5 FORMATION OF RAD51 FILAMENTS

VISUALIZATION AND QUANTIFICATION OF RAD51 FILAMENTS AT SINGLE-MONOMER RESOLUTION

The RAD51 nucleoprotein filament is the molecular species responsible for the reaction of homologous recombination (HR), an essential pathway of double-strand break DNA repair. Fundamental aspects of the mechanism of RAD51 filament-formation concerning its nucleation, selectivity for single-stranded DNA and growth rate, remain unclear. Here we combine fluorescence microscopy, optical tweezers and microfluidics to quantify the assembly of RAD51 filaments with single-monomer resolution. We show that the filaments are seeded from RAD51 nuclei that are heterogeneous in size, reflecting the presence of a range of RAD51 oligomeric species in solution. We propose that the heterogeneity arises from the energetic balance between RAD51 self-assembly in solution and the size-dependent interaction of the nuclei with DNA. The process of RAD51 nucleation is intrinsically selective, strongly favoring filament formation on single-stranded DNA. Furthermore, using a new single-molecule fluorescence recovery after photobleaching (FRAP) assay we can independently observe filament nucleation and growth, permitting direct measurement of the cooperativity of RAD51 filament formation. Our findings yield a comprehensive quantitative understanding of RAD51 function in HR and could make the suggested regulation by recombination mediators such as BRCA2 redundant. Finally, the strong preference of RAD51 for ssDNA implies a structural basis for recently proposed break-independent role of RAD51 at active replication forks.
5.1 INTRODUCTION

The formation of RAD51-nucleoprotein filaments on ssDNA is a crucial step in homologous recombination (HR), an important DNA-repair pathway activated in response to DNA double-stranded breaks. In vivo, nucleoprotein filament assembly and disassembly are regulated by accessory proteins [159, 142]. Defects in HR are involved in carcinogenesis [73] and are targeted by anti-tumour drugs [110]. The tumour suppressor protein BRCA2, for example, directly interacts with RAD51 and is involved in the assembly of RAD51 filaments on ssDNA. It has been proposed that RAD51-filament formation follows a two-step mechanism, consisting of a nucleation and a growth phase [109, 169, 6, 72]. Nucleation, the minimal form of stable RAD51 binding, has been proposed to represent the critical step of filament formation [76]. Currently, the physical mechanism underlying the minimal size requirement is unknown and the number of RAD51 monomers constituting a single nucleus has never been measured directly.

Also, RAD51 can form filaments on both ssDNA and dsDNA. A proposed role for recombination mediators, is to remodel RAD51 filaments [124] and selectively load RAD51 onto ssDNA [27, 151, 83, 165, 28]. However, experiments comparing RAD51-filament formation rates on ssDNA and dsDNA and the effect of recombinase mediators in modulating RAD51 binding selectivity have yielded contrasting outcomes [105, 109, 27, 162, 68, 28]. The persistence of these controversies is due to the fact that the direct visualization and counting of individual RAD51 proteins binding on ssDNA and dsDNA of identical sequence and length has, until now, remained impossible.

In this study, we combined fluorescence microscopy, dual-trap optical tweezers, micro-fluidics [25] and force-induced DNA melting [60] to monitor individual RAD51 nuclei on both ssDNA and dsDNA with single-monomer resolution. From this analysis, we have built a physical model that explains both the minimal size requirement for RAD51 nucleation and the reported non-linearity in the binding rate observed at the single-molecule level [55, 169, 72, 12]. Next, we adapted the FRAP technique to our single-molecule assay and determined nucleation and growth rates separately, therefore measuring directly the cooperativity of RAD51 polymerization for ssDNA and dsDNA at different forces without the need of any mathematical modeling. Using force-induced melting of duplex DNA
and gapped DNA constructs we revealed unequivocally a strong intrinsic binding preference of RAD51 for ssDNA. Finally, our approach was used to clarify the effect of BRCA2 BRC4 on RAD51 filament formation on ss-DNA in different assembly conditions.

5.2 MATERIALS & METHODS

5.2.1 Combined optical tweezers, fluorescence microscopy and micro-fluidics setup

The experimental setup has been described elsewhere in detail. In brief, our combined dual optical trapping and single-molecule fluorescence setup is built around an inverted microscope (Nikon Eclipse TE2000-U). Optical traps are generated by a powerful near-IR laser (Ventus 1064 nm, 3W, Laser Quantum, Cheshire, Uk). An optical isolator (Newport 1030-1080 ISO-FRDY-OPT) is placed directly in front of the laser output to prevent coupling of back-reflections into the laser cavity. A combination of a zero order half-wave plate (WPH10M-1064, Thorlabs) and a polarizing beam splitter (10BC16PC.9, Newport, Irvine, CA) is used as a power regulator. A second zero order half-wave plate (WPH10M-1064, Thorlabs) and polarizing beam splitter (10BC16PC.9, Newport, Irvine, CA) are used to set the linear polarization and split the laser into two independent beams. After splitting, a Galilean telescope (1:2.67) is used to obtain the desired beam diameter (approximately 8 mm) and to steer the trap position. In one path, computer-controlled steering is achieved by displacing the first telescope lens with motorized actuators (T-LA28, Zaber Technologies Inc., Richmond, BC, Canada). The two beams are recombined using a second polarizing beam splitter (10BC16PC.9, Newport, Irvine, CA) and coupled into a high-NA water-immersion objective (PlanApo 60X, NA 1.2, Nikon) via a dichroic mirror (950 dcsp, Chroma Tech Corp, Rockingham, VT). The focusing power of the objective produces a three-dimensional optical trap where individual beads can be trapped and manipulated. Force detection is achieved by collecting the transmitted trapping light via a high NA oil-immersion condenser (Achromat/Aplanat, NA 1.4, Nikon) and imaged on a position-sensitive diode (DL100-7PCBA3 - Pacific Silicon Sensor) using a
single achromatic lens. Rejection of the unwanted polarization is achieved by a polarizing beam splitter (10BC16PC.9, Newport, Irvine, CA).

The end-to-end distance of the DNA is monitored by acquiring bright-field images of the trapped beads. Illumination of the sample is performed by a blue LED (LXHL-NB98 Luxeon Star/O, Lumileds) a single lens and a dichroic mirror and images are recorded using a CCD camera (CCD-902K, Wattec) and digitized by a frame grabber.

Fluorescence excitation is performed using wide-field strategy. AlexaFluor 555 is excited by a 532 nm laser (GCL-0.25L, 25 mW, Crystal laser, Reno, NV). The fluorescence excitation beam is first expanded and then coupled into the microscope via a dichroic mirror (z532rdc, Chroma Technology Corp.) and a lens. The emitted fluorescence signal is passed through a band-pass filter (hq575/50m, Chroma Technology Corp.) and imaged on an EMCCD camera (Cascade 512B, Princeton instruments, Monmouth Junction, NJ) and read using the Winview software.

A glass, custom-fabricated, multi-channel laminar flow-system is used to obtain parallel flow and to rapidly exchange buffer during the experiment. Image analysis is performed using a custom-written program in Labview (National Instruments).

5.2.2 RAD51 fluorescent labeling

RAD51 (isoform Q313, variant C319S) fluorescent labelling with Alexa555 was performed as previously described [114]. The degree of labelling, as characterized using mass spectrometry was 1.3. This method allowed us to exclude the presence of unlabelled RAD51 proteins in the preparation. Therefore we expect that 1 out 4 RAD51 proteins contains a non-specifically attached fluorophore. Biochemical characterization showed that RAD51 (C319S) is proficient in ATP hydrolysis, strand-exchange and DNA binding. Also this fluorescent RAD51 mutant has been extensively used and characterized in previously published single-molecule work [135, 114, 173].

5.2.3 Preparation of DNA construct

To produce a 38.412 kbp construct having biotin labels on the ends of the same strand we started our reaction by digesting Phage Lambda dsDNA (0.25 mg/mL) with Apal Fast Digest (30 minutes at 37°C). This reaction
yielded two different products, a 10 kbp and a 38.4 kbp dsDNA. Afterwards, in the same tube, we added Klenow buffer (final concentration 0.28x), nucleotides (dGTP, dTTP, final concentration 0.067 mM), biotin-modified nucleotides (biotin-14-dACTP and biotin-14-dCTP, final concentration 0.043 mM) and Klenow DNA polymerase exo- (final concentration 0.07 U/µL). The polymerization was carried out at 37°C for 30 minutes. The reaction was heat-inactivated (10 minutes at 75°C) and ethanol precipitated. The pellet was resuspended in 50 µL of 10 mM Tris pH 7.5.

The second part of the protocol consisted in annealing of a primer containing 4 biotin d-TTP (incubation 15 minutes at 65°C, primer sequence: 5'-cTcTcTcTcTctctctctcttgccc 3', final concentrations 1 µM, capital letters indicate biotin-positions). The hybridized primer was ligated with T4 DNA Ligase (45 minutes at RT, T4 DNA Ligase final concentration: 1 U/µl). The reaction was heat inactivated (5 minutes at 65°C) and ethanol precipitated.

To produce a 48,512 nt ssDNA melting Lambda DNA was biotinylated at the 3’ and 5’ end of the same strand using three oligonucleotides. First, the 5’-end of Lambda DNA and oligonucleotide 1 (5’- ggg cgg cga cct gga caa-3’) and 2 (5’- agg tgc ccc tcc ttt tgt TtT-3’) were phosphorylated for 30 min at 37°C in a reaction containing 14 nM Lambda DNA or 10 µM of the oligonucleotide and 0.25 U/µl of T4 Polynucleotide Kinase in 1x T4 Ligase buffer (Fermentas). Next, oligonucleotides 1 and 3 (5’- TtT tTt ttt aga gta ctg tac gat cta gca tca atc ttg tcc-3’) were annealed to the overhangs of Lambda DNA in a 10:1 oligonucleotide: DNA ratio (total volume 500 µl) by heating the reaction to 65°C and slowly cooling down to room temperature. The ligation reaction was then initiated by adding T4 DNA ligase (0.02 U/µl) and carried out for two hours. Next, oligonucleotide 2 was annealed in a 100:1 ratio to the Lambda DNA construct by incubation at 45°C for 30 minutes. Subsequently, the oligonucleotide was ligated to the DNA at room temperature. Finally, the DNA was purified by ethanol precipitation.

5.2.4 Experimental Conditions

Beads catching, DNA tethering, RAD51 incubation and fluorescence imaging were performed in the following conditions (unless otherwise mentioned): 25 mM Tris pH 7.5, 100 mM KCl, 1 mM CaCl2, 0.5 mM ATP and 10 mM DTT. DNA overstretching and melting was performed in 10 mM Tris pH 7.5, 25 mM KCl, 1 mM CaCl2, 0.5 mM ATP and 10 mM DTT.
5.2.5 Fluorescence Anisotropy measurements

Fluorescence anisotropy measurements were recorded in a PHERAstar FS plate reader (BMG Labtech; Germany) equipped with a fluorescence polarisation optic module (excitation = 485 nm; emission = 520 nM). The anisotropy was measured at 37°C in black 96-well half area, flat bottom, NBS plates (Corning; USA). The instrument was set to top optic measurement mode and the gain and focal height were set using the instrument software against free fluorescein (Fluorescein Sodium; Fluka) in 50 mM Tris-HCl pH 8.0 (mP = 35).

Wild-type human RAD51 (100 nM) was incubated with increasing concentrations of human BRC4 peptide at room temperature for 10 minutes in binding buffer (50 mM Tris-Acetic Acid pH 8.0, 150 mM KCl, 2 mM MgCl₂, 2.2 mM ATP, 1 mM DTT). After equilibration for 5 minutes at 37°C, the binding reaction was initiated by the addition of 10 nM (final concentration) Fluorescein-labelled ssDNA, using the PHERAstar FS built-in injectors. Reading intervals were 0.2 second for the first 2.5 seconds, followed by 1 second intervals for 1 minute, using 50 flashes per data point. Steady-state measurements were taken every 5 minutes, using 200 flashes per time point until the signal was stable.

RAD51 concentrations were determined by UV spectroscopy using theoretical extinction coefficients determined from the amino acids sequence (http://www.expasy.ch/tools/protparam.html). Human BRC4 peptide (LKE PTL LGF HTA SGK KVK ESL DKV KNL FDE KEQ M) concentration was determined by amino acid analysis (Protein and Nucleic Acid Chemistry Facility, Department of Biochemistry, University of Cambridge). A 60 nucleotide long 5'-fluorescein-labeled ssDNA was purchased from Integrated DNA Technologies (IDT, USA). The DNA sequence (5' F-ATG GTG TGT GTA GGT TAA TGT GAG GAG GAG AGG TGA AGA AGG AGG AGA GAA GAA GGA GGC 3') was designed to minimize secondary structure and dimer formation using the NUPACK nucleic acid package (http://www.nupack.org).
5.3 RESULTS

5.3.1 RAD51 nucleation on ssDNA

To quantify directly the rate and size of RAD51 nucleation on ssDNA, we used force-induced melting to denature a dsDNA molecule labelled with biotins at the 3’ and 5’ ends of the same strand using dual-trap optical tweezers in order to obtain a ssDNA tether [60]. Next, we employed a multi-channel microfluidic system [22, 59, 25, 51] and incubated the ssDNA molecule at a defined tension (in the 5 to 50 pN range) with fluorescently labelled (Alexa555) RAD51 [114] in the presence of ATP and Ca$^{2+}$. After incubation, the ssDNA molecule was repositioned in the imaging channel and inspected using fluorescence microscopy. The number of fluorescent RAD51 spots appearing after a given incubation period were counted in a low-coverage regime to insure single-nucleus resolution (less than 10 spots per ssDNA molecule at 100 mM KCl, 1 mM CaCl$_2$, 0.5 mM ATP) and divided by the length of the ssDNA molecule. The thus determined nucleation rates were found to depend strongly on RAD51 concentration (see Figure 34.A), ranging from $10^{-7}$ s$^{-1}$nt$^{-1}$ to $10^{-5}$ s$^{-1}$nt$^{-1}$ ([RAD51]=7.5 to 75 nM). The nucleation rates ($k_{nucl}$) did not depend linearly on the RAD51 concentration but could be well described by a power-law ($k_{nucl}=k_0[RAD51]^n$). The fitting parameter n, the Hill coefficient, has been interpreted before to represent the minimum number of monomers required to obtain a stable nucleus[v55, 72, 12]. Our fit yields $k_0=(1\pm1)\times10^{-8}$ s$^{-1}$ nt$^{-1}$ and $n=1.5\pm0.3$. According to the canonical interpretation, this would indicate that RAD51 nucleation requires a dimer [55, 169, 72, 12]. In order to get a deeper understanding of the nucleation process, we determined the nucleus size directly, by counting the number of RAD51 monomers in a nucleus using calibrated fluorescence intensity [25]. The calibration was performed using intensity drops due to photobleaching of the fluorescent labels. Under our experimental conditions, a single Alexa555 produced 460±160 counts per 500 ms (mean±standard deviation, see Supplementary Figure S38) [25]. To estimate the size of RAD51 nucleation, care was taken to discriminate between filament nucleation and growth by using low RAD51 concentrations (12.5 nM) and short incubation times (70 seconds). In this regime, we detected on average a single RAD51 nucleus per incubation. Experiments addressing filament growth
and cooperativity (discussed later) allowed us to conclude that under these conditions, 90% of the observed clusters resulted from an individual binding event from solution, hence they reflect dominantly RAD51 nucleation. We found that RAD51 nuclei are heterogeneously distributed in size (see Figure 34.B), ranging from individual RAD51 monomers and dimers (see supplementary figure S40) up to a significant population of hexamers, indicating that preassembled species of RAD51 can readily nucleate on ssDNA without the need of remodelling factors. We confirmed that RAD51 in solution is distributed in a collection of distinct oligomeric states using cross-linking gel experiments (see supplementary figure S41), these oligomeric states in solution have also been reported before [189, 44, 122].

The observation that the concentration dependence of the RAD51 nucleation rate follows a 1.5 power-law, while the size of RAD51 nuclei is broadly distributed might indicate that the stability of RAD51 nuclei on ssDNA is size-dependent. To test this, nucleation experiments were performed at low RAD51 concentration (12.5 nM RAD51, 0.5 mM ATP, 1 mM CaCl2, 100 mM KCl and 70 seconds incubation time) and the resulting RAD51 nuclei were observed for extended periods of time (up to 30 minutes, in the absence of RAD51 in solution) by taking fluorescence snapshots every 30 seconds in order to minimize photo-bleaching. In Figure 34.C an example illustrating the intrinsic instability of RAD51 nuclei is shown. One of the nuclei stayed bound on the ssDNA for the whole measurement, while the other one disappeared suddenly. It is highly unlikely that the sudden fluorescence intensity drop was caused by the simultaneous photobleaching of both fluorophores at once. Therefore, we attributed such intensity drop to the release into solution of the entire RAD51 nucleus. Next, we correlated the lifetime of the nucleus to the number of RAD51 monomers (see Figure 34.D), finding that RAD51-nucleus stability (cumulative probability of remaining bound to ssDNA for longer than 6 minutes) depends on size, in agreement with our initial hypothesis. The clear correlation between size and lifetime indicates that the initial phase of filament formation is characterized by the unstable binding of small RAD51 nuclei, needing additional monomers or mediator proteins to stabilize the nascent filament. Next, we tested whether the RAD51-nucleation rate is influenced by monovalent salt concentration. We found that the rate of RAD51 nucleation on ssDNA decreases approximately by 10-fold by increasing the salt concentration from 50 mM to 400 mM KCl. Finally, the RAD51-nucleation
Figure 34. RAD51 nucleation on ssDNA. [A] RAD51-concentration dependence of nucleation rate; black dots: experimental data (error bars represent the standard deviation); red line: power-law fit ($k_{\text{nuc}} = k_0 [\text{RAD51}]^n$) yielding an exponent $n$ of 1.5±0.3. [B] Histogram of nucleus sizes; red bars: experimental data (N=105); black, solid circles: model fit of distribution of number of fluorophores per nucleus; blue circles: distribution of expected number of RAD51 monomers per nucleus obtained from model fit (see Supplementary Information). [C] Fluorescence intensity time traces of individual RAD51 nuclei bound to ssDNA. Solid circles: RAD51 nucleus consisting of 2 fluorophores, detaching between 300 and 330 seconds. Hollow circles: RAD51 nucleus consisting of three fluorophores remaining ssDNA bound for more than 9 minutes. [D] Bar diagram showing how stable nucleus fraction depends on nucleus size. Error bars: standard deviation according to binomial distribution statistics. [E] Kinetic model of RAD51 filament formation in which assembly and disassembly are regulated by the free energy parameters $\epsilon_{\text{sol}}$ and $\epsilon_{\text{DNA}}$.

rate on ssDNA was determined in the presence of Mg$^{2+}$. The nature of the divalent cation (Ca$^{2+}$ or Mg$^{2+}$) has important consequences for filament stability [174] and ATP-hydrolysis [166]. We observed that the rate of RAD51 nucleus formation on ssDNA was in the same order of magni-
tude when either Ca$^{2+}$ or Mg$^{2+}$ is used ($1.9\times10^{-5}$ s$^{-1}$ nt$^{-1}$ in Ca$^{2+}$ and $0.7\times10^{-5}$ s$^{-1}$ nt$^{-1}$ in Mg$^{2+}$, [RAD51]=75 nM).

5.3.2 RAD51 filament growth

Next, we focussed on the growth phase of RAD51-filament formation on ss-DNA. We disentangled the nucleation from the growth phase by adopting a single-molecule fluorescence recovery after photobleaching method (sm-FRAP): after a first incubation step, nuclei were visualized in the imaging channel, their position recorded and afterwards completely photobleached. This was followed by subsequent incubation, detection and photobleaching cycles. Fluorescence images obtained from consecutive incubations were superimposed and growth events were scored when fluorescent patches colocalized (see Figure 35). The number of RAD51 monomers added per incubation was determined from the fluorescence intensity, revealing that also filament growth occurs by incorporation of various types of RAD51 multimers, including monomers (see supplementary figure S42). Also, the probability of RAD51 growth was determined, which was defined as the fraction of filaments exhibiting growth, divided by the incubation time. The measurements obtained should be taken as a lower estimate, since some of the nuclei might have disassembled from the DNA during subsequent incubations, but were still accounted for in the normalization. Following this procedure, we estimated the probability of RAD51 nuclei growth to be $(8\pm6)\times10^{-4}$ s$^{-1}$ filament$^{-1}$ (mean$\pm$SD, N=32) at 12.5 nM RAD51. The ratio between growth and nucleation rate, referred to as the cooperativity, could therefore be directly measured without the need of mathematical modelling. We found that the cooperativity value for RAD51 filament formation on ssDNA is about 2500 (considering that the nucleation rate at 12.5 nM is $3.3\times10^{-7}$ s$^{-1}$nt$^{-1}$).

5.3.3 RAD51-filament formation is highly substrate-specific

We also tested how the rate of RAD51-filament assembly is affected by the DNA substrate. ssDNA and dsDNA molecules were kept under a tension of 20 pN, incubated in 75 nM RAD51, moved to the observation channel and fluorescence images were taken. Figure 36 shows images of RAD51
Figure 35. Sm-FRAP allows detection of RAD51 growth on ssDNA. [A] Fluorescence image showing 3 individual fluorescent RAD51 nuclei on ssDNA. Subsequent continuous laser illumination resulted in complete photo-bleaching of the nuclei (not shown). [B] Fluorescence image of the same ssDNA-RAD51 complex after an additional incubation period. Fluorescent image shows the appearance of three distinct fluorescent patches. [C] Superposition of image [A] and [B] allows distinguishing new nucleation events from RAD51 growth. In the yellow circle we show that two of the fluorescent patches obtained from consecutive incubations co-localize exactly. [D] Line profile and Gaussian fitting of image A and B confirm the co-localization of the two patches with a resolution below 20 nm. This confirms the direct separate detection of RAD51 nucleation and growth on ssDNA.

filaments bound to ssDNA and dsDNA after incubations of 15 and 480 seconds, respectively.

Compared to the ssDNA, the dsDNA contained a far lower number of filaments, directly demonstrating that the nucleation rate of RAD51 for
Figure 36.: Selectivity of RAD51 binding. [A] ssDNA molecule after RAD51 incubation ([RAD51]=75 nM); [B] dsDNA molecule after RAD51 incubation ([RAD51]=75 nM). [C] Rate of RAD51 nucleation versus applied tension for ssDNA (solid circles) and dsDNA (hollow circles). In the inset the proposed free energy diagram of the RAD51 nucleation process. [D] Rate of RAD51 filament growth versus applied tension for ssDNA (solid circles) and dsDNA (hollow circles).

ssDNA is much higher than for dsDNA. This result was confirmed in different buffer conditions using a gapped DNA construct containing both an ssDNA and a dsDNA segment (see supplementary figure S43). To quan-
titatively describe the differential affinity, nucleation and growth rates (using the sm-FRAP method) were determined for both ssDNA and dsDNA at various applied tensions (see Figure 36.C and Figure 36.D).

From the graphs, two key aspects are immediately clear: (i) nucleation and growth are systematically faster on ssDNA than on dsDNA, within the force regime explored; (ii) nucleation and growth are strongly force-dependent on dsDNA, but are not on ssDNA. At zero tension nucleation is 5000-fold faster on ssDNA than on dsDNA; in addition, growth is also faster on ssDNA than on dsDNA (approximately 100 times). Values for the cooperativity could be obtained both for ssDNA (∼2500) and dsDNA (∼ 10⁶ at zero tension down to 6x10⁴ at 50 pN). We independently confirm that RAD51 binding is faster on ssDNA substrates using fluorescence anisotropy measurements (see supplementary figure S44). Taken together, these results show that RAD51 has a strong intrinsic preference for forming filaments on ssDNA over dsDNA in the absence of recombinase mediators, such as BRCA2.

The force dependence of nucleation and growth on dsDNA could be well fitted with an Arrhenius-law based model \( k(F) = k(0) \exp[-F \cdot \delta x/k_B T] \), where \( F \) is the force acting on the dsDNA, \( \delta x \) the distance to the transition state along the reaction coordinate and \( k_B T \) the thermal energy [174]. The fits yield \( \delta x_{nucl} = 0.45 \pm 0.05 \) nm and \( k_{nucl}(0) = (4 \pm 3) \times 10^{-9} \) s⁻¹ bp⁻¹ for nucleation (see Figure 36.C) and \( \delta x_{growth} = 0.27 \pm 0.03 \) nm and \( k_{growth}(0) = (3 \pm 1) \times 10^{-3} \) s⁻¹ filament⁻¹ for filament growth (see Figure 36.D).

5.3.4 Effect of BRC4 on RAD51 filament formation

In cells, homologous recombination is tightly regulated by so-called recombinase mediator proteins [73, 159, 142].

We studied the kinetics of RAD51-filament assembly on ssDNA molecules of various lengths and sequences in the presence of the BRC-repeat peptide BRC4, one of the RAD51-binding domains of BRCA2. RAD51-filament formation was measured at the single-molecule level using long random sequences of ssDNA (48’512 nt) in both sub-stoichiometric (1:0.5 (RAD51:BRC4)) and excess quantities of BRC4 (1:1 and 1:5). When pre-incubated with RAD51, BRC4 blocked RAD51 filament formation at all concentrations tested and in different buffers (see Figure 37).
Figure 37.: Effect of BRC4 on RAD51-filament formation. [A] Inhibitory effect of BRC4 on RAD51-nucleoprotein-filament assembly on ssDNA. From left to right: filament formation in absence and in the presence of BRC4 at indicated RAD51:BRC4 ratios. RAD51 concentration was 20 nM in all measurements. A clear inhibitory effect was observed in Mg\(^{2+}\), Mg\(^{2+}\)/Ca\(^{2+}\) (20 mM Tris pH 7.5, 10 mM MgO(Ac)\(_2\), 2 mM CaCl\(_2\), 2 mM ATP) and Ca\(^{2+}\) buffer. [B] Fluorescence polarization anisotropy measurements of the rate of RAD51 (100 nM) assembly on fluorescein-labelled 60-nucleotide long ssDNA (10 nM) confirmed that BRC4 repeats inhibits RAD51 binding to ssDNA. Each data point represents the average of three independent experiments.

Using sub-stoichiometric amounts of BRC4, we observed approximately a two-fold reduction in RAD51 binding to ssDNA, while inhibition was complete when BRC4 was present in excess. These observations were confirmed by ensemble fluorescence-anisotropy experiments on a short heteropolymeric ssDNA (60 nt), which showed a purely inhibitory effect of
BRC4 on RAD51 filament formation for all RAD51:BRC4 ratios tested (see Figure 37). Exposure of pre-formed RAD51 filaments to BRC4 did not result in any noticeable disassembly. Finally, using a fluorescently labeled BRC4 (BRC4-Alexa488), we found that stable association to RAD51-ssDNA filaments does not occur, differently to what has previously reported.

5.4 Discussion

Several kinetic schemes have been put forward to describe the mechanism of RAD51 filament formation, differing mainly in the proposed fundamental units of nucleation, growth and cooperativity number [109, 169, 6, 72].

Our approach shows that RAD51 nucleation does not involve a unique RAD51 species as previously anticipated, but is intrinsically heterogeneous. The observation that RAD51 is present in solution as oligomeric species that are competent for filament nucleation removes the requirement for a potential involvement of BRCA2 in topological remodelling of RAD51 into a form that is suitable for DNA binding.

We propose a novel kinetic scheme (see Figure 34.E) for RAD51 nucleation based on our findings that (i) in solution a range of RAD51 oligomeric species is present that are competent for binding ssDNA and (ii) the stability of RAD51 nuclei bound to ssDNA is size-dependent. These observations can be modelled introducing two free energy parameters with a direct physical interpretation: (i) $\epsilon_{sol}$, describing the RAD51 monomer-monomer interaction in solution and (ii) $\epsilon_{DNA}$, the contribution of each RAD51 monomer to the energy barrier of dissociation of the nucleus from the ssDNA. The probability of observing a RAD51 nucleus of $n$ monomers after an incubation time period $t$ then can be expressed as (for derivation, see Appendix A):

$$p(n, t) = \frac{k_{on}(n)}{k_{off}(n)} \left(1 - e^{-tk_{off}(n)}\right)$$

(2)

where $k_{off}(n)$ and $k_{off}(n)$ are the unbinding and binding rate of RAD51 nuclei with size $n$. $k_{on}(n)$ is proportional to the concentration of RAD51 oligomers with size $n$ in solution ($k_{on} \propto C(n)$). For an equilibrated solution, we have: $k_{on}(n) = k_0 e^{-n\epsilon_{sol}}$. To unbind, RAD51 has to overcome an
energy barrier linearly dependent on nucleus size, therefore: \( k_{\text{eff}}(n) = k_0 e^{-n\epsilon_{\text{DNA}}} \).

Maximum-likelihood fitting of this model to our data (See Supplementary Information) resulted in a nucleus-size distribution that matches our experimental observations (see Figure 34A). According to our fit, every RAD51 monomer contributes to the overall stability of the nucleus) with an amount equal to \( \epsilon_{\text{DNA}} = 1.8 \pm 0.5 \kappa B T \). The interaction energy between RAD51 monomers in solution depends on the concentration and is \( \epsilon_{\text{sol}} = 1.1 \pm 0.3 \kappa B T \) at [RAD51]=12.5 nM (see Supplementary Information). In the nucleation phase, therefore, RAD51 monomers and dimers nucleate onto ssDNA very frequently because more abundant, but relatively inefficiently due to their small interaction energy with ssDNA. Larger RAD51 multimers (3-6 protomers), although less abundant, nucleate readily, forming stable seeds for filament formation.

Following our formalism, we investigated the expected effect of varying protein concentration on the observed nucleation rate and apparent Hill coefficient. We predict that the apparent Hill coefficient in the concentration range used in our experiments (7.5 to 75 nM) should vary between 1.2 and 2.1, with an average value of 1.6, which is in agreement with our experimental observation. Previous studies on recombinase proteins [55, 72, 12] interpreted the Hill coefficient as the minimal nucleation unit. Here we show that this simplistic interpretation may be in general not valid. Our thermodynamic model provides a more general framework for the interpretation of nucleation on a template by proteins capable of self-assembling in solution.

Next, we investigated the mechanism leading to selective filament formation. Previous studies comparing RAD51 affinity for ssDNA or dsDNA present some incongruities. In gel-mobility shift experiments RAD51 displayed higher affinity for dsDNA [105], whilst in another study ssDNA was the preferred substrate [149]. Magnetic-tweezers experiments have resulted in contradicting models, in which RAD51 does not display substrate specificity [169] or binds faster to ssDNA [109]. Here, we directly quantified RAD51 nucleation and growth rates on single ssDNA and dsDNA molecules. We detected two striking differences in the DNA-binding kinetics of RAD51: (i) nucleation is highly substrate selective, favouring ssDNA over dsDNA; (ii) nucleation rate is independent of tension on ssDNA, while it increases exponentially with tension on dsDNA. In the case
of dsDNA, RAD51 filament formation requires a substantial amount of lengthening and therefore is energetically costly. The assisting load tilts the energy landscape of RAD51 nucleation and growth on dsDNA, decreasing template selectivity in the high force regimes. Quantitatively, the application of force on dsDNA corresponds to lowering the energy barrier for nucleation and growth up to $\sim 7 k_B T$ and $\sim 4 k_B T$ respectively, at the maximum force dsDNA can sustain (65 pN). These values, in turn, represent a lower limit for the actual energy barriers for binding dsDNA in a relaxed conformation.

The strong intrinsic preference of RAD51 for ssDNA uncovered by our findings represents a simple yet powerful means of targeting the recombinase towards its physiological nucleic-acid substrate and limiting non-productive binding to dsDNA. Recent findings suggest a novel role for RAD51, independent of homologous recombination, in coating regions of ssDNA that form at replication forks in normal or stressed conditions [68, 146]. Our observation that RAD51 possesses an intrinsic affinity for ssDNA provides a physical basis for this novel biological role of RAD51. Recombination mediators such as BRCA2 and RAD51 paralogues have been often implicated to overcome the alleged lack of binding preference by directing RAD51 specifically to ssDNA [27, 151, 83, 165, 28]. Our observation that RAD51 has a strong intrinsic preference for filament formation on ssDNA indicates that BRCA2 might be less important in regulating RAD51 substrate specificity. Indeed, using both single-molecule and ensemble kinetic measurements we found that BRC4 domain of BRCA2 acts as an inhibitor of RAD51-filament formation in a concentration-dependent way. Hence, the role of BRCA2 might be restricted in promoting filament formation by displacing RPA from ssDNA [83, 98] and/or regulating ATP hydrolysis and stabilizing RAD51 filaments [28]. These putative roles of BRCA2 can now be tested, in our experimental setup, by observing directly the impact of recombinase mediators such as BRCA2 on RAD51-filament formation. Moreover, another class of mediators including Srs2, RTEL1 and RecQ5 helicases, can be investigated at the single-molecule level, in order to obtain a full picture of the regulation of RAD51-filament formation.

Our sm-FRAP approach allowed us to quantify RAD51-filament growth on both ssDNA and dsDNA. Past attempts to discriminate between nucleation and growth was hindered by experimental limitations, since the accessible experimental readout, the coverage fraction in magnetic-tweezers
Cooperativity of RAD51 filament formation and comparison with RecA

assays, represents an average behaviour of the RAD51 binding reaction. Cooperativity is a fundamental property of the RAD51-filament formation process, determining the final size and structure of the filaments. Separate detection of filament growth permitted us to measure for the first time RAD51 cooperativity. We found that the rate of RAD51 growth is \(\sim 3\) (ss-DNA) and \(\sim 5\) (dsDNA) orders of magnitude larger than the nucleation. These results confirm that nucleation is a rate-limiting step of filament formation and, as such, could represent an important regulatory step during HR. A recent single-molecule study indicated that the RecA filament formation on SSB-coated ssDNA [12] is highly cooperative (cooperativity was determined to be on the order of \(10^7\), considering a growth rate of 30 monomer min\(^{-1}\) and a nucleation rate of \(2 \times 10^{-6}\) nt\(^{-1}\) min\(^{-1}\) at 200 nM of RecA). Whether this high value for the cooperativity stems from the intrinsic binding kinetics of RecA on ssDNA or from the inhibitory effect of SSB on the nucleation step of RecA is not clear. The striking difference between the reported cooperativity values for RecA and RAD51 (approximately 4 orders of magnitude) might indicate that RAD51-filament growth is rate-limiting during filament formation, thus requiring the intervention of recombinase mediators at this critical step of recombinational repair. To conclude, the experimental approaches developed in this study elucidated the mechanism and regulation of the crucial homologous recombination process and, in addition, provide a demonstration that single-molecule methods can now visualize directly and precisely multi-protein interactions at single-monomer resolution on individual ssDNA molecules.
We here set up a minimal model coupling RAD51 filament association/dissociation to/from, a single DNA molecule.

### A.1 Filament-DNA Binding Dynamics in the Low Coverage Regime

The total coverage of the DNA molecule remains very low throughout the experiments, and we can ignore the possibility of existing filament blocking the binding of new filaments. If we let \( p(n,t) \) to be the probability that a binding position contains the left end of a filament of length \( n \) at time \( t \), the corresponding Master equation for each binding site reads

\[
\frac{\partial p(n,t)}{\partial t} = -k_{\text{off}}(n)p(n,t) + k_{\text{on}}(n) \tag{3}
\]

Here, \( k_{\text{off}}(n) \) is the rate at which a filament of length \( n \) dissociates from the DNA strand, and \( k_{\text{on}}(n) \) is the rate at which filaments of length \( n \) attach to a binding position from solution. Solving the Master equation we have:

\[
p(n,t) = \frac{k_{\text{on}}(n)}{k_{\text{off}}(n)} \left( 1 - e^{-tk_{\text{off}}(n)} \right) \tag{4}
\]

For times and filament lengths such that \( tk_{\text{off}}(n) \ll 1 \), few filaments that have attached to the DNA have also had time to fall off, and the occupation probability grows linearly with time \( p(n,t) = tk_{\text{on}}(n) \). If instead times and filament lengths are such that \( tk_{\text{off}}(n) \gg 1 \), the filaments on the DNA have had time to equilibrate with the buffer, and we have the time-independent distribution \( p(n,t) = k_{\text{on}}(n) / k_{\text{off}}(n) \).
A.2 EQUILIBRIUM DISTRIBUTION OF FILAMENTS IN SOLUTION

We now assume that RAD51 proteins can form linear filaments in solution, with a constant interaction energy gained for every protein-protein association. These filaments are continuously formed and broken up by thermal forces in equilibrium. For the concentration of filaments of length \( n+m \), \([\text{RAD51}]_{n+m} \), to be equilibrated with the concentrations \([\text{RAD51}]_{n} \) and \([\text{RAD51}]_{m} \), the law of mass action dictates that \([\text{RAD51}]_{n+m} = \alpha [\text{RAD51}]_{n} [\text{RAD51}]_{m} \). Here \( \alpha \) is the association constant between filament ends, and the resulting exponential distribution of filaments is \([\text{RAD51}]_{n} = (\alpha [\text{RAD51}])^{n} / \alpha \). If we use that the total concentration of protein is \([\text{RAD51}] = \sum n [\text{RAD51}]_{n} \), we can rewrite the filament concentrations as:

\[
[\text{RAD51}]_{n} = 4 [\text{RAD51}] \sinh^{2} \left( -n \epsilon_{\text{sol}} / k_{B}T \right) \tag{5}
\]

\[
\epsilon_{\text{sol}} = k_{B}T \ln \left( \frac{2 \alpha [\text{RAD51}]}{1 + 2 \alpha - \sqrt{1 + 4 \alpha [\text{RAD51}]}} \right) \tag{6}
\]

A.3 ASSOCIATION AND DISSOCIATION RATES BETWEEN DNA AND RAD51 FILAMENTS

We take the on-rate of filaments of length \( n \) to be directly proportional to the filament concentration in the buffer, and write \( k_{on} (n) = k_{on} e^{-n \epsilon_{\text{sol}} / k_{B}T} \). We further assume each RAD51 molecule in a filament contributes a fixed amount \( \epsilon_{\text{DNA}} \) to the dissociation barrier from DNA, giving \( k_{off} (n) = k_{off} e^{-n \epsilon_{\text{DNA}} / k_{B}T} \). This defines the model in terms of two rate constants \( (k_{on} \) and \( k_{off} \)) and two energy scales \( (\epsilon_{\text{sol}} \) and \( \epsilon_{\text{DNA}} \)) with direct physical interpretations.

For any given time \( t \), the maximum filament length that has had time to equilibrate with the buffer is \( n^{*}(t) \approx (k_{B}T / \epsilon_{\text{DNA}}) \ln \left( tk_{off} \right) \). Interestingly, if \( \epsilon_{\text{DNA}} > \epsilon_{\text{sol}} \) the nuclei length distribution on the DNA is peaked somewhere around \( n^{*}(t) \), resulting in a typical nucleus size that shifts logarithmically with the time since incubation.
A.4 NUCLEATION RATE AND THE APPARENT HILL COEFFICIENT

The apparent nucleation rate of filaments measured in the experiment only accounts for the filaments that remain bound at the end of the incubation. In terms of our model, this is given by

\[ k_{on}^{app} = \sum_{n=1}^{\infty} p(n,t) / t_{inc} \]  

(7)

The actual total on-rate is inferred from our model parameters as:

\[ k_{on}^{true} = \sum_{n=1}^{\infty} k_{on} e^{-\frac{n \epsilon_{sol}}{k_B T}} = \frac{k_{on}}{e^{\frac{\epsilon_{sol}}{k_B T}} - 1} \]  

(8)

If we estimate the model parameters for any one known concentration, we can predict the filament distribution at any other concentration. Of special interest is the apparent Hill coefficient, since it has previously been used as a measure of the (assumed) degree of cooperative binding. We are working in the low-coverage regime—well before any saturation effects become important—and therefore define the concentration dependent apparent Hill coefficients as the derivative:

\[ n_{Hill}^{app} ([RAD51]) = \frac{d \ln k_{on}^{app} ([RAD51])}{d [RAD51]} \]  

(9)

This is the local slope in a log-log-diagram of the apparent binding rate (or nucleation rate) vs. concentration.

A.5 MAXIMUM LIKELIHOOD FITTING OF FLUOROPHORE DISTRIBUTION

We here fit our proposed model to the data using Maximum likelihood methods, yielding both estimators for the parameters themselves and the errors.

We are faced with the additional problem that the number of fluorophores \( f \), is not unique on each monomer, but distributed according to, say, \( q(f) \). The
observed number of fluorophores \( f \) in a fiber of \( n \) monomers is then distributed according to

\[
q_m (f) = \sum_{f_1, \ldots, f_m} q (f_1) q (f_2) \ldots q (f_m) \delta_{f_1 + \cdots + f_m = f} = (q^*)^m (f)
\]  

(10)

Here \( \delta \) is the Kronecker delta function. Letting \( \theta = \{ k_{on}, k_{off}, \epsilon_{sol}, \epsilon_{DNA} \} \), the probability of finding \( f \) fluorophores attached to a filament starting at any particular position is given by

\[
P (f, t \mid \theta) = \sum_{m=1}^{\infty} q_m (f) \ p (m, t \mid \theta)
\]  

(11)

In our system the only non-zero fluorophore probabilities are \( q (1) = 3/4 \) and \( q (2) = 1/4 \).

### A.6 FITTING THE FLUOROPHORE DISTRIBUTION

To fit the model to our data, we utilized a maximum-likelihood scheme. The fit is generated by maximizing the likelihood of the observed fluorescence of the DNA molecules, given our model

\[
L \left( \theta \mid \left\{ N_f \right\}_{f=0}^{\infty} \right) = \exp \left( - \sum_{f=0}^{\infty} N_f \ln P (f, t_{inc} \mid \theta) \right)
\]  

(12)

Here \( t_{inc} = 70s \) is the incubation time, and \( N_f \) is the number of filaments observed with \( f \) fluorophores. The total number of binding sites exposed during all experiments are \( N_0 = N_{DNA} N_{Exp} \), where \( N_{DNA} \) is the number of binding sites per tether and \( N_{Exp} \) is the number of experiments. To estimate the model parameters we numerically minimize \(- \ln N \left( \theta \mid \left\{ N_f \right\}_{f=0}^{\infty} \right)\) with respect to them. To estimate the error in the fit parameters, we perform the fit on 1000 bootstrapped data sets. We report the best fit of the original data, together with the standard deviation in the fit parameters as calculated over the bootstrapped data set:

\[
k_{on} = 10^{-4.7 \pm 0.8} n t^{-1} s^{-1}
\]  

(13)
\[ k_{\text{off}} = 10^{1.5\pm0.8} \text{s}^{-1} \quad (14) \]

\[ \epsilon_{\text{sol}} = (1.1 \pm 0.3) k_B T \quad (15) \]

\[ \epsilon_{\text{DNA}} = (1.8 \pm 0.5) k_B T \quad (16) \]

The fit of our model to the measured data set is presented in Figure 2B, and indicates that the average RAD51 oligomer size in solution was \( n_{\text{sol}} = \frac{1}{(1 - e^{-\epsilon_{\text{sol}}/k_B T})} \) monomers—in good agreement with chemical cross-linking experiments performed in a similar concentration range (Supplementary Figure 3). From the fit parameters we calculated the association constant between RAD51 filaments by using our expression for \( \epsilon_{\text{sol}} \). The experiments were performed at [RAD51]=12.5 nM, giving \( \alpha = 0.060 \pm 0.036 / \text{nM} \). With this we have estimates for all the microscopic parameters.

Apparent vs. true nucleation rate The apparent total nucleation rate in the experiments is the number of observed nuclei divided by the total number of bases incubated across experiments, and then further divided by the incubation time, giving \( k_{\text{tot}*} = 3.7 \times 10^{-7} \text{nt}^{-1} \text{s}^{-1} \). From above we know that the true total on-rate is given by:

\[ k_{\text{on}}^{\text{tot}} = \frac{k_{\text{on}}}{e^{\epsilon_{\text{sol}}} - 1} \approx 1.1 \times 10^{-5} \text{nt}^{-1} \text{s}^{-1} \quad (17) \]

\section*{A.7 Apparent Hill Coefficient}

Since we have estimates for all the microscopic parameters, we can make predictions regarding the expected dynamics at any concentration. The apparent Hill-coefficient calculated in the concentration range around [RAD51]=12.5 nM is 2.0. The average Hill coefficient in the concentration range used in our experiment is 1.6.
Figure S38: Force-induced melting of dsDNA was employed to produce ssDNA templates. Force-stretching curve of a 38,412 bp dsDNA (Black line). The molecule was pulled beyond the overstretched plateau (F>85 pN). The force-stretching curve (red line) confirmed the generation of a single ssDNA molecule tethered between the beads.

Figure S39: DNA-protein complexes can absorb onto polystyrene beads. [A] Force distance curve of a naked DNA molecule (black) and after incubation with XRCC4-XLF (red). In the red curve a force rise and drop is detected, resembling the rupture of an intra-molecular bridge. [B] Fluorescent visualization during bead separation reveals the unspecific absorption of DNA-protein complexes onto the polystyrene bead.
Figure S40.: RAD51 nucleus on ssDNA. Example of stable RAD51 nucleus on ssDNA composed of 2 fluorophores. [A] Fluorescent snapshot of an individual ssDNA molecule kept at 40 pN after a 70 seconds incubation in the enzyme channel containing 12.5 nM of RAD51. [B] Kymograph. [C] Fluorescence intensity shows staircase pattern. Two bleaching events can be easily recognized. In this experiment we set the camera exposure to 1000 ms and we increased the laser excitation for obtaining high S/N. The presence of two steps in the photo-bleaching trace indicates that the RAD51 nucleus was composed of two monomers, each labelled with one Alexa555 dye. Alternatively there is the possibility that we observed the presence of a single RAD51 monomers labelled with two Alexa555 dyes.
Figure S41.: Alexa-RAD51 cross-linking in solution with BS2. The cross-linked complexes were resolved by polyacrylamide gel electrophoresis. Detection of Alexa-RAD51 by anti-RAD51 Western blot [A] and typhoon scan of the Alexa fluorescence signal [B] show that with increasing cross-linker concentration more defined multimeric species of RAD51 are cross-linked. The example shows cross-linking at 75 nM RAD51. [C] The percentage of different forms of RAD51 determined from densitometry scanning of protein cross-linked at 7.5 nM, 25 nM and 75 nM RAD51 and run on SDS-PAGE.
Figure S42: RAD51 growth on dsDNA  
[A] Averaged fluorescence image of λ-DNA molecule (48502 bp, length 16.4 µm) held at 40 pN after incubation with 75 nM RAD51 for 70 seconds. After visualization, continuous exposure resulted in complete photo-bleaching of all the fluorescent spots (data not shown).  
[B] Averaged fluorescent image of the same dsDNA molecule, re-incubated with RAD51 for a second cycle of 70 seconds.  
[C] Overlaying figure A and B allows distinguishing between growth and new nucleation events. In the yellow circles are example of patch co-localization.  
[D] A line profile over a region along the dsDNA is plotted. We found that in two cases the co-localization of the patches was with sub-20 nanometers precision.  
[E] A photo-bleaching trace of the fluorescent spot at position 7472nm detected in the second incubation cycle (growth event) shows one individual step in agreement with our single-molecule fluorescence calibration and demonstrates in a direct way that extension of RAD51 can take place in form of individual RAD51 monomers.
Figure S43.: Hybrid ss-ds DNA constructs reveal structural specificity of human RAD51. RAD51 filament formation rate is affected by the DNA structure. To generate the hybrid templates a λ-DNA molecule was held at 22 µm at 65 pN for 60 seconds. When single nicks are present on the unlabelled strand of the DNA, a hybrid ss-dsDNA structure can be generated. Force-distance curves were taken to verify and quantify the extent of the ssDNA and dsDNA region. The incubation time in both cases was 60 seconds and the RAD51 concentration was 75 nM. [A] Experiment in 20 mM Tris pH 7.6, 100 mM KCl, 1 mM CaCl2, 0.5 mM ATP, 10 mM DTT. [B] Experiment in 20 mM Tris pH 7.6, 10 mM MgOAc, 2 mM CaCl2, 1 mM ATP, 10 mM DTT.

Figure S44.: Fluorescence polarization anisotropy measurement on RAD51 selectivity. Fluorescence anisotropy measurements of RAD51 binding to fluorescein labelled ssDNA (black filled circles) and to dsDNA (white filled circles). The ssDNA was a 60-nucleotide long DNA sequence. The dsDNA was constructed by annealing the fluorescein – labelled ssDNA with its complementary strand. A concentration series of RAD51 in binding buffer was titrated into wells using the PHERAstar FS built-in injectors. Binding was initiated by injecting fluorescein labelled DNA (10 nM final) and measurements were taken after 10 minutes with 200 flashes. The data points represent the average of three independent experiments (error bars = 1SD).