RAD52-mediated
single-strand DNA annealing:
a single-molecule approach

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von
Maria Kharlamova
aus Lobnja, Russland

Tübingen
2023

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

To my parents and to my husband

"If you thought that science was certain - well,
that is just an error on your part."
Richard P. Feynman

Abstract

The integrity of the genetic information is a fundamental requirement for the maintenance of somatic cell lineages and crucially, the maintenance and transfer of genetic traits within a gene pool. However, there is a plethora of exogenous factors, such as UV-light, radiation, and various chemicals, and endogenous factors including reactive oxygen species, failed DNA replication, and meiosis that can affect genome integrity. The variety of factors and damaging agents result in a multitude of DNA damage types with different levels of cytotoxicity. Amongst all of them, DNA double-strand breaks are the most severe type of DNA damage. Since both strands are affected, it results in a more complex process of information restoration. Cells utilize two major pathways to repair DNA double-strand breaks: non-homologous end joining and homologous recombination. While the non-homologous end joining relies only on short homologous stretches of DNA, it leads to the accumulation of errors and point mutations within the genome. On the contrary, DNA repair *via* homologous recombination utilizes a sister chromatid to ensure faithful restoration of the genetic information. Conventional homologous recombination in mammalian cells is dependent on RAD51, which requires a variety of interaction partners to ensure its localization on DNA strands. In yeast, Rad52 has been identified as a mediator that displaces DNA end-protection proteins and facilitates RAD51 loading. In mammalian cells, this function is taken over by BRCA2. However, in the case of BRCA2-deficient cells, which is the hallmark of many cancers, human RAD52 can overcome this deficiency, secure RAD51 loading *in vitro* and promote the survival of cancer cells. Thus, RAD52 has become a very prominent target molecule in anti-cancer therapeutic studies. Nevertheless, the exact mechanism by which RAD52 performs DNA annealing remains unknown. The latest annealing model, utilizing the intrinsic propensity of the protein to self-oligomerize, has proposed that multiple ring structures would load onto single-stranded DNA and search for the homology while sliding along each other. This model has been established by studying only the N-terminal domain of RAD52 and by utilizing very high protein concentrations in the presence of large amounts of salt that could artificially favor the formation of ring structures. The inability of this model to explain the proof-reading mechanism together with other inconsistencies has raised a need for an alternative annealing model. By combining biochemical assays with single-molecule

mass photometry, I investigated the dependency between the DNA binding and annealing activity of the full-length RAD52 and its N-terminal truncation and their oligomeric states under various experimental conditions. These experiments revealed a striking difference between the proteins. While RAD52(209) primarily formed undecameric rings, RAD52, which was known to form ring structures consisting of seven subunits, unexpectedly, exhibited the formation of a heterogeneous pool of rings with different numbers of subunits. Moreover, the mass photometry revealed monomers and lower-order oligomers that have been seldom observed before. Upon binding to single-stranded DNA, both proteins formed higher-order oligomers. Interestingly, the DNA excess caused the dissociation of these oligomers as well as single rings increasing the fraction of short oligomers. Overall, these experiments bring new information that will contribute to the revision of RAD52's annealing mechanism and help the design of more efficient therapeutic pathways in anti-cancer research.

Zusammenfassung

Die Unverfälschtheit der genetischen Information ist eine grundlegende Voraussetzung für die Aufrechterhaltung von somatischen Zellen und insbesondere für die Erhaltung und Vererbung genetischer Merkmale innerhalb eines Genpools. Es gibt jedoch eine Fülle von exogenen Faktoren wie UV-Licht, Strahlung und verschiedene Chemikalien sowie endogene Faktoren wie reaktive Sauerstoffspezies, fehlgeschlagene DNS-Replikation oder Meiose, die die Integrität des Genoms beeinträchtigen können. Diese Vielfalt von Faktoren und schädigenden Agenzien führt zu einer Vielzahl von Arten von DNS-Schäden mit unterschiedlichem Graden an Zytotoxizität. Unter allen DNS-Schädigungen sind DNS-Doppelstrangbrüche die schwerwiegendsten. Da in diesem Fall beide Stränge betroffen sind, ist hier ein komplexerer Prozess der Informationswiederherstellung nötig um den Schaden zu beheben. Zellen nutzen hauptsächlich zwei Varianten zur Reparatur von DNS-Doppelstrangbrüchen: nicht-homologe Endverknüpfung und homologe Rekombination. Die nicht-homologe Endverknüpfung beruht zwar nur auf kurzen und homologen DNS-Abschnitten, kann aber zu einer Anhäufung von Fehlern und Punktmutationen im Genom führen. Im Gegensatz dazu wird bei der DNS-Reparatur durch homologe Rekombination ein Schwesterchromatid verwendet, um eine getreue Wiederherstellung der genetischen Information zu gewährleisten. Die konventionellehomologe Rekombination in Säugetierzellen hängt von RAD51 ab, dass eine Vielzahl von Interaktionspartnern benötigt, um seine Lokalisierung auf den DNS-Strängen sicherzustellen. In Hefe wurde Rad52 als Mediator identifiziert, der DNS-Endschutzproteine verdrängt und die Beladung von RAD51 erleichtert. In Säugetierzellen wird diese Funktion von BRCA2 übernommen. Bei Zellen mit BRCA2-Mangel, dem Kennzeichen vieler Krebsarten, kann menschliches RAD52 diesen Mangel jedoch überwinden, die RAD51-Beladung *in vitro* sicherstellen und damit das Überleben von Krebszellen fördern. Dementsprechend ist RAD52 zu einem bedeutenden Zielmolekül in therapeutischen Studien zur Krebsbekämpfung geworden. Dennoch ist nach wie vor der genaue Mechanismus wie RAD52 die DNS-Hybridisierungherbeiführt unbekannt. Das jüngste Hybridisierungs-Modell, beruht auf der Proteinen intrinsischen Eigenschaft der Selbstoligomerisierung. Es theoretisiert, dass sich mehrere Ringstrukturen auf den einzelsträngigen DNS- Sequenzen bilden und auf der Suche nach der Homologie aneinander entlang gleiten. Bei der Erstellung dieses Modells wurde je-

doch nur die N-terminale Domäne von RAD52 verwendet. Des Weiteren wurden sehr hohe Proteinkonzentrationen in Gegenwart großer Salzmengen verwendet. Gemeinsam könnte dies die Bildung von Ringstrukturen künstlich begünstigt haben. Die Unfähigkeit dieses Modells, den Korrekturlese-Mechanismus zu erklären, hat zusammen mit anderen Ungereimtheiten den Bedarf an einem alternativen Hybridisierung-Modell geweckt. Durch die Kombination von biochemischer Assays mit Einzelmolekül-Massenphotometrie habe ich die Abhängigkeit zwischen der DNS-Bindungs- und Hybridisierungsaktivität von RAD52 und seinen oligomeren Formen unter verschiedenen Versuchsbedingungen untersucht. Hierbei verwendete ich sowohl das volllängen RAD52 als auch eine verkürzte, N-Terminale Variante (RAD52(209)). Diese Experimente zeigten einen auffälligen Unterschied zwischen den beiden Proteinvarianten. Während RAD52(209) in erster Linie undekamere Ringe bildete, zeigte RAD52, von dem bekannt war, dass es aus sieben Untereinheiten bestehende Ringstrukturen bildet, überraschenderweise die Bildung eines heterogenen Pools von Ringen mit einer unterschiedlichen Anzahl von Untereinheiten. Im Weiteren zeigte die massenphotometrischen Messungen Monomere und Oligomere niedrigerer Ordnungen, die bisher nur selten beobachtet wurden. Nach Bindung an einzelsträngigen DNS-Molekülen bildeten beide Proteine Oligomere höherer Ordnungen. Interessanterweise führte ein DNS-Überschuss zur Dissoziation dieser Oligomere sowie den einzelnen Ringen und erhöhte damit den Anteil an kürzeren Oligomeren. Insgesamt liefern diese Experimente neue Informationen, die zur Revision des RAD52-Hybridisierungs-Mechanismus beitragen und damit die Entwicklung effizienterer, therapeutischen Behandlungsmöglichkeiten in der Krebsforschung unterstützen.

Contents

Chapter 1

Introduction

1.1   DNA damage and DNA repair mechanisms

Genetic information is essential at all stages in the complexity of life: from individual cells to entire organisms. Genetic information is stored in a double-helix formed of two chains of nucleic acids. These antiparallel chains form a very stable double-helical structure, widely known as Deoxyribonucleic Acid (DNA) double-helix [25]. The chains are held together *via* hydrogen bonds formed between complementary base pairs (bp): adenine with thymine and guanine with cytosine [152] (Figure 1.1, a). It is a matter of great importance for cellular viability to keep the genetic information, intact during the entire cellular lifetime and faithfully transmit it to daughter cells in order to ensure the continuous viability of an organism and the transfer of genetic traits to future progeny.

However, a plethora of factors perturb and challenge DNA integrity by inducing DNA damage [40]. These factors have been classified into three major groups [37, 75]. The first group encompasses what is qualified as exogenous factors, like X-radiation, UV exposure, medical drugs, and various toxic chemicals that may originate from the organism's environment. The second group is referred to as reactive oxygen species (ROS), which may be generated by the cellular energetic machinery and other cellular processes. The third group is the hydrolysis of nucleotides. Moreover, certain cellular processes may induce different types of DNA damage, by for instance stalling replication forks, inducing DNA catabolism, and promoting incorrect recombination [155, 79, 85]. It is estimated that each cell may experience between 10 and 100 thousand lesions per day [40, 151]. Such a massive number of errors leads to the malfunction of many cellular processes, including DNA transcription and replication. Moreover, the accumulation of point mutations in the genome as a result of the unrepaired lesions can speed up the aging process, lead to severe neurodegenerative diseases, or cause the development of cancer [10, 91, 1, 129]. There is a

broad spectrum of DNA damage types, like base-pairs mismatches, nucleotide deletions, single-strand breaks, and inter- and intra-strands crosslinks [40] (Figure 1.1, b). All of these lesions should be recognized by the cell through the DNA damage response (DDR) machinery and, consequently, faithfully repaired. Mutations or malfunction of the DDR machinery can eventually lead to the same severe diseases caused by unrepaired DNA lesions [51, 52]. DNA repair is a complex symphony of processes provided by an orchestra of many protein complexes, of which many have yet to be revealed and deciphered. Matching to the diversity of DNA lesions, cells may utilize various repair pathways [51]. For example, chemical modifications of individual bases are less harmful to the stability of the DNA double-helix and they are repaired *via* the Base Excision Repair (BER) pathway [3, 47]. Single-strand breaks are recovered *via* the Single-Strand Break Repair pathway (SSBR) [18]. More damaging types of DNA lesions that can cause the destabilization of the DNA helical structure are repaired *via* the Nucleotide Excision Repair pathway (NER) [50, 41]. Despite the impact of single-strand breaks on the DNA structure, the presence of the intact complementary strand secures the faithful repair and the full restoration of the damaged information [40]. Nevertheless, there is another type of lesion - DNA Double-Strand Breaks (DSB) that is considered to be the most harmful as both strands are affected, limiting the amount of template available for self-repair.
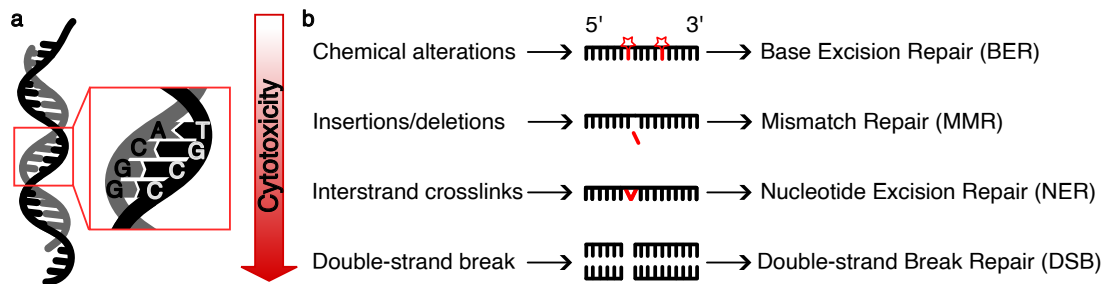
Figure 1.1: Schematic representation of the DNA structure and different types of DNA damage with the corresponding repair pathways. **(a)** DNA is a double-helix consisting of two antiparallel chains held by hydrogen bonds between complementary bases (insert). **(b)** Schematic representation of potential DNA damaging factors sorted by their cytotoxicity with corresponding repair mechanisms.

1.2 DNA double-strand break repair pathways

There are two major DNA DSB repair pathways: Non-Homologous End Joining (NHEJ) and Homologous Recombination (HR) [17, 157]. The choice of which pathway to follow depends primarily on the cell cycle, the number of breaks, and the length of the DNA overhangs [26] (Figure 1.2). For a successful HR, a sister chromatid is required, therefore this pathway is utilized during S- and G2-phases, while NHEJ can be performed during the G1 phase (Figure 1.2, a).
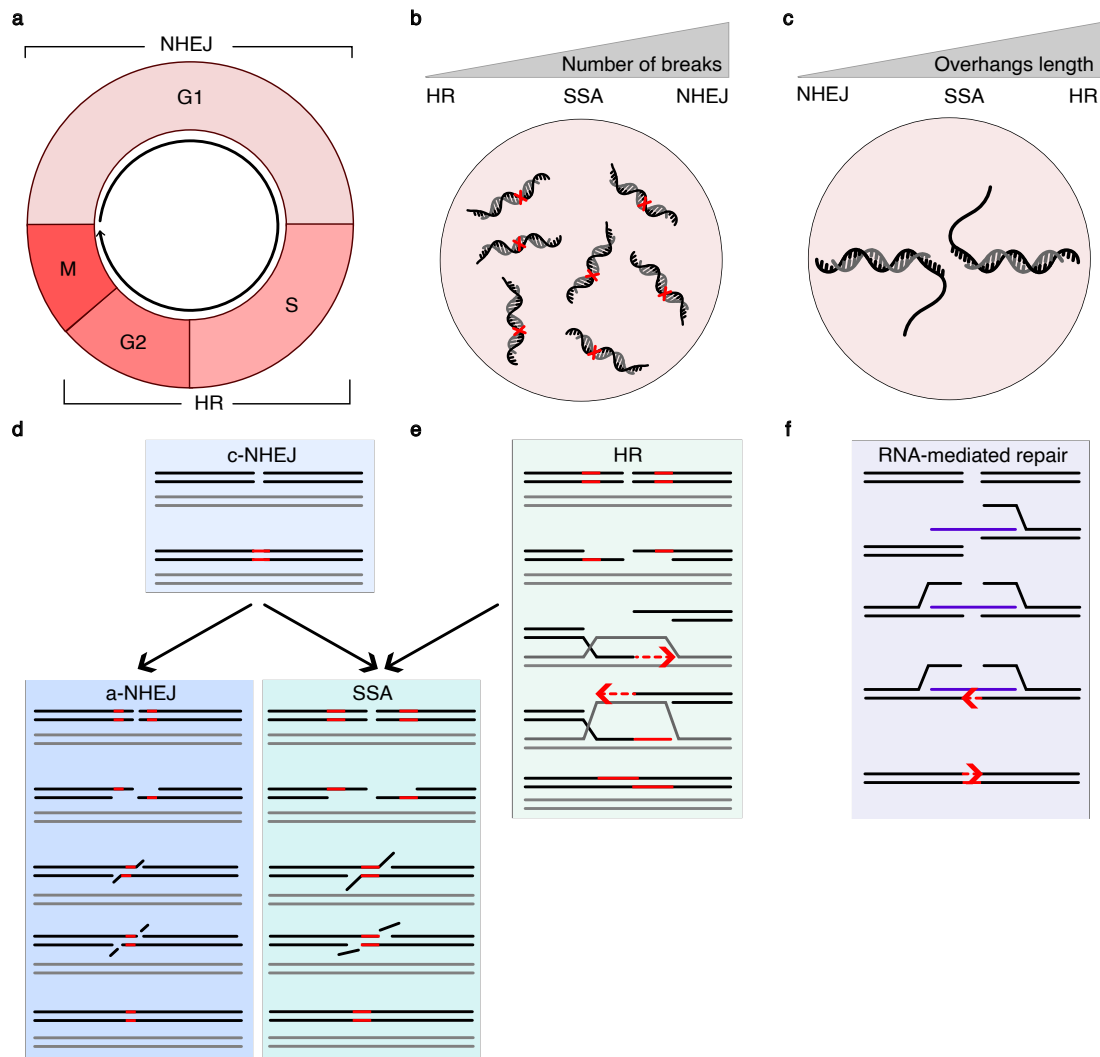
Figure 1.2: Factors affecting the choice of the DNA DSB repair mechanism. **(a)** The phase of the cell cycle. Depending on the presence of the sister chromatid DNA DSB will be repaired via c-NHEJ conventional-NHEJ or HR. **(b)** The number of double-strand breaks. HR is able to repair 5 DSBs per cell cycle. Once the number of DSBs is higher, c-NHEJ takes over the repair function. **(c)** The length of the DNA overhangs. **(d)** During the c-NHEJ no DNA overhangs are formed, and the DNA machinery ligates the blunt ends. **(e)** HR relies on the long DNA overhangs to probe the homologous regions and ensure error-free repair. NHEJ can be transformed into a-NHEJ (alternative-NHEJ) that mediates the repair based on the micro-homologous sequence alignment, or into SSA where the single-strand annealing is performed by RAD52. SSA is sometimes referred to as the RAD51-independent HR pathway. **(f)** The DSB DNA repair mediated by short single-stranded RNA molecules does not require a homologous chromatid, but the template sequence provided by RNA molecules ensures faithful repair.

1.2.1 Non-homologous End-Joining

Once DSB occurs, KU70/KU80 will respond quickly and bind to the newly-formed DNA ends as well as recruit PI3-kinase to the lesion site [89] (Figure 1.2, d). Consequently, broken ends will be processed by the Artemis nuclease and MRN (MRE11, RAD50, NBS1) complex. Finally, single-strands will be ligated by the XRCC4-Ligase IV complex [155, 16]. NHEJ is a quick solution that does not require sophisticated DNA machinery to ensure faithful repair. However, during the initial end resection, a certain amount of the base pairs can be deleted and, therefore, permanently lost from the genome. As a result, NHEJ leads to potential mismatches and loss of genetic information. Some of them will be corrected at later stages, while the rest will go undisturbed and will be transmitted to daughter cells. Thus, NHEJ is an error-prone mechanism.

1.2.2 Homologous Recombination

On the contrary, the HR pathway ensures that DNA DSBs are repaired faithfully by utilizing the intact sister chromatid as a template to restore the hereditary information, as well as *via* homology proof-reading proteins (Figure 1.2, e).

An example of this sophisticated repair machinery can be demonstrated *via* a response to DSB in mammalian cells [65]. The MRN and Ku complexes will be recruited to the lesion site and hold it together [27]. Subsequently, both MRN and Ku recruit ATM DNA-PK kinases, respectively [134]. Kinases are essential for the H2AX histone phosphorylation, as it is considered a major mark to initiate the recruitment of the DNA DSB response machinery to the lesion site [138, 109]. Post-translational modifications and rearrangements in the structure of the chromatin are essential for successful repair [40, 87]. Additionally, the formation of 3' overhangs is a requisite for HR. MRE11, CtIP and EXO1 nucleases perform the DNA-ends resections [145, 108, 98, 124]. To protect the newly formed single-stranded DNA overhangs from further degradation and in order to avoid the formation of secondary structure elements, Replication Protein A (RPA) tightly binds the overhangs [117, 26]. To ensure faithful annealing, RAD51 has to be loaded on the single-stranded DNA [80]. However, RAD51 is incompetent in displacing RPA from the DNA single-strand. Therefore, in mammalian cells, a complex machinery will recruit BRCA2 that serves as a mediator. BRCA2 displaces RPA and facilitates the loading of RAD51 [147]. For a successful performance, RAD51 additionally requires RAD54, RAD51B, RAD51C, RAD51D, XRCC2 and XRCC3 [150]. This protein assembly also contains WRN helicase, which interacts with RAD51, RAD54B, and RAD52 [8, 103]. Once the complex has been assembled, RAD51 forms a nucleoprotein filament on the DNA and initiates a single-strand invasion from the sister chromatid to copy the information [158]. This process is also known as a D-loop formation. Once one of the strands has been restored, the D-loop

is released by RAD54, and RAD52 performs the annealing of the newly restored strand and the single-stranded overhang at the lesion site [148, 15]. The single-stranded DNA gap will be eventually filled and sealed by a DNA ligase. The DNA repair machinery will be released from the restored double-stranded DNA. It is important to note that all the players of this complex machinery have yet to be deciphered. [77].

1.2.3 Alternative DNA DSB repair pathways

Although the RAD51-dependent pathway is preferential due to its high accuracy, the variety and the coordination of the protein complexes slow down the process. RAD51-dependent HR is capable of repairing around 5 lesions at a time [92]. However, there are almost 50 DNA DSBs occur during a single-cell cycle [151]. Therefore, cells need to avoid such complex DNA machinery recruitment and utilize faster pathways (Figure 1.2, b). One of them is an alternative End-Joining (a-EJ) [23] (Figure 1.2, d). At the beginning of a-EJ, MRN-CtIP nucleases resect DNA ends and form short overhangs (20 bp). Furthermore, PARP1 and θ polymerases perform the annealing of the short homologous regions (10 bp, micro-homology). Finally, DNA ends are sealed by XRCC1 and DNA ligase III [11]. RAD52 has been also shown to affect the outcome of this pathway, but its exact role has yet to be revealed [48].

Moreover, RAD52 plays a crucial role in the Single-Strand Annealing (SSA) pathway (Figure 1.2 d). The initial step is similar to previously described pathways - generation of single-stranded overhangs by BLM and WRN helicases and by the nuclease complex DNA2-CtIP-EXO1 [21]. RAD52 binds newly-formed DNA overhangs (around 30 bp) and performs their annealing [44]. After the annealing has been completed, the ERCC1-XPF complex binds RAD52 aiming to inhibit the annealing activity, while facilitating its own nuclease activity [93]. As a result, the ERCC1-XPF-RAD52 complex cleaves the leftover 3' overhangs. Similar to other pathways, missing nucleotides are filled by the DNA polymerase, and DNA ends are sealed by DNA ligase I [11]. In analogy to NHEJ, SSA is an error-prone repair pathway as the single-stranded DNA that was not annealed by RAD52 will be degraded and, therefore, this genetic information will be permanently lost.

There is one additional mechanism that despite not utilizing the information from the sister chromatid, ensures faithful repair: RNA-templated DNA DSB repair (Figure 1.2, f). Short single-stranded RNA molecules are transported to the DNA lesions by RAD52. In this case, single-stranded RNA serves a double function: as a bridge to bring broken DNA ends together and as a template to restore the lost genetic information. Remarkably, RAD52 is the only protein that has been shown to be capable of transporting RNA molecules to the lesion site [153, 86, 59].

1.2.4 DNA damage foci formation and liquid phase separation

In situ studies revealed that the DDR machinery assembles into multiple foci, remarkably, some of these factors are recruited simultaneously, like BRCA1 and MRN, and co-localize, some are not, like RAD51 and RAD52 [84, 9, 87]. These foci are very different kinetically. The first echelon of checkpoint mediators, including MRN, BRCA1, and 53BP1, will be loaded onto DNA immediately after the break. The second group, including DNA DSB-activated proteins, such as RAD54 and RAD52, is slower and loaded only after the resection of DNA ends, potentially in the response to RPA [87]. These factors form micro-foci with a lower amount of protein molecules in comparison to the "first echelon" factors foci that contain several hundreds of copies of each protein [40]. The foci's prime function is to increase the local concentration of DDR proteins at the lesion site [40, 121]. However, it has been suggested that, additionally, foci formation facilitates tethering of damaged DNA ends and prevents telomere inclusions [72]. Moreover, in yeast, the DNA damage foci are highly mobile in order to collect all the DNA lesions together [76]. This has been also demonstrated for bacteria in the case of RecA [73]. On the contrary, foci in mammalian cells appeared to be immobile [133].

Interestingly, it has been also demonstrated that some of the DDR factors, including 53BP1, RPA, and yeast Rad52 undergo liquid-liquid phase separation in response to DNA damage [7, 64, 102, 135], reinforcing the idea that the formation of foci could be a way to compartmentalize certain factors in order to ensure proper function.

1.3 Single-strand annealing proteins: from phages to humans

Annealing of the damaged DNA strands is the most crucial step in the DNA DSB repair mechanism [59]. Recently, a novel grouping of Single-strand annealing proteins (SSAP) has been proposed based on their crystal structures, revealing a common oligonucleotide-binding domain (OBD) consisting of three anti-parallel β-sheets crossed over by a single α-helix [111, 2]. Within this new family, 3 distinct groups, according to their most notorious members, are defined: Redβ/RecT, Erf, and Rad52 [56, 2]. Aside from their core function in DNA repair, that they perform *via* their N-terminal domain, single-strand annealing proteins participate in distinct cellular processes that involve the formation of single-stranded DNA such as DNA metabolism, replication, and recombination. They bind single-stranded DNA overhangs to prevent them from further degradation by nucleases or the formation of secondary-structure elements. Additionally, *via* their disordered C-terminal domain they recruit other DDR proteins to the lesion site to ensure faithful homologous recombination [111, 139, 131, 160, 31].

1.4 Single-strand annealing proteins in prokaryotes

1.4.1 Redβ and RecT

Redβ is a single-strand annealing protein from the λ phage (Figure 1.3, a-c). Its N-terminal domain carries DNA annealing and oligomerization properties, while its C-terminal domain ensures the interactions with partner proteins essential to homologous recombination [139, 131]. Moreover, the C-terminal adjusts the interaction with DNA[156]. The N-terminal domain has higher cooperativity with single-stranded DNA, while the full-length protein more stably binds newly synthesized double-stranded product [131]. Interestingly, Redβ is not capable of binding to preformed double-stranded DNA [97]. Redβ has a remarkable ability to form different types of filaments depending on the type of the DNA substrate. It can switch between helical filaments and ring structures with different numbers of protomers [106, 33]. Recently, its footprint has been finalized based on its X-ray crystal structure: Redβ binds 4 nt and 2 more are pointing out from the adjacent monomer [97]. It has been proposed that the monomers of the protein establish the original homology *via* conformational changes and oligomerization. Once the dimer has been formed, Redβ constitutes a nucleoprotein filament [33, 4]. However, recent crystal structures of the N-terminal domain demonstrated that once the filament-shaped complex of the protein with one DNA single-strand has been established, the second complementary strand binds weakly to the filament. Such a weak binding allows the diffusion of the second strand along the preformed protein-DNA filament to establish the homology region [97]. The shortest DNA length that Redβ can anneal is 16-20 bp, suggesting that at least 4 to 5 protomers should bind to the DNA to initiate the formation of the filament structure [33, 4]. This proposition can be supported by the established cooperativity binding of Redβ to the DNA molecule [131].

Similarly to Redβ, the crystal structure of RecT has recently been obtained [19] (Figure 1.3, d-f). RecT solely forms short oligomers up to 8 subunits, with the monomers being the most preferential state. On the single-stranded DNA, it forms a helical structure consisting of 8 to 10 subunits, irrespective of the length of the DNA substrate. The protein is capable of accumulating two similar DNA single-strands that cause the extension of the filament to up to 16-17 subunits. Similar to Redβ, to obtain a double-stranded DNA product, the protein has to be initially preincubated with one strand, which will bind inside the deep, narrow groove, running outside of the helical structure of the protein [97, 19]. The consequent addition of the complementary strand results in both strands being bound to this groove where the annealing will occur. The final DNA product appeared to be highly extended rather than in the helical form. Interestingly, RecT is capable of accumulating both the single- and double-stranded DNA, while the presence of Mg^{2+} inhibits the DNA binding to the protein [99]. RecT has a similar footprint as Redβ,

7

5 base pairs per monomer [19]. Overall, these 2 proteins are functionally very similar.
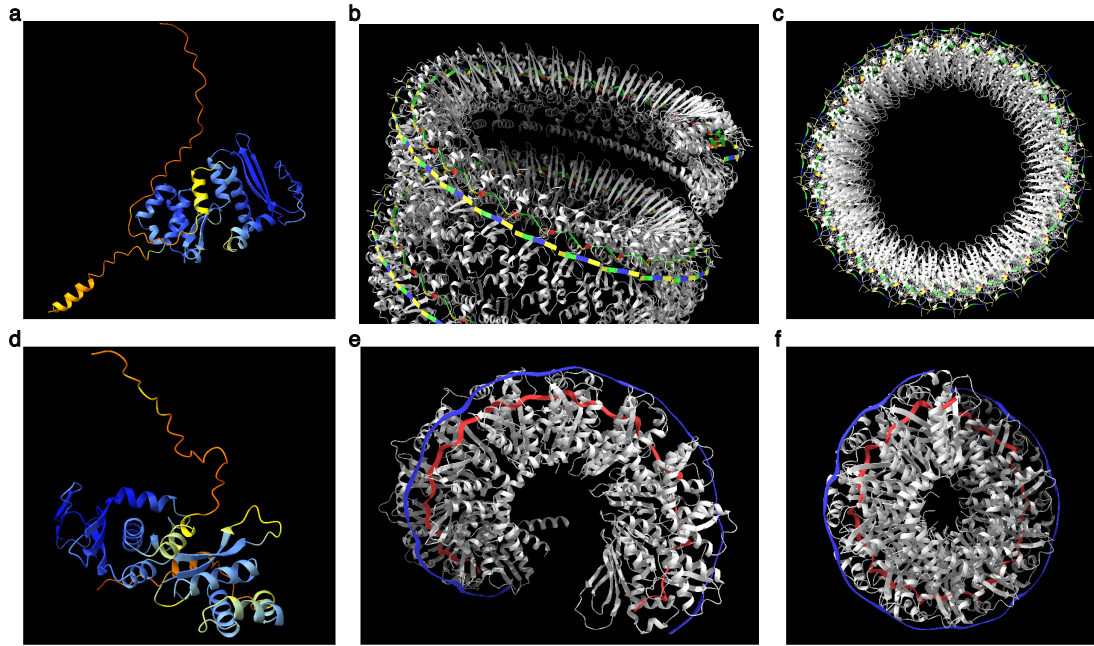
Figure 1.3: Structures of Redβ and RecT.(**a**) An AlphaFold prediction of Redβ monomer. The N-terminal domain is bound to the double-stranded DNA during the annealing process (PDB 7UJL), side view (**b**), and top view (**c**). (**d**) An AlphaFold predicted structure of RecT monomer. The filament of the RecT N-terminus is bound to the double-stranded DNA during the annealing, side view (PDB 7UB2) (**e**), and the top view (**f**). The blue color indicates the highest prediction confidence of the protein regions, and the yellow color is the lowest one. Images were created in ChimeraX [112].

1.4.2 DdrB

Unlike Redβ, the single-strand binding protein DdrB from *Deinococcus radiodurans* does not belong to the SSAP superfamily, due to topological differences of its OD [140] (Figure 1.4). The OD of DdrB established another class of single-strand annealing proteins. Nevertheless, it has certain DNA binding similarities to bacterial RecA and human RAD52 [100] and it is annealing mechanism has been revealed. *D. radiodurans* is known for an extreme survival rate after intense ionic radiation. Ddrb has been assigned, similarly to RAD52, as a key survival factor to ionizing radiation (IR) [146, 100]. DdrB monomer binds 10 nucleotides, has preferences for single-stranded DNA over double-stranded DNA, and weakly binds single-stranded RNA. DdrB does not have an ATPase activity. Moreover, it competes with *Dr*RecA for single-stranded DNA binding. *Dr*RecA fails to bind ssDNA when DdrB is already present [100]. Interestingly, DdrB has been originally proposed as an alternative SSB protein, however, its single-stranded DNA annealing activity was identified soon after [159]. It is involved in both RecA-dependent and -independent DNA

annealing pathways [159]. Similarly to RAD52, DdrB is present transiently after the IR and forms discrete foci [28]. Moreover, a truncated DdrB, lacking its disordered C-terminus, was shown the same ability to bind single-stranded DNA as its full-length counterpart [28]. Intermediate structures occurring during DNA annealing have been successfully crystallized demonstrating two pentameric rings bound with a single DNA strand that are coming in close proximity to each other while the hydrogen bonds between the complementary bases are being formed. It has been proposed that DdrB ensures the correct sequence alignment in a two-step mechanism. Once two strands, previously bound to DdrB rings, approach each other *via* ring-ring interaction, a certain fraction (around 30%) of the bases are hidden. Once the complementarity between the exposed bases is established, DdrB rings undergo conformational changes to expose the hidden bases and complete the annealing [141]. Interestingly, this study also showed that DdrB inhibits the annealing at higher protein concentrations.
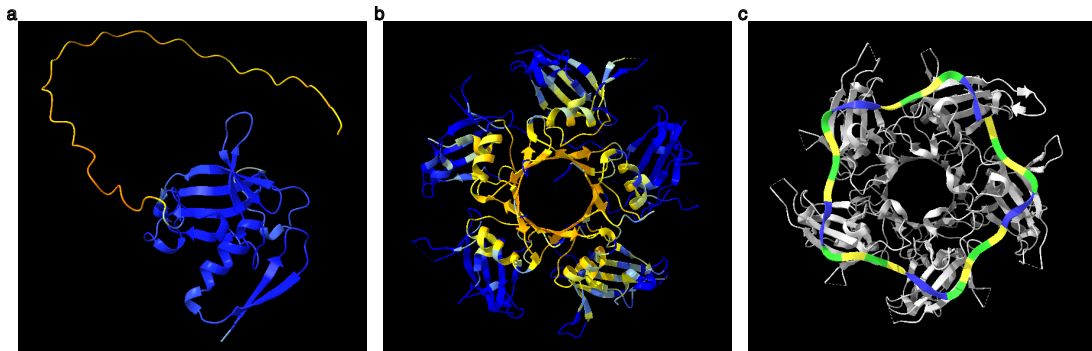
Figure 1.4: Structure of DdrB. (**a**) An AlphaFold prediction of DdrB monomer, the pentameric ring structure of DdrB alone (PDB 4EXW) (**b**), and in a complex with 30 nt DNA (PDB 4NOE) (**c**). The blue color indicates the highest prediction confidence of the protein regions, and the yellow color is the lowest one. Images were created in ChimeraX [112].

1.4.3 Erf

Erf from the bacteriophage P22 is a transcriptional repressor that is involved in the suppression of prostate cancer development [115, 12] (Figure 1.5). As the majority of the SSA protein, Erf protects plasmid DNA from the nuclease activity [115]. In solution the full-length protein forms ring structures with long protrusions, while its N-terminal domain forms rings without detectable extensions [116]. Erf binds both single- and double-stranded DNA. Interestingly, when binding the DNA molecules at physiological temperatures, Erf forms discrete filaments, while at high temperatures it forms aggregates. Recently, it has been shown that Erf forms tetrameric structures on the double-stranded DNA with GGAA recognition sequence [55]. Remarkably, Erf is considered a distinct

member of a SSA proteins superfamily due to its OBD, while its annealing activity has not yet been shown [2].
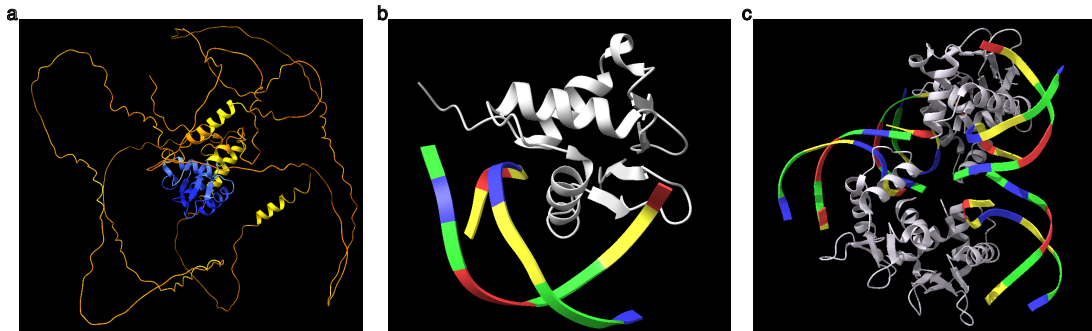
Figure 1.5: Structure of Erf. An AlphaFold prediction of a monomer (**a**), the N-terminal domain of Erf bound to the double-stranded DNA (PDB 7JSA) (**b**). Erf tetrameric structure in the oxidized form bound to two double-stranded DNA molecules (PDB 7JSL) (**c**). The blue color indicates the highest prediction confidence of the protein regions, and the yellow color is the lowest one. Images were created in ChimeraX.

1.4.4 Sak

Another single-strand annealing protein has been identified in *Lactococcus lactis* - bacterial homolog of RAD52, Sak [13, 125]. Sak preferentially binds single-stranded DNA over double-stranded DNA, performs DNA annealing, and participates in the homologous recombination *via* interaction with RecA [114]. 6His-tagged Sak forms undecameric rings in the solution and stacks thereof. It has been demonstrated that high concentrations of the protein inhibit DNA annealing [114]. Interestingly, mutated forms of Sak, lacking ring stacking properties, exhibited a reduced ability to bind DNA, suggesting that Sak oligomerization is required for protein-DNA binding [114]. A 3D model of Sak demonstrated a ring structure with a groove on the outer surface with a negatively charged central opening [125]. Sak has been proposed as a bacterial ancestor of Rad52 [114]. Nevertheless, Sak's exact mechanism of DNA annealing has yet to be elucidated.

1.5 Single-strand annealing proteins in eukaryotes

1.5.1 *RAD52* in the green lineage

Another big group of single-stand annealing proteins is presented by RAD52. This gene is highly conserved within eukaryotes. RAD52 orthologs have been identified in land plants

and algae, with the presence of at least 2 homologs in most species analyzed, suggesting a duplication event that occurred with the advent of non-vascular plants [122, 95, 29]. In *Arabidopsis thaliana*, T-DNA insertion mutants of RAD52-1 and RAD52-2, as well as knock-down lines for both homologs, resulted in reduced fertility as well as inhibition of the intrachromosomal recombination [122]. In *Oryza sativa* a RAD52 ortholog has been identified that is able to bind single- and double-stranded DNA. *Os*Rad52-2a is capable of annealing complementary single strands. Analysis of the protein's oligomeric state suggested that OsRad52-2a forms undecameric rings [95]. In Cyanidiophyceae algae, a RAD52 homolog is suggested to enhance resistance to high temperatures and acidic environments [29]. However, the research on RAD52 in plants has so far been limited, as the majority of studies have focused on studying *S. cerevisiae* and mammalian RAD52 orthologs.

1.5.2   Yeast Rad52

Yeast Rad52 carries the most crucial function during DNA repair *via* homologous recombination - a mediator activity [38] (Figure 1.6). It displaces RPA from newly-formed DNA overhangs and facilitates loading of Rad51 and stabilizes the presynaptic complex formation [88, 127, 143, 96, 113, 126]. Interestingly, Rad52-RPA interactions are species-specific. Rad52 is not capable of annealing DNA strands covered with human RPA [63]. However, it mediates the recruitment of human RAD51 but does not promote the strand exchange *in vitro* [63]. In yeast, Rad52 assists Rad51 during the D-loop formation [5]. Despite this mediator function, Rad51 competes with Rad52. For instance, Rad51 inhibits the annealing *via* Rad52 and "decides" on the final repair pathway [156]. Similarly, only in the absence of Rad51, Rad52 inhibits *de novo* telomere addition which is important for the chromosomal maintenance and stability [101, 32]. As the majority of the single-strand annealing proteins, it protects the DNA overhangs by interfering with the Sgs1 activity *via* the C-terminal domain [160]. Yeast Rad52 has two DNA binding sites, one within the N-terminal domain and the second one in the C-terminal. Rad52 C-terminal truncation inhibits its single-strand annealing activity indicating that both DNA binding sites are essential for the annealing [126, 63]. Based on this observation, the annealing mechanism by Rad52 utilizes a single-ring structure as the reaction unit [63]. Recently, a high-resolution cryo-EM structure of Rad52 has been obtained, demonstrating a decameric ring (10 subunits) [31]. This data has been complemented by the mass photometry technique showing homodecameric rings at 100 nM of Rad52 with a small fraction of short oligomers and stacks of rings. Even with a high excess of the protein over the DNA, the majority of the complexes formed had a 1:1 stoichiometry. Remarkably, a certain fraction of the protein appeared in the form of high-order oligomers. This effect has been attributed to the presence of a disordered C-terminal domain that could

11

cause Rad52 to undergo liquid-liquid phase separation *in vivo* [102, 31]. The disordered C-terminal domain was shown to interact only with a few subunits of the N-terminal domain. Moreover, it has been demonstrated that the C-terminus could inhibit the diffusion of the protein on single-stranded DNA which could be essential to identify the homology regions. Additionally, Rad51 binds to Rad52 asymmetrically and exhibits two distinct modes of interaction with Rad52: *via* the C-terminal or N-terminal domain. It has been proposed that the C-terminal domain orchestrates the initial binding and the following nucleation of Rad51 filaments [31].
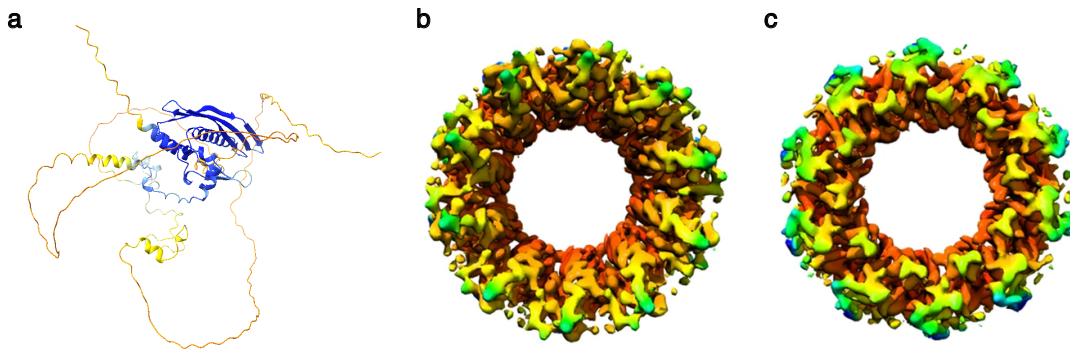
Figure 1.6: Structure of yeast Rad52. An AlphaFold prediction of a monomer (**a**). Cryo-EM maps of the N-terminal domain of Rad52, top view (**b**), bottom view (**c**). The images were adapted from Deveryshetty, *et al.* [31]. The blue color indicates the highest prediction confidence of the protein regions, and the yellow color is the lowest one. Images were created in ChimeraX [112].

1.5.3 Human RAD52

In *Homo sapiens*, single-stranded DNA annealing is performed by RAD52. Remarkably, a backup function of RAD52 has been proposed based on its yeast homology, however, it has yet to be demonstrated *in vivo*. For many years, mammalian RAD52 was out of the research focus due to its minor function in homologous recombination - as a backup for BRCA2. In the case of BRCA2-deficiency, RAD52 can partially restore the DNA repair [36]. BRCA-2 deficiency is usually a symptom of tumor cells, therefore, RAD52 facilitates tumor progression. Moreover, contrary to its yeast homolog, the mediator activity of RAD52 has never been observed *in vivo* [63]. It has been shown recently, that RAD52 instead of facilitating the load of RAD51 on the single-stranded DNA, actually competes with it [94]. As a result of such competition, cells switch from a RAD51-mediated error-free pathway to a highly-mutagenic RAD52-mediated single-strand annealing pathway that has been described above (section "Alternative DNA DSB repair pathways"). This additional mutagenic activity of RAD52 made it a very prominent target for anti-cancer research [49]. By blocking RAD52 annealing activity, it can be possible to specifically target the proliferation of malicious BRCA2-deficient cells [22]. Recently, additional functions

of RAD52 have been revealed that raised the interest to study the protein [22, 59, 119]. Apart from its single-strand annealing function, RAD52 participates in D-loop formation during RAD51-dependent homologous recombination. RAD52 also ensures the release of stall replication forks [81, 71, 132]. Moreover, as it was mentioned above (section "Alternative DNA DSB repair pathways"), RAD52 has a major function in the RNA-mediated repair mechanism [153, 86, 154].
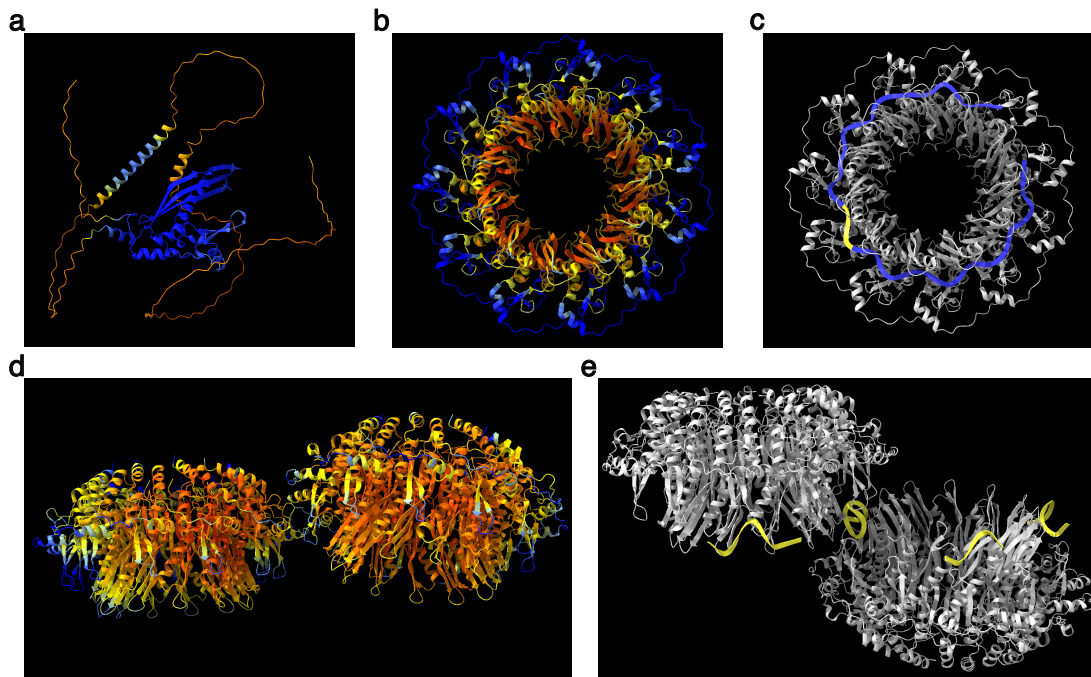
Figure 1.7: Structure of human RAD52. (**a**) An AlphaFold prediction of a monomer. Crystal structures of the N-terminal domain of RAD52 undecameric ring alone (PDB 5JRB) (**b**) and in a complex with single-stranded DNA bound to the inner groove of the ring structure (blue, PDB 5XRZ) (**c**). Crystal structures of two N-terminal domains stacking solely (PDB 1H2I) (**d**) and on the single-stranded DNA bound to the outer stem of the rings (yellow, PDB 5XS0) (**e**). The blue color indicates the highest prediction confidence of the protein regions, and the yellow color is the lowest one. Images were created in ChimeraX [112].

Despite all the functions that RAD52 carries in mammalian cells, the exact mechanism of its interaction with DNA remains debatable. As human RAD52 is the major focus of this work, the protein structure will be described as circumstantial. The first 209 to 212 residues of RAD52 are highly ordered and conserved within the species [56] (Figure 1.7). The crystal structure of the N-terminal revealed an undecameric ring with two DNA binding sites. The first DNA binding site is located in the inner groove on the outer surface of the ring and accumulates only single-stranded DNA. Upon binding to the inner groove, DNA is extended into B-form-like DNA with base pairs facing outside [61, 130, 83]. The second DNA-binding site is suitable for both single-stranded and double-stranded DNA molecules. Interestingly, if double-stranded DNA binds to the protein first, the single-stranded DNA is incapable of binding to the inner groove [62, 60, 5]. The possibility to

bind both single and double-stranded DNA molecules is essential for the D-loop formation [94]. Mutagenesis studies revealed that the second DNA binding site is essential for DNA annealing [60]. However, early biochemical studies utilizing distinct RAD52 truncations showed that the first 85 amino acids are able to facilitate the annealing process despite lacking the second DNA binding site and the self-oligomerization domain [82]. The undecameric rings of the N-terminal domain exhibit strong self-oligomerization behavior, they form double- and triple-ring structures [123]. Interestingly, the mutation of Tyr79 responsible for the ring formation resulted in complete inhibition of the DNA binding ability of RAD52 [62, 130, 83]. As for the majority of the single-strand annealing proteins, the C-terminal domain is disordered and responsible for protein-protein interaction. It has RPA and RAD51 recognition domains, as well as the Nuclear Localization Sequence (NLS) [104, 149, 136, 128, 142]. Electron-microscopy (EM) images have depicted the full-length RAD52 as a ring formed by seven monomers, a heptamer, with long protrusions [136]. The protrusions are formed by the highly disordered C-terminus. It was also shown that the protein forms massive clusters in the solution [118]. Further attempts to obtain EM images of RAD52 with higher resolution were not successful [123, 137]. However, recently two groups have independently crystallized the full-length RAD52 [68, 6]. Despite of low resolution and almost complete absence of the C-terminal domain, the core structure of the full-length protein was presented as an undecameric ring.

1.6 Models of single-stranded DNA annealing by RAD52

Over the years of studying RAD52, there have been many attempts to reveal the mechanism by which RAD52 anneals DNA single-strands. The majority of the proposed models have suggested rings as the main annealing entity [60, 120, 43, 123].
The most recent annealing model proposes ring stacks as a key complex promoting the annealing [123]. Once the DSB is formed, rings of RAD52 accumulate on the single-stranded overhangs. Single-stranded DNA is bound to the inner groove of the ring structures. Neighboring rings slide along each other to identify the homology region. Two rings with complementary sequences inside the grooves come in close proximity and stack to establish the HB [123]. However, as the distance from one groove to another is too far to sense the homology, the authors suggested that the protein might undergo a conformational change to release single-stranded DNA from the groove in order to complete the formation of an HB with a complementary strand [123]. This process could be cognate to DdrB or Redβ securing base-pairs [141, 97]. A similar model has also been previously suggested [60, 43]. Moreover, the same groups proposed an alternative mechanism utilizing a single-ring structure as the most prominent annealing species [60, 43, 123]. Interestingly, a crystal structure of the N-terminal domain of RAD52 bound to two single-stranded DNA molecules has been obtained [123]. According to this crystal structure, one DNA

has an extended confirmation inside the groove, while the second DNA is bound to the outer stem and partially retains the helicity. Such positioning of complementary strands on the protein is reminiscent of the mechanism of homology search of RecA [70]. Recently the same mechanism of DNA pairing has been suggested based on the crystal structures of RecT and Redβ [97, 19]. An alternative model utilizing multiple non-stacking rings suggested that the annealing is promoted by overlapping two nucleoprotein filaments carrying complementary strands [120].

The above-described ring models raise several questions. The model proposing the ring stacks was solely based on the N-terminal domain of RAD52 obtained at micromolar concentrations with point mutations abolishing the activity of one of the DNA-binding sites [123]. As previously mentioned, the second DNA-binding site is essential for proper annealing [62]. Therefore, the mutation of the second DNA-binding site could affect the annealing and therefore the proper function and behavior of the protein. The use of such mutants to reveal the annealing mechanism is questionable. The sliding of the rings on the single-stranded DNA could be possible on the N-terminal domain, but not on the full-length protein. The single-molecule studies performed on the fluorescently-tagged full-length RAD52 demonstrated that the protein is statically bound to the single-stranded DNA, while it diffuses on the double-stranded DNA [14]. Similarly, the C-terminal domain of yeast Rad52 has been proposed as the inhibitor for protein diffusion on the DNA [31].

The relevance of the utilized micromolar concentrations with respect to the actual *in vivo* situation has to be thoroughly analyzed [123]. It has been demonstrated in yeast and mammalian cells that after the DNA DSB, the protein forms micro-foci in order to increase the local concentration and bring the broken DNA ends in close proximity [34, 76, 102]. Taking into consideration an intrinsic property of RAD52 to self-oligomerize, one would expect the propensity of RAD52 to form ring structures and stacks thereof inside the foci [34]. However, an estimation of the exact concentration of human RAD52 within one focus is challenging. It has been shown that yeast cells form foci containing 25-90 % (600 - 2100 protein molecules) of all the protein molecules [76]. The cells form only one or two foci per 80 breaks. Thus, there is a certain excess of the protein concentration over the concentration of DNA breaks. Nevertheless, the application of micromolar concentrations appears somewhat extreme with respect to the approximate *in vivo* concentrations. It should be revised carefully, as the excess of the protein inhibits DNA annealing [143, 82].

Another important question is how RAD52 performs homology proofreading. Each monomer of RAD52 binds 4 nt [105]. Following the ring model, two rings with single strands bound can probe only 4 base pairs at once. However, 4 base pairs are not sequence-unique for the mammalian genome [45]. According to bioinformatic studies, the shortest unique sequence for the mammalian genome is 11 base pairs [46]. That implies that three monomers from one ring should approach three monomers from another ring. Considering the circularity of RAD52 undecamer, this even seems rarely challenging. Contrary to the models

utilizing ring structures, an alternative annealing mechanism designating short oligomers as the species promoting the annealing has been proposed [58, 33, 4]. Alternatively, one model suggested that RAD52 monomers forming with an excess of DNA or as a result of RPA-RAD52 interaction, are the key annealing species [58, 30]. It has been demonstrated *in vitro* that upon the excess of single-stranded DNA, the protein primarily exists in a monomeric form, following the originally proposed Redβ annealing mechanism *via* dimers [58, 33, 4]. However, *in vitro* results based on the DNA excess appeared opposite to the *in vivo* protein-DNA ratio [76]. Additionally, anti-cancer research studies utilizing distinct pathways to disturb the rings observed impaired or fully inhibited annealing due to the inability of lower oligomers of RAD52 to bind DNA [22, 49]. With respect to the involvement of RPA in the DNA DSB repair mechanism, it is known to inhibit RAD52 annealing on short single-stranded DNA that mimics DNA overhangs *in vivo* [143]. The exact stoichiometry of the RAD52-RPA complex remains unknown. It has been proposed that yeast Rad52 binds to RPA in the ring form [39]. Therefore, the exact impact of RPA on RAD52 oligomeric state is obscure.

1.7   The aim of this project

There were many attempts to reveal the precise mechanism of DNA annealing by RAD52. Yet, it remains a point of many questions and debates due to some contradictory results obtained by different groups. The major scope of this work was to give another trial, by utilizing a combination of a novel single-molecule technique with conventional biochemical assays, and understand the mechanism of single-stranded DNA annealing by human RAD52. Taking into account the most debated points within the process, I aimed to answer the following questions. How different is the full-length RAD52 from its widely used N-terminal truncated form? The above-mentioned annealing models promoting the rings were based on the N-terminal domain, neglecting the C-terminal and the different number of subunits reported for the full-length ring structure. Secondly, considering the propensity of the protein to self-oligomerize, I inquired whether the protein concentration affects its oligomeric state or/and its annealing activity. How does DNA binding affect the oligomeric state of the protein? And most importantly, what are the oligomeric species responsible for the annealing reaction?

Chapter 2

Materials & Methods

2.1   DNA molecules

Single-stranded DNA molecules were purchased from Sigma-Aldrich and Eurofins in dry form after high-performance liquid chromatography (HPLC) purification (Table S1). Upon arrival, oligonucleotides were spun thoroughly and diluted in TE buffer (Tris, EDTA, pH 7.0) to a final concentration of 100 μM. Oligonucleotides were aliquoted (5 μl) and stored at -80°C. For a short time storage, the utilized aliquotes were kept at 4°C. Each oligonucleotide had a phosphate group at the 5'-end and some had various fluorophores at the 3'-end (Table S1). Each fluorescently labeled oligonucleotide had a so-called "STOP partner" for the termination of the annealing reaction (Section "Terminated annealing assays"). The utilized single-stranded DNA molecules were designed to ensure the absence of self-dimerization as well as to prevent the formation of secondary structure elements (SSE).

2.2   Proteins

All the proteins were purified by Dr. Sivaraman Subramaniam (Dresden, BIOTEC, RG of Prof. Francis Stewart). The proteins were expressed and purified in *E.coli* with a 6x His-tag at the N-terminal domain linked via a thrombin recognition sequence. During the experiments, the full-length RAD52 and the truncated RAD52(209) were utilized at most. However, the full-length RAD52 with a GFP-tag and truncated RAD52(209) chemically labeled with DyLight488 were tested for future optical tweezer experiments. The protein stock concentrations were measured by Nanodrop and confirmed on the SDS PAGE with the Bovine Serum Albumin (BSA, Sigma-Aldrich #P0834) as the mass standard. Proteins

were aliquoted (5 μl) and stored at -80°C.

2.3 Biochemical assays

2.3.1 SDS-PAGE

To confirm the molecular weight of the proteins, Sodium Dodecyl Sulfate Polyacrylamide
Gel Electrophoresis (SDS-PAGE) was performed. Since the protein concentrations were
previously estimated, each sample was diluted to 1 μM in 10 μl volume. Sequentially,
2 μl of the 6x Gel Loading Dye (New England Biolabs, #B7024S) was mixed with β-
Mercaptoethanol (BME) and added in a 1:4 ratio to the sample. Samples were incubated
at 98°C for 7 minutes. Later the samples were centrifuged shortly and transferred on
ice. Samples were loaded on the 10% SDS-PAGE (#4561033) together with 5 μl of the
protein standard (New England Biolabs, P7719S). The SDS-PAGE running buffer (10x)
was purchased from BioRad, #1610744. Gels ran at 200 V for 30-40 minutes, until the
gel loading dye reached the bottom of the gasket. After the run was completed, gels were
washed 3 times, 5 minutes each wash, in MilliQ water and stained with Imperial Protein
Stain (Thermo Fischer Scientific, #24615) overnight. Gels were washed several times in
MiLLiQ until the protein bands were easily resolvable. The destained gels were imaged
on the Epson Perfection V700 Photo scanner. Obtained images were processed in ImageJ.

2.3.2 Blue Native protein PAGE

Blue Native gels were performed at 1, 2, 5, and 10 μM concentrations. Samples were
diluted in the annealing buffer (AB, Tris-Acetate 25 mM, Mg-Acetate 2 mM, DTT 1 mM,
pH = 7.5) or the mass photometry (MP, 20 mM Tris-HCl, 50 mM NaCl, pH 7.5) buffer
with the 2% of n-Dodecyl-β-D-Maltosid (DDM, Thermo Fischer, BN2005) and 0.5% of
Digitionin (Thermo Fischer, BN2006). Additionally, 3 μl of the 4x Native PAGE Sample
Buffer (Thermo Fischer, BN2003) was added. Right before loading, 0.5% of the Sample
Additive (Thermo Fischer, BN2004) was added. The cathode and the anode buffers
were prepared in advance from the concentrated stocks (Native PAGE Running Buffer
(Thermo Fischer, BN2001), Native PAGE Cathode Additive (Thermo Fischer, BN2002))
following the protocol for the Light Blue Native and stored at 4°C. The NativeMark
Unstained Protein Standard was the mass standard (Thermo Fischer, LC0725). Gels ran
on ice at 150 V for 1 hour, and at 250 V for another 2-3 hours, until the loading front
reached the bottom of the gasket. After the run was complete, gels were quickly washed in
MilliQ, fixed in 40% Ethanol and 20% Acetic Acid solution, and stained with Coommasie

overnight. Afterward, gels were thoroughly washed in MilliQ to obtain prominent protein bands. The destained gels were imaged on the Epson Perfection V700 Photo scanner and processed in ImageJ.

2.3.3   Single-stranded DNA binding assays

Single-stranded DNA binding affinity of the proteins was tested with oligonucleotides of various lengths: 32-, 48-, 60-, and 70-nt that had various fluorescent modifications at 3'-end - Atto565, Alexa488, Atto680, and Atto532, respectively. There were two different types of experiments performed.

In the first type of experiment, to quantify the binding cooperativity of the proteins, the oligonucleotide concentration was kept constant at 10 nM, while the protein concentration varied from 10 nM to 1 μM. Proteins were mixed with the single-stranded DNA in the AB in 10 μl reaction volume. The mixture was incubated either at 25°C for 15 minutes or at 37°C for 5 minutes. Before the loading on the gel, 2 μl of the DNA gel loading dye 6x (Thermo Fisher, R0611) were added. Samples were resolved on 10% Tris-Glycine gels (BioRad). Gels ran on ice in a cold TAE buffer at 100 V for 1 hour. Once the run was complete, the gels were imaged on Typhoon 9500FLA (General Electrics) at 650 PMT with 100 μm per pixel resolution. Band intensities were quantified in ImageJ. Obtained values were normalized to the control reaction (no protein) and analyzed in Python. Weighted data points were fitted to the Hill or Michaelis-Menten equations. The resulting fit parameters are present in Table S3.

In the second type of experiment, the protein concentration was 500 nM, while the single-stranded DNA concentration was gradient - 200, 500, and 1000 nM. Proteins were mixed with the DNA and incubated at 37°C for 5 minutes in the annealing buffer or in the mass photometry buffer. The process of the gel loading and run was described above. However, after the imaging on the Typhoon scanner, gels were stained with the Pierce Silver Stain Kit following the standard protocol (Thermo Fisher, 24612) to obtain the signal from the protein. Stained gels were imaged on the Epson Perfection V700 Photo scanner and processed in ImageJ.

2.3.4   Single-stranded DNA annealing

Single-stranded DNA annealing assays were performed to quantify the activity of the proteins. Proteins were tested at different concentrations, at various protein-to-DNA ratios, and the oligonucleotides with different lengths were used as annealing pairs. Moreover, some of the assays were terminated, while the others were mimicking the native conditions.

Terminated annealing assays.

Two types of terminated single-stranded DNA annealing assays were performed. In the first type of assay, the reaction rate was quantified. An examined protein at 100 nM was preincubated with one unlabeled ssDNA (at 10 nM) for 5 minutes at 37°C. Following the incubation, the fluorescently labeled complementary strand (at 10 nM) was added. The reaction was terminated after 1, 3, and 8 minutes at 25°C by adding the STOP buffer (5 μM of the unlabeled DNA (sequence is identical to the fluorescently-labeled one), 1% SDS, 100 mM EDTA, 10 units of Proteinase K (New England Biolabs, #P8107S)) to the corresponding fraction from the total reaction volume in 1:4 ratio. For instance, 8 μl of the reaction were terminated with 2 μl of the STOP buffer. The mixture was incubated for 15 minutes at 37°C. DNA bands were resolved on 10% Tris-Glycine gels (Biorad). The following procedure is identical to the single-stranded DNA binding assay (as described above). There were two ways to analyze the obtained intensities of the bands. First, to normalize the intensities of the single-stranded DNA at different time points to the control reaction (no protein). By quantifying the amount of unreacted ssDNA, one can derive the amount of the single-stranded DNA that participated in the annealing reaction. Second, by normalizing the intensities of the double-stranded DNA bands to the first minute of the control reaction (intrinsic DNA annealing, no protein), one can quantify how many times the protein facilitated the reaction yield (normalized annealing product).

In the second type of the terminated annealing assay, the concentration of the protein that gives the highest amount of the reaction product (double-stranded DNA) was identified. The DNA concentration was 10 nM. The protein activity was analyzed over a broad concentration range, from 10 nM to 10 μM. The terminated annealing assays were performed at various incubation strategies. One strategy consisted of, as described above, preincubation of the protein diluted to the desired concentration with one unlabeled ssDNA at 37°C for 5 minutes. Subsequently, a second fluorescently labeled single-stranded DNA was added, and the reaction was terminated after 15 minutes at 25°C. Following the second incubation strategy, the protein was mixed simultaneously with both single-stranded DNA without prior incubation. As in the previous case, the reaction was terminated after 15 minutes at 25°C. According to the third type of incubation strategy, a protein diluted to a desired concentration was split into two fractions. Each fraction was incubated with one of the oligonucleotides at 37°C for 5 minutes. Later these fractions were mixed and the reaction was terminated as described above. The amount of double-stranded DNA product formed at different protein concentrations and after distinct incubation strategies was normalized to the amount of the double-stranded DNA product formed in the control reaction (intrinsic DNA annealing, no protein) and plotted using a custom-written Python notebook.

Native annealing assays.

To visualize the annealing intermediate species, the annealing assays described above were performed without termination. The proteins were tested at concentrations from 10 nM to 2 μM. The ssDNA concentrations were 10 nM. In this type of experiment, both ssDNA had a distinct fluorescent modification at the 3'-end, e.g. ATTO565 and ALEXA488, or ATTO680 and ALEXA488. All the native annealing assays were performed using the first incubation strategy - the protein was preincubated with one strand with the subsequent addition of the complementary strand. The obtained intensities of the bands were normalized to the ss- and dsDNA control reactions (intrinsic annealing, no protein).

2.4 Mass Photometry experiments

2.4.1 Experimental setup and working principle

The mass photometry experiments were performed on the OneMP system ((REFEYN Inc., Oxford, United Kingdom) in the RG of Dr. John Weir (MPI-FML). The schematic of the setup is demonstrated in Figure 2.1, a (adapted from [161]). The laser beam is focused *via* the wide-field lens (WFL) onto the back focal plane (BFP) of the objective (OBJ). The partial reflector (PR) couples the illumination light directed to and from the OBJ and mitigates the reflected light before it reaches the camera (CMOS) (Figure 2.1, a). The working principle of the system is based on the fact that molecules (proteins, nucleic acids, lipids) scatter light proportionally to their masses (Figure 2.1, b) [24, 161]. Due to the difference (ratio) between the scattered and the reflected light, the binding events of the molecules can be visualized (ratiometric contrast) (Figure 2.1, c). Native (raw) images are shown in Figure 2.1, d. To obtain the ratiometric contrast image capturing the actual binding event, the software (AcquireMP) divides the averages from several frames before (x_1) and after (x_2) the binding event (Figure 2.1, d-e). By utilizing the averaging method during the entire experiment, multiple binding events can be observed (Figure 2.1, c). The intensity of the grey circles is proportional to the molecular mass of the corresponding molecules (Figure 2.1, b and c, marked with blue circles). It is important to mention that several overlapping events or events that appeared at the borders of the imaging field will be removed from the following so-called fitting procedure. Therefore, the number of fitted events is always lower than the total number of events bound to the glass surface.
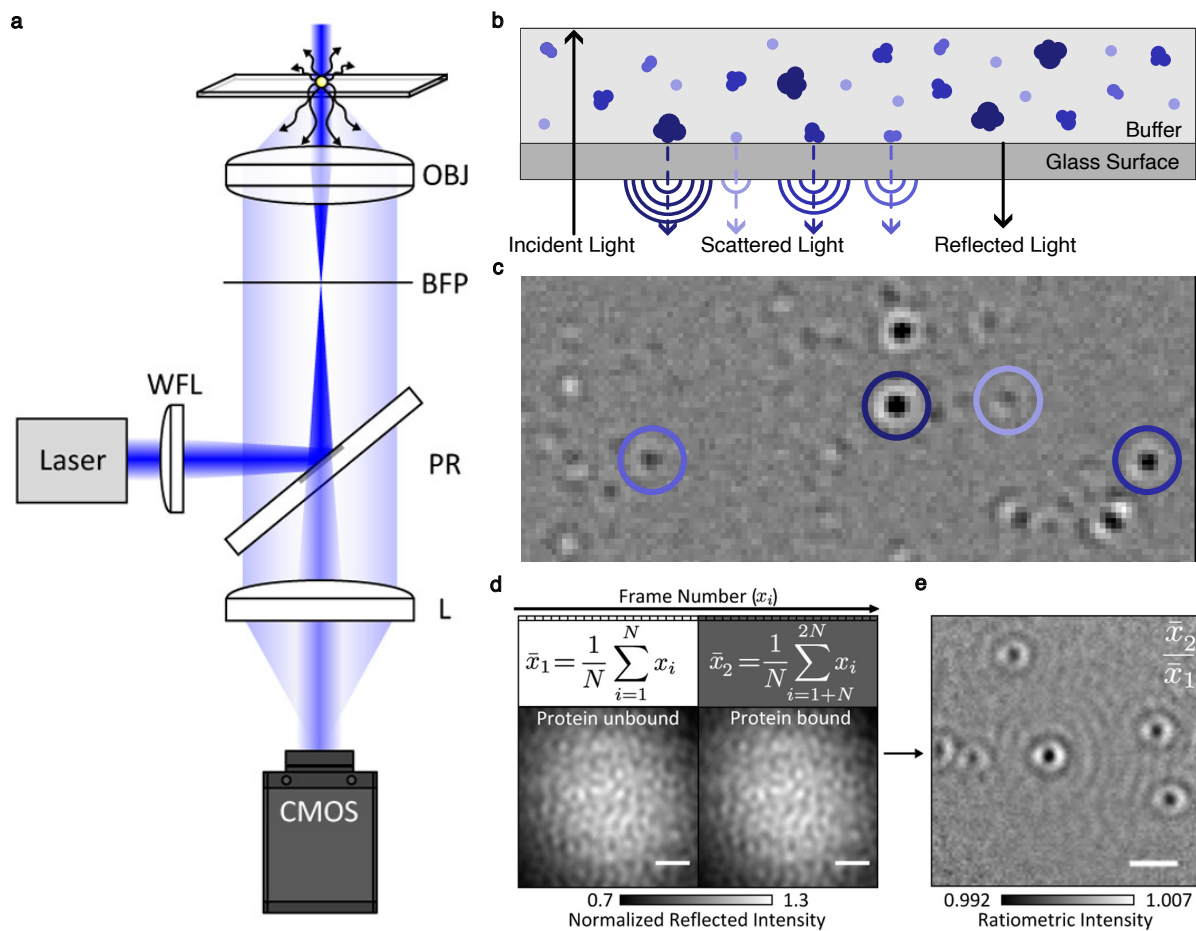
Figure 2.1: Schematic of the mass photometry setup and the data acquisition. **(a)** The schematic representation of the OneMP setup. The wide-field lens (WFL) focuses the laser beam into the back focal plane (BFP) of the oil immersion objective (OBJ). The illumination light in both directions is being combined *via* the partial reflector (PR). PR also mitigates the reflected light by the additional lens (L) before it reaches the camera (CMOS). The schematic has been adapted from [24]. **(b)** Schematic representation of the mass photometry principle. Proteins scatter the incident light proportionally to their masses. **(c)** The difference (ratio) between the reflected light and the scattered light appeared as grey circles with different intensities proportionally to the masses of the protein mixture. **(d)** The principle of ratiometric imaging is based on dividing averages from a certain number of frames before (x_1) and after (x_2) the binding event took place. The image has been adapted from [24].

2.4.2 Experimental accessories

Mass photometry experiments were executed in the annealing buffer utilized for the biochemical assays as well as in the mass photometry buffer. Before measurements, buffers were filtered two times with 200 nm filters (Sigma-Aldrich, #SLGS033) and degassed for 30 min in the desiccator while stirring. Between the measurements, the buffers were stored at 4°C.

To obtain the easy-cleaned slides, coverslips (25 mm x 50 mm, Marienfeld Superior, #0117520) were located in a slide holder (Wash-N-Dry, Carl Roth) and sonicated for 5 min in MilliQ water, Isopropanol-2, and in MilliQ again. The slides were dried under the hood with filtered air (200 nM filter unit). For storage or transportation, the slides were kept in the rack inside the glass cylinder and sealed with parafilm.
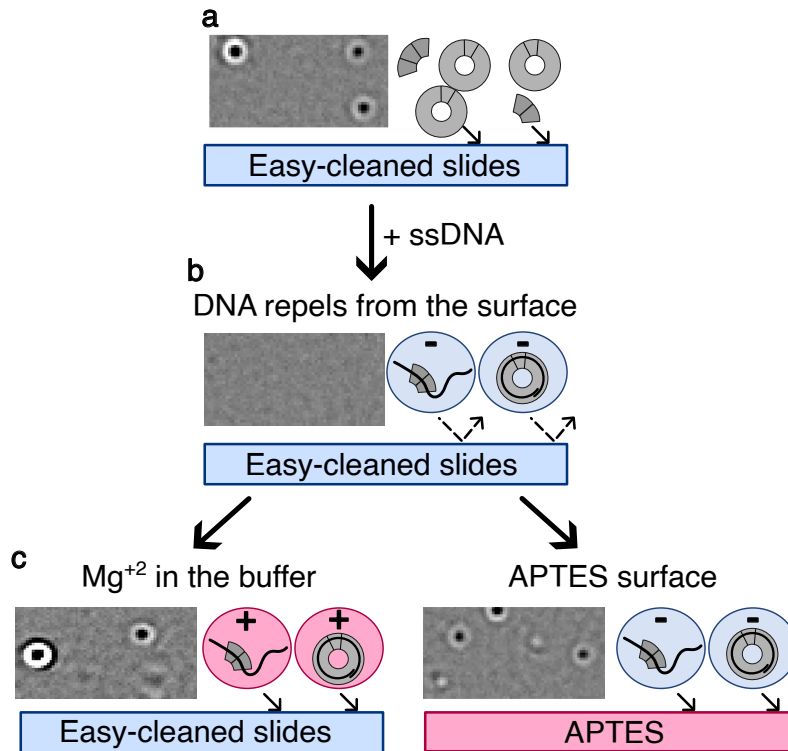
Figure 2.2: Schematic representation of protein binding to the glass in mass photometry. **(a)** Easy-cleaned glass slides might retain some negatively charged -OH groups, therefore, positively charged proteins bind freely. **(b)** Upon binding to the ssDNA, the protein-DNA complex repels from the surface due to the negatively charged DNA. **(c)** The latter, could be conquered by utilizing a buffer with Mg^{+2} or modified positively-charged APTES slides, or both.

To obtain a positively charged glass surface, coverslips were treated with (3-Aminopropyl) triethoxysilan (APTES, Sigma-Aldrich, #440140) [74]. First, coverslips were sonicated in MilliQ water, Isopropanol-2, and in MilliQ, for 5 minutes each wash. Later, the slides were submerged in HCl and sonicated at 65° for 90 minutes. After the coverslips cooled down, they were submerged in acetone to remove the water excess, then into the 2% APTES diluted in acetone for 1 minute, and, finally, the excess of APTES was washed out in fresh acetone. Immediately after that, coverslips were transferred to an oven and baked at 110° for 2 hours. Subsequently, the slides were cooled down and easy-cleaned again. Finally, the slides were dried and sealed with parafilm.

CultureWell gaskets (Sigma-Aldrich #GBL103250) were utilized as a reaction chamber (Figure 2.3, a). Gaskets were kept in a falcon tube in pure Isopropanol-2. Before the

experiment, they were washed two times in MilliQ water upon shaking for 30 seconds. After the wash, gaskets were dried on Kimtech paper under the hood and sealed in a bag. If properly stored, gaskets can be reused a few times.

2.4.3 Experimental procedure

Two fields of view were used to acquire the data - the regular field of view (RFV, 10.8 μm $\times$ 2.9 μm) and the medium one (MFV, 10.8 μm $\times$ 6.8 μm) (Figure 2.3, b).
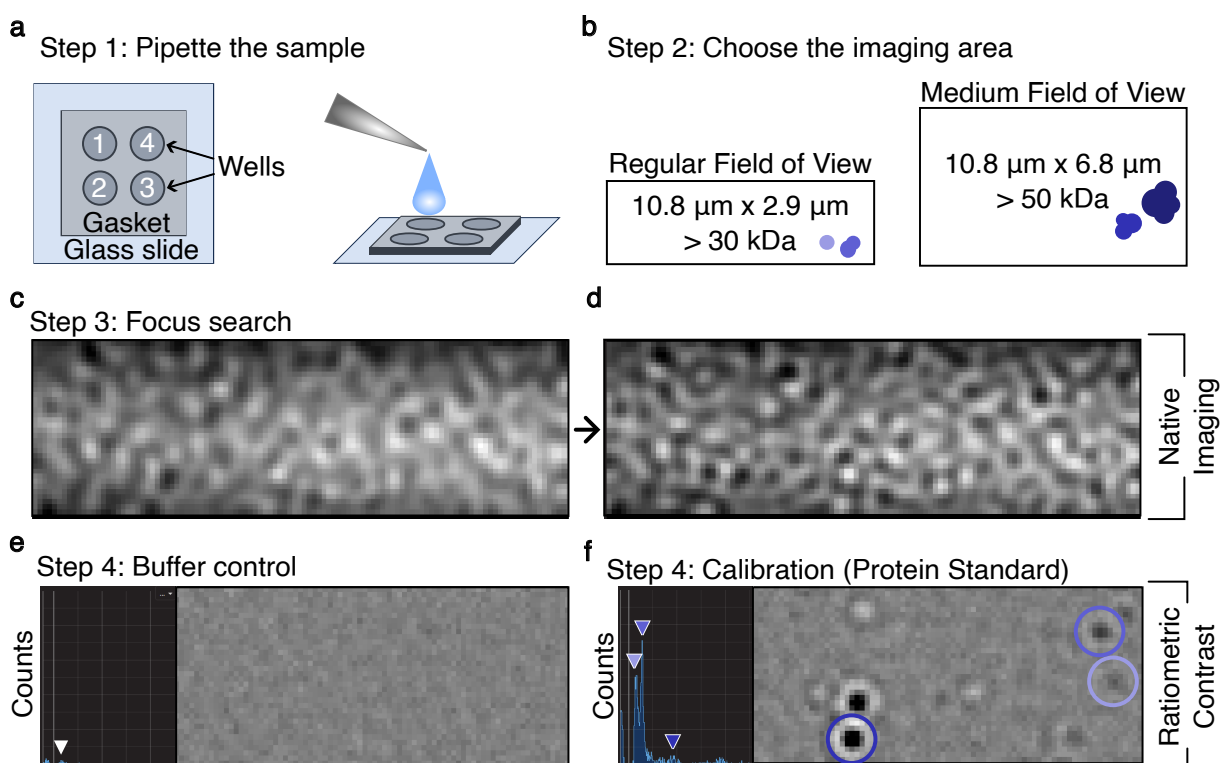
Figure 2.3: Focus search and the mass standard measurement using mass photometry. **(a)** Schematic representation of a gasket with four wells on top of the glass slide. **(a)** Regular field of view (RFV) is utilized to detect small proteins (around 30 kDa), while the medium field of view (MFV) cannot detect samples below 50 kDa. Exemplary native images of the glass acquired in RFV out of focus **(c)** and in focus **(d)**. Ratiometric contrast images (RFV) of the mass photometry buffer **(e)** and the protein standard **(f)**. The filtered buffer should ideally produce a low number of counts (marked with a white arrow). The protein standard contains proteins with different masses that appear as distinct peaks. Blue circles mark the binding events, and the corresponding peaks are labeled with blue arrows.

RFV has a lower resolution limit of 30 kDa and is better suited for the detection of small proteins, while MFV cannot detect species below 50 kDa, but it acquires more

binding events and, therefore, provides better statistics. A silicone gasket was placed on top of the coverslip, and positioned on top of the oil-emerged objective. To focus on the surface, a 5 μl of the buffer was pipetted into the well (Figure 2.3, a). The focus was identified based on the Sharpness monitor in the AcquireMP software (REFEYN). To focus on the glass surface, the Sharpness values should be between 5 and 7, additionally, the native image of the glass surface should not contain any impurities (Figure 2.3, c-d). Once the focus was found, it was saved and locked. Afterward, 5 μl of the sample mixture was added to the gasket, the focus was quickly unlocked to ensure the correct positioning and locked again. The native imaging was switched to the ratiometric and a movie with protein binding events was recorded (Figure 2.3, e-f). Acquired movies were analyzed using the DiscoverMP (REFEYN), resulting files were saved in .h5 format and further analysis was performed using the custom-written Python script [57].

To avoid buffer contamination, it was measured on each experimental day before the main measurements (Figure 2.3, e). NativeMark Unstained Protein Standard was used as a protein standard diluted in the buffer in a 1:400 ratio (Figure 2.3, f). Samples were diluted in 50 μl reaction volume. From this volume, several fractions were taken and measured, further referred to as readings. During the analysis, readings from the same reaction mixture were combined to obtain a cumulative distribution from the same day.

2.4.4 Data analysis

Mass photometry data was analyzed using a custom-written package in Python [57]. The acquired contrast values of the protein standard were fitted to a set of Gaussian functions to obtain a better fit for all the peaks. Resulting Gaussian mean values were plotted against the known molecular masses of the protein standard (Figure 2.4, a). The data was further fitted using the linear regression function implemented in Python (Figure 2.4, b). The obtained slope and intercept values (calibration parameters) were applied to convert the ratiometric values into the masses for the proteins of interest. A similar fitting procedure utilizing a set of Gaussian functions was performed for the proteins of interest by providing the initial guess (molecular mass, Gaussian mean value) for the distinct peaks corresponding to the protein protomers (Figure 2.4, c). The best fit was chosen based on diminishing Akaike criteria. Resulting fit parameters including sigma, areas under the Gaussian, and mean values were further used for the statistical analysis. The number of counts was calculated as the number of binding events in each reading (each experiment). Peaks with the Gaussian areas smaller than 1% were discarded from further analysis.
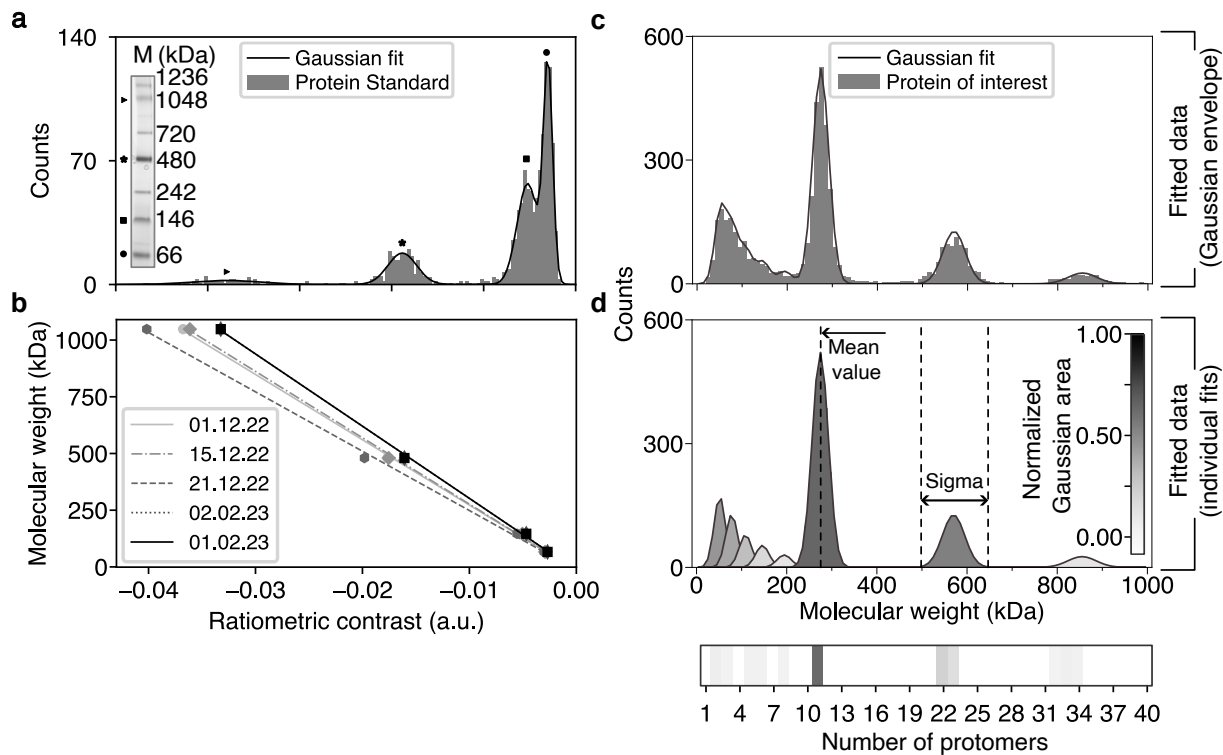
Figure 2.4: Calibration procedure and data analysis of acquired MP data. **(a)** An exemplary mass histogram (grey) of the protein standard (Native Mark) in ratiometric values fitted to a set of Gaussian functions corresponding to each peak and presented as the fitted envelope (solid black line). Masses of the protein standard corresponding to the peaks and labeled with symbols are shown in the insert. **(b)** Linear fit of the ratiometric contrast values (Gaussian peaks) to their corresponding masses. The resulting slope and intercept are utilized for further analysis. **(c)** An exemplary mass histogram (grey) of the protein of interest fitted to a set of Gaussian functions and presented as the envelope (solid black line). The total area under the Gaussian envelope is 1. **(d)** Individual Gaussian functions corresponding to different protein species are color-coded according to their fraction within the total area under the Gaussian envelope. Darker colors reflect the protein species with the biggest area under the Gaussian. Below an exemplary heatmap is presented showing color-coded areas corresponding to different types of protein species.

Chapter 3

Results

The biochemical assays were carried out simultaneously with the mass photometry experiments in order to combine and complement the obtained results. Therefore, all the results will be presented similarly as a mixture of different approaches following the same questions.

3.1 Overview of the protein variants

All the experiments were performed on both full-length RAD52 and an N-terminal truncation hereby referred to as RAD52(209) (Figure 3.1). For future optical tweezers experiments, two fluorescently-labeled variants were tested as well - the full-length RAD52 tagged with Green Fluorescent Protein (GFP) and RAD52(209) that was chemically labeled with DyLight488 (Figure S1). All proteins have a 6x-His-tag at the N-terminus linked to the main sequence *via* a thrombin recognition sequence (Figure 3.1, a).

The full-length RAD52 is 439 amino acids (aa) in length, including the His-tag and the linker sequence. The sequence-predicted molecular weight of a monomer is 48.46 kDa (Table S2). RAD52 carries several domains essential for its functionality. The self-oligomerization domain (residues 85-59), labeled in Figure 3.1, a, as RAD52, an RPA recognition domain (residues 221-280), RAD51 recognition domain (residues 291-330), and Nuclear Localization Sequence (NLS, residues 405-414). RAD52(209) is constituted of 230 aa, with a monomer size of 25.35 kDa. RAD52(209) carries the self-oligomerization domain and two DNA-binding sites.

For the optical tweezers combined with Total Internal Reflection Fluorescence microscopy experiments, GFP-RAD52 has been designed and purified. It is a 678 aa protein, which results in a monomer size of 75.51 kDa (Figure S1, a). To obtain fluorescently labeled RAD52(209), a fraction of the protein was chemically labeled with DyLight488. Dy-

Light488 is known to bind lysine residues *via* their primary amine groups [35]. According to its primary structure, RAD52(209) has 13 lysine residues. Most importantly, four of these residues are located at the two DNA-binding sites (Figure S1, a, yellow arrows). Moreover, carrying an overall negative charge, DyLight488 could be a potential obstacle for the DNA binding to the protein. We have not assessed where exactly DyLight488 was bound. Nevertheless, the biochemical analysis of this variant and its comparison to the unlabeled RAD52(209) was a point of interest.
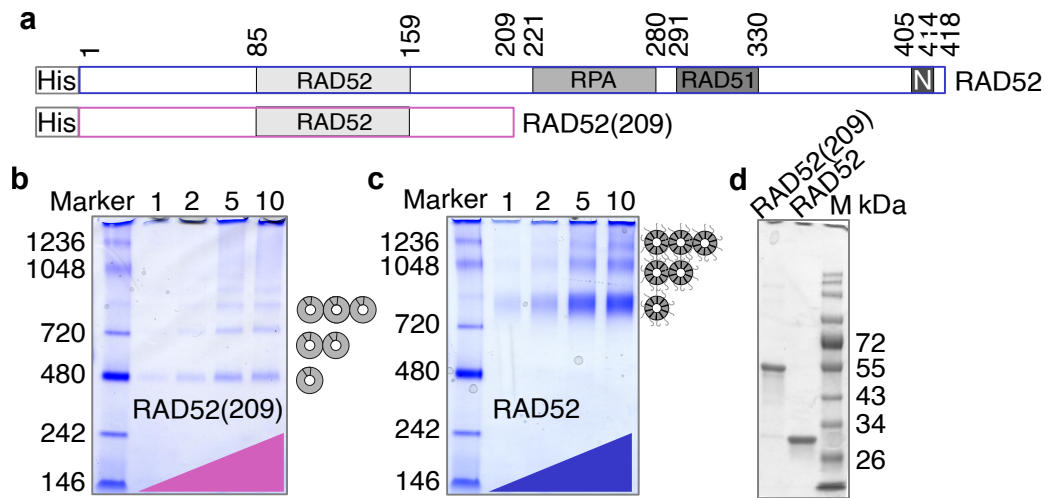
Figure 3.1: An overview of RAD52 and RAD52(209) proteins. **(a)** Schematics of the protein sequences with the identified domains and tags. **(b)** An exemplary Blue Native PAGE of RAD52(209) and RAD52 **(c)** at 1, 2, 5, and 10 μM. NativeMark was used as the protein standard. **(d)** An exemplary SDS-PAGE of the proteins with the protein marker.

3.2   The oligomeric state of the proteins

### 3.2.1   The biochemical approach

Sequence-predicted molecular weights of the proteins were verified by SDS-PAGE (Figure 3.1, d). However, all the proteins ran slightly above their expected masses. Notably, RAD52 and GFP-RAD52 displayed a smear on the gels (Figure 3.1, d and Figure S1, d). To obtain a better result, SDS-PAGE with RAD52 was repeated, and β-mercaptoethanol (βME) was added to the sample buffer to ensure complete denaturation. Only in the presence of βME, it was possible to obtain a sharp band of the full-length protein, indicating the presence of high-order complexes which required extensive denaturation. The oligomerization of the proteins at μM concentrations was visualized with Native Blue PAGE (Figure 3.1, b-c and Figure S1, b-c). All the proteins ran above their predicted

molecular weights, as had been previously observed [53]. Upon increasing protein concentrations, the formation of higher-order oligomers was observed with all the protein variants. Interestingly, unlabeled and labeled RAD52(209) displayed no differences in terms of oligomerization (Figure 3.1, b and Figure S1, b). Notably, the full-length RAD52 exhibited a different oligomerization pattern in comparison to its truncated versions (Figure 3.1, b and c). The major band that would correspond to a single ring structure appeared much thicker than the bands that would correspond to stacks of multiple rings. This pattern was even more pronounced for GFP-RAD52, which showed a smear at the predicted ring size and only a weak band corresponding to potential stacks of rings. It is important to mention that at such concentrations no trace of smaller-order oligomers, including monomers, was detected.

3.2.2 Mass photometry data

Oligomeric states of the protein variants were additionally examined with mass photometry. One of the key advantages of this technique is its ability to detect species at concentrations as low as 1 nM. However, to obtain a significant number of counts necessary for a correct fitting of the Gaussian function corresponding to each supra-molecular species, samples were measured over a concentration range from 10 to 200 nM of protein (Figure 3.2). Above 200 nM, the regular field of view (RFV, 10.8 μm $\times$ 2.9 μm, the lower limit is 30 kDa) appeared too crowded, and individual peaks could not be resolved properly. Therefore, RAD52(209) was measured at 500 nM in the medium field of view (MFV, 10.8 μm $\times$ 6.8 μm, the lower limit is 50 kDa), however, at the cost of not detecting lower oligomeric species (Figure 3.12). Figure 3.2 presents a cumulative sum of the protein distributions from different days. The individual data sets from different days are shown in Figures S2 and S3.

Within the inspected concentration range, RAD52(209) formed ring structures with a molecular weight of 279.6 $\pm$ 1.3 kDa (Figure 3.2, a). Based on the monomer size (25.35 kDa) and expected ring structure formed by 11 monomers, RAD52(209) ring size is expected to be 278.85 kDa. Therefore, the mass photometry results of RAD52(209) were in good agreement with the literature. Besides single rings, the protein also formed double and triple rings, which will be further referred to as stacks of rings. The number of rings and stacks thereof increased proportionally to the protein concentration. RAD52(209) also formed dimers and trimers, and their number also increased linearly with the concentration. Due to the lower detection limit of the system (30-40 kDa), monomers were only partially detected at 10 nM as a shoulder of the Gaussian distribution. From 10 to 20 nM, the number of dimers and trimers dropped with the following increase at higher protein concentrations (Figure 3.3, c).
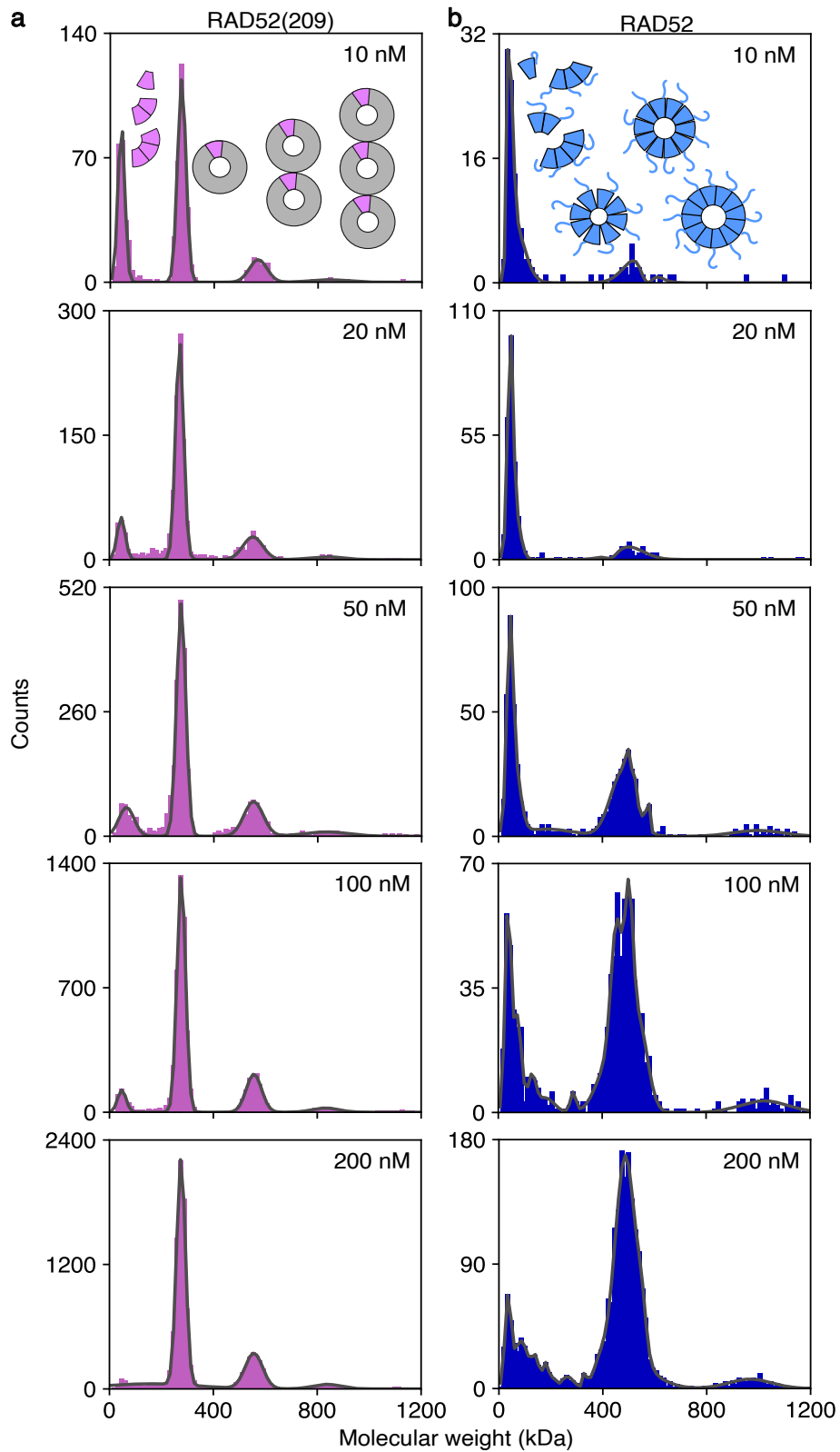
Figure 3.2: The oligomeric state of the proteins. Mass photometry data of RAD52(209) **(a)** and RAD52 **(b)** at 10, 20, 50, 100, and 200 nM. Each histogram is a cumulative distribution of several repeats (N=4). The Gaussian envelope is shown as a solid black line. The individual distributions from each repeat are demonstrated in Figures S2 and S3.

The oligomeric state of RAD52 appeared to be very distinct not only in comparison to RAD52(209) but also to previously published results [136]. RAD52 has been reported to form ring structures consisting of 7 monomers (a heptamer). Therefore, a ring peak would be expected at around 340 kDa. Unexpectedly, RAD52 formed structures with molecular weights varying from 300 to 600 kDa (Figure 3.2, b). Based on the sigma values of the mass standard and RAD52(209) (Figure S4), such a broad distribution indicated the presence of distinct structures, which would correspond to rings with different numbers of subunits - from 7 to 12 subunits. Moreover, RAD52 did not show as many ring stacks as RAD52(209). Additionally, a significant number of lower-order oligomers and monomers were present. With increasing protein concentrations, the number of monomers and dimers decreased, while the number of rings increased drastically (Figure 3.3, d). Due to the appearance of trimers and tetramers, the overall number of lower-order oligomers did not change significantly at different concentrations.

The above-described mass photometry readings of RAD52 and RAD52(209) from different days at different concentrations were plotted together and sorted according to the number of protomers (subunits) present in each type of protein species (Figure 3.3). As previously mentioned, within the examined concentration range, the major peak of RAD52(209) belonged to a single-ring structure (Figure 3.3, a). At low concentrations, the protein formed a small fraction of lower-order oligomers that initially decreased and remained unchanged upon increasing protein concentrations. On the contrary, the fraction of single rings increased proportionally to the protein concentration (Figure 3.3, c). RAD52 formed a broad variety of species from monomers and lower-order oligomers to rings with different numbers of subunits (Figure 3.3, b and d). While the protein concentration increased, the number of lower-order oligomers slightly decreased and the amount of ring structures increased following the concentration (Figure 3.3, d).

Labeled proteins, GFP-RAD52 and RAD52(209)-DyLight488, were not intended to be used for the label-free mass photometry experiments, yet it was interesting to reveal potential oligomeric alterations incurred by their fluorescent tags (Figure S5, a, upper panel). RAD52(209)-DyLight488's oligomeric state appeared to be very similar to its unlabeled version with the ring peak value corresponding to an undecamer (284 kDa). Similarly to RAD52, the GFP-RAD52 formed a variety of species up to 1 MDa (Figure S5, a, upper panel). However, it was challenging to distinguish the border between the short oligomers and rings. Assuming that the smallest ring could be formed by 7 monomers, a molecular weight of 528.57 kDa would be expected. However, there was no significant differentiation, as in the case of RAD52, observed. Additionally, there was a peak at around 50 kDa, which could correspond to a dimer of GFP cleaved during the purification procedure.

The data obtained with mass photometry confirmed the oligomeric state of all the proteins, yet complemented the biochemical data with essential information regarding the presence of short oligomeric species at nanomolar concentrations. Next, I investigated how such different oligomeric states affect DNA binding and annealing.

31
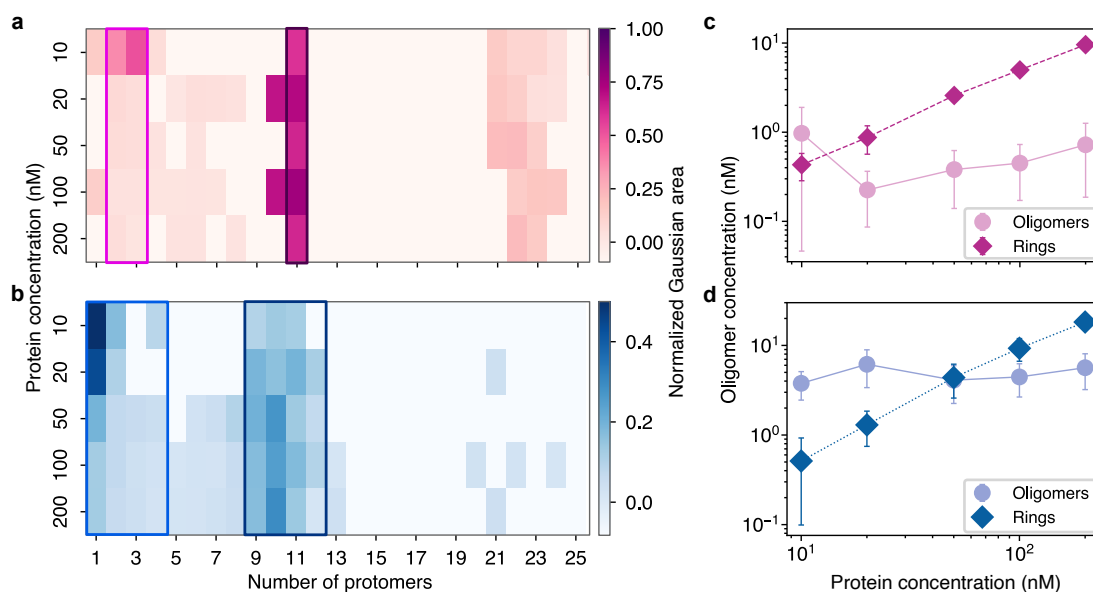
Figure 3.3: Heatmaps demonstrating the oligomeric states of RAD52(209) and RAD52. The heatmaps of RAD52(209) **(a)** and RAD52 **(b)** contain the normalized areas under the Gaussian envelope (described in "Materials & Methods") plotted versus the number of protomers in the corresponding protein species. Each heatmap is the cumulative distribution of several readings combined from different days. The protein concentrations were 10, 20, 50, 100, and 200 nM. The number of short oligomers (dimers and trimers of RAD52(209) (marked in magenta), and from monomers to tetramers for RAD52 (light blue)), as well as rings (dark magenta for RAD52(209) and dark blue for RAD52), is plotted as the function of the protein concentration **(c-d)**.
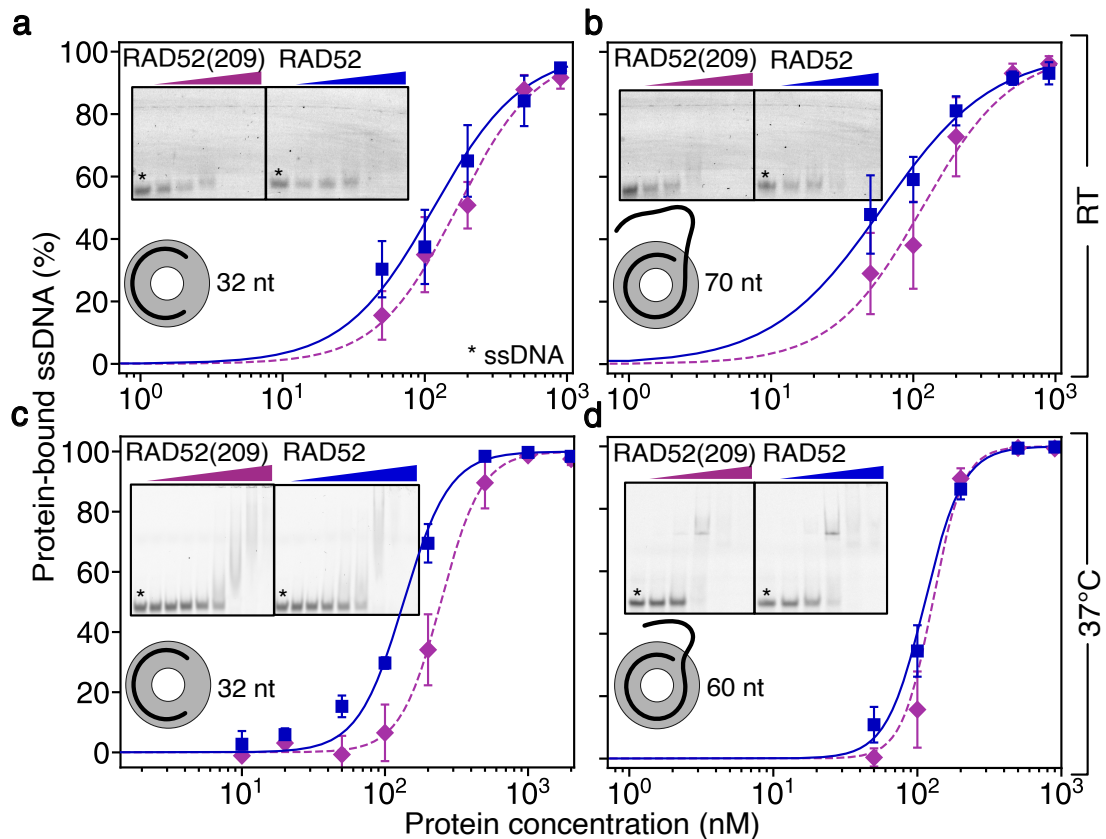
3.3 Single-stranded DNA binding experiments

3.3.1 Electrophoretic Mobility Shift Assays

The protein binding affinity was tested with 32-, 60-, and 70-oligonucleotide-long (nt) single-stranded DNA (ssDNA) (Figure 3.4). The concentration of the ssDNA was 10 nM, while proteins were tested over a broad concentration range (from 10 nM to 10 μM). The oligonucleotide's lengths were chosen to be either shorter (32-nt) or much longer (60- and 70-nt) than the circumference of a single ring (44 nt for RAD52(209), and 28-48 nt for RAD52) in order to estimate whether the DNA-protein interaction would depend on DNA length. The binding affinities were tested at 25°C and 37°C.

At 25°C both proteins exhibited a cooperative binding to 32-nt DNA, with a Hill coefficient of 1.5 ± 0.2 and 1.4 ± 0.2 for RAD52(209) and RAD52, respectively. Longer ssDNA (70-nt) did not have a strong effect on RAD52(209) binding affinity, with a Hill coefficient of 1.6 ± 0.3. RAD52 displayed no cooperative binding, with a Hill coefficient below 1, therefore only the Michaelis-Menten constant is listed (Table S3). Higher temperatures

significantly increased the binding affinities of both proteins. Moreover, the cooperativity was stronger with longer ssDNA. The Hill coefficient for RAD52(209) was 3.1 ± 0.5 for 32-nt and 4.1 ± 0.4 for 70-nt. For RAD52, these coefficients were 2.7 ± 0.2 and 3.3 ± 0.2, respectively. In conclusion, both proteins demonstrated a cooperative binding behavior to single-stranded DNA. The affinity appeared to be stronger at physiological incubation temperatures and with longer DNA molecules.

Figure 3.4: Single-stranded DNA binding affinities. The binding was examined at 25°C with 32- **(a)** and 70-nt oligonucleotides **(b)**, and at 37°C with 32- **(c)** and 60-nt oligonucleotides **(d)**. Each plot contains normalized mean values of the band intensities with SEM. Exemplary gels are shown in the inserts. The control unbound single-stranded DNA is labeled with an asterisk. The data acquired from different repeats was averaged and fitted to the Hill and Michaelis-Menten equations. The fits are presented as solid and dashed lines for RAD52 and RAD52(209), respectively. The fit parameters are listed in Table S3.

RAD52(209)-DyLight488 and GFP-RAD52 single-stranded DNA binding affinities were tested with 32-nt single-stranded DNA at 6.25 nM and at 25° (Figure S6). With these experimental conditions, GFP-RAD52 demonstrated a similar affinity to ssDNA as RAD52, with a Hill coefficient of 1.13 ± 0.3, which would correspond to either a near non-cooperative or a very weak binding affinity. However, the DyLight488 version of RAD52(209) surprisingly exhibited a very strong non-cooperative behavior with

a Michaelis-Menten constant of 33.576 ± 4.202 M. RAD52(209)-DyLight488 binds ssDNA much stronger than the rest of the RAD52 variants, even at 25°.
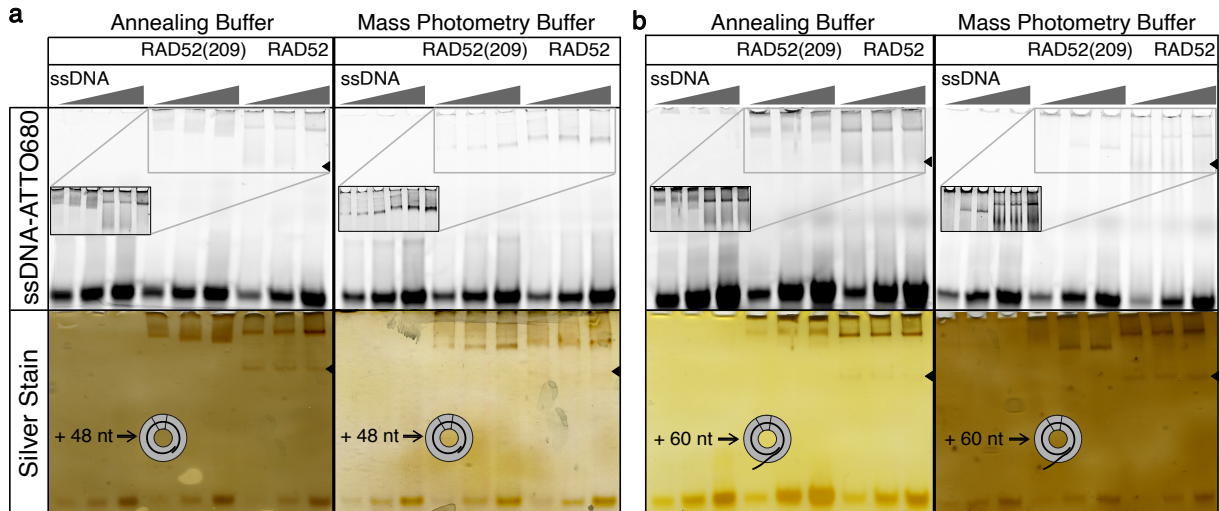
Figure 3.5: Single-stranded DNA binding experiments. RAD52 and RAD52(209) at 500 nM were incubated with 200, 500, and 1000 nM (indicated as a grey triangle) of single-stranded DNA of 48- **(a)** and 60-nt **(b)** in the annealing and the mass photometry buffers. The upper panel contains the fluorescent signal from the DNA molecules, and on the lower panel, the same gels after the Silver Staining procedure are displayed. Due to the DNA's high concentrations, its excess appeared on the gels as well (yellow bands at the bottom). The inserts contain oversaturated images of the wells. Black arrows indicate short oligomers of RAD52 and the corresponding signal from the single-stranded DNA.

3.3.2 Revealing the protein-DNA complexes with Silver Stain

To visualize potential protein binding intermediates, proteins were incubated with 48- or 60-nt long fluorescently-labeled ssDNA at concentrations of 200, 500, and 1000 nM (Figure 3.5, a-b). The protein concentration was fixed at 500 nM in order to obtain visible bands with Silver staining. Proteins were incubated with the DNA in both the annealing buffer and the mass photometry buffer and no significant dependence on the buffer conditions was identified. Additionally, 48- and 60-nt oligonucleotides had no detectable effect on the abundance of protein species. However, RAD52(209) and RAD52 exhibited certain differences in their binding patterns. For RAD52(209) no lower-order oligomers were observed, only high molecular weight species, presumably single rings, stably binding DNA. While DNA concentration was increasing, the major protein band was getting broader and diffusing lower on the gel. RAD52, on the contrary, demonstrated a visible band running lower than the single-ring species, presumably corresponding to a monomer or a dimer (black arrow). The intensity of this band did not change upon increasing DNA concentrations, suggesting that there is a certain threshold in the number of DNA

molecules that these species can accumulate. However, no sharp DNA band corresponding to this protein band was detected, only a smear. On some of the gels, weak bands that could represent other protomers (trimers and tetramers) were visible. Similarly to RAD52(209), RAD52 demonstrated another strong band potentially corresponding to the size of a single ring that exhibited an increased intensity proportional to the DNA concentration. However, while the RAD52(209) single-ring band appeared as a smear, the RAD52 single-ring band was very sharp. It is necessary to mention that even after 5 minutes of the incubation at 37°C there was a significant amount of free single-stranded DNA that migrated through the gel (yellow bans on the bottom). In conclusion, RAD52 and RAD52(209) demonstrated different oligomeric states upon ssDNA binding. Both proteins exhibited single-ring structures with bound DNA molecules. Additionally, while lower-order oligomers of RAD52(209) were not detected, RAD52 demonstrated a broad distribution with the strongest band corresponding to the size of a monomer or dimer.
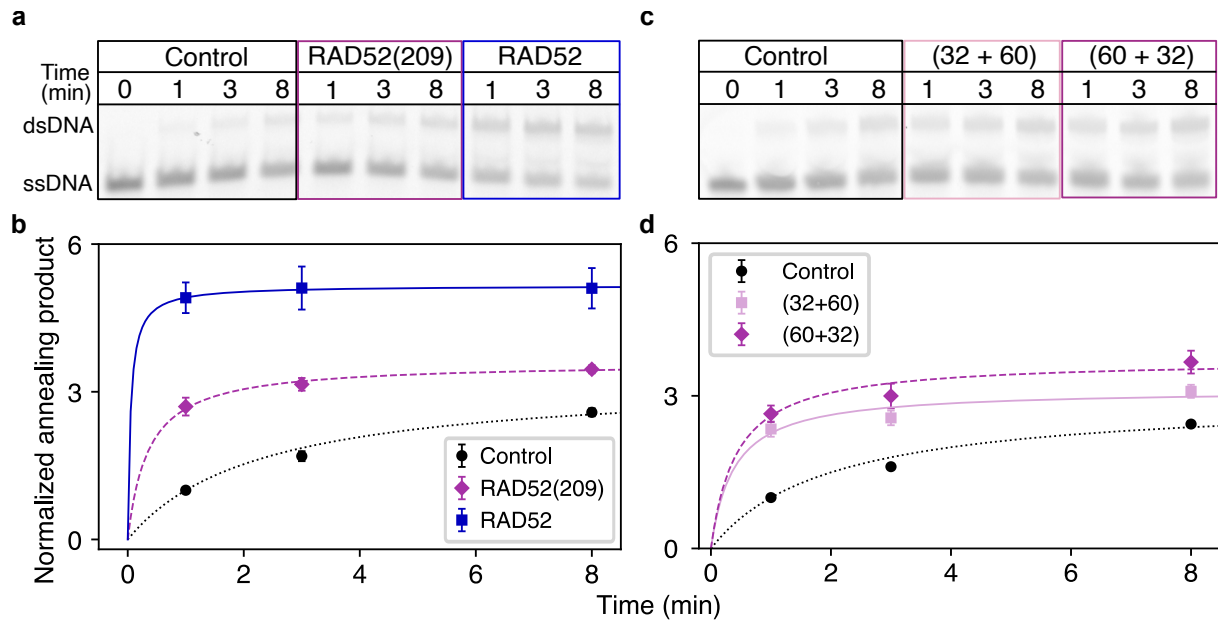
3.4 Single-stranded DNA annealing

3.4.1 Rates of the annealing reactions

Single-stranded DNA annealing reactions were performed to quantify the functionality of the proteins. The corresponding annealing rates were determined at 100 nM of the protein. One pair of oligonucleotides was utilized (32- and 60-nt) at 10 nM concentration. The annealing rates were examined in both annealing and mass photometry buffers (Figure 3.6). All the protein constructs accelerated the reaction in comparison to the intrinsic annealing of the oligonucleotides. The data for RAD52(209) and RAD52 was fitted to the second-order reaction kinetics, and the resulting rates are listed in Table S4. According to the fitted curves, RAD52 accelerates the annealing reaction by 72 times in the annealing buffer, and by 80 times in the mass photometry buffer. RAD52(209) exhibited a modest increase in the reaction rate as well, by 7 and 4 times for the annealing and the mass photometry buffer, respectively. Remarkably, the annealing rates of RAD52(209) are drastically lower in comparison with the full-length protein. Interestingly, the annealing rate of RAD52(209) did not change if it was incubated first with the short (32-nt) or the long (60-nt) DNA strands, indicating that the DNA length does not affect the reaction rate (Figure 3.6, c-d).

RAD52(209)-DyLight488 and GFP-RAD52 showed facilitation of the annealing reaction similar to their unlabeled variants (Figure S7, a). Moreover, at the conditions tested, GFP-RAD52 almost fully annealed all the single-stranded DNA molecules within the first minute. RAD52(209)-DyLight488 did not exhibit such a striking effect, however, the reaction was still occurring faster than the intrinsic annealing (*i.e.* without the presence

of any protein). Thus, the presence of the fluorescent modifications had little to no effect on the annealing rates of the protein variant.

Figure 3.6: Rates of the single-stranded DNA annealing reactions. The rates were tested with and without RAD52 and RAD52(209) **(a-b)**. The annealing pair was 32-ATTO680 and 60-nt at 10 nM. The protein concentration was 100 nM. The reaction fractions were terminated after 1, 3, and 8 minutes. The reaction rates were fitted to the second-order reaction kinetics, obtained parameters are listed in Table S4. The reaction rates were tested with RAD52(209) at 100 nM, upon initial preincubated either with 32-nt (pink squares) or 60-nt (magenta diamonds) **(c-d)**.

3.4.2   DNA annealing at different protein concentrations

To investigate whether annealing rates are dependent on the protein concentration, the annealing activities of both proteins were examined over a broad range of protein concentrations, from 10 nM to 10 μM (Figure 3.7). Different incubation strategies were tested, and as for Redβ, the incubation order appeared to be very important for the annealing capacity of the protein [97]. For the first incubation strategy, the protein was mixed with both single-stranded DNA molecules (32- and 60-nt at 10 nM) at once (Figure 3.7, b-e). Both proteins facilitated the formation of dsDNA products only slightly above the level of intrinsic annealing (Figure 3.7, c-d). The amount of unannealed single-stranded DNA was above 40% for RAD52 and above 60% for RAD52(209), implying that almost half of the single-stranded DNA did not participate in the reaction. In the second incubation strategy, the protein was pre-incubated with one of the strands, and the complementary strand was added afterward (Figure 3.7, f-i).
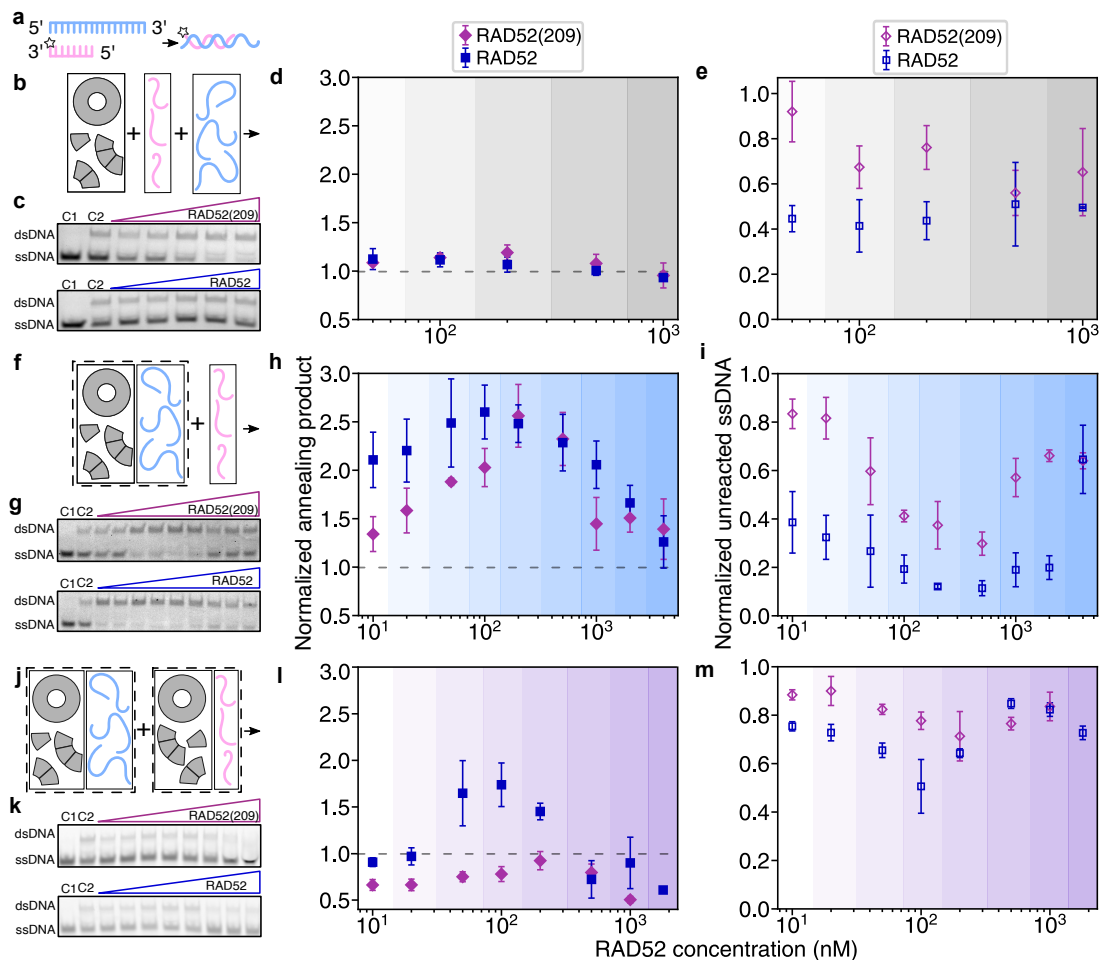
Figure 3.7: Incubation strategies for the single-stranded DNA annealing reactions. The left column contains the schematics of the incubation procedures with the exemplary gels. The middle column contains normalized values of the dsDNA product intensities plotted as the function of the protein concentration. The right column contains the normalized values of the unreacted single-stranded DNA intensities as the function of the protein concentration. **(a)** 32- and 60-nt complementary single-stranded DNA molecules were utilized as the annealing pair. The first strategy was a simultaneous mix of all the reactants - the protein and both strands **(b-e)**. The second strategy implied a pre-incubation of one strand with the protein and a consequent addition of the second strand **(f-i)**. During the third strategy, the protein fractions were pre-incubated with one of the complementary single-stranded DNA **(j-m)** and mixed. The data is plotted as mean values with SEM from several repeats. The first strategy was repeated 5 times, the second one - 6, and the third one - 4 times.

It resulted in a significant increase in the quantity of reaction products for both proteins. The amount of unannealed DNA was consequently drastically reduced. However, as the protein concentration was increased up to the μM range, the amount of product started to decrease and almost reached the level of the intrinsic annealing. In good agreement with the reduction of the quantity of dsDNA, the amount of unreacted single-stranded DNA increased.

37

In the last incubation strategy, two separate fractions of the protein were pre-incubated separately with one of the strands and subsequently combined with each other (Figure 3.7, j-m). Surprisingly, the amount of the reaction product dropped for RAD52, while in the case of RAD52(209), the annealing reaction was inhibited over the entire concentration range. Most of the single-stranded DNA appeared to not participate in the annealing reaction (Figure 3.7, k and m). Longer reaction times at higher temperatures were also tested to check whether it would improve the annealing. Therefore, reactions following the second and the third strategies were repeated at 37°C for 1 hour (Figure S9). However, the previously described trends were observed again. The overall increase in the amount of the dsDNA product was reduced and a higher amount of intrinsically annealed strands was observed (Figure S9, c and f). Remarkably, the annealing efficiency of both RAD52 and RAD52(209) was reduced at μM concentrations with the second incubation strategy (Figure S9, c). Moreover, with the third incubation strategy, the annealing efficiency was drastically inhibited (Figure S9, f).

The protein annealing activity following the second incubation strategy was also examined with the mass photometry buffer (S8, c-f). Both proteins facilitated the annealing reaction and exhibited a drop in the activity at μM concentrations (S8, c). The initial pre-incubation of RAD52(209) with either a short or a long oligonucleotide had no significant effect (S8, e-f).

In conclusion, the different incubation strategies demonstrated a significant impact on the outcome of the annealing reaction. The amount of the product was only slightly elevated in comparison to the intrinsic annealing, once the protein had not been incubated with any of the strands. However, once the fractions of the protein were preincubated with one of the strands, the annealing reaction was drastically inhibited for both proteins. This effect was so severe for RAD52(209) that even a higher incubation temperature and a longer reaction time did not facilitate the annealing. Only the initial pre-incubation of one strand with the protein and subsequent addition of the following strand demonstrated a drastic increase in the annealing reaction, similarly to RecT and Redβ [97, 19]. All the annealing experiments described below were carried out following the latter strategy providing the best annealing outcome.

3.4.3 Protein annealing efficiency at different DNA concentrations and with various DNA lengths

All above-mentioned experiments were performed with 32- and 60-nt DNA at 10 nM. I examined next how the length and concentration of a single-stranded DNA affect the outcome of the annealing reaction. Several annealing pairs have been chosen - (12- + 26-nt), (16- + 26-nt), (32- + 60-nt), and (48- + 60-nt) (Figure 3.8, a and b). First of all, both RAD52 and RAD52(209) demonstrated facilitation of the annealing reaction with

each pair of ssDNA. Secondly, RAD52(209) at 200 nM exhibited the highest amount of dsDNA (dsDNA) product, independently of the length of the oligonucleotides (Figure 3.8, a).
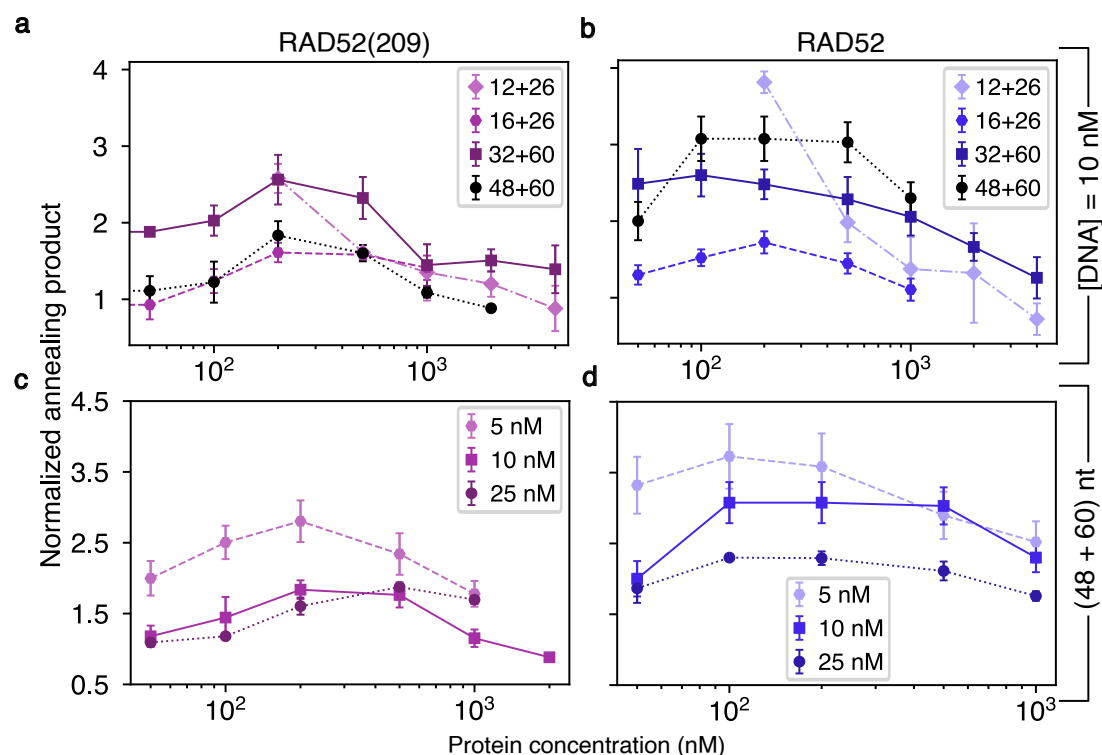
Figure 3.8: Annealing reactions with different DNA pairs and DNA concentrations. The examined annealing pairs were - (12+26), (16+26), (32+60), and (48+60) **(a)** and **(b)**. DNA concentration was 10 nM, and the proteins were tested from 50 to 4000 nM. One annealing pair (48+60) has been utilized at different concentrations - 5, 10, and 25 nM, while the protein concentration was within the range (as mentioned above) **(c)** and **(d)**. In each experiment, the protein was initially preincubated with the longest single-stranded DNA, and a complementary strand was added thereafter. The intensities of the dsDNA bands were normalized to the control reaction (intrinsic annealing) and averaged. The plots contain mean values with SEM. Each experiment was performed at least 3 times.

Interestingly, with the different pairs of DNA strands, RAD52(209) did not reveal any strong dependency between the annealing activity and the DNA lengths. On the contrary, RAD52 appeared to anneal faster longer strands ((32- + 60-nt) and (48- + 60-nt)) (Figure 3.8, b), however the annealing of ((12- + 26-nt) strands) was off. In the future, these annealing reactions should be repeated at a broader range of protein concentrations. Importantly, both proteins successfully annealed single-stranded DNA as short as 12-nt.

The effect of the DNA concentration was examined with one annealing pair: 48- and 60-nt at 5, 10, and 25 nM (Figure 3.8, c and d). The amount of the DNA product promoted by RAD52(209) was dependent on the DNA concentration. The highest amount of the DNA

product at 10 nM ssDNA was between 100 and 200 nM of RAD52(209), while at 25 nM ssDNA, it shifted to 500-1000 nM of the protein (Figure 3.8, c). Notably, at higher ssDNA concentrations, the intrinsic annealing is faster, therefore, the normalized amount of the DNA product synthesized in the presence of the protein might decrease. However, at 10 and 25 nM of ssDNA, the amount of the dsDNA annealed by RAD52(209) is almost the same within the tested concentration range. On the contrary, this dependency is visible for RAD52, however, RAD52 did not reveal a certain concentration of the protein that would result in the highest amount of the DNA product (Figure 3.8, d).

In summary, the length of the utilized oligonucleotides had no effect on the annealing activity of both proteins. RAD52(209) displayed a dependency between the concentration resulting in the highest amount of the annealing product and the concentration of the ssDNA in the reaction. However, no such dependency was observed for RAD52.

3.4.4 Native annealing reactions

The inhibition of the annealing activity at high protein concentrations had been previously demonstrated, however, not been investigated thoroughly [130, 82, 63]. To address this question, I performed native annealing reactions with both oligonucleotides carrying distinct fluorophores to observe potential intermediate annealing species over a broad range of protein concentrations (Figure 3.9). The native annealing reactions were examined primarily with 32- and 60-nt ssDNA at 10 nM (Figure 3.9). Upon increasing protein concentrations, the amount of both unreacted ssDNA molecules drastically decreased. Similarly, the amount of the free dsDNA product dropped significantly as the protein concentration increased. Partially, ssDNA accumulated in the wells, and a certain fraction appeared bound to protein (white arrows). However, whether unreacted ssDNA molecules or the already formed dsDNA product were bound to the intermediate-sized protein oligomers is unclear, as both possibilities can appear plausible. Quantitative and statistical analysis of the gels confirmed these observations, demonstrating that at high protein concentrations, both ssDNA molecules are bound to the protein species while the amount of the dsDNA product decreased significantly (Figure 3.9, b and d).

Native annealing reactions promoted by RAD52(209) were repeated with 48-nt 3'-ALEXA488 and 60-nt 3'-ATTO680 (Figure S10) oligonucleotides. However, the overall trends were the same: reduced amounts of the unreacted ssDNA and dsDNA product upon increasing protein concentrations, confirming previously obtained results that DNA length does not affect the annealing reaction. Nevertheless, these experiments revealed important information: as mentioned above, it was impossible to distinguish previously whether single-stranded or dsDNA were bound to the protein structures, in these experiments it became apparent that both unreacted ssDNA and the newly formed dsDNA products accumulated on distinct protein species and ran separately on the gel.
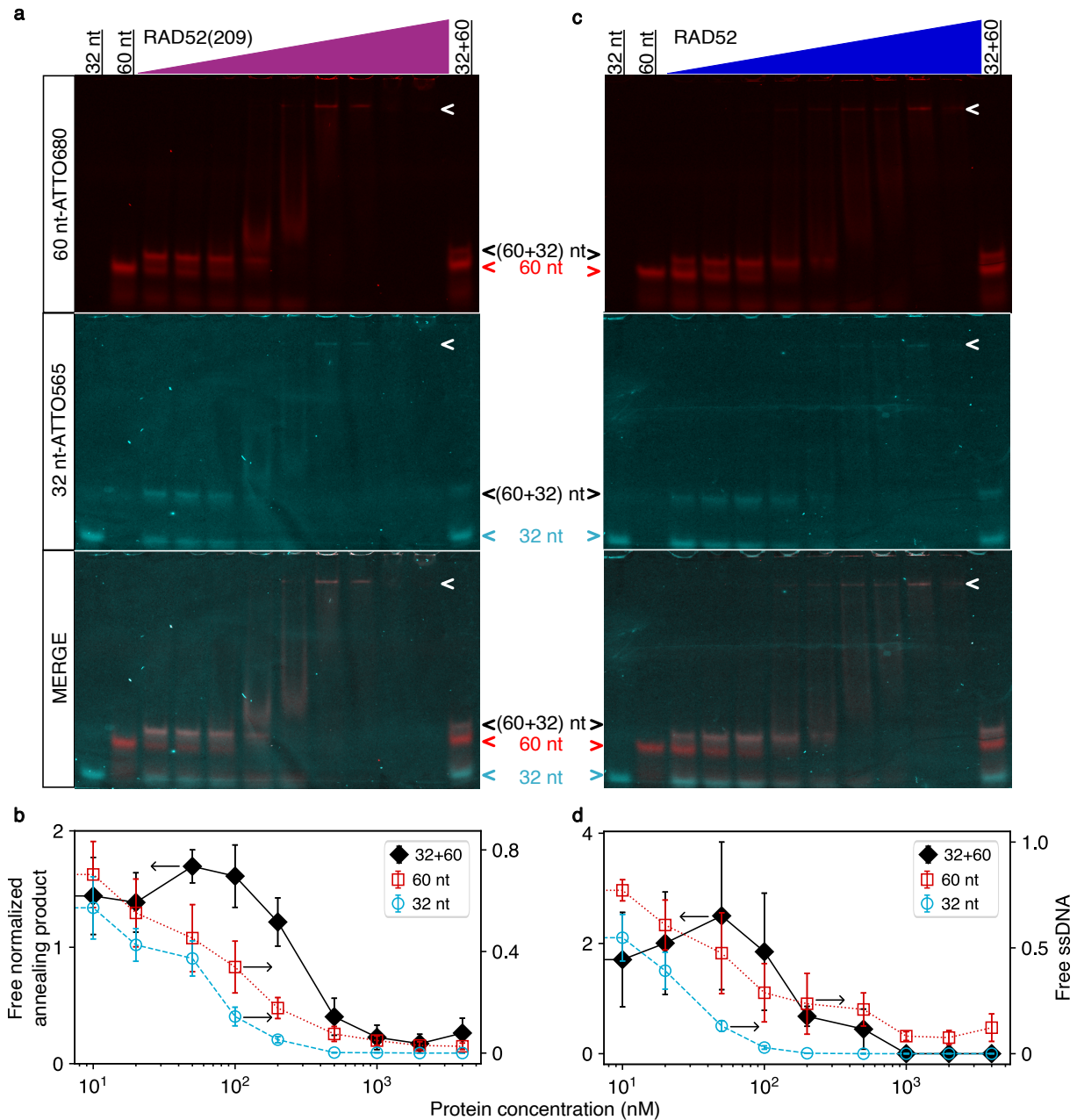
Figure 3.9: Native annealing reactions. The reactions were performed with RAD52(209) **(a-b)** and RAD52 **(c-d)**. The annealing pair was 32-nt 3'-ATTO565 (cyan open circles **(b)**) and 60-nt 3'-ATTO680 (red open squares **(d)**). The oligonucleotide concentration was 10 nM. Protein concentrations were 10, 20, 50, 100, 200, 500, 1000, 2000, and 4000 nM. Gels demonstrate the red channel (60-nt), cyan channel (32-nt), and a combined image (MERGE). White arrows point to the oligonucleotides bound to protein structures. The intensities of the dsDNA bands were normalized to the control reaction (32+60), and the intensities of the single-stranded DNA were normalized to the corresponding control: 32- or 60-nt. Data points were averaged from several repeats (5 for RAD52(209) and 4 for RAD52) and plotted as mean values with SEM. Black arrows direct data points to the corresponding axis.

RAD52(209) was incubated first with a short oligonucleotide (32- or 48-nt) before the addition of the longer complementary strand to identify whether it would affect the formation of the intermediate protein species (Figure S12). In the case of both annealing pairs, unreacted 60-nt ssDNA (Figure S11, red channel, white arrows) was running separately from the dsDNA product visible in both channels. At such low DNA and protein concentrations, the Silver Stain did not reveal any protein complexes corresponding to the DNA bands. Therefore, whether it was an unbound oligonucleotide or a protein-DNA complex remains unclear.

To understand whether the newly formed DNA product can bind to protein complexes, the annealing reaction was traced over 40 minutes at 25°C (Figure S12). During the intrinsic annealing (*i.e.* with no protein), the amount of dsDNA product increased over time. However, once the reaction was promoted by 200 nM of RAD52(209), the amount of the free DNA product drastically decreased over time, as it appeared bound to the protein complexes. Moreover, at 2 μM of RAD52(209), no ssDNA molecules or the newly formed dsDNA product ran through the gel.

Utilization of the fluorescently-labeled protein variants, RAD52(209)-DyLight488 and GFP-RAD52, was useful to overcome the limitations of the Silver Stain resolution and visualize the protein complexes bound to the DNA (Figure S13). However, in the case of RAD52(209)-DyLight488, all the DNA molecules appeared to be stuck in the wells (Figure S13, a). But with GFP-RAD52, it was remarkable how the newly formed dsDNA product migrated within distinct protein structures upon increasing protein concentrations (Figure S13, b). Notably, lower-order oligomers of GFP-RAD52 appeared as a smear with no trace of single- or double-stranded DNA bound.

To summarize, the experiments with native annealing reactions demonstrated that at high protein concentrations, ssDNA molecules were bound to distinct protein structures and therefore appeared excluded from the annealing reaction. Additionally, newly formed dsDNA products were observed bound to the protein complexes. However, the exact size of the protein oligomers that interacted with single- or double-stranded DNA molecules was not clearly identifiable. Therefore, all the above-mentioned experiments were repeated using mass photometry.

3.5   Mass photometry experiments with RAD52(209)

In an effort to clarify and avoid the switching between RAD52 data and RAD52(209), the results of the mass photometry experiments are divided into two parts. The first part will address the mass photometry results with RAD52(209), followed by the results for RAD52.

3.5.1 Single-stranded DNA binding in the mass photometry buffer

One set of experiments was performed in the mass photometry buffer in order to estimate how the absence of bivalent ions affects protein-DNA binding. Based on the annealing efficiency of RAD52(209) (Figure 3.7) and in order to obtain a significant number of counts for the mass histograms, 100 nM was chosen as the concentration for the majority of the experiments. The DNA concentrations were 10 and 100 nM.

To visualize the oligomeric state of the protein upon the binding of DNA molecules, RAD52(209) was mixed with ssDNA molecules of various lengths (Figure 3.10). The DNA molecules were chosen to be either shorter or equal to the ring circumference (11 monomers bind 44 nt): 32-nt (**a**) and 48-nt (**b**), or to be longer: 60-nt (**c**) and 83-nt (**d**). As previously described (see section "Mass photometry data"), RAD52(209) organizes as an undecameric ring with stacks thereof and a small fraction of lower-order oligomers of various lengths. The cumulative distribution displaying the normalized area under the Gaussian fit from different repeats of the experiment is presented as a heatmap on the lower panel (Figure 3.10, b, d, f, h). Upon mixing DNA with the protein in a 1:10 ratio, the number of counts for molecular weights corresponding to single-rings decreased significantly independently of the DNA length (Figure S14, c). The averaged number of RAD52(209) single rings was 206 ± 19 (mean $\pm$ SEM), after the protein was mixed with ssDNA the counts for single rings were 65 ± 20 (for 32 nt), 22 ± 3 (48 nt), 80 ± 18 (60 nt), and 64 ± 11 (83 nt). The peaks corresponding to stacks of rings diminished drastically and contained only a small fraction within all protein species. The variety of lower-order oligomers increased. However, the total number of counts in the presence of DNA was drastically reduced in comparison to the control reaction (*i.e.* no DNA, upper panel) (Figure S19). Additionally, the single-ring peak and the detected peaks of the ring stacks were shifted towards higher molecular weights (marked with pink arrows). These shifts were quantified and they correspond to two single-stranded DNA molecules bound to the same ring (Figure S14, a). Remarkably, the presence of 83-nt DNA caused a significant broadening of the area corresponding to the peak of a single ring (*i.e.* undecamer). For the rest of the oligonucleotides, the obtained sigma values of the single ring peak were not significantly different from the control reaction (Figure S14, b).

Once the protein was saturated with single-stranded DNA in a 1:1 ratio, the number of counts corresponding to the single-ring peak increased, which resulted in higher (less significant) p-values (Figure S14, f). To compare these conditions with the 1:10 case described above, the number of single rings at 100 nM DNA became 97 ± 15 (for 32 nt), 40 ± 4 (48 nt), 118 ± 30 (60 nt), and 92 ± 13 (83 nt). For some replicates, counts for a single ring peak appeared to be even higher than in the control (Figure 3.10, a and c). Moreover, the counts of the ring stacks did not change and were barely detectable. In the case of protein binding to 32-, 48-, and 60-nt ssDNA, the diversity of the lower-order oligomers was reduced. However, the number of species presumably corresponding to dimers and

trimers strongly increased. In the case of 83-nt binding to RAD52(209), the number of protein dimers and trimers also increased, however, the lower-order oligomers of various lengths were detected as well. Quantification of the molecular weight shifts corresponding to single rings bound to DNA showed that two single-stranded DNA molecules were bound to the same ring (Figure S14, e), similarly to the 1:10 ratio. However, the sigma values were significantly different from the control reaction for each tested DNA length (Figure S14, e) indicating that the single ring peak became broader and could contain different populations of species with a different number of DNA molecules.
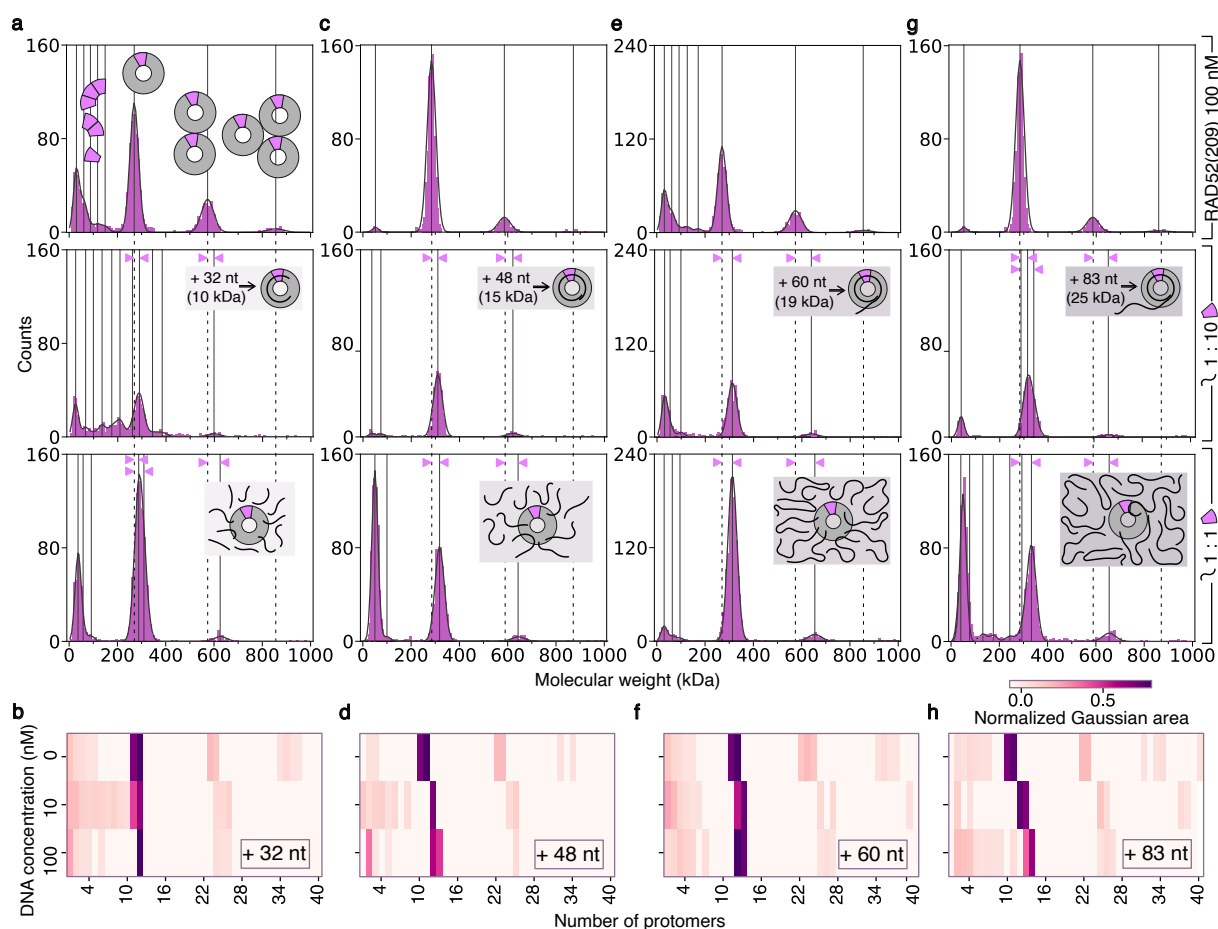

Figure 3.10: Mass histograms depicting RAD52(209) binding to single-stranded DNA. Utilized DNA molecules had various lengths: 32-nt **(a-b)**, 48-nt **(c-d)**, 60-nt **(e-f)**, and 83-nt **(g-h)**, DNA concentrations were 10 and 100 nM. The protein concentration was 100 nM. The oligomeric state of the protein alone is shown on the upper panel, followed by the protein incubated with single-stranded DNA at 1:10 and 1:1 DNA to protein ratio. Dashed lines depict the ring peaks and stacks thereof at the control reaction (no DNA) as references. Solid vertical lines demonstrate the individual protomers. The heatmaps demonstrate the normalized Gaussian areas corresponding to different protomers (vertical lines) as the cumulative distributions of all the repeats **(b)**, **(d)**, **(f)**, and **(h)**. The Gaussian envelopes are depicted as solid black lines. Inserts contain schematics of the protein binding to DNA molecules at different conditions.

44

The annealing activity of the proteins has been tested at protein concentrations significantly exceeding the DNA concentration (Figure 3.7, h). It was therefore important to also test whether the proteins are able to promote the annealing with the exceeding DNA concentrations. To address this question, I acquired the mass photometry data of RAD52(209) at 100 nM mixed with single-stranded DNA at 200 nM (Figure S15) to detect any possible alterations in the oligomeric state of the protein. The DNA molecules used were 48- and 83-nt in size. Strikingly, a strong reduction in the number of counts corresponding to rings was observed. Additionally, very strong peaks in the range of MW that could correspond to protein dimers and trimers appeared. However, the Silver Stain experiment performed with low protein-to-DNA ratios showed a significant amount of unbound DNA. To reveal whether mass photometry peaks corresponded to protein dimers or free DNA, I measured DNA alone (Figure S28). Indeed, at high concentrations, both 48- and 83-nt DNA produced a strong signal with an estimated MW peak similar to a dimer of RAD52(209). Therefore, it was not possible to distinguish whether the detected low MW peaks corresponded to free ssDNA or protein dimers. Nevertheless, the observed higher MW peaks most definitely corresponded to different oligomeric protein species. Notably, almost all ring stacks disappeared (Figure S15, a-d). The analysis of the MW shifts corresponding to single rings revealed that, in these conditions, a single ring could accumulate three 48-nt DNA and two 83-nt DNA (Figure S15, e). This resulted in a broader distribution of the peaks corresponding to rings and significantly larger sigma values indicating different populations (Figure S15, f). Next, I performed ssDNA annealing to estimate how the excess of the DNA affected RAD52(209)'s annealing activity (Figure S16). RAD52(209) at 100 nM was preincubated with 200 nM of 60-nt ssDNA and subsequently, the complementary 32-nt strand (10 nM) was added to the reaction (Figure S16, a). In these conditions, RAD52(209) was not facilitating the annealing (colored in grey) (Figure S16, b). The amount of the newly formed dsDNA product was similar to the intrinsic annealing. The slight facilitation of the annealing reaction only occurred at around 1 μM of the protein.

3.5.2 RAD52(209) forms higher-order oligomers

Occasionally, in the regular field of view (RFV), the high molecular weight (MW) species were detected, as in the case of RAD52(209) mixed with 48-nt (Figure 3.11, b). However, as the sensitivity of the system in the RFV decreases upon increasing the ability to detect higher MW, the above-mentioned experiments were repeated using the medium field of view (MFV) (Figure 3.11). It is important to mention that MFV has a lower detection sensitivity for species with molecular weights below 60 kDa.
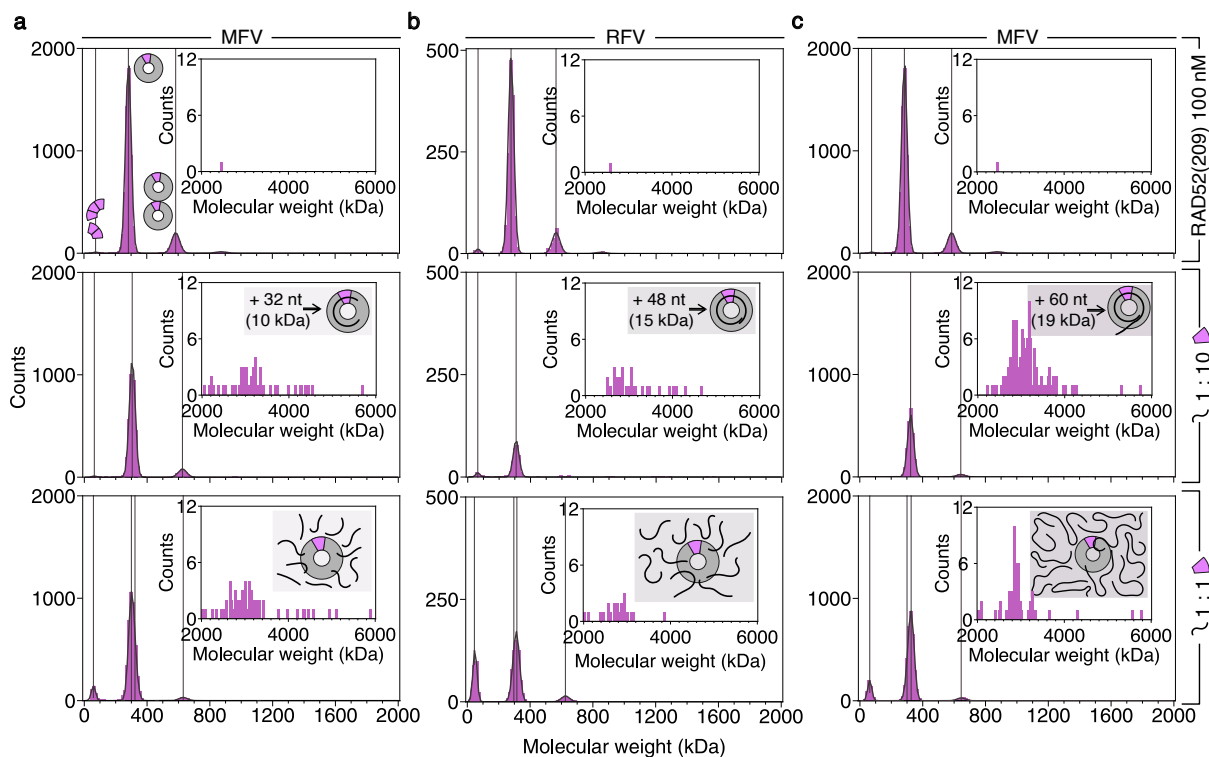
Figure 3.11: The formation of RAD52(209) higher-order oligomers upon single-stranded DNA binding. Utilized DNA molecules were - 32- **(a)**, 48- **(b)**, and 60-nt **(c)**. The experiments with 48-nt were performed in the regular field of view (RFV), and the rest were measured in the medium field of view (MFV). RAD52(209) concentration was 100 nM, and DNA concentrations were 10 and 100 nM. The upper panel depicts the oligomeric state of the protein, followed by the protein incubated with DNA at 1:10 and 1:1 DNA to protein ratio. The inserts present the graphics of the protein binding to the DNA of various lengths and concentrations. Each dataset was fitted to many Gaussian functions, and the resulting mean values (protomer masses) are depicted as vertical lines. Dashed lines are references to the ring peaks and stacks thereof at the control reaction (no DNA). The Gaussian envelopes are shown as solid black lines. Each mass histogram is a cumulative distribution of several readings acquired on the same day.

Therefore, RAD52(209) monomers could not be properly detected in the MFV. However, in these experiments, the major focus was on high MW structures. Upon binding to ssDNA (10 nM), independently of its length, RAD52(209) (100 nM) formed oligomers with molecular masses in the range of 3 to 5 MDa. Due to the wide distribution of the detected species, it was not possible to fit the data and obtain the exact peak values. Upon the saturation of the protein with single-stranded DNA (1:1 ratio), the masses of these higher-order oligomers shifted towards lower molecular weights indicating that the excess of the DNA appeared to dissociate these complexes (Figure 3.11, lower panel). Interestingly, experiments performed on RAD52(209)-DyLight488 (100 nM) in RFV demonstrated the opposite trend upon ssDNA binding (10 nM) (Figure S5, a). First of all, the protein alone formed dimers, single rings, and stacks thereof. And, remarkably, single rings

46

were not as dominant as in the case of RAD52(209) at similar concentrations. Moreover, higher-order oligomers above 1 MDa were observed. Once the ssDNA was added, these higher-order structures disappeared and counts corresponding to single rings increased. Remarkably, the number of dimers and double rings was the same, indicating that the higher-order oligomers potentially contained only single rings. However, the MW shift of the single ring was similar to unlabeled RAD52 and corresponded to two DNA molecules bound.

Compellingly, RAD52(209) at 500 nM increased the amount of the annealing product only up to 2.3 times (Figure 3.7, h), while 100 nM already doubled it. To understand, why such an excess of the protein did not result in the proportional increase of the DNA product, I performed mass photometry measurements with RAD52(209) at 500 nM alone and in the presence of ssDNA: 32-, 60-, and 83-nt at 10 nM in the MFV (Figure 3.12). Surprisingly, a small fraction of lower-order oligomers including monomers was still detected in the MFV (Figure 3.12, b). No single-stranded DNA of 32- or 60-nt appeared bound to single rings and only rings with one 83-nt DNA bound were detected (Figure 3.12, c). Therefore, the width of the single ring peak did not change and the obtained sigma values were not significantly different from the control measurement (Figure 3.12, d). Interestingly, the total number of counts for the single ring peak decreased proportionally to the DNA length (Figure 3.12, e). Similarly, the total number of counts diminished upon ssDNA binding correlating with its length (Figure S20). RAD52(209) at 500 nM had 1193 ± 185 counts for the single-ring peak, upon binding to the 32-nt DNA, the ring counts were 779 ± 176. Once the 60-nt DNA was mixed with RAD52(209), the counts decreased to 531 ± 142, and the binding of the 83-nt DNA reduced the ring counts to $64, \pm 19$.

To summarize, the binding of ssDNA molecules significantly altered the oligomeric state of RAD52(209). A 1:10 DNA-to-protein ratio caused the formation of a variety of lower-order oligomers, while the number of rings decreased significantly. The binding of the DNA resulted in the MW shift of the single-ring peaks. The quantification of these shifts revealed that a single ring can accumulate up to two ssDNA molecules. The stacks of rings disappeared almost completely. However, higher-order oligomers of 3-5 MDa were detected, suggesting that the ring stacks accumulated into these substantially sized oligomers. Interestingly, the excess of ssDNA dissociated some of the higher-order oligomers which potentially caused a slight increase in the total number of single rings, however, the stacks of rings were not recovered. Additionally, the excess of DNA did not change significantly the number of DNA molecules bound to a ring. Only at 200 nM DNA, three 48-nt molecules were detected on the same ring. This was surprising considering that there were two DNA molecules per monomer and a ring structure containing 11 subunits theoretically having 22 DNA binding sites. However, negatively charged DNA molecules could be repelled from the negatively charged surface of the glass slide. To exclude this effect, I repeated mass photometry experiments in the annealing buffer that contains Mg^{+2}. Mg^{+2} binds DNA molecules and, therefore, facilitates landing on the glass surface of the

protein complexes bound to the DNA.
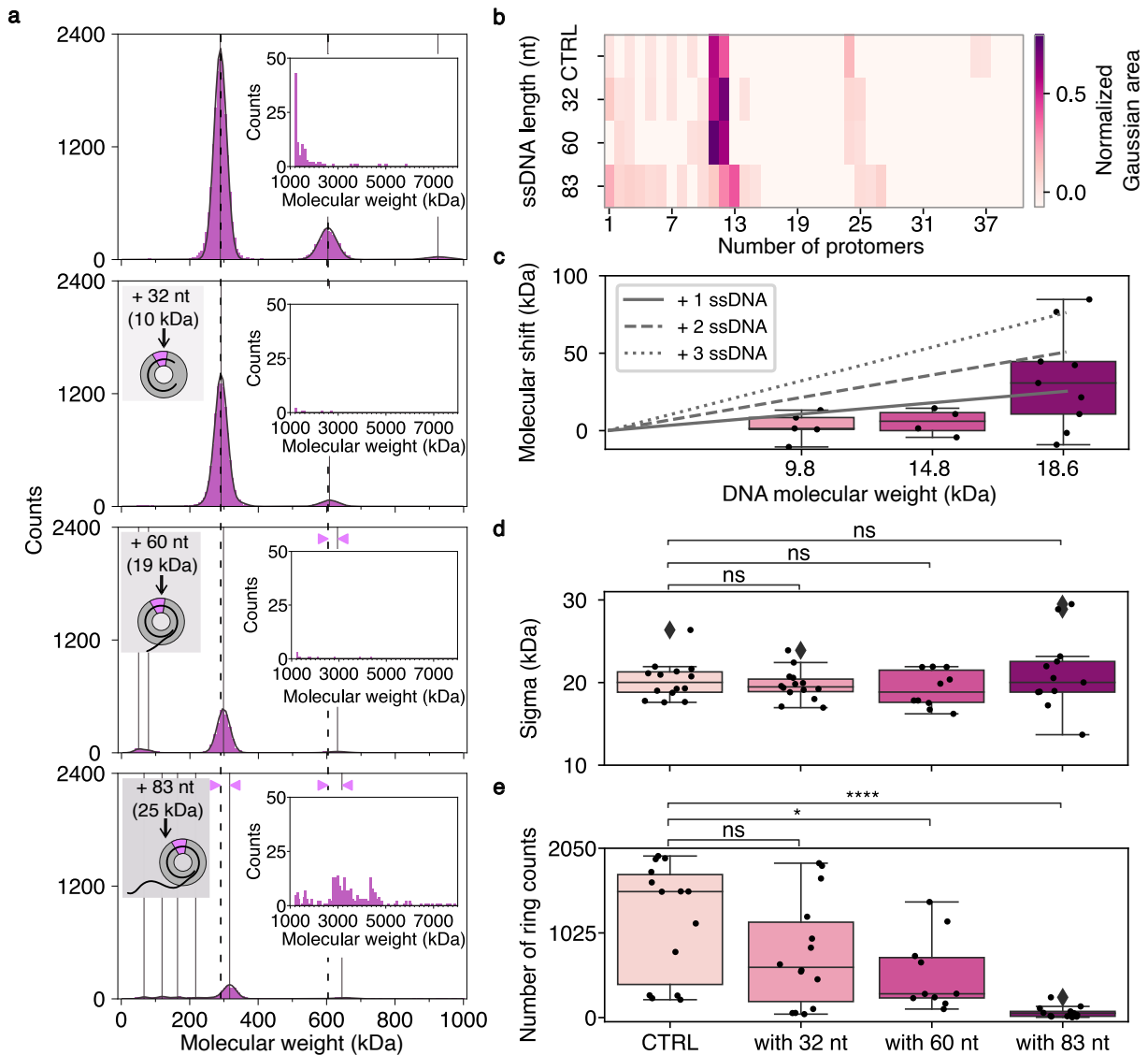
Figure 3.12: Mass histograms RAD52(209) at 500 nM upon single-stranded DNA binding. Utilized DNA molecules had various lengths - 32-, 60-, and 83-nt at 10 nM. All the experiments were performed in the MFV (Medium Field of View). **(a)** The mass histograms show cumulative sums of several readings performed on the same day. The individual protomers peaks and the Gaussian envelopes are shown as before. **(b)** The heatmap shows the normalized Gaussian areas of various protomers as the cumulative distribution from all the repeats (N=4). The boxplots depict the analysis of the single ring peak - MW shifts **(c)**, sigma values **(d)**, and the number of counts **(e)**. The data is averaged and plotted as mean values with SEM. Each black circle corresponds to one reading, the filled diamonds are the outliers. The two-sided Mann-Whitney test was performed, and the corresponding p-values are depicted as asterisks or "ns" above the graphs.
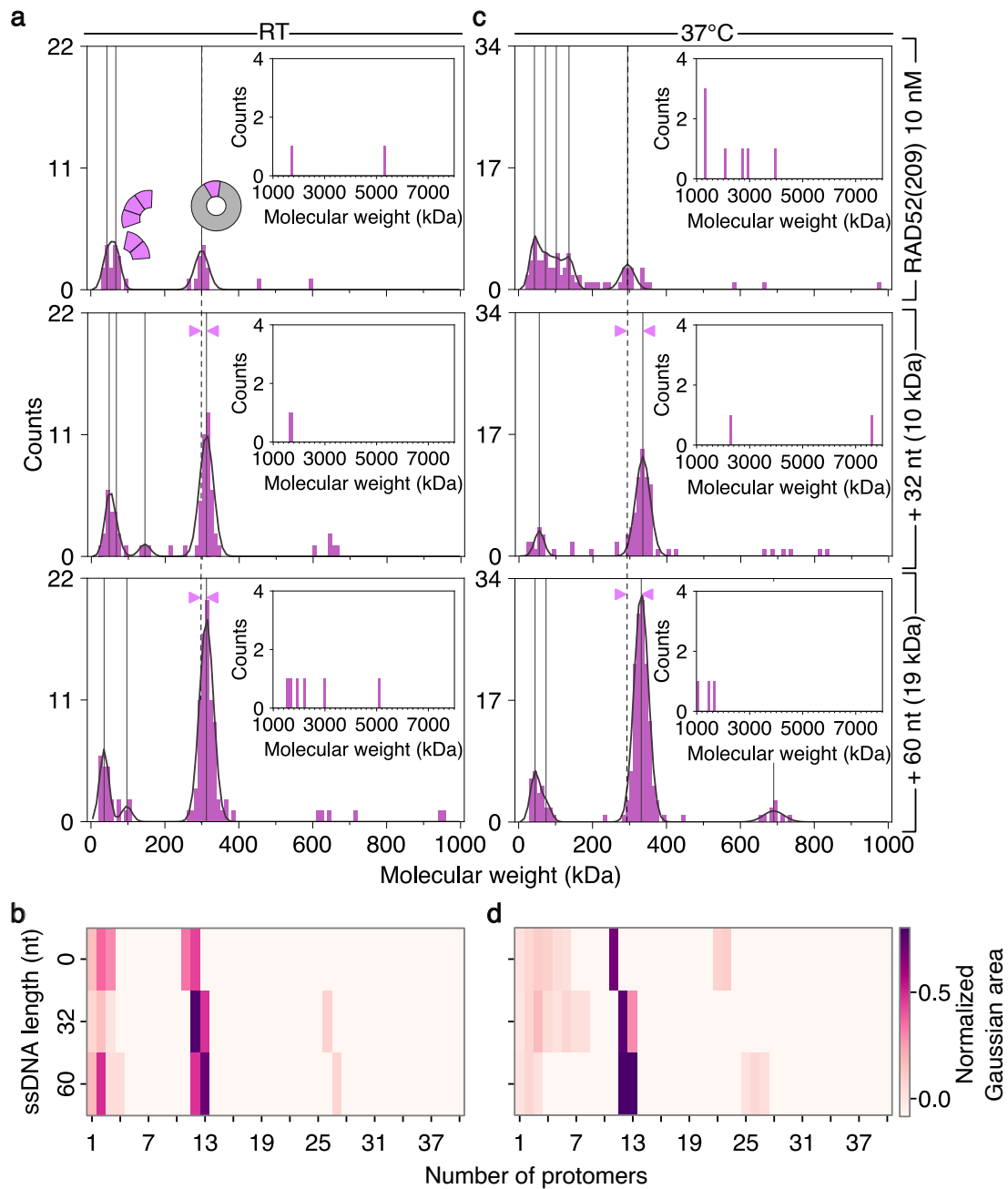
3.5.3 Single-stranded DNA binding in the annealing buffer

Mass photometry experiments in the annealing buffer were performed with RAD52(209) at 100 nM alone and in the presence of 32- or 60-nt ssDNA at 10 nM, to compare to the annealing experiments (Figure 3.7). Moreover, the measurements were performed at 25°C and 37°C to estimate the levels of the protein titration upon longer incubation at higher temperatures (Figure 3.13). RAD52(209) at 100 nM formed undecameric rings, stacks thereof, and a small fraction of short oligomers. However, the total number of counts was lower than in the mass photometry buffer, while higher-order oligomers with masses around 2-4 MDa were detected (Figure 3.13, upper panel). Surprisingly, upon adding the single-stranded DNA, the total number of counts increased especially the single ring counts (Figure 3.13, a and Figure S17, c and f). The average number of RAD52(209) single rings was 45 ± 5 (mean $\pm$ SEM), after the protein was mixed with ssDNA the counts for single rings were 102 ± 18 (for 32 nt) and 115 ± 20 (60 nt). The absence of the higher-order oligomers could be due to the RFV limitations. These results are opposite to the ones acquired in the mass photometry buffer. Moreover, the sigma values of the single ring peaks did not change significantly upon DNA binding (Figure S17, b and e). Quantification of the MW shifts of the single ring peak demonstrated that, independently of the temperature and incubation time, single rings accumulated three 32-nt DNA and only two 60-nt DNA (Figure S17, a and d). Notable, the longer incubation at higher temperatures had very little to almost no effect (Figure 3.13, c).

Due to a strong increase in the total number of counts in the annealing buffer, I repeated the above-mentioned experiments at 10 nM of RAD52(209) to observe the oligomeric species at these, considerably low, protein concentrations (Figure 3.14). The 1:1 DNA-to-protein ratio was chosen to mimic the annealing assays. Two incubation temperatures were tested again as described above. Due to the low protein concentration, the overall number of counts of RAD52(209) alone was very low, and stacks of the rings as well as the higher-order oligomers were rarely observed (Figure 3.14, inserts). Despite this, RAD52(209) in the absence of DNA still formed undecameric rings and a fraction of lower-order oligomers. As observed before, in the annealing buffer upon binding to the ssDNA, the total number of counts significantly increased, and even some double-rings were detected. The average number of RAD52(209) single rings at 10 nM was 5 ± 1 (mean $\pm$ SEM), after the protein was mixed with ssDNA the counts for single rings were 18 ± 3 (for 32 nt) and 15 ± 3 (60 nt). The single ring counts increased drastically independently of the incubation temperature (Figure S18, c and f). The DNA binding to the single ring shifted its mass by two and sometimes three masses of the corresponding DNA (Figure S18, a and d). In most of the cases, sigma values were not significantly different (Figure S18, b and e) suggesting that there was a single population of rings present with two or three ssDNA bound.
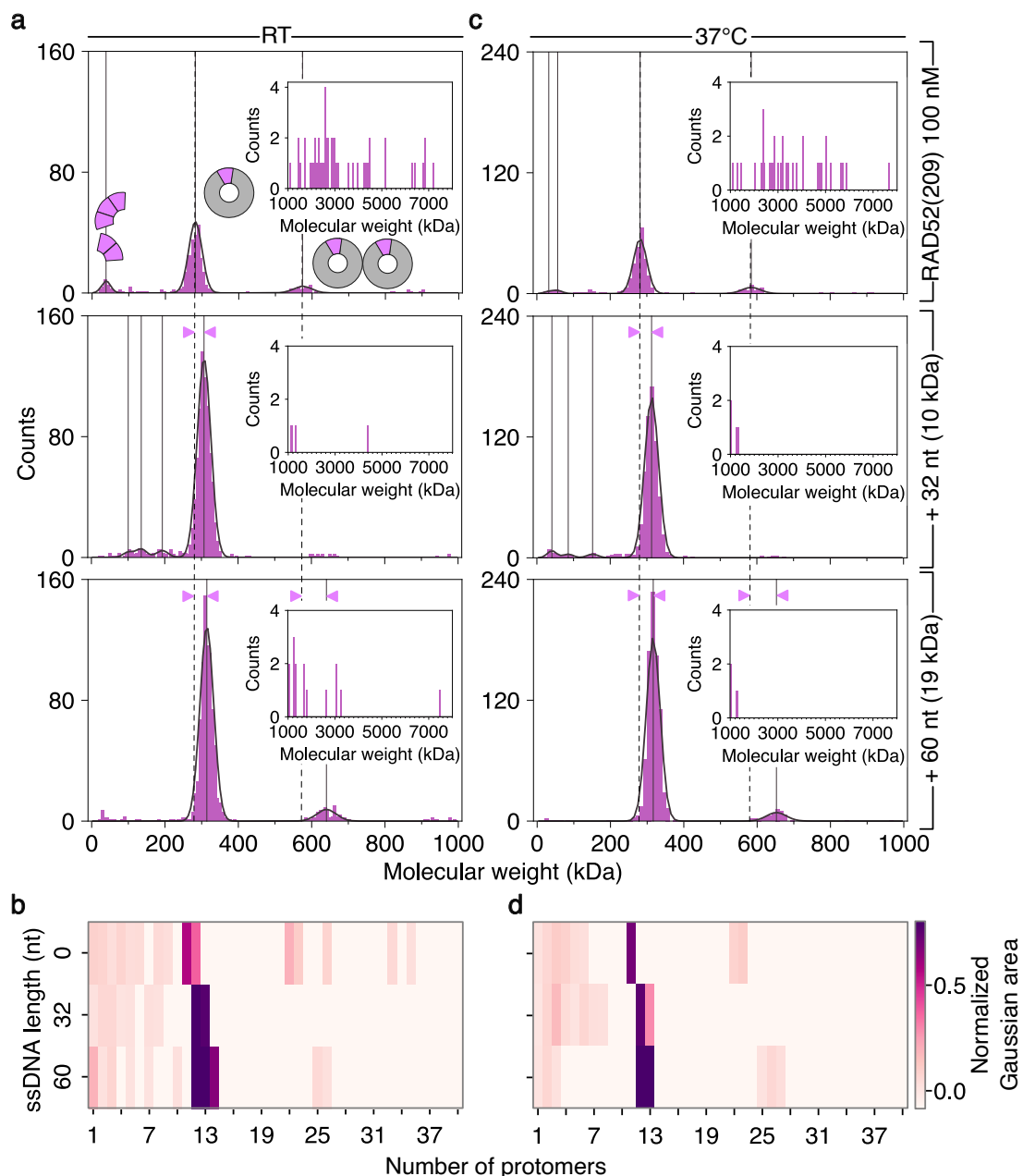
Figure 3.13: The oligomeric state of RAD52(209) at 100 nM upon single-stranded DNA binding in the annealing buffer. DNA molecules were 32- and 60-nt long. RAD52(209) concentration was 100 nM, and the DNA concentration was 10 nM. The experiments were performed at 25°C **(a-b)** and 37°C **(c-d)**. The upper panel depicts the oligomeric state of the protein alone, followed by the protein incubated with 32- and 60-nt DNA The inserts contain mass histograms at high MW. The black solid lines are the Gaussian envelopes. The vertical solid lines depict the masses of each protomer, dashed lines are the references of the control measurement (no DNA). Each mass histogram is a cumulative distribution of several readings acquired on the same day. The heatmaps show the normalized Gaussian area of each protomer as the cumulative distribution from different days. The data was acquired in the RFV.

Figure 3.14: The oligomeric state of RAD52(209) at 10 nM upon single-stranded DNA binding in the annealing buffer. DNA molecules were 32- and 60-nt long. RAD52(209) concentration was 10 nM, and the DNA concentration was 10 nM. The experiments were performed at 25°C **(a-b)** and 37°C **(c-d)**. The upper panel depicts the oligomeric state of the protein alone, followed by the protein incubated with 32- and 60-nt DNA The inserts contain mass histograms at high MW. The black solid lines are the Gaussian envelopes. The vertical solid lines depict the masses of each protomer dashed lines are the references of the control measurement (no DNA). Each mass histogram is a cumulative distribution of several readings acquired on the same day. The heatmaps show the normalized Gaussian area of each protomer as the cumulative distribution from different days. The data was acquired in the RFV.

Mg^{+2} in the reaction buffer affected the oligomeric state of RAD52(209). While forming undecameric rings and stacks thereof, higher-order oligomers with molecular masses of around 2-6 MDa were also detected. Upon binding to the ssDNA, independently of its length, these higher-order oligomers were no longer present. However, as expected, Mg^{+2} significantly increased the binding of the protein-DNA complexes to the glass surface. Additionally, the increased number of rings could be the result of dissociated higher-order oligomers. The incubation temperatures caused a minor reduction in the total number of counts while distinct types of protein species were still present. Binding of the single-molecule of DNA to RAD52(209) in the mass photometry buffer altered its oligomeric state significantly. Notably, only concentrations of the DNA had an effect on the oligomerization of RAD52(209), while their lengths had no essential effect. With the protein present in excess relative to the DNA (10:1 ratio), RAD52(209) formed higher-order oligomers with molecular weights between 3-5 MDa, irrespective of the DNA length. A variety of lower-order oligomers were detected, with an exception for 83-nt binding. At a 10:1 ratio, the shift in the mass of the single ring peak was equal to two DNA molecules bound, with an exception again for 83-nt. Once the protein was saturated with DNA, at a 1:1 ratio, these shifts were equal to two DNA molecules bound to a ring for all the DNA lengths tested. The saturating amount of DNA molecules increased the number of rings, however, not the number of stacks. Higher DNA concentrations reduced the variety of lower-order oligomers with dominant populations of dimers. Importantly, the distribution of protein monomers at high DNA concentrations could not be completely distinguished from the signal of unbound DNA. Interestingly, once the protein was incubated with 83-nt DNA, a broad variety of different lower-order oligomers were detected. Moreover, once RAD52(209) at 500 nM was mixed with ssDNA at a 50:1 molar ratio, only the 83-nt DNA caused significant changes, such as a massive reduction of the total number of counts, and detectable shift in the MW of the single ring peak (1-2 DNA molecules). In the case of the other two ssDNA sizes tested, 32-nt and 60-nt, no DNA bound to the ring was detected as well as no higher-order oligomers. However, considering that 60-nt caused a significant reduction in the total number of counts, certain alterations of the protein's oligomeric state could potentially occur with longer DNA molecules.

3.6 Double-stranded DNA binding in the mass photometry buffer

RAD52(209) is capable of accumulating dsDNA, although with a lower affinity compared to ssDNA [59]. To visualize how dsDNA binding alters the oligomeric state of the protein, mass photometry experiments were carried out with pre-annealed dsDNA molecules. One pair was 48- and 48-nt to obtain DNA molecules with no overhangs, and another pair

was 48- and 60-nt with 6-nt overhangs at the 3' and 5' ends (Figure 3.15). The protein was mixed with the DNA at 10:1 and 1:1 ratios. In the presence of the dsDNA without overhangs single-ring structures were almost completely absent (Figure 3.15, a and g). RAD52(209) control measurements produced 152 ± 4 counts for the single ring, while upon dsDNA binding it became 4 ± 1. Interestingly, the double rings were still present which was not the case once the protein was incubated with single-stranded DNA. Notably, the protein formed higher-order oligomers with molecular masses of up to 5 MDa (Figure 3.15, a, inserts). These higher-order oligomers were different from the ones that formed with the single-stranded DNA molecules. The complexes formed with the single-stranded DNA were around 3-5 MDa. In comparison, the protein oligomers that appeared on the double-stranded DNA had molecular masses around 1, 1.5, 2, and 2.5 MDa, indicating a certain periodicity that resembles a filament-like complex when the DNA is assembled on the multiple protein rings. The most recent cryo-EM structure demonstrated two undecameric rings of RAD52 assembled on the double-stranded part of the DNA replication fork [54]. The total number of binding events was significantly reduced (Figure S21, a). Once the protein was mixed with the dsDNA at a 1:1 ratio, a certain fraction of higher-order oligomers dissociated, at the same time the number of single rings slightly increased (23 ± 5) and the total number of counts was not significantly different from the control reaction (Figure S21, b). Moreover, strong peaks resembling protein dimers and trimers appeared. However, these peaks could potentially be signals from free dsDNA (Figure S28, d-f). The quantification of the molecular weights for the single-ring peaks demonstrated a broad distribution of species populations. Some of these masses could correspond to one or two dsDNA bound to a ring (Figure 3.15, e). Conclusively, the corresponding sigma values were significantly different at a 1:1 protein-to-DNA ratio (Figure 3.15, f). The double-ring structures were easily detected with a significant shift towards the higher molecular weights upon DNA binding.

The presence of the dsDNA with 6-nt overhangs resulted in alterations of the protein oligomeric state similar to the observed behavior occurring when the protein is incubated with single-stranded DNA (Figure 3.15, c-d). At the 10:1 protein-to-DNA ratio, the overall number of single rings was reduced (69 ± 3) but not as drastically as in the case of double-stranded DNA with no overhangs (Figure 3.15, g). The total number of counts did not change significantly either (Figure S21, a). The double rings were almost gone. The higher-order oligomers were also detected, but with lower masses and reduced number of counts. Similarly, to the dsDNA with no overhangs, at higher DNA concentrations (1:1 ratio), the number of single rings increased (74 ± 12) together with a variety of lower-order oligomers that potentially resulted in a higher number of total binding events (Figure S21, b). Presumably, the lower molecular weight peaks corresponded to the unbound dsDNA. The quantification of the molecular weight of the single ring peaks revealed a broad distribution with distinct protein fractions bound to one or two dsDNA (Figure 3.15, e). Corresponding sigma values were significantly different from the control, especially at a

1:1 protein-to-DNA ratio.
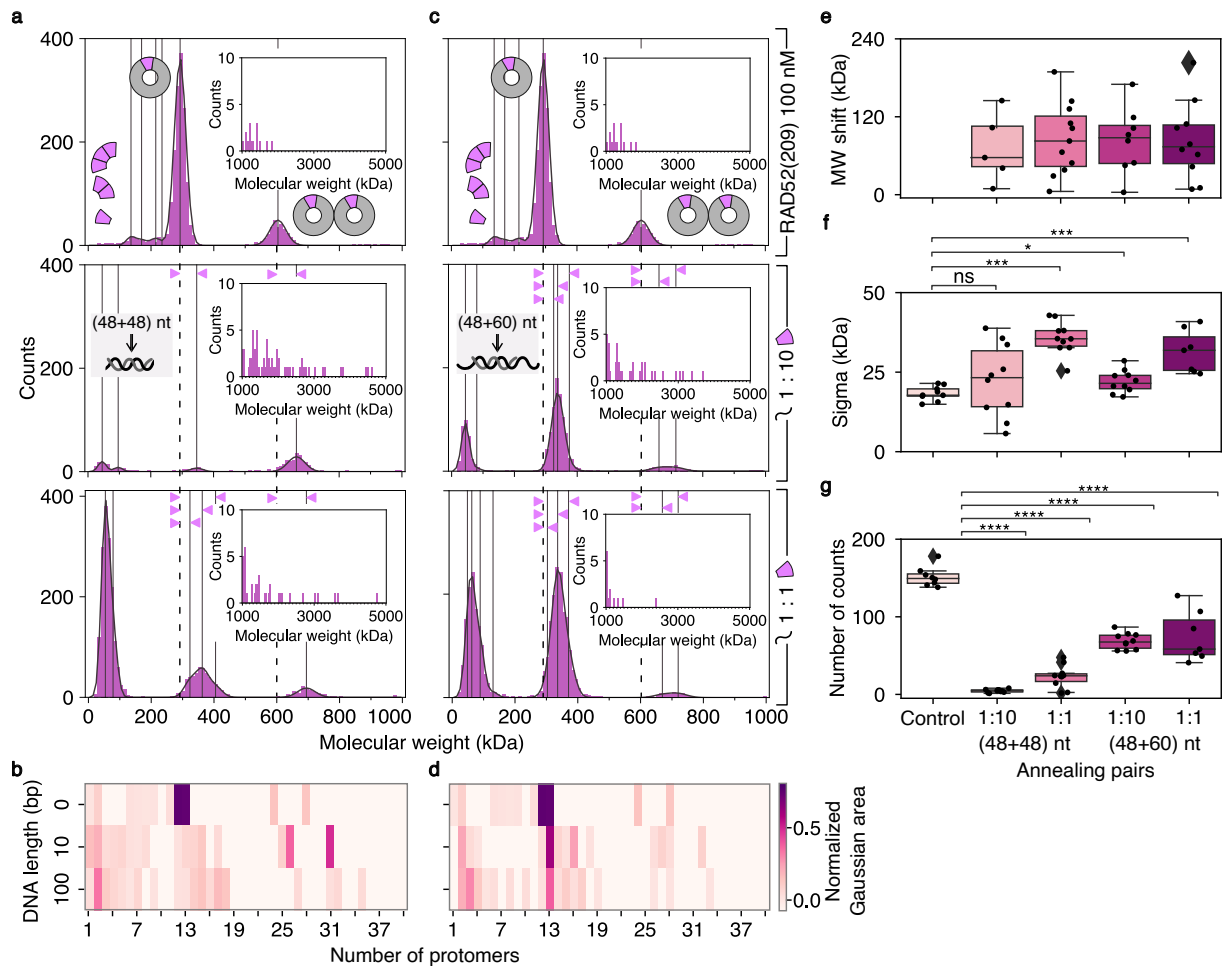
Figure 3.15: RAD52(209) binding to double-stranded DNA. The dsDNA without the overhangs was formed by 48- and 48-nt single strands **(a-b)**, and the DNA with the overhangs was formed by 48- and 60-nt strands **(c-d)**. DNA concentrations were 10 and 100 nM. The protein concentration was 100 nM. The oligomeric state of the protein **(a)** and **(c)** is followed by the corresponding heatmaps **(b)** and **(d)**. Inserts demonstrate protein structures above 1 MDa. Plots indicate the single ring analysis - MW shifts **(e)**, sigma values **(f)**, and the total number of counts **(g)**. The two-sided Mann-Whitney test was performed and the corresponding p-values marked as asterisks or "ns". Each black circle is a single data point, the black diamonds are the outliers.

To summarise, the binding of the dsDNA with and without the overhangs caused distinct alterations in the protein oligomeric state. Upon the binding of the DNA without overhangs, the single rings disappeared almost completely followed by the formation of higher-order oligomers of up to 5 MDa. In comparison, the binding of the DNA with the overhangs was similar to the single-stranded DNA binding experiments. However, the type of higher-order oligomeric species was different and exhibited a certain periodicity. Many single-ring structures were present, while the double rings were gone. In both cases,

the ring peak shifted towards higher molecular weights demonstrating a broad distribution of protein fractions potentially bound to one or two dsDNA molecules. Knowing how the binding of single- and double-stranded DNA alters the oligomeric state of RAD52(209), the next step should be an attempt to observe the single-stranded DNA using mass photometry.

3.7 Single-stranded DNA annealing

The DNA annealing process promoted by RAD52(209) was monitored on mass photometry setup (Figure 3.16). The protein concentration was 100 nM, and the DNA concentration was 10 nM. The annealing activity of the protein was tested with the complementary single-stranded DNA (32- and 60-nt), with either the short or the long strand being preincubated with the protein first, labeled as (32+60)- or (60+32)-nt, respectively (Figure 3.16, a and b). Moreover, two pairs of non-complementary strands, labeled as (32+83)- and (60+83)-nt, were also examined to reveal whether the oligomeric state of the protein is affected by mismatches in the DNA sequences (Figure 3.16, c and d). During the annealing reaction promoted by RAD52(209) with two complementary strands, the total number of counts was significantly reduced (Figure 3.16, a and g). A small fraction of the short oligomers was still detected, while the amount of single- and double-rings was reduced. The single-ring counts for RAD52(209) alone were 394 ± 50, during (32+60)-nt annealing, the ring counts reached 45 ± 11 and 78 ± 8 in case of (60+32)-nt reaction. Notably, there was a significant amount of higher-order oligomers with molecular masses formed from 1 to 6 MDa (Figure 3.16, inserts). In the case of the mismatching strands, the number of counts was reduced (51 ± 14 for (32+83)-nt and 77 ± 6 for (60+83)-nt). However, not as drastically as when the strands were complementary, suggesting that this effect was due to the binding of the newly formed dsDNA to the protein (Figure 3.16, c and d). Moreover, a very distinct single-ring peak was present, while the double rings were almost completely absent. Rarely, the high-order oligomers were also detected, however, they had lower molecular masses and reduced number of counts. Remarkably, the sigma values for the single-ring structures appeared to be significantly higher than the control reaction only when the long single-stranded DNA was initially preincubated with the protein (60 and 32, 60 and 83, 60-nt was preincubated with RAD52(209)) (Figure 3.16, f). The molecular weights of the single-ring peaks were higher and broader, however, they did not display any dependency on the annealing pairs (Figure 3.16, e). The averaged MW shifts of a single-ring structure are listed in Table S7. I analyzed the annealing process on the complementary and non-complementary strands at different time points to examine the cause of the broad single-ring peak (Figure S31). Based on the MW shifts, one could see that there was a constant association and dissociation happening on the single- and double-ring structures. Interestingly, the first strand that was initially preincubated with

RAD52(209) (60-nt) appeared to stay associated with the ring over almost 10 minutes, while the second strand was binding transiently by one or two molecules (Figure S31, a). Notably, the annealing process at 100 nM RAD52(209) with the complementary strands (Figure 3.16, a) resulted in distinct oligomeric states compared with the protein mixed with the pre-annealed dsDNA with overhangs (Figure 3.15, c). This indicates that the variety of the protein species detected was a signature of the single-strand annealing reaction and not simply newly-formed double-stranded DNA bound to RAD52(209).
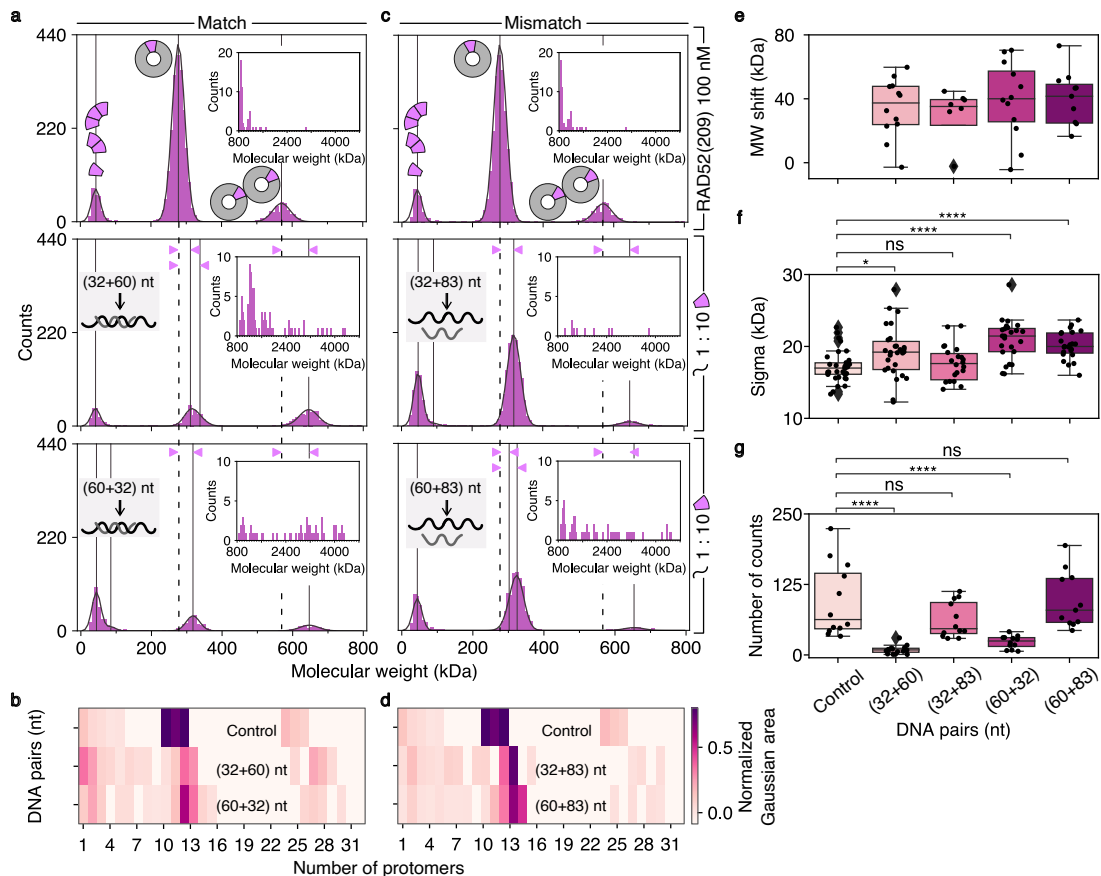
Figure 3.16: Single-stranded DNA annealing in the presence of RAD52(209). The annealing pair was - 32- and 60-nt **(a-b)**, as well as two non-complementary pairs - 32- and 83-nt, 60- and 83-nt **(c-d)**. The DNA concentrations were 10 nM, RAD52(209) concentration was 100 nM. **(a)** and **(c)** The oligomeric state of the protein alone and during the annealing. **(b)** and **(d)** The corresponding heatmaps demonstrate a cumulative distribution from all the repeats acquired on different days. Inserts depict protein species above 1 MDa. Plots demonstrate the single ring analysis - MW shifts **(e)**, sigma values **(f)**, and the total number of counts **(g)**. The two-sided Mann-Whitney test was performed and the corresponding p-values marked as asterisks or "ns". Each black circle is a single data point, the black diamonds are the outliers.

RAD52(209) at 500 nM produced the highest amount of the dsDNA product in comparison to 100 nM (Figure S8). Therefore, I performed the annealing reaction with 32- and

60-nt ssDNA in the presence of RAD52(209) at 500 nM (Figure 3.17). The experiments were performed in the MFV, therefore, the possibility to detect lower-order oligomers was reduced. The overall trend was similar to RAD52(209) at 100 nM case (Figure 3.16).
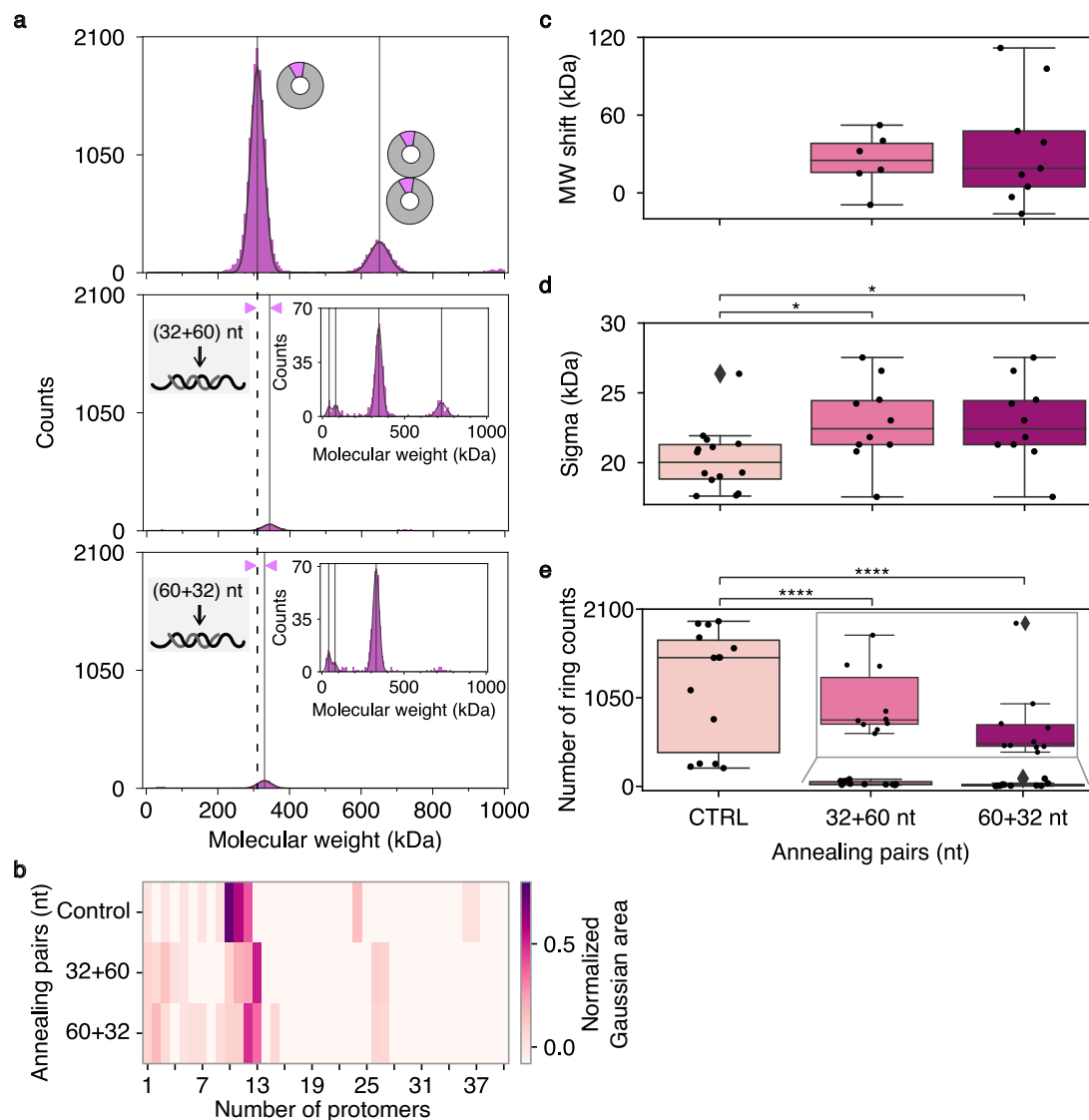
Figure 3.17: Mass histograms during DNA annealing at 500 nM RAD52(209). Annealing pairs were - (32+60)- and (60+32)-nt at 10 nM. All the experiments were performed in the MFV. **(a)** Mass histograms demonstrate a cumulative distribution of several readings acquired on the same day. **(b)** The heatmap shows the normalized Gaussian area corresponding to distinct protomers as the cumulative distribution from all the repeats (N=4). The boxplots depict the analysis of the single ring peak - MW shifts **(c)**, sigma values **(d)**, and the number of counts **(e)**. The data is averaged and plotted as mean values with SEM. Each black circle corresponds to one reading, the filled diamonds are the outliers. The two-sided Mann-Whitney test was performed, and the corresponding p-values are depicted as asterisks or "ns" above the graphs.

However, the reduction of the total number of counts was very drastic (Figure 3.17, a

and Figure S20, b). Single-ring counts for RAD52(209) at 500 nM alone were 1193 ± 185, they reduced to 38 ± 7 and 22 ± 8 during the annealing of (32+60)- and (60+32)-nt DNA, respectively. Upon incubation of RAD52(209) either with the short (32-nt) or along (60-nt) nucleotide with the following addition of the complementary strand, no significant difference in the protein oligomeric state was detected (Figure 3.17, a). As previously observed, the residual single-ring structures had a broad distribution of MW shifts that could not be properly assigned to a particular protein-DNA complex (Figure 3.17, c-d). Interestingly, at these conditions no higher-order oligomers were detected, indicating that there were potential complexes formed with the masses exceeding the upper detection limit of the system. However, it was not clear whether the formation of such massive complexes was essential for the annealing and resulted in a higher amount of the dsDNA product (Figure S8).

To conclude, RAD52(209) demonstrated distinct oligomeric states during the annealing process of the complementary and non-complementary strands. In the case of complementary DNA pairs, single-ring structures were not detected, rather a significant number of higher-order oligomers were present. On the contrary, in the case of mismatching oligonucleotides, the single-ring structures were present together with the reduced number of higher-order oligomers.

3.8 Mass photometry experiments with RAD52

3.8.1 Single-stranded DNA binding

RAD52 binding to the single-stranded DNA was also analyzed using mass photometry (Figure 3.18). The protein concentration used was 100 nM. The measurements were carried out in the mass photometry buffer on the "easy-cleaned" slides. The potential alterations in the oligomeric state of RAD52 were tested with 32-, 48-, and 60-nt DNA molecules at 10 and 100 nM (Figure 3.18). As described previously, RAD52 at 100 nM formed a variety of lower-order oligomers, as well as rings with various numbers of monomers. In the presence of single-stranded DNA, irrespective of its length, ring complexes were shifted towards higher molecular weights (marked with blue arrows), indicating the presence of DNA (Figure 3.18, a). However, unlike RAD52(209), the absence of distinct peaks made the analysis of the MW shifts of rings impossible. Similarly, estimating whether the short oligomers were bound to the single-stranded DNA was challenging. Despite these obstacles, a reduction of rings, either by their number or variety, was observed. The total number of binding events was significantly reduced, irrespective of the DNA length and its concentration (Figure S25). The total binding events for RAD52 at 100 nM alone were 335 ± 46.
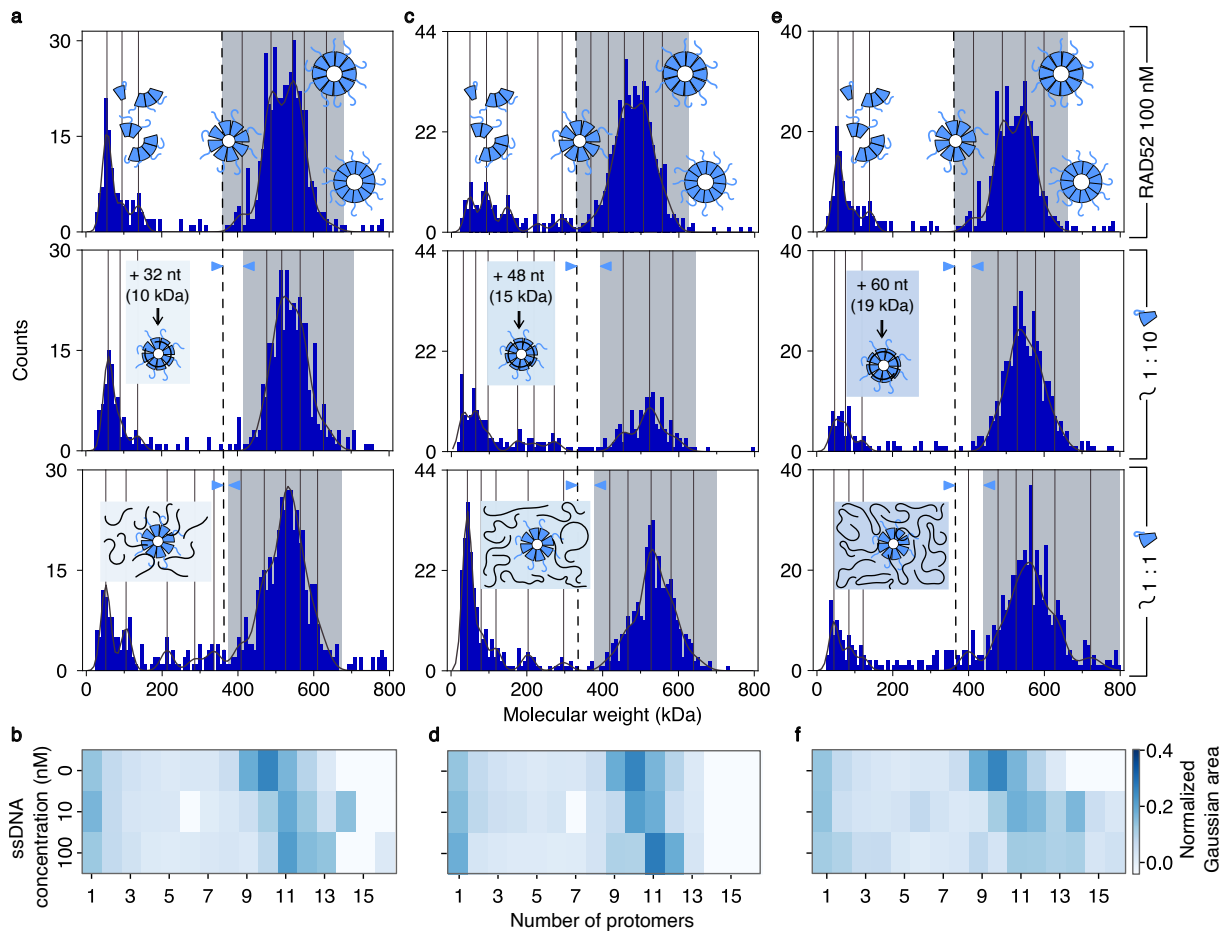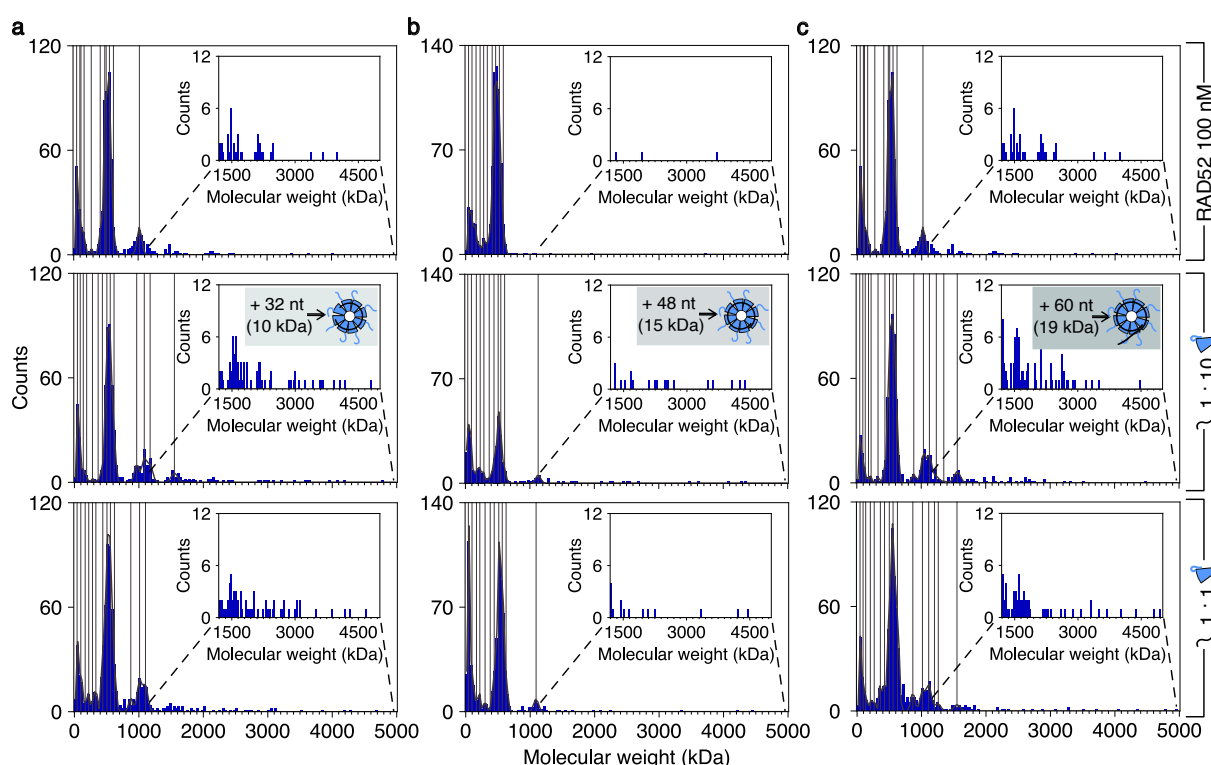
Figure 3.18: RAD52 binding to the single-stranded DNA. DNA molecules were 32-nt **(a-b)**, 48-nt **(c-d)**, and 60-nt **(e-f)** at 10 and 100 nM. The protein concentration was 100 nM. The oligomeric state of the protein alone is shown on the upper panel, followed by the protein incubated with single-stranded DNA at 1:10 and 1:1 ratios. Vertical lines correspond to individual protomers. Gaussian envelopes are depicted as solid black lines. Grey areas indicate the variety of the ring structures. Dashed lines indicate the first ring type formed by RAD52 alone (without DNA). Blue arrows depict the MW shifts. Each mass histogram is a cumulative distribution of a few readings acquired on the same day. The heatmaps corresponding to individual protomers are shown as the cumulative distributions from all the repeats **(b)**, **(d)**, and **(f)**. Schematics of the protein binding to various DNA molecules are shown as blue inserts. The data was acquired in the RFV.

Upon binding to the 32-nt DNA the binding events were 212 ± 36, 105 ± 16 in the case of 48-nt DNA binding, and 172 ± 24 in the case of 60-nt binding. Interestingly, once the protein was mixed with an excess of the DNA in a 1:1 ratio, the number of ring strictures increased (158 ± 24 for 32-, 303 ± 31 for 48-, 164 ± 43 for 60-nt DNA binding experiments) (Figure 3.18, a, lower panel). The number of short oligomers became slightly higher with a strong peak that could correspond either to monomers or to unbound single-stranded DNA (Figure S28). Interestingly, upon mixing GFP-RAD52 with 48-nt DNA at 10:1 ratio, the overall protein species present shifted towards higher masses with the formation

of higher-order oligomers (Figure S5, b). However, the total number of counts was higher once DNA was added, which was somewhat unexpected based on the outcome of the same type of experiment performed on RAD52(209) and RAD52.

RAD52 at 100 nM was also tested with 48- and 83-nt at 200 nM (Figure S23). Upon the addition of an excess of the single-stranded DNA, the number of ring structures decreased. However, the total number of counts was reduced significantly only in the case of 83-nt DNA (Figure S25, c). The appeared peaks could be assigned to protein monomers, dimers, and trimers. However, as has been previously discussed, the lower molecular weight peaks could be mistaken for unbound single-stranded DNA, especially in the case of long (83-nt) DNA. Nevertheless, upon excessive binding of the single-stranded DNA, the oligomeric state was strikingly different from everything observed before. I performed the annealing experiments to reveal any potential correlation between it and the protein annealing activity (Figure S16). In comparison to RAD52(209), RAD52 promoted the annealing slightly even with the massive DNA excess, however, not as efficient as the 10:1 protein-to-DNA ratio that resulted in the highest amount of the DNA product.

Figure 3.19: Higher-order oligomers of RAD52 forming upon single-stranded DNA binding. The DNA molecules were 32- **(a-b)**, 48- **(c-d)**, and 60-nt **(e-f)** at 10 and 100 nM. RAD52 concentration was 100 nM. The upper panel shows the oligomeric state of RAD52 alone, followed by the protein incubated with single-stranded DNA at 1:10 and 1:1 ratios. The inserts present high-order oligomers above 1 MDa. Schematics of the protein binding to various DNA molecules are shown as blue inserts. Each mass histogram is a cumulative distribution of a few readings collected on the same day. The data was acquired in the RVF.

60

Notably, the total number of counts for RAD52 was significantly lower than for RAD52(209) at similar concentrations (Figure 3.2). During the data acquisition, several massive protein complexes were visible on the setup, however, their molecular masses were above the detection limit (Figure S26). Additionally, the mass histograms acquired in the RFV revealed that RAD52 only formed structures above 1 MDa (Figure 3.19). The presence of the single-stranded DNA increased the number and the molecular masses of these higher-order oligomers. Remarkably, while RAD52(209) formed higher-order oligomers at 3-6 MDa that were easily segregated from single rings and stacks thereof, the high-oligomers of RAD52 appeared more like filamentous structures. Due to the insignificant number of counts and their continuity, it was not possible to quantify properly the exact molecular masses of these structures. However, this observation agrees with the latest cryo-EM structure of RAD52 bound to the double-stranded region of the DNA replication fork [54]. The obtained structure shows a RAD52 double-ring wrapped around dsDNA. Potentially similar protein structures can be formed on the ssDNA as well.

As previously described, negatively charged DNA in a complex with the protein could repel from the glass surface of the microscopy slide. To avoid this artifact, RAD52 binding to 60-nt DNA was tested in the buffer containing Mg^{+2} and on modified, positively-charged APTES slides (Figure 3.20). In the presence of Mg^{+2}, RAD52's oligomeric state was similar to previously described experiments, as well as upon DNA binding (Figure 3.20, a-b). The protein alone exhibited a broad distribution of short oligomers and a heterogeneous pool of rings. Moreover, the monomers appeared to be the dominant species and in this case, they could not be mistaken with DNA molecules, as none were added. Upon binding to the single-stranded DNA with a 1:10 ratio, the total number of counts as well as the variety of the protein species drastically reduced (494 ± 102 for RAD52 alone and 211 ± 24 with 60-nt DNA bound) (Figure S25, e). The remaining fraction of the ring structures was shifted towards higher molecular masses. Once the protein was mixed with a saturating amount of 60-nt ssDNA (1:1 ratio), the ring structures reappeared, suggesting the DNA excess dissociated some higher-order oligomers. The distribution of the short oligomers was roughly the same. The total number of counts was 348 ± 38. Overall, analogous trends were observed with APTES slides (Figure 3.20, d-e). The total number of RAD52 counts at 100 nM was 380 ± 39. Upon binding to the 60-nt DNA in a 1:10 ratio, the total number of binding events decreased to 226 ± 29. However, upon DNA excess, the total number of counts increased above the control level - 458 ± 67 suggesting that DNA caused dissociation of some high molecular weight protein structures that were not detected previously. The most significant difference between the previously described results and the APTES slide experiments was the higher amount of detected short oligomers. In the mass photometry buffer at 100 nM RAD52 primarily formed ring structures with a detectable, but rather small, fraction of short oligomers (Figure 3.18). While in the presence of Mg^{+2} and/or on the positively-charged slides, the number of detected short oligomers was visibly higher. However, it is important to mention that the

acquisition of RAD52 species can be improved by utilizing the MFV (Figures S24, a and c), at the price of losing the distribution of the monomers.
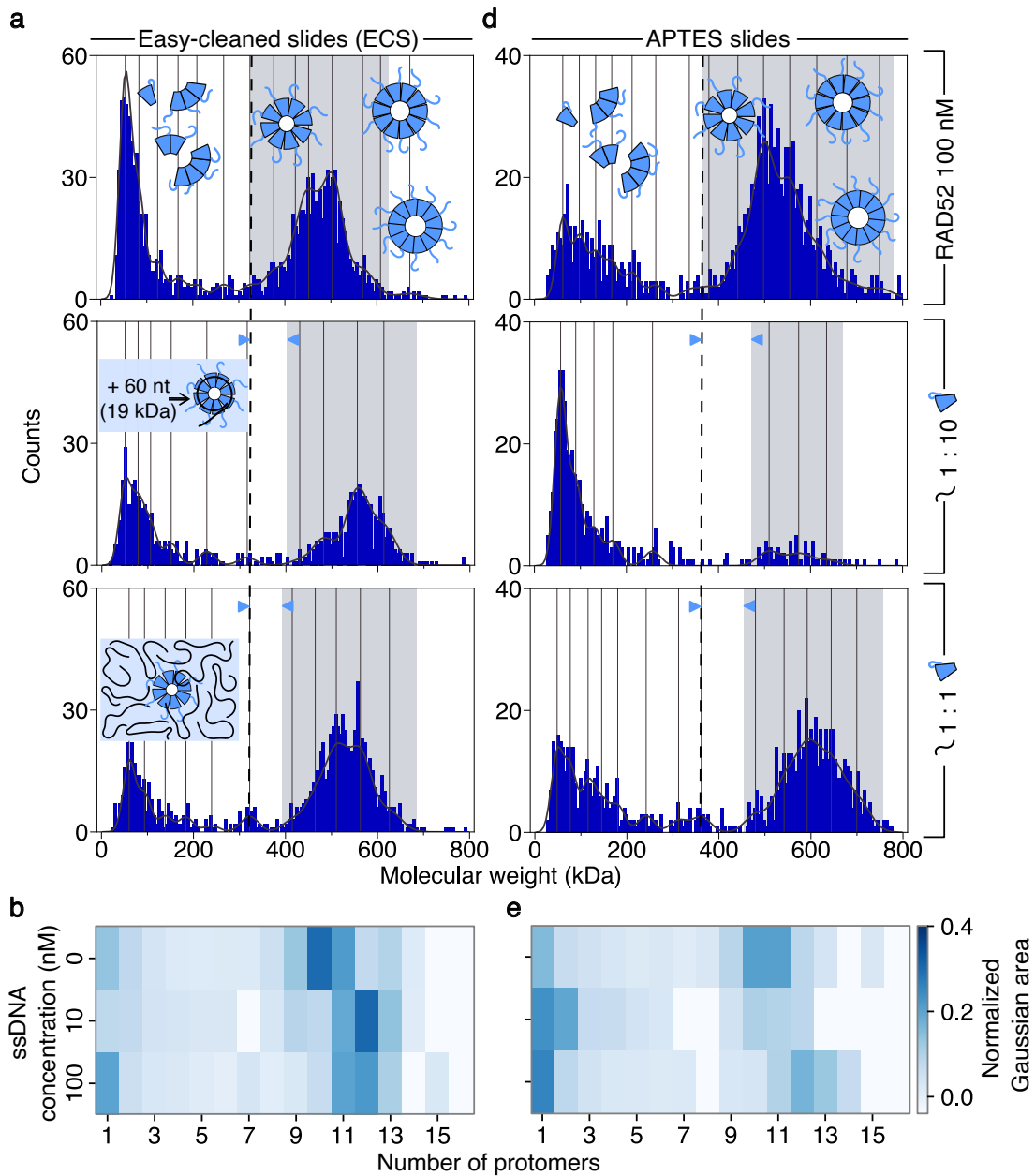

Figure 3.20: RAD52 binding to single-stranded DNA in the annealing buffer. The 60-nt DNA was utilized at 10 and 100 nM concentrations. The protein concentration was 100 nM. The experiments were performed on the easy-cleaned slides (ECS) **(a-b)** and APTES slides **(c-d)**. Each mass histogram is a cumulative distribution of a few readings from the same day. Blue arrows demonstrate the MW shifts of the ring structures (dashed lines are the control references). The corresponding heatmaps are shown on the lower panel. Schematics of the protein binding to DNA at different concentrations are shown in the blue inserts. The data was acquired in the RFV.

The temperature effect was tested in the annealing buffer at 37°C on the easy-cleaned

slides (Figures S24). Nevertheless, the higher temperature facilitated the formation of higher-order oligomers, as the variety of the detected protein species reduced (Figure S25). Apart from the expected protein titration due to longer incubation at higher temperatures, the above-described trends were present. Notably, the same experiments were performed in the MFV to acquire better statistics (Figure S24, c). The total number of counts doubled, however, no additional structures were detected, confirming previous observations. Importantly, due to the lower limit of the mass detection, the dominant peak within the short oligomers belonged to the dimer, not the monomer as detected before.

In summary, RAD52, irrespective of the buffer or type of the glass slide used, formed a variety of ring structures as well as a broad distribution of lower-order oligomers. Therefore, the quantification of the molecular weight shifts caused by the DNA binding was scarcely possible. Nevertheless, it was detected that upon DNA binding, the ring structures shifted towards higher molecular masses. However, the correlation between these shifts and the DNA length or its concentration was not observed. The total number of binding events dropped at all tested DNA lengths once the single-stranded DNA was added to the protein mixture in a 1:10 ratio. Resulting higher-order oligomers, resembling protein filaments, were detected in the 1-5 MDa range. Moreover, a certain number of large protein structures were observed during these experiments with the molecular masses above the detection limit of the mass photometry setup. The excess of the DNA appeared to be able to dissociate some of the higher-order oligomers which resulted in a slight increase of the number of single rings.

3.8.2   Single-stranded DNA annealing

The annealing reaction promoted by RAD52 was examined with 32- and 60-nt complementary strands at 10 nM and 100 nM of the protein (Figure 3.21). According to the annealing gels, at these DNA-to-protein ratios, the amount of newly formed dsDNA product was higher than the intrinsic annealing (Figure S8). RAD52 was preincubated with 60-nt ssDNA and the complementary strand was subsequently added. As previously observed, upon binding to the single-stranded DNA, a variety of higher-order oligomers above 1 MDa were formed (Figure 3.21, c). Interestingly, unlike the excess of single-stranded DNA, the addition of the complementary strand resulted in the further formation of higher-order oligomers with higher masses. Potentially there were super-structures formed with molecular weights above the detection limit of the system, as the variety and number of short oligomers and single rings almost diminished after the addition of the complementary DNA (Figure 3.21, a). This effect could be explained by the interaction of RAD52 and newly synthesized double-stranded DNA, similar to how it was observed for RAD52(209).
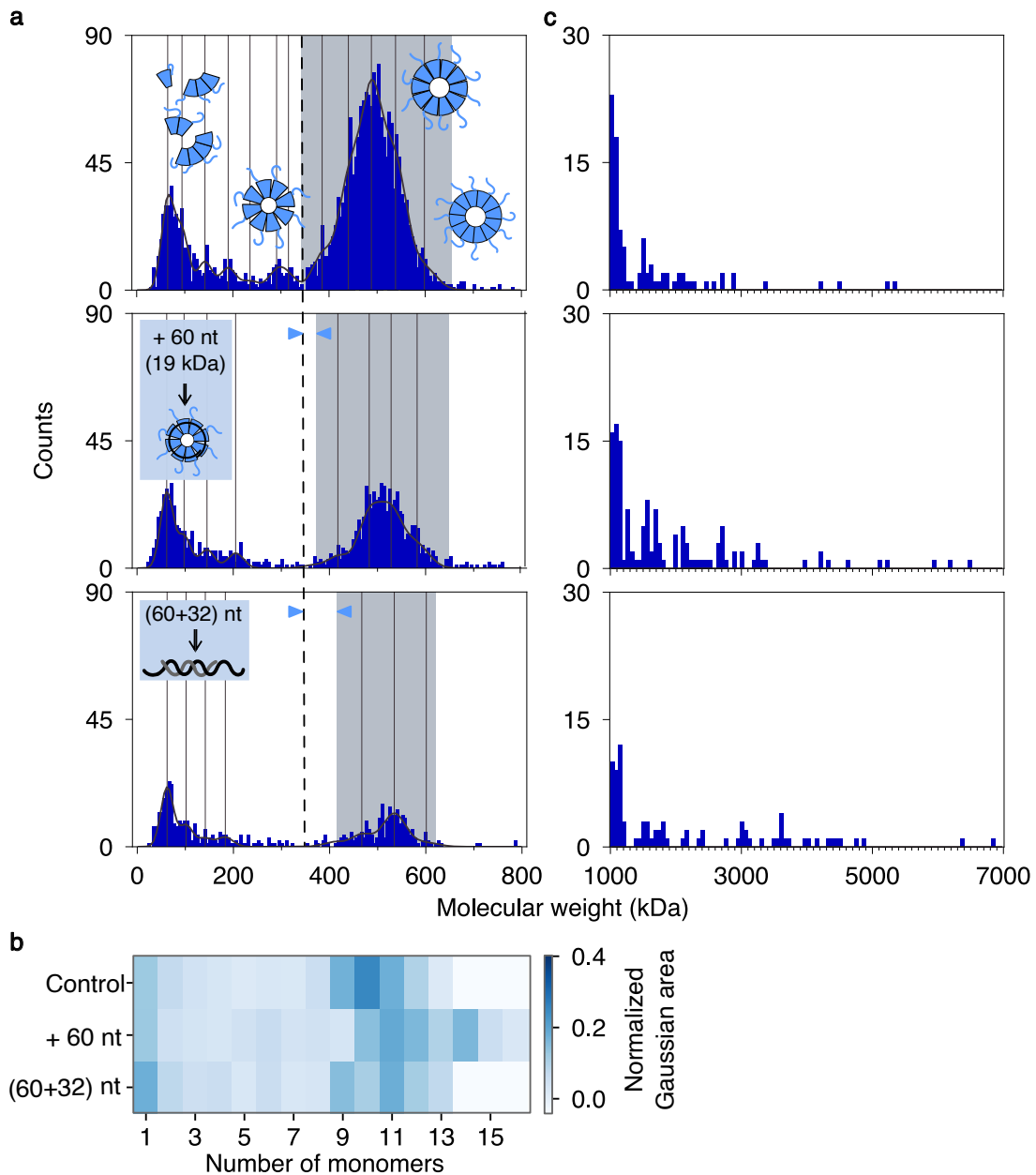
Figure 3.21: Mass histograms of RAD52 the annealing reaction. The annealing pair was 60 and 32 nt, at 10 nM. The protein concentration was 100 nM. **(a)** Exemplary mass histograms are a cumulative distribution of a few readings acquired on the same day. Blue arrows demonstrate the MW shift of the ring structures and the inserts describe the steps of the annealing reaction **(b)** The heatmaps contain the data from several repeats (N=8). **(c)** The high-order oligomers are detected at each step of the annealing reaction.

3.9   Closure

The extensive analysis of the full-length RAD52 and its truncated variant RAD52(209) demonstrated that both proteins bind single-stranded DNA in a cooperative manner most

of the time, with the cooperativity depending on the incubation temperature and the DNA length (Table 3.1 and S3).

Table 3.1: Table presents an overview and a comparison of RAD52(209) and RAD52 single-stranded DNA binding and annealing activities. Single-stranded DNA binding was tested on different DNA molecules smaller (32-nt) or bigger (60- and 70-nt) than the RAD52(209) ring's circumference (44-nt). The detailed analysis of the binding constants is presented in Table S3. The efficiency of the single-stranded DNA annealing was estimated *via* the reaction rate and the increase in the amount of the newly-formed double-stranded DNA product upon utilizing different incubation strategies. The detailed analysis of the reaction rates is shown in Table S4.

		RAD52(209)	RAD52
Single-stranded DNA binding			
Temperature (T°C)	DNA length (nt)		
25	32	Cooperative	Cooperative
	70	Cooperative	Non-cooperative
37	32	Cooperative	Cooperative
	60	Cooperative	Cooperative
Single-stranded DNA annealing			
Reaction rate increase AB		4 times	72 times
Reaction rate increase NT		7 times	80 times
Annealing strategy:			
1. P + 1 ssDNA + 2 ssDNA		Weak annealing	Weak annealing
2. (P + 1 ssDNA) + 2 ssDNA		Strong annealing	Strong annealing
3. (P + 1 ssDNA) + (P + 2 ssDNA)		No annealing	Weak annealing

The presence of the disordered C-terminal domain did not impair the protein's ability to promote the annealing reaction. On the contrary, at certain experimental conditions, the presence of the full-length protein resulted in a higher amount of the newly synthesized DNA product than in the case of RAD52(209). Both proteins facilitate the annealing at most, once one of the single strands is preincubated with the subsequent addition of the complementary strand (the second strategy, Table 3.1). Both proteins demonstrated diminishing annealing activity at micromolar concentrations. The annealing activity was not dependent on the DNA length or the DNA concentration for RAD52. However, RAD52(209) demonstrated a slight correlation between the DNA concentration and the protein concentration promoting the annealing reaction at most.

Table 3.2: The table presents the oligomeric state of RAD52(209) and RAD52 measured *via* mass photometry. Proteins were examined at different concentrations (10, 20, 50, 100, and 200 nM) alone and in the presence of single-stranded DNA. DNA molecules were 32-, 48-, 60-, and 83-nt long at 10, 100, and 200 nM. For the single-stranded DNA binding experiments, the protein concentration was 100 nM. The oligomeric state is indicated as a fraction of protein species. The fraction of short oligomers, including monomers, were collected up to tetramers. The ring structures of RAD52(209) contained undecamers, and RAD52 ring distribution was combined from 8 to 12 protomers.

Oligomeric state of the protein alone				
Protein fraction (%)	RAD52(209)		RAD52	
Protomers number	1 - 4	11	1- 4	8 - 12
Protein concentration (nM)				
10	38	52	83	16
20	8	63	64	34
50	12	66	34	63
100	4	72	27	71
200	5	67	26	70
Oligomeric state upon single-stranded DNA binding				
DNA length (nt) DNA concentration (nM)				
32 10	23	61	33	52
100	11	85	29	47
48 10	20	69	50	43
100	37	57	31	62
200	52	46	47	49
60 10	21	70	30	47
100	11	81	34	34
83 10	10	75	27	56
100	25	66	38	56
200	47	35	74	24

Despite similarities in the annealing activity, the mass photometry results revealed striking differences in the oligomeric states of the protein. While RAD52(209) appeared as an undecameric ring with the stacks thereof and a small fraction of short oligomers, the full-length protein exhibited a heterogeneous pool of ring structures with different numbers of subunits and a significant amount of short oligomers (Table 3.2).

The experiments performed on RAD52(209) and RAD52 upon binding to the single-stranded DNA revealed the diversity of the interaction modes with DNA. Upon binding to the single-stranded DNA in a 1:10 ratio, both proteins formed high-order oligomers. RAD52(209) formed complexes at around 3-5 MDa, while RAD52 demonstrated high-order oligomers with a certain periodicity, similar to a filament formation as if multiple

protein rings assemble on the DNA. The excess of the single-stranded DNA (1:1 protein-to-DNA ratio) caused the dissociation of these oligomers, and even single rings, giving rise to short oligomers including monomers. It is important to remember, that upon DNA binding at 10:1 protein-to-DNA ratio, the total number of binding events was significantly reduced. Therefore, the obtained protein fractions with similar percentages should be compared with caution as they correspond to different numbers of counts prior to and after the DNA binding.

Table 3.3: The table presents the oligomeric state of RAD52(209) during the double-stranded DNA binding and annealing. The protein concentration was 100 in all the experiments. For the annealing experiments four pairs of single strands were utilized - complementary strands were (32+60)- and (60+32)-nt, and non-complementary pairs were (32+83)- and (60+83)-nt. The DNA concentration was 10 nM. For the double-stranded DNA binding experiments pre-formed dsDNA were utilized - (48+48)- and (48+60)-nt at 10 and 100 nM. The fraction of short oligomers, including monomers, were collected up to tetramers, and the ring structures of RAD52(209) due to the broad distribution of species were collected from 11 up to 16 protomers.

Double-stranded DNA binding			
Protomers number		1 - 4	11-16
DNA length (nt)	DNA concentration (nM)	RAD52(209) fraction (%)	
48+48	10	32	13
	100	51	37
48+60	10	27	65
	100	62	29
Single-stranded DNA annealing			
Annealing pair (nt)	DNA concentration (nM)	RAD52(209) fraction (%)	
32+60	10	40	44
60+32	10	27	58
32+83	10	16	71
60+83	10	10	81

Due to the single-ring structure formed by RAD52(209), it was possible to study the annealing process as well as the double-stranded DNA binding (Table 3.3). Upon annealing the complementary strands, irrespective of the incubation order, the total number of counts significantly decreased giving rise to high molecular weight species. Moreover, the MW peak of the RAD52(209) single-ring structure appeared much broader. Therefore, I analyzed the annealing process on the complementary and non-complementary strands at different time points (Figure S31). Based on the MW shifts, one could see that there was a constant association and dissociation of single strands on the single-ring structure. Interestingly, the first strand that was preincubated with RAD52(209) appeared to stay attached to the ring over 9 minutes, while the second strand was binding transiently. The

presence of a complementary DNA strand further facilitates the formation of massive superstructures that could not be analyzed due to the upper-resolution limit of the mass photometry setup. Despite that, the mass photometry results brought a lot of information that is essential for a better understanding of the DNA annealing mechanism by RAD52.

Chapter 4

Discussion & Conclusion

Human RAD52 is an important player in various cellular pathways, such as DNA DSB repair and the release of stalled replication forks. RAD52 participates in distinct sub-pathways within the DNA DSB repair, including HR via RAD51, SSA pathway, and RNA-mediated DNA DSB repair [80, 40, 86]. However, despite numerous attempts, the exact mechanism of RAD52 core function - DNA single-strand annealing, remains obscure. In this work, combining the biochemical assays with the single-molecule approach, I investigated the protein's oligomeric state under various conditions including broad protein concentration ranges, temperatures, presence of bivalent ions, and, most importantly, supplemented with various DNA molecules that were added to the reaction volume following different incubation strategies. The obtained results brought new evidence about the oligomeric state of the full-length protein at nanomolar concentrations and upon DNA binding, explained distinct results present in the literature, and, at last, gave an important insight into the process of the single-stranded DNA annealing by RAD52.

4.1 RAD52(209) is an undecamer

In order to understand the annealing process performed by human RAD52, the majority of research groups have used its truncated version - the N-terminal domain RAD52(209) [60, 43, 123]. Crystallographic studies have determined that RAD52(209) forms symmetrical ring structures that contain 11 subunits - an undecamer [62, 130]. Additionally, it has been demonstrated that the rings of the N-terminal domain stack and form double and triple rings [123]. The mass photometry experiments performed on RAD52(209) confirmed that the protein exists as an undecameric ring and, upon increasing protein concentrations, stacks of these rings were detected. Additionally, a small fraction of lower-order oligomers and, very seldomly, monomers were observed. This was an important insight

into the oligomeric state of RAD52(209), previously unreported. Importantly, monomers and lower-order oligomers of RAD52(209) cannot be identified by conventional imaging techniques like scanning electron microscopy or cryo-electron microscopy, due to their small molecular weight and the high protein concentrations required for these types of imaging.

4.2   RAD52 is not a homoheptamer

The full-length RAD52 additionally carries RPA and RAD51 recognition sites, as well as a nuclear localization sequence in its C-terminal domain [104, 127, 128]. Due to its highly disordered structure, crystallization was known to be impossible until recently [123]. In 2023, two research groups independently obtained crystal structures of the full-length RAD52 [68, 6]. The main part of the crystal contained the undecameric N-terminal domain and the C-terminal domain was reconstituted only partially with very poor resolution. One research group pointed out that while undecameric rings were the dominant structures observed and chosen for further analysis, there were other structures present [68]. Until now, the major knowledge about RAD52 structure comes from the TEM images introducing RAD52 as a heptameric ring [136]. These studies also demonstrated images of rings containing various numbers of subunits (6-10). Moreover, hydrogen peroxide treatment of RAD52 bound to a single-stranded DNA demonstrated a 4 nt pattern going up to 40 nt, suggesting that the protein could form decameric complexes [105]. Putting all this evidence together, the exact structure of the full-length protein remains a mystery.

Results obtained in this study have shed light on this problem. They revealed a broad distribution of species with a main molecular weight peak corresponding to a decameric (10) ring. Most importantly, based on the analysis of the obtained sigma values from the Gaussian fits of the mass histograms, it appeared that the full-length is able to form ring structures with different numbers of subunits (6-13). Strikingly, at low nanomolar concentrations, RAD52 exhibited predominantly in the monomeric and dimeric forms with a small fraction of rings. Contrary to RAD52(209), the full-length protein exhibited only weak stacking of rings which could be explained by the heterogeneity of the rings and their distorted structures. Additionally, the number of binding events of RAD52 was significantly lower in comparison to RAD52(209). During the experiments with RAD52 alone, the formation of high molecular weight structures was observed in agreement with previous studies reporting clustering of the protein [149, 118, 68]. These structures could be condensates, as it is known that disordered domains lead to cluster formation and can promote aggregation or liquid-liquid phase separation [7, 110]. Studies on yeast Rad52 have recently revealed the formation of large condensates, driven by the disordered C-terminal domain [31]. Liquid-liquid phase separation of the yeast Rad52 has also been

observed *in vivo* [102]. Therefore, despite lacking quantitative analyses of the observed condensates, the obtained mass photometry results have confirmed the C-terminal domain of human RAD52 induced the formation of large clusters that could potentially be liquid-liquid phase separation or aggregation.

The analysis of the oligomeric state of RAD52 and RAD52(209) brought unexpected, yet exciting, results. While confirming precisely the undecameric nature of RAD52(209), mass photometry revealed a broad distribution of rings rather than homoheptameric rings of RAD52. Additionally, it was possible to detect monomers of RAD52 and lower-order oligomers formed by both proteins at nanomolar concentrations. The diversity in the literature in regard to the ring structures can be explained by the ability of RAD52 to form rings with different numbers of subunits, while the decameric (10) and undecameric (11) rings were the dominant species. Moreover, in order to reconstitute the full-length ring, an undecameric structure of the N-terminal domain was utilized in the cryo-EM studies, potentially imposing the 11-fold symmetry [69]. Nevertheless, the ring diversity could imply that the exact number of subunits is not essential for the annealing activity. Moreover, the presence of a significant fraction of RAD52 lower-order oligomers as well as high molecular weight clusters suggests that these species could also participate in the annealing mechanism.

4.3 DNA concentrations affect the protein annealing activity, while the DNA lengths do not

Taking into consideration the significantly different oligomeric states of RAD52 and RAD52(209), one could expect that these differences would affect the annealing reaction. Despite that, the overall annealing activity of both proteins was similar. Interestingly, RAD52 appeared to facilitate the annealing more efficiently at lower concentrations than RAD52(209). The annealing might be facilitated by RAD52 lower-order oligomers apart from the main mechanism that would be suitable for both proteins. In agreement with previously published studies, at high concentrations, both proteins failed to facilitate the annealing [130, 82, 63].

I tested different factors that could potentially affect the outcome of the annealing reaction. It has been shown that the annealing activity of RAD52 bacterial homolog - Redβ, is capable of only promoting the annealing of DNA strands above 16-20 bp, meaning that for a stable DNA-protein complex, at least a tetramer is required [33, 97]. Interestingly, both RAD52 and RAD52(209) were able to anneal DNA strands as short as 12-16 nt. Taking into consideration that RAD52's footprint is 4 nt, this would imply that a trimer or a tetramer is sufficient to initiate or promote the annealing reaction [105]. The cooperativity constants of both proteins binding to single-stranded DNA were also in a

range between 3 and 4, implying that either a trimer or a tetramer is able to immobilize the DNA, or binding to a trimer-tetramer within the ring structure is sufficient for DNA attachment. Following this lead, short DNA strands should be faster to bind and, consequently, to anneal. However, there was no correlation between the DNA length and the amount of the newly synthesized DNA product. RAD52 did not reveal any dependence on the DNA concentration, only RAD52(209) demonstrated a slight correlation. These results indicate that once the minimum required number of protomers is bound to the DNA, the annealing is promoted, irrespective of the DNA length. Both proteins formed trimers, however, only very few were detected. Therefore, potentially a trimer or a tetramer is essential to establish the initial contact on a single ring and secure the DNA localization. Furthermore, an Alphafold-predicted structure, using ColabFold, of RAD52 and RAD52(209) resembles the ring circumference suggesting that, unlike Redβ, both proteins tend to form ring structures rather than filaments (Figure 4.1) [90]. However, these predictions have to be confirmed experimentally. To this date, RAD52 filaments have not been crystallized, nor have RAD52 tetramers or trimers.
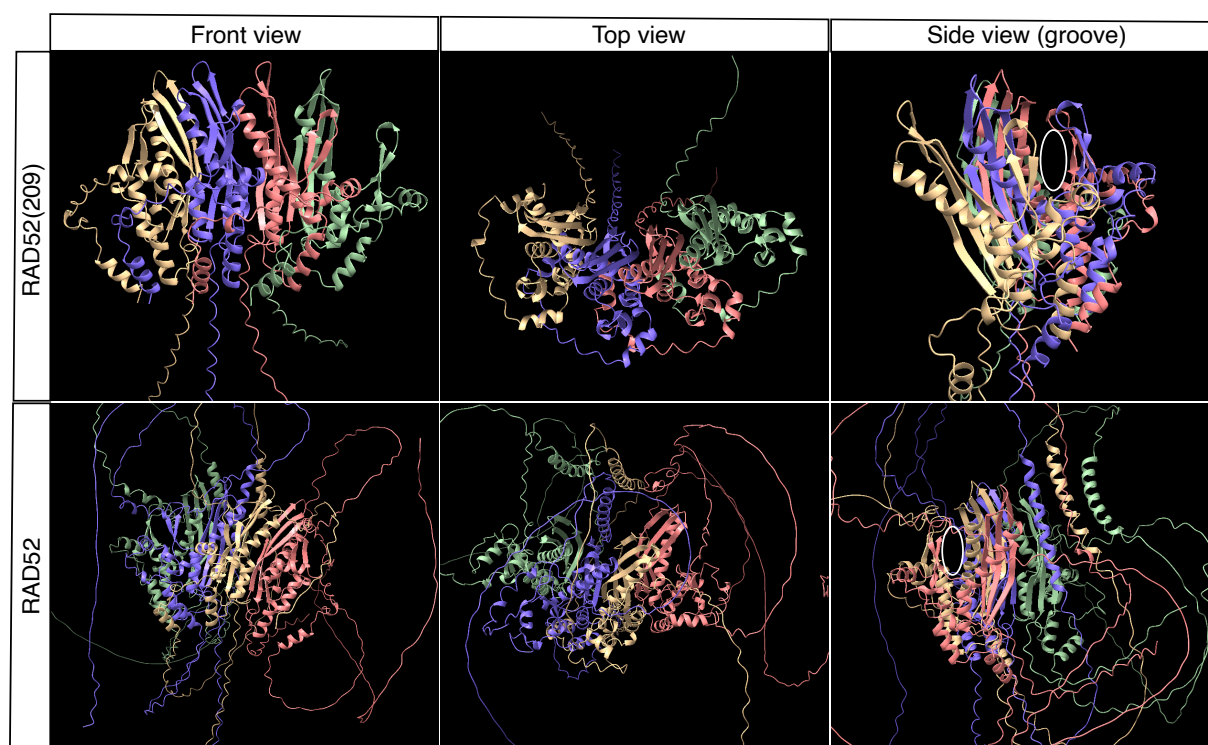
Figure 4.1: AlphaFold-predicted structures of RAD52(209) and RAD52 tetramers. Each structure is shown from three views - front, top, and side view depicting the first DNA binding site (white oval). Monomers are color-coded in yellow, violet, pink, and magenta. The AlphaFold prediction was performed by Dr. Paul Gouguet. The images were colorized and processed in ChimeraX 1.6.

4.4 The incubation strategies influence the annealing outcome

Similar to Redβ and RecT, RAD52 and RAD52(209) facilitate the annealing reaction only once the protein was preincubated with one strand prior to the addition of a complementary strand. The inhibition of the annealing at high protein concentrations could be verified *via* native annealing gels (Figure 3.9). At high protein concentrations, the stacks of rings were formed, in the case of RAD52(209), or higher-order oligomers of RAD52. Once the protein is preincubated with the first strand, it forms higher-order oligomers that are visible with mass photometry. But, most importantly, there is a significant amount of the protein still present in the reaction volume that was not bound to DNA. Once the second complementary strand was added, it might associate with these free protein complexes and, therefore, become hindered from completing the annealing reaction. However, another puzzling point is what is an explanation of the annealing inhibition at the lowest tested protein concentrations, where both proteins have not demonstrated any high-order oligomers? The answer might be hidden in the altered oligomeric state of the protein upon ssDNA binding. Even at such low protein concentrations, ssDNA binding triggered the formation of high-order oligomers of both RAD52 and RAD52(209). Due to the insignificant number of counts, it was not possible to characterize these oligomers. However, the potential explanation for the annealing inhibition might be similar to the case with very high protein concentrations.

4.5 DNA binding alters the protein oligomeric states

Upon binding to single-stranded DNA, independently of its length, both RAD52 and RAD52(209) demonstrated the formation of higher-order oligomers containing single rings and stacks thereof. These results were obtained from what was determined as the most favorable annealing conditions, suggesting that the formation of higher-order oligomers is essential for the annealing reaction. Moreover, it has been explicitly shown that both human RAD52 and yeast Rad52 form a small number of foci to accumulate and process multiple DNA breaks simultaneously [78, 76]. My results demonstrated that the DNA excess caused the dissociation of these higher-order oligomers as well as some of the single rings giving rise to lower-order oligomers. The annealing gels performed at these DNA-to-protein ratios confirmed that the annealing was drastically diminished. The Silver Stain DNA-to-protein binding assays clearly demonstrated that with an excess of the DNA, there was a clear DNA band running along with the protein ring structures for both proteins. However, no clear bands corresponding to the DNA bound to lower-order

oligomers were resolved. Only a smear appeared in the case of RAD52. This could indicate that single-stranded DNA strongly binds to the ring structures and only transiently to the shorter oligomers.

It is known that Mg^{+2} facilitates oligomerization of the protein and affects DNA-to-protein interaction. However, the presented mass photometry results indicate that magnesium had no to very little effect on RAD52's oligomeric state, while RAD52(209)'s oligomeric state changed drastically. As expected, the protein formed more stacks of rings and higher-order oligomers that dissociate upon single-stranded DNA binding, giving rise to single rings. This behavior is opposite to what was observed in the buffer without magnesium, yet in both buffers, RAD52(209) successfully performed the annealing reaction. However, these deviations could arise from the fact that binding of the single-stranded DNA to a single ring of RAD52(209) gives an overall negative charge to this complex, and, as a result, it repels from the negatively-charged glass surface. Magnesium binds to DNA and facilitates its binding to the surface, including the DNA molecules that are bound to the protein. In other words, the protein-DNA complexes are always present in solution, however, without magnesium, they could not bind to the glass and, therefore, remained undetected as they are not present in the focal plane. Although, it does not explain why, following the same hypothesis, there was no significant difference observed for RAD52 in buffers with and without magnesium. In the case of RAD52, its highly disordered C-terminus might hinder single-stranded DNA from the surface, therefore, the binding of such a protein-DNA complex was not affected [6].

4.6 A single ring of RAD52(209) can accumulate two single-stranded DNA

Remarkably, RAD52(209) can accumulate up to two ssDNA molecules on a single ring simultaneously. This was observed independently of the DNA-to-protein ratio, buffer composition, incubation temperatures, and, in most cases, independently of the length of the DNA molecules. The only exception was the 83-nt strand that rarely was bound as a single DNA on the same ring. However, most importantly, these bound DNA molecules were identical and non-complementary. Thus, after the preincubation of the protein with one of the strands, both DNA binding sites of a single ring are occupied by the two identical DNA molecules. This observation challenges the annealing model utilizing a single-ring structure as the key species of the annealing process. Moreover, in the case of RAD52(209) at 500 nM incubated with 32- or 60-nt, the mass photometry results clearly showed that no single-stranded DNA was bound to single rings. Only in the case of the longer strand (83-nt), the shift was observed. On the other hand, in the experiments with RAD52(209) at 500 nM, upon binding to the single-stranded DNA the

total number of counts was reduced drastically implying that only an insignificant fraction of the single rings that were not bound to the DNA remained and the rest of the protein formed a complex with MW above the detection limit. This could be an additional hint that the formation of higher-order oligomers could be essential for the annealing process. And, potentially, an accumulation of DNA strands of the single-ring structures could be simply a mechanism to ensure the formation of these massive complexes to bring all the DNA molecules in close proximity while increasing the local protein concentration. Additionally, the oligomeric state of RAD52(209) during the annealing of ssDNA indicated the formation of higher-order oligomers and the almost complete absence of single rings. The absence of the single rings could be explained by their interaction with the newly-formed DNA product. However, upon binding to the pre-formed double-stranded DNA with overhangs, a significant fraction of single rings was still detected. Therefore, different experiments pointed to the same conclusion that the disappearance of single rings along with the formation of higher-order oligomers must be a signature of the annealing reaction. Even in the case of RAD52, despite the complexity of the analysis, the annealing reaction revealed the formation of higher-order oligomers together with a low amount of detectable single-ring structures during annealing.

It has been observed that yeast Rad52 undergoes liquid-liquid phase separation *in vivo* [102]. Preliminary single-molecule TIRF experiments confirmed mass photometry results and demonstrated that at similar protein concentrations both RAD52 and GFP-RAD52 form clusters in the presence of single-stranded DNA (Figure S27). Additionally, the possibility that these clusters could be functional condensates essential for the annealing was supported by the fact that the annealing activity of both RAD52 and RAD52(209) was drastically reduced in the presence of increasing concentrations of HD. Unfortunately, it was not possible to determine the oligomeric state of the protein in the presence of HD *via* mass photometry. One of the possible explanations of this observation might be due to the fact that HD carries negatively-charged -OH groups and, upon binding to the protein, it prevents its binding to the glass surface which can refrain some of the -OH groups after the easy cleaning procedure. Therefore, this observation requires further experiments. Until then, there is a possibility that human RAD52 undergoes liquid-liquid phase separation or aggregation during the annealing process.

4.7 Alternative models of the DNA annealing by RAD52

In one of the early studies, a mutant of RAD52(85) that did not contain the second DNA binding site was analyzed [82]. The gel filtration profile revealed its ability to form a heterogeneous pool of oligomers with different numbers of protomers without a dominant type of species [82]. With a massive excess of the protein over the DNA, this mutant bound ssDNA and promoted the annealing reaction very poorly. This observation implies

that a ring structure is not essential for the DNA binding and annealing but rather reflects the potential of the protein to self-oligomerize. This hypothesis is supported by the alanine mutants demonstrating the absence of the ring formation and a weak, but still detectable, ability of these mutants to bind the DNA [83]. Interestingly, these mutants (F79A, Y81A) had a similar gel filtration profile to RAD52(85) with many short oligomers and no dominant structure. The results of both studies indicate that RAD52 alone is able to form filament-like structures, similar to RAD51, and not only the rings [31]. However, the filament-like oligomeric state of the protein demonstrates a significantly lower propensity of binding and annealing DNA molecules. Nevertheless, considering the fraction of short oligomers of RAD52, it is possible that these species facilitate the annealing reaction alongside other protein structures (Figure 4.2, b). If this is the case, the mechanism of the annealing and the homology proofreading could be similar to Redβ [33, 4]. According to this model, Redβ dimers probe the DNA for the homology search *via* a weak interaction. This interaction allows the protein to dissociate quickly from the DNA in case the homology is not established. Moreover, the dimer binds 8 bp which is not yet a genome-unique sequence.
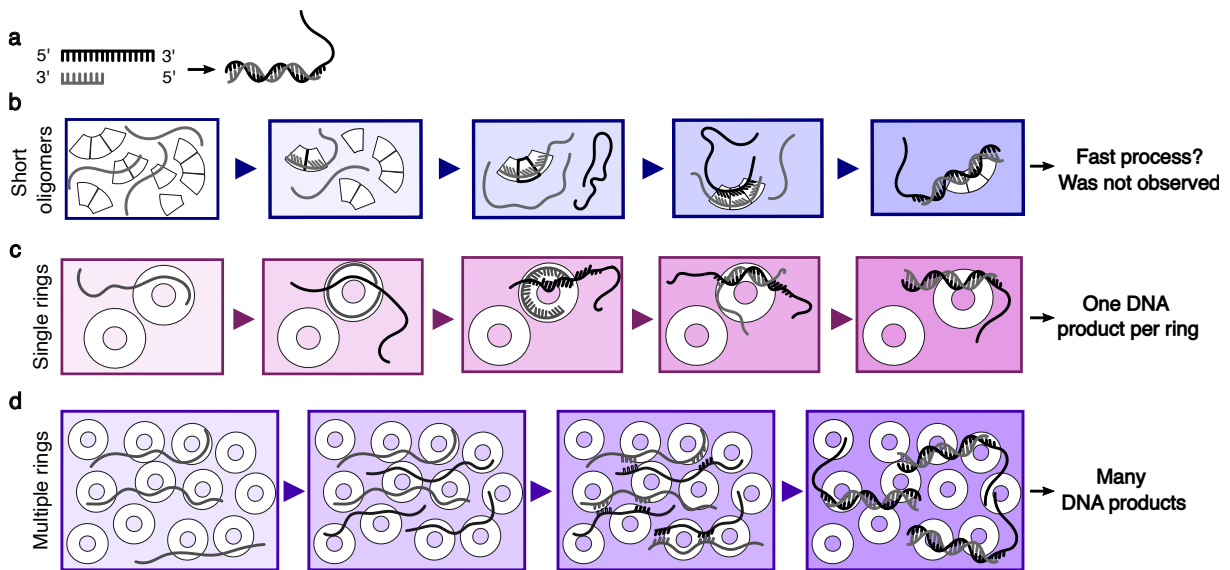
Figure 4.2: Alternative models of the DNA annealing by RAD52. Complementary strands are depicted in black and grey **(a)**. **(b)** The annealing mechanism *via* short oligomers of RAD52. **(c)** The annealing could be facilitated *via* single ring structures. **(d)** Many ssDNA could be paired on the multiple rings.

However, once another monomer or a dimer binds to the strands securing 12-16 bases that are sequence-unique, Redβ trimer-tetramer locks the DNA and initiates the annealing reaction [4]. In analogy, a trimer of RAD52 binds to single-stranded DNA. Upon binding to the DNA, the protein trimer undergoes conformational changes that cause DNA to bind to the inner groove of the complex, where it is securely locked and extended into

B-form DNA exposing the bases for the homologous pairing (Figure 4.2, b) [123]. Once a complementary strand binds to the outer side of the stem, the formation of hydrogen bonds between the complementary strands gives sufficient energy to release the dsDNA from the protein trimer. A trimeric complex binds 12 nt which is sufficient to ensure homologous sequence recognition, establishing proofreading. Cancer-research studies utilizing chemical compounds aiming to impair RAD52 annealing activity by occupying the DNA binding site, demonstrated that upon binding of these compounds, RAD52 rings dissociated into dimers that were unable to promote the annealing reaction [22, 144, 49]. Additionally, the monomers and dimers of RAD52 and RAD52(209) observed with an excess of single-stranded DNA failed to facilitate the annealing reaction (Figure S15 and S23). However, there are a few points that need to be clarified regarding this annealing mechanism. Firstly, according to the mass photometry results, trimers of RAD52 were present only above 50 nM, while for RAD52(209), their fraction was almost undetectable. Additionally, such a trimeric protein complex has never been observed or crystallized implying its high instability and the propensity of the protein to form ring structures. Secondly, according to the crystal structures of RAD52(209) with DNA bound to the outer side of the stem, the DNA retains a helical structure, suggesting that not all 12 base pairs will be in high proximity to the base pairs of another strand extended inside the inner groove. Therefore, it is not exactly clear how the base-pairing will occur and, most importantly, how the proofreading will be achieved. It is important to mention, that the latest published cryo-EM structure of the full-length protein bound to the DNA replication fork showed that the ssDNA part of the fork bound to the protein did not retain the helicity [54]. Taking together, these results indicate that the protein's ability to oligomerize, either to form rings or filaments, is essential, and the monomers or dimers are not sufficient for promoting the annealing.

The ability of RAD52(209) to accumulate simultaneously two ssDNA revealed a very interesting possibility that the annealing is promoted *via* single-ring structures (Figure 4.2, c) [60, 123]. In this case, the annealing mechanism could be similar to RecA [70]. The first ssDNA bound to the inner groove of the RAD52 ring is extended into B-form DNA exposing the base pairs. Upon binding to the protein, the second strand establishes contact *via* only a few base pairs on one monomer. Following this scenario, the helicity of the second DNA molecule bound to the outer stem of the protein is not an obstacle, but rather an advantage. Upon binding to the regions shorter than 7-nt, the protein can transiently probe for micro-homology and release easily when there is none [42]. This transient binding *via* short regions is essential to overcome the potential incapacity of the full-length protein to diffuse on the ssDNA [14, 31]. However, the exact quantification of the number of DNA molecules bound to the same ring in the case of RAD52 was challenging and it requires further investigation. According to the mechanism, one ring would promote the annealing of only one pair of complementary strands. Potentially both strands will stay on the same ring until the homologous region is established *via* constant

binding and unbinding of the ssDNA bound to the outer stem of the ring as it appeared in the case of RAD52(209) when the annealing process was investigated into details (Figure S31). Moreover, this model can be modified assuming that the crystal structure of the mutant RAD52(209) bound to two DNA molecules is not fully correct [123]. Once the second DNA binding site of RAD52(209) was identified, it was considered primarily with respect to the protein's ability to bind dsDNA upon the D-loop formation [60]. There were two sets of residues identified, the first one was located on the outer stem of the protein, while the second set of residues was running outside the inner groove. Assuming that the second complementary strand is able to bind outside of the groove, the structural positioning of the strands on the protein will resemble the $Red\beta$ and RecT annealing models [97, 20]. One strand is extended and partially protected within the groove, while the second strand is only transiently bound from the outside probing for the homology region. This model has been already proposed for the yeast Rad52 [63].

The third possible annealing mechanism could be related to the formation of higher-order oligomers or condensates, where multiple ring structures are bound to a few (two or three) single-stranded DNA molecules and by bringing them in close proximity, the hydrogen bonds between the complementary bases are formed (Figure 4.2, d). This mechanism can be similar to Rad51 and RecA which establish multiple transient contacts on the DNA in order to probe for homologous regions [42]. It has been shown that RecA bundles DNA breaks similarly to RAD52 [73]. The annealing inhibition in the presence of HD supports the idea that the high-order oligomers are essential for the annealing. However, these oligomers might only serve the function of a local condensate to bundle the ssDNA molecules and protect them from degradation. Therefore, this hypothesis requires further investigation.

The diversity in the oligomeric states of RAD52 and its truncated variant, as well as the variety of ring structures formed by RAD52, resembles the human recombinase Dmc1. It has been shown that Dmc1 forms ring structures with different numbers of subunits, as well as filaments [107, 66, 67]. It has been also proposed that Dmc1 exhibits different association modes with DNA *via* short oligomers and rings. Therefore, there is also a possibility that RAD52 and RAD52(209) utilize different annealing models based on different interaction modes with DNA.

Chapter 5

Outlook

The combination of biochemical assays and the mass photometry approach revealed new information about the oligomeric states of both RAD52 and RAD52(209) alone and upon binding to the DNA. The lower oligomers of both RAD52(209) and RAD52 have never been identified before, as well as the diversity of RAD52 ring structures was never thoroughly investigated. This new information has to be taken into account to create a new model of the single-strand annealing by RAD52. Moreover, it is important to remember that single- or double-stranded DNA binding alters significantly the oligomeric state of the protein. The results of this project brought a new insight into the protein oligomerization and its interaction modes with DNA, however, it also raised additional questions. Therefore, further experiments are required to answer these questions and mimic as closely as possible the *in vivo* conditions in order to design a new model explaining the mechanism of DNA annealing by RAD52.

First of all, it is necessary to utilize the proteins without the His-tag. All the above-described experiments were carried out with the His-tag at the N-terminal domain on both proteins. Although the majority of the annealing models published up-to-date also utilized His-tagged proteins, since its presence did not inhibit the ability of the protein to bind or anneal the DNA. Nevertheless, the His-tag might facilitate or alter the oligomerization of the protein. As a side project, I investigated the oligomeric state of the Sak protein, which is proposed to be a RAD52 ancestor, with and without the His-tag. Sak without the tag did not form any ring structures. Therefore, it is important to examine the oligomeric state of untagged RAD52. The utilized constructs carry a Thrombin cleavage site, however, several attempts to cleave it were unsuccessful. I proposed to replace the Thrombin cleavage site with TEV, and Dr. Zhihao Jiang (RG Prof. Lozano-Durán) created RAD52 and RAD52(209) modified constructs that could be utilized for further research.

Additionally, it would be very important to study the RAD52 oligomeric state and its

DNA binding mode in the presence of RPA. The preincubation of the ssDNA with RPA will resemble more closely the *in vivo* situation, and ensure the visibility of the DNA in mass photometry. Although the short ssDNA has MW below the lower detection limit of mass photometry (30 kDa), its binding to the trimeric RPA would be detectable since the total MW of the DNA-RPA complex will be above 116 kDa. Moreover, the exact mechanism of RPA displacement by RAD52 is not known, as well as the structure of the RPA-DNA-RAD52 complex. Mass photometry is a powerful tool to investigate these questions and provide the missing information. Furthermore, one could think of adding RAD51 to the system, however, the simultaneous combination of DNA and three proteins exhibiting different oligomeric states might be challenging.

In order to investigate the possibility of RAD52 undergoing liquid-liquid phase separation, the mass photometry experiments should be further conducted on both RAD52 and RAD52(209) in the presence of HD and upon binding to positively charged APTES slides. The preliminary experiments performed on the standard easy-cleaned slides only seldomly revealed the formation of RAD52 monomers upon increasing HD concentration, as the total number of counts was very low. Additionally, the fusion of protein droplets upon increasing salt concentrations should be investigated in order to confirm the liquid-like behavior.

Another important piece of information is the ability of RAD52 to perform the proofreading while aligning the homology regions. Is the protein capable of discriminating the mismatching base pairs? Or, perhaps, the quick annealing of broken DNA ends is more essential than the ideally matching strands for cellular viability during the DNA DSB. Point mistakes in the genome are less cytotoxic and could be resolved via the single-strand break repair (Figure 1.1). The ability of RAD52 to perform proofreading has never been challenged. My preliminary experiments with two and four mismatches at the 3' and 5' ends revealed the incapacity of the protein to sense the errors (Figure S29). However, further experiments with different numbers and the locations of the mismatches should be conducted, in analogy to DdrB [141]. DdrB, similarly to RAD52, hides a certain number of base pairs, releasing them only once the first alignment of the homologous regions occurs [141]. For RAD52, Arg55 and Tyr65 would be the best candidates to challenge protein proofreading. These residues secure ssDNA inside the groove and they were suggested to undergo conformational changes which release the strand and enable the base-pairing [60, 123]. A modification of these residues might impact the protein proofreading mechanism in case it is confirmed for the wild-type RAD52. The results of the above-proposed experiments will contribute drastically to our understanding of RAD52 interaction modes with DNA and its annealing mechanism. Furthermore, this knowledge could be applied to design more efficient anti-cancer therapeutic pathways aiming to inhibit RAD52.

Supplementary information

Table S1: The list of single-stranded DNA molecules. All the oligonucleotides have a phosphate group at the 5'-end and some have various fluorescent modifications at the 3'-end. DNA lengths (nt) are mentioned in the oligonucleotide name, followed by its sequence, molecular weight (kDa), and presence of the secondary structure elements (SSE).

Oligo name	Sequence	MW (kDa)	SSE
12-Atto565	GGAAGCGATGTG	3.541	None
16-Atto565	CACATCGCTTCCTAGT	4.792	None
26	GCGGAAGCGATGTGAATTTTTTTA	7.437	None
32-Atto565/ Atto680	GCTCTAAGCCATCCGCAAAAA TGACCTCTTAT	9.712	Weak
48-Alexa488	GCAATTAAGCTCTAAGCCATC CGCAAAAATGACCTCTTATCAAAAGGA	14.705	Weak
60-Atto680	TCCTTTTGATAAGAGGTCATT TTTGCGGATGGCTTAGAGCTT AATTGCTTTTTTTTTTTT	18.475	Weak
60-2mismatch	TCCTTTTGCTAAGAGGTCATT TTTGCGGATGGCTTAGAGATT AATTGCTTTTTTTTTTTT	18.560	Weak
60-4mismatch	TCCTTTTGCTAAGGGGTCATT TTTGCGAATGGCTTAGAGATT AATTGCTTTTTTTTTTTT	18.560	Moderate
70-Atto532	TTTAATAATTTACTTTATTTT CTATGTCTATTCATTTACTTAC TTGTGTATTATCCTTATACTTACTTAC		
83	AAATAAACATAAAGTAAGTAA GTATAAGGATAATACACAATA AGTAAATGAATAGACATAGAA AATAAAGTAAATTATATAAA	25.827	Weak

Table S2: Protein variants. The table contains information about all the protein variants utilized in this work. The protein lengths (number of amino acids) as followed by the sequence-predicted molecular weight of a monomer (kDa), the isoelectric point pI value, and the extinction coefficient.

Protein	Length (a.a.)	MW (kDa)	pI	Exc. Coefficient
RAD52(209)	230	25.352	8.48	20587.5
RAD52(209)-DyLight488	230	26.174	-	20587.5
RAD52	439	48.463	8.64	40130
GFP-RAD52	678	75.513	6.72	63572.5

Table S3: Protein binding affinity. The table contains fitting constants of RAD52(209) and RAD52 binding to single-stranded DNA. Binding affinities were examined with different lengths of single-stranded DNA (32-, 60-, and 70-nt) at 10 nM. The protein concentration was 100 nM. The DNA binding was examined at different temperatures (25°C and 37°C). The number of repeats is N. Mean values weighted with SEM were fitted to Hill and Michaelis-Menten equations, V_{max} was fixed to 100 %.

	RAD52(209)			RAD52		
Number of nt (T in °C)	K_d (M)	H	N	K_d (M)	H	N
32-nt (25)	169 ± 24	1.5 ± 0.2	5	115 ± 22	1.4 ± 0.2	6
70-nt (25)	109 ± 23	1.6 ± 0.3	4	53 ± 8	-	5
32-nt (37)	246 ± 33	3.1 ± 0.5	3	135 ± 4	2.7 ± 0.2	3
60-nt (37)	127 ± 10	4.1 ± 0.4	3	112.4 ± 0.1	3.3 ± 0.2	3

Table S4: Rates of the single-strand DNA annealing. The table contains the rate parameters of the annealing reactions performed with and without the protein. Annealing rates were tested with 32- and 60-nt DNA at 10 nM in the annealing buffer and the mass photometry buffer. The protein concentrations were 100 nM. The number of repeats is N. Mean values weighted with SEM were fitted to the second-order kinetic reaction $r = k[A][B]$.

	Annealing Buffer		Mass Photometry Buffer	
Protein	r (nM/min)	N	r (nM/min)	N
No protein	0.722	5	0.976	3
RAD52(209)	5.093	5	3.997	3
RAD52	52.272	5	77.912	3

Table S5: RAD52(209) binding to the single-stranded DNA. The table contains averaged shifts in molecular weights (mean with SEM) of the single-ring peak of RAD52(209). RAD52(209) concentrations were 100 and 500 nM. DNA concentrations were 10, 100, and 200 nM (first column). Tested DNA lengths were 32-, 48-, 60-, and 83-nt, their molecular weights are listed in the second column. The number of repeats is N.

RAD52(209) (nM)			100	N	500	N
DNA length (nt)	Conc.(nM)	MW (kDa)				
32	10	9.8	19.592 ± 1.39	3	5.982 ± 2.976	3
	100	9.8	20.634 ± 0.549	3		
48	10	14.8	31.48 ± 1.732	3		
	100	14.8	34.534 ± 0.514	3		
	200	14.8	33.421 ± 9.601	3		
60	10	18.6	42.184 ± 1.435	3	8.884 ± 3.866	3
	100	18.6	42.448 ± 0.315	3		
83	10	25.3	35.583 ± 1.709	3		
	100	25.3	53.534 ± 5.374	3	44.466 ± 10.368	3
	200	25.3	41.039 ± 7.088	3		

Table S6: RAD52(209) binding single-stranded DNA in the annealing buffer. The table contains averaged molecular weight shifts (mean with SEM) of the RAD52(209) single-ring peak. RAD52(209) concentration was 100 nM. DNA concentration was 10 nM. DNA lengths were 32- and 60-nt, their molecular weights are in the second column. The experiments were performed at 25°C and 37 °C with several repeats (N).

Temperature °C		25		37	
RAD52(209) (nM)		100 (N)	10 (N)	100 (N)	10 (N)
ssDNA (nt)	MW (kDa)				
32	9.8	37.0 ± 9.8 (4)	21.0 ± 5.6 (3)	27.5 ± 5.7 (4)	21.2 ± 9.5 (3)
60	18.6	44.4 ± 8.2 (4)	31.1 ± 9.3 (3)	36.0 ± 3.4 (4)	34.0 ± 2.8 (2)

Table S7: Single-stranded DNA annealing in the presence of RAD52(209). The table contains molecular weight shifts (mean with SEM) of the RAD52(209) single-ring peak during the annealing reactions. The reactions were examined in the presence of RAD52(209) at 100 and 500 nM. The annealing pairs were: (32+60)-nt and (60+32)-nt - complementary strands, and (32+83)-nt and (60+83)-nt - non-complementary strands. For reference, the sum of MW of utilized oligonucleotides is listed in the second column. The DNA concentrations were 10 nM. The number of repeats is N.

RAD52(209) (nM)		100	N	500	N
Ann.pair (nt)	MW (kDa)				
(32+60)	28.4	34.217 ± 5.329	8	35.7 ± 7.2	3
(60+32)	28.4	39.306 ± 6.93	6	62.7 ± 17.6	3
(32+83)	35.1	26.283 ± 7.71	6		
(60+83)	35.1	36.027 ± 7.105	6		

Table S8: RAD52(209) binding to the double-stranded DNA. The table contains molecular weight shifts (mean with SEM) of the RAD52(209) single-ring peaks upon binding to the double-stranded DNA. RAD52(209) concentration was 100 nM, and DNA concentrations were 10 and 100 nM. The double-stranded DNA without overhangs was (48+48)-nt, and the DNA with overhangs was (48+60)-nt. Corresponding MW (kDa) are listed in the third column. The number of repeats is N.

dsDNA (nt)	conc.(nM)	MW (kDa)	MW shift (kDa)	N
(48+48)	10	29.6	71.102 ± 23.886	4
	100	29.6	86.06 ± 16.761	4
(48+60)	10	33.4	83.096 ± 18.03	4
	100	33.4	83.228 ± 18.995	4
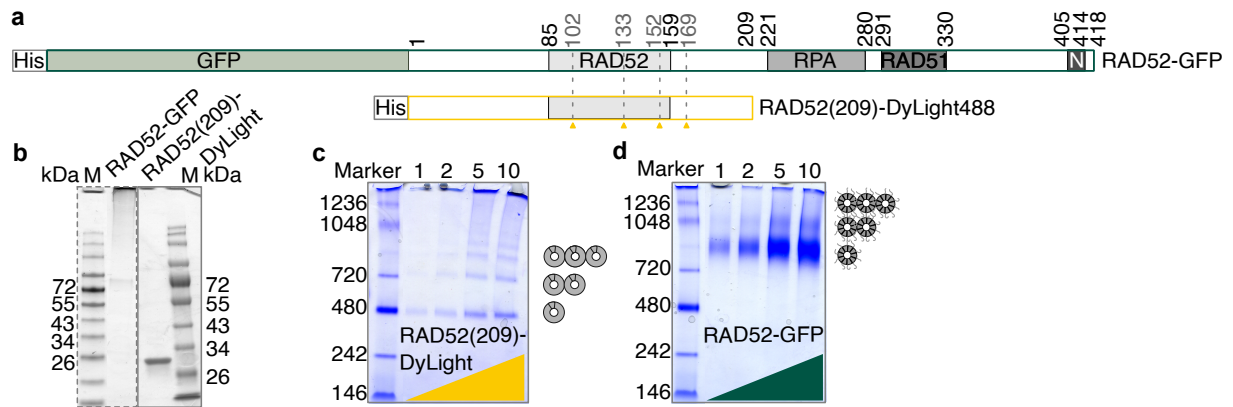
Figure S1: Overview of GFP-RAD52 and RAD52(209)-DyLight488. **(a)** Schematics of the protein sequences with the identified domains, tags, and chemical labeling. Lysine residues in the RAD52(209)-DyLight488 are potential docking sites for DyLight488 (yellow arrows). **(b)** SDS-PAGE of the proteins with the marker (M). **(c)** Blue Native PAGE of RAD52(209)-DyLight488 and **(d)** GFP-RAD52 and GFP-RAD52(209) at 1, 2, 5, and 10 μM.
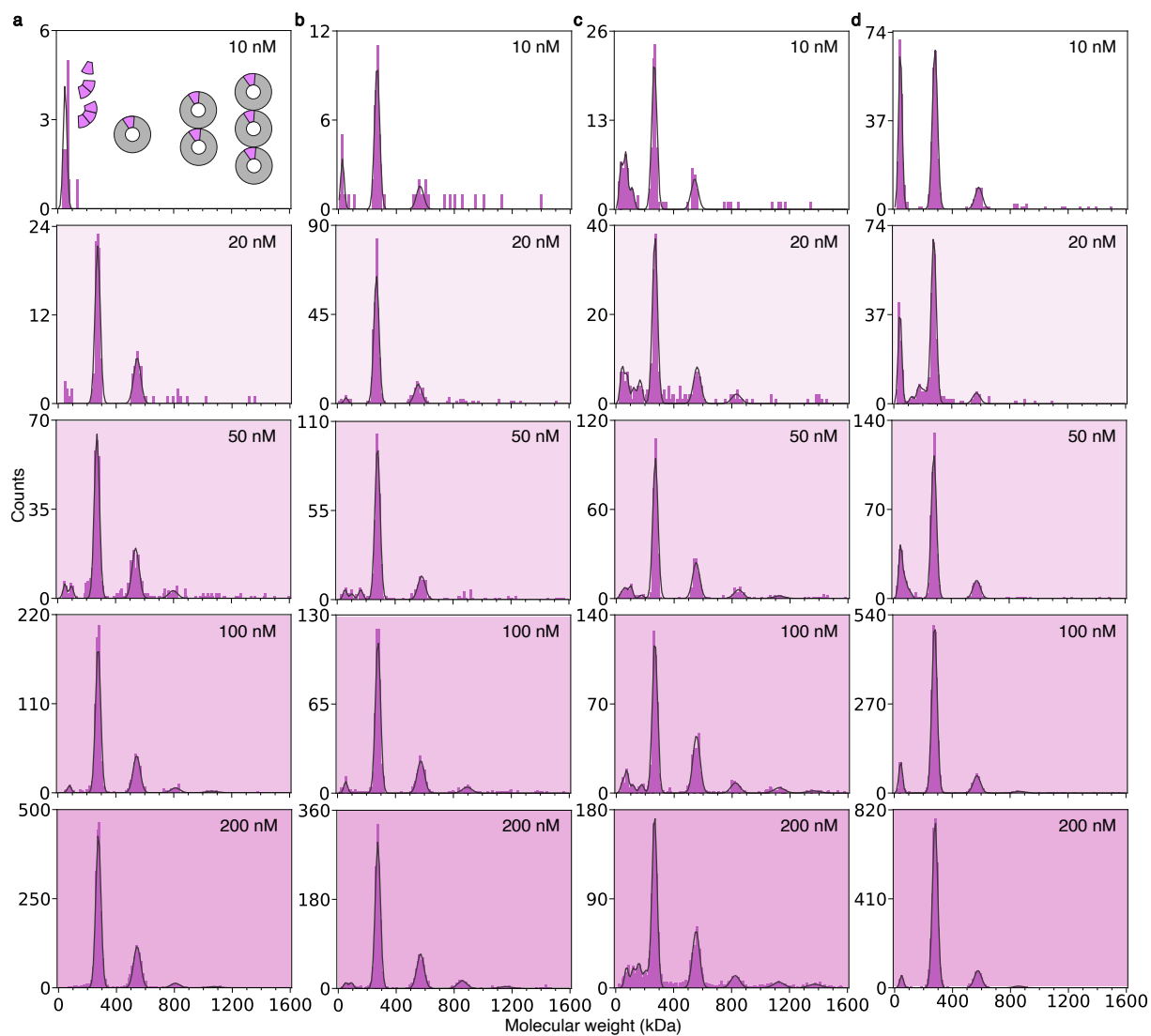
Figure S2: Mass histograms of RAD52(209) at different days. The protein was measured at 10, 20, 50, 100, and 200 nM. Each column presents one day of experiments **(a-d)**. Each mass histogram contains several readings collected on the same day. The data was acquired in the RFV.
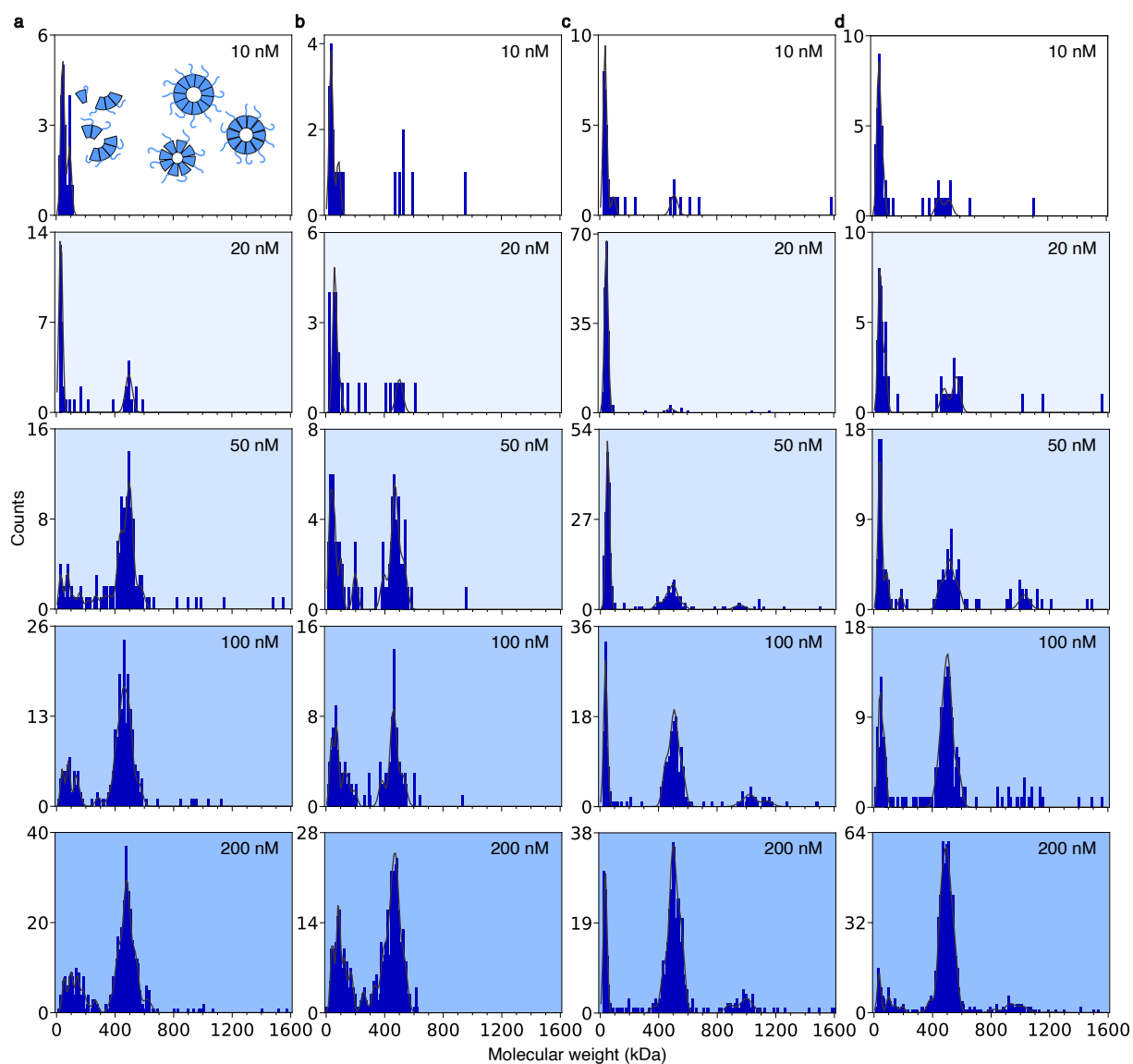
Figure S3: Mass histograms of RAD52 at different days. The protein was measured at 10, 20, 50, 100, and 200 nM. Each column presents one day of experiments **(a-d)**. Each mass histogram contains several readings collected on the same day. The data was acquired in the RFV.
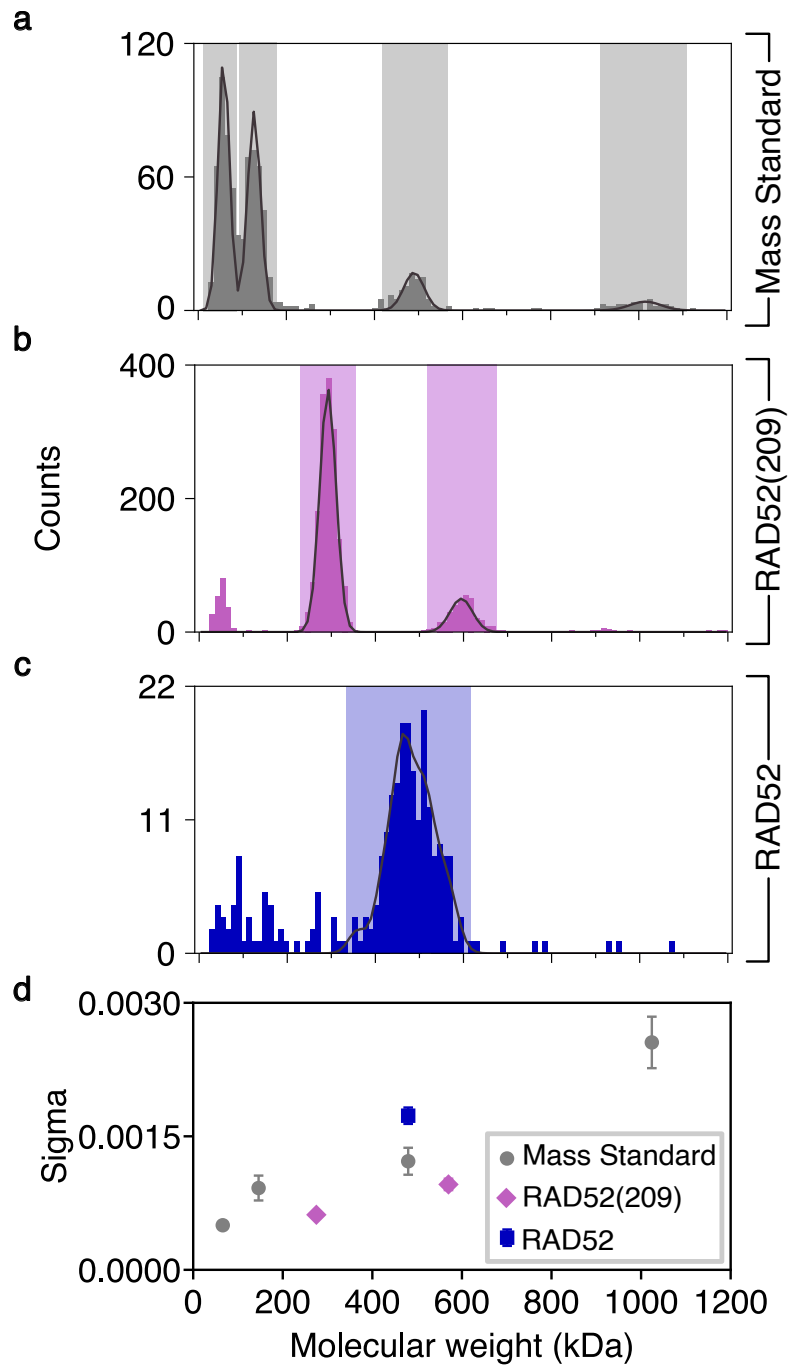
Figure S4: Constriction of the sigma values. The mass histograms of the protein standard **(a)**, RAD52(209) **(b)**, and RAD52 **(c)** with Gaussian envelopes (solid black lines). Distributions of the major peaks are color-coded - grey, magenta, and blue, respectively. **(d)** Mean sigma values with SEM were averaged from several readings and plotted for each protein as the function of its molecular weight.
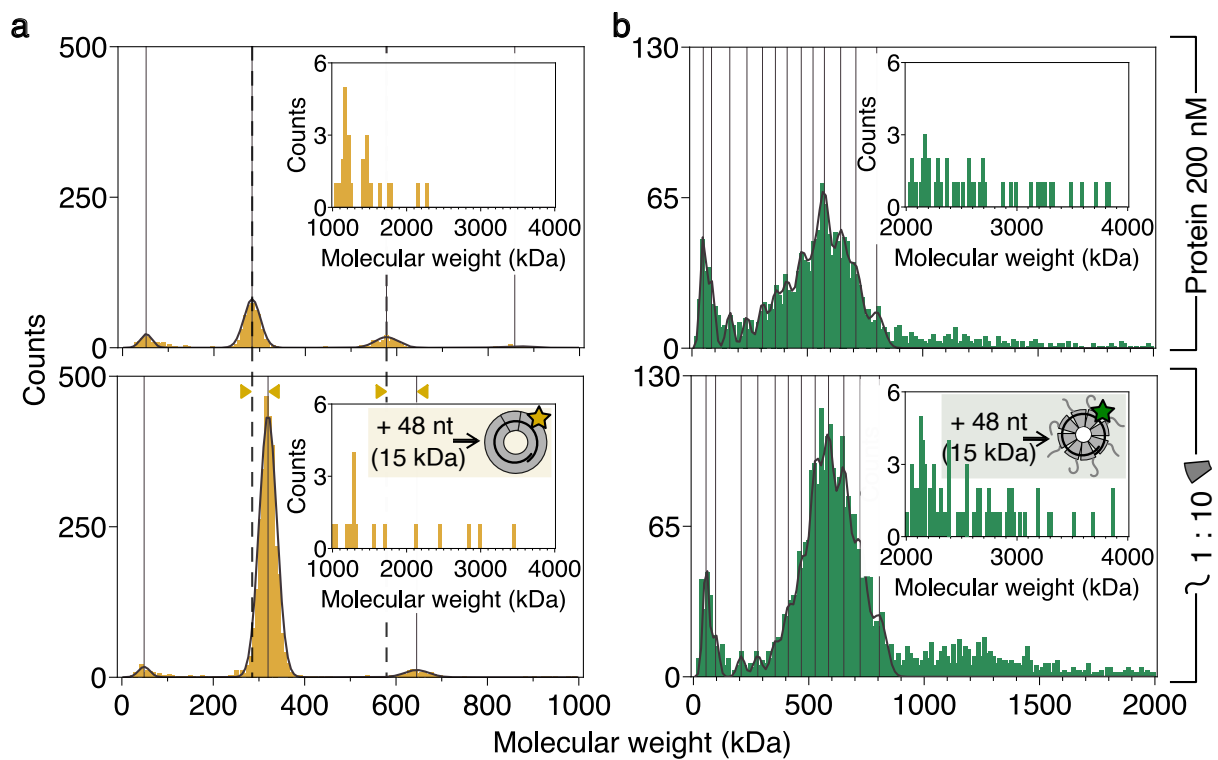
Figure S5: Mass histograms of RAD52(209)-DyLight and GFP-RAD52. The oligomeric state of RAD52(209)-DyLight **(a)** and GFP-RAD52 **(b)** at 200 nM (the upper panel) and after the incubation with 10 nM of the 48-nt DNA (lower panel). The mass histograms contain the cumulative distributions of several readings (N=3) measured on the same day. The data was acquired in the RFV.
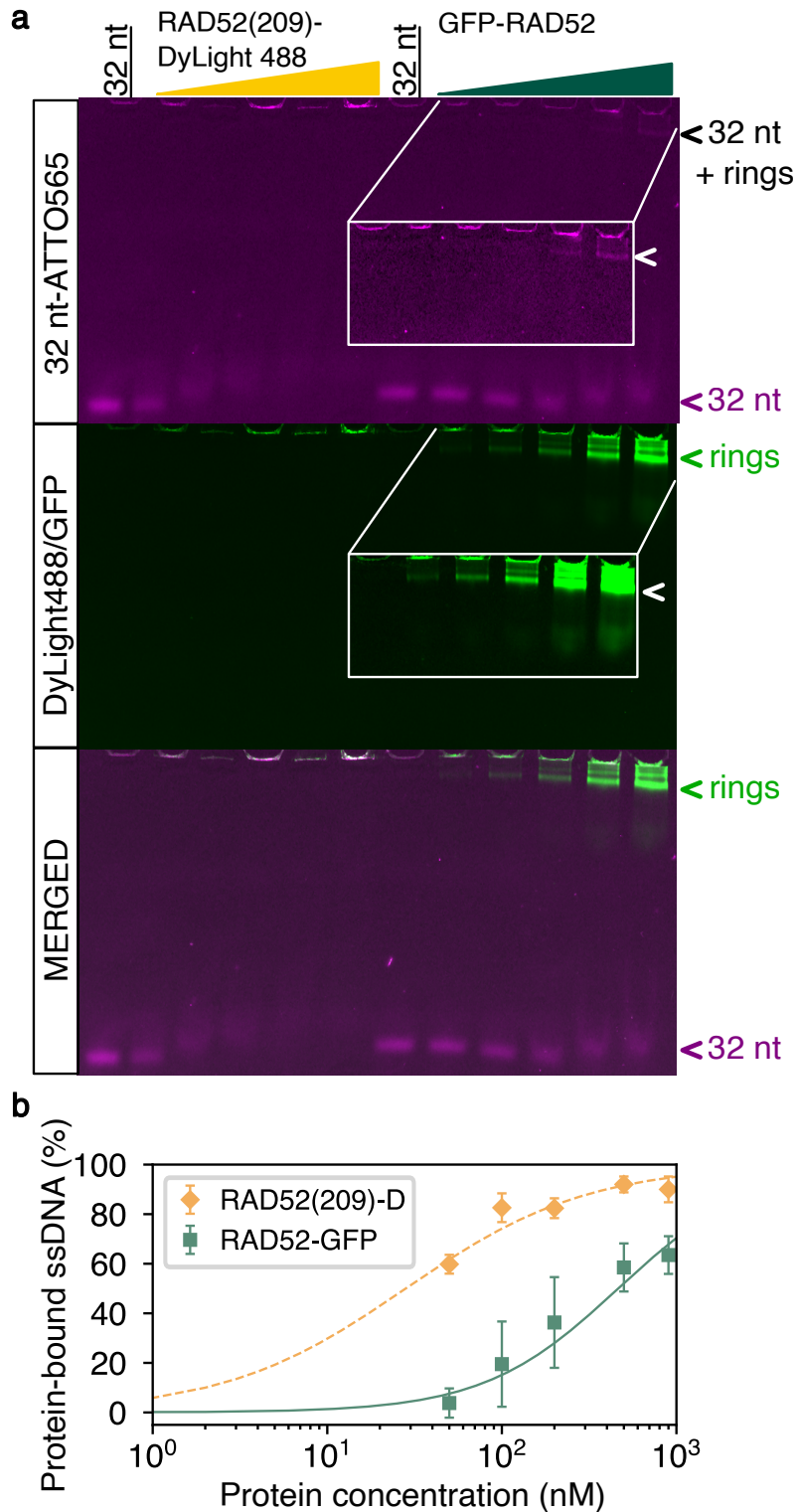
Figure S6: Single-stranded DNA binding of RAD52(209)-DyLight488 and GFP-RAD52. The affinities were tested at 50, 100, 200, 500, and 1000 nM protein. The DNA was 32-nt DNA long at 6.25 nM. **(a)** EMSA shows 32-ATTO565 (red channel), proteins (green channel), and the overlap (MERGE). The inserts contain the saturated images of the wells. Color-coded arrows indicate individual DNA or DNA-protein complexes. **(b)** Normalized values of the 32-nt band intensities (mean with SEM) were averaged from 3 repeats and plotted as the function of the protein concentration. The data was fitted to the Hill and Michaelis-Menten equations.
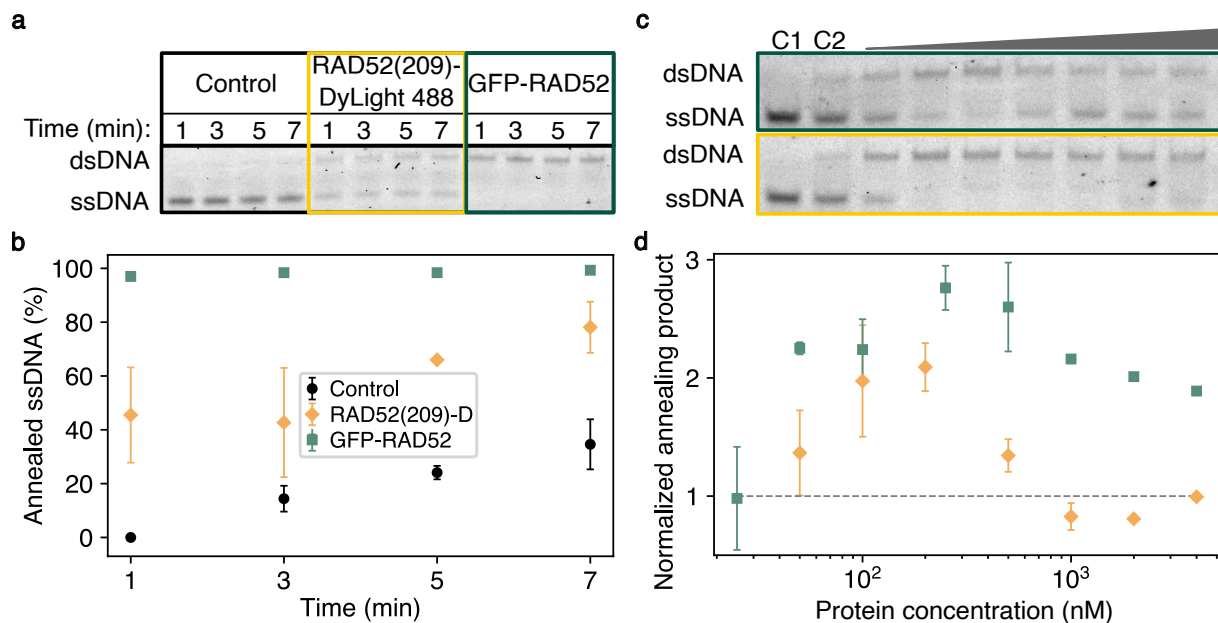
Figure S7: Single-stranded DNA annealing promoted by RAD52(209)-DyLight488 and GFP-RAD52. The protein concentration was 100 nM. The annealing pair was 32- and 60-nt at 6.25 nM. **(a)** An exemplary annealing gel. The reaction was terminated after 1, 3, 5, and 7 min. The intensity of the dsDNA product formed in the presence of GFP-RAD52 after 7 minutes was used as 100%, and the rest of the dsDNA bands were normalized to it. **(b)** Normalized mean values with SEM of the dsDNA band intensities averaged from 3 repeats are plotted as the function of the protein concentration. **(c)** An exemplary gel demonstrating the annealing activity of both proteins at 50, 100, 200, 500, 1000, 2000, and 4000 nM. C1 is the control reaction with only single-stranded DNA present, and C2 is the intrinsic annealing reaction (no protein). **(d)** Normalized to C2 mean values with SEM of the dsDNA band intensities averaged from 3 repeats are plotted as the function of the protein concentration.
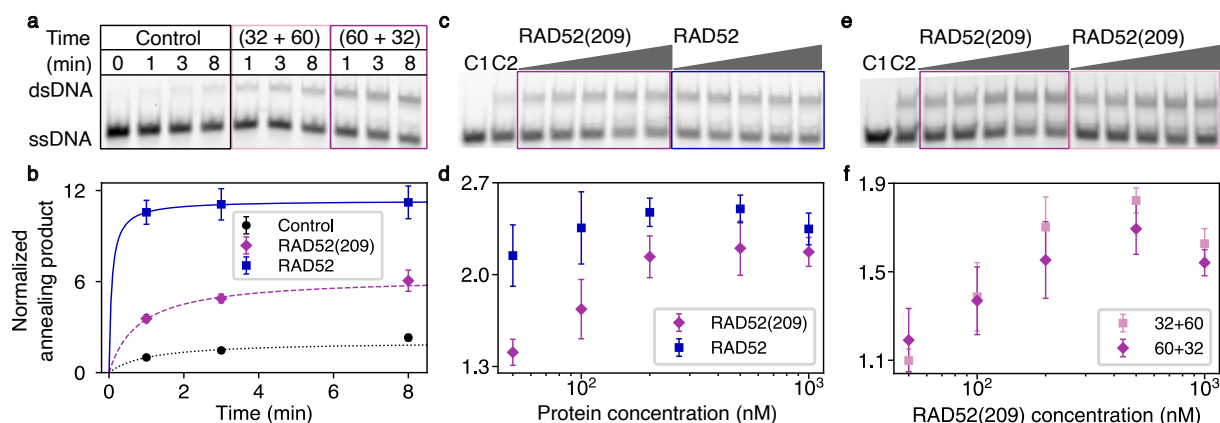
Figure S8: Single-stranded DNA annealing in the mass photometry buffer. The annealing pair was 32- and 60-nt at 10 nM concentration. The annealing rates were tested at 100 nM of the protein **(a-b)**. The reaction fractions were terminated after 1, 3, and 8 minutes. The reaction rates were fitted to the second-order reaction kinetics, obtained parameters are listed in Table S4. **(c-d)** The annealing activity was tested at 50, 100, 200, 500, and 1000 nM of the proteins. **(e-f)** RAD52(209) annealing activity was tested once the protein was initially preincubated either with the short (32-nt, pink) or the long DNA (60-nt, magenta).
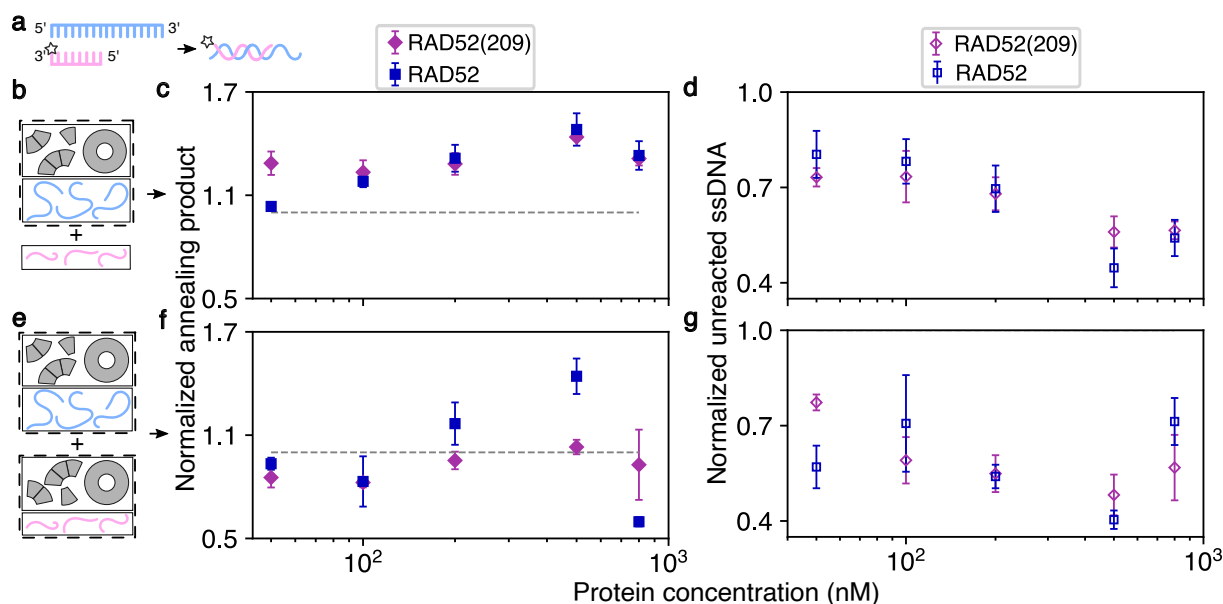
Figure S9: Single-stranded DNA annealing promoted by RAD52 and RAD52(209) at 37°C for 1 hour. **(a)** The annealing pair was 32- and 60-nt at 10 nM concentration. The two incubation strategies were examined. In the first strategy, the protein was preincubated with one single-stranded DNA, and the second strand was consequently added to the mixture **(b-d)**. According to the second strategy, two protein fractions were preincubated with one of the strands and mixed after that **(e-g)**. The intensity of the dsDNA bands was normalized to the control band (no protein, intrinsic annealing). The intensity of the single-stranded DNA bands was normalized to the intensity of the single-stranded DNA. Obtained mean values with SEM were calculated from 3 repeats and plotted as the function of the protein concentration (for each incubation strategy). The dashed grey lines indicate the level of the intrinsic annealing (control).
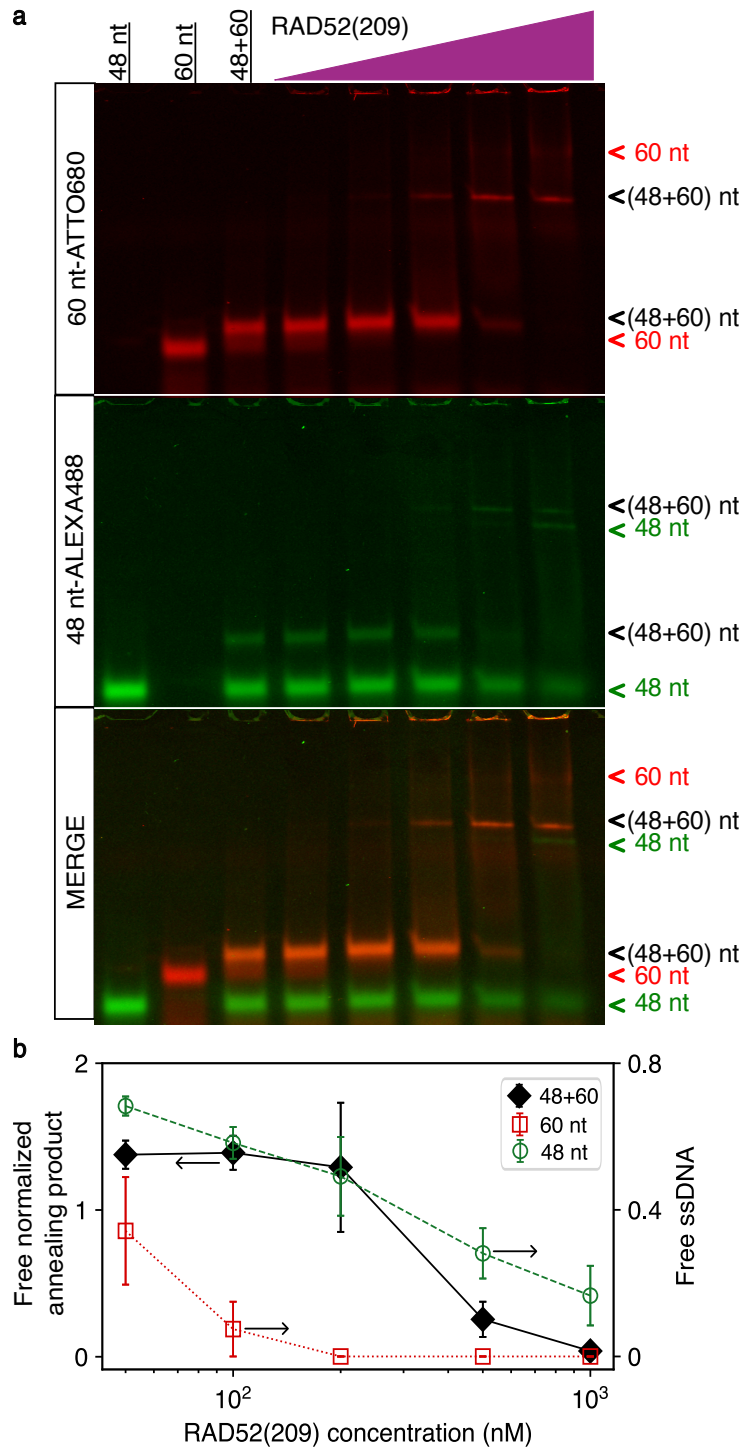
Figure S10: Native annealing promoted by RAD52(209). The annealing pair was 48-nt 3'-ALEXA488 (green open circles) and 60-nt 3'-ATTO680 (red open squares) **(b)**. The oligonucleotide concentration was 10 nM. Protein concentrations were 50, 100, 200, 500, and 1000 nM. **(a)** Gels demonstrate the 60-nt (red channel), 32-nt (green channel), and a combined image (MERGE). Color-coded arrows point to the intermediate species. The intensities of the dsDNA were normalized to the control reaction (48+60), and the intensities of the single-stranded DNA were normalized to the corresponding control - 48- or 60-nt. Data points were averaged from 4 repeats and plotted as mean values with SEM. Black arrows direct data points to the corresponding axis.
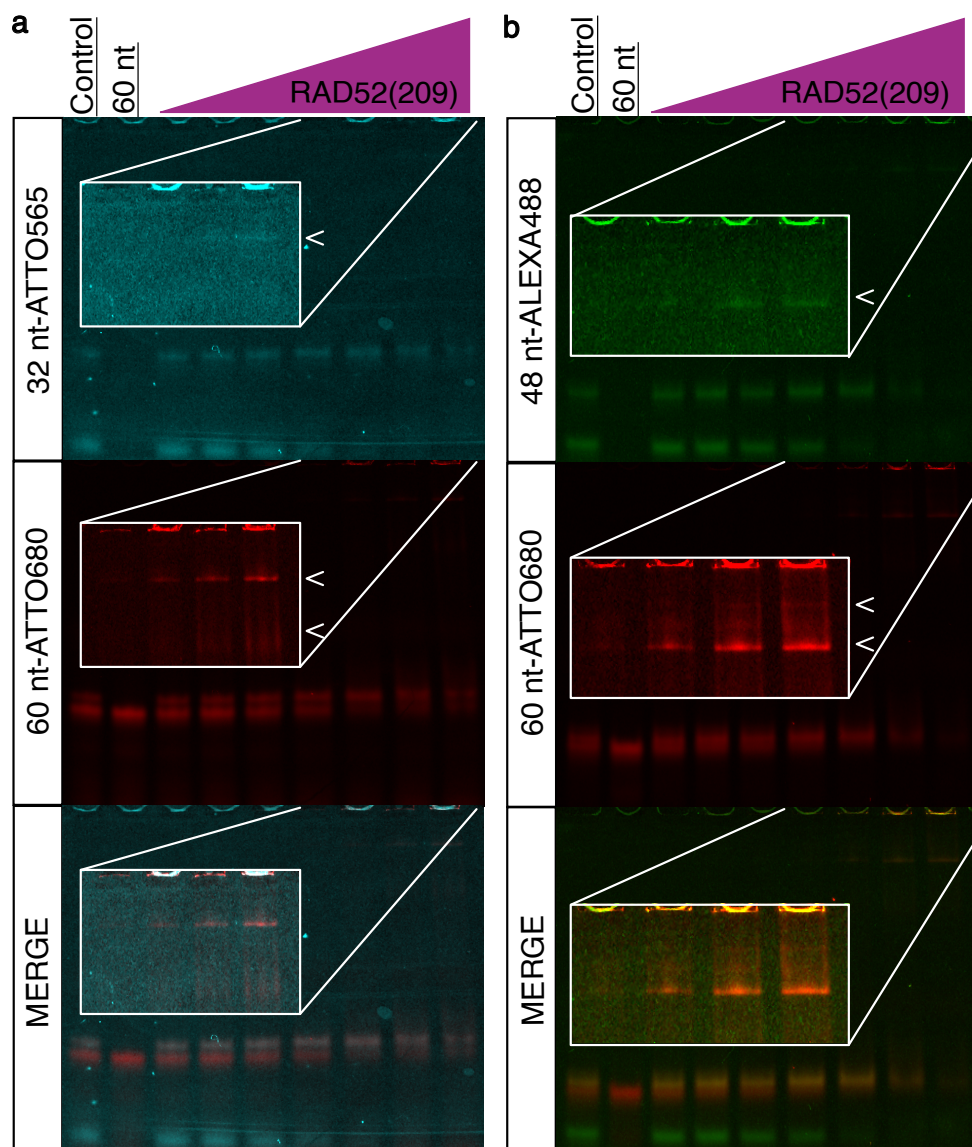
Figure S11: Native annealing promoted by RAD52(209) on different DNA pairs. The annealing pairs were 32-nt 3'-ATTO565 and 60-nt 3'-ATTO680 **(a)**, and 48-nt 3'-ALEXA488 with 60-nt 3'-ATTO680 **(b)**. The DNA concentration was 10 nM. The protein was initially incubated with a shorter single-stranded DNA (32- or 48-nt). The protein concentrations were 50, 100, 200, 500, 1000, and 2000 nM. Gels demonstrate 60-nt (red channel), 32- and 48-nt (cyan and green channels), and corresponding combined images (MERGE). Each experiment was performed only once. The inserts demonstrate saturated images of the wells. The white arrows indicate the annealing intermediate species.
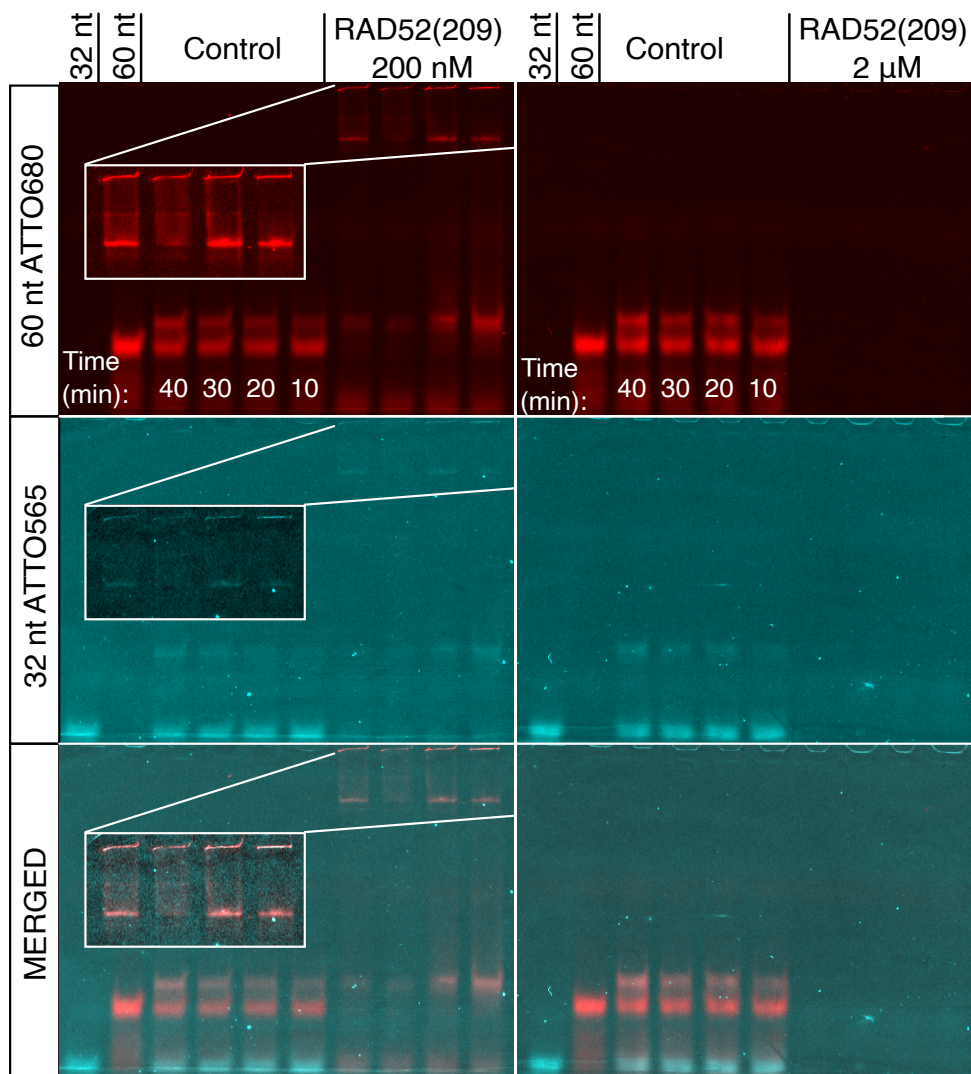
Figure S12: The formation of the dsDNA product during the native annealing reaction over 40 min. The annealing pair was 32-nt 3'-ATTO565 and 60-nt 3'-ATTO680 at 10 nM. RAD52(209) was tested at 200 and 2000 nM. The amount of dsDNA product was examined after 10, 20, 30, and 40 min at 25°C. The inserts demonstrate saturated images of the wells showing the annealing intermediates. The experiment was repeated 3 times, however, due to the absence of the bands at 2000 nM RAD52(209), no statistical analysis was performed.
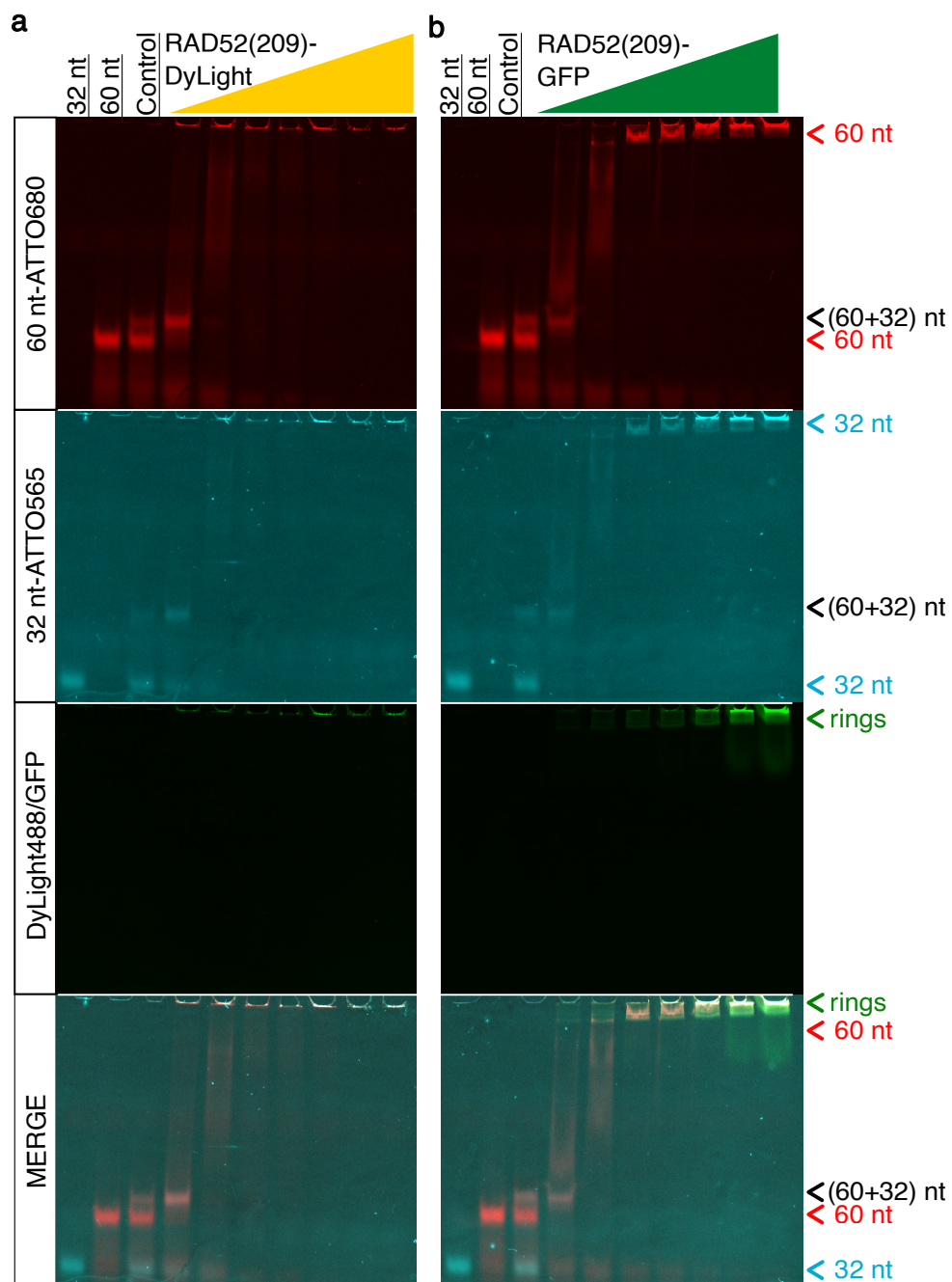
Figure S13: Native annealing promoted by RAD52(209)-DyLight488 **(a)** and GFP-RAD52 **(b)**. The annealing pair was 32-nt 3'-ATTO565 and 60-nt 3'-ATTO680 at 10 nM. Protein concentrations were 100, 200, 400, 500, 1000, 2000, and 4000 nM. Gels demonstrate the red channel (60-nt), the cyan channel (32-nt), and the green channel (DyLight488 and GFP), as well as the combined images (MERGE). Each experiment was performed only once.
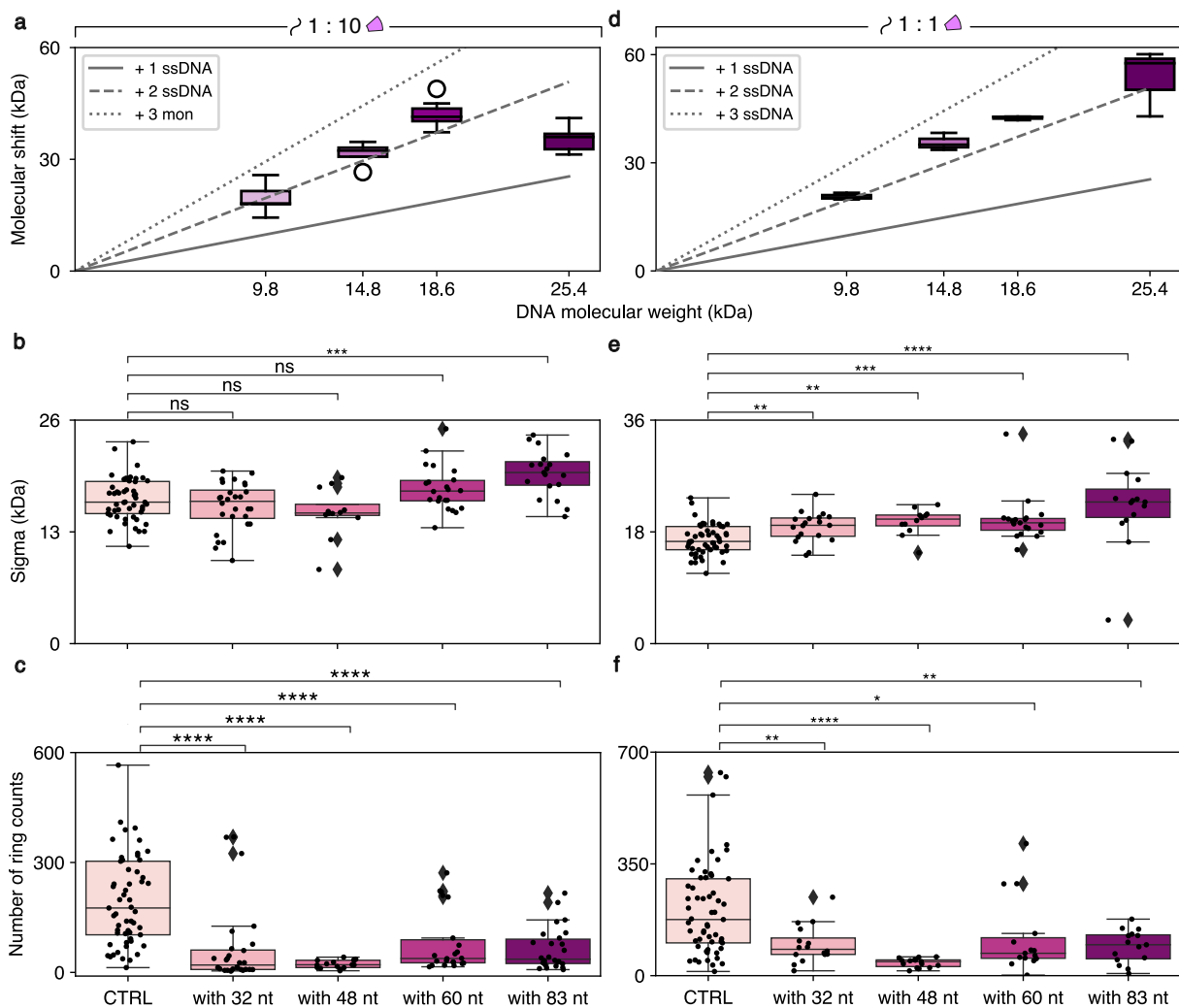
Figure S14: Statistical analysis of the RAD52(209) single ring upon single-stranded DNA binding. Corresponding data is shown in Figure 3.10. RAD52(209) concentration was 100 nM. The DNA concentrations were 10 (**a-c**) and 100 nM (**d-f**). The upper panel (**a**) and (**d**)) depicts the molecular weight shifts as the function of the DNA molecular weight. Lines indicate where the data point should be in case MW shifts are proportional to one (solid line), two (dashed line), or three (dotted line) single-stranded DNA bound to the same ring. The middle panel shows the evolution of sigma values at different DNA lengths and concentrations. The lower panel contains information about the total number of counts. A single black circle is one reading, black filled diamonds (**b**) and empty circles (**a**) are the outliers. Sigma values and the total number of counts were statistically compared to the control data (CTRL, no DNA). For the statistical analysis, the two-sided Mann-Whitney test was applied, corresponding p-values are labeled with asterisks or "ns" and listed in Table S5.
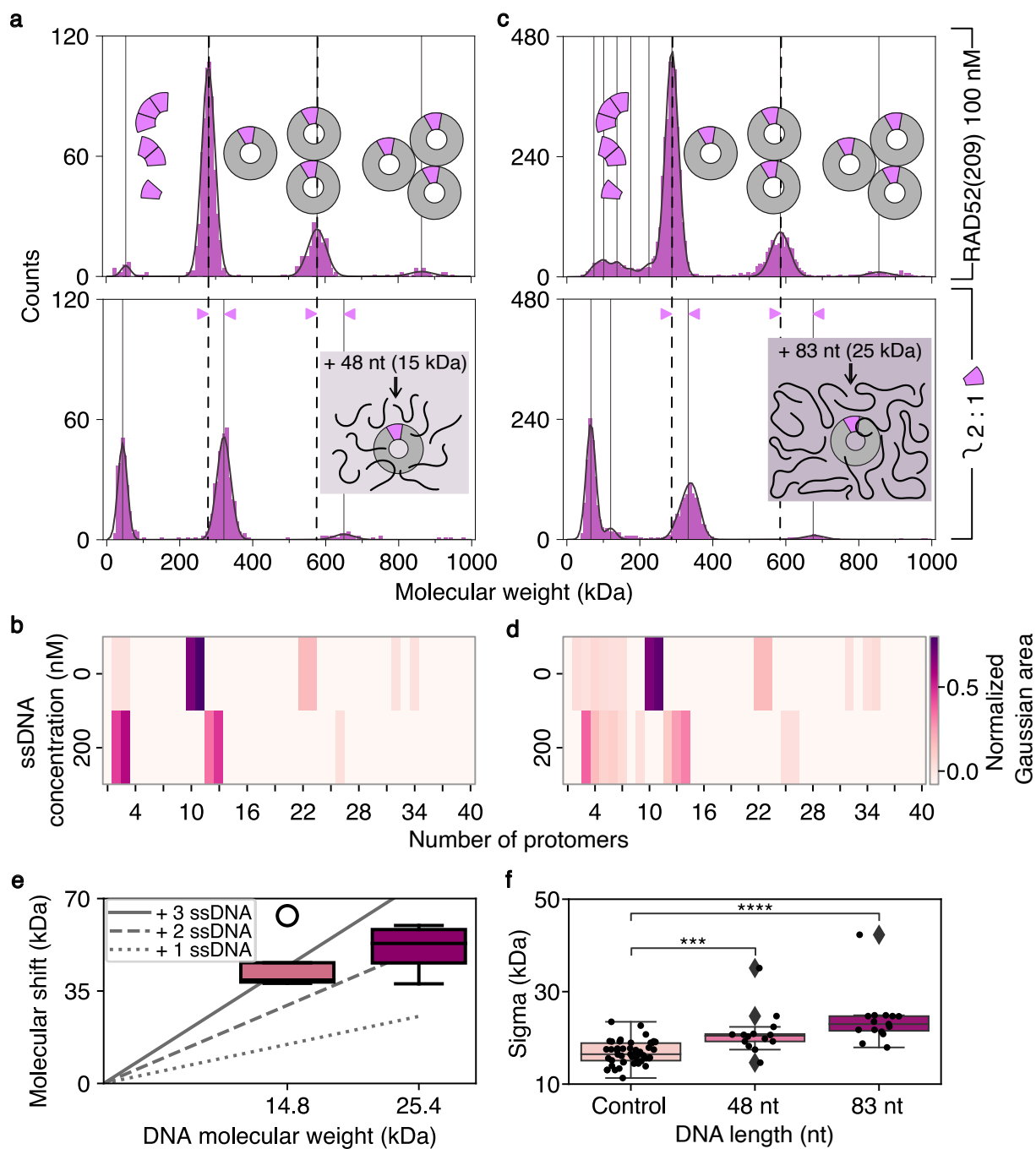
Figure S15: RAD52(209) binding to the excess of the single-stranded DNA. The DNA molecules were 48- **(a-b)** and 83-nt long **(c-d)**. RAD52(209) concentration was 100 nM, and the DNA concentration was 200 nM. Mass histograms display a cumulative distribution of several readings acquired on the same day. The heatmaps **(b)** and **(d)** display normalized areas of distinct protomers as a cumulative sum of all the repeats. The MW shifts of the single-ring peak **(e)** are plotted versus the DNA molecular weights. Sigma values of the single ring peak **(f)** were compared to the control reaction (no DNA) and the two-sided Mann-Whitney test was performed. Corresponding p-values depicted as asterisks. Each black circle represents one reading, the black filled diamonds **(f)** and empty circles **(e)** are the outliers.
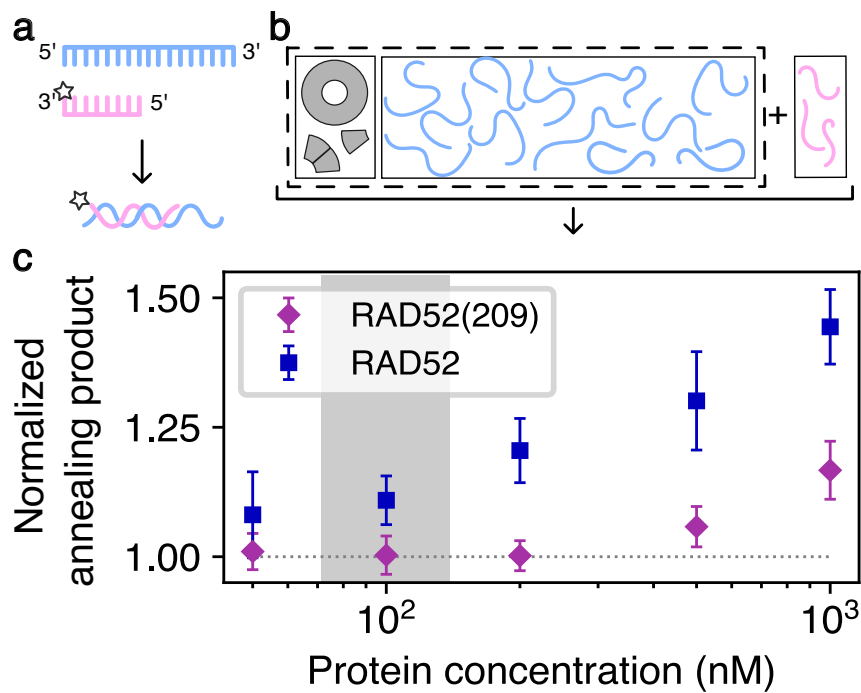
Figure S16: Single-stranded DNA annealing promoted by RAD52 and RAD52(209) with the excess of the DNA. The annealing pair was 32- and 60-nt at 10 nM and 200 nM concentration, respectively **(a-b)**. The protein was preincubated with 60-nt at 37°C for 5 min, subsequently the complementary 32-nt strand was added. The reaction was stopped after 15 min at 25°C. The intensities of the dsDNA bands were normalized to the control band (no protein, intrinsic annealing). The reaction was repeated 7 times. Mean values and SEM are plotted as the function of the protein concentration. The grey area indicates the DNA-to-protein ratio (2:1) that was utilized for the mass photometry measurements (Figure S15).
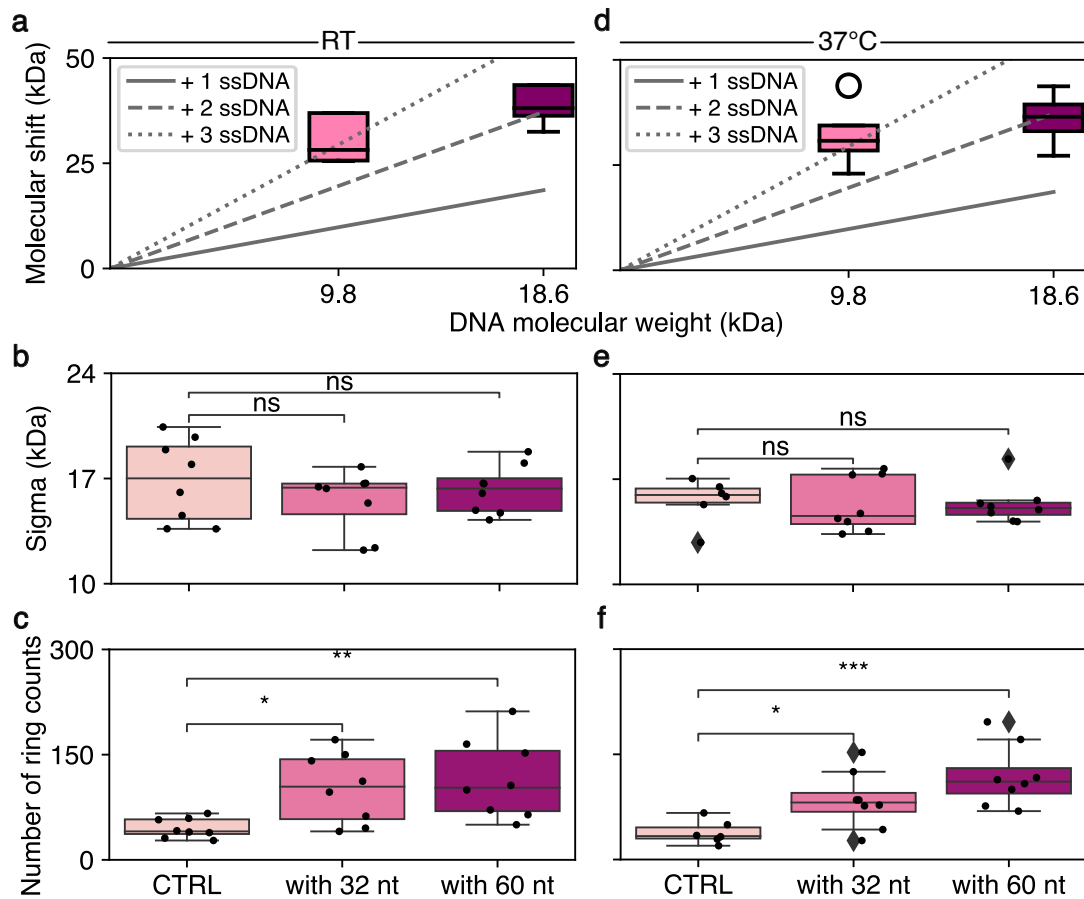
Figure S17: RAD52(209) at 100 nM binding to the single-stranded DNA in the annealing buffer. DNA molecules were 32- and 60-nt long, and DNA concentration was 10 nM. Protein-to-DNA binding was tested at 25°C **(a-c)** and 37°C **(d-f)**. The upper panel demonstrates the molecular weight shifts of the single ring peak as the function of the DNA MW. Lines indicate the predicted positioning of the data in case the shifts would be strictly proportional to one (solid line), two (dashed line), or three (dotted line) single-stranded DNA bounded. The middle panel depicts sigma values of the single ring peak at different DNA lengths. The lower panel presents the total number of ring counts with and without DNA (CTRL). Each black circle is one reading, black filled diamonds **(e)** and empty circles **(d)** are the outliers. Sigmas and the total number of counts are plotted as mean values with SEM. The two-sided Mann-Whitney test was performed, and corresponding p-values are depicted as asterisks and "ns" and listed in Table S6.
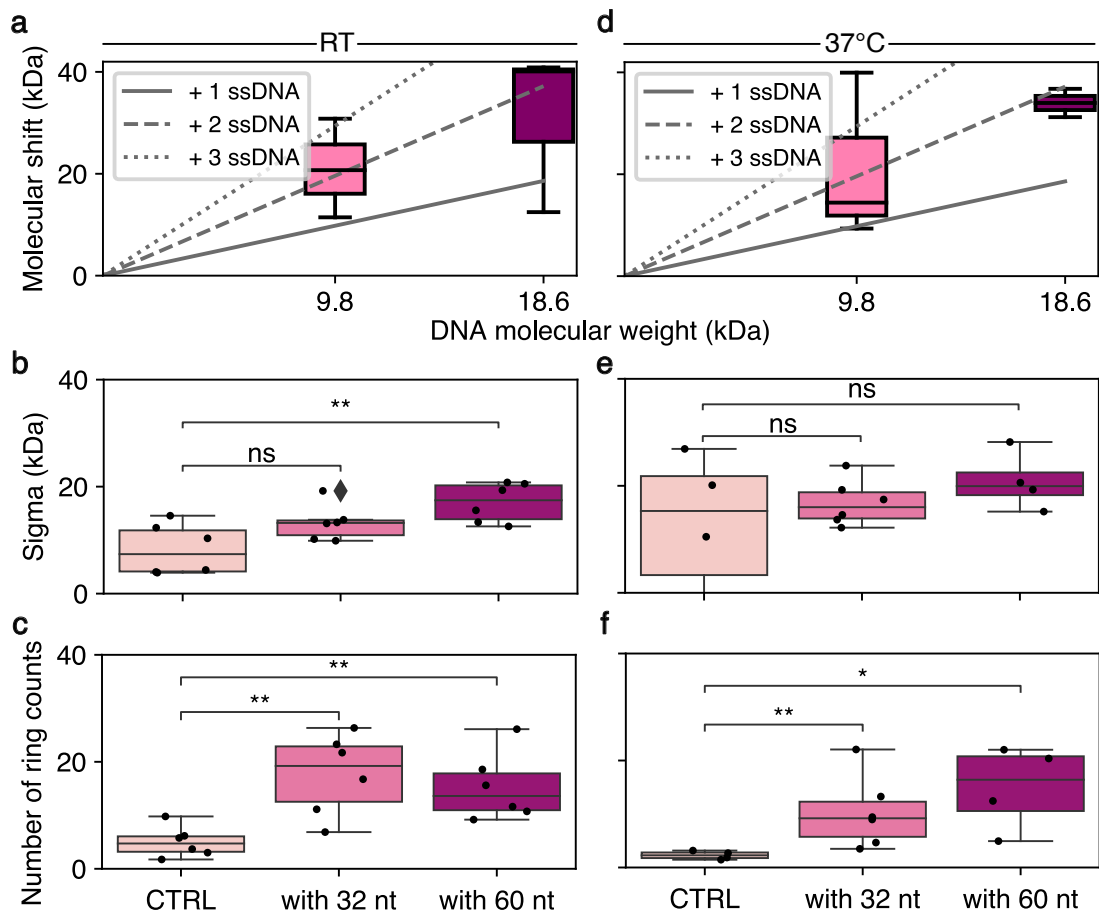
Figure S18: RAD52(209) at 10 nM binding to the single-stranded DNA in the annealing buffer. DNA molecules were 32- and 60-nt long, and DNA concentration was 10 nM. Protein-to-DNA binding was tested at 25°C **(a-c)** and 37°C **(d-f)**. All further details and data analysis are exactly as described in Figure S17.
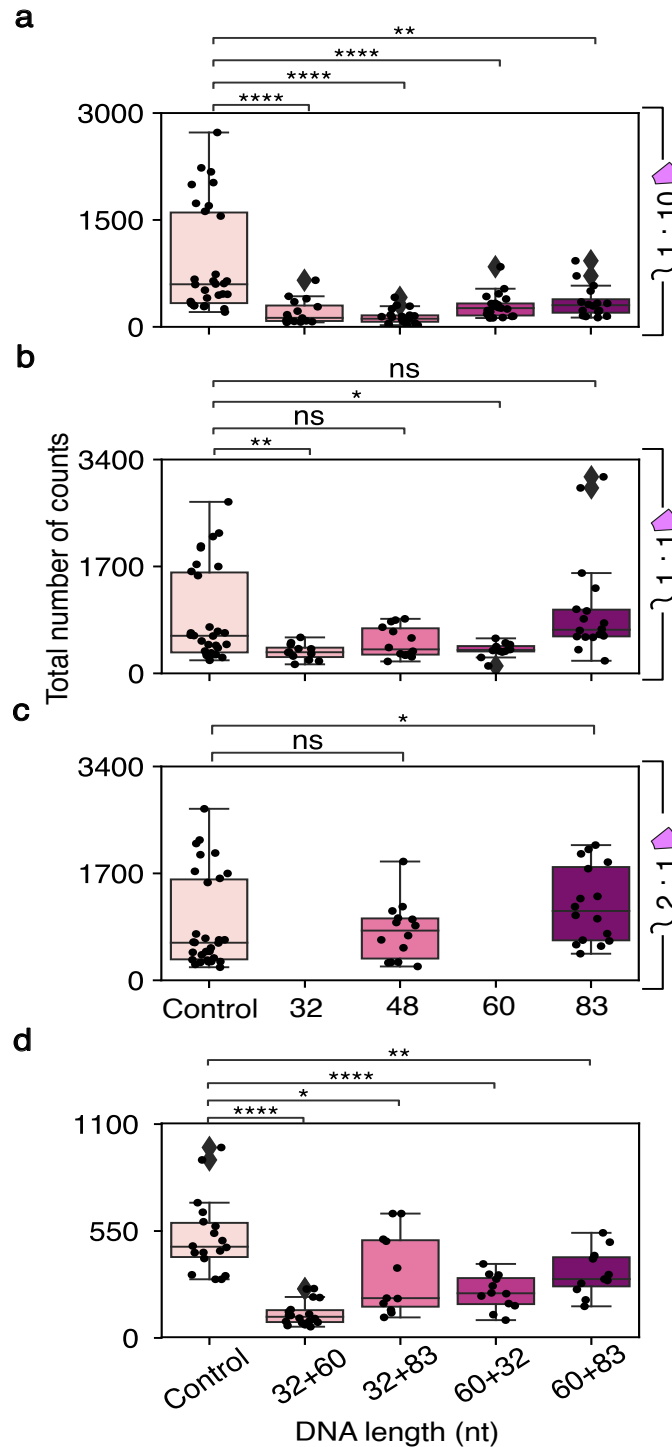
Figure S19: The total number of RAD52(209) binding events upon single-stranded DNA binding and during the annealing reactions. The single-stranded DNA molecules were 32-, 48-, 60-, and 83 nt at 10 nM **(a)**, 100 nM **(b)**, and 200 nM **(c)**. The protein concentration was 100 nM. **(d)** The total number of binding events during the annealing reaction. The protein concentration was 100 nM, and the DNA concentration was 10 nM. The boxplots depict mean values with SEM averaged from several repeats. The number of repeats is listed in Table S7. Each black circle is one reading, the black diamonds are the outliers. The two-sided Mann-Whitney test was used for the statistical analysis. The corresponding p-values are shown as asterisks or "ns".
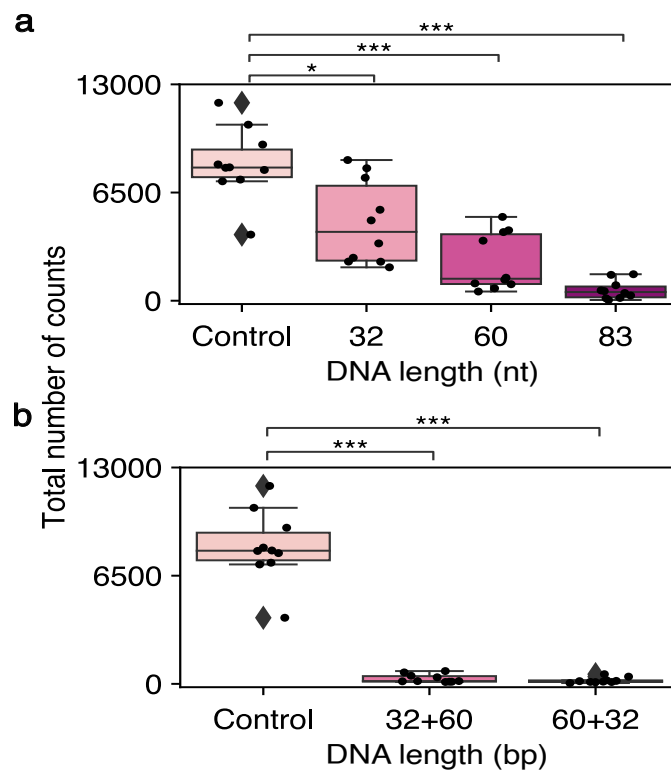
Figure S20: The total number of binding events for RAD52(209) at 500 nM upon single-stranded DNA binding and during the annealing reaction. The single-stranded DNA molecules were 32-, 60-, and 83-nt at 10 nM **(a)**. The number of binding events during the annealing reactions (32+60)- and (60+32)-nt **(b)**. In all the experiments the DNA concentration was 10 nM. The boxplots depict mean values with SEM averaged from different repeats. The number of repeats is listed in Table S5 and S7. Each black circle is one reading, the black diamonds are the outliers. The two-sided Mann-Whitney test was used for the statistical analysis. The corresponding p-values are shown as asterisks or "ns".
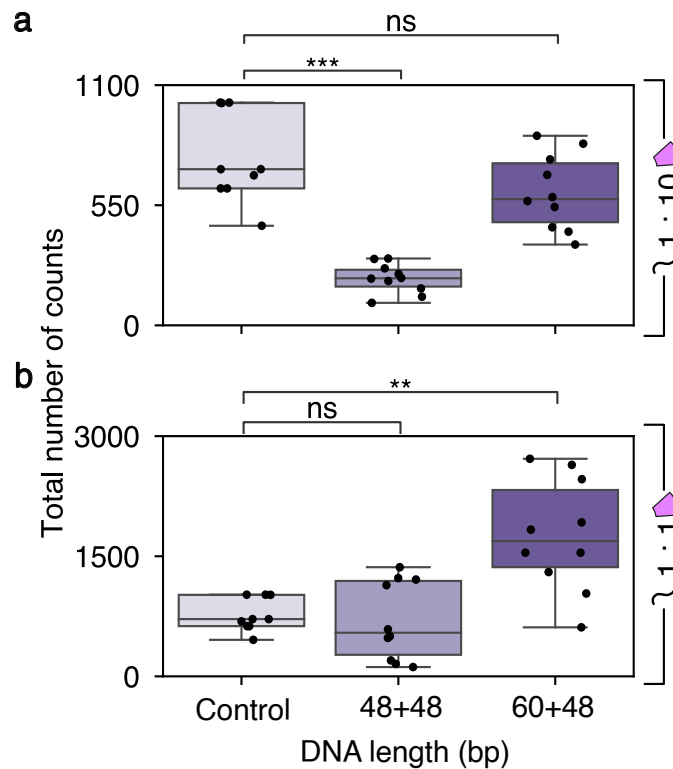
Figure S21: The total number of RAD52(209) counts upon binding to double-stranded DNA. The dsDNA molecules were (48+48)-nt (no overhangs) and (48+60)-nt (with overhangs). The DNA concentrations were 10 **(a)** and 100 nM **(b)**. The protein concentration was 100 nM. The boxplots depict mean values with SEM averaged from different repeats. The number of repeats is listed in Table S8. Each black circle is one reading, the black diamonds are the outliers. The two-sided Mann-Whitney test was used for the statistical analysis. The corresponding p-values are shown as asterisks or "ns".
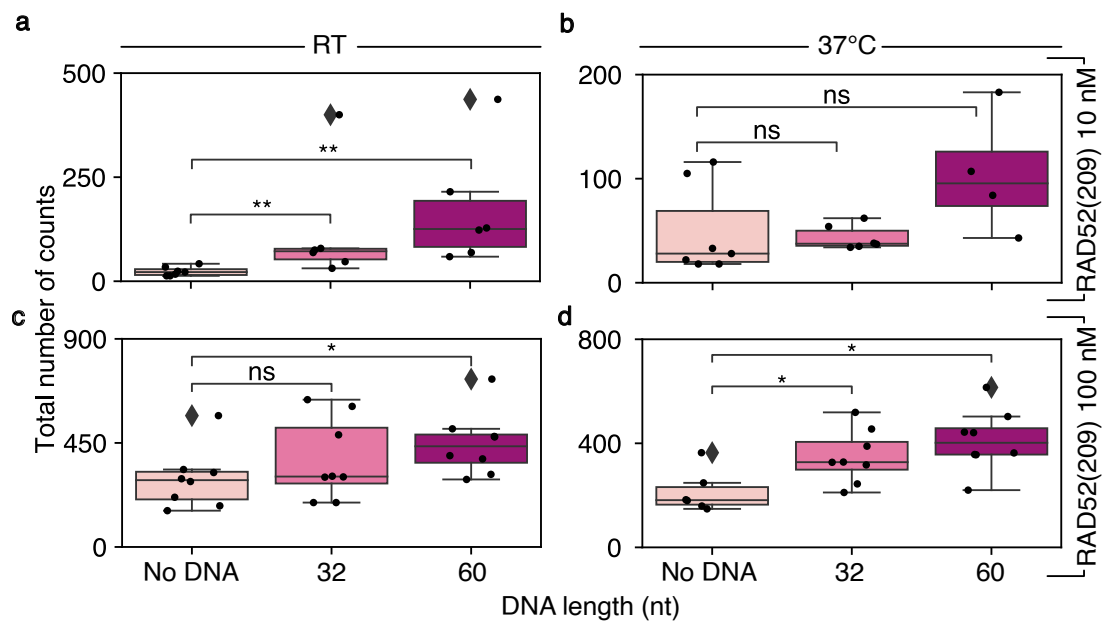
Figure S22: The total number of RAD52(209) binding events upon single-stranded DNA binding in the annealing buffer. The protein concentrations were 10 **(a)**-**(b)** and 100 nM **(c)**-**(d)**. DNA molecules were 32- and 60-nt long at 10 nM. The experiments were performed at 25∘C **(a)** and **(b)** and at 37∘C **(b)** and **(d)**. The two-sided Mann-Whitney test was utilized for the statistical analysis. Corresponding p-values are shown as asterisks or "ns". The number of repeats is listed in Table S6. One black circle is one reading, black diamonds are the outliers.
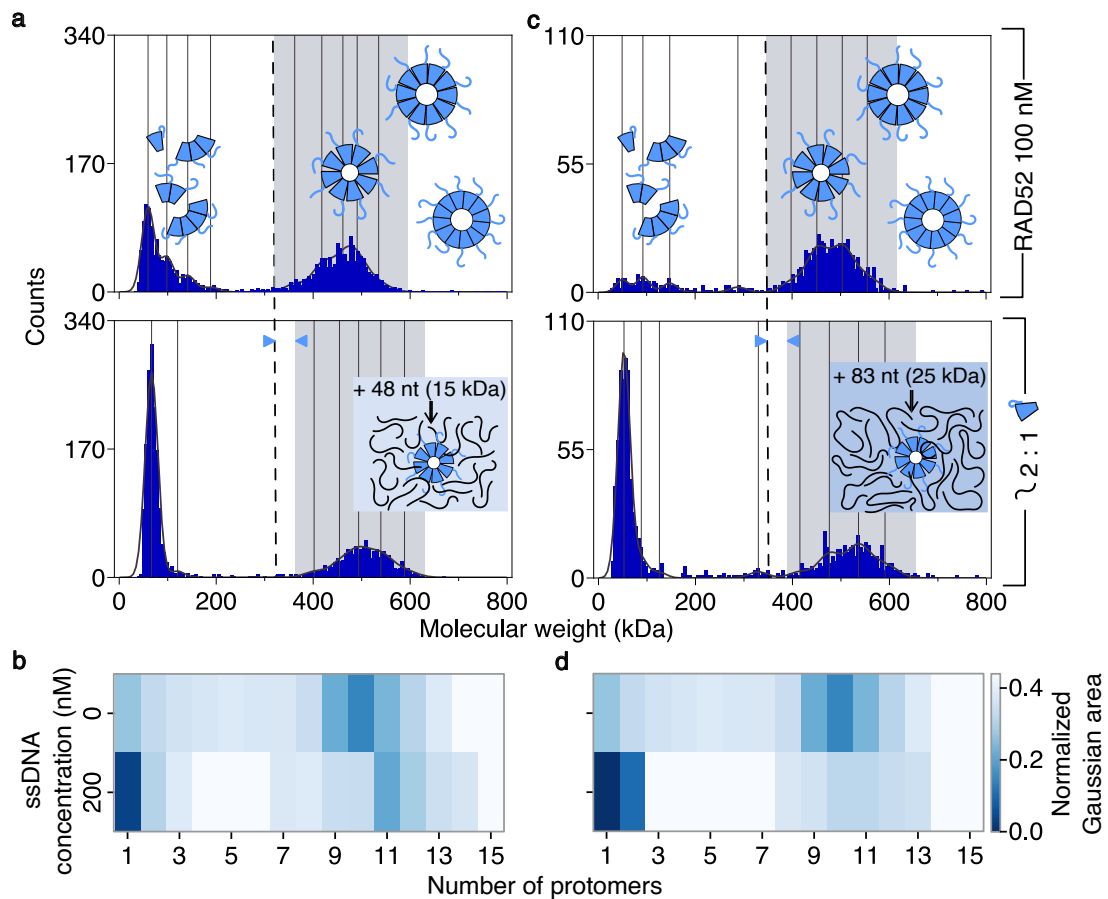
Figure S23: RAD52 binding the excess of the single-stranded DNA. The DNA molecules were 48-**(a-b)** and 83-nt long **(c-d)** at 200 nM. RAD52 concentration was 100 nM. Mass histograms display a cumulative distribution of several readings acquired on the same day. Grey areas indicate the variety of the ring structures. The blue arrows show the MW shift of the ring distribution. The dashed lines are the references from the control reaction (no DNA). The heatmaps **(b)** and **(d)** display normalized areas of distinct protomers as a cumulative sum of all the repeats.
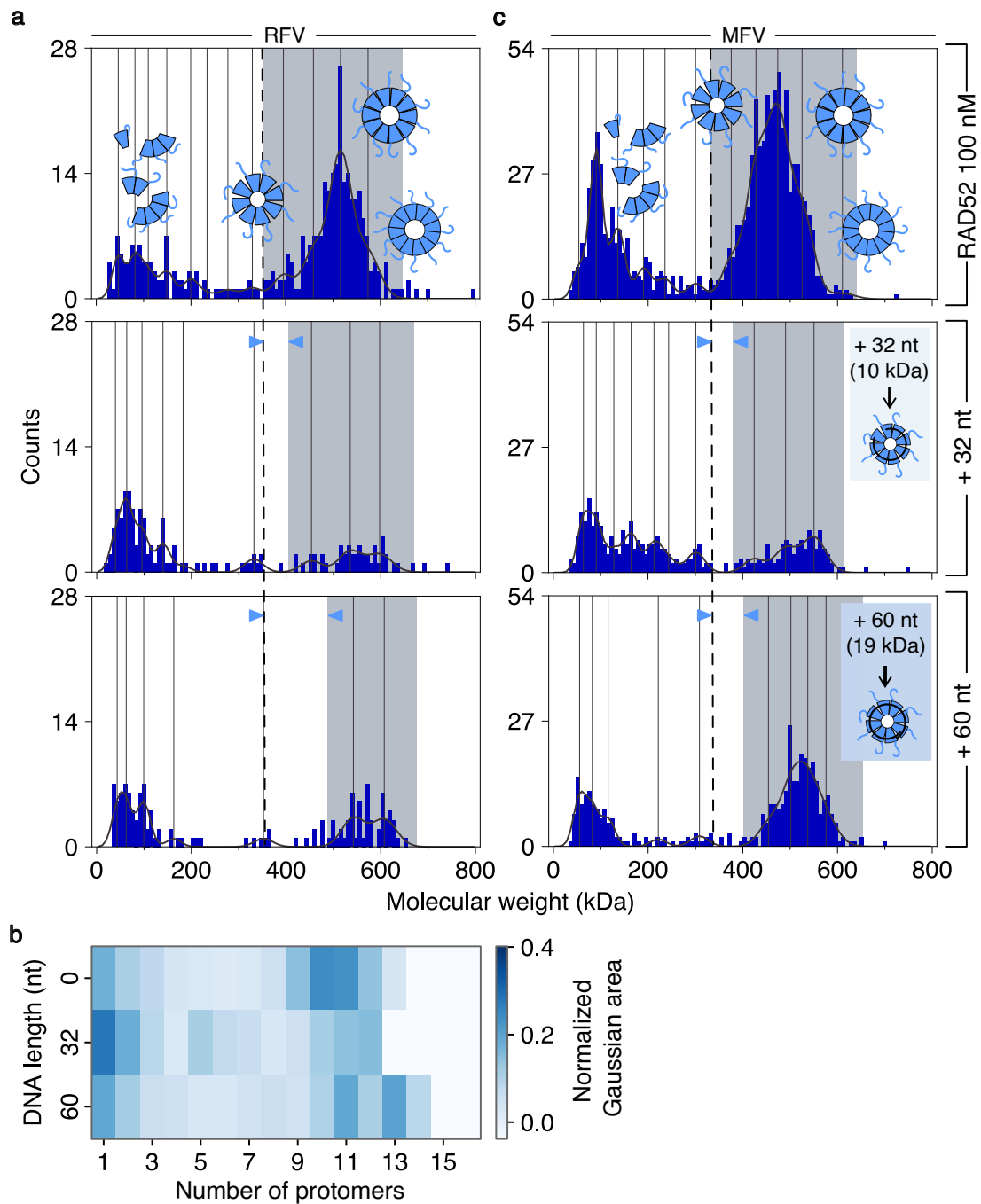
Figure S24: The oligomeric state of RAD52 upon binding to the single-stranded DNA in the annealing buffer. The utilized DNA molecules were 32- and 60-nt at 10 nM. The protein concentration was 100 nM. The experiments were performed at 37°C either in the RVF **(a)** or in the MVF **(c**. For demonstration purposes only, the dataset acquired in the MFV is presented, however, it was excluded from the analysis. **(b)** The heatmap displays normalized areas of distinct protomers as a cumulative sum of all the repeats.
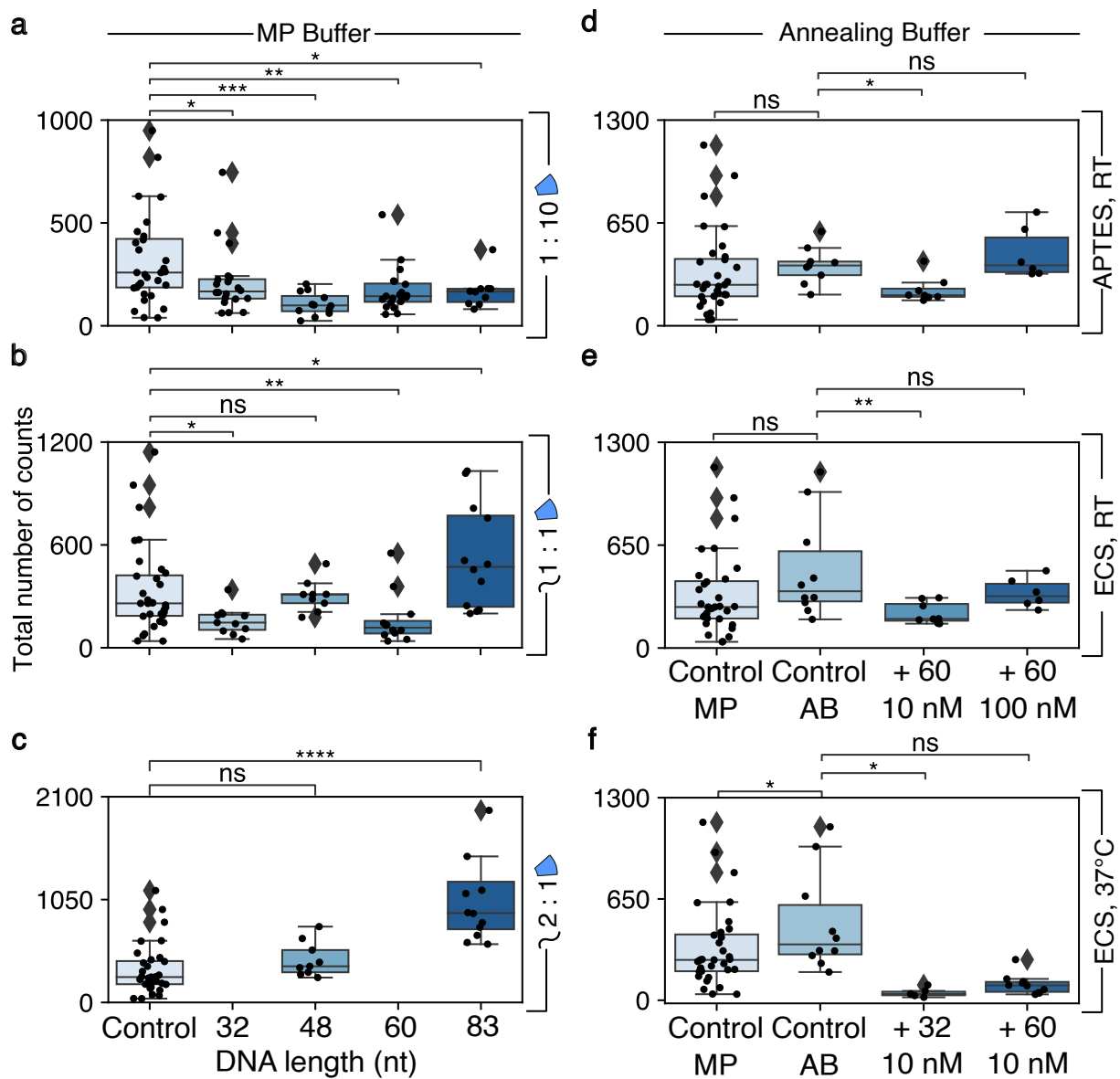
Figure S25: The total number of RAD52 binding events upon single-stranded DNA binding. The experiments were carried out in the mass photometry buffer **(a)**-**(c)** and in the annealing buffer **(d)**-**(f)**. The protein concentration was 100 nM. The utilized DNA molecules were - 32-, 48-, 60-, and 83-nt. DNA concentrations were 10 **(a)**, **(d)**, and **(f)**, 100 **(b)**, and 200 nM **(c)**. The protein binding was tested on the APTES **(d)** and easy-cleaned slides (ECS) **(e-f)**, at 25°C **(d-e)** and at 37°C **(f)**. A two-sided Mann-Whitney test was applied for the statistical analysis. The corresponding p-values are shown as asterisks or "ns". One black circle is one reading, black diamonds are the outliers.
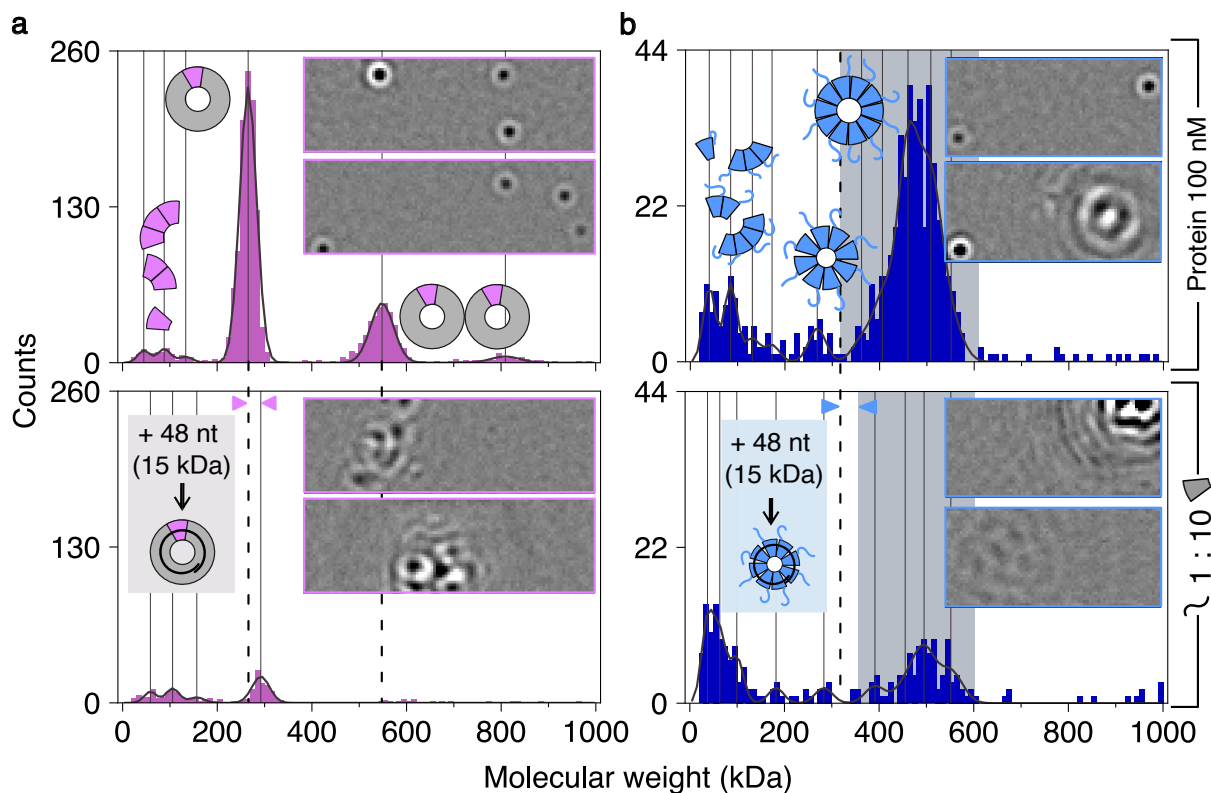
Figure S26: Protein clusters observed via mass photometry. RAD52(209) **(a)** and RAD52 **(b)** at 100 nM were mixed with the single-stranded DNA molecule (48-nt) at 10 nM. The proteins formed massive structures that were not quantified, however, they were visible in the RFV (inserts). Remarkably, at 100 nM RAD52 formed such structures alone, without DNA molecules. The vertical lines are the individual protomers, and the dashed lines are the references from the control reaction (no DNA). Grey areas demonstrate the variety of the ring structures formed by RAD52 **(b)**. The pink and blue arrows indicate the MW shifts of the ring structures upon DNA binding.
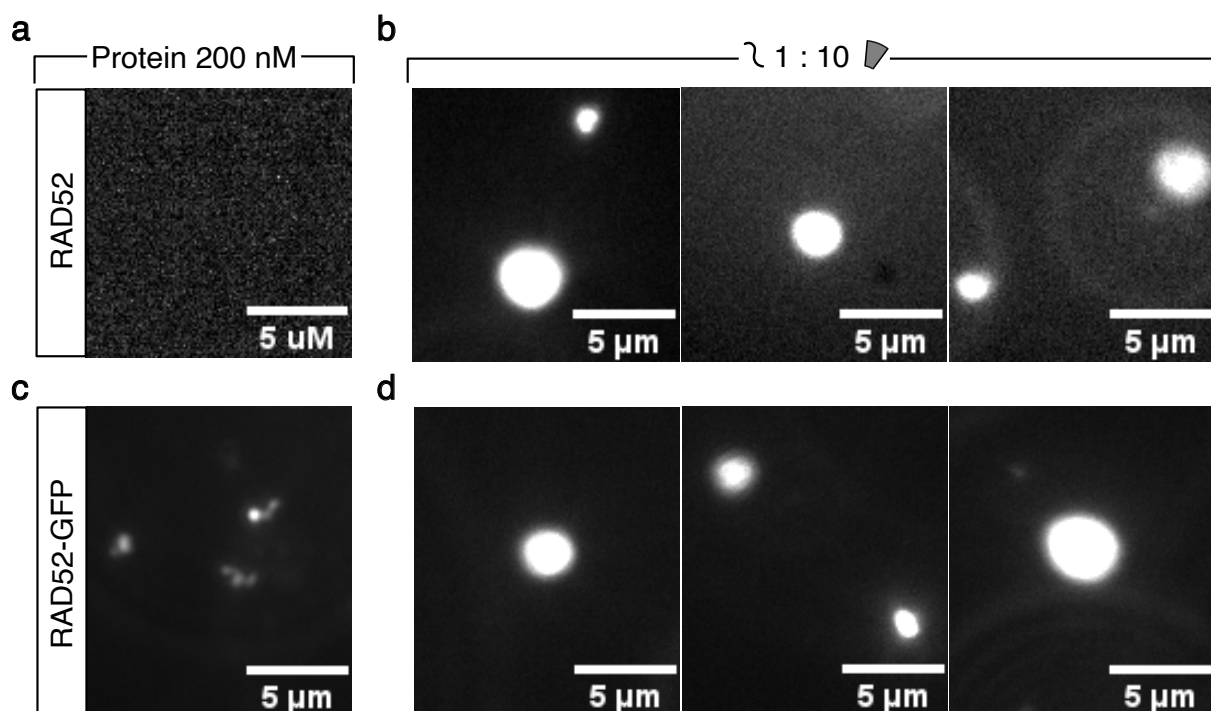
Figure S27: Clusters of RAD52 and GFP-RAD52 observed in the total internal reflection fluorescence microscopy. **(a)** RAD52 at 100 nM was imaged alone and in the presence of 48-ALEXA488 at 10 nM **(b)**. **(c)** GFP-RAD52 at 100 nM was imaged alone and in the presence of 32-ATTO565 at 10 nM **(d)**. All images demonstrate the green channel with depicted protein clusters. The scale bar is 5 μM.
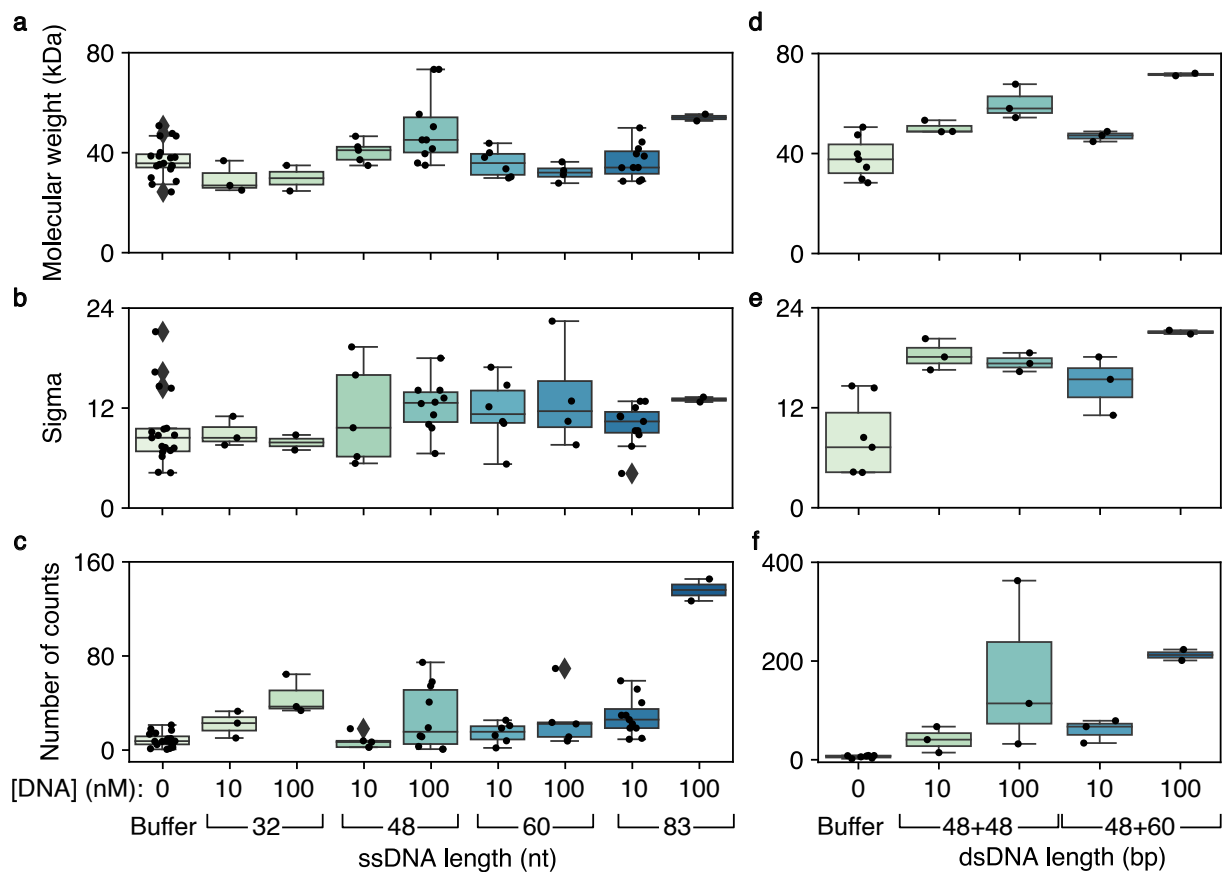
Figure S28: The mass photometry data of the single- and double-stranded DNA and the buffer. Single-stranded DNA molecules were 32-, 48-, 60-, and 83-nt long. The DNA concentrations were 10, 100, and 200 nM **(a-c)**. The dsDNA molecules were (48+48)- and (48+60)-nt long at 10 and 100 nM **(d-f)**. The boxplots show DNA molecular weights **(a)** and **(d)**, sigma values **(b)** and **(e)**, and the total number of counts **(c)** and **(f)**. Each boxplot contains corresponding parameters obtained from the mass photometry buffer (Buffer).
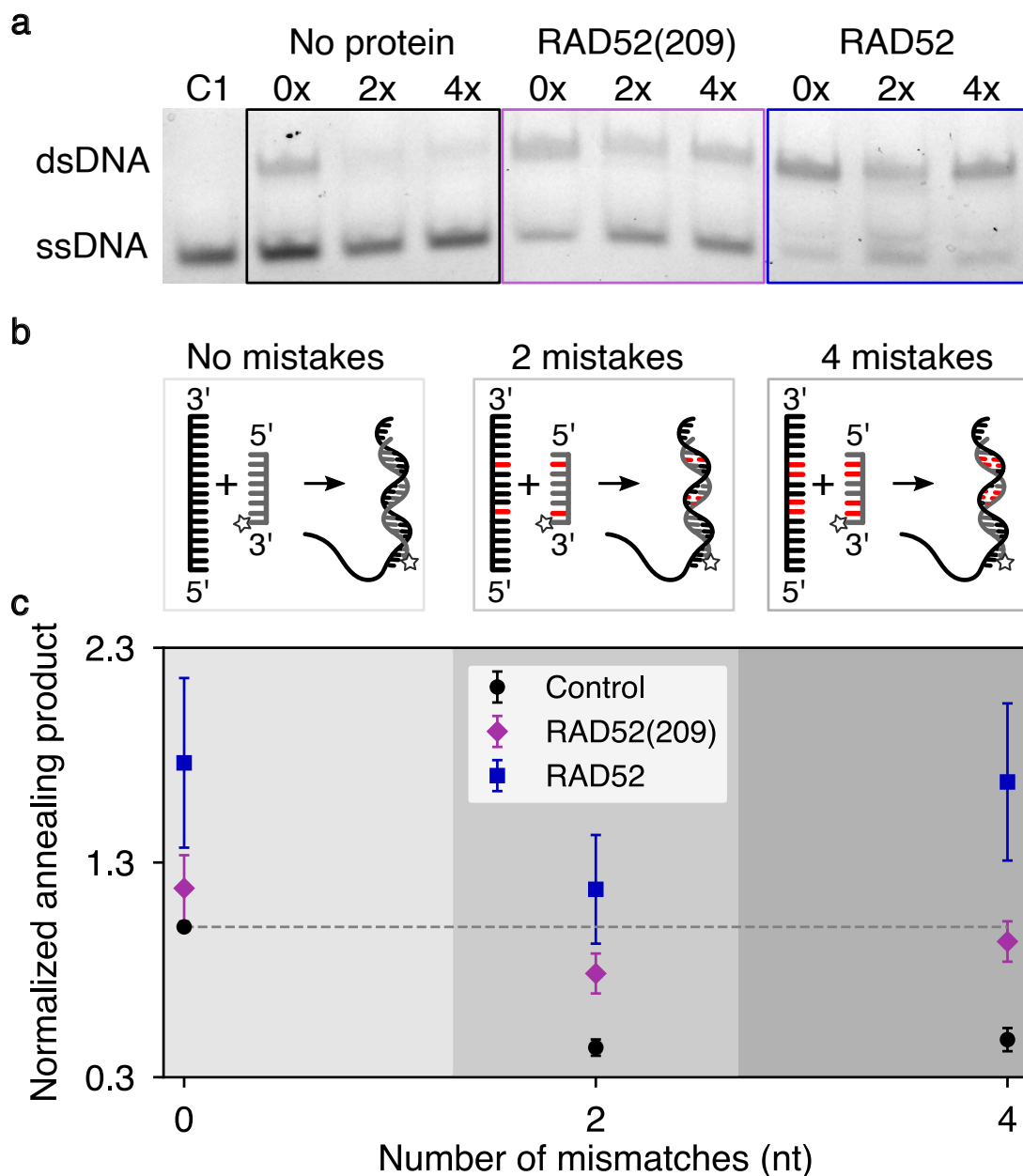
Figure S29: Single-stranded DNA annealing with mismatching oligonucleotides. 60-nt DNA molecules with two and four mismatches were purchased from Sigma-Aldrich, their sequences are listed in Table S1. The DNA concentration was 10 nM, and the protein concentration was 100 nM. **(a)** An exemplary gel demonstrating the intrinsic annealing (black rectangle), annealing promoted by RAD52(209) (magenta), and promoted by RAD52 (blue) with and without the mismatches. **(b)** Schematics demonstrates the location of the mismatches at 3'- and 5'-ends. **(c)** The intensities of the dsDNA bands were normalized to the control reaction (intrinsic annealing, no mismatches) and plotted as mean values with SEM. The experiment was repeated 4 times.
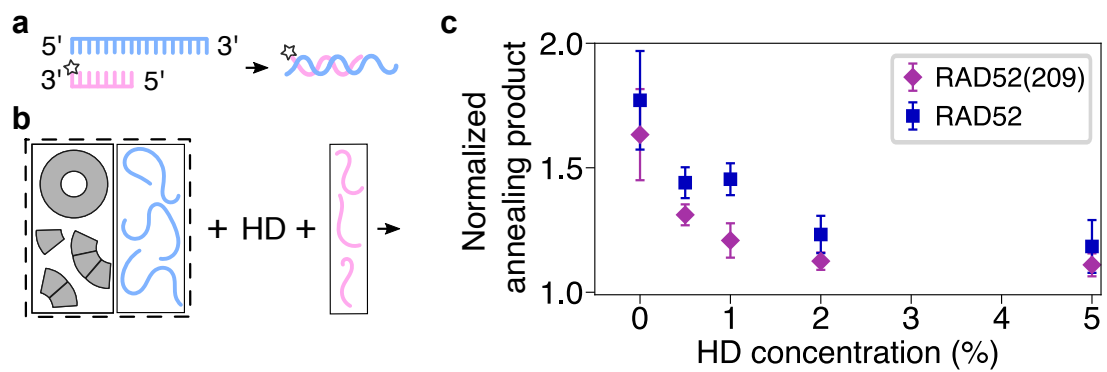
Figure S30: Single-stranded DNA annealing in the presence of 1,6-Hexanediol (HD). **(a)** The utilized DNA molecules were 32- and 60-nt at 10 nM. The protein concentration was 100 nM. **(b)** The protein was preincubated with 60-nt DNA at 37°C for 5 min following the incubation with HD at 37°C for 10 min. The HD concentration was utilized at 1, 2, and 5 %. **(c)** The intensities of the dsDNA bands were normalized to the control reaction (intrinsic annealing, no HD) and plotted as mean values with SEM. The experiment was repeated 9 times.
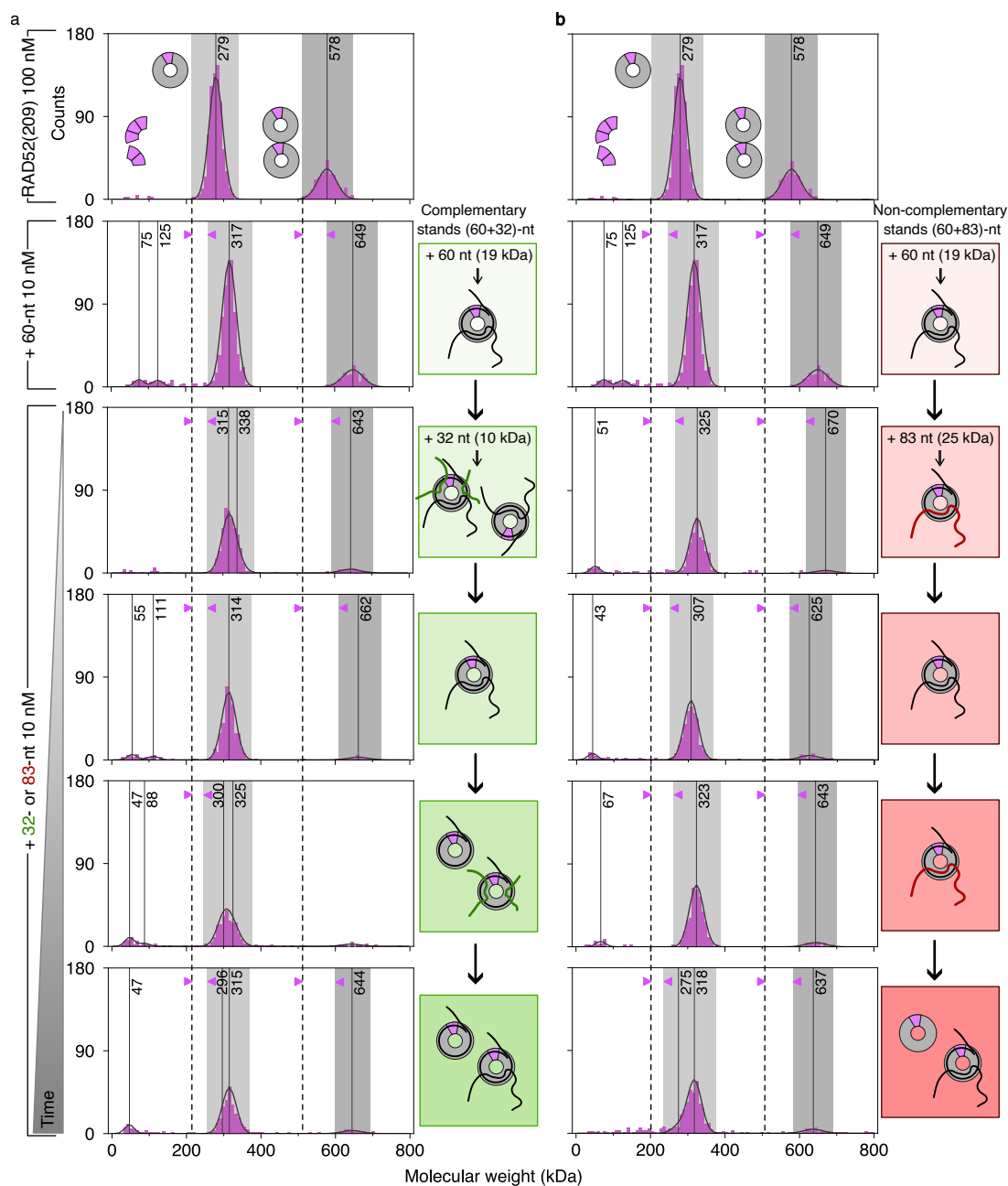
Figure S31: Single-stranded DNA annealing at different reaction points. The dynamics of the annealing reaction was tested with a pair of complementary strands (60+32)-nt **(a)**, and non-complementary ones (60+83)-nt **(b)**. DNA concentration was 10 nM. RAD52(209) concentration was 100 nM. The protein was preincubated with 60-nt for 5 min at 37°C and the second strand was added. Reaction fractions were measured after 2, 5, 7, and 9 min at 25°C. Grey areas indicate the distribution of the ring and the double-ring structures at different steps of the experiment. Dashed lines are references from the control reaction (no DNA). Pink arrows indicate the MW shifts. Schematics explain the individual steps of the experiments. 60-nt DNA is colored in black, the complementary 32-nt is green, and the non-complementary 83-nt is red.

114

Abbreviations

AB	annealing buffer
APTES	3-Aminopropyltriethoxysilan
ATP	Adenosine triphosphate
BER	Base-excision repair
bp	base-pair
BRCA	Breast Cancer protein
DDR	DNA damage response
DNA	Deoxyribonucleic acid
DSB	DNA double-strand break
dsDNA	double-stranded DNA
ECS	easy-cleaned slides
HB	hydrogen bonds
HR	homologous recombination
MFV	medium field of view
MP	mass photometry
NER	nucleotide excision repair
NHEJ	non-homologous end-joining
nt	nucleotide
OBD	oligonucleotide binding domain
PAGE	Polyacrylamide gel electrophoresis
RAD	Radiation sensitive
RFV	regular field of view
ROS	reactive oxygen species
RPA	Replication protein A
SDS	sodium dodecyl sulfate
SSA	single-strand annealing
SSB	single-strand binding
SSBR	Single-strand break repair
ssDNA	single-stranded DNA
TAE	Tris-acetate-EDTA

Contributions

The design of various constructs described in this work with subsequent protein purification was performed by Dr. Sivaraman Subramaniam (Dresden, BIOTEC, RG of Prof. Francis Stewart). With my suggestion, Dr. Sivaraman Subramaniam designed the constructs for RAD52 and GFP-RAD52. The alternative protein constructs for RAD52(209) and RAD52 with TEV recognition site were designed by Dr. Zhihao Jiang (ZMBP, RG of Prof. Lozano-Durán). The AlphaFold prediction models of RAD52(209) and RAD52 were created by Dr. Paul Gouguet (RUB, RG of Prof. Suayib Üstün).

The original Python script for the analysis of binding events acquired *via* mass photometry was written by Tobias J. Jachowski. The updated version of the software including the calibration procedure, simplified fitting procedure, and different ways to present the data has been written by me.

The preliminary mass photometry measurements of RAD52(209) were performed in the RG of Prof. Philipp Kukura (Oxford University) by Dr. Manish S. Kushwah and by Tobias J. Jachowski. Dr. Manish S. Kushwah has introduced me to the MP setup and to the analysis software.

All the mass photometry experiments presented in this work were performed and analyzed by me. Python scripts for the statistical analysis of the acquired data were written by me. All the biochemical assays, including various versions of single-stranded DNA binding and annealing, were adapted from the literature and performed by me with the following statistical analysis.

List of publications

- **ISCAM Analysis: A Python package to plot and analyze iSCAM data.**
  Tobias J.Jachowski and Maria A. Kharlamova.
  GitHub, **2020**.

- **Short oligomers of human RAD52 promote single-strand annealing.**
  Maria A. Kharlamova, Manish S. Kushwah, Tobias J. Jachowski, Sivaraman Sub-
  ramaniam, A. Francis Stewart, Philipp Kukura, Erik Schäffer.
  BioRxiv, **2023**.

- **TRAK1 enhances the motility of all human Kinesin-1 isoforms to pro-
  mote long-range transport.**
  Benedikt Fischer, Maria A. Kharlamova, Julius Dangelmaier, Malte Lemmin, Jo-
  hannes Busch, Daniel Lambton, Erik Schäffer.
  (in preparation)

- **His-tag promotes ring formation for the bacterial ancestor of human
  RAD52.**
  Maria A. Kharlamova, Sivaraman Subramaniam, A. Francis Stewart, Erik Schäffer.
  (in preparation)

List of Figures

List of Tables

Bibliography

[1] M. Akbari and H. E. Krokan. Cytotoxicity and mutagenicity of endogenous DNA base lesions as potential cause of human aging. *Mechanisms of ageing and development*, 129(7-8):353–365, 2008.

[2] A. Al-Fatlawi, M. Schroeder, and A. F. Stewart. The Rad52 SSAP superfamily and new insight into homologous recombination. *Communications Biology*, 6(1):1–4, 2023.

[3] K. H. Almeida and R. W. Sobol. A unified view of base excision repair: lesion-dependent protein complexes regulated by post-translational modification. *DNA repair*, 6(6):695–711, 2007.

[4] M. Ander, S. Subramaniam, K. Fahmy, A. Francis Stewart, and E. Schäffer. A single-strand annealing protein clamps DNA to detect and secure homology. *PLoS Biology*, 13(8):1–23, 2015.

[5] N. Arai, W. Kagawa, K. Saito, Y. Shingu, T. Mikawa, H. Kurumizaka, and T. Shibata. Vital roles of the second DNA-binding site of Rad52 protein in yeast homologous recombination. *Journal of Biological Chemistry*, 286(20):17607–17617, 2011.

[6] B. Balboni, R. Marotta, F. Rinaldi, S. Girotto, A. Cavalli, S. Biophysics, and T. P. Facility. Novel structural insights on full-length human RAD52: Cryo-EM and beyond. *BioRxiv*, pages 5–8, 2023.

[7] S. F. Banani, H. O. Lee, A. A. Hyman, and M. K. Rosen. Biomolecular condensates: Organizers of cellular biochemistry. *Nature Reviews Molecular Cell Biology*, 18(5):285–298, 2017.

[8] K. Baynton, M. Otterlei, M. Bjøras, C. von Kobbe, V. A. Bohr, and E. Seeberg. WRN interacts physically and functionally with the recombination mediator protein RAD52. *Journal of Biological Chemistry*, 278(38):36476–36486, 2003.

[9] S. Bekker-Jensen, C. Lukas, R. Kitagawa, F. Melander, M. B. Kastan, J. Bartek, and J. Lukas. Spatial organization of the mammalian genome surveillance machinery in response to DNA strand breaks. *Journal of Cell Biology*, 173(2):195–206, 2006.

[10] C. Bernstein, A. R., V. Nfonsam, and H. Bernstei. DNA Damage, DNA Repair and Cancer. In *New Research Directions in DNA Repair*. InTech, may 2013.

[11] R. Bhargava, D. O. Onyango, and J. M. Stark. Regulation of Single Strand Annealing. *Trends in Genetics*, 32(9):566–575, 2016.

[12] R. Bose, W. R. Karthaus, W. Abida, J. Phillip, Z. Zhang, J. Wongvipat, E. V. Wasmuth, N. Shah, S. Patrick, M. G. Doran, P. Wang, A. Patruno, Y. Zhao, D. Zheng, N. Schultz, C. L. Sawyers, P. Program, M. Sloan, K. Cancer, Y. Avenue, M. Sloan, K. Cancer, M. Sloan, K. Cancer, V. Gerstner, M. S. Kettering, and C. Chase. HHS Public Access. *Nature*, 546(7660):671–675, 2018.

[13] J. D. Bouchard and S. Moineau. Lactococcal phage genes involved in sensitivity to AbiK and their relation to single-strand annealing proteins. *Journal of Bacteriology*, 186(11):3649–3652, 2004.

[14] I. Brouwer, H. Zhang, A. Candelli, D. Normanno, E. J. Peterman, G. J. Wuite, and M. Modesti. Human RAD52 Captures and Holds DNA Strands, Increases DNA Flexibility, and Prevents Melting of Duplex DNA: Implications for DNA Recombination. *Cell Reports*, 18(12):2845–2853, mar 2017.

[15] D. V. Bugreev, O. M. Mazina, and A. V. Mazin. Rad54 protein promotes branch migration of Holliday junctions. *Nature*, 442(7102):590–593, 2006.

[16] S. Burma, B. P. C. Chen, and D. J. Chen. Role of non-homologous end joining (NHEJ) in maintaining genomic integrity. *DNA repair*, 5(9-10):1042–1048, 2006.

[17] D. Cahill, B. Connor, and J. P. Carney. Radiation Oncology Research Laboratory and Marlene and Stewart Greenebaum Cancer Center, University of Maryland School of Medicine, Baltimore, Maryland 21201. *Oncology Research*, pages 1958–1976, 2006.

[18] K. W. Caldecott. Mammalian single-strand break repair: mechanisms and links with chromatin. *DNA repair*, 6(4):443–453, 2007.

[19] B. J. Caldwell, A. Norris, V. H. Wysocki, and C. E. Bell. Structure of a Rad52 homolog from bacteriophage in complex with a novel duplex intermediate of DNA annealing. *Nature Communications*, pages 1–29, 2022.

[20] B. J. Caldwell, A. S. Norris, C. F. Karbowski, A. M. Wiegand, V. H. Wysocki, and C. E. Bell. Structure of a RecT/Redβ family recombinase in complex with a duplex intermediate of DNA annealing. *Nature Communications*, 13(1):1–14, 2022.

[21] R. Ceccaldi, B. Rondinelli, and A. D. D 'andrea. Repair Pathway Choices and Consequences at the Double- Strand Break Mechanisms of DNA DSB Repair. *Trends in Biochemical Sciences*, 26(1):52–64, 2016.

[22] G. Chandramouly, S. McDevitt, K. Sullivan, T. Kent, A. Luz, J. F. Glickman, M. Andrake, T. Skorski, and R. T. Pomerantz. Small-Molecule Disruption of RAD52 Rings as a Mechanism for Precision Medicine in BRCA-Deficient Cancers. *Chemistry and Biology*, 22(11):1491–1504, 2015.

[23] H. H. Chang, N. R. Pannunzio, N. Adachi, and M. R. Lieber. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nature Reviews Molecular Cell Biology*, 18(8):495–506, 2017.

[24] D. Cole, G. Young, A. Weigel, A. Sebesta, and P. Kukura. Label-Free Single-Molecule Imaging with Numerical-Aperture-Shaped Interferometric Scattering Microscopy. *ACS Photonics*, 4(2):211–216, 2017.

[25] F. Crick and J. Watson. Molecular structure of nucleic acids. *Nature*, 1953.

[26] J. M. Daley and T. E. Wilson. Rejoining of DNA Double-Strand Breaks as a Function of Overhang Length. *Molecular and Cellular Biology*, 25(3):896–906, 2005.

[27] M. de Jager, J. van Noort, D. C. van Gent, C. Dekker, R. Kanaar, and C. Wyman. Human Rad50/Mre11 is a flexible complex that can tether DNA ends. *Molecular cell*, 8(5):1129–1135, 2001.

[28] C. B. de la Tour, S. Boisnard, C. Norais, M. Toueille1, E. Bentchikou, F. Vannier, M. M. Cox2, S. Sommer, and P. Servant. The deinococcal DdrB protein is involved in an early step of DNA double strand break repair and in plasmid transformation through its single-strand annealing activity. *Bone*, 23(1):1–7, 2011.

[29] A. Del Mondo, M. Iovinella, M. Petriccione, A. Nunziata, S. J. Davis, D. Cioppa, and C. Ciniglia. A spotlight on Rad52 in cyanidiophytina (Rhodophyta): A relic in algal heritage. *Plants*, 8(2):1–11, 2019.

[30] X. Deng, A. Prakash, K. Dhar, G. S. Baia, C. Kolar, G. G. Oakley, and G. E. Borgstahl. Human replication protein A-Rad52-single-stranded DNA complex: Stoichiometry and evidence for strand transfer regulation by phosphorylation. *Biochemistry*, 48(28):6633–6643, 2009.

[31] J. Deveryshetty, R. Chadda, J. R. Mattice, S. Karunakaran, M. J. Rau, K. Basore, N. Pokhrel, N. Englander, J. A. J. Fitzpatrick, B. Bothner, and E. Antony. Yeast Rad52 is a homodecamer and possesses BRCA2-like bipartite Rad51 binding modes. *Nature Communications*, 2023.

[32] E. A. Epum, M. J. M. Id, N. P. Ruppe, and K. L. F. Id. Interaction of yeast Rad51 and Rad52 relieves Rad52-mediated inhibition of de novo telomere addition. *PLOS Genetics*, pages 1–27, 2020.

[33] A. Erler, S. Wegmann, C. Elie-Caille, C. R. Bradshaw, M. Maresca, R. Seidel, B. Habermann, D. J. Muller, and A. F. Stewart. Conformational Adaptability of Redβ during DNA Annealing and Implications for Its Structural Relationship with Rad52. *Journal of Molecular Biology*, 391(3):586–598, 2009.

[34] J. Essers, A. B. Houtsmuller, L. Van Veelen, C. Paulusma, A. L. Nigg, A. Pastink, W. Vermeulen, J. H. Hoeijmakers, and R. Kanaar. Nuclear dynamics of RAD52 group homologous recombination proteins in response to DNA damage. *EMBO Journal*, 21(8):2030–2037, 2002.

[35] D. N. H. S. Ester. DyLight Amine-Reactive Dyes. *Thermo Fisher Scientific*, 0747(46427), 2011.

[36] Z. Feng, S. P. Scott, W. Bussen, G. G. Sharma, G. Guo, T. K. Pandita, and S. N. Powell. Rad52 inactivation is synthetically lethal with BRCA2 deficiency. *Proceedings of the National Academy of Sciences of the United States of America*, 108(2):686–691, 2011.

[37] E. C. Friedberg, G. C. Walker, W. Siede, and R. D. Wood. *DNA repair and mutagenesis.* American Society for Microbiology Press, 2005.

[38] S. L. Gasior, A. K. Wong, Y. Kora, A. Shinohara, and D. K. Bishop. Rad52 associates with RPA and functions with Rad55 and Rad57 to assemble meiotic recombination complexes. *Genes and Development*, 12(14):2208–2221, 1998.

[39] B. Gibb, L. F. Ye, Y. Kwon, H. Niu, P. Sung, and C. Eric. Protein dynamics during presynaptic complex assembly on individual ssDNA molecules. *Nature Structural Molecular Biology*, 21(10):893–900, 2015.

[40] G. Giglia-Mari, A. Zotter, and W. Vermeulen. DNA damage response. *Cold Spring Harbor Perspectives in Biology*, 3(1):1–19, 2011.

[41] L. C. J. Gillet and O. D. Schärer. Molecular mechanisms of mammalian global genome nucleotide excision repair. *Chemical reviews*, 106(2):253–276, 2005.

[42] E. C. Greene. DNA sequence alignment during homologous recombination. *Journal of Biological Chemistry*, 291(22):11572–11580, 2016.

[43] J. M. Grimme, M. Honda, R. Wright, Y. Okuno, E. Rothenberg, A. V. Mazin, T. Ha, and M. Spies. Human Rad52 binds and wraps single-stranded DNA and mediates annealing via two hRad52-ssDNA complexes. *Nucleic Acids Research*, 38(9):2917–2930, 2010.

[44] K. Hanamshet, O. M. Mazina, and A. V. Mazin. Reappearance from obscurity: Mammalian Rad52 in homologous recombination. *Genes*, 7(9):1–18, 2016.

[45] B. Haubold, N. Pierstorff, F. Möller, and T. Wiehe. Genome comparison without alignment using shortest unique substrings. *BMC Bioinformatics*, 6:1–11, 2005.

[46] B. Haubold, N. Pierstorff, F. Möller, and T. Wiehe. Genome comparison without alignment using shortest unique substrings. *BMC Bioinformatics*, 6:1–11, 2005.

[47] M. L. Hegde, T. K. Hazra, and S. Mitra. Early steps in the DNA base excision/single-strand interruption repair pathway in mammalian cells. *Cell research*, 18(1):27–47, 2008.

[48] E. A. Hendrickson. RAD52: Viral friend or foe? *Cancers*, 12(2):1–11, 2020.

[49] S. R. Hengel, M. A. Spies, and M. Spies. Small-Molecule Inhibitors Targeting DNA Repair and DNA Repair Deficiency in Research and Cancer Therapy. *Cell Chemical Biology*, 24(9):1101–1119, 2017.

[50] J. H. J. Hoeijmakers. Nucleotide excision repair II: from yeast to mammals. *Trends in Genetics*, 9(6):211–217, 1993.

[51] J. H. J. Hoeijmakers. The consequences of DNA injury Genome maintenance mechanisms for preventing cancer. *Nature*, 411(6835):366–374, 2001.

[52] J. H. J. Hoeijmakers. DNA damage, aging, and cancer. *New England Journal of Medicine*, 361(15):1475–1485, 2009.

[53] M. Honda, Y. Okuno, J. Yoo, T. Ha, and M. Spies. Tyrosine phosphorylation enhances RAD52-mediated annealing by modulating its DNA binding. *EMBO Journal*, 30(16):3368–3382, 2011.

[54] M. Honda, M. Razzaghi, P. Gaur, E. Malacaria, L. D. Biagi, F. A. Aiello, E. A. Paintsil, A. Stanfield, B. J. Deppe, L. Gakhar, N. J. Schnicker, M. A. Spies, and P. Pichierri. Human RAD52 double-ring remodels replication forks restricting fork reversal. *BioRxiv*, 2023.

[55] C. Hou, C. McCown, D. N. Ivanov, and O. V. Tsodikov. Structural Insight into the DNA Binding Function of Transcription Factor ERF. *Biochemistry*, 59(47):4499–4506, 2020.

[56] L. M. Iyer, E. V. Koonin, and L. Aravind. Classification and evolutionary history of the single-strand. *BMC Genomics*, 11(1):1–11, 2002.

[57] T. Jachowski and M. Kharlamova. ISCAM Analysis: A Python package to plot and analyze iSCAM data. *Github*, 2020.

[58] D. Jackson, K. Dhar, J. K. Wahl, M. S. Wold, and G. E. Borgstahl. Analysis of the human replication protein A:Rad52 complex: Evidence for crosstalk between RPA32, RPA70, Rad52 and DNA. *Journal of Molecular Biology*, 321(1):133–148, 2002.

[59] M. Jalan, K. S. Olsen, and S. N. Powell. Emerging roles of RAD52 in genome maintenance. *Cancers*, 11(7), 2019.

[60] W. Kagawa, A. Kagawa, K. Saito, S. Ikawa, T. Shibata, H. Kurumizaka, and S. Yokoyama. Identification of a second DNA binding site in the human Rad52 protein. *Journal of Biological Chemistry*, 283(35):24264–24273, aug 2008.

[61] W. Kagawa, H. Kurumizaka, S. Ikawa, S. Yokoyama, and T. Shibata. Homologous Pairing Promoted by the Human, Rad52 Protein. *Journal of Biological Chemistry*, 276(37):35201–35208, 2001.

[62] W. Kagawa, H. Kurumizaka, R. Ishitani, S. Fukai, O. Nureki, T. Shibata, and S. Yokoyama. Crystal Structure of the Homologous-Pairing Domain from the Human Rad52 Recombinase in the Undecameric Form. Technical report, 2002.

[63] N. V. Khade and T. Sugiyama. Roles of C-terminal region of yeast and human rad52 in rad51-nucleoprotein filament formation and ssDNA annealing. *PLoS ONE*, 11(6), jun 2016.

[64] S. Kilic, A. Lezaja, M. Gatti, E. Bianco, J. Michelena, R. Imhof, and M. Altmeyer. Phase separation of 53 BP 1 determines liquid-like behavior of DNA repair compartments . *The EMBO Journal*, 38(16):1–17, 2019.

[65] J.-S. Kim, T. B. Krasieva, H. Kurumizaka, D. J. Chen, A. M. R. Taylor, and K. Yokomori. Independent and sequential recruitment of NHEJ and HR factors to DNA damage sites in mammalian cells. *The Journal of cell biology*, 170(3):341–347, 2005.

[66] 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. *Molecular Cell*, 14:363–374, 2004.

[67] T. Kinebuchi, W. Kagawa, H. Kurumizaka, and S. Yokoyama. Role of the N-terminal domain of the human DMC1 protein in octamer formation and DNA binding. *Journal of Biological Chemistry*, 280(31):28382–28387, 2005.

[68] C. Kinoshita, Y. Takizawa, M. Saotome, S. Ogino, H. Kurumizaka, and W. Kagawa. The cryo-EM structure of full-length RAD52 protein contains an undecameric ring. *FEBS Open Bio*, pages 1–11, 2023.

[69] C. Kinoshita, Y. Takizawa, M. Saotome, S. Ogino, H. Kurumizaka, and W. Kagawa. The cryo-EM structure of full-length RAD52 protein contains an undecameric ring. *FEBS Open Bio*, 13:408–418, 2023.

[70] K. Klapstein, T. Chou, and R. Bruinsma. Physics of RecA-mediated homologous recognition. *Biophysical Journal*, 87(3):1466–1477, 2004.

[71] O. B. Kramara, J. and A. Malkova. Break Induced Replication: the where, the why, and the how. *Physiology & behavior*, 176(1):139–148, 2016.

[72] K. Lee, Y. Zhang, and S. E. Lee. Saccharomyces cerevisiae ATM orthologue suppresses break-induced chromosome translocations. *Nature*, 454(7203):543–546, 2008.

[73] C. Lesterlin, G. Ball, L. Schermelleh, and D. J. Sherratt. RecA bundles mediate homology pairing between distant sisters during DNA break repair. *Nature*, 2014.

[74] Y. Li, W. B. Struwe, and P. Kukura. Single molecule mass photometry of nucleic acids. *Nucleic Acids Research*, 48(17):E97, 2020.

[75] T. Lindahl. Instability and decay of the primary structure of DNA. *Nature*, 362(6422):709–715, 1993.

[76] M. Lisby, U. H. Mortensen, and R. Rothstein. Colocalization of multiple DNA double-strand breaks at a single Rad52 repair center. *Nature Cell Biology*, 5(6):572–577, 2003.

[77] M. Lisby and R. Rothstein. Choreography of recombination proteins during the DNA damage response. *DNA Repair*, 8(9):1068–1076, 2009.

[78] M. Lisby, R. Rothstein, and U. H. Mortensen. Rad52 forms DNA repair and recombination centers during S phase. *Proceedings of the National Academy of Sciences of the United States of America*, 98(15):8276–8282, 2001.

[79] M. Liu and D. G. Schatz. Balancing AID and DNA repair during somatic hypermutation. *Trends in Immunology*, 30(4):173–181, 2009.

[80] Y. Liu and N. Maizels. Coordinated response of mammalian Rad51 and Rad52 to DNA damage. *EMBO Reports*, 1(1):85–90, 2000.

[81] B. Llorente, C. E. Smith, and L. S. Symington. Break-induced replication: What is it and what is it for? *Cell Cycle*, 7(7):859–864, 2008.

[82] J. A. Lloyd, A. L. Forget, and K. L. Knight. Correlation of biochemical properties with the oligomeric state of human Rad52 protein. *Journal of Biological Chemistry*, 277(48):46172–46178, 2002.

[83] J. A. Lloyd, D. A. McGrew, and K. L. Knight. Identification of residues important for DNA binding in the full-length human Rad52 protein. *Journal of Molecular Biology*, 345(2):239–249, 2005.

[84] C. Lukas, J. Falck, J. Bartkova, J. Bartek, and J. Lukas. Distinct spatiotemporal dynamics of mammalian checkpoint regulators induced by DNA damage. *Nature Cell Biology*, 5(3):255–260, 2003.

[85] B. L. Mahaney, K. Meek, and S. P. Lees-Miller. Repair of ionizing radiation-induced DNA double-strand breaks by non-homologous end-joining. *Biochemical Journal*, 417(3):639–650, 2009.

[86] O. M. Mazina, H. Keskin, K. Hanamshet, F. Storici, and A. V. Mazin. Rad52 Inverse Strand Exchange Drives RNA-Templated DNA Double-Strand Break Repair. *Molecular Cell*, 67(1):19–29.e3, 2017.

[87] L. Michael and R. Rodney. Choreography of Recombination Proteins during the DNA Damage Response. *Encyclopedia of Cancer*, 8(9):1307–1307, 2017.

[88] G. T. Milne and D. T. Weaver. Dominant negative alleles of RAD52 reveal a DNA repair/recombination complex including Rad51 and Rad52. *Genes and Development*, 7(9):1755–1765, 1993.

[89] T. Mimori and J. A. Hardin. Mechanism of interaction between Ku protein and DNA. *Journal of Biological Chemistry*, 261(22):10375–10379, 1986.

[90] M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, and M. Steinegger. ColabFold: making protein folding accessible to all. *Nature Methods*, 19(6):679–682, 2022.

[91] J. R. Mitchell, J. H. Hoeijmakers, and L. J. Niedernhofer. Divide and conquer: Nucleotide excision repair battles cancer and ageing. *Current Opinion in Cell Biology*, 15(2):232–240, 2003.

[92] E. Mladenov, C. Staudt, A. Soni, T. Murmann-Konda, M. Siemann-Loekes, and G. Iliakis. Strong suppression of gene conversion with increasing DNA double-strand break load delimited by 53BP1 and RAD52. *Nucleic Acids Research*, 48(4):1905–1924, 2020.

[93] T. A. Motycka, T. Bessho, S. M. Post, P. Sung, and A. E. Tomkinson. Physical and Functional Interaction between the XPF/ERCC1 Endonuclease and hRad52. *Journal of Biological Chemistry*, 279(14):13634–13639, 2004.

[94] A. A. Muhammad, T. Peterlini, J. Guirouilh-Barbat, X. Veaute, D. Busso, B. Lopez, J.-Y. Masson, G. Mazon, E. L. Cam, and P. Dupaigne. Human RAD52 stimulates the RAD51-mediated homology search. *bioRxiv*, page 2022.06.27.497340, 2022.

[95] A. Nair, R. Agarwal, and R. K. Chittela. Biochemical characterization of plant Rad52 protein from rice (Oryza sativa). *Plant Physiology and Biochemistry*, 106:108–117, 2016.

[96] J. H. New and S. C. Kowalczykowski. Rad52 protein has a second stimulatory role in DNA strand exchange that complements replication protein-A function. *Journal of Biological Chemistry*, 277(29):26171–26176, 2002.

[97] T. P. Newing, J. L. Brewster, L. J. Fitschen, J. C. Bouwer, N. P. Johnston, H. Yu, and G. Tolun. Redβ177 annealase structure reveals details of oligomerization and λ Red-mediated homologous DNA recombination. *Nature Communications*, 13(1):1–14, 2022.

[98] A. V. Nimonkar, A. Z. Özsoy, J. Genschel, P. Modrich, and S. C. Kowalczykowski. Human exonuclease 1 and BLM helicase interact to resect DNA and initiate DNA repair. *Proceedings of the National Academy of Sciences*, 105(44):16906–16911, 2008.

[99] P. Noirot and R. D. Kolodner. DNA Strand Invasion Promoted by Escherichia coli RecT Protein. *Journal of Biological Chemistry*, 273(20):12274–12280, 1998.

[100] C. A. Norais, S. Chitteni-Pattu, E. A. Wood, R. B. Inman, and M. M. Cox. DdrB protein, an alternative Deinococcus radiodurans SSB induced by ionizing radiation. *Journal of Biological Chemistry*, 284(32):21402–21411, 2009.

[101] U. C. Obodo, E. A. Epum, M. H. Platts, J. Seloff, N. A. Dahlson, S. M. Velkovsky, S. R. Paul, and K. L. Friedman. Endogenous Hot Spots of De Novo Telomere Addition in the Yeast Genome Contain Proximal Enhancers That Bind Cdc13 . *Molecular and Cellular Biology*, 36(12):1750–1763, 2016.

[102] R. Oshidari, R. Huang, M. Medghalchi, E. Y. Tse, N. Ashgriz, H. O. Lee, H. Wyatt, and K. Mekhail. DNA repair by Rad52 liquid droplets. *Nature Communications*, 11(1), dec 2020.

[103] M. Otterlei, P. Bruheim, B. Ahn, W. Bussen, P. Karmakar, K. Baynton, and V. A. Bohr. Werner syndrome protein participates in a complex with RAD51, RAD54, RAD54B and ATR in response to ICL-induced replication arrest. *Journal of cell science*, 119(24):5137–5146, 2006.

[104] M. S. Park, D. L. Ludwig, E. Stigger, and S. H. Lee. Physical interaction between human RAD52 and RPA is required for homologous recombination in mammalian cells. *Journal of Biological Chemistry*, 271(31):18996–19000, 1996.

[105] C. A. Parsons, P. Baumann, E. V. Dyck, and S. C. West. Precise binding of single-stranded DNA termini by human RAD52 protein. *The EMBO Journal*, 19(15):4175–4181, 2000.

[106] S. I. Passy, X. Yu, Z. Li, C. M. Radding, and E. H. Egelman. Rings and filaments of β protein from bacteriophage λ suggest a superfamily of recombination proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 96(8):4279–4284, 1999.

[107] S. I. Passy, X. Yu, Z. Li, C. M. Radding, J. Y. Masson, S. C. West, and E. H. Egelman. Human Dmc1 protein binds DNA as an octameric ring. *Proceedings of the National Academy of Sciences of the United States of America*, 96(19):10684–10688, 1999.

[108] T. T. Paull and M. Gellert. The 3 to 5 exonuclease activity of Mre11 facilitates repair of DNA double-strand breaks. *Molecular cell*, 1(7):969–979, 1998.

[109] T. T. Paull, E. P. Rogakou, V. Yamazaki, C. U. Kirchgessner, M. Gellert, and W. M. Bonner. A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. *Current Biology*, 10(15):886–895, 2000.

[110] F. Pessina, F. Giavazzi, Y. Yin, U. Gioia, V. Vitelli, A. Galbiati, S. Barozzi, M. Garre, A. Oldani, A. Flaus, R. Cerbino, D. Parazzoli, E. Rothenberg, and F. d'Adda di Fagagna. Functional transcription promoters at DNA double-strand breaks mediate RNA-driven phase separation of damage-response factors. *Nature Cell Biology*, 21(10):1286–1299, 2019.

[111] P. E. Pestryakov and O. I. Lavrik. Mechanisms of single-stranded DNA-binding protein functioning in cellular DNA metabolism. *Biochemistry (Moscow)*, 73(13):1388–1404, 2008.

[112] E. F. Pettersen, T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, T. I. Croll, J. H. Morris, and T. E. Ferrin. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science*, 30(1):70–82, 2021.

[113] I. Plate, L. Albertsen, M. Lisby, S. C. Hallwyl, Q. Feng, C. Seong, R. Rothstein, P. Sung, and U. H. Mortensen. Rad52 multimerization is important for its nuclear localization in Saccharomyces cerevisiae. *DNA Repair*, 7(1):57–66, 2008.

[114] M. Ploquin, A. Bransi, E. R. Paquet, A. Z. Stasiak, A. Stasiak, A. M. Cieslinska, E. H. Egelman, and S. Moineau. Functional and Structural Basis for a Bacteriophage Homolog of Human RAD52. *Current Biology*, 18(15):1142–1146, 2008.

[115] A. R. Poteete and A. C. Fenton. DNA-binding properties of the Erf protein of bacteriophage P22. *Journal of Molecular Biology*, 163(2):257–275, 1983.

[116] A. R. Poteete, R. T. Sauer, and R. W. Hendrix. Domain structure and quaternary organization of the bacteriophage P22 Erf protein. *Journal of Molecular Biology*, 171(4):401–418, 1983.

[117] E. Raderschall, E. I. Golub, and T. Haaf. Nuclear foci of mammalian recombination proteins are located at single-stranded DNA regions formed after DNA damage. *Proceedings of the National Academy of Sciences of the United States of America*, 96(5):1921–1926, 1999.

[118] W. Ranatunga, D. Jackson, J. A. Lloyd, A. L. Forget, K. L. Knight, and G. E. Borgstahl. Human RAD52 Exhibits Two Modes of Self-association. *Journal of Biological Chemistry*, 276(19):15876–15880, may 2001.

[119] M. J. Rossi, S. F. DiDomenico, M. Patel, and A. V. Mazin. RAD52: Paradigm of Synthetic Lethality and New Developments. *Frontiers in Genetics*, 12(November):1–16, 2021.

[120] E. Rothenberg, J. M. Grimme, M. Spies, and T. Ha. Human Rad52-mediated homology search and annealing occurs by continuous interactions between overlapping nucleoprotein complexes. *Proceedings of the National Academy of Sciences of the United States of America*, 105(51):20274–20279, 2008.

[121] K. Rothkamm, S. B. JayneMoquet, M. Ellender, Z. Rana, and S. Burdak-Rothkamm. DNA Damage Foci: Meaning and Significance. *Environmental and Molecular Mutagenesis*, 405(April):391–405, 2010.

[122] A. Samach, C. Melamed-Bessudo, N. Avivi-Ragolski, S. Pietrokovski, and A. A. Levy. Identification of plant RAD52 homologs and characterization of the arabidopsis thaliana RAD52-like genes. *Plant Cell*, 23(12):4266–4279, 2011.

[123] M. Saotome, K. Saito, T. Yasuda, H. Ohtomo, S. Sugiyama, Y. Nishimura, H. Kurumizaka, and W. Kagawa. Structural Basis of Homology-Directed DNA Repair Mediated by RAD52. *iScience*, 3:50–62, may 2018.

[124] A. A. Sartori, C. Lukas, J. Coates, M. Mistrik, S. Fu, J. Bartek, R. Baer, J. Lukas, and S. P. Jackson. Human CtIP promotes DNA end resection. *Nature*, 450(7169):509–514, 2007.

[125] E. Scaltriti, S. Moineau, H. Launay, J. Y. Masson, C. Rivetti, R. Ramoni, V. Campanacci, M. Tegoni, and C. Cambillau. Deciphering the function of lactococcal phage ul36 Sak domains. *Journal of Structural Biology*, 170(3):462–469, 2010.

[126] C. Seong, M. G. Sehorn, I. Plate, I. Shi, B. Song, P. Chi, U. Mortensen, P. Sung, and L. Krejci. Molecular anatomy of the recombination mediator function of Saccharomyces cerevisiae Rad5. *Journal of Biological Chemistry*, 283(18):12166–12174, 2008.

[127] Z. Shen, K. G. Cloud, D. J. Chen, and M. S. Park. Specific interactions between the human RAD51 and RAD52 proteins. *Journal of Biological Chemistry*, 271(1):148–152, 1996.

[128] 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. *Genes to Cells*, 3(3):145–156, 1998.

[129] D. A. Sinclair and P. Oberdoerffer. The ageing epigenome: damaged beyond repair? *Ageing research reviews*, 8(3):189–198, 2009.

[130] M. R. Singleton, L. M. Wentzell, Y. Liu, S. C. West, and D. B. Wigley. Structure of the single-strand annealing domain of human RAD52 protein. Technical Report 21, 2002.

[131] C. E. Smith and C. E. Bell. Domain Structure of the Redβ Single-Strand Annealing Protein: The C-terminal Domain is Required for Fine-Tuning DNA-binding Properties, Interaction with the Exonuclease Partner, and Recombination in vivo. *Journal of Molecular Biology*, 428(3):561–578, 2016.

[132] S. K. Sotiriou, I. Kamileri, N. Lugli, K. Evangelou, C. Da-Ré, F. Huber, L. Padayachy, S. Tardy, N. L. Nicati, S. Barriot, F. Ochs, C. Lukas, J. Lukas, V. G. Gorgoulis, L. Scapozza, and T. D. Halazonetis. Mammalian RAD52 Functions in Break-Induced Replication Repair of Collapsed DNA Replication Forks. *Molecular Cell*, 64(6):1127–1134, dec 2016.

[133] E. Soutoglou, J. F. Dorn, K. Sengupta, M. Jasin, A. Nussenzweig, T. Ried, G. Danuser, and T. Misteli. Positional stability of single double-strand breaks in mammalian cells. *Nature Cell Biology*, 9(6):675–682, 2007.

[134] L. Spagnolo, A. Rivera-Calzada, L. H. Pearl, and O. Llorca. Three-dimensional structure of the human DNA-PKcs/Ku70/Ku80 complex assembled on DNA and its implications for DNA DSB repair. *Molecular cell*, 22(4):511–519, 2006.

[135] V. Spegg, A. Panagopoulos, M. Stout, A. Krishnan, G. Reginato, R. Imhof, B. Roschitzki, P. Cejka, and M. Altmeyer. Phase separation properties of RPA combine high-affinity ssDNA binding with dynamic condensate functions at telomeres. *Nature Structural and Molecular Biology*, 30(4):451–462, 2023.

[136] A. Z. Stasiak, E. Larquet, A. Stasiak, S. Müller, A. Engel, E. V. Dyck, S. C. West, and E. H. Egelman. The human Rad52 protein exists as a heptameric ring. Technical report, 2000.

[137] B. Stefanovie, S. R. Hengel, J. Mlcouskova, J. Prochazkova, M. Spirek, F. Nikulenkov, D. Nemecek, B. G. Koch, F. E. Bain, L. Yu, M. Spies, and L. Krejci. DSS1

interacts with and stimulates RAD52 to promote the repair of DSBs. *Nucleic Acids Research*, 48(2):694–708, 2020.

[138] T. Stiff, M. O'Driscoll, N. Rief, K. Iwabuchi, M. Lobrich, and P. A. Jeggo. ATM and DNA-PK function redundantly to phosphorylate H2AX after exposure to ionizing radiation. *Cancer research*, 64(7):2390–2396, 2004.

[139] S. Subramaniam, E. Axel, J. Fu, A. Kranz, J. Tang, M. Gopalswamy, S. Ramakrishnan, A. Keller, G. Grundmeier, D. Müller, M. Sattler, and A. F. Stewart. DNA annealing by Red β is insufficient for homologous recombination and the additional requirements involve intra- and inter-molecular interactions. *Nature Publishing Group*, (August):1–14, 2016.

[140] S. Sugiman-Marangos and M. S. Junop. The structure of DdrB from Deinococcus: A new fold for single-stranded DNA binding proteins. *Nucleic Acids Research*, 38(10):3432–3440, 2010.

[141] S. N. Sugiman-marangos, Y. M. Weiss, and M. S. Junop. Mechanism for accurate, protein-assisted DNA annealing by Deinococcus radiodurans DdrB. *PNAS*, 113(16):1–6, 2016.

[142] 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. *Journal of Biological Chemistry*, 277(35):31663–31672, 2002.

[143] T. Sugiyama, J. H. New, and S. C. Kowalczykowski. DNA annealing by Rad52 protein is stimulated by specific interaction with the complex of replication protein A and single-stranded DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 95(11):6049–6054, 1998.

[144] K. Sullivan, K. Cramer-Morales, D. L. McElroy, D. A. Ostrov, K. Haas, W. Childers, R. Hromas, and T. Skorski. Identification of a Small Molecule Inhibitor of RAD52 by Structure-Based Selection. *PLoS ONE*, 11(1):1–11, 2016.

[145] S. Takeda, K. Nakamura, Y. Taniguchi, and T. T. Paull. Ctp1/CtIP and the MRN complex collaborate in the initial steps of homologous recombination. *Molecular cell*, 28(3):351–352, 2007.

[146] M. Tanaka, A. M. Earl, H. A. Howell, M. J. Park, J. A. Eisen, S. N. Peterson, and J. R. Battista. Analysis of Deinococcus radiodurans's transcriptional response to ionizing radiation and desiccation reveals novel proteins that contribute to extreme radioresistance. *Genetics*, 168(1):21–33, 2004.

[147] M. Tarsounas, D. Davies, and S. C. West. BRCA2-dependent and independent formation of RAD51 nuclear foci. *Oncogene*, 22(8):1115–1123, 2003.

[148] A. J. Van Brabant, T. Ye, M. Sanz, J. L. German, N. A. Ellis, and W. K. Holloman. Binding and melting of D-loops by the Bloom syndrome helicase. *Biochemistry*, 39(47):14617–14625, 2000.

[149] E. Van Dyck, N. M. Hajibagheri, A. Stasiak, and S. C. West. Visualisation of human Rad52 protein and its complexes with hRad51 and DNA. *Journal of Molecular Biology*, 284(4):1027–1038, 1998.

[150] L. R. van Veelen, J. Essers, M. W. M. M. van de Rakt, H. Odijk, A. Pastink, M. Z. Zdzienicka, C. C. Paulusma, and R. Kanaar. Ionizing radiation-induced foci formation of mammalian Rad51 and Rad54 depends on the Rad51 paralogs, but not on Rad52. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 574(1-2):34–49, 2005.

[151] M. M. Vilenchik and A. G. Knudson. Endogenous DNA double-strand breaks: Production, fidelity of repair, and induction of cancer. *Proceedings of the National Academy of Sciences of the United States of America*, 100(22):12871–12876, 2003.

[152] J. D. Watson and F. H. Crick. Genetical implications of the structure of deoxyribonucleic acid, 1953.

[153] L. Wei, S. Nakajima, S. Böhm, K. A. Bernstein, Z. Shen, M. Tsang, A. S. Levine, and L. Lan. DNA damage during the G0/G1 phase triggers RNA-templated, Cockayne syndrome B-dependent homologous recombination. *Proceedings of the National Academy of Sciences of the United States of America*, 112(27):E3495–E3504, 2015.

[154] S. Welty, Y. Teng, Z. Liang, W. Zhao, L. H. Sanders, J. T. Greenamyre, M. E. Rubio, A. Thathiah, R. Kodali, R. Wetzel, A. S. Levine, and L. Lan. RAD52 is required for RNA-templated recombination repair in post-mitotic neurons. *Journal of Biological Chemistry*, 293(4):1353–1362, 2018.

[155] E. Weterings and D. C. Van Gent. The mechanism of non-homologous end-joining: a synopsis of synapsis. *DNA repair*, 3(11):1425–1435, 2004.

[156] Y. Wu, N. Kantake, T. Sugiyama, and S. C. Kowalczykowski. Rad51 protein controls Rad52-mediated DNA annealing. *Journal of Biological Chemistry*, 283(21):14883–14892, 2008.

[157] C. Wyman and R. Kanaar. DNA double-strand break repair: all's well that ends well. *Annu. Rev. Genet.*, 40:363–383, 2006.

[158] C. Wyman, D. Ristic, and R. Kanaar. Homologous recombination-mediated double-strand break repair. *DNA repair*, 3(8-9):827–833, 2004.

[159] G. Xu, H. Lu, L. Wang, H. Chen, Z. Xu, Y. Hu, B. Tian, and Y. Hua. DdrB stimulates single-stranded DNA annealing and facilitates RecA-independent DNA repair in Deinococcus radiodurans. *DNA Repair*, 9(7):805–812, 2010.

[160] Z. Yan, C. Xue, S. Kumar, J. B. Crickard, Y. Yu, W. Wang, N. Pham, Y. Li, H. Niu, P. Sung, E. C. Greene, and G. Ira. Rad52 Restrains Resection at DNA Double-Strand Break Ends in Yeast. *Molecular Cell*, 76(5):699–711.e6, 2019.

[161] G. Young, N. Hundt, D. Cole, A. Fineberg, J. Andrecka, A. Tyler, A. Olerinyova, A. Ansari, E. G. Marklund, M. P. Collier, S. A. Chandler, O. Tkachenko, J. Allen, M. Crispin, N. Billington, Y. Takagi, J. R. Sellers, C. Eichmann, P. Selenko, L. Frey, R. Riek, M. R. Galpin, W. B. Struwe, J. L. P. Benesch, and P. Kukura. Quantitative mass imaging of single molecules HHS Public Access. *Science*, 360(6387):423–427, 2018.

Acknowledgements

At the end of this long, sophisticated, and yet fascinating journey, I would like to express my gratitude to many people I met on this way.

I would like to thank Prof. Dr. Erik Schäffer for assigning me to this project and supervising me. It was very challenging and equally amusing work. I am grateful for this. I would like to express my sincere gratitude to the committee members, Dr. Ana J. García-Sáez, Dr. John Weir, and Prof. Dr. Klaus Harter. Prof. Dr. Isabel Monte, Dr. Hauke Drechsler, Dr. Serapion Pyrpassopoulos, Dr. Carolina Carrasco Pulido, and Dr. Paul Gouguet for supporting me in my PostDoc application journey, for always finding time to listen and give good advice. I cannot express enough how much I appreciate your input! Additional acknowledgments to Ionas, Carolina, and Paul for proofreading my thesis. I am grateful for all the valuable comments and feedback from you. You made this work better. I thank Dr. Anita Jannasch for the thesis formatting comments and for all administrative support and advice during these years.

I want to thank my fantastic lab members who made me feel great at work and outside the lab! I was happy going to work every day because of the people I saw in the lab and at the Institute. I am grateful I have experienced that! My special appreciation goes to the lab members with whom I started and who navigated me in the lab and shared their experiences - Steve, Gero, and, especially, Tobi! Carolina, Ionas, Viktoria, Bene, Shu Yao, Yannic, and Aleksandr, you made every day brighter. I appreciate beyond anything the great time we spent together! I will cherish our friendship as a priceless gift I was not expecting to have. Many warm thanks to all the lab members! Thanks to our curious Bachelor and Master students for a great scientific and working environment! Thanks to all the colleagues from the 6th floor and the ZMBP for a friendly and collaborative atmosphere. Special big thanks to our fantastic secretary, Charlotte-Consuelo Rehm, for always being there for us, for constant care and support! And many thanks to our new secretary, Kevin, for stepping in and making this transition so smooth and comfortable!

I want to express my gratitude to my collaborators in Oxford and Dresden for always their feedback and suggestions on the project - Prof. Dr. Philipp Kukura, Dr. Manish S. Kushwah, Prof. Dr. Francis Stewart, and Dr. Sivaraman Subramaniam.

I want to thank Dr. John Weir and his lab members, especially Dr. Veronika Altmannova,

Elizaveta Selezneva, and Linda Chen, for being so welcoming and supportive. For always making me feel at their lab like in my own! I want to express my gratitude and thank the MPI of Biology, IMPRS program, and Ph.D. Coordination Office, especially Sarah, Sibylle, Jeanette, and Alejandra, for their enormous help and support!

I want to thank my good friends and IMPRS mates, most of whom I met because of the welcoming and collaborative environment at the MPI - Mikel, Nidhi, Layla, Lisa, and Saskia. Additionally, I would like to acknowledge my good friends whom I met through the network at ZMBP, Erin and Dennis, Irina and Luba. Thank you for sharing this time with me.

In this part, I would like to express my deep appreciation to people who might not be doing science with me every day, but without their love and encouraging words, I would have never made it - to my family and to my very close friends. I am very grateful to my fantastic parents, who supported me at every step and made me believe I can achieve everything. I am thankful to my little brother, who always makes me feel that being a scientist is like being a superhero. My special acknowledgment goes to my husband, who always encourages me to go forward and supports me along the way. Thank you for being by my side all the time! I would not make it without you. Thank you for all the valuable suggestions and advice during my PhD, for all scientific discussions and debates. I thank my beautiful best friends, Anastasia and Maria, for always being there for me! We have lived through this experience together, and I will remember that.

Statement of authorship

<u>Statement of authorship</u>

I hereby certify that

- I have composed this thesis by myself,

- all references and verbatim extracts have been quoted, and all sources of information have been specifically acknowledged,

- this thesis has not been accepted in any previous application for a degree, neither in total nor in substantial parts.

<u>Eidesstattlicher Versicherung</u>

Ich versichere hiermit, dass ich

- die vorliegende Arbeit selbstständig verfasst habe,

- keine anderen als die angegebenen Quellen benutzt und alle wörtlich oder sinngemäß aus anderen Werken übernommenen Aussagen als solche gekennzeichnet habe,

- description und die eingereichte Arbeit weder vollständig noch in wesentlichen Teilen Gegenstand eines anderen Prüfungsverfahrens gewesen ist.

....
Ort, Datum   Unterschrift/Signature
Place, Date   Maria Kharlamova