCA1–BARD1 with Mg^{2+} and ATP) or n = 2 (all other conditions). *P < 0.05; **P < 0.01.

6.3.5 Cellular Role of the BRCA1-BARD1-RAD51 Complex

We conducted cell-based studies to investigate the association between BRCA1–BARD1 and RAD51 and to ascertain the importance of the BRCA1–BARD1–RAD51 complex. The amount of RAD51 that co-immunoprecipitated with wild-type siRNA-resistant BARD1 (BARD1^{WT}*res*) was increased by treatment of cells with mitomycin C (MMC) (Fig. 6.12A), and the BARD1^{AAE}*res* mutation impaired the DNA damage-induced association with RAD51 (Fig. 6.12A). Cellular fractionation confirmed that the nuclear localization of BRCA1 and BARD1 is not affected by the BARD1AAEres mutation (Fig. 6.13A).

Next, we used the direct repeat-green fluorescent protein (DR-GFP) reporter [217, 456], which measures DSB-induced homologous recombination, and a CRISPR–Cas9-stimulated gene-targeting assay [457, 458] to investigate whether the BARD1^{AAE}*res* mutation affects homologous recombination. As expected, knockdown of endogenous BRCA1 or BARD1 using small inhibitory RNA (siRNA) impaired homologous recombination in both systems (Fig. 6.13B-D). Notably, although the ectopic expression of BARD1^{WT}*res* in BARD1-deficient cells fully restored homologous recombination, ectopic expression of BARD1^{AAE}*res* resulted in only partial complementation (Fig. 6.12B and Fig. 6.14A,B). Moreover, in clonogenic cell survival assays, BARD1-deficient cells expressing BARD1^{AAE}*res* were markedly more sensitive than cells expressing BARD1^{WT}*res* to MMC and to the poly-ADP ribose polymerase (PARP) inhibitor olaparib (Fig. 6.12D and Fig. 6.14C).

We also investigated whether the BARD1^{AAE}*res* mutation would affect the DNA damage-induced assembly of RAD51 nuclear foci. As expected, knockdown of endogenous BRCA1 diminished RAD51 focus formation, either spontaneously or after γ-ray exposure

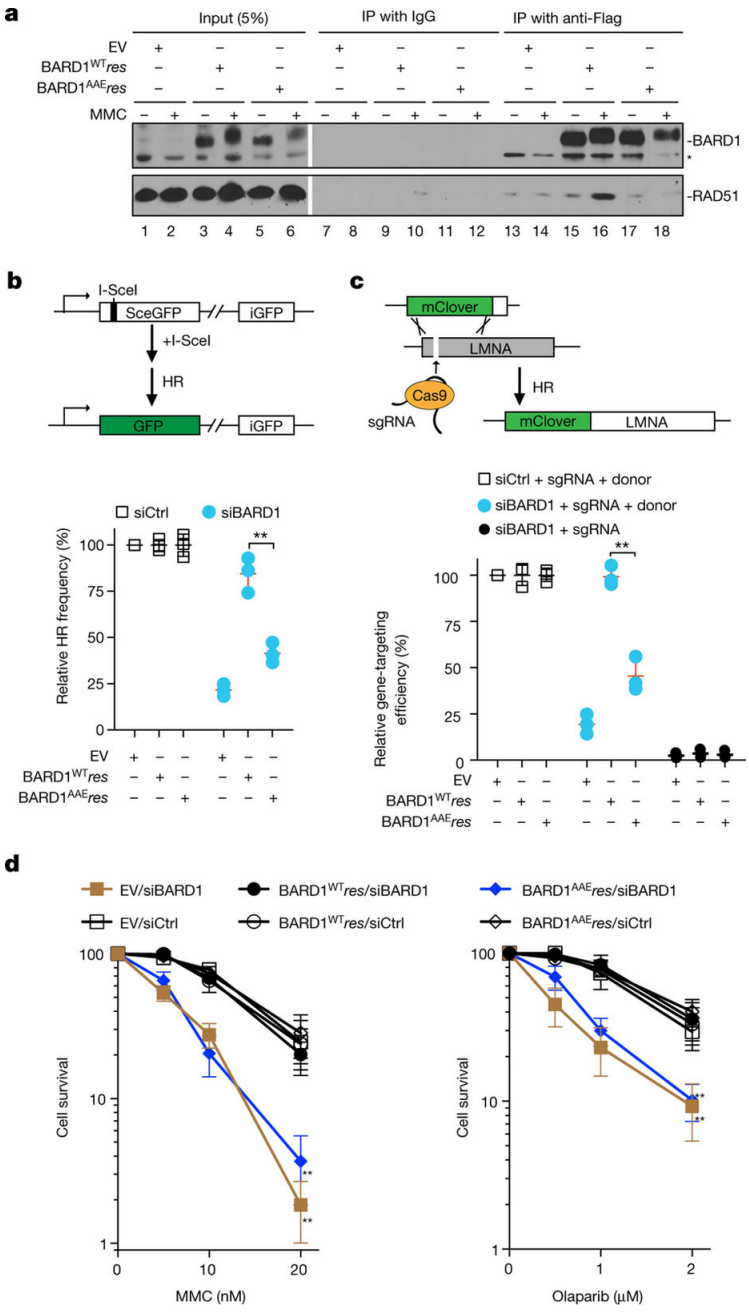

Figure 6.12: **Biological relevance of the BARD1-RAD51 complex.** (**A**) Immunoprecipitation to test BARD1WT and BARD1^{AAE} for RAD51 association after treatment with MMC. Asterisk denotes a non-specific band. (**B**) Schematic of the DR-GFP reporter assay (top). Results obtained with cells expressing BARD1^{WT}*res* or BARD1^{AAE}*res* upon treatment with BARD1 siRNA or control siRNA (siCtrl) (bottom). Data are means ± s.d., n = 3. (**C**) Schematic of the CRISPR–Cas9 gene targeting assay (top). Results obtained with cells expressing BARD1^{WT}*res* or BARD1^{AAE}*res* upon treatment with BARD1 siRNA or siCtrl. Data are means ± s.d., n = 3. (**D**) Clonogenic survival of cells expressing BARD1WTres or BARD1AAEres after treatment with olaparib or MMC. Data are means ± s.d., n = 3. EV, empty vector. P values were calculated using two-way ANOVA and multiple comparisons were corrected by the Bonferroni method. **P < 0.01.

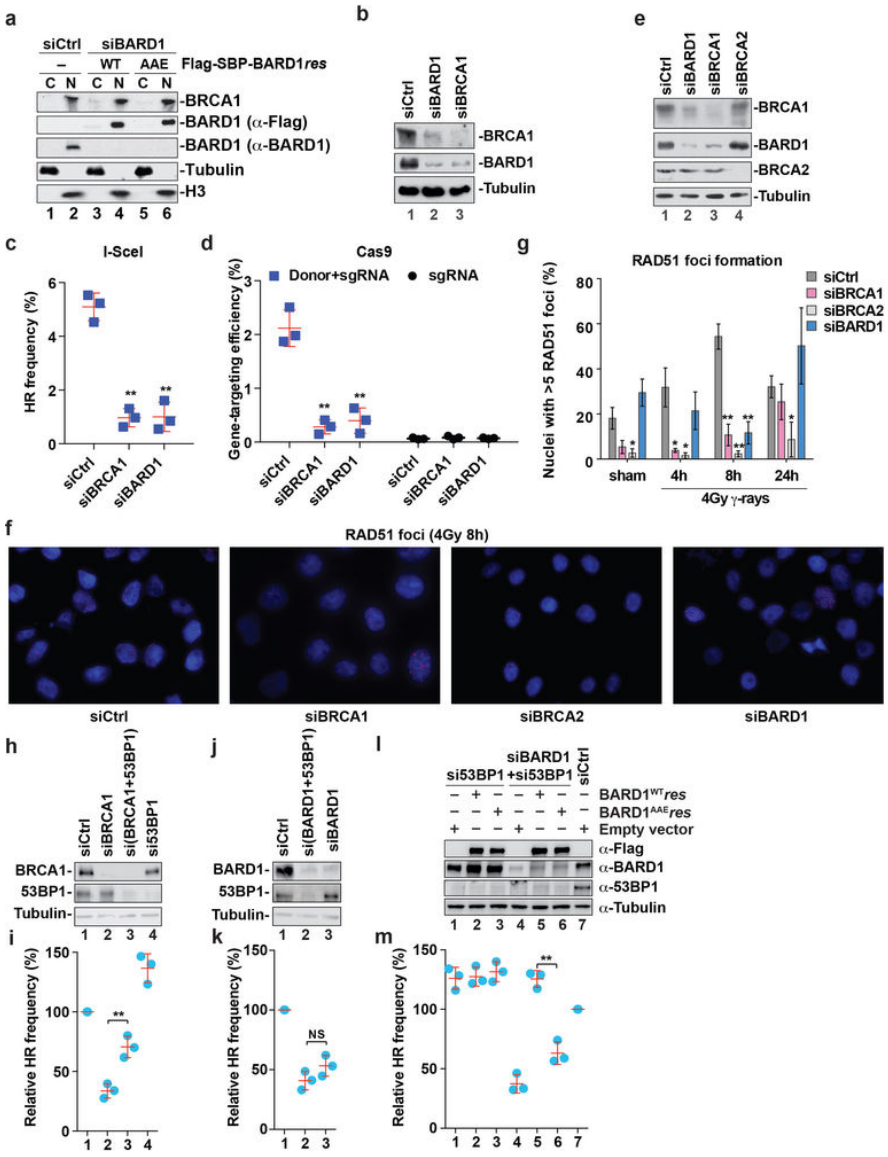

Figure 6.13: **Role of BRCA1 and BARD1 in homologous recombination and RAD51 focus formation.** (**A**) Western blot to verify the nuclear localization of endogenous BRCA1 and ectopically expressed Flag-SBP-tagged BARD1 or the AAE mutant in HeLa cells. The cytoplasmic (C) and nuclear (N) fractions were also analyzed for their alpha-tubulin and histone H3 contents. (**B**) Western blot analysis to detect endogenous BRCA1 and BARD1 after treatment of DR-U2OS cells with BRCA1 or BARD1 siRNA. (**C**) Homologous recombination frequency in DR-U2OS cells with siRNA-mediated knockdown of BRCA1 or BARD1. Data are means ± s.d., n = 3. (**D**) Gene-targeting efficiency of CRISPR–CAS9 in U2OS cells with siRNA-mediated knockdown of BRCA1 or BARD1. Data are means ± s.d., n = 3.

Figure 6.13: (cont.) (**E**) Western blot analysis to detect endogenous BRCA1, BARD1 and BRCA2 after treatment of HeLa cells with siRNA against BRCA1, BARD1 or BRCA2. Alpha-tubulin serves as loading control. (**F**) Representative micrographs of RAD51 foci (red) in the nuclei of HeLa cells treated with BRCA1, BARD1, BRCA2 or control siRNA 8 h after exposure to 4 Gy γ-rays. Blue, DAPI. (**G**) Quantification of RAD51 foci at various time points after exposure to 4 Gy γ-rays or sham irradiation. The mean values ± s.e.m. of 4 (siBRCA2 and siBARD1), 6 (siBRCA1) or 7 (siControl) independent experiments are shown. (**H**) Western blot analysis to detect endogenous BRCA1 and 53BP1 after treatment of DR-U2OS cells with BRCA1 or TP53BP1 siRNA. (**I**) Homologous recombination frequency in DR-U2OS cells with siRNA-mediated knockdown of BRCA1 and/or TP53BP1. Data are means ± s.d., n = 3. (**J**) Western blot analysis to detect endogenous BARD1 and 53BP1 after treatment of DR-U2OS cells with BARD1 and/or TP53BP1 siRNA. (**K**) Homologous recombination frequency in DR-U2OS cells with siRNA-mediated knockdown of BARD1 or TP53BP1. Data are means ± s.d., n = 3. (**L**) Western blot analysis to detect ectopically expressed and endogenous BARD1 after treatment of U2OS cells with BARD1 and/or TP53BP1 siRNA. As the abundance of ectopically expressed Flag-SBP-tagged wild-type and mutant BARD1 was lower than that of endogenous BARD1, we revealed it with anti-Flag antibodies in western blot analysis. (**M**) Homologous recombination frequency in DR-U2OS cells treated with siRNA against BARD1 and/or TP53BP1 and stably expressing BARD1^{WT}*res* or BARD1^{AAE}*res*. Data are means ± s.d., n = 3. *P < 0.05; **P < 0.01; NS, non-significant.

(Fig. 6.13E-G). However, treatment with siRNA targeting BARD1 impaired RAD51 focus formation to a lesser extent (Fig. 6.13G). In cells depleted of endogenous BARD1 and expressing BARD1^{WT}*res* or BARD1^{AAE}*res*, RAD51 focus formation occurred similarly, both spontaneously and after γ-irradiation (Fig. 6.14D,E). However, as indicated by S4/S8 phosphorylation of RPA32, BARD1AAEres cells retained a much higher level of DNA damage 72 h after release from MMC treatment (Fig. 6.14F). These results showed that homologous recombination-mediated repair is deficient in BARD1^{AAE}*res*-expressing cells despite the fact that RAD51 focus formation is not affected. Even though depletion of 53BP1, an inhibitor of DNA end resection [459], partially overcame the homologous recombination defect associated with BRCA1 deficiency (Fig. 6.13H,I), it did not suppress the homologous recombination defect in BARD1-deficient cells (Fig. 6.13J,K). In cells lacking both BARD1 and 53BP1, BARD1^{AAE}*res* was less able than BARD1^{WT}*res*to reverse the homologous recombination deficiency (Fig. 6.13L,M). Together, our results help to establish the biological importance

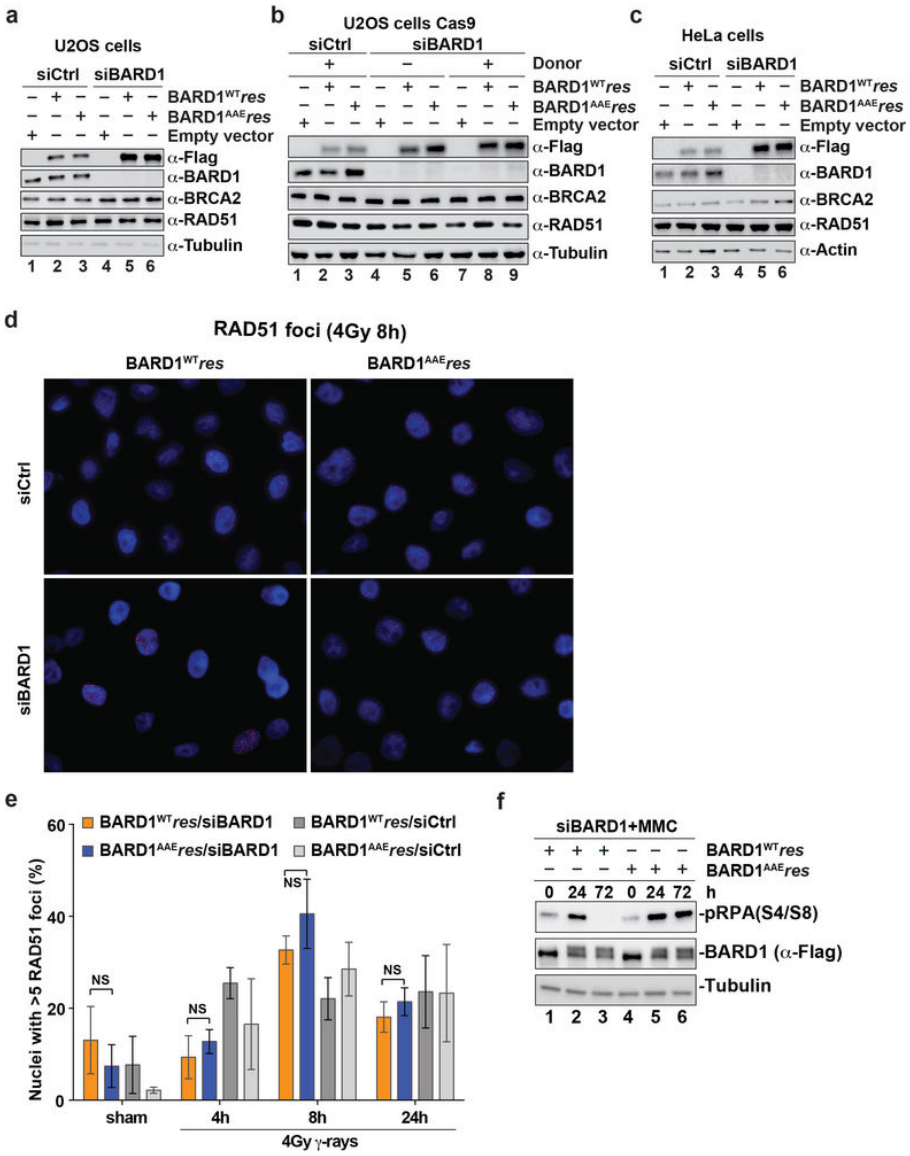

Figure 6.14: **Characterizations of human cells expressing BARD1 mutants.** (**A**) Western blot analysis to detect ectopically expressed and endogenous BARD1 after treatment of U2OS cells with BARD1 or control siRNA for the experiments in Fig. 6.12B. (**B**) Western blot analysis to detect ectopically expressed and endogenous BARD1 after treatment of U2OS cells with BARD1 or control siRNA for the experiments in Fig. 6.12C. (**C**) Western blot analysis to detect ectopically expressed and endogenous BARD1 after treatment of HeLa cells with BARD1 or control siRNA for the experiments in Fig. 6.12D. In a–c, as the abundance of ectopically expressed Flag-SBP-tagged wild-type and mutant BARD1 was lower than that of endogenous BARD1, we revealed it with anti-Flag antibodies in western blot analysis. (**D**) Representative micrographs of RAD51 foci (red) in the nuclei of HeLa cells expressing Flag-SBP-tagged BARD1WTres or BARD1AAEres 8 h after exposure to 4 Gy γ-rays. Blue, DAPI. (**E**) Quantification of RAD51 foci at various time points after exposure to 4 Gy γ-rays or sham irradiation. The mean values $\pm$ s.e.m. of 5 (8-h time point) or 3 (all other time points) independent experiments are shown. NS, non-significant. (**F**) Western blot to reveal pRPA32(S4/S8) (with tubulin as the loading control) at various time points (0, 24 and 72 h) after a 1-h treatment with 2 μM MMC.

of the BRCA1–BARD1–RAD51 complex in DNA damage repair via homologous recombination and provide cellular evidence for a role of BRCA1–BARD1 in the DNA strand invasion step of homologous recombination.

6.3.6 Cellular Role of BRCA1 in RAD51-Mediated DNA Pairing

To investigate the role of BRCA1 in RAD51-mediated reactions, we expressed and purified BRCA1$^{1–500}$–BARD1, which lacks the RAD51 interaction and DNA binding domains of BRCA1 [432, 446] and also BRCA1$^{\Delta 758–1064}$–BARD1, in which the RAD51-interaction domain of BRCA1 has been deleted. These mutant complexes appeared to be proficient in DNA binding (Fig. 6.15A-E) but were weakened for RAD51 interaction (Fig. 6.15F-H). Importantly, neither mutant complex could strongly enhance RAD51-mediated D-loop formation (Fig. 6.15I,J) or synaptic complex assembly (Fig. 6.7G and 6.11C-F). Thus, BRCA1 is also indispensable for the functional integrity of BRCA1–BARD1 as a co-factor of RAD51.

6.4 Discussion

Our study has revealed that BRCA1–BARD1 enhances DNA invasion in homologous recombination by interacting directly with RAD51 (Fig. 6.16). Mechanistically, BRCA1–BARD1 functions with the RAD51 presynaptic filament in the assembly of the synaptic complex, a critical precursor to D-loop formation (Fig. 6.16). Both BRCA1 and BARD1 are indispensable for this attribute. It is likely that physical association with RAD51 facilitates dsDNA engagement by the presynaptic filament and that the specific recognition of unwound

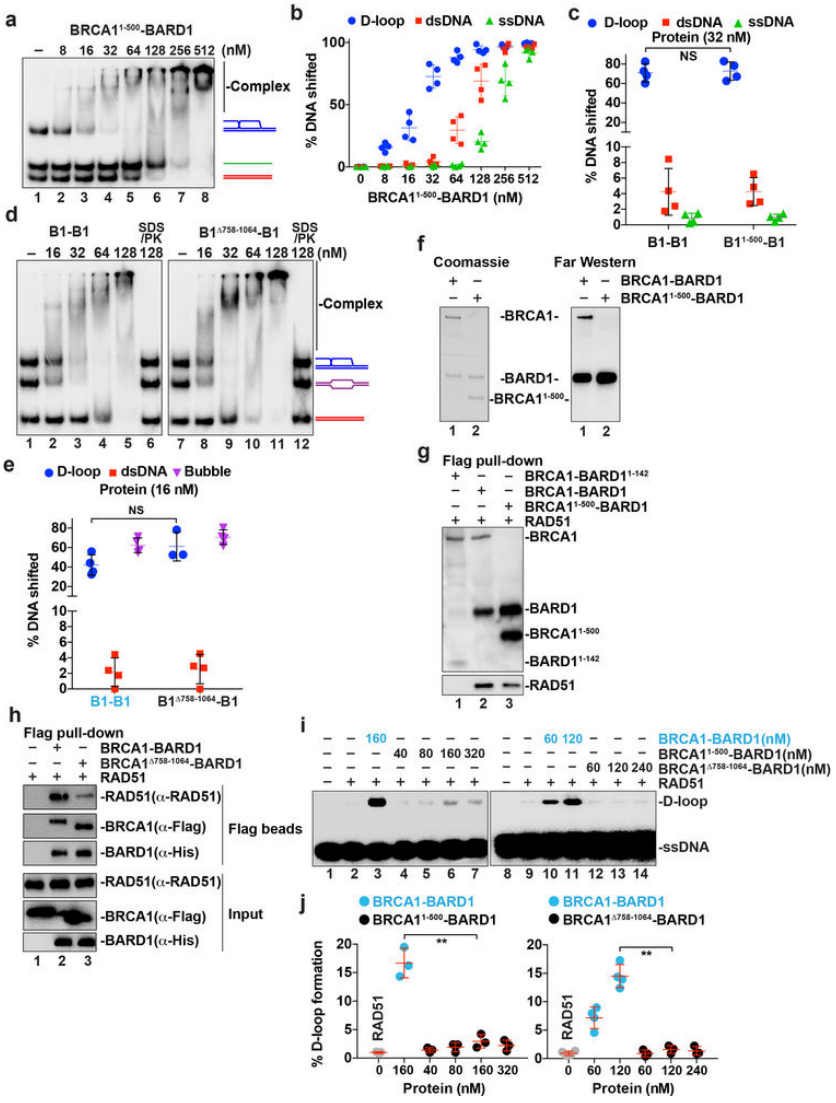

Figure 6.15: **Characterizations of BRCA1$^{1–500}$-BARD1 and BRCA1$^{\Delta 758–1064}$–BARD1.** (**A**) BRCA1$^{1–500}$–BARD1 was tested for DNA binding using a mixture of radiolabelled D-loop, dsDNA, and ssDNA as substrates. (**B**) Quantification of data from experiments in A. Data are means ± s.d., n = 4. (**C**) Comparison of results obtained using 32 nM BRCA1–BARD1 (from **??**G) and BRCA1$^{1–500}$–BARD1 (from A). Data are means ± s.d., n = 3. NS, non-significant. (**D**)BRCA1–BARD1 and BRCA1$^{\Delta 758–1064}$–BARD1 were tested for DNA binding using a mixture of radiolabelled D-loop, bubble, and dsDNA as substrates. (**R**)Comparison of results obtained using 16 nM BRCA1–BARD1 and BRCA1$^{\Delta 758–1064}$–BARD1. Data are means ± s.d., n = 4. NS, non-significant. (**F**) Far western analysis to detect RAD51 association with BRCA1$^{1–500}$ and BARD1 immobilized on nitrocellulose membrane. (**G**) Pull-down assay to test for the interaction of RAD51 with BRCA1$^{1–500}$–BARD1, BRCA1–BARD1$^{1–142}$ and BRCA1–BARD1 via the Flag tag on the BRCA1 species. The eluates from the various anti-Flag resin fractions were subjected to immunoblot analysis with anti-Flag (for BRCA1), anti-His (for BARD1) and anti-RAD51 antibodies. (**H**)Pull-down assay to test for the interaction between RAD51 and BRCA1–BARD1 or BRCA1$^{\Delta 758–1064}$–BARD1 via the Flag tag on the BRCA1 species. (**I**) BRCA1$^{1–500}$–BARD1 and BRCA1$^{\Delta 758–1064}$–BARD1 were tested along with the wild-type complex for the ability to enhance RAD51-mediated D-loop formation. (**J**) Quantification of data from experiments in I. Data are means ± s.d., n = 3 or 4. **P < 0.01..

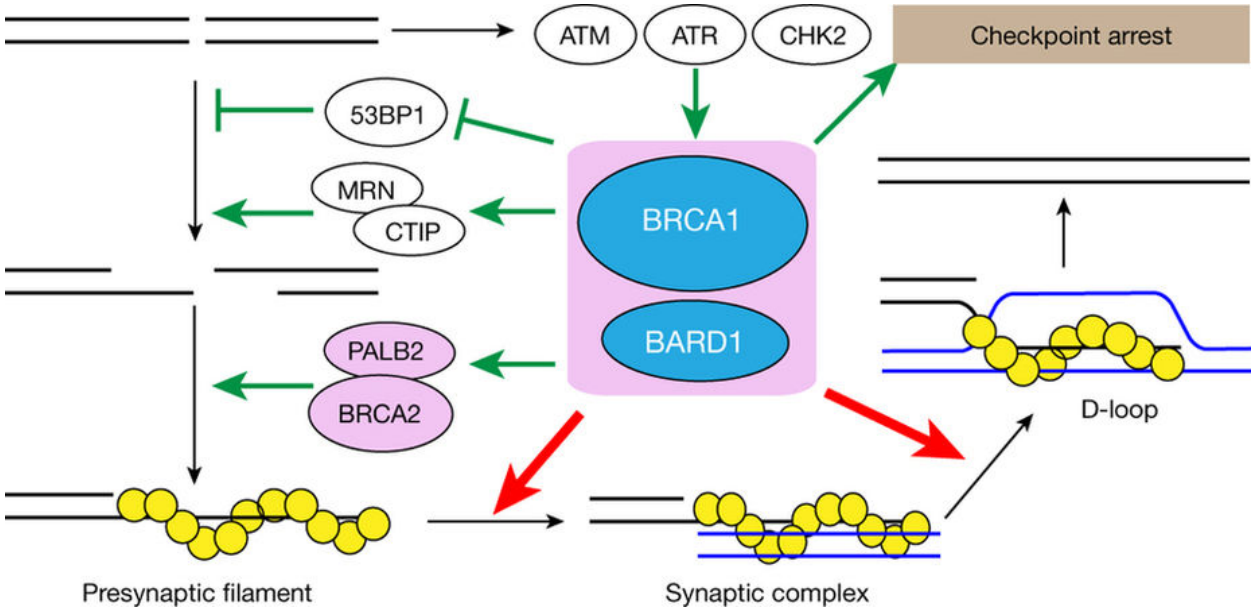

Figure 6.16: **Role of BRCA1 and BARD1 in homologous recombination and RAD51 focus formation.** (**A**) Western blot to verify the nuclear localization of endogenous BRCA1 and ectopically expressed Flag-SBP-tagged BARD1 or the AAE mutant in HeLa cells. The cytoplasmic (C) and nuclear (N) fractions were also analyzed for their alpha-tubulin and histone H3 contents. (**B**) Western blot analysis to detect endogenous BRCA1 and BARD1 after treatment of DR-U2OS cells with BRCA1 or BARD1 siRNA. (**C**) Homologous recombination frequency in DR-U2OS cells with siRNA-mediated knockdown of BRCA1 or BARD1. Data are means ± s.d., n = 3. (**D**) Gene-targeting efficiency of CRISPR–CAS9 in U2OS cells with siRNA-mediated knockdown of BRCA1 or BARD1. Data are means ± s.d., n = 3.

DNA by BRCA1–BARD1 enhances the formation of the nascent DNA joint in the D-loop reaction. As we have not observed significant stimulation by BRCA1–BARD1 of DNA strand exchange between an oligonucleotide and a short linear duplex, it remains possible that the complex facilitates DNA homology search within a long duplex DNA target. A model incorporating this new function and the known roles of BRCA1–BARD1 in DNA end resection and RAD51 presynaptic filament assembly is shown in (Fig. 6.16).

The findings from our study open up a new avenue towards understanding how mutations in BRCA1–BARD1 affect its DNA damage repair and tumor suppression functions. Indeed, we have provided evidence that the cancer-associated mutation K140N in BARD1

compromises the physical and functional interactions of BRCA1–BARD1 with RAD51. We note that the region of BRCA1 (amino acid residues 758–1,064) that harbors the RAD51-interaction domain [432] is frequently mutated in cancer (cBioPortal for Cancer Genomic) [454, 455] and that deletion of this domain abolishes the activity of BRCA1–BARD1 in RAD51-mediated DNA strand invasion. The biochemical systems established in our work should be valuable for determining the impact of pathogenic mutations on BRCA1–BARD1 functions. Moreover, our findings may guide the development of targeted therapies for breast, ovarian, and other cancers.

6.5 Methods

6.5.1 Construction of plasmids

A His6 affinity tag was fused to BARD1 in pFastbac-BARD1 (from J. Parvin) using the QuikChange mutagenesis kit (Stratagene). The mammalian pS-Flag-SBP-BARD1res expression vector was modified by removing the GFP coding sequence from the pS-Flag-SBP-BARD1 vector (from X. Yu) and introducing silent mutations into the siRNA target regions of BARD1using oligos 1 (5'-GAT GAT AAT ATG GCC ACA ACC AGC GGC CGC GAC TCT AGA TC-3') and 2 (5'-GAT CTA GAG TCG CGG CCG CTG GTT GTG GCC ATA TTA TCA TC-3') and oligos 3 (5'-GAA AGT CAG ATA TGT TGT GAG CAA GGC AAG TGT CCA GAC CCA GCC TGC AAT AAA AA-3') and 4 (5'-TTT TTA TTG CAG GCT GGG TCT GGA CAC TTG CCT TGC TCA CAA CAT ATC TGA CTT TC-3'), respectively. QuikChange site-directed mutagenesis was used to construct the mutant

forms of BARD1: BARD1$^{1–122}$, BARD1$^{1–162}$, BARD1$^{1–261}$, BARD1$^{\Delta 123–162}$, BARD1$^{\Delta 123–261}$, BARD1$^{\Delta 163–261}$, BARD1^{AAE} and BARD1^{K140N} (the sequences of the primers used are available upon request). BARD1$^{123–162}$was introduced into pDEST15 for expression of the GST-tagged form of this BARD1 fragment in E. coli. BARD1$^{124–270}$ was cloned into pE-SUMO vector (LifeSensors Inc.) for expression of the SUMO-tagged form of this BARD1 domain in *E. coli*.

6.5.2 Protein purification: purification of BRCA1–BARD1 from insect cells

pFastbac-Flag-BRCA1 (from J. Parvin) and pFastbac-His-BARD1 were introduced into E. coli strain DH10Bac for bacmid generation. The bacmids were used to transfect SF9 insect cells to generate recombinant baculoviruses. After amplification in SF9 cells, the viruses were used to infect Hi5 insect cells for expression of BRCA1 and BARD1 (10 ml BRCA1 and 10 ml BARD1 P3 viruses for 600 ml culture). After a 44-h incubation at 27 °C, cells were harvested by centrifugation, frozen in liquid nitrogen, and stored at −80 °C. All purification steps were carried out at 0–4 °C. To prepare extract, the frozen cell pellet (8 g, from 600 ml culture) was thawed and suspended in 40 ml cell breakage buffer A (50 mM Tris-HCl, pH 7.5, 500 mM KCl, 1 mM 2-mercaptoethanol, 0.5% NP-40, 5 mM $MgCl_2$, 2 mM ATP and the following protease inhibitors: aprotinin, chymostatin, leupeptin, and pepstatin A at 3 μg ml^{-1} each, and 1 mM PMSF) for cell lysis using a Dounce homogenizer type B pestle (30 strokes). The lysate was cleared by centrifugation at 10,000g for 15 min, and the supernatant was incubated with 3 ml anti-Flag M2 affinity resin (Sigma) for 2 h. The

resin was transferred to a column (1.5×15 cm), washed with 50 ml lysis buffer and then with 50 ml buffer B (25 mM Tris-HCl, pH 7.5, 300 mM KCl, 10% glycerol, 0.5 mM EDTA, 0.01% Igepal CA-630, 1 mM 2-mercaptoethanol, 5 mM $MgCl_2$ and 2 mM ATP), before the bound proteins were eluted four times with 2 ml buffer B containing the single Flag peptide (200 μg ml^{-1}). The eluates were combined and mixed with 32 ml buffer C (25 mM Tris-HCl, pH 7.5, 10% glycerol, 0.5 mM EDTA, 0.01% Igepal CA-630, 1 mM 2-mercaptoethanol) before being further fractionated in a 1 ml HiTrap SP Sepharose HP column (GE Healthcare) using a 12 ml gradient of 75–500 mM KCl in buffer C. The pooled BRCA1–BARD1 fractions (250–350 mM KCl) were further fractionated in a gel filtration column of Superose 6 10/300 GL (GE Healthcare), which was developed with 24 ml buffer C containing 300 mM KCl. The peak fractions were pooled, divided into 10- l portions, frozen in liquid nitrogen, and stored at −80 °C. The mutant forms of BRCA1–BARD1 were expressed and purified using the same procedures. The yield of highly purified BRCA1–BARD1 from 600 ml insect cell culture ranged from 150 to 300 g with a final concentration of 300–500 μg ml^{-1}.

Protein purification: purification of BARD1$^{123–162}$ and BARD1$^{124–270}$ from *E. coli* The GST–BARD1$^{1–162}$ expression plasmid pDEST15–BARD1$^{1–162}$ or the BARD1124–270 expression plasmid pET-SUMO-BARD1124–270 was introduced into E. coli Rosetta (DE3) cells. An overnight culture derived from a single colony in 50 ml LB medium grown at 37 °C was used to inoculate 2 l fresh LB medium. IPTG was added to 0.4 mM when the cell density had reached OD600 = 0.8, and cells were harvested after a 16-h incubation at 16 °C. All the subsequent steps were carried out at 0–4 °C. The cell pellet (8 g) was suspended in 50 ml buffer D (20 mM KH2PO4, pH 7.5, 10% glycerol, 0.5 mM EDTA, 0.01% Igepal CA-630, 1 mM 2-mercaptoethanol and 300 mM KCl) containing the protease inhibitors (aprotinin,

chymostatin, leupeptin, and pepstatin A at 3 g ml−1 each, and 1 mM PMSF) and cell lysate was prepared by sonication. After centrifugation (100,000g for 90 min), the clarified lysate was incubated with 2 ml Glutathione Sepharose 4 Fast Flow resin (GE Healthcare; for GST-BARD1$^{123–162}$) or Ni-NTA resin (GE healthcare; for BARD1$^{124–270}$) for 2 h. The affinity resin was transferred to a glass column (1.5×15 cm) and washed with 20 ml buffer D before being eluted three times with 3 ml of 20 mM glutathione or 150 mM imidazole in buffer D. For BARD1$^{124–270}$, the His6-SUMO tag was cleaved by the Ulp1 protease by overnight incubation at 4 °C. The eluates were pooled and concentrated in a Centricon-10K concentrator (Amicon) to 0.5 ml before being further fractionated in a Superdex 200 10/300 GL column (GE Healthcare) with 24 ml of buffer C containing 300 mM KCl. The peak fractions were pooled, concentrated to ~ 100 μl as above, divided into 5 l portions, frozen in liquid nitrogen, and stored at −80 °C.

6.5.3 Protein purification: other recombination proteins

BRCA2–DSS1, RAD51, RPA and yeast Rad51 were purified to near homogeneity using our previously described procedures [185, 409, 449].

6.5.4 DNA substrates and DNA binding assay

D-loop, DNA bubble, replication fork and double-stranded DNA were assembled from oligonucleotides 5/6/7, oligonucleotides 5/6, oligonucleotides 8/9/10/11 and oligonucleotides 12/13, respectively; the asterisk identifies the oligonucleotide that was 32P-labelled at its 5 end in each substrate. The single-stranded DNA substrate was 5 32P-labelled oligonu-

cleotide 12. These DNA substrates (10 nM each) were incubated with wild-type or the specified mutant form of BRCA1–BARD1 at 37 °C in 10 μl buffer E (25 mM Tris-HCl, pH 7.5, 90 mM KCl, 1 mM DTT, and 100 μg ml^{-1} bovine serum albumin (BSA)) for 10 min. After the addition of loading buffer (50% glycerol, 20 mM Tris-HCl, pH 7.4, 0.5 mM EDTA, 0.05% Orange G), the reaction mixtures were resolved by 6% native polyacrylamide gel electrophoresis in TAE buffer (30 mM Tris-acetate, pH 7.4 and 0.5 mM EDTA) at 4 °C. The gels were dried, and DNA species were visualized by autoradiography and quantified using the Personal Molecular Imager and Quantity One software (Bio-Rad). Since the nucleoprotein complexes formed by BRCA–1-BARD1 do not always migrate as well-defined species, we quantified DNA binding by measuring the disappearance of the DNA substrate.

Oligo 5: 5'-CAT TGC ATA TTT AAA ACA TGT TGG AA GGC TCG ATG CAT GCT GAT AGC CTA CTA GTG CT G CTG GCT TTC AAA TGA CCT CTT ATC AAG TGA C-3'

Oligo 6: 5'-GTC ACT TGA TAA GAG GTC ATT TGA ATT CAT GGC TTA GAG CTT AAT TGC TGA ATC TGG T CT GGG ATC CAA CAT GTT TTA AAT ATG CAA TG-3'

Oligo 7: 5'-CTG CTA CGA TGC TAG TCG TAG CTC GGC AGT CGT AGC AGG TTC CCA GCA CCA GAT TCA GC A ATT AAG CTC TAA GCC ATG AA-3'

Oligo 8: 5'-GAC GCT GCC GAA TTC TAC CAG TGC CTT GCT AGG ACA TCT TTG CCC ACC TGC AGG TTC ACC C-3'

Oligo 9: 5'-GGA CAT CTT TGC CCA CCT GCA GGT TCA CCC-3'

Oligo 10: 5'-TGG GTG AAC CTG CAG GTG GGC AAA GAT GTC C-3'

Oligo 11: 5'-GGG TGA ACC TGC AGG TGG GCA AAG ATG TCC CAG CAA

GGC ACT GGT AGA ATT CGG CAG CGT C-3'

Oligo 12: 5'-TTA TAT CCT TTA CTT TGA ATT CTA TGT TTA ACC TTT TAC TTA TTT TGT ATT AGC CGG ATC CTT ATT TCA ATT ATG TTC AT-3'

Oligo 13: 5'-ATG AAC ATA ATT GAA ATA AGG ATC CGG CTA ATA CAA AAT AAG TAA AAG GTT AAA CAT AGA ATT CAA AGT AAA GGA TAT AA-3'

6.5.5 Affinity pull-down

RAD51, yRad51 or RecA (5 μM) was incubated with 0.5 M Flag–BRCA1–BARD1 or 3 M GST–BARD1$^{123–162}$ at 4 °C for 30 min in 30 μl buffer F (25 mM Tris-HCl pH 7.5, 10% glycerol, 0.5 mM EDTA, 0.05% Igepal CA-630, 1 mM 2-mercaptoethanol, 150 mM KCl). Then the reaction mixture was mixed with 12 μl anti-Flag M2 affinity resin or Glutathione Sepharose 4 Fast Flow resin at 4 °C for 30 min to capture protein complexes through the Flag tag on BRCA1 or the GST tag on BARD1, respectively. After the resin was washed three times with 200 l buffer F, bound proteins were eluted with 20 l 2% SDS at 37 °C for 5 min. The supernatant (S), last wash (W) and SDS eluate (E), 8 l each, were analyzed by SDS–PAGE and Coomassie blue staining.

6.5.6 Southwestern analysis

BRCA1–BARD1 was resolved in a 7.5% SDS–PAGE gel and transferred onto a nitrocellulose membrane (Bio-Rad) at 4 °C in transfer buffer (25 mM Tris-HCl, 192 mM glycine, pH 8.3, 20% methanol). After being soaked in buffer G (25 mM Tris-HCl, pH 7.5, 2 mg ml^{-1} BSA, 1 mM DTT, 0.1% Triton X-100, 10% glycerol and 100 mM KCl) at 4 °C for 20 h,

the membrane was rinsed twice with buffer H (25 mM Tris-HCl, pH 7.5, 200 μg ml^{-1} BSA, 90 mM KCl, 4 mM $MgCl_2$ and 1 mM DTT) and then incubated in 10 ml buffer H containing 32P-labelled D-loop DNA (2 nM) at 25 °C for 1 h. The membrane was washed four times with 10 ml buffer H before analysis by phosphorimaging.

6.5.7 Far western analysis

After SDS–PAGE, BRCA1 and BARD1 were transferred onto a nitrocellulose membrane, as described for the southwestern analysis. The membrane was soaked in buffer I (10 mM KH2PO4 at pH 7.4, 150 mM KCl, 15 mg ml^{-1} BSA, 2 mM 2-mercaptoethanol, 0.05% Tween 20) at 25 °C for 2 h and then incubated with 5 μg ml^{-1} RAD51 in buffer I at 25 °C for 2 h. Then, the membrane was washed with 10 ml buffer I three times, incubated with anti-RAD51–HRP antibodies (Abcam, ab195548) for 1 h in buffer I, washed again with 10 ml buffer I three times, and developed with the Super Signal Substrate Kit (Pierce).

6.5.8 Homologous DNA pairing assay

The homologous DNA pairing assay was conducted as described [448, 449]. The reaction was assembled in buffer J (25 mM Tris-HCl, pH 7.5, 60 mM KCl, 1 mM DTT and 100 g ml−1 BSA) containing 1 mM ATP and 2 mM $MgCl_2$ in a final volume of 12.5 μl. For mediator activity, the 150-mer oligonucleotide 14 (6 μM nucleotides) was first incubated with RPA (600 nM) at 37 °C for 5 min, and then RAD51 (2 μM) with or without the indicated concentration of BRCA1–BARD1 or BRCA2–DSS1 was incorporated into the reaction. Following a 5-min incubation at 37 °C, 32P-labelled homologous dsDNA (40 bp; oligos 15 (5'-TAA

TAC AAA ATA AGT AAA TGA ATA AAC AGA GAA AAT AAA G-3') and 16 (5'-CTT TAT TTT CTC TGT TTA TTC ATT TAC TTA TTT TGT ATT A-3'); 1.6 μM base pairs) and 4 mM spermidine hydrochloride were added. For testing of ssDNA targeting activity, RAD51 was incubated with the mixture of ssDNA, 32P-labelled dsDNA and spermidine hydrochloride with and without BRCA1–BARD1 or BRCA2–DSS1 for 30 min. The reaction was terminated by adding an equal volume of 1% SDS containing 1 mg ml^{-1} proteinase K. Following a 5-min incubation at 37 °C °C, the deproteinized reaction mixtures were resolved in an 8% non-denaturing polyacrylamide gel in TAE buffer. The gel was dried onto 3MM CHR cellulose chromatography papers (GE Healthcare), and DNA species were visualized by autoradiography and quantified, as above.

6.5.9 D-loop assay

The D-loop assay was conducted as described [453, 460]. In brief, the 32P-labelled 90-mer oligonucleotide 17 (5'-AAA TCA ATC TAA AGT ATA TAT GAG TAA ACT TGG TCT GAC AGT TAC CA A TGC TTA ATC AGT GAG GCA CCT ATC TCA GCG ATC TGT CTA TTT-3'; 2.4 M nucleotides) was incubated with RAD51 (1 M) at 37 °C for 5 min in buffer J containing 1 mM MgCl2 and 2 mM ATP or AMP-PNP. Following the incorporation of the indicated concentration of BRCA1–BARD1 and a 5-min incubation at 37 °C, the D-loop reaction was initiated by adding pBluescript SK replicative form I DNA (37 M base pairs) and was incubated at 37 °C for 7 min. The molar ratio of the 90-mer to pBluescript plasmid in the reactions was 2.1 to 1. The reaction was terminated by adding an equal volume of 1% SDS containing 1 mg ml−1 proteinase K and a 5-min incubation at

37 °C. The deproteinized reaction mixtures were resolved by electrophoresis in a 1% agarose gel, which was dried onto Hybond-N membrane (GE Healthcare). Phosphorimaging analysis was used to visualize and quantify the radiolabelled DNA species.

6.5.10 Synaptic complex assay

The synaptic complex assay was conducted at 37 °C as described5,6. In brief, RAD51 (4 M) was incubated with the 60-mer oligonucleotide 18 (12 M nucleotide; 5'-AAT GTT GAA TAC TCA TAC TCT TCC TTT TTC AAT ATT ATT GAA GCA TTT ATC AGG GTT ATT-3'), which is homologous to the SspI restriction site in the target pUC19 dsDNA, in 8 l buffer K (35 mM Tris-HCl, pH 7.5, 50 mM KCl, 2 mM ATP, 2 mM MgCl2, 100 g ml−1 BSA, and 1 mM DTT) for 5 min. After adding the indicated amounts of BRCA1–BARD1 in 1 l volume, the reaction mixture was incubated for 5 min. Then, linear pUC19 plasmid DNA (85 M nucleotides) was added in 1 l, followed by a 5-min incubation and treatment with 2.5 units of SspI for 10 min. The reaction mixtures were resolved by agarose gel electrophoresis in TAE buffer, and DNA species were stained with ethidium bromide. The heterologous oligonucleotide 19 (5'-CAG AAT CAG GGG ATA ACG CAG GAA AGA ACA TGT GAG CAA AAG GCC AGC AAA AGG CCA GGA-3') was used as a control.

6.5.11 DNA curtain imaging analysis

RAD51 filaments were assembled on ssDNA curtains, and the dsDNA-binding properties of the resulting RAD51–ssDNA filaments were measured as described39,40. To determine the number of dsDNA-binding events, BRCA1–BARD1 was diluted from a 1.5 M stock

with buffer K (30 mM Tris–acetate pH 7.5, 1 mM MgCl2, 5 mM CaCl2, 100 mM KCl, 2 mM ATP, 1 mM DTT, and 0.2 mg ml−1 BSA) to concentrations ranging from 10 to 100 nM, introduced into the flow cell chamber, and incubated for 10 min at 37 °C. After washing with 1 ml buffer K (at 1 ml min−1), 2.16 nM Atto565-labelled dsDNA (70 bp) with 9 nt of homology (oligos 20 (5'-Atto565-CCG GAG GCC TTA GGC CTT AGG CCT TAG GCC TTC AGC TGT TAG CCT TAG CTA GCT AGC TAG CTA GCT AGC T-3'; the underlined sequence is homologous to the ssDNA substrate) and 21 (5'-AGC TAG CTA GCT AGC TAG CTA GCT AAG GCT AAC AGC TGA AGG CCT AAG GCC TAA GGC CTA AGG CCT CCG G-3') to the RAD51–ssDNA filaments was introduced into the chamber, followed by a 10-min incubation at 37 °C. Then, the cell was washed with 0.5 ml buffer K (at 1 ml min−1) and three images were taken. The length of each of the RAD51–ssDNA filaments and the number of labeled dsDNA molecules bound were recorded and normalized to a length of 50 pixels (40 kb). The weighted average and standard deviation based on the length of each filament were calculated. Confidence intervals of 95% are represented as error bars. For survival probabilities, experiments were conducted without or with 100 nM BRCA1–BARD1 and 100 ms exposures were recorded every 30 s over 90 min. Dwell times of 180 molecules were determined for each experiment by kymograph and survival probabilities were plotted on a semi-log plot. Error bars represent 70% confidence as measured by bootstrap analysis, a close approximation of one standard deviation from the mean.

6.5.12 Cell culture and transfection

U2OS and HeLa cells from ATCC were grown in Dulbecco's modified Eagles medium (DMEM) supplemented with 10% fetal bovine serum (Sigma), 100 g ml−1 streptomycin, and 100 U ml−1 penicillin (Sigma). The cells were tested for mycoplasma contamination by Bionique testing labs (http://www.bionique.com/). Control siRNA (UAGCCGGUAGACUUAGGUCUG), BARD1 siRNA (AAGAGUAAAGCUUCAGUGCAA), BRCA1 siRNA (AAGCUCCUCUCACUCUUCAGU) and BRCA2 siRNA (UUGGAGGAAUAUCGUA GGUAA) oligonucleotides were purchased from Qiagen. TP53BP1 siRNA (s14313) was purchased from Ambion-Thermo Fisher Scientific. Transfection of siRNA, pS-Flag-SBP-BARD1resand pCMV-I-SceI-3×NLS was carried out using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. To generate stable HeLa and U2OS cell lines expressing Flag–SBP–BARD1 or its mutants, cells were transfected with their respective plasmids (pS-Flag–SBP–BARD1WTres, pS-Flag–SBP–BARD1AAEres, and pS-Flag–SBP–BARD1K140Nres) and individual clones were selected with 800 g ml−1G418.

6.5.13 Co-immunoprecipitation analysis

HeLa cells grown on 15-cm cell culture dishes were treated with or without 1 M MMC overnight before collection. Following a wash with PBS, cells were scraped off and transferred to Eppendorf tubes. Whole cell lysate was prepared by adding 1 ml lysis buffer (300 mM NaCl, 1.0% Triton X-100, 5 mM EDTA, 2 mM NaVO4, 2 mM Na4O7P2, 0.02% NaN3, and 50 mM Tris-HCl, pH7.4) with protease inhibitors (Roche Complete Protease Inhibitor Cock-

tail Tablet) to cell pellets. Following a 12-s sonication, the cell extract was cleared by centrifugation at 18,400g for 15 min at 4 °C. The supernatant fraction (2 mg protein in total) was incubated with DNase I (20 U) for 15 min at room temperature and 15 min at 37 °C. Then, 50 l anti-Flag resin (Sigma) or anti-mouse IgG resin (Santa Cruz) was added, followed by a 12 h incubation at 4 °C overnight. After the resin was washed four times with lysis buffer, bound proteins were eluted with 100 l SDS gel loading buffer (50 mM Tris-HCl pH 6.8, 2% SDS, 0.1% bromophenol blue, 10% glycerol, 10% 2-mercaptoethanol) and 100 l of the eluates were subjected to western blot analysis with anti-Flag and anti-RAD51 antibodies.

6.5.14 Immunoblot analysis

Protein was extracted from cells collected two days after transfection with the indicated siRNAs using NETN buffer (20 mM Tris-HCl pH 8, 420 mM NaCl, 1 mM EDTA, 0.5% Igepal CA-630, 1 mM DTT, and Roche Protease Inhibitor Cocktail) and 8 freeze–thaw cycles. Blots (20–50 g total protein) were probed with the following antibodies: BARD1 (Bethyl, A300-263A; Santa Cruz Biotech, Sc11438), BRCA1 (Abcam, ab16780), 53BP1 (Abcam, ab36823), Flag M2-HRP (Sigma, A8592), Phospho RPA32 S4/S8 (Bethyl, A300-245A), BRCA2 (EMD Millipore, OP95-100UG), RAD51 (Santa Cruz Biotech, sc-8349), Actin (Abcam, ab3280), Tubulin (Santa Cruz Biotech, sc-53030), HA.11 (16B12) (Covance, MMS-101P), or GST-HRP (NEB, E2624S) according to the instructions provided by the manufacturers. If needed, the blots were incubated with HRP-conjugated secondary antibodies (Pierce 31450 for rabbit anti-mouse IgG-HRP; Sigma A6154 for goat anti-rabbit IgG-HRP; Santa Cruz Biotech Sc-

2032 for goat anti-rat IgG-HRP) before visualization of protein signals using the ECL kit (Thermo Scientific Pierce).

6.5.15 DR-GFP reporter assay

The DR-U2OS cell line containing a single integrated copy of the DR-GFP reporter was used [217, 456]. Exponentially growing cells were seeded in 6-well plates at 2×105 cells per well before transfection with 2 l siRNA (20 M) and 5 l Lipofectamine 2000 (Invitrogen). One day after siRNA transfection, cells were transfected with 2 g I-SceI expression vector (pCBASce) and 5 l Lipofectamine 2000. Homologous recombination proficiency was determined by counting the fraction of GFP-positive cells using a BD FACScalibur S 72 h after I-SceI transfection. The results were derived from between three and five transfections of at least three independent experiments.

6.5.16 CRISPR–Cas9-induced gene targeting assay

The assay was conducted as described [458]. U2OS cells were seeded in 6-well plates at 2×105 cells per well before transfection with 2 l siRNA (20 M) and 5 l Lipofectamine 2000 (Invitrogen). One day after siRNA transfection, cells were co-transfected with 1.6 g sgRNA plasmid pX330-LMNA1 (from G. Dellaire) and 0.4 g donor template pCR2.1-CloverLamin (from G. Dellaire) and 5 l Lipofectamine 2000 (Invitrogen). Gene targeting efficiency was determined by counting the per cent of Clover-positive cells using a BD FACScalibur S 72 h after plasmid co-transfection. The results were derived from between three and five transfections of at least three independent experiments.

6.5.17 Immunofluorescence microscopy and image analysis

HeLa cells in exponential growth were transfected on two consecutive days in Opti-MEM medium using RNAiMAX (Invitrogen) with 20 nM BARD1 or control siRNA, as recommended by the manufacturer. Exposure to -rays was performed using a 137Cs -irradiator (J.L. Shepherd, model 81-14) and a dose rate of 1.05 Gy min−1. Immunohistochemistry was performed as previously described [461], except that cells were fixed in 4% paraformaldehyde at room temperature for 10 min and permeabilized in 0.4% Triton X-100 in PBS for 5 min. Rabbit anti-RAD51 (H-92; Santa Cruz Biotechnology; 1:2,000) and goat anti-rabbit AlexaFluor-488 secondary antibodies (Invitrogen; 1:750) were used. For image capture of RAD51 foci, Z-stack section images consisting of 20 stacks (0.2- m intervals) from 100 to 150 nuclei per sample were taken using a 63× oil objective and a Zeiss Axio-Imager.Z2 microscope equipped with Zen Blue software (Carl Zeiss Microscopy). For computational analyses of foci, Z-stacks were collapsed down to the maximum intensity projections, and a combination of ImageJ (http://rsb.info.nih.gov/ij/) and Cell Profiler (http://www.cellprofiler.org/) software programs was used with the following custom program settings for image processing: minimum object size = 3; maximum object size = 35; despeckle ratio = 0.3; rolling ball size = 5. A custom-built pipeline for automated cell (80–300 pixel units) and foci counting with settings for shape (0.5) and dimensions (5 pixels diameter) was employed. The threshold for foci detection was determined based on sham-irradiated samples, and nuclei with >5 foci per nucleus were counted positive. Group allocation and outcome assessment were done in a fully blinded manner.

6.5.18 Clonogenic survival assay

HeLa cells were transiently transfected with siRNA as described above. After 48 h, cells were seeded into 6-well plates at 50–32,000 cells per well, and treated with 0, 5, 10 and 20 nM MMC (Sigma) or 0, 0.5, 1 and 2 M Olaparib (Selleckchem) in regular growth medium for 14 days. Cells were fixed with 10% methanol and 10% acetic acid, and stained with 1% crystal violet in methanol before colonies were counted. Clonogenic survival was determined for a given concentration of cells that were plated by dividing the number of colonies on each treated plate by the number of colonies on the untreated plate, taking the plating efficiency of untreated cells into account.

6.5.19 Preparation of cytoplasmic and nuclear extracts

The Dignam method for the preparation of cytoplasmic and nuclear extracts was followed [462]. In brief, 109 cells were washed with PBS and Dignam buffer A (10 mM Hepes pH 7.9, 1.5 mM MgCl2, 10 mM KCl, 0.5 mM DTT and 0.5 mM PMSF), collected by centrifugation, and lysed in two packed cell volumes of Dignam buffer A using a Dounce homogenizer (50 strokes) with the type A pestle. After centrifugation, the supernatant containing cytoplasmic proteins was saved for analysis. The pelleted nuclei were resuspended and lysed in 3 ml Dignam buffer C (20 mM Hepes pH 7.9, 1.5 mM MgCl2, 420 mM NaCl, 0.2 mM EDTA, 0.5 mM DTT, 25% glycerol, and 0.5 mM PMSF) using a Dounce homogenizer (80 strokes) with the type B pestle. Debris was removed by centrifugation to yield the nuclear extract fraction. The cytoplasmic and nuclear fractions, 20 g each, were analysed by immunoblotting for their content of BRCA1, Flag–SBP–BARD1, tubulin and histone H3.

6.5.20 Statistics and reproducibility

The statistical analysis was performed using Prism 7 (GraphPad Software, Inc.; http://www.graphpad.com/quickcalcs/ttest1.cfm) on the data from at least three independent experiments, as specified. Unless stated otherwise, statistical significance was assessed by two-tailed unpaired Student's t-test. *P 0.05 and **P 0.01 were considered significant.

Discussion and Future Perspectives

A mixture of curiosity and logical rigor has allowed humanity to ask and answer questions about how collections of colliding molecules bond, break, and interact in unfathomably complex ways to coalesce into the miraculous life forms that exists on this planet today. This relentless pursuit of knowledge has led to the branching of many fields including molecular genetics, biochemistry, structural biology, and single-molecule biophysics. Over the course of this work, we have seen the power of combining these various scientific techniques to answer a question of why organisms evolved seemingly redundant proteins. We arrived at this question because evolution is a brutally efficient force that rarely leaves room for the unnecessary consumption of limited resources. These concluding remarks will seek to demonstrate how this work has begun to answer the question why eukaryotes evolved two recombinases for mitotic and meiotic recombination and on the future work that can bring us further answers.

Summary

In the first chapter, we explored the novel technique high-throughput single-molecule DNA curtains developed in the Eric Greene Lab which has allowed us to answer questions in the field of homologous recombination that have otherwise been unsolved. The basic premise

of single-molecule biophysics is that the macroscale metabolic and catabolic pathways that govern the lifecycle of an organism are made up of countless individual molecules that move and interact in seemingly infinite ways but in total, result in the workings of the cell. By only looking at the summation of the molecules, we miss the finer details that can lie in the subpopulations of molecular behaviors. Exploring these subpopulations can provide greater mechanistic understanding in how these pathways work.

Although chapter one features an analogy to explain this concept of subpopulations, I often provide individuals trying understand the power of single-molecule biophysics with the everyday event of traffic. On a typical Friday afternoon, millions of commuters leave work within a range of a couple hours, producing the phenomena of rush hour. This unstable equilibrium of moderate to heavy congestion generally allows cars to move a decent pace, however, small perturbations, such as a long weekend motivating a subpopulation of people looking to get away can quickly lead to complete standstills on that same road and in that same time window. A subpopulation of savvy commuters, knowing that a long weekend is approaching might try to leave work earlier and thus expanding the rush hour window all simply by the influence of another subpopulation. Thus, we will use single-molecule research to explore the inner workings and subpopulations of behaviors that govern homologous recombination towards our better mechanistic understanding of the pathway.

In chapter two, we use single-molecule DNA curtains to understand shared principles in how recombinases across life sample partially homologous DNA molecules towards understanding how recombinases are so efficiently able to search the entire genome for their desired target. A set of universal principles emerged showing that all members of the RecA/Rad51 family of recombinases required at least 8 nt of homology in order to stably bind dsDNA

and that increases in binding stability occurred in 3 nt steps. Since it was previously known that recombinase bind to 3-nt, this 8 nt benchmark hints at how recombinases might communicate with adjacent recombinase molecules to create a cooperative energetic barrier to sampling every possible homologous triplet in the genome. By setting this lower bound to 8 nt, the number of possible substrates across the genome for a presynaptic filament to search greatly reduces this search. The 3 nt stepping points to the way in which recombinases bind nucleotide triplets in a very unique way. Each triplet is held in a B-form-like DNA structure with large stretches in between, creating an asymmetric or recombinase-stretched (RS)-DNA. The exact reason why it might be advantageous for recombinases to bind DNA in this manner for the homology search and eventual homologous pairing is still a topic of exploration. However, by using this triplet stepping behavior, we were able to show the first significant biophysical or biochemical difference between Dmc1 and Rad51. Specifically, Dmc1 appears to be more tolerant of mismatches in internal triplets along the captured dsDNA, whereas RecA and Rad51 could not. This could potentially explaining by eukaryotes evolved two recombinases. Since, Dmc1's role during meiotic recombination is to form crossover events with heterologous homologs, it might beneficial for Dmc1 to be tolerant of mismatches.

In chapter three, we test the limits of Dmc1's tolerance to dsDNA imperfections towards understanding how Dmc1 mediates this novel behavior. We found that although Dmc1 was intolerant of insertions in triplets, it was still able to tolerate multiple mismatches within a triplet as long as there were flanking non-mismatched triplets on each side. In addition, Dmc1 was able to stabilize DNA with abasic sites suggesting a potential role in interacting with the phosphate backbone. However, if this is the case, it is most likely not through non-

bridging oxygen atoms on the DNA backbone since it was still able to stabilize a mismatched triplet with a methylphosphonate backbone.

In chapter four, we combine structure alignments with principles of structure/function to find lineage-specific amino acids in Dmc1 loops, L1 and L2, that differ from those Rad51's loops and due to their three-dimensional proximity to the presynaptic and captured DNA, potentially mediate Dmc1's mismatch stabilization behavior. By using basic cloning, purification and bulk biochemistry techniques, we then made chimeras and tested their basic function to show that most chimeras retained the DNA-dependent ATPase and strand exchange activities that define recombinases. We then used the same single-molecule technique discussed in chapters 1 through 3 to show that, in fact, we can create a Dmc1 mutant with Rad51-like amino acids in L1 that can no longer tolerate mismatches and a Rad51 mutant with Dmc1-like amino acids in L1 that is able to tolerate mismatches. We pair down these mutations to determine the minimal amount of amino acids necessary for producing this effect and found that three in particular are most critical. We then venture into the world of *in vivo* yeast genetics to see how this mismatch tolerance could affect mitotic recombination by introducing mismatches in the homology target of MAT-type switching, the Z box, and see whether mutant Rad51's with Dmc1-like amino acids in L1 MAT-type switch with greater efficiency than WT. Although the full loop swap mutants were somewhat deficient both *in vitro* and *in vivo*, the Rad51 M301Q point mutant proved to more efficient than WT once the Z box had at least mismatches every 1 in 5 bps. This might not seem biologically relevant since homologs will certainly not diverge by as much as 80%. In addition, one of the reasons we needed to introduce so many mutations is that we found that Rad51 is inherently very tolerant of mismatches in vivo and work from other labs have found the same [378]. Thus, in

order to capture any discernible difference between mutant Rad51 and WT, we needed WT begin to struggle with the mutated templates we were providing. Again, from a bulk average perspective, this subtle difference between mutant and WT seems irrelevant. But what we argue, is that this subtle difference we've now seen *in vitro* and *in vivo* might represent a huge effect over the course of generations. Small differences in a single 3-hour repair process could snowball into massive differences when talking on evolutionary scales. For example, the primordial yeast with only primordial Rad51 might have been able to sporulate with 91.4% efficiency while another yeast had an extra Rad51 copy in its genome with a single mutation in L1 that allowed 91.7% sporulation efficiency. Over the course of millions of generations, that has huge implications in evolutionary fitness.

Finally, in chapter 5, we find that the protein expressed by tumor suppressor gene, BRCA1, when in combination with BARD1 is capable of aiding RAD51 in dsDNA capture without affecting the stability of the captured DNA on the filament. This research helps us to draw the link between those with BRCA1 mutations and why they might be more susceptible to cancer. This is an excellent example of how basic science research can aid in our understanding of disease, while simultaneously, the ardent pursuit in understanding this particular gene was because of the disease is produces. The fact that BRCA1 mutations specifically lead to cancer in the third-decade of light, shows how slight defects in a proteins Rad51 mediating activity could lead to RAD51 repairing DNA in a slightly less efficient manner, leading to more effects on genomic integrity.

Future Work

There are many avenues for this research moving forward. One potential avenue is using the method we demonstrated worked for L1 and use bioinformatics, primary structure alignment, and known tertiary and quaternary structure information to explore other lineage-specific amino acids that might be responsible for differences in how Rad51 and Dmc1 interact with mediators and anti-recombinases. For example, there may be particular amino acids or groups of amino acids that make Dmc1 resistance to binding Hed1 or Srs2 stripping Dmc1 from the presynaptic filaments. Using the addition of the *in vivo* assay developed, this information can go from the microscope to the agar plate towards understanding the impact such mutations might have on HR and cell function in general.

In terms of my primary findings with L1 and L2, I had findings that showed Dmc1 had significantly higher affinity for labeled dsDNA than Rad51 and L2 appeared to be responsible for this behavior. One issue is effect was not conserved between yeast and humans, however, one could explore why that is the case or try to pair down the amino acids that are most critical for this effect as done with L1 and mismatch stabilization. Since the Rad51 with a Dmc1-like L2 was extremely deficient *in vivo*, finding the minimal mutations necessary could allow for understanding how greater dsDNA capture by Dmc1 might be important for meiotic recombinase. Additionally, the analogous L1 point mutations could be made in Dmc1 and see whether mismatch intolerance persists once those three amino acids are mutated. Additionally, in vivo yeast genetics can be done with Dmc1, looking at meiotic recombination as well as in vivo work in mice and C. elegans, looking at the effect of these mutations in Rad51 and Dmc1 in multicellular organisms with longer lifespans.

Finally, the recent intersection between DSBR and neurology is a nascent field and quite fascinating. Looking at the relationship between DNA tensions as a regulator for transcription could be an excellent question for DNA curtains to answer. Overall, the of the main themes of this is work is that by combining, techniques across fields and using all available clues from both medicine and the lab bench, significant progress can be made in understanding the world and the organisms that occupy it.

Bibliography

[1] A. Mehta and J. E. Haber. "Sources of DNA double-strand breaks and models of recombinational DNA repair." In: *Cold Spring Harb Perspect Biol* 6.9 (2014), a016428.

[2] D. C. van Gent, J. H. Hoeijmakers, and R. Kanaar. "Chromosomal stability and the DNA double-stranded break connection." In: *Nat Rev Genet* 2.3 (2001), pp. 196–206.

[3] M. M. Vilenchik and A. G. Knudson. "Endogenous DNA double-strand breaks: production, fidelity of repair, and induction of cancer." In: *Proc Natl Acad Sci U S A* 100.22 (2003), pp. 12871–6.

[4] E. Sonoda, C. Morrison, Y. M. Yamashita, M. Takata, and S. Takeda. "Reverse genetic studies of homologous DNA recombination using the chicken B-lymphocyte line, DT40." In: *Philos Trans R Soc Lond B Biol Sci* 356.1405 (2001), pp. 111–7.

[5] A. H. Syeda, M. Hawkins, and P. McGlynn. "Recombination and replication." In: *Cold Spring Harb Perspect Biol* 6.11 (2014), a016550.

[6] R. Madabhushi, F. Gao, A. R. Pfenning, L. Pan, S. Yamakawa, J. Seo, R. Rueda, T. X. Phan, H. Yamakawa, P. C. Pao, R. T. Stott, E. Gjoneska, A. Nott, S. Cho, M. Kellis, and L. H. Tsai. "Activity-Induced DNA Breaks Govern the Expression of Neuronal Early-Response Genes." In: *Cell* 161.7 (2015), pp. 1592–605.

[7] B. J. Lamarche, N. I. Orazio, and M. D. Weitzman. "The MRN complex in double-strand break repair and telomere maintenance." In: *FEBS Lett* 584.17 (2010), pp. 3682–95.

[8] M. Lisby and R. Rothstein. "Choreography of recombination proteins during the DNA damage response." In: *DNA Repair (Amst)* 8.9 (2009), pp. 1068–76.

[9] J. M. Daley, P. L. Palmbos, D. Wu, and T. E. Wilson. "Nonhomologous end joining in yeast." In: *Annu Rev Genet* 39 (2005), pp. 431–51.

[10] K. Lobachev, E. Vitriol, J. Stemple, M. A. Resnick, and K. Bloom. "Chromosome fragmentation after induction of a double-strand break is an active process prevented by the RMX repair complex." In: *Curr Biol* 14.23 (2004), pp. 2107–12.

[11] E. Gobbini, C. Cassani, M. Villa, D. Bonetti, and M. P. Longhese. "Functions and regulation of the MRX complex at DNA double-strand breaks." In: *Microb Cell* 3.8 (2016), pp. 329–337.

[12] K. Fukunaga, Y. Kwon, P. Sung, and K. Sugimoto. "Activation of protein kinase Tel1 through recognition of protein-bound DNA ends." In: *Mol Cell Biol* 31.10 (2011), pp. 1959–71.

[13] A. Javaheri, R. Wysocki, O. Jobin-Robitaille, M. Altaf, J. Cote, and S. J. Kron. "Yeast G1 DNA damage checkpoint regulation by H2A phosphorylation is independent of chromatin remodeling." In: *Proc Natl Acad Sci U S A* 103.37 (2006), pp. 13771–6.

[14] A. Decottignies. "Alternative end-joining mechanisms: a historical perspective." In: *Front Genet* 4 (2013), p. 48.

[15] L. S. Symington and J. Gautier. "Double-strand break end resection and repair pathway choice." In: *Annu Rev Genet* 45 (2011), pp. 247–71.

[16] A. J. Davis and D. J. Chen. "DNA double strand break repair via non-homologous end-joining." In: *Transl Cancer Res* 2.3 (2013), pp. 130–143.

[17] Z. Dudasova, A. Dudas, and M. Chovanec. "Non-homologous end-joining factors of Saccharomyces cerevisiae." In: *FEMS Microbiol Rev* 28.5 (2004), pp. 581–601.

[18] D. Wu, L. M. Topper, and T. E. Wilson. "Recruitment and dissociation of nonhomologous end joining proteins at a DNA double-strand break in Saccharomyces cerevisiae." In: *Genetics* 178.3 (2008), pp. 1237–49.

[19] Y. Zhang, M. L. Hefferin, L. Chen, E. Y. Shim, H. M. Tseng, Y. Kwon, P. Sung, S. E. Lee, and A. E. Tomkinson. "Role of Dnl4-Lif1 in nonhomologous end-joining repair complex assembly and suppression of homologous recombination." In: *Nat Struct Mol Biol* 14.7 (2007), pp. 639–46.

[20] K. K. Chiruvella, B. M. Renard, S. R. Birkeland, S. Sunder, Z. Liang, and T. E. Wilson. "Yeast DNA ligase IV mutations reveal a nonhomologous end joining function of BRCT1 distinct from XRCC4/Lif1 binding." In: *DNA Repair (Amst)* 24 (2014), pp. 37–45.

[21] S. H. Teo and S. P. Jackson. "Lif1p targets the DNA ligase Lig4p to sites of DNA double-strand breaks." In: *Curr Biol* 10.3 (2000), pp. 165–8.

[22] R. A. Deshpande and T. E. Wilson. "Modes of interaction among yeast Nej1, Lif1 and Dnl4 proteins and comparison to human XLF, XRCC4 and Lig4." In: *DNA Repair (Amst)* 6.10 (2007), pp. 1507–16.

[23] H. Yang, Y. Matsumoto, K. M. Trujillo, S. P. Lees-Miller, M. A. Osley, and A. E. Tomkinson. "Role of the yeast DNA repair protein Nej1 in end processing during the repair of DNA double strand breaks by non-homologous end joining." In: *DNA Repair (Amst)* 31 (2015), pp. 1–10.

[24] X. Chen and A. E. Tomkinson. "Yeast Nej1 is a key participant in the initial end binding and final ligation steps of nonhomologous end joining." In: *J Biol Chem* 286.6 (2011), pp. 4931–40.

[25] M. R. Lieber, J. Gu, H. Lu, N. Shimazaki, and A. G. Tsai. "Nonhomologous DNA end joining (NHEJ) and chromosomal translocations in humans." In: *Subcell Biochem* 50 (2010), pp. 279–96.

[26] C. A. Waters, N. T. Strande, D. W. Wyatt, J. M. Pryor, and D. A. Ramsden. "Non-homologous end joining: a good solution for bad ends." In: *DNA Repair (Amst)* 17 (2014), pp. 39–51.

[27] B. O. Krogh and L. S. Symington. "Recombination proteins in yeast." In: *Annu Rev Genet* 38 (2004), pp. 233–71.

[28] J. San Filippo, P. Sung, and H. Klein. "Mechanism of eukaryotic homologous recombination." In: *Annu Rev Biochem* 77 (2008), pp. 229–57.

[29] Jason Sylvan James E. Haber Sang Eun Lee Frédéric Pâques. "Role of yeast SIR genes and mating type in directing DNA double-strand breaks to homologous and non-homologous repair paths." In: *Curr Biol* 9.14 (1999), pp. 767–770.

[30] Z. Zhu, W. H. Chung, E. Y. Shim, S. E. Lee, and G. Ira. "Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends." In: *Cell* 134.6 (2008), pp. 981–94.

[31] K. J. Lee, J. Saha, J. Sun, K. R. Fattah, S. C. Wang, B. Jakob, L. Chi, S. Y. Wang, G. Taucher-Scholz, A. J. Davis, and D. J. Chen. "Phosphorylation of Ku dictates DNA double-strand break (DSB) repair pathway choice in S phase." In: *Nucleic Acids Res* 44.4 (2016), pp. 1732–45.

[32] P. Huertas, F. Cortes-Ledesma, A. A. Sartori, A. Aguilera, and S. P. Jackson. "CDK targets Sae2 to control DNA-end resection and homologous recombination." In: *Nature* 455.7213 (2008), pp. 689–92.

[33] E. P. Mimitou and L. S. Symington. "Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing." In: *Nature* 455.7214 (2008), pp. 770–4.

[34] E. P. Mimitou and L. S. Symington. "Nucleases and helicases take center stage in homologous recombination." In: *Trends Biochem Sci* 34.5 (2009), pp. 264–72.

[35] H. Chen, M. Lisby, and L. S. Symington. "RPA coordinates DNA end resection and prevents formation of DNA hairpins." In: *Mol Cell* 50.4 (2013), pp. 589–600.

[36] Z. Chen, H. Yang, and N. P. Pavletich. "Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures." In: *Nature* 453.7194 (2008), pp. 489–4.

[37] A. B. Conway, T. W. Lynch, Y. Zhang, G. S. Fortin, C. W. Fung, L. S. Symington, and P. A. Rice. "Crystal structure of a Rad51 filament." In: *Nat Struct Mol Biol* 11.8 (2004), pp. 791–6.

[38] B. Gibb, L. F. Ye, Y. Kwon, H. Niu, P. Sung, and E. C. Greene. "Protein dynamics during presynaptic-complex assembly on individual single-stranded DNA molecules." In: *Nat Struct Mol Biol* 21.10 (2014), pp. 893–900.

[39] J. Xu, L. Zhao, Y. Xu, W. Zhao, P. Sung, and H. W. Wang. "Cryo-EM structures of human RAD51 recombinase filaments during catalysis of DNA-strand exchange." In: *Nat Struct Mol Biol* 24.1 (2017), pp. 40–46.

[40] M. McVey, V. Y. Khodaverdian, D. Meyer, P. G. Cerqueira, and W. D. Heyer. "Eukaryotic DNA Polymerases in Homologous Recombination." In: *Annu Rev Genet* 50 (2016), pp. 393–421.

[41] T. Miura, Y. Yamana, T. Usui, H. I. Ogawa, M. T. Yamamoto, and K. Kusano. "Homologous recombination via synthesis-dependent strand annealing in yeast requires the Irc20 and Srs2 DNA helicases." In: *Genetics* 191.1 (2012), pp. 65–78.

[42] J. Matos and S. C. West. "Holliday junction resolution: regulation in space and time." In: *DNA Repair (Amst)* 19 (2014), pp. 176–81.

[43] R. Prakash, D. Satory, E. Dray, A. Papusha, J. Scheller, W. Kramer, L. Krejci, H. Klein, J. E. Haber, P. Sung, and G. Ira. "Yeast Mph1 helicase dissociates Rad51-made D-loops: implications for crossover control in mitotic recombination." In: *Genes Dev* 23.1 (2009), pp. 67–79.

[44] K. Zakharyevich, S. Tang, Y. Ma, and N. Hunter. "Delineation of joint molecule resolution pathways in meiosis identifies a crossover-specific resolvase." In: *Cell* 149.2 (2012), pp. 334–47.

[45] J. Matos, M. G. Blanco, S. Maslen, J. M. Skehel, and S. C. West. "Regulatory control of the resolution of DNA recombination intermediates during meiosis and mitosis." In: *Cell* 147.1 (2011), pp. 158–72.

[46] S. Sarbajna, D. Davies, and S. C. West. "Roles of SLX1-SLX4, MUS81-EME1, and GEN1 in avoiding genome instability and mitotic catastrophe." In: *Genes Dev* 28.10 (2014), pp. 1124–36.

[47] S. L. Andersen and J. Sekelsky. "Meiotic versus mitotic recombination: two different routes for double-strand break repair: the different functions of meiotic versus mitotic DSB repair are reflected in different pathway usage and different outcomes." In: *Bioessays* 32.12 (2010), pp. 1058–66.

[48] K. P. Kohl and J. Sekelsky. "Meiotic and mitotic recombination in meiosis." In: *Genetics* 194.2 (2013), pp. 327–34.

[49] R. C. Von Borstel, K. T. Cain, and C. M. Steinberg. "Inheritance of spontaneous mutability in yeast." In: *Genetics* 69.1 (1971), pp. 17–27.

[50] R. C. Vonborstel, D. E. Graham, K. J. Labrot, and M. A. Resnick. "Mutator Activity of a X-Radiation-Sensitive Yeast." In: *Genetics* 60.1p2 (1968), pp. 233–+.

[51] J. C. Game and R. K. Mortimer. "A genetic study of x-ray sensitive mutants in yeast." In: *Mutat Res* 24.3 (1974), pp. 281–92.

[52] R. H. McKee and C. W. Lawrence. "Genetic analysis of gamma-ray mutagenesis in yeast. I. Reversion in radiation-sensitive strains." In: *Genetics* 93.2 (1979), pp. 361–73.

[53] R. H. McKee and C. W. Lawrence. "Genetic analysis of gamma-ray mutagenesis in yeast. II. Allele-specific control of mutagenesis." In: *Genetics* 93.2 (1979), pp. 375–81.

[54] K. S. Ho. "The gene dosage effect of the rad52 mutation on X-ray survival curves of tetraploid yeast strains." In: *Mutat Res* 33.2-3 (1975), pp. 165–72.

[55] M. A. Resnick and P. Martin. "The repair of double-strand breaks in the nuclear DNA of Saccharomyces cerevisiae and its genetic control." In: *Mol Gen Genet* 143.2 (1976), pp. 119–29.

[56] L. S. Symington. "Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair." In: *Microbiol Mol Biol Rev* 66.4 (2002), 630–70, table of contents.

[57] L. S. Symington. "Mechanism and regulation of DNA end resection in eukaryotes." In: *Crit Rev Biochem Mol Biol* 51.3 (2016), pp. 195–212.

[58] C. R. Wobbe, L. Weissbach, J. A. Borowiec, F. B. Dean, Y. Murakami, P. Bullock, and J. Hurwitz. "Replication of simian virus 40 origin-containing DNA in vitro with purified proteins." In: *Proc Natl Acad Sci U S A* 84.7 (1987), pp. 1834–8.

[59] M. S. Wold and T. Kelly. "Purification and characterization of replication protein A, a cellular protein required for in vitro replication of simian virus 40 DNA." In: *Proc Natl Acad Sci U S A* 85.8 (1988), pp. 2523–7.

[60] S. J. Brill and B. Stillman. "Yeast replication factor-A functions in the unwinding of the SV40 origin of DNA replication." In: *Nature* 342.6245 (1989), pp. 92–5.

[61] T. Shibata, C. DasGupta, R. P. Cunningham, and C. M. Radding. "Purified Escherichia coli recA protein catalyzes homologous pairing of superhelical DNA and single-stranded fragments." In: *Proc Natl Acad Sci U S A* 76.4 (1979), pp. 1638–42.

[62] R. Kolodner, D. H. Evans, and P. T. Morrison. "Purification and characterization of an activity from Saccharomyces cerevisiae that catalyzes homologous pairing and strand exchange." In: *Proc Natl Acad Sci U S A* 84.16 (1987), pp. 5560–4.

[63] W. D. Heyer and R. D. Kolodner. "Purification and characterization of a protein from Saccharomyces cerevisiae that binds tightly to single-stranded DNA and stimulates a cognate strand exchange protein." In: *Biochemistry* 28.7 (1989), pp. 2856–62.

[64] E. Alani, R. Thresher, J. D. Griffith, and R. D. Kolodner. "Characterization of DNA-binding and strand-exchange stimulation properties of y-RPA, a yeast single-strand-DNA-binding protein." In: *J Mol Biol* 227.1 (1992), pp. 54–71.

[65] L. F. Erdile, W. D. Heyer, R. Kolodner, and T. J. Kelly. "Characterization of a cDNA encoding the 70-kDa single-stranded DNA-binding subunit of human replication protein A and the role of the protein in DNA replication." In: *J Biol Chem* 266.18 (1991), pp. 12090–8.

[66] W. D. Heyer, M. R. Rao, L. F. Erdile, T. J. Kelly, and R. D. Kolodner. "An essential Saccharomyces cerevisiae single-stranded DNA binding protein is homologous to the large subunit of human RP-A." In: *EMBO J* 9.7 (1990), pp. 2321–9.

[67] X. Wang and J. E. Haber. "Role of Saccharomyces single-stranded DNA-binding protein RPA in the strand invasion step of double-strand break repair." In: *PLoS Biol* 2.1 (2004), E21.

[68] N. Kantake, T. Sugiyama, R. D. Kolodner, and S. C. Kowalczykowski. "The recombination-deficient mutant RPA (rfa1-t11) is displaced slowly from single-stranded DNA by Rad51 protein." In: *J Biol Chem* 278.26 (2003), pp. 23410–7.

[69] K. Umezu, N. Sugawara, C. Chen, J. E. Haber, and R. D. Kolodner. "Genetic analysis of yeast RPA1 reveals its multiple functions in DNA metabolism." In: *Genetics* 148.3 (1998), pp. 989–1005.

[70] B. Gibb, L. F. Ye, S. C. Gergoudis, Y. Kwon, H. Niu, P. Sung, and E. C. Greene. "Concentration-dependent exchange of replication protein A on single-stranded DNA revealed by single-molecule imaging." In: *PLoS One* 9.2 (2014), e87922.

[71] C. Kim, R. O. Snyder, and M. S. Wold. "Binding properties of replication protein A from human and yeast cells." In: *Mol Cell Biol* 12.7 (1992), pp. 3050–9.

[72] C. J. Ma, B. Gibb, Y. Kwon, P. Sung, and E. C. Greene. "Protein dynamics of human RPA and RAD51 on ssDNA during assembly and disassembly of the RAD51 filament." In: *Nucleic Acids Res* 45.2 (2017), pp. 749–761.

[73] M. S. Wold. "Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism." In: *Annu Rev Biochem* 66 (1997), pp. 61–92.

[74] E. Cannavo, P. Cejka, and S. C. Kowalczykowski. "Relationship of DNA degradation by Saccharomyces cerevisiae exonuclease 1 and its stimulation by RPA and Mre11-Rad50-Xrs2 to DNA end resection." In: *Proc Natl Acad Sci U S A* 110.18 (2013), E1661–8.

[75] P. Ruff, R. A. Donnianni, E. Glancy, J. Oh, and L. S. Symington. "RPA Stabilization of Single-Stranded DNA Is Critical for Break-Induced Replication." In: *Cell Rep* 17.12 (2016), pp. 3359–3368.

[76] A. Georgaki and U. Hubscher. "DNA unwinding by replication protein A is a property of the 70 kDa subunit and is facilitated by phosphorylation of the 32 kDa subunit." In: *Nucleic Acids Res* 21.16 (1993), pp. 3659–65.

[77] A. Georgaki, B. Strack, V. Podust, and U. Hubscher. "DNA unwinding activity of replication protein A." In: *FEBS Lett* 308.3 (1992), pp. 240–4.

[78] K. Treuner, U. Ramsperger, and R. Knippers. "Replication protein A induces the unwinding of long double-stranded DNA regions." In: *J Mol Biol* 259.1 (1996), pp. 104–12.

[79] L. Safa, N. M. Gueddouda, F. Thiebaut, E. Delagoutte, I. Petruseva, O. Lavrik, O. Mendoza, A. Bourdoncle, P. Alberti, J. F. Riou, and C. Saintome. "5' to 3' Unfolding Directionality of DNA Secondary Structures by Replication Protein A: G-QUADRUPLEXES AND DUPLEXES." In: *J Biol Chem* 291.40 (2016), pp. 21246–21256.

[80] T. R. Salas, I. Petruseva, O. Lavrik, A. Bourdoncle, J. L. Mergny, A. Favre, and C. Saintome. "Human replication protein A unfolds telomeric G-quadruplexes." In: *Nucleic Acids Res* 34.17 (2006), pp. 4857–65.

[81] B. Gibb, T. D. Silverstein, I. J. Finkelstein, and E. C. Greene. "Single-stranded DNA curtains for real-time single-molecule visualization of protein-nucleic acid interactions." In: *Anal Chem* 84.18 (2012), pp. 7607–12.

[82] R. Chen, S. Subramanyam, A. H. Elcock, M. Spies, and M. S. Wold. "Dynamic binding of replication protein a is required for DNA repair." In: *Nucleic Acids Res* 44.12 (2016), pp. 5758–72.

[83] C. A. Brosey, C. Yan, S. E. Tsutakawa, W. T. Heller, R. P. Rambo, J. A. Tainer, I. Ivanov, and W. J. Chazin. "A new structural framework for integrating replication protein A into DNA processing machinery." In: *Nucleic Acids Res* 41.4 (2013), pp. 2313–27.

[84] S. Kumaran, A. G. Kozlov, and T. M. Lohman. "Saccharomyces cerevisiae replication protein A binds to single-stranded DNA in multiple salt-dependent modes." In: *Biochemistry* 45.39 (2006), pp. 11958–73.

[85] C. A. Brosey, S. E. Soss, S. Brooks, C. Yan, I. Ivanov, K. Dorai, and W. J. Chazin. "Functional dynamics in replication protein A DNA binding and protein recruitment domains." In: *Structure* 23.6 (2015), pp. 1028–38.

[86] P. Sung. "Function of yeast Rad52 protein as a mediator between replication protein A and the Rad51 recombinase." In: *J Biol Chem* 272.45 (1997), pp. 28194–7.

[87] L. Krejci, V. Altmannova, M. Spirek, and X. Zhao. "Homologous recombination and its regulation." In: *Nucleic Acids Res* 40.13 (2012), pp. 5795–818.

[88] K. Adzuma, T. Ogawa, and H. Ogawa. "Primary structure of the RAD52 gene in Saccharomyces cerevisiae." In: *Mol Cell Biol* 4.12 (1984), pp. 2735–44.

[89] D. Schild, B. Konforti, C. Perez, W. Gish, and R. Mortimer. "Isolation and characterization of yeast DNA repair genes : I. Cloning of the RAD52 gene." In: *Curr Genet* 7.2 (1983), pp. 85–92.

[90] G. T. Milne and D. T. Weaver. "Dominant negative alleles of RAD52 reveal a DNA repair/recombination complex including Rad51 and Rad52." In: *Genes Dev* 7.9 (1993), pp. 1755–65.

[91] A. Shinohara, H. Ogawa, and T. Ogawa. "Rad51 protein involved in repair and recombination in S. cerevisiae is a RecA-like protein." In: *Cell* 69.3 (1992), pp. 457–70.

[92] U. H. Mortensen, C. Bendixen, I. Sunjevaric, and R. Rothstein. "DNA strand annealing is promoted by the yeast Rad52 protein." In: *Proc Natl Acad Sci U S A* 93.20 (1996), pp. 10729–34.

[93] H. Zou and R. Rothstein. "Holliday junctions accumulate in replication mutants via a RecA homolog-independent mechanism." In: *Cell* 90.1 (1997), pp. 87–96.

[94] A. V. Nimonkar, R. A. Sica, and S. C. Kowalczykowski. "Rad52 promotes second-end DNA capture in double-stranded break repair to form complement-stabilized joint molecules." In: *Proc Natl Acad Sci U S A* 106.9 (2009), pp. 3077–82.

[95] A. Shinohara and T. Ogawa. "Stimulation by Rad52 of yeast Rad51-mediated recombination." In: *Nature* 391.6665 (1998), pp. 404–7.

[96] S. L. Hays, A. A. Firmenich, P. Massey, R. Banerjee, and P. Berg. "Studies of the interaction between Rad52 protein and the yeast single-stranded DNA binding protein RPA." In: *Mol Cell Biol* 18.7 (1998), pp. 4400–6.

[97] T. Sugiyama and S. C. Kowalczykowski. "Rad52 protein associates with replication protein A (RPA)-single-stranded DNA to accelerate Rad51-mediated displacement of RPA and presynaptic complex formation." In: *J Biol Chem* 277.35 (2002), pp. 31663–72.

[98] A. Shinohara, M. Shinohara, T. Ohta, S. Matsuda, and T. Ogawa. "Rad52 forms ring structures and co-operates with RPA in single-strand DNA annealing." In: *Genes Cells* 3.3 (1998), pp. 145–56.

[99] I. Shi, S. C. Hallwyl, C. Seong, U. Mortensen, R. Rothstein, and P. Sung. "Role of the Rad52 amino-terminal DNA binding activity in DNA strand capture in homologous recombination." In: *J Biol Chem* 284.48 (2009), pp. 33275–84.

[100] K. S. Ho and R. K. Mortimer. "Two mutations which confer temperature-sensitive radiation sensitivity in the yeast Saccharomyces cerevisiae." In: *Mutat Res* 33.2-3 (1975), pp. 157–64.

[101] S. T. Lovett and R. K. Mortimer. "Characterization of null mutants of the RAD55 gene of Saccharomyces cerevisiae: effects of temperature, osmotic strength and mating type." In: *Genetics* 116.4 (1987), pp. 547–53.

[102] R. D. Johnson and L. S. Symington. "Functional differences and interactions among the putative RecA homologs Rad51, Rad55, and Rad57." In: *Mol Cell Biol* 15.9 (1995), pp. 4843–50.

[103] M. Lisby, J. H. Barlow, R. C. Burgess, and R. Rothstein. "Choreography of the DNA damage response: spatiotemporal relationships among checkpoint and repair proteins." In: *Cell* 118.6 (2004), pp. 699–713.

[104] N. Sugawara, X. Wang, and J. E. Haber. "In vivo roles of Rad52, Rad54, and Rad55 proteins in Rad51-mediated recombination." In: *Mol Cell* 12.1 (2003), pp. 209–19.

[105] P. Sung. "Yeast Rad55 and Rad57 proteins form a heterodimer that functions with replication protein A to promote DNA strand exchange by Rad51 recombinase." In: *Genes Dev* 11.9 (1997), pp. 1111–21.

[106] G. S. Fortin and L. S. Symington. "Mutations in yeast Rad51 that partially bypass the requirement for Rad55 and Rad57 in DNA repair by increasing the stability of Rad51-DNA complexes." In: *EMBO J* 21.12 (2002), pp. 3160–70.

[107] J. Liu, L. Renault, X. Veaute, F. Fabre, H. Stahlberg, and W. D. Heyer. "Rad51 paralogues Rad55-Rad57 balance the antirecombinase Srs2 in Rad51 filament formation." In: *Nature* 479.7372 (2011), pp. 245–8.

[108] M. Budd and R. K. Mortimer. "Repair of Double-Strand Breaks in a Temperature Conditional Radiation-Sensitive Mutant of Saccharomyces Cerevisiae." In: *Mutation Research* 103.1 (1982), pp. 19–24.

[109] G. M. Cole, D. Schild, S. T. Lovett, and R. K. Mortimer. "Regulation of RAD54- and RAD52-lacZ gene fusions in Saccharomyces cerevisiae in response to DNA damage." In: *Mol Cell Biol* 7.3 (1987), pp. 1078–84.

[110] B. Clever, H. Interthal, J. Schmuckli-Maurer, J. King, M. Sigrist, and W. D. Heyer. "Recombinational repair in yeast: functional interactions between Rad51 and Rad54 proteins." In: *EMBO J* 16.9 (1997), pp. 2535–44.

[111] H. Jiang, Y. Xie, P. Houston, K. Stemke-Hale, U. H. Mortensen, R. Rothstein, and T. Kodadek. "Direct association between the yeast Rad51 and Rad54 recombination proteins." In: *J Biol Chem* 271.52 (1996), pp. 33181–6.

[112] F. Palladino and H. L. Klein. "Analysis of mitotic and meiotic defects in Saccharomyces cerevisiae SRS2 DNA helicase mutants." In: *Genetics* 132.1 (1992), pp. 23–37.

[113] G. Petukhova, S. Stratton, and P. Sung. "Catalysis of homologous DNA pairing by yeast Rad51 and Rad54 proteins." In: *Nature* 393.6680 (1998), pp. 91–4.

[114] J. A. Eisen, K. S. Sweder, and P. C. Hanawalt. "Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions." In: *Nucleic Acids Res* 23.14 (1995), pp. 2715–23.

[115] S. J. Ceballos and W. D. Heyer. "Functions of the Snf2/Swi2 family Rad54 motor protein in homologous recombination." In: *Biochim Biophys Acta* 1809.9 (2011), pp. 509–23.

[116] G. Petukhova, S. Van Komen, S. Vergano, H. Klein, and P. Sung. "Yeast Rad54 promotes Rad51-dependent homologous DNA pairing via ATP hydrolysis-driven change in DNA double helix conformation." In: *J Biol Chem* 274.41 (1999), pp. 29453–62.

[117] A. V. Mazin, C. J. Bornarth, J. A. Solinger, W. D. Heyer, and S. C. Kowalczykowski. "Rad54 protein is targeted to pairing loci by the Rad51 nucleoprotein filament." In: *Mol Cell* 6.3 (2000), pp. 583–92.

[118] A. V. Mazin, A. A. Alexeev, and S. C. Kowalczykowski. "A novel function of Rad54 protein. Stabilization of the Rad51 nucleoprotein filament." In: *J Biol Chem* 278.16 (2003), pp. 14029–36.

[119] S. Van Komen, G. Petukhova, S. Sigurdsson, S. Stratton, and P. Sung. "Superhelicity-driven homologous DNA pairing by yeast recombination factors Rad51 and Rad54." In: *Mol Cell* 6.3 (2000), pp. 563–72.

[120] A. Alexeev, A. Mazin, and S. C. Kowalczykowski. "Rad54 protein possesses chromatin-remodeling activity stimulated by the Rad51-ssDNA nucleoprotein filament." In: *Nat Struct Biol* 10.3 (2003), pp. 182–6.

[121] B. Wolner and C. L. Peterson. "ATP-dependent and ATP-independent roles for the Rad54 chromatin remodeling enzyme during recombinational repair of a DNA double strand break." In: *J Biol Chem* 280.11 (2005), pp. 10855–60.

[122] M. Jaskelioff, S. Van Komen, J. E. Krebs, P. Sung, and C. L. Peterson. "Rad54p is a chromatin remodeling enzyme required for heteroduplex DNA joint formation with chromatin." In: *J Biol Chem* 278.11 (2003), pp. 9212–8.

[123] K. Kiianitsa, J. A. Solinger, and W. D. Heyer. "Rad54 protein exerts diverse modes of ATPase activity on duplex DNA partially and fully covered with Rad51 protein." In: *J Biol Chem* 277.48 (2002), pp. 46205–15.

[124] J. A. Solinger, K. Kiianitsa, and W. D. Heyer. "Rad54, a Swi2/Snf2-like recombinational repair protein, disassembles Rad51:dsDNA filaments." In: *Mol Cell* 10.5 (2002), pp. 1175–88.

[125] P. M. Kim, K. S. Paffett, J. A. Solinger, W. D. Heyer, and J. A. Nickoloff. "Spontaneous and double-strand break-induced recombination, and gene conversion tract lengths, are differentially affected by overexpression of wild-type or ATPase-defective yeast Rad54." In: *Nucleic Acids Res* 30.13 (2002), pp. 2727–35.

[126] D. V. Bugreev, O. M. Mazina, and A. V. Mazin. "Rad54 protein promotes branch migration of Holliday junctions." In: *Nature* 442.7102 (2006), pp. 590–3.

[127] I. Amitani, R. J. Baskin, and S. C. Kowalczykowski. "Visualization of Rad54, a chromatin remodeling protein, translocating on single DNA molecules." In: *Mol Cell* 23.1 (2006), pp. 143–8.

[128] J Brooks Crickard, Kyle Kaniecki, YoungHo Kwon, Patrick Sung, Michael Lisby, and Eric C Greene. "Regulation of Hed1 and Rad54 binding during maturation of the meiosis-specific presynaptic complex." In: *The EMBO Journal* 37.7 (2018). eprint: `http://emboj.embopress.org/content/37/7/e98728.full.pdf`.

[129] N. H. Thoma, B. K. Czyzewski, A. A. Alexeev, A. V. Mazin, S. C. Kowalczykowski, and N. P. Pavletich. "Structure of the SWI2/SNF2 chromatin-remodeling domain of eukaryotic Rad54." In: *Nat Struct Mol Biol* 12.4 (2005), pp. 350–6.

[130] M. S. Esposito and R. E. Esposito. "The genetic control of sporulation in Saccharomyces. I. The isolation of temperature-sensitive sporulation-deficient mutants." In: *Genetics* 61.1 (1969), pp. 79–89.

[131] M. S. Esposito and R. E. Esposito. "Mutants of meiosis and ascospore formation." In: *Methods Cell Biol* 11 (1975), pp. 303–26.

[132] M. S. Esposito, R. E. Esposito, M. Arnaud, and H. O. Halvorson. "Conditional mutants of meiosis in yeast." In: *J Bacteriol* 104.1 (1970), pp. 202–10.

[133] R. E. Esposito, N. Frink, P. Bernstein, and M. S. Esposito. "The genetic control of sporulation in Saccharomyces. II. Dominance and complementation of mutants of meiosis and spore formation." In: *Mol Gen Genet* 114.3 (1972), pp. 241–8.

[134] H. Sun, D. Treco, N. P. Schultes, and J. W. Szostak. "Double-strand breaks at an initiation site for meiotic gene conversion." In: *Nature* 338.6210 (1989), pp. 87–90.

[135] L. Cao, E. Alani, and N. Kleckner. "A pathway for generation and processing of double-strand breaks during meiotic recombination in S. cerevisiae." In: *Cell* 61.6 (1990), pp. 1089–101.

[136] L. W. Thorne and B. Byers. "Stage-specific effects of X-irradiation on yeast meiosis." In: *Genetics* 134.1 (1993), pp. 29–42.

[137] A. Bergerat, B. de Massy, D. Gadelle, P. C. Varoutas, A. Nicolas, and P. Forterre. "An atypical topoisomerase II from Archaea with implications for meiotic recombination." In: *Nature* 386.6623 (1997), pp. 414–7.

[138] S. Keeney, C. N. Giroux, and N. Kleckner. "Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family." In: *Cell* 88.3 (1997), pp. 375–84.

[139] S. Keeney and N. Kleckner. "Covalent protein-DNA complexes at the 5' strand termini of meiosis-specific double-strand breaks in yeast." In: *Proc Natl Acad Sci U S A* 92.24 (1995), pp. 11274–8.

[140] S. Keeney. "Spo11 and the Formation of DNA Double-Strand Breaks in Meiosis." In: *Genome Dyn Stab* 2 (2008), pp. 81–123.

[141] M. J. Neale and S. Keeney. "Clarifying the mechanics of DNA strand exchange in meiotic recombination." In: *Nature* 442.7099 (2006), pp. 153–8.

[142] N. Manfrini, I. Guerini, A. Citterio, G. Lucchini, and M. P. Longhese. "Processing of meiotic DNA double strand breaks requires cyclin-dependent kinase and multiple nucleases." In: *J Biol Chem* 285.15 (2010), pp. 11628–37.

[143] M. J. Neale, J. Pan, and S. Keeney. "Endonucleolytic processing of covalent protein-linked DNA double-strand breaks." In: *Nature* 436.7053 (2005), pp. 1053–7.

[144] J. Pan, M. Sasaki, R. Kniewel, H. Murakami, H. G. Blitzblau, S. E. Tischfield, X. Zhu, M. J. Neale, M. Jasin, N. D. Socci, A. Hochwagen, and S. Keeney. "A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation." In: *Cell* 144.5 (2011), pp. 719–31.

[145] A. H. McKee and N. Kleckner. "Mutations in Saccharomyces cerevisiae that block meiotic prophase chromosome metabolism and confer cell cycle arrest at pachytene identify two new meiosis-specific genes SAE1 and SAE3." In: *Genetics* 146.3 (1997), pp. 817–34.

[146] K. P. Rabitsch, A. Toth, M. Galova, A. Schleiffer, G. Schaffner, E. Aigner, C. Rupp, A. M. Penkner, A. C. Moreno-Borchart, M. Primig, R. E. Esposito, F. Klein, M. Knop, and K. Nasmyth. "A screen for genes required for meiosis and spore formation based on whole-genome expression." In: *Curr Biol* 11.13 (2001), pp. 1001–9.

[147] A. Hayase, M. Takagi, T. Miyazaki, H. Oshiumi, M. Shinohara, and A. Shinohara. "A protein complex containing Mei5 and Sae3 promotes the assembly of the meiosis-specific RecA homolog Dmc1." In: *Cell* 119.7 (2004), pp. 927–40.

[148] H. Tsubouchi and G. S. Roeder. "The budding yeast mei5 and sae3 proteins act together with dmc1 during meiotic recombination." In: *Genetics* 168.3 (2004), pp. 1219–30.

[149] S. R. Ferrari, J. Grubb, and D. K. Bishop. "The Mei5-Sae3 protein complex mediates Dmc1 activity in Saccharomyces cerevisiae." In: *J Biol Chem* 284.18 (2009), pp. 11766–70.

[150] A. F. Say, L. L. Ledford, D. Sharma, A. K. Singh, W. K. Leung, H. A. Sehorn, H. Tsubouchi, P. Sung, and M. G. Sehorn. "The budding yeast Mei5-Sae3 complex interacts with Rad51 and preferentially binds a DNA fork structure." In: *DNA Repair (Amst)* 10.6 (2011), pp. 586–94.

[151] V. Cloud, Y. L. Chan, J. Grubb, B. Budke, and D. K. Bishop. "Rad51 is an accessory factor for Dmc1-mediated joint molecule formation during meiosis." In: *Science* 337.6099 (2012), pp. 1222–5.

[152] M. S. Brown and D. K. Bishop. "DNA strand exchange and RecA homologs in meiosis." In: *Cold Spring Harb Perspect Biol* 7.1 (2014), a016659.

[153] H. Tsubouchi and G. S. Roeder. "Budding yeast Hed1 down-regulates the mitotic recombination machinery when meiotic recombination is impaired." In: *Genes Dev* 20.13 (2006), pp. 1766–75.

[154] S. Chu, J. DeRisi, M. Eisen, J. Mulholland, D. Botstein, P. O. Brown, and I. Herskowitz. "The transcriptional program of sporulation in budding yeast." In: *Science* 282.5389 (1998), pp. 699–705.

[155] J. P. Lao, V. Cloud, C. C. Huang, J. Grubb, D. Thacker, C. Y. Lee, M. E. Dresser, N. Hunter, and D. K. Bishop. "Meiotic crossover control by concerted action of Rad51-Dmc1 in homolog template bias and robust homeostatic regulation." In: *PLoS Genet* 9.12 (2013), e1003978.

[156] V. Busygina, M. G. Sehorn, I. Y. Shi, H. Tsubouchi, G. S. Roeder, and P. Sung. "Hed1 regulates Rad51-mediated recombination via a novel mechanism." In: *Genes Dev* 22.6 (2008), pp. 786–95.

[157] J. Y. Leu, P. R. Chua, and G. S. Roeder. "The meiosis-specific Hop2 protein of S. cerevisiae ensures synapsis between homologous chromosomes." In: *Cell* 94.3 (1998), pp. 375–86.

[158] C. Zierhut, M. Berlinger, C. Rupp, A. Shinohara, and F. Klein. "Mnd1 is required for meiotic interhomolog repair." In: *Curr Biol* 14.9 (2004), pp. 752–62.

[159] H. Tsubouchi and G. S. Roeder. "The Mnd1 protein forms a complex with hop2 to promote homologous chromosome pairing and meiotic double-strand break repair." In: *Mol Cell Biol* 22.9 (2002), pp. 3078–88.

[160] Y. K. Chen, C. H. Leng, H. Olivares, M. H. Lee, Y. C. Chang, W. M. Kung, S. C. Ti, Y. H. Lo, A. H. Wang, C. S. Chang, D. K. Bishop, Y. P. Hsueh, and T. F. Wang. "Heterodimeric complexes of Hop2 and Mnd1 function with Dmc1 to promote meiotic homolog juxtaposition and strand assimilation." In: *Proc Natl Acad Sci U S A* 101.29 (2004), pp. 10572–7.

[161] R. J. Pezza, O. N. Voloshin, F. Vanevski, and R. D. Camerini-Otero. "Hop2/Mnd1 acts on two critical steps in Dmc1-promoted homologous pairing." In: *Genes Dev* 21.14 (2007), pp. 1758–66.

[162] H. A. Kang, H. C. Shin, A. S. Kalantzi, C. P. Toseland, H. M. Kim, S. Gruber, M. D. Peraro, and B. H. Oh. "Crystal structure of Hop2-Mnd1 and mechanistic insights into its role in meiotic recombination." In: *Nucleic Acids Res* 43.7 (2015), pp. 3841–56.

[163] C. S. Eichinger and S. Jentsch. "Synaptonemal complex formation and meiotic checkpoint signaling are linked to the lateral element protein Red1." In: *Proc Natl Acad Sci U S A* 107.25 (2010), pp. 11370–5.

[164] N. Fasten. "Spermatogenesis of the American crayfish, Cambarus virilis and Cambarus immunis (?), with special reference to synapsis and the chromatoid bodies." In: *Journal of Morphology* 25.4 (1914), pp. 587–649.

[165] M. J. Moses. "Chromosomal structures in crayfish spermatocytes." In: *J Biophys Biochem Cytol* 2.2 (1956), pp. 215–8.

[166] D. W. Fawcett. "The fine structure of chromosomes in the meiotic prophase of vertebrate spermatocytes." In: *J Biophys Biochem Cytol* 2.4 (1956), pp. 403–6.

[167] F. M. Engels and A. F. Croes. "The synaptinemal complex in yeast." In: *Chromosoma* 25.1 (1968), pp. 104–6.

[168] C. B. Gillies. "Synaptonemal complex and chromosome structure." In: *Annu Rev Genet* 9 (1975), pp. 91–109.

[169] 2013.

[170] M. Sym, J. A. Engebrecht, and G. S. Roeder. "ZIP1 is a synaptonemal complex protein required for meiotic chromosome synapsis." In: *Cell* 72.3 (1993), pp. 365–78.

[171] S. L. Page and R. S. Hawley. "The genetics and molecular biology of the synaptonemal complex." In: *Annu Rev Cell Dev Biol* 20 (2004), pp. 525–58.

[172] J. Loidl, F. Klein, and H. Scherthan. "Homologous pairing is reduced but not abolished in asynaptic mutants of yeast." In: *J Cell Biol* 125.6 (1994), pp. 1191–200.

[173] N. M. Hollingsworth and B. Byers. "HOP1: a yeast meiotic pairing gene." In: *Genetics* 121.3 (1989), pp. 445–62.

[174] B. Rockmill and G. S. Roeder. "RED1: a yeast gene required for the segregation of chromosomes during the reductional division of meiosis." In: *Proc Natl Acad Sci U S A* 85.16 (1988), pp. 6057–61.

[175] B. Rockmill and G. S. Roeder. "Meiosis in asynaptic yeast." In: *Genetics* 126.3 (1990), pp. 563–74.

[176] A. Schwacha and N. Kleckner. "Interhomolog bias during meiotic recombination: meiotic functions promote a highly differentiated interhomolog-only pathway." In: *Cell* 90.6 (1997), pp. 1123–35.

[177] A. V. Smith and G. S. Roeder. "The yeast Red1 protein localizes to the cores of meiotic chromosomes." In: *J Cell Biol* 136.5 (1997), pp. 957–67.

[178] D. Woltering, B. Baumgartner, S. Bagchi, B. Larkin, J. Loidl, T. de los Santos, and N. M. Hollingsworth. "Meiotic segregation, synapsis, and recombination checkpoint functions require physical interaction between the chromosomal proteins Red1p and Hop1p." In: *Mol Cell Biol* 20.18 (2000), pp. 6646–58.

[179] N. Humphryes and A. Hochwagen. "A non-sister act: recombination template choice during meiosis." In: *Exp Cell Res* 329.1 (2014), pp. 53–60.

[180] M. S. Brown, J. Grubb, A. Zhang, M. J. Rust, and D. K. Bishop. "Small Rad51 and Dmc1 Complexes Often Co-occupy Both Ends of a Meiotic DNA Double Strand Break." In: *PLoS Genet* 11.12 (2015), e1005653.

[181] J. B. Crickard, K. Kaniecki, Y. Kwon, P. Sung, and E. C. Greene. "Spontaneous self-segregation of Rad51 and Dmc1 DNA recombinases within mixed recombinase filaments." In: *J Biol Chem* (2018).

[182] A. J. Clark and A. D. Margulies. "Isolation and Characterization of Recombination-Deficient Mutants of Escherichia Coli K12." In: *Proc Natl Acad Sci U S A* 53 (1965), pp. 451–9.

[183] E. M. Witkin. "The mutability toward ultraviolet light of recombination-deficient strains of Escherichia coli." In: *Mutat Res* 8.1 (1969), pp. 9–14.

[184] D. K. Bishop, D. Park, L. Xu, and N. Kleckner. "DMC1: a meiosis-specific yeast homolog of E. coli recA required for recombination, synaptonemal complex formation, and cell cycle progression." In: *Cell* 69.3 (1992), pp. 439–56.

[185] P. Sung. "Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast RAD51 protein." In: *Science* 265.5176 (1994), pp. 1241–3.

[186] E. L. Hong, A. Shinohara, and D. K. Bishop. "Saccharomyces cerevisiae Dmc1 protein promotes renaturation of single-strand DNA (ssDNA) and assimilation of ssDNA into homologous super-coiled duplex DNA." In: *J Biol Chem* 276.45 (2001), pp. 41906–12.

[187] M. H. Lee, Y. C. Chang, E. L. Hong, J. Grubb, C. S. Chang, D. K. Bishop, and T. F. Wang. "Calcium ion promotes yeast Dmc1 activity via formation of long and fine helical filaments with single-stranded DNA." In: *J Biol Chem* 280.49 (2005), pp. 40980–4.

[188] L. Du and Y. Luo. "Structure of a filament of stacked octamers of human DMC1 recombinase." In: *Acta Crystallogr Sect F Struct Biol Cryst Commun* 69.Pt 4 (2013), pp. 382–6.

[189] T. Kinebuchi, W. Kagawa, R. Enomoto, K. Tanaka, K. Miyagawa, T. Shibata, H. Kurumizaka, and S. Yokoyama. "Structural basis for octameric ring formation and DNA interaction of the human homologous-pairing protein Dmc1." In: *Mol Cell* 14.3 (2004), pp. 363–74.

[190] J. Y. Masson and S. C. West. "The Rad51 and Dmc1 recombinases: a non-identical twin relationship." In: *Trends Biochem Sci* 26.2 (2001), pp. 131–6.

[191] L. Brocchieri and S. Karlin. "A symmetric-iterated multiple alignment of protein sequences." In: *J Mol Biol* 276.1 (1998), pp. 249–64.

[192] S. D. Sheridan, X. Yu, R. Roth, J. E. Heuser, M. G. Sehorn, P. Sung, E. H. Egelman, and D. K. Bishop. "A comparative analysis of Dmc1 and Rad51 nucleoprotein filaments." In: *Nucleic Acids Res* 36.12 (2008), pp. 4057–66.

[193] R. M. Story, I. T. Weber, and T. A. Steitz. "The structure of the E. coli recA protein monomer and polymer." In: *Nature* 355.6358 (1992), pp. 318–25.

[194] C. E. Bell. "Structure and mechanism of Escherichia coli RecA ATPase." In: *Mol Microbiol* 58.2 (2005), pp. 358–66.

[195] X. Xing and C. E. Bell. "Crystal structures of Escherichia coli RecA in a compressed helical filament." In: *J Mol Biol* 342.5 (2004), pp. 1471–85.

[196] L. T. Chen, T. P. Ko, Y. W. Chang, K. A. Lin, A. H. Wang, and T. F. Wang. "Structural and functional analyses of five conserved positively charged residues in the L1 and N-terminal DNA binding motifs of archaeal RADA protein." In: *PLoS One* 2.9 (2007), e858.

[197] V. A. Malkov and R. D. Camerini-Otero. "Photocross-links between single-stranded DNA and Escherichia coli RecA protein map to loops L1 (amino acid residues 157-164) and L2 (amino acid residues 195-209)." In: *J Biol Chem* 270.50 (1995), pp. 30230–3.

[198] Y. Wang and K. Adzuma. "Differential proximity probing of two DNA binding sites in the Escherichia coli recA protein using photo-cross-linking methods." In: *Biochemistry* 35.11 (1996), pp. 3563–71.

[199] F. Maraboeuf, O. Voloshin, R. D. Camerini-Otero, and M. Takahashi. "The central aromatic residue in loop L2 of RecA interacts with DNA. Quenching of the fluorescence of a tryptophan reporter inserted in L2 upon binding to DNA." In: *J Biol Chem* 270.52 (1995), pp. 30927–32.

[200] T. Selmane, P. Wittung-Stafshede, F. Maraboeuf, O. N. Voloshin, B. Norden, D. R. Camerini-Otero, and M. Takahashi. "The L2 loop peptide of RecA stiffens and restricts base motions of single-stranded DNA similar to the intact protein." In: *FEBS Lett* 446.1 (1999), pp. 30–4.

[201] T. Shinohara, S. Ikawa, W. Iwasaki, T. Hiraki, T. Hikima, T. Mikawa, N. Arai, N. Kamiya, and T. Shibata. "Loop L1 governs the DNA-binding specificity and order for RecA-catalyzed reactions in homologous recombination and DNA repair." In: *Nucleic Acids Res* 43.2 (2015), pp. 973–86.

[202] P. S. Malik and L. S. Symington. "Rad51 gain-of-function mutants that exhibit high affinity DNA binding cause DNA damage sensitivity in the absence of Srs2." In: *Nucleic Acids Res* 36.20 (2008), pp. 6504–10.

[203] Y. Matsuo, I. Sakane, Y. Takizawa, M. Takahashi, and H. Kurumizaka. "Roles of the human Rad51 L1 and L2 loops in DNA binding." In: *FEBS J* 273.14 (2006), pp. 3148–59.

[204] L. A. Bannister, R. J. Pezza, J. R. Donaldson, D. G. de Rooij, K. J. Schimenti, R. D. Camerini-Otero, and J. C. Schimenti. "A dominant, recombination-defective allele of Dmc1 causing male-specific sterility." In: *PLoS Biol* 5.5 (2007), e105.

[205] K. Hortnagel, O. N. Voloshin, H. H. Kinal, N. Ma, C. Schaffer-Judge, and R. D. Camerini-Otero. "Saturation mutagenesis of the E. coli RecA loop L2 homologous DNA pairing region reveals residues essential for recombination and recombinational repair." In: *J Mol Biol* 286.4 (1999), pp. 1097–106.

[206] M. Silva, N. Villanueva, T. Mandigo, X. X. Nguyen, and S. Morrical. "Role of Yeast Rad51 Interface Residue F187A and DNA Binding Loop 1 Residue F290 in Coordinating DNA Binding, ATPase, and DNA Strand Exchange Activities." In: *Faseb Journal* 29 (2015).

[207] L. Krejci, J. Damborsky, B. Thomsen, M. Duno, and C. Bendixen. "Molecular dissection of interactions between Rad51 and members of the recombination-repair group." In: *Mol Cell Biol* 21.3 (2001), pp. 966–76.

[208] X. P. Zhang, V. E. Galkin, X. Yu, E. H. Egelman, and W. D. Heyer. "Loop 2 in Saccharomyces cerevisiae Rad51 protein regulates filament formation and ATPase activity." In: *Nucleic Acids Res* 37.1 (2009), pp. 158–71.

[209] D. E. Rao and Y. Luo. "pH-dependent activities and structural stability of loop-2-anchoring helix of RadA recombinase from Methanococcus voltae." In: *Protein Pept Lett* 21.7 (2014), pp. 679–87.

[210] T. K. Prasad, C. C. Yeykal, and E. C. Greene. "Visualizing the assembly of human Rad51 filaments on double-stranded DNA." In: *J Mol Biol* 363.3 (2006), pp. 713–28.

[211] A. Fiser, R. K. Do, and A. Sali. "Modeling of loops in protein structures." In: *Protein Sci* 9.9 (2000), pp. 1753–73.

[212] A. Reymer, K. Frykholm, K. Morimatsu, M. Takahashi, and B. Norden. "Structure of human Rad51 protein filament from molecular modeling and site-specific linear dichroism spectroscopy." In: *Proceedings of the National Academy of Sciences of the United States of America* 106.32 (2009), pp. 13248–13253.

[213] C. Simmerling, B. Strockbine, and A. E. Roitberg. "All-atom structure prediction and folding simulations of a stable protein." In: *J Am Chem Soc* 124.38 (2002), pp. 11258–9.

[214] O. Bougie and J. I. Weberpals. "Clinical Considerations of BRCA1- and BRCA2-Mutation Carriers: A Review." In: *Int J Surg Oncol* 2011 (2011), p. 374012.

[215] B. B. Roa, A. A. Boyd, K. Volcik, and C. S. Richards. "Ashkenazi Jewish population frequencies for common mutations in BRCA1 and BRCA2." In: *Nat Genet* 14.2 (1996), pp. 185–7.

[216] R. Prakash, Y. Zhang, W. Feng, and M. Jasin. "Homologous recombination and human health: the roles of BRCA1, BRCA2, and associated proteins." In: *Cold Spring Harb Perspect Biol* 7.4 (2015), a016600.

[217] B. Xia, Q. Sheng, K. Nakanishi, A. Ohashi, J. Wu, N. Christ, X. Liu, M. Jasin, F. J. Couch, and D. M. Livingston. "Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2." In: *Mol Cell* 22.6 (2006), pp. 719–29.

[218] W. Zhao, J. B. Steinfeld, F. Liang, X. Chen, D. G. Maranon, C. Jian Ma, Y. Kwon, T. Rao, W. Wang, C. Sheng, X. Song, Y. Deng, J. Jimenez-Sainz, L. Lu, R. B. Jensen, Y. Xiong, G. M. Kupfer, C. Wiese, E. C. Greene, and P. Sung. "BRCA1-BARD1 promotes RAD51-mediated homologous DNA pairing." In: *Nature* 550.7676 (2017), pp. 360–365.

[219] A. Carreira and S. C. Kowalczykowski. "Two classes of BRC repeats in BRCA2 promote RAD51 nucleoprotein filament function by distinct mechanisms." In: *Proc Natl Acad Sci U S A* 108.26 (2011), pp. 10448–53.

[220] R. T. Neff, L. Senter, and R. Salani. "BRCA mutation in ovarian cancer: testing, implications and treatment considerations." In: *Ther Adv Med Oncol* 9.8 (2017), pp. 519–531.

[221] C. Della Pepa, G. Tonini, C. Pisano, M. Di Napoli, S. C. Cecere, R. Tambaro, G. Facchini, and S. Pignata. "Ovarian cancer standard of care: are there real alternatives?" In: *Chin J Cancer* 34.1 (2015), pp. 17–27.

[222] I. Godet and D. M. Gilkes. "BRCA1 and BRCA2 mutations and treatment strategies for breast cancer." In: *Integr Cancer Sci Ther* 4.1 (2017).

[223] Y. Hu. "BRCA1, hormone, and tissue-specific tumor suppression." In: *Int J Biol Sci* 5.1 (2009), pp. 20–7.

[224] M. Widschwendter, A. N. Rosenthal, S. Philpott, I. Rizzuto, L. Fraser, J. Hayward, M. P. Intermaggio, C. K. Edlund, S. J. Ramus, S. A. Gayther, L. Dubeau, E. O. Fourkala, A. Zaikin, U. Menon, and I. J. Jacobs. "The sex hormone system in carriers of BRCA1/2 mutations: a case-control study." In: *Lancet Oncol* 14.12 (2013), pp. 1226–32.

[225] E. Y. Lee and S. Abbondante. "Tissue-specific tumor suppression by BRCA1." In: *Proc Natl Acad Sci U S A* 111.12 (2014), pp. 4353–4.

[226] T. J. Broering, K. G. Alavattam, R. I. Sadreyev, Y. Ichijima, Y. Kato, K. Hasegawa, R. D. Camerini-Otero, J. T. Lee, P. R. Andreassen, and S. H. Namekawa. "BRCA1 establishes DNA damage signaling and pericentric heterochromatin of the X chromosome in male meiosis." In: *J Cell Biol* 205.5 (2014), pp. 663–75.

[227] J. S. Martinez, C. von Nicolai, T. Kim, A. Ehlen, A. V. Mazin, S. C. Kowalczykowski, and A. Carreira. "BRCA2 regulates DMC1-mediated recombination through the BRC repeats." In: *Proc Natl Acad Sci U S A* 113.13 (2016), pp. 3515–20.

[228] K. R. Smith, H. A. Hanson, G. P. Mineau, and S. S. Buys. "Effects of BRCA1 and BRCA2 mutations on female fertility." In: *Proc Biol Sci* 279.1732 (2012), pp. 1389–95.

[229] O. Uziel, R. Yerushalmi, L. Zuriano, S. Naser, E. Beery, J. Nordenberg, I. Lubin, Y. Adel, D. Shepshelovich, H. Yavin, I. Ben Aharon, S. Pery, S. Rizel, M. Pasmanik-Chor, D. Frumkin, and M. Lahav. "BRCA1/2 mutations perturb telomere biology: characterization of structural and functional abnormalities in vitro and in vivo." In: *Oncotarget* 7.3 (2016), pp. 2433–54.

[230] T. Tsuzuki, Y. Fujii, K. Sakumi, Y. Tominaga, K. Nakao, M. Sekiguchi, A. Matsushiro, Y. Yoshimura, and MoritaT. "Targeted disruption of the Rad51 gene leads to lethality in embryonic mice." In: *Proceedings of the National Academy of Sciences* 93.13 (1996), pp. 6236–6240.

[231] M. Kato, K. Yano, F. Matsuo, H. Saito, T. Katagiri, H. Kurumizaka, M. Yoshimoto, F. Kasumi, F. Akiyama, G. Sakamoto, H. Nagawa, Y. Nakamura, and Y. Miki. "Identification of Rad51 alteration in patients with bilateral breast cancer." In: *J Hum Genet* 45.3 (2000), pp. 133–7.

[232] T. Ishida, Y. Takizawa, I. Sakane, and H. Kurumizaka. "Altered DNA binding by the human Rad51-R150Q mutant found in breast cancer patients." In: *Biol Pharm Bull* 30.8 (2007), pp. 1374–8.

[233] J. Chen, M. D. Morrical, K. A. Donigan, J. B. Weidhaas, J. B. Sweasy, A. M. Averill, J. A. Tomczak, and S. W. Morrical. "Tumor-associated mutations in a conserved structural motif alter physical and biochemical properties of human RAD51 recombinase." In: *Nucleic Acids Res* 43.2 (2015), pp. 1098–111.

[234] M. J. Smith and R. Rothstein. "Poetry in motion: Increased chromosomal mobility after DNA damage." In: *DNA Repair (Amst)* 56 (2017), pp. 102–108.

[235] L. Hasselbach, S. Haase, D. Fischer, H. C. Kolberg, and H. W. Sturzbecher. "Characterisation of the promoter region of the human DNA-repair gene Rad51." In: *Eur J Gynaecol Oncol* 26.6 (2005), pp. 589–98.

[236] D. Sekhar, S. Pooja, S. Kumar, and S. Rajender. "RAD51 135G>C substitution increases breast cancer risk in an ethnic-specific manner: a meta-analysis on 21,236 cases and 19,407 controls." In: *Sci Rep* 5 (2015), p. 11588.

[237] M. C. Silva, M. D. Morrical, K. E. Bryan, A. M. Averill, J. Dragon, J. P. Bond, and S. W. Morrical. "RAD51 variant proteins from human lung and kidney tumors exhibit DNA strand exchange defects." In: *DNA Repair (Amst)* 42 (2016), pp. 44–55.

[238] N. Ameziane, P. May, A. Haitjema, H. J. van de Vrugt, S. E. van Rossum-Fikkert, D. Ristic, G. J. Williams, J. Balk, D. Rockx, H. Li, M. A. Rooimans, A. B. Oostra, E. Velleuer, R. Dietrich, O. B. Bleijerveld, A. F. Maarten Altelaar, H. Meijers-Heijboer, H. Joenje, G. Glusman, J. Roach, L. Hood, D. Galas, C. Wyman, R. Balling, J. den Dunnen, J. P. de Winter, R. Kanaar, R. Gelinas, and J. C. Dorsman. "A novel Fanconi anaemia subtype associated with a dominant-negative mutation in RAD51." In: *Nat Commun* 6 (2015), p. 8829.

[239] B. P. Alter, P. S. Rosenberg, and L. C. Brody. "Clinical and molecular features associated with biallelic mutations in FANCD1/BRCA2." In: *J Med Genet* 44.1 (2007), pp. 1–9.

[240] H. Dong, D. W. Nebert, E. A. Bruford, D. C. Thompson, H. Joenje, and V. Vasiliou. "Update of the human and mouse Fanconi anemia genes." In: *Hum Genomics* 9 (2015), p. 32.

[241] F. Vaz, H. Hanenberg, B. Schuster, K. Barker, C. Wiek, V. Erven, K. Neveling, D. Endt, I. Kesterton, F. Autore, F. Fraternali, M. Freund, L. Hartmann, D. Grimwade, R. G. Roberts, H. Schaal, S. Mohammed, N. Rahman, D. Schindler, and C. G. Mathew. "Mutation of the RAD51C gene in a Fanconi anemia-like disorder." In: *Nat Genet* 42.5 (2010), pp. 406–9.

[242] Y. Liu, J. Y. Masson, R. Shah, P. O'Regan, and S. C. West. "RAD51C is required for Holliday junction processing in mammalian cells." In: *Science* 303.5655 (2004), pp. 243–6.

[243] S. Sigurdsson, S. Van Komen, W. Bussen, D. Schild, J. S. Albala, and P. Sung. "Mediator function of the human Rad51B-Rad51C complex in Rad51/RPA-catalyzed DNA strand exchange." In: *Genes Dev* 15.24 (2001), pp. 3308–18.

[244] J. Chun, E. S. Buechelmaier, and S. N. Powell. "Rad51 paralog complexes BCDX2 and CX3 act at different stages in the BRCA1-BRCA2-dependent homologous recombination pathway." In: *Mol Cell Biol* 33.2 (2013), pp. 387–95.

[245] H. Kurumizaka, S. Ikawa, M. Nakada, R. Enomoto, W. Kagawa, T. Kinebuchi, M. Yamazoe, S. Yokoyama, and T. Shibata. "Homologous pairing and ring and filament structure formation activities of the human Xrcc2*Rad51D complex." In: *J Biol Chem* 277.16 (2002), pp. 14315–20.

[246] H. E. Shamseldin, M. Elfaki, and F. S. Alkuraya. "Exome sequencing reveals a novel Fanconi group defined by XRCC2 mutation." In: *J Med Genet* 49.3 (2012), pp. 184–6.

[247] Ondrej Beláň Jarmila Mlčoušková Mário Špírek Vincenzo Costanzo Lumír Krejčí Karina Zadorozhny Vincenzo Sannino. "Fanconi-Anemia-Associated Mutations Destabilize RAD51 Filaments and Impair Replication Fork Protection." In: *Cell Rep* 21.2 (2017), pp. 333–340.

[248] D. L. Croteau, V. Popuri, P. L. Opresko, and V. A. Bohr. "Human RecQ helicases in DNA repair, recombination, and replication." In: *Annu Rev Biochem* 83 (2014), pp. 519–52.

[249] T. Wechsler, S. Newman, and S. C. West. "Aberrant chromosome morphology in human cells defective for Holliday junction resolution." In: *Nature* 471.7340 (2011), pp. 642–6.

[250] O. Bischof, S. H. Kim, J. Irving, S. Beresten, N. A. Ellis, and J. Campisi. "Regulation and localization of the Bloom syndrome protein in response to DNA damage." In: *J Cell Biol* 153.2 (2001), pp. 367–80.

[251] S. Sakamoto, K. Nishikawa, S. J. Heo, M. Goto, Y. Furuichi, and A. Shimamoto. "Werner helicase relocates into nuclear foci in response to DNA damaging agents and co-localizes with RPA and Rad51." In: *Genes Cells* 6.5 (2001), pp. 421–30.

[252] M. Goto, Y. Ishikawa, M. Sugimoto, and Y. Furuichi. "Werner syndrome: a changing pattern of clinical manifestations in Japan (1917 2008)." In: *Biosci Trends* 7.1 (2013), pp. 13–22.

[253] J. M. Lauper, A. Krause, T. L. Vaughan, and Jr. Monnat R. J. "Spectrum and risk of neoplasia in Werner syndrome: a systematic review." In: *PLoS One* 8.4 (2013), e59709.

[254] S. Kaiser, F. Sauer, and C. Kisker. "The structural and functional characterization of human RecQ4 reveals insights into its helicase mechanism." In: *Nat Commun* 8 (2017), p. 15907.

[255] L. L. Wang and S. E. Plon. "Rothmund-Thomson Syndrome." In: *GeneReviews((R))*. Ed. by M. P. Adam, H. H. Ardinger, R. A. Pagon, S. E. Wallace, L. J. H. Bean, K. Stephens, and A. Amemiya. Seattle (WA), 1993.

[256] L. Larizza, G. Roversi, and L. Volpi. "Rothmund-Thomson syndrome." In: *Orphanet J Rare Dis* 5 (2010), p. 2.

[257] A. R. Gennery. "Primary immunodeficiency syndromes associated with defective DNA double-strand break repair." In: *Br Med Bull* 77-78 (2006), pp. 71–85.

[258] Charles Janeway. *Immunobiology 5 the immune system in health and disease*. 5th. Garland Pub., 2001.

[259] 2001.

[260] A. Kwan, R. S. Abraham, R. Currier, A. Brower, K. Andruszewski, J. K. Abbott, M. Baker, M. Ballow, L. E. Bartoshesky, F. A. Bonilla, C. Brokopp, E. Brooks, M. Caggana, J. Celestin, J. A. Church, A. M. Comeau, J. A. Connelly, M. J. Cowan, C. Cunningham-Rundles, T. Dasu, N. Dave, M. T. De La Morena, U. Duffner, C. T. Fong, L. Forbes, D. Freedenberg, E. W. Gelfand, J. E. Hale, I. C. Hanson, B. N. Hay, D. Hu, A. Infante, D. Johnson, N. Kapoor, D. M. Kay, D. B. Kohn, R. Lee, H. Lehman, Z. Lin, F. Lorey, A. Abdel-Mageed, A. Manning, S. McGhee, T. B. Moore, S. J. Naides, L. D. Notarangelo, J. S. Orange, S. Y. Pai, M. Porteus, R. Rodriguez, N. Romberg, J. Routes, M. Ruehle, A. Rubenstein, C. A. Saavedra-Matiz, G. Scott, P. M. Scott, E. Secord, C. Seroogy, W. T. Shearer, S. Siegel, S. K. Silvers, E. R.

Stiehm, R. W. Sugerman, J. L. Sullivan, S. Tanksley, M. L. th Tierce, J. Verbsky, B. Vogel, R. Walker, K. Walkovich, J. E. Walter, R. L. Wasserman, M. S. Watson, G. A. Weinberg, L. B. Weiner, H. Wood, A. B. Yates, J. M. Puck, and V. R. Bonagura. "Newborn screening for severe combined immunodeficiency in 11 screening programs in the United States." In: *JAMA* 312.7 (2014), pp. 729–38.

[261] 2017.

[262] N. Kutukculer, N. Gulez, N. E. Karaca, G. Aksu, and A. Berdeli. "Novel mutations and diverse clinical phenotypes in recombinase-activating gene 1 deficiency." In: *Ital J Pediatr* 38 (2012), p. 8.

[263] C. C. Dvorak and M. J. Cowan. "Radiosensitive severe combined immunodeficiency disease." In: *Immunol Allergy Clin North Am* 30.1 (2010), pp. 125–42.

[264] Y. Ma, U. Pannicke, K. Schwarz, and M. R. Lieber. "Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination." In: *Cell* 108.6 (2002), pp. 781–94.

[265] M. van der Burg, H. Ijspeert, N. S. Verkaik, T. Turul, W. W. Wiegant, K. Morotomi-Yano, P. O. Mari, I. Tezcan, D. J. Chen, M. Z. Zdzienicka, J. J. van Dongen, and D. C. van Gent. "A DNA-PKcs mutation in a radiosensitive T-B- SCID patient inhibits Artemis activation and nonhomologous end-joining." In: *J Clin Invest* 119.1 (2009), pp. 91–8.

[266] D. Punwani, M. Kawahara, J. Yu, U. Sanford, S. Roy, K. Patel, D. A. Carbonaro, A. D. Karlen, S. Khan, K. Cornetta, M. Rothe, A. Schambach, D. B. Kohn, H. L. Malech, R. S. McIvor, J. M. Puck, and M. J. Cowan. "Lentivirus Mediated Correction of Artemis-Deficient Severe Combined Immunodeficiency." In: *Hum Gene Ther* 28.1 (2017), pp. 112–124.

[267] T. Altmann and A. R. Gennery. "DNA ligase IV syndrome; a review." In: *Orphanet J Rare Dis* 11.1 (2016), p. 137.

[268] P. Ahnesorg, P. Smith, and S. P. Jackson. "XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining." In: *Cell* 124.2 (2006), pp. 301–13.

[269] U. Grawunder, M. Wilm, X. Wu, P. Kulesza, T. E. Wilson, M. Mann, and M. R. Lieber. "Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells." In: *Nature* 388.6641 (1997), pp. 492–5.

[270] D. Cagdas, T. T. Ozgur, G. T. Asal, P. Revy, J. P. De Villartay, M. van der Burg, O. Sanal, and I. Tezcan. "Two SCID cases with Cernunnos-XLF deficiency success-

fully treated by hematopoietic stem cell transplantation." In: *Pediatr Transplant* 16.5 (2012), E167–71.

[271] K. E. Orii, Y. Lee, N. Kondo, and P. J. McKinnon. "Selective utilization of nonhomologous end-joining and homologous recombination DNA repair pathways during nervous system development." In: *Proc Natl Acad Sci U S A* 103.26 (2006), pp. 10017–22.

[272] E. Suberbielle, P. E. Sanchez, A. V. Kravitz, X. Wang, K. Ho, K. Eilertson, N. Devidze, A. C. Kreitzer, and L. Mucke. "Physiologic brain activity causes DNA double-strand breaks in neurons, with exacerbation by amyloid-beta." In: *Nat Neurosci* 16.5 (2013), pp. 613–21.

[273] X. Fernandez, O. Diaz-Ingelmo, B. Martinez-Garcia, and J. Roca. "Chromatin regulates DNA torsional energy via topoisomerase II-mediated relaxation of positive supercoils." In: *EMBO J* 33.13 (2014), pp. 1492–501.

[274] Y. Shiloh and Y. Ziv. "The ATM protein kinase: regulating the cellular response to genotoxic stress, and more." In: *Nat Rev Mol Cell Biol* 14.4 (2013), pp. 197–210.

[275] 2017.

[276] J. H. Hoeijmakers. "Genome maintenance mechanisms for preventing cancer." In: *Nature* 411.6835 (2001), pp. 366–74.

[277] G. Rotman and Y. Shiloh. "ATM: from gene to function." In: *Hum Mol Genet* 7.10 (1998), pp. 1555–63.

[278] A. Marechal and L. Zou. "DNA damage sensing by the ATM and ATR kinases." In: *Cold Spring Harb Perspect Biol* 5.9 (2013).

[279] G. K. Alderton, H. Joenje, R. Varon, A. D. Borglum, P. A. Jeggo, and M. O'Driscoll. "Seckel syndrome exhibits cellular features demonstrating defects in the ATR-signalling pathway." In: *Hum Mol Genet* 13.24 (2004), pp. 3127–38.

[280] A. Verloes, S. Drunat, P. Gressens, and S. Passemard. "Primary Autosomal Recessive Microcephalies and Seckel Syndrome Spectrum Disorders." In: *GeneReviews((R))*. Ed. by M. P. Adam, H. H. Ardinger, R. A. Pagon, S. E. Wallace, L. J. H. Bean, K. Stephens, and A. Amemiya. Seattle (WA), 2013.

[281] K. H. Chrzanowska, H. Gregorek, B. Dembowska-Baginska, M. A. Kalina, and M. Digweed. "Nijmegen breakage syndrome (NBS)." In: *Orphanet J Rare Dis* 7 (2012), p. 13.

[282] J. Kobayashi, A. Antoccia, H. Tauchi, S. Matsuura, and K. Komatsu. "NBS1 and its functional role in the DNA damage response." In: *DNA Repair (Amst)* 3.8-9 (2004), pp. 855–61.

[283] 2017.

[284] R. Varon, C. Vissinga, M. Platzer, K. M. Cerosaletti, K. H. Chrzanowska, K. Saar, G. Beckmann, E. Seemanova, P. R. Cooper, N. J. Nowak, M. Stumm, C. M. Weemaes, R. A. Gatti, R. K. Wilson, M. Digweed, A. Rosenthal, K. Sperling, P. Concannon, and A. Reis. "Nibrin, a novel DNA double-strand break repair protein, is mutated in Nijmegen breakage syndrome." In: *Cell* 93.3 (1998), pp. 467–76.

[285] R. Waltes, R. Kalb, M. Gatei, A. W. Kijas, M. Stumm, A. Sobeck, B. Wieland, R. Varon, Y. Lerenthal, M. F. Lavin, D. Schindler, and T. Dork. "Human RAD50 deficiency in a Nijmegen breakage syndrome-like disorder." In: *Am J Hum Genet* 84.5 (2009), pp. 605–16.

[286] H. Tauchi, J. Kobayashi, K. Morishima, D. C. van Gent, T. Shiraishi, N. S. Verkaik, D. vanHeems, E. Ito, A. Nakamura, E. Sonoda, M. Takata, S. Takeda, S. Matsuura, and K. Komatsu. "Nbs1 is essential for DNA repair by homologous recombination in higher vertebrate cells." In: *Nature* 420.6911 (2002), pp. 93–8.

[287] Y. Saito and K. Komatsu. "Functional Role of NBS1 in Radiation Damage Response and Translesion DNA Synthesis." In: *Biomolecules* 5.3 (2015), pp. 1990–2002.

[288] G. A. Cromie, J. C. Connelly, and D. R. Leach. "Recombination at double-strand breaks and DNA ends: conserved mechanisms from phage to humans." In: *Mol Cell* 8.6 (2001), pp. 1163–74.

[289] W. D. Heyer, K. T. Ehmsen, and J. Liu. "Regulation of homologous recombination in eukaryotes." In: *Annu Rev Genet* 44 (2010), pp. 113–39.

[290] M. Sasaki, J. Lange, and S. Keeney. "Genome destabilization by homologous recombination in the germ line." In: *Nat Rev Mol Cell Biol* 11.3 (2010), pp. 182–95.

[291] J. E. Haber. "Mating-type genes and MAT switching in Saccharomyces cerevisiae." In: *Genetics* 191.1 (2012), pp. 33–64.

[292] F. Paques and J. E. Haber. "Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae." In: *Microbiol Mol Biol Rev* 63.2 (1999), pp. 349–404.

[293] L. S. Symington, R. Rothstein, and M. Lisby. "Mechanisms and regulation of mitotic recombination in Saccharomyces cerevisiae." In: *Genetics* 198.3 (2014), pp. 795–835.

[294] P. R. Bianco, R. B. Tracy, and S. C. Kowalczykowski. "DNA strand exchange proteins: a biochemical and physical comparison." In: *Front Biosci* 3 (1998), pp. D570–603.

[295] S. W. Morrical. "DNA-pairing and annealing processes in homologous recombination and homology-directed repair." In: *Cold Spring Harb Perspect Biol* 7.2 (2015), a016444.

[296] M. Lisby and R. Rothstein. "Cell biology of mitotic recombination." In: *Cold Spring Harb Perspect Biol* 7.3 (2015), a016535.

[297] J. Mine-Hattab and R. Rothstein. "DNA in motion during double-strand break repair." In: *Trends Cell Biol* 23.11 (2013), pp. 529–36.

[298] T. Fazio, M. L. Visnapuu, S. Wind, and E. C. Greene. "DNA curtains and nanoscale curtain rods: high-throughput tools for single molecule imaging." In: *Langmuir* 24.18 (2008), pp. 10524–31.

[299] J. Gorman, T. Fazio, F. Wang, S. Wind, and E. C. Greene. "Nanofabricated racks of aligned and anchored DNA substrates for single-molecule imaging." In: *Langmuir* 26.2 (2010), pp. 1372–9.

[300] J. Gorman, A. J. Plys, M. L. Visnapuu, E. Alani, and E. C. Greene. "Visualizing one-dimensional diffusion of eukaryotic DNA repair factors along a chromatin lattice." In: *Nat Struct Mol Biol* 17.8 (2010), pp. 932–8.

[301] A. Graneli, C. C. Yeykal, T. K. Prasad, and E. C. Greene. "Organized arrays of individual DNA molecules tethered to supported lipid bilayers." In: *Langmuir* 22.1 (2006), pp. 292–9.

[302] D. Duzdevich, M. D. Warner, S. Ticau, N. A. Ivica, S. P. Bell, and E. C. Greene. "The dynamics of eukaryotic replication initiation: origin specificity, licensing, and firing at the single-molecule level." In: *Mol Cell* 58.3 (2015), pp. 483–94.

[303] J. Gorman, F. Wang, S. Redding, A. J. Plys, T. Fazio, S. Wind, E. E. Alani, and E. C. Greene. "Single-molecule imaging reveals target-search mechanisms during DNA mismatch repair." In: *Proc Natl Acad Sci U S A* 109.45 (2012), E3074–83.

[304] J. Y. Lee, I. J. Finkelstein, L. K. Arciszewska, D. J. Sherratt, and E. C. Greene. "Single-molecule imaging of FtsK translocation reveals mechanistic features of protein-protein collisions on DNA." In: *Mol Cell* 54.5 (2014), pp. 832–43.

[305] J. Y. Lee, T. Terakawa, Z. Qi, J. B. Steinfeld, S. Redding, Y. Kwon, W. A. Gaines, W. Zhao, P. Sung, and E. C. Greene. "DNA RECOMBINATION. Base triplet stepping by the Rad51/RecA family of recombinases." In: *Science* 349.6251 (2015), pp. 977–81.

[306] Z. Qi, S. Redding, J. Y. Lee, B. Gibb, Y. Kwon, H. Niu, W. A. Gaines, P. Sung, and E. C. Greene. "DNA sequence alignment by microhomology sampling during homologous recombination." In: *Cell* 160.5 (2015), pp. 856–69.

[307] S. Redding, S. H. Sternberg, M. Marshall, B. Gibb, P. Bhat, C. K. Guegler, B. Wiedenheft, J. A. Doudna, and E. C. Greene. "Surveillance and Processing of Foreign DNA by the Escherichia coli CRISPR-Cas System." In: *Cell* 163.4 (2015), pp. 854–65.

[308] T. D. Silverstein, B. Gibb, and E. C. Greene. "Visualizing protein movement on DNA at the single-molecule level using DNA curtains." In: *DNA Repair (Amst)* 20 (2014), pp. 94–109.

[309] S. H. Sternberg, S. Redding, M. Jinek, E. C. Greene, and J. A. Doudna. "DNA interrogation by the CRISPR RNA-guided endonuclease Cas9." In: *Nature* 507.7490 (2014), pp. 62–7.

[310] F. Wang, S. Redding, I. J. Finkelstein, J. Gorman, D. R. Reichman, and E. C. Greene. "The promoter-search mechanism of Escherichia coli RNA polymerase is dominated by three-dimensional diffusion." In: *Nat Struct Mol Biol* 20.2 (2013), pp. 174–81.

[311] D. Axelrod. "Total internal reflection fluorescence microscopy." In: *Methods Cell Biol* 30 (1989), pp. 245–70.

[312] E. C. Greene, S. Wind, T. Fazio, J. Gorman, and M. L. Visnapuu. "DNA curtains for high-throughput single-molecule optical imaging." In: *Methods Enzymol* 472 (2010), pp. 293–315.

[313] S. Broderick, K. Rehmet, C. Concannon, and H. P. Nasheuer. "Eukaryotic single-stranded DNA binding proteins: central factors in genome stability." In: *Subcell Biochem* 50 (2010), pp. 143–63.

[314] S. K. Deng, B. Gibb, M. J. de Almeida, E. C. Greene, and L. S. Symington. "RPA antagonizes microhomology-mediated repair of DNA double-strand breaks." In: *Nat Struct Mol Biol* 21.4 (2014), pp. 405–12.

[315] T. Ha. "Single-molecule approaches embrace molecular cohorts." In: *Cell* 154.4 (2013), pp. 723–6.

[316] S. Cocco, J. F. Marko, and R. Monasson. "Stochastic ratchet mechanisms for replacement of proteins bound to DNA." In: *Phys Rev Lett* 112.23 (2014), p. 238101.

[317] J. S. Graham, R. C. Johnson, and J. F. Marko. "Concentration-dependent exchange accelerates turnover of proteins bound to double-stranded DNA." In: *Nucleic Acids Res* 39.6 (2011), pp. 2249–59.

[318] N. Hadizadeh, R. C. Johnson, and J. F. Marko. "Facilitated Dissociation of a Nucleoid Protein from the Bacterial Chromosome." In: *J Bacteriol* 198.12 (2016), pp. 1735–42.

[319] C. P. Joshi, D. Panda, D. J. Martell, N. M. Andoy, T. Y. Chen, A. Gaballa, J. D. Helmann, and P. Chen. "Direct substitution and assisted dissociation pathways for turning off transcription by a MerR-family metalloregulator." In: *Proc Natl Acad Sci U S A* 109.38 (2012), pp. 15121–6.

[320] S. Kunzelmann, C. Morris, A. P. Chavda, J. F. Eccleston, and M. R. Webb. "Mechanism of interaction between single-stranded DNA binding protein and DNA." In: *Biochemistry* 49.5 (2010), pp. 843–52.

[321] U. H. Mortensen, M. Lisby, and R. Rothstein. "Rad52." In: *Curr Biol* 19.16 (2009), R676–7.

[322] P. Sung and H. Klein. "Mechanism of homologous recombination: mediators and helicases take on regulatory functions." In: *Nat Rev Mol Cell Biol* 7.10 (2006), pp. 739–50.

[323] J. P. Lao, S. D. Oh, M. Shinohara, A. Shinohara, and N. Hunter. "Rad52 promotes postinvasion steps of meiotic double-strand-break repair." In: *Mol Cell* 29.4 (2008), pp. 517–24.

[324] M. J. McIlwraith and S. C. West. "DNA repair synthesis facilitates RAD52-mediated second-end capture during DSB repair." In: *Mol Cell* 29.4 (2008), pp. 510–6.

[325] A. V. Nimonkar and S. C. Kowalczykowski. "Second-end DNA capture in double-strand break repair: how to catch a DNA by its tail." In: *Cell Cycle* 8.12 (2009), pp. 1816–7.

[326] B. Efron and R. Tibshirani. *An introduction to the bootstrap*. New York: Champman and Hall, Inc., 1993.

[327] M. M. Cox, M. F. Goodman, K. N. Kreuzer, D. J. Sherratt, S. J. Sandler, and K. J. Marians. "The importance of repairing stalled replication forks." In: *Nature* 404.6773 (2000), pp. 37–41.

[328] C. Fraser, W. P. Hanage, and B. G. Spratt. "Recombination and the nature of bacterial speciation." In: *Science* 315.5811 (2007), pp. 476–80.

[329] S. C. Kowalczykowski. "Structural biology: snapshots of DNA repair." In: *Nature* 453.7194 (2008), pp. 463–6.

[330] K. Ragunathan, C. Joo, and T. Ha. "Real-time observation of strand exchange reaction with high spatiotemporal resolution." In: *Structure* 19.8 (2011), pp. 1064–73.

[331] P. Yakovchuk, E. Protozanova, and M. D. Frank-Kamenetskii. "Base-stacking and base-pairing contributions into thermal stability of the DNA double helix." In: *Nucleic Acids Res* 34.2 (2006), pp. 564–74.

[332] E. M. Selva, L. New, G. F. Crouse, and R. S. Lahue. "Mismatch correction acts as a barrier to homeologous recombination in Saccharomyces cerevisiae." In: *Genetics* 139.3 (1995), pp. 1175–88.

[333] W. Chen and S. Jinks-Robertson. "The role of the mismatch repair machinery in regulating mitotic and meiotic recombination between diverged sequences in yeast." In: *Genetics* 151.4 (1999), pp. 1299–313.

[334] E. Evans, N. Sugawara, J. E. Haber, and E. Alani. "The Saccharomyces cerevisiae Msh2 mismatch repair protein localizes to recombination intermediates in vivo." In: *Mol Cell* 5.5 (2000), pp. 789–99.

[335] J. P. Doye, T. E. Ouldridge, A. A. Louis, F. Romano, P. Sulc, C. Matek, B. E. Snodin, L. Rovigatti, J. S. Schreck, R. M. Harrison, and W. P. Smith. "Coarse-graining DNA for simulations of DNA nanotechnology." In: *Phys Chem Chem Phys* 15.47 (2013), pp. 20395–414.

[336] C. Danilowicz, A. Peacock-Villada, J. Vlassakis, A. Facon, E. Feinstein, N. Kleckner, and M. Prentiss. "The differential extension in dsDNA bound to Rad51 filaments may play important roles in homology recognition and strand exchange." In: *Nucleic Acids Res* 42.1 (2014), pp. 526–33.

[337] A. L. Forget and S. C. Kowalczykowski. "Single-molecule imaging of DNA pairing by RecA reveals a three-dimensional homology search." In: *Nature* 482.7385 (2012), pp. 423–7.

[338] C. Joo, S. A. McKinney, M. Nakamura, I. Rasnik, S. Myong, and T. Ha. "Real-time observation of RecA filament dynamics with single monomer resolution." In: *Cell* 126.3 (2006), pp. 515–27.

[339] R. Roy, A. G. Kozlov, T. M. Lohman, and T. Ha. "SSB protein diffusion on single-stranded DNA stimulates RecA filament formation." In: *Nature* 461.7267 (2009), pp. 1092–7.

[340] T. Sugiyama, E. M. Zaitseva, and S. C. Kowalczykowski. "A single-stranded DNA-binding protein is needed for efficient presynaptic complex formation by the Saccharomyces cerevisiae Rad51 protein." In: *J Biol Chem* 272.12 (1997), pp. 7940–5.

[341] D. V. Bugreev and A. V. Mazin. "Ca2+ activates human homologous recombination protein Rad51 by modulating its ATPase activity." In: *Proc Natl Acad Sci U S A* 101.27 (2004), pp. 9988–93.

[342] V. Busygina, W. A. Gaines, Y. Xu, Y. Kwon, G. J. Williams, S. W. Lin, H. Y. Chang, P. Chi, H. W. Wang, and P. Sung. "Functional attributes of the Saccharomyces cerevisiae meiotic recombinase Dmc1." In: *DNA Repair (Amst)* 12.9 (2013), pp. 707–12.

[343] T. E. Ouldridge, A. A. Louis, and J. P. Doye. "Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model." In: *J Chem Phys* 134.8 (2011), p. 085101.

[344] C. Danilowicz, E. Feinstein, A. Conover, V. W. Coljee, J. Vlassakis, Y. L. Chan, D. K. Bishop, and M. Prentiss. "RecA homology search is promoted by mechanical stress along the scanned duplex DNA." In: *Nucleic Acids Res* 40.4 (2012), pp. 1717–27.

[345] I. De Vlaminck, M. T. van Loenhout, L. Zweifel, J. den Blanken, K. Hooning, S. Hage, J. Kerssemakers, and C. Dekker. "Mechanism of homology recognition in DNA recombination from dual-molecule experiments." In: *Mol Cell* 46.5 (2012), pp. 616–24.

[346] Y. Savir and T. Tlusty. "RecA-mediated homology search as a nearly optimal signal detection system." In: *Mol Cell* 40.3 (2010), pp. 388–96.

[347] G.M. Torrie and J.P. Valleau. "Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling." In: *J. Comput. Phys.* 23.2 (1977), pp. 187–199.

[348] S. Whitelam, E. H. Feng, M. F. Hagan, and P. L. Geissler. "The role of collective motion in examples of coarsening and self-assembly." In: *Soft Matter* 5.6 (2009), pp. 1251–1262.

[349] G. Coop and M. Przeworski. "An evolutionary view of human recombination." In: *Nat Rev Genet* 8.1 (2007), pp. 23–34.

[350] D. Ortiz-Barrientos, J. Engelstadter, and L. H. Rieseberg. "Recombination Rate Evolution and the Origin of Species." In: *Trends Ecol Evol* 31.3 (2016), pp. 226–36.

[351] H. Flores-Rozas and R. D. Kolodner. "Links between replication, recombination and genome instability in eukaryotes." In: *Trends Biochem Sci* 25.4 (2000), pp. 196–200.

[352] R. D. Kolodner, C. D. Putnam, and K. Myung. "Maintenance of genome stability in Saccharomyces cerevisiae." In: *Science* 297.5581 (2002), pp. 552–7.

[353] A. Malkova and J. E. Haber. "Mutations arising during repair of chromosome breaks." In: *Annu Rev Genet* 46 (2012), pp. 455–73.

[354] S. M. Soucy, J. Huang, and J. P. Gogarten. "Horizontal gene transfer: building the web of life." In: *Nat Rev Genet* 16.8 (2015), pp. 472–82.

[355] N. Hunter. "Meiotic Recombination: The Essence of Heredity." In: *Cold Spring Harb Perspect Biol* 7.12 (2015).

[356] S. C. Kowalczykowski. "An Overview of the Molecular Mechanisms of Recombinational DNA Repair." In: *Cold Spring Harb Perspect Biol* 7.11 (2015).

[357] L. Aravind, D. R. Walker, and E. V. Koonin. "Conserved domains in DNA repair proteins and evolution of repair systems." In: *Nucleic Acids Res* 27.5 (1999), pp. 1223–42.

[358] Z. Lin, H. Kong, M. Nei, and H. Ma. "Origins and evolution of the recA/RAD51 gene family: evidence for ancient gene duplication and endosymbiotic gene transfer." In: *Proc Natl Acad Sci U S A* 103.27 (2006), pp. 10328–33.

[359] M. Prentiss, C. Prevost, and C. Danilowicz. "Structure/function relationships in RecA protein-mediated homology recognition and strand exchange." In: *Crit Rev Biochem Mol Biol* 50.6 (2015), pp. 453–76.

[360] E. H. Egelman and A. Stasiak. "Structure of helical RecA-DNA complexes. Complexes formed in the presence of ATP-gamma-S or ATP." In: *J Mol Biol* 191.4 (1986), pp. 677–97.

[361] T. Ogawa, X. Yu, A. Shinohara, and E. H. Egelman. "Similarity of the yeast RAD51 filament to the bacterial RecA filament." In: *Science* 259.5103 (1993), pp. 1896–9.

[362] X. Yu, S. A. Jacobs, S. C. West, T. Ogawa, and E. H. Egelman. "Domain structure and dynamics in the helical filaments formed by RecA and Rad51 on DNA." In: *Proc Natl Acad Sci U S A* 98.15 (2001), pp. 8419–24.

[363] J. Y. Lee, T. Terakawa, Z. Qi, J. B. Steinfeld, S. Redding, Y. Kwon, W. A. Gaines, W. Zhao, P. Sung, and E. C. Greene. "DNA RECOMBINATION. Base triplet stepping by the Rad51/RecA family of recombinases." In: *Science* 349.6251 (2015), pp. 977–81.

[364] J. Y. Lee, Z. Qi, and E. C. Greene. "ATP hydrolysis Promotes Duplex DNA Release by the RecA Presynaptic Complex." In: *J Biol Chem* 291.42 (2016), pp. 22218–22230.

[365] Z. Qi and E. C. Greene. "Visualizing recombination intermediates with single-stranded DNA curtains." In: *Methods* 105 (2016), pp. 62–74.

[366] J. I. Kim, M. M. Cox, and R. B. Inman. "On the role of ATP hydrolysis in RecA protein-mediated DNA strand exchange. I. Bypassing a short heterologous insert in one DNA substrate." In: *J Biol Chem* 267.23 (1992), pp. 16438–43.

[367] W. Rosselli and A. Stasiak. "The ATPase activity of RecA is needed to push the DNA strand exchange through heterologous regions." In: *EMBO J* 10.13 (1991), pp. 4391–6.

[368] M. M. Cox and I. R. Lehman. "Directionality and polarity in recA protein-promoted branch migration." In: *Proc Natl Acad Sci U S A* 78.10 (1981), pp. 6018–22.

[369] Y. Murayama, Y. Tsutsui, and H. Iwasaki. "The fission yeast meiosis-specific Dmc1 recombinase mediates formation and branch migration of Holliday junctions by preferentially promoting strand exchange in a direction opposite to that of Rad51." In: *Genes Dev* 25.5 (2011), pp. 516–27.

[370] P. Sung and D. L. Robberson. "DNA strand exchange mediated by a RAD51-ssDNA nucleoprotein filament with polarity opposite to that of RecA." In: *Cell* 82.3 (1995), pp. 453–61.

[371] S. C. West, E. Cassuto, and P. Howard-Flanders. "Heteroduplex formation by recA protein: polarity of strand exchanges." In: *Proc Natl Acad Sci U S A* 78.10 (1981), pp. 6149–53.

[372] E. A. Namsaraev and P. Berg. "Branch migration during Rad51-promoted strand exchange proceeds in either direction." In: *Proc Natl Acad Sci U S A* 95.18 (1998), pp. 10477–81.

[373] E. A. Namsaraev and P. Berg. "Rad51 uses one mechanism to drive DNA strand exchange in both directions." In: *J Biol Chem* 275.6 (2000), pp. 3970–6.

[374] M. M. Cox. "Motoring along with the bacterial RecA protein." In: *Nat Rev Mol Cell Biol* 8.2 (2007), pp. 127–38.

[375] C. Danilowicz, D. Yang, C. Kelley, C. Prevost, and M. Prentiss. "The poor homology stringency in the heteroduplex allows strand exchange to incorporate desirable mismatches without sacrificing recognition in vivo." In: *Nucleic Acids Res* 43.13 (2015), pp. 6473–85.

[376] R. D. Camerini-Otero and P. Hsieh. "Homologous recombination proteins in prokaryotes and eukaryotes." In: *Annu Rev Genet* 29 (1995), pp. 509–52.

[377] N. S. Persky and S. T. Lovett. "Mechanisms of recombination: lessons from E. coli." In: *Crit Rev Biochem Mol Biol* 43.6 (2008), pp. 347–70.

[378] R. Anand, A. Beach, K. Li, and J. Haber. "Rad51-mediated double-strand break repair and mismatch correction of divergent substrates." In: *Nature* 544.7650 (2017), pp. 377–380.

[379] L. S. Symington, R. Rothstein, and M. Lisby. "Mechanisms and regulation of mitotic recombination in Saccharomyces cerevisiae." In: *Genetics* 198.3 (2014), pp. 795–835.

[380] M. S. Brown and D. K. Bishop. "DNA strand exchange and RecA homologs in meiosis." In: *Cold Spring Harb Perspect Biol* 7.1 (2014), a016659.

[381] J. Y. Lee, J. B. Steinfeld, Z. Qi, Y. Kwon, P. Sung, and E. C. Greene. "Sequence imperfections and base triplet recognition by the Rad51/RecA family of recombinases." In: *J Biol Chem* 292.26 (2017), pp. 11125–11135.

[382] M. V. Borgogno, M. R. Monti, W. Zhao, P. Sung, C. E. Argarana, and R. J. Pezza. "Tolerance of DNA Mismatches in Dmc1 Recombinase-mediated DNA Strand Exchange." In: *J Biol Chem* 291.10 (2016), pp. 4928–38.

[383] S. V. Chintapalli, G. Bhardwaj, J. Babu, L. Hadjiyianni, Y. Hong, G. K. Todd, C. A. Boosalis, Z. Zhang, X. Zhou, H. Ma, A. Anishkin, D. B. van Rossum, and R. L. Patterson. "Reevaluation of the evolutionary events within recA/RAD51 phylogeny." In: *BMC Genomics* 14 (2013), p. 240.

[384] R. M. Story, D. K. Bishop, N. Kleckner, and T. A. Steitz. "Structural relationship of bacterial RecA proteins to recombination proteins from bacteriophage T4 and yeast." In: *Science* 259.5103 (1993), pp. 1892–6.

[385] C. J. Ma, J. B. Steinfeld, and E. C. Greene. "Single-Stranded DNA Curtains for Studying Homologous Recombination." In: *Methods Enzymol* 582 (2017), pp. 193–219.

[386] M. R. G. Taylor, M. Spirek, C. Jian Ma, R. Carzaniga, T. Takaki, L. M. Collinson, E. C. Greene, L. Krejci, and S. J. Boulton. "A Polar and Nucleotide-Dependent Mechanism of Action for RAD51 Paralogs in RAD51 Filament Remodeling." In: *Mol Cell* 64.5 (2016), pp. 926–939.

[387] P. Hsieh, C. S. Camerini-Otero, and R. D. Camerini-Otero. "The synapsis event in the homologous pairing of DNAs: RecA recognizes and pairs less than one helical repeat of DNA." In: *Proc Natl Acad Sci U S A* 89.14 (1992), pp. 6492–6.

[388] J. M. Short, Y. Liu, S. Chen, N. Soni, M. S. Madhusudhan, M. K. Shivji, and A. R. Venkitaraman. "High-resolution structure of the presynaptic RAD51 filament on single-stranded DNA by electron cryo-microscopy." In: *Nucleic Acids Res* 44.19 (2016), pp. 9017–9030.

[389] Dana Branzei and Barnabas Szakal. "Building up and breaking down: mechanisms controlling recombination during replication." In: *Critical Reviews in Biochemistry and Molecular Biology* 52.4 (2017), pp. 381–394.

[390] Alexander Lorenz. “Modulation of meiotic homologous recombination by DNA helicases.” In: *Yeast* 34.5 (2017), pp. 195–203.

[391] Rachelle Miller Spell and Sue Jinks-Robertson. “Examination of the Roles of Sgs1 and Srs2 Helicases in the Enforcement of Recombination Fidelity in Saccharomyces cerevisiae.” In: *Genetics* 168.4 (2004), pp. 1855–1865.

[392] Maria Spies and Richard Fishel. “Mismatch Repair during Homologous and Homeologous Recombination.” In: *Cold Spring Harbor Perspectives in Biology* 7.3 (2015), a022657.

[393] N. Sugawara, T. Goldfarb, B. Studamire, E. Alani, and J. E. Haber. “Heteroduplex rejection during single-strand annealing requires Sgs1 helicase and mismatch repair proteins Msh2 and Msh6 but not Pms1.” In: *Proceedings of the National Academy of Sciences* 101.25 (2004), pp. 9315–9320.

[394] M. A. Ramesh, S. B. Malik, and Jr. Logsdon J. M. “A phylogenomic inventory of meiotic genes; evidence for sex in Giardia and an early eukaryotic origin of meiosis.” In: *Curr Biol* 15.2 (2005), pp. 185–91.

[395] A. M. Villeneuve and K. J. Hillers. “Whence meiosis?” In: *Cell* 106.6 (2001), pp. 647–50.

[396] Ana María León-Ortiz, Stephanie Panier, Grzegorz Sarek, Jean-Baptiste Vannier, Harshil Patel, Peter J. Campbell, and Simon J. Boulton. “A Distinct Class of Genome Rearrangements Driven by Heterologous Recombination.” In: *Molecular Cell* 69.2 (2018), 292–305.e6.

[397] J. L. Gerton and R. S. Hawley. “Homologous chromosome interactions in meiosis: diversity amidst conservation.” In: *Nat Rev Genet* 6.6 (2005), pp. 477–87.

[398] A. Storlazzi, S. Tesse, S. Gargano, F. James, N. Kleckner, and D. Zickler. “Meiotic double-strand breaks at the interface of chromosome movement, chromosome remodeling, and reductional division.” In: *Genes Dev* 17.21 (2003), pp. 2675–87.

[399] F. M. Cohan and S. Aracena. “Prokaryotic sex: eukaryote-like qualities of recombination in an Archaean lineage.” In: *Curr Biol* 22.15 (2012), R601–2.

[400] M. Mevarech and R. Werczberger. “Genetic transfer in Halobacterium volcanii.” In: *J Bacteriol* 162.1 (1985), pp. 461–2.

[401] A. Naor and U. Gophna. “Cell fusion and hybrids in Archaea: prospects for genome shuffling and accelerated strain development for biotechnology.” In: *Bioengineered* 4.3 (2013), pp. 126–9.

[402] A. Naor, P. Lapierre, M. Mevarech, R. T. Papke, and U. Gophna. "Low species barriers in halophilic archaea and the formation of recombinant hybrids." In: *Curr Biol* 22.15 (2012), pp. 1444–8.

[403] R. T. Papke, J. E. Koenig, F. Rodriguez-Valera, and W. F. Doolittle. "Frequent recombination in a saltern population of Halorubrum." In: *Science* 306.5703 (2004), pp. 1928–9.

[404] I. Rosenshine, R. Tchelet, and M. Mevarech. "The mechanism of DNA transfer in the mating system of an archaebacterium." In: *Science* 245.4924 (1989), pp. 1387–9.

[405] LLC Schrodinger. "The PyMOL Molecular Graphics System, Version 1.8." 2015.

[406] J. S. Papadopoulos and R. Agarwala. "COBALT: constraint-based alignment tool for multiple protein sequences." In: *Bioinformatics* 23.9 (2007), pp. 1073–9.

[407] X. Robert and P. Gouet. "Deciphering key features in protein structures with the new ENDscript server." In: *Nucleic Acids Res* 42.Web Server issue (2014), W320–4.

[408] C. D. Lee, H. C. Sun, S. M. Hu, C. F. Chiu, A. Homhuan, S. M. Liang, C. H. Leng, and T. F. Wang. "An improved SUMO fusion protein system for effective production of native proteins." In: *Protein Sci* 17.7 (2008), pp. 1241–8.

[409] S. Sigurdsson, K. Trujillo, B. Song, S. Stratton, and P. Sung. "Basis for avid homologous DNA strand exchange by human Rad51 and RPA." In: *J Biol Chem* 276.12 (2001), pp. 8798–806.

[410] P. Chi, S. Van Komen, M. G. Sehorn, S. Sigurdsson, and P. Sung. "Roles of ATP binding and ATP hydrolysis in human Rad51 recombinase function." In: *DNA Repair (Amst)* 5.3 (2006), pp. 381–91.

[411] Y. Kwon, W. Zhao, and P. Sung. "Biochemical studies on human Rad51-mediated homologous recombination." In: *Methods Mol Biol* 745 (2011), pp. 421–35.

[412] N. Erdeniz, U. H. Mortensen, and R. Rothstein. "Cloning-free PCR-based allele replacement methods." In: *Genome Res* 7.12 (1997), pp. 1174–83.

[413] S. Scherer and R. W. Davis. "Replacement of chromosome segments with altered DNA sequences constructed in vitro." In: *Proc Natl Acad Sci U S A* 76.10 (1979), pp. 4951–5.

[414] C. S. Hoffman and F. Winston. "A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of Escherichia coli." In: *Gene* 57.2-3 (1987), pp. 267–72.

[415] S. A. Narod and W. D. Foulkes. "BRCA1 and BRCA2: 1994 and beyond." In: *Nat Rev Cancer* 4.9 (2004), pp. 665–76.

[416] R. Roy, J. Chun, and S. N. Powell. "BRCA1 and BRCA2: different roles in a common pathway of genome protection." In: *Nat Rev Cancer* 12.1 (2011), pp. 68–78.

[417] S. L. Sawyer, L. Tian, M. Kahkonen, J. Schwartzentruber, M. Kircher, Genomics University of Washington Centre for Mendelian, Forge Canada Consortium, J. Majewski, D. A. Dyment, A. M. Innes, K. M. Boycott, L. A. Moreau, J. S. Moilanen, and R. A. Greenberg. "Biallelic mutations in BRCA1 cause a new Fanconi anemia subtype." In: *Cancer Discov* 5.2 (2015), pp. 135–42.

[418] D. P. Silver and D. M. Livingston. "Mechanisms of BRCA1 tumor suppression." In: *Cancer Discov* 2.8 (2012), pp. 679–84.

[419] P. A. Futreal, Q. Liu, D. Shattuck-Eidens, C. Cochran, K. Harshman, S. Tavtigian, L. M. Bennett, A. Haugen-Strano, J. Swensen, Y. Miki, and et al. "BRCA1 mutations in primary breast and ovarian carcinomas." In: *Science* 266.5182 (1994), pp. 120–2.

[420] A. K. Godwin, L. Vanderveer, D. C. Schultz, H. T. Lynch, D. A. Altomare, K. H. Buetow, M. Daly, L. A. Getts, A. Masny, N. Rosenblum, and et al. "A common region of deletion on chromosome 17q in both sporadic and familial epithelial ovarian tumors distal to BRCA1." In: *Am J Hum Genet* 55.4 (1994), pp. 666–77.

[421] Y. Miki, J. Swensen, D. Shattuck-Eidens, P. A. Futreal, K. Harshman, S. Tavtigian, Q. Liu, C. Cochran, L. M. Bennett, W. Ding, and et al. "A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1." In: *Science* 266.5182 (1994), pp. 66–71.

[422] S. Chang and S. K. Sharan. "BRCA1 and microRNAs: emerging networks and potential therapeutic targets." In: *Mol Cells* 34.5 (2012), pp. 425–32.

[423] S. Kawai and A. Amano. "BRCA1 regulates microRNA biogenesis via the DROSHA microprocessor complex." In: *J Cell Biol* 197.2 (2012), pp. 201–8.

[424] F. E. Kleiman and J. L. Manley. "Functional interaction of BRCA1-associated BARD1 with polyadenylation factor CstF-50." In: *Science* 285.5433 (1999), pp. 1576–9.

[425] F. E. Kleiman and J. L. Manley. "The BARD1-CstF-50 interaction links mRNA 3' end formation to DNA damage and tumor suppression." In: *Cell* 104.5 (2001), pp. 743–53.

[426] K. I. Savage, J. J. Gorski, E. M. Barros, G. W. Irwin, L. Manti, A. J. Powell, A. Pellagatti, N. Lukashchuk, D. J. McCance, W. G. McCluggage, G. Schettino, M.

Salto-Tellez, J. Boultwood, D. J. Richard, S. S. McDade, and D. P. Harkin. "Identification of a BRCA1-mRNA splicing complex required for efficient DNA repair and maintenance of genomic stability." In: *Mol Cell* 54.3 (2014), pp. 445–59.

[427] C. X. Deng. "BRCA1: cell cycle checkpoint, genetic instability, DNA damage response and cancer evolution." In: *Nucleic Acids Res* 34.5 (2006), pp. 1416–26.

[428] E. Hatchi, K. Skourti-Stathaki, S. Ventz, L. Pinello, A. Yen, K. Kamieniarz-Gdula, S. Dimitrov, S. Pathania, K. M. McKinney, M. L. Eaton, M. Kellis, S. J. Hill, G. Parmigiani, N. J. Proudfoot, and D. M. Livingston. "BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair." In: *Mol Cell* 57.4 (2015), pp. 636–47.

[429] K. Schlacher, H. Wu, and M. Jasin. "A distinct replication fork protection pathway connects Fanconi anemia tumor suppressors to RAD51-BRCA1/2." In: *Cancer Cell* 22.1 (2012), pp. 106–16.

[430] K. W. Caestecker and G. R. Van de Walle. "The role of BRCA1 in DNA double-strand repair: past and present." In: *Exp Cell Res* 319.5 (2013), pp. 575–87.

[431] M. E. Moynahan, J. W. Chiu, B. H. Koller, and M. Jasin. "Brca1 controls homology-directed DNA repair." In: *Mol Cell* 4.4 (1999), pp. 511–8.

[432] R. Scully, J. Chen, A. Plug, Y. Xiao, D. Weaver, J. Feunteun, T. Ashley, and D. M. Livingston. "Association of BRCA1 with Rad51 in mitotic and meiotic cells." In: *Cell* 88.2 (1997), pp. 265–75.

[433] I. Irminger-Finger and C. E. Jefford. "Is there more to BARD1 than BRCA1?" In: *Nat Rev Cancer* 6.5 (2006), pp. 382–91.

[434] L. C. Wu, Z. W. Wang, J. T. Tsan, M. A. Spillman, A. Phung, X. L. Xu, M. C. Yang, L. Y. Hwang, A. M. Bowcock, and R. Baer. "Identification of a RING protein that can interact in vivo with the BRCA1 gene product." In: *Nat Genet* 14.4 (1996), pp. 430–40.

[435] R. M. Densham, A. J. Garvin, H. R. Stone, J. Strachan, R. A. Baldock, M. Daza-Martin, A. Fletcher, S. Blair-Reid, J. Beesley, B. Johal, L. H. Pearl, R. Neely, N. H. Keep, F. Z. Watts, and J. R. Morris. "Human BRCA1-BARD1 ubiquitin ligase activity counteracts chromatin barriers to DNA resection." In: *Nat Struct Mol Biol* 23.7 (2016), pp. 647–55.

[436] M. Laufer, S. V. Nandula, A. P. Modi, S. Wang, M. Jasin, V. V. Murty, T. Ludwig, and R. Baer. "Structural requirements for the BARD1 tumor suppressor in chromosomal stability and homology-directed DNA repair." In: *J Biol Chem* 282.47 (2007), pp. 34325–33.

[437] E. E. McCarthy, J. T. Celebi, R. Baer, and T. Ludwig. "Loss of Bard1, the heterodimeric partner of the Brca1 tumor suppressor, results in early embryonic lethality and chromosomal instability." In: *Mol Cell Biol* 23.14 (2003), pp. 5056–63.

[438] U. K. Westermark, M. Reyngold, A. B. Olshen, R. Baer, M. Jasin, and M. E. Moynahan. "BARD1 participates with BRCA1 in homology-directed repair of chromosome breaks." In: *Mol Cell Biol* 23.21 (2003), pp. 7926–36.

[439] C. Ghimenti, E. Sensi, S. Presciuttini, I. M. Brunetti, P. Conte, G. Bevilacqua, and M. A. Caligo. "Germline mutations of the BRCA1-associated ring domain (BARD1) gene in breast and breast/ovarian families negative for BRCA1 and BRCA2 alterations." In: *Genes Chromosomes Cancer* 33.3 (2002), pp. 235–42.

[440] M. Ishitobi, Y. Miyoshi, S. Hasegawa, C. Egawa, Y. Tamaki, M. Monden, and S. Noguchi. "Mutational analysis of BARD1 in familial breast cancer patients in Japan." In: *Cancer Lett* 200.1 (2003), pp. 1–7.

[441] S. M. Karppinen, K. Heikkinen, K. Rapakko, and R. Winqvist. "Mutation screening of the BARD1 gene: evidence for involvement of the Cys557Ser allele in hereditary susceptibility to breast cancer." In: *J Med Genet* 41.9 (2004), e114.

[442] T. H. Thai, F. Du, J. T. Tsan, Y. Jin, A. Phung, M. A. Spillman, H. F. Massa, C. Y. Muller, R. Ashfaq, J. M. Mathis, D. S. Miller, B. J. Trask, R. Baer, and A. M. Bowcock. "Mutations in the BRCA1-associated RING domain (BARD1) gene in primary breast, ovarian and uterine cancers." In: *Hum Mol Genet* 7.2 (1998), pp. 195–202.

[443] J. Y. Wu, A. T. Vlastos, M. F. Pelte, M. A. Caligo, A. Bianco, K. H. Krause, G. J. Laurent, and I. Irminger-Finger. "Aberrant expression of BARD1 in breast and ovarian cancers with poor prognosis." In: *Int J Cancer* 118.5 (2006), pp. 1215–26.

[444] L. S. Symington. "DNA repair: Making the cut." In: *Nature* 514.7520 (2014), pp. 39–40.

[445] J. M. Daley and P. Sung. "53BP1, BRCA1, and the choice between recombination and end joining at DNA double-strand breaks." In: *Mol Cell Biol* 34.8 (2014), pp. 1380–8.

[446] T. T. Paull, D. Cortez, B. Bowers, S. J. Elledge, and M. Gellert. "Direct DNA binding by Brca1." In: *Proc Natl Acad Sci U S A* 98.11 (2001), pp. 6086–91.

[447] A. M. Simons, A. A. Horwitz, L. M. Starita, K. Griffin, R. S. Williams, J. N. Glover, and J. D. Parvin. "BRCA1 DNA-binding activity is stimulated by BARD1." In: *Cancer Res* 66.4 (2006), pp. 2012–8.

[448] J. San Filippo, P. Chi, M. G. Sehorn, J. Etchin, L. Krejci, and P. Sung. "Recombination mediator and Rad51 targeting activities of a human BRCA2 polypeptide." In: *J Biol Chem* 281.17 (2006), pp. 11649–57.

[449] W. Zhao, S. Vaithiyalingam, J. San Filippo, D. G. Maranon, J. Jimenez-Sainz, G. V. Fontenay, Y. Kwon, S. G. Leung, L. Lu, R. B. Jensen, W. J. Chazin, C. Wiese, and P. Sung. "Promotion of BRCA2-Dependent Homologous Recombination by DSS1 via RPA Targeting and DNA Mimicry." In: *Mol Cell* 59.2 (2015), pp. 176–87.

[450] R. B. Jensen, A. Carreira, and S. C. Kowalczykowski. "Purified human BRCA2 stimulates RAD51-mediated recombination." In: *Nature* 467.7316 (2010), pp. 678–83.

[451] L. Pellegrini, D. S. Yu, T. Lo, S. Anand, M. Lee, T. L. Blundell, and A. R. Venkitaraman. "Insights into DNA recombination from the structure of a RAD51-BRCA2 complex." In: *Nature* 420.6913 (2002), pp. 287–93.

[452] O. V. Kovalenko, C. Wiese, and D. Schild. "RAD51AP2, a novel vertebrate- and meiotic-specific protein, shares a conserved RAD51-interacting C-terminal domain with RAD51AP1/PIR51." In: *Nucleic Acids Res* 34.18 (2006), pp. 5081–92.

[453] W. Zhao and P. Sung. "Significance of ligand interactions involving Hop2-Mnd1 and the RAD51 and DMC1 recombinases in homologous DNA repair and XX ovarian dysgenesis." In: *Nucleic Acids Res* 43.8 (2015), pp. 4055–66.

[454] E. Cerami, J. Gao, U. Dogrusoz, B. E. Gross, S. O. Sumer, B. A. Aksoy, A. Jacobsen, C. J. Byrne, M. L. Heuer, E. Larsson, Y. Antipin, B. Reva, A. P. Goldberg, C. Sander, and N. Schultz. "The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data." In: *Cancer Discov* 2.5 (2012), pp. 401–4.

[455] J. Gao, B. A. Aksoy, U. Dogrusoz, G. Dresdner, B. Gross, S. O. Sumer, Y. Sun, A. Jacobsen, R. Sinha, E. Larsson, E. Cerami, C. Sander, and N. Schultz. "Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal." In: *Sci Signal* 6.269 (2013), pl1.

[456] K. Nakanishi, Y. G. Yang, A. J. Pierce, T. Taniguchi, M. Digweed, A. D. D'Andrea, Z. Q. Wang, and M. Jasin. "Human Fanconi anemia monoubiquitination pathway promotes homologous DNA repair." In: *Proc Natl Acad Sci U S A* 102.4 (2005), pp. 1110–5.

[457] A. Orthwein, S. M. Noordermeer, M. D. Wilson, S. Landry, R. I. Enchev, A. Sherker, M. Munro, J. Pinder, J. Salsman, G. Dellaire, B. Xia, M. Peter, and D. Durocher. "A mechanism for the suppression of homologous recombination in G1 cells." In: *Nature* 528.7582 (2015), pp. 422–6.

[458] J. Pinder, J. Salsman, and G. Dellaire. “Nuclear domain ’knock-in’ screen for the evaluation and identification of small molecule enhancers of CRISPR-based genome editing.” In: *Nucleic Acids Res* 43.19 (2015), pp. 9379–92.

[459] S. F. Bunting, E. Callen, N. Wong, H. T. Chen, F. Polato, A. Gunn, A. Bothmer, N. Feldhahn, O. Fernandez-Capetillo, L. Cao, X. Xu, C. X. Deng, T. Finkel, M. Nussenzweig, J. M. Stark, and A. Nussenzweig. “53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks.” In: *Cell* 141.2 (2010), pp. 243–54.

[460] W. Zhao, D. Saro, M. Hammel, Y. Kwon, Y. Xu, R. P. Rambo, G. J. Williams, P. Chi, L. Lu, R. J. Pezza, R. D. Camerini-Otero, J. A. Tainer, H. W. Wang, and P. Sung. “Mechanistic insights into the role of Hop2-Mnd1 in meiotic homologous DNA pairing.” In: *Nucleic Acids Res* 42.2 (2014), pp. 906–17.

[461] C. Wiese, E. Dray, T. Groesser, J. San Filippo, I. Shi, D. W. Collins, M. S. Tsai, G. J. Williams, B. Rydberg, P. Sung, and D. Schild. “Promotion of homologous recombination and genomic stability by RAD51AP1 via RAD51 recombinase enhancement.” In: *Mol Cell* 28.3 (2007), pp. 482–90.

[462] J. D. Dignam, R. M. Lebovitz, and R. G. Roeder. “Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei.” In: *Nucleic Acids Res* 11.5 (1983), pp. 1475–89.