Molecular flexibility of DNA as a major determinant of RAD51 recruitment

Federico Paoletti [1], Afaf El-Sagheer [2,3], Jun Allard [4],

Tom Brown [2], Omer Dushek [1,†], Fumiko Esashi [1,†]

1. Sir William Dunn School of Pathology, University of Oxford, OX1 3RE, UK.

2. Department of Chemistry, University of Oxford, OX1 3TA, UK.

3. Department of Science and Mathematics, Suez University, Suez 43721, Egypt.

4. Department of Mathematics, University of California, Irvine, CA 92697-4575, US.

[†] Equal contribution and corresponding authors:

Omer Dushek: omer.dushek@path.ox.ac.uk

Fumiko Esashi: fumiko.esashi@path.ox.ac.uk

Keywords: Homologous recombination, RAD51, ssDNA, dsDNA, Mathematical model, SPR

Abstract

The timely activation of homologous recombination is essential for the maintenance of genome stability, in which the RAD51 recombinase plays a central role. Biochemically, human RAD51 binds to both single-stranded DNA (ssDNA) and double-stranded DNA (dsDNA) with similar affinities, raising a key conceptual question: how does it discriminate between them? In this study, we tackled this problem by systematically assessing RAD51 binding kinetics on ssDNA and dsDNA differing in length and flexibility using surface plasmon resonance. By fitting detailed polymerisation models informed by our experimental datasets, we show that RAD51 is a mechano-sensor that exhibits a larger polymerisation rate constant on flexible ssDNA compared to rigid ssDNA or dsDNA. This model presents a new general framework suggesting that the flexibility of DNA, which may increase locally as a result of DNA damage, plays an important role in recruiting repair factors that multimerise at sites of DNA damage.

Introduction

DNA double-strand breaks (DSBs) are cytotoxic lesions that can lead to chromosomal breaks, genomic instability and tumourigenesis in mammalian cells (Tubbs and Nussenzweig, 2017). Homologous recombination (HR) can offer an error-free DNA repair mechanism to restore genetic information at DSB sites, and in this way, contribute to genome stability. During HR-mediated repair, single-stranded DNA (ssDNA) overhangs are generated and rapidly coated with the ssDNA-binding replication protein A (RPA) (Chen and Wold, 2014). ssDNA-bound RPA is then exchanged for RAD51, the central ATP-dependent recombinase that catalyses HR-mediated repair. RAD51 polymerises on ssDNA to form a ssDNA-nucleoprotein filament and guide homologous strand invasion and DSB repair (Baumann et al., 1996).

The central mechanism of HR is evolutionarily highly conserved, and the bacterial RAD51 ortholog RecA has clear preference to polymerise on ssDNA over dsDNA (Benson et al., 1994). However, human RAD51 shows weaker binding affinity to ssDNA compared to RecA, and the mechanism by which RAD51 polymerises on ssDNA in preference to dsDNA remains enigmatic. Earlier studies using electrophoretic mobility shift assay (EMSA) have suggested that RAD51 binds both ssDNA and dsDNA with similar affinities (Benson et al., 1994), albeit its preferential ssDNA binding was visible in the presence of ammonium sulphate (Shim et al., 2006). These endpoint assays, however, do not provide information as to whether binding kinetics may contribute to a potential RAD51-dependent ssDNA / dsDNA discrimination mechanism. Indeed, more recent kinetic studies have revealed that RAD51 polymerisation consists of two phases: a rate-limiting nucleation phase and a consecutive growth phase (Hilario et al., 2009; Miné et al., 2007; Van der Heijden et al., 2007). A minimal polymer nucleus, with a length of either two-to-three or four-to-five RAD51 protomers, is proposed to elicit the growth phase of RAD51 polymerisation. Nonetheless, it remains unknown whether there are any mechanistic differences between RAD51 polymerisation on ssDNA and dsDNA. Intriguingly, RAD51 is shown to display faster association kinetics on

ssDNA (Candelli et al., 2014), whilst it also exhibits slower dissociation kinetics on dsDNA, indicating that the RAD51-dsDNA complex is stable once formed (Miné et al., 2007).

A key difference between ssDNA and dsDNA is their molecular flexibility: ssDNA is known to be more flexible compared to dsDNA. The flexibility of a DNA molecule can be characterised by its persistence length (L_p), a mechanical parameter quantifying polymer rigidity: the higher the persistence length, the more rigid the polymer. In the presence of monovalent or divalent salt, dsDNA displays an L_p of ~ 30 - 55 nm (Baumann et al., 1997; Brunet et al., 2015), while ssDNA is much more flexible, with an L_p of 1.5 - 3 nm (Chi et al., 2013; Kang et al., 2014; Murphy et al., 2004). These observations imply that ssDNA can explore a much larger configurational space compared to dsDNA. It follows that the formation of a structured (less flexible) RAD51 polymer on ssDNA will incur a large entropic energy penalty compared to the formation of a RAD51 polymer on dsDNA. Despite the clear thermodynamic implications of RAD51 polymerisation on DNA, the impact of DNA flexibility on RAD51 nucleoprotein filament formation has been largely overlooked.

In this study, we describe how RAD51 polymerises on DNA. Using a combination of surface plasmon resonance (SPR) and small-angle x-ray scattering (SAXS), we have assessed the RAD51 binding kinetics on DNA using ssDNA and dsDNA oligonucleotides differing in length and flexibility. Analyses of the SPR data using biochemical mathematical models revealed that RAD51 polymerisation required a minimal nucleus of four and two molecules on ssDNA and dsDNA, respectively. Interestingly, our analyses further uncovered that RAD51 is a mechano-sensor that polymerises faster on flexible DNA. We hypothesised that the entropic penalty to polymerisation on flexible DNA is offset by a large enthalpic contribution and show this to be the case by perturbing the RAD51 protomer-protomer interaction. We propose the DUET model (**D**na molec**U**lar fl**E**xibili**T**y), which explains the ability of RAD51 to preferentially and stably polymerise on highly flexible ssDNA compared to dsDNA, enabled by the enthalpic energy of RAD51 polymerisation.

Results

To evaluate how human RAD51 discriminates between ssDNA and dsDNA, we assessed the binding kinetics of untagged recombinant human RAD51 to a 50-mer mixed base ssDNA molecule (dN-50) and a 50-mer mixed base paired dsDNA molecule (dN-50p) by SPR. These analyses performed in the presence of ATP and Ca^{2+}, which blocks RAD51 ATP hydrolysis (Bugreev and Mazin, 2004), showed 1) faster RAD51 association to ssDNA compared to dsDNA, and 2) similar RAD51 lifetimes on both ssDNA and dsDNA (**Fig. S1A, B**). These observations validated that RAD51 distinguishes ssDNA from dsDNA through faster polymerisation on ssDNA (Candelli et al., 2014)**.**

For an advanced understanding of the mechanism by which RAD51 polymerises differentially on ssDNA and dsDNA, we set out a streamlined study comprises of the following processes; 1) perform SPR measurements to identify the kinetics of RAD51 binding to ssDNA and dsDNA, 2) develop a linear, ordinary differential equation (ODE) model, which infers RAD51 polymer formation on DNA, informed by respective SPR dataset, and 3) use Approximate Bayesian Computation coupled to Sequential Monte Carlo (ABC-SMC) (Toni et al., 2009) to fit the ODE model to all related SPR datasets. This strategy enables us to summarise our experimental SPR dataset with a mechanistic ODE model describing how RAD51 associates with naked DNA, forms a stable nucleus, and elongates to form polymers on DNA.

RAD51 nucleates poorly on ssDNA, but elongates efficiently to form long polymers.

We first aimed to define the initial steps of RAD51 association with ssDNA. To this end, we generated a series of mixed base short ssDNA oligos with different lengths, each of which can be bound by a restricted number of RAD51 molecules. As a single RAD51 molecule engages with three nucleotides of DNA (Short et al., 2016), a ssDNA consists of 5 nucleotides (dN-5), 8 nucleotides (dN-8), 11 nucleotides (dN-11), 14 nucleotides (dN-14) and 17 nucleotides (dN-17) can accommodate up to one, two, three, four and five RAD51 molecules, respectively (**Fig. 1A**). Our

systematic SPR measurements of RAD51 binding kinetics revealed no RAD51 binding to the dN-5, dN-8 and dN-11, slow association and moderate dissociation with dN-14, and slightly faster association and slow dissociation with dN-17 (**Fig. 1B**). These observations suggested that three or less RAD51 molecules are unable to generate stable nucleus on ssDNA, four RAD51 molecules form a quasi-stable nucleus, and five or more RAD51 molecules are able to form a highly stable nucleus.

We then moved on to develop ODE models describing RAD51 polymerisation on ssDNA, informed by these SPR datasets. The distinct RAD51 nucleation states on ssDNA led us to hypothesise that RAD51 polymerisation is achieved by three sequential DNA-binding steps, namely adsorption, elongation and stabilisation (**Fig. 1C**). The adsorption step is characterised by the forward and reverse adsorption rate constants ($k_{ad.f}$ and $k_{ad.r}$) and describes the formation of an unstable RAD51 nucleus on ssDNA. The elongation step is characterised by the forward and reverse elongation rate constants ($k_{el.f}$ and $k_{el.r}$), and describes the formation of a quasi-stable nucleus. Finally, the stabilisation step is characterised by the same forward elongation rate constant ($k_{el.f}$) but a different stabilisation reverse rate constant (k_s). This step describes stable nucleus formation, and also nucleus elongation up to a maximum length of 16, which corresponds to the predicted maximum RAD51 molecules binding to a 50-mer ssDNA molecule. Additionally, we assumed that 1) RAD51 polymerisation mainly occurs *via* the addition of RAD51 dimers, and to a lesser extent *via* the addition of RAD51 monomers (Subramanyam et al., 2016)**,** and that 2) RAD51 dissociation only occurs *via* the removal of monomers.

Having established the ODE model, we used ABC-SMC to simultaneously fit this model to all the RAD51 binding kinetics datasets and obtained values for the five model parameters (**Fig. 1B, C, Fig. S2A, B)**. This model describes the experimental data accurately compared to two simpler models which respectively describe RAD51 polymerisation only *via* monomer addition (**Fig. S2C**), or without taking account of the reverse elongation rate constant (i.e., $k_{el.r} = k_s$) (**Fig. S2D),** as summarised in **Fig. S2E**. Importantly, our model identified a high forward elongation rate constant ($k_{el.f}$ = 53.7 ± 10.66 /μM/s) as the factor driving the rapid RAD51 polymerisation on

ssDNA. Indeed, although our experimental data measured an average of six RAD51 molecules per dN-50 at the end of the SPR injection, our model predicts that most of the RAD51 forms long polymers of 16, 15 and 14 molecules on a fraction of DNA molecules during each injection (**Fig. S2F**). Similarly, the model predicts that RAD51 forms mainly pentamers and tetramers on the dN-17 (**Fig. S2G**), and only tetramers on the dN-14 (**Fig. S2H**). These results support the notion that RAD51 polymer formation on ssDNA requires the nucleation of at least four RAD51 protomers, but is facilitated by a high elongation rate constant.

RAD51 nucleates efficiently on dsDNA, but elongates poorly to form long polymers.

We next evaluated the initial steps of RAD51 association with dsDNA. We generated an analogous series of mixed base dsDNA consists of 5 base pairs (dN-5p), 8 base pairs (dN-8p) and 11 base pairs (dN-11p), which can accommodate up to one, two or three RAD51 molecules, respectively (**Fig. 2A**). Our SPR measurement detected no RAD51 binding to the dsDNA dN-5p, as was the case for the ssDNA dN-5. To our surprise, however, we detected slow association and moderate dissociation of RAD51 with the dN-8p, and even faster association and slower dissociation with the dN-11p (**Fig. 2B**). These observations suggested that one RAD51 molecule is unable to associates stably with ssDNA, but two RAD51 molecules form a quasi-stable nucleus, and three or more RAD51 molecules are able to form a highly stable nucleus.

An ODE model describing RAD51 polymerisation on dsDNA was then developed using the same assumptions applied to the RAD51-ssDNA ODE model, with one exception: the model is simplified to include two, instead of three, sequential DNA-binding steps, namely adsorption and stabilisation (i.e. with $k_{el.f} = k_s$) (**Fig. 2C**). Our ABC-SMC simulation was able to provide defined values for the four parameters of the model (**Fig. 2C, Fig. S3A, B**), validating that this model explains the SPR data well. Conversely, the assumption of RAD51 polymerisation only *via* monomer addition (**Fig. S3C**) or the addition of a reverse elongation step ($k_{el.r} \neq k_s$) (**Fig. S3D**), do not improve the model (**Fig. S3E**). Importantly, our model highlighted a considerably low forward

elongation rate constant ($k_{el.f}$ = 0.953 ± 0.48 /μM/s) compared to ssDNA, describing the unproductive RAD51 polymerisation on dsDNA. This is reflected by our model prediction that most RAD51 does not form long polymers on the dN-50p (**Fig. S3F**). Moreover, our model predicted that RAD51 forms dimers with a small amount of trimers on the dN-11p (**Fig. S3G**), and both monomers and dimers on the dN-8p (**Fig. S3H**). Collectively, these analyses revealed that a RAD51 dimer is sufficient to form a stable nucleus on dsDNA, but its polymer formation is restricted by a lower elongation rate constant.

RAD51 polymerises faster on flexible DNA

Overall, our analyses suggest that RAD51 nucleates more efficiently on dsDNA but elongates significantly faster on ssDNA (56-fold higher ssDNA $k_{el.f}$ compared to dsDNA $k_{el.f}$). We speculate that the difference in RAD51 elongation speed is due to the higher flexibility of ssDNA compared to dsDNA. Explicitly, we propose a Bend-To-Capture (BTC) mechanism to explain how DNA flexibility can impact polymerisation kinetics (**Fig. 3A**). A free RAD51 molecule in solution would need to generate two sequential, non-covalent interactions to be incorporated into the growing polymer: the interaction with the exposed interface of existing RAD51 polymer and with the exposed scaffold DNA. To avoid steric clashes, we reasoned that each of these interactions requires naked DNA immediately next to the existing RAD51 polymer in a configuration that can bend away from the preferred direction of the polymer. In this way, flexible DNA is expected to explore more conformations compatible with the further addition of RAD51 per unit time.

To test this notion, we designed an experiment to measure the kinetics of RAD51 binding to DNA oligos of varying flexibility. It has been shown that poly-dT ssDNA, which is widely used for RAD51 binding assays, is highly flexible, while poly-dA ssDNA is highly rigid due to base stacking interactions (Sim et al., 2012)**.** Consistently, our SAXS-derived persistence length measurements showed ssDNA dT-50 is the most flexible oligo, followed by dN-50, dA-50 and dsDNA dN-50p (**Fig. S4 and Table 1**). By measuring RAD51 binding kinetics to these DNA oligos (**Fig. 3B)**, we found that RAD51 indeed displayed faster association to the dT-50 compared to the dN-50, and the

model fit suggests this is due to a higher elongation rate ($k_{el.f}$) (**Fig. 3C, D**). Furthermore, RAD51 displayed very slow association to the dA-50, comparable to that of dsDNA dN-50p, which is explained by a lower elongation rate (**Fig. 3C, D**). Similarly, individually fitting the adsorption forward rate ($k_{ad.f}$) or adsorption reverse rate ($k_{ad.r}$) can explain the ssDNA dT-50 and dA-50 data (**Fig. S5A, B**). Conversely, individually fitting the elongation reverse rate constant ($k_{el.r}$) or the stabilisation rate constant (k_s) was unable to explain the dT-50 and dA-50 data (**Fig. S5C, D**). Meantime, we observed that the lifetime of the RAD51 polymers was independent of the flexibility of the underlying DNA (**Fig. 3E**). Taken together, these observations suggest that although RAD51 is able to form a stable nucleus with fewer molecules on dsDNA compared to ssDNA (two versus four), it elongates and/or absorbs more efficiently on ssDNA because of its higher flexibility, consistent with our proposed BTC mechanism.

RAD51 protomer-protomer interaction compensates for the entropic penalty of flexible DNA

The BTC mechanism explains how higher DNA flexibility can induce faster RAD51 polymerisation. However, the formation of a rigid RAD51 polymer on highly flexible DNA is expected to be associated with a large entropic penalty, because RAD51 polymerisation can dramatically reduce the number of spatial configurations that a flexible DNA polymer can explore. Hence, RAD51 polymerisation on flexible ssDNA is expected to incur a higher entropic penalty compared to its polymerisation on rigid dsDNA. To explain this paradox, we reasoned that the favourable enthalpic energy of RAD51 polymerisation is sufficient to compensate for any entropic penalty, referred to as the entropy-enthalpy compensation (EEC) mechanism (**Fig. 4A**).

To directly test this hypothesis, we assessed the kinetics of a RAD51 mutant, which confers a reduced enthalpic contribution compared to wild-type (WT) RAD51. We predicted that, with such a mutation, the stability of the mutant RAD51 polymers on DNA would become more dependent on the flexibility of the underlying DNA (**Fig. 4A, B**). We took advantage of a phenylalanine to glutamate substitution at RAD51 residue 86 (F86E), which reduces the RAD51 protomer-protomer interaction affinity (Esashi et al., 2007; Yu et al., 2003). Indeed, using size exclusion

chromatography with multi-angle light scattering (SEC-MALS), we confirmed that F86E RAD51 is primarily monomeric in solution (**Fig. S6A-C**). We then used SPR to measure the binding kinetics of F86E RAD51 to ssDNA dT-50, dN-50, dA-50, and dsDNA dN-50p, and fit our models to these data under the assumption that F86E RAD51 polymerisation takes place only *via* monomer addition (**Fig. 4C, D**). It was immediately evident that F86E RAD51 shows significantly reduced affinity for ssDNA and dsDNA compared to WT RAD51, with no detectable binding to DNA at the concentration of 300 nM or 3 μM (**Fig. S6D, E**). Nonetheless, at 30 μM, we observed that F86E RAD51 displays faster elongation on more flexible DNA, as we observed for WT RAD51 (**Fig. 4D**). Strikingly, F86E RAD51 displayed reduced lifetimes ($1/k_s$) that inversely correlated with the flexibility of DNA (**Fig. 4E**). This observation is in sharp contrast with WT RAD51, which formed stable, long-lived polymers independently of DNA flexibility (**Fig. 3E**). Taken together, these observations demonstrate that the enthalpic contribution of WT RAD51 polymer formation is sufficient to offset the large entropic penalty associated with polymerisation on flexible DNA. As a result, RAD51 is able to polymerise more rapidly on flexible DNA despite incurring a larger entropic penalty.

Discussion

In this study, we have analysed the kinetics of RAD51 binding to ssDNA and dsDNA oligos of varying length and flexibility *via* a combination of SPR, SAXS and mathematical modelling to understand how RAD51 discriminates between ssDNA and dsDNA. Based on our observations, we propose the DUET (**D**na molec**U**lar fl**E**xibili**T**y) model, which describes how RAD51 polymerises on ssDNA in preference to dsDNA: 1) RAD51 has a faster polymer elongation rate on flexible DNA because flexible DNA explores more conformations compatible with RAD51 binding per unit time (the BTC mechanism) (**Fig. 3A**), and 2) the enthalpic contribution of RAD51 polymerisation enables RAD51 to overcomes the entropic penalty to binding to flexible DNA (the EEC mechanism) (**Fig. 4A**).

The diploid human genome consists of 6.4 billion base pairs and ~ 50 endogenous DSBs are estimated to occur in every cell cycle (Vilenchik and Knudson, 2003). In human cells, the resection step of HR-mediated DSB repair can generate ssDNA overhangs up to 3.5 k nucleotides in length (Zhou et al., 2014), which then serve as a platform for RAD51 polymerisation. Given that ssDNA overhangs would constitute only ~ 0.005 % of the total genomic DNA (3.5 k nucleotides x two overhangs x 50 DSBs / 6.4 billion base pairs), the polymerisation of RAD51 on resected ssDNA needs to be greatly directed. The average nuclear concentration of RAD51 is estimated at ~ 100 nM (Reuter et al., 2014), and several RAD51 mediators, such as BRCA2, PALB2 and RAD52, contribute to increase the RAD51 local concentration at DSBs and/or RAD51 binding to ssDNA, while preventing its association with dsDNA (Buisson et al., 2010; Carreira and Kowalczykowski, 2011; Jensen et al., 2010; Ma et al., 2017a; Miyazaki et al., 2004; Zhao et al., 2015). Nonetheless, to date, *in vivo* evidence that these RAD51 mediators are sufficient to promote RAD51 polymerisation on DSB-derived ssDNA in preference to the bulk of undamaged dsDNA is limited.

It is important to note that resected ssDNA is first bound by RPA prior to RAD51 polymerisation. RPA is a heterotrimer complex with six-OB folds, four of which can associate tightly with ssDNA in a stepwise manner (Fan and Pavletich, 2012; Zou et al., 2006). RPA binding to ssDNA is believed to eliminate ssDNA secondary structures to facilitate RAD51 polymerisation. Significantly, previous kinetic studies have demonstrated a rapid exchange of ssDNA-bound RPA with free RPA available in solution, a phenomenon known as 'facilitated exchange' (Ma et al., 2017b). Given that one RPA heterotrimer has a footprint of 30 nucleotides and one RAD51 protomer binds to 3 nucleotide of ssDNA (Short et al., 2016), the dissociation of a single RPA heterotrimer would expose enough ssDNA to accommodate up to ten RAD51 molecules. Hence, it is reasonable to speculate that RPA dissociation provides highly flexible ssDNA, which in turn can promote the adsorption and elongation steps of RAD51 polymerisation.

We initially aimed to model RAD51 polymerisation by stepwise addition of RAD51 monomers, but with this approach we were unable to explain the ssDNA data or the dsDNA data

(**Fig. S2C and S3C**, respectively). The revised assumption of RAD51 polymerisation *via* dimers addition, presented in this study, is in line with the model proposed by Subrayaman et al. (Subramanyam et al., 2016)**.** Although the multimeric state of RAD51 in solution is not fully characterised, it is conceivable that RAD51 mediators and/or the increased local concentration RAD51 may support the availability of soluble RAD51 dimers at DSB sites.

Another major unanswered question regarding RAD51 polymerisation is how RAD51 can distinguish ssDNA generated by DSB resection from ssDNA formed during normal cellular processes, such as DNA replication and transcription. Spontaneous RAD51 polymerisation on all ssDNA would be problematic, as it may trigger undesired, toxic recombination events or disruption of DNA replication and transcription, causing genomic instability. In this context, the DUET model is particularly appealing: ssDNA generated during transcription and replication does not have free ends and could therefore be less flexible compared to resected ssDNA with a free 3' overhang. This in turn may limit RAD51 polymerisation on transcription- and replication-derived ssDNA, while promoting it on flexible ssDNA at DSBs.

This study also suggests a more general mechanism for the recruitment of DSB repair factors. Indeed, numerous DNA repair machineries are comprised of multiple subunits with several binding interfaces to DNA and other co-factors. Broken DNA ends may enhance recruitment of such repair complexes simply due to the increased DNA flexibility. Further studies in combination with RAD51 mediators, such as RAD51 paralogs and BRCA2, which are shown to stimulate RAD51 filament formation (Shahid et al., 2014; Taylor et al., 2016), will be important to further advance our full understanding of the RAD51 polymerisation mechanism. This study provides evidence for a previously unappreciated role of DNA flexibility in the process of DNA repair.

Material and methods

RAD51 Mutagenesis

The bacterial expression vector pET11d (Merck-Millipore) carrying the human WT RAD51 (WT-pET11d) was used as a template for the PCR-mediated QuikChange site-directed mutagenesis (Agilent Technologies) to introduce F86E substitution (F86E-pET11d) with a forward primer 5'-GCTAAATTAGTTCCAATGGGTGAGACC ACTGCAACTGAATTCCACC - 3' and a reverse primer 5'- GGTGGAATTCAGTTGCAGTGGTCT CACCCATTGGAACTAATTTAGC - 3'.

RAD51 Protein Purification

To express RAD51, Rosetta 2 (DE3) pLysS cells (Novagen) carrying the WT-pET11d or F86E-pET11d were grown in LB media containing 100 µg/ml ampicillin and 25 µg/ml chloroamphenicol, and by adding 0.5 mM IPTG at OD_{595} = 0.6, protein expression was induced. Cell pellets were resuspended in PBS and mixed with the equal volume of lysis buffer (3 M NaCl, 100 mM Tris-HCl pH 7.5, 4 mM EDTA pH 8, 20 mM ß-mercaptoethanol, Sigma Protease Inhibitor Cocktail (Sigma)). The suspension was sonicated and spun at 20k rpm with a 45 Ti rotor (Beckman). The supernatant was slowly mixed with 0.1 % polyethylenimine at 4°C for 1 hour and spun at 20K rpm using a 45 Ti rotor to remove DNA. The supernatant was then slowly mixed with an equal volume of 4 M ammonium sulphate (2 M final concentration) at 4°C for 1 hour and spun at 10 K rpm using a JA-17 rotor (Beckman). Pellets containing RAD51 were suspended in 25 ml of resuspension buffer (0.5 M KCl, 50 mM Tris-HCl pH 7.5, 1 mM EDTA, 2 mM DTT, 10 % glycerol) and spun again at 20k rpm to remove residual DNA. The supernatant was then dialysed overnight at 4°C in TEG buffer (50 mM Tris pH 7.5, 1 mM EDTA, 2 mM DTT, 10 % glycerol) containing 200 mM KCl (TEG200). For F86E RAD51, the cell lysate was prepared as for WT RAD51, and dialysed overnight at 4°C in TEG buffer containing 50 mM KCl (TEG75).

RAD51 purification was carried out by chromatography at 4°C using the AKTA Pure Protein Purification System (GE Healthcare). For WT RAD51, the dialysed WT RAD51 containing sample was loaded onto a 5 ml HiTrap Heparin column (GE Healthcare) and eluted *via* a linear gradient of 200 mM – 600 mM KCl. The peak fractions were pooled and dialysed in TEG200, and concentrated on a 1 ml HiTrap Q column (GE Healthcare) followed by the isocratic elusion with

TEG buffer containing 600 mM KCl (TEG600). Peak fractions containing WT RAD51 were dialysed in SPR buffer (150 mM KCl, 20 mM Hepes pH 7.5, 2 mM DTT, 10 % glycerol), aliquoted, snap frozen and stored at -80°C. Similarly, the dialysed F86E RAD51 containing cell lysate was fractionated through a 5 ml HiTrap Heparin column, but with a linear gradient of 75 mM – 600 mM KCl. The flow-through and F86E RAD51 peak fractions were pooled and dialysed in TEG buffer containing 50 mM KCl (TEG50) and reloaded onto a 5 ml HiTrap Heparin column. Following a 50 mM – 600 mM KCl linear gradient elution, F86E RAD51 peak fractions were pooled and dialysed in TEG buffer containing 100 mM KCl (TEG100). The sample was concentrated on a 1 ml HiTrap Q column followed by TEG600 isocratic elution. Peak fractions were applied on a 24 ml Superdex200 10/300 GL size-exclusion column (GE Healthcare), and fractions containing monometic F86E RAD51 were pooled and dialysed in TEG100. F86E RAD51 was then re-concentrated using a 1 ml HiTrap Q column and TEG600 isocratic elution. The peak fractions were dialysed in SPR buffer, aliquoted, snap frozen and stored at -80°C.

Multi-Angle Light Scattering

To confirm the monomeric status of F86E RAD51, the peak size-exclusion chromatography F86E RAD51 elution fraction was serially diluted (5, 1:1 serial dilutions) and loaded onto a Superdex200 10/300 GL size exclusion column (GE Healthcare) equilibrated with MALS Buffer (150 mM KCl, 50 mM Tris pH 7.5, 1 mM EDTA, 2 mM DTT). Each elution was analysed using a Wyatt Heleos8+ 8-angle light scatterer linked to a Shimadzu HPLC system comprising LC-20AD pump, SIL-20A Autosampler and SPD20A UV/Vis detector. Data collection was carried out at the Department of Biochemistry, University of Oxford. Data analysis was carried out using the ASTRA software (Wyatt).

DNA Oligo and Duplex Synthesis

The dN-5p, dN-8p and dN-11p dsDNA sequences were designed as two complimentary ssDNA sequences connected *via* two units of hexaethylene glycol (flexible linker). The dN-5, dN-8, dN-11

ssDNA molecules were designed using one of the two corresponding dsDNA annealing sequences. The dN-14 and dN-17 ssDNA molecules were designed by extending one of the two dN-11p dsDNA annealing sequences. Standard DNA phosphoramidites, solid supports, 3'-Biotin-TEG CPG and additional reagents were purchased from Link Technologies Ltd and Applied Biosystems Ltd. All oligonucleotides were synthesized on an Applied Biosystems 394 automated DNA / RNA synthesizer using a standard 1.0 *μ*mole phosphoramidite cycle of acid-catalyzed detritylation, coupling, capping, and iodine oxidation. Stepwise coupling efficiencies and overall yields were determined by the automated trityl cation conductivity monitoring facility and in all cases were >98.0%. All β-cyanoethyl phosphoramidite monomers were dissolved in anhydrous acetonitrile to a concentration of 0.1 M immediately prior to use. The coupling time for normal A, G, C, and T monomers was 60 s, and the coupling time for the hexaethylene glycol phosphoramidite monomer (from link) was extended to 600 s. Cleavage of the oligonucleotides from the solid support and deprotection was achieved by exposure to concentrated aqueous ammonia solution for 60 min at room temperature followed by heating in a sealed tube for 5 h at 55 °C. Purification of oligonucleotides was carried out by reversed-phase HPLC on a Gilson system using a Brownlee Aquapore column (C8, 8 mm x 250 mm, 300 Å pore) with a gradient of acetonitrile in triethylammonium bicarbonate (TEAB) increasing from 0% to 50% buffer B over 20 min with a flow rate of 4 ml/min (buffer A: 0.1 M triethylammonium bicarbonate, pH 7.0, buffer B: 0.1 M triethylammonium bicarbonate, pH 7.0 with 50% acetonitrile). Elution of oligonucleotides was monitored by ultraviolet absorption at 295 or 300 nm. After HPLC purification, oligonucleotides were freeze dried then dissolved in water without the need for desalting. All oligonucleotides were characterised by negative-mode HPLC-mass spectrometry using either a Bruker micrOTOFTM II focus ESI-TOF mass spectrometer with an Acquity UPLC system, equipped with a BEH C18 column (Waters) or a Waters Xevo G2-XS QT mass spectrometer with an Acquity UPLC system, equipped with an Acquity UPLC oligonucleotide BEH C18 column (particle size: 1.7 μm; pore size: 130 Å; column dimensions: 2.1 x 50 mm). Data were analysed using Waters MassLynx software or Waters UNIFI Scientific Information System software.

Small Angle X-Ray Scattering

The flexibilities of the ssDNA dT-50, ssDNA dN-50, ssDNA dA-50 and the dsDNA dN-50p molecules were assessed using small-angle x-ray scattering (SAXS) at the Diamond Light Source (Harwell, UK). SAXS data were collected using a size exclusion KW 402.5 (2.4 ml) column (Shodex). 50 µl of DNA sample was injected and elution was carried out at 37°C at 75 µl/min using SPR running buffer in the absence of glycerol and BSA. The flexible cylinder model was fit to the four scattering data sets within the 0.0037 - 0.27 [1/Å] range to derive the persistence lengths. Data plotting and fitting was carried out using the SasView software for SAXS data analysis.

Surface Plasmon Resonance

The binding kinetics of WT and F86E RAD51 on ssDNA and dsDNA were assessed *via* surface plasmon resonance (SPR) using a Biacore T200 (GE Healthcare). CM5 SPR chip flow cells were activated by injecting 100 µl of a 1:1 N-Hydroxysuccinimide (NHS), ethyl(dimethylaminopropyl) carbodiimide (EDC) mix at 10 µl/min. 100 µl of purified streptavidin (1 mg/ml) was then injected over the flow cells at 10 µl/min. Unbound amine groups were then deactivated by injecting 50 µl of 1 M ethanolamine-HCl. 50 µl of 10 mM glycine pH 2.5 was injected to remove uncoupled, sterically bound streptavidin. For all experiments, ~ 8 - 15 RU of biotinylated DNA substrates were immobilised onto either flow cells 2, 3 or 4. After DNA immobilisation, 10 µl of purified biotin was injected over all four flow cells at 10 µl/min to block the remaining free streptavidin binding sites. Chip activation and ligand immobilisation steps were carried out at 25°C. Using high performance injections, the CM5 chip surfaces were then primed *via* 10 injections of 10 µl SPR running buffer (150 mM KCl, 20 mM Hepes pH 7.5, 2 mM DTT, 5 mg/ml BSA, 2.5 mM ATP pH 7.5, 10 mM $CaCl_2$, 10 % glycerol) at 30 µl/min. WT or F86E RAD51 diluted in SPR running buffer to the specified concentration was injected at 30 µl/min (association), followed by the injection of SPR running buffer for 20 minutes at 30 µl/min (dissociation) at 37°C.

SPR Data Processing

SPR traces were processed using the BIAEvaluation software (GE Healthcare) as following. First, the negative control flow cell trace (Fc1 or Fc3) was subtracted from the experimental flow cell trace. The trace was then vertically and horizontally aligned, such that the start of the protein injection occurs at time = 0 s, Response = 0 RU. Subsequently, the drift trace between the negative control flow cell and the experimental flow cell during the final priming injection was subtracted from the experimental flow cell trace (double-referencing) (Myszka, 1999). All experimental curves were normalised using the following equation:

$$\boldsymbol{N} = \boldsymbol{S} / (\boldsymbol{L} * (\boldsymbol{M_{RAD51}} / \boldsymbol{M_{DNA}}))$$

N is the normalised signal (mean number of RAD51 molecules per DNA oligo), ***S*** is the signal in RU units (1 RU ~ 50 pg/mm^2), ***L*** is the amount of DNA ligand immobilised onto the experimental flow cell (RU), $\boldsymbol{M_{RAD51}}$ is the molecular weight (kDa) of RAD51 (~ 37 kDa) and $\boldsymbol{M_{DNA}}$ is the molecular weight (kDa) of the immobilised DNA molecule.

Mathematical Model Development and Analysis

During SPR injections, the concentration of free analyte (i.e. RAD51) in solution in the flow cell remains constant because only a very small proportion of analyte binds to the surface. For this reason, the ordinary differential equation (ODE) models were formulated using first order mass action kinetics *via* the rule-based modelling language BioNetGen (Harris et al., 2016) , and analysed using MATLAB 2016b lsqcurvefit unless otherwise specified. The ABC-SMC simulations were carried out in MATLAB 2016b to determine a unique set of parameters that can explain the SPR data of ssDNA (Fig. S2A, B) and of dsDNA (Fig. S3A, B). The ABC-SMC simulation for the ssDNA was unable to determine a unique value for $k_{ad.r}$ and $k_{ad.f}$, but explained the data when either higher or lower values are concurrently applied to both $k_{ad.r}$ and $k_{ad.f}$. This result is due to the fact that only a lower bound can be determined for the adsorption dissociation constant ($K_{ad} = k_{ad.r}$ /

$k_{ad.f}$), as RAD51 binding to the ssDNA dN-5, dN-8 or dN-11 was undetectable. Nonetheless, the accurate fit for the ssDNA dN-50 (R^2 = 0.95) suggests that the model captures the essential features of RAD51 polymerisation on ssDNA. The ABC-SMC simulation for the dsDNA model suggests all four model parameters are well defined, as the $k_{ad.f}$ and $k_{ad.r}$ can be fitted directly based on the dsDNA dN-8p data. All scripts are freely available at:

https://www.dropbox.com/s/xh1z2h7eo9rmybw/RAD51%20Polymerisation%20Models.zip?dl=0

Statistical Analysis

Unpaired, two-tailed t-tests were carried out to test for significant differences between the different fitted RAD51 elongation rates and lifetimes (Fig. 3D, E, 4D and E). All t-tests were carried out using GraphPad Prism 7.

Table 2. The list of DNA oligos used in this study.

Oligo Name	Mass (kDa)	Sequence
ssDNA dN-5	2.057	5'-CGGAC-Biotin TEG-3'
dsDNA dN-5p	4.286	5'-CGGAC **LL** GTCCG-Biotin TEG-3'
ssDNA dN-8	2.594	5'-CTGACTGC- Biotin TEG-3'
dsDNA dN-8p	6.139	5'-CTGACTGC **LL** GCAGTCAG-Biotin TEG-3'
ssDNA dN-11	3.926	5'-CGTCGATAGGC-Biotin TEG-3'
dsDNA dN-11p	7.993	5'-CGTCGATAGGC **LL** GCCTATCGACG-Biotin TEG-3'
ssDNA dN-14	4.872	5'-CGTCGATGAGCAGT-Biotin TEG-3'
ssDNA dN-17	5.795	5'-CGTCGATGAGCAGTGTC-Biotin TEG-3'
ssDNA dN-50	15.682	5'-Biotin-TTGAGAGAGCAGACCACAATTATCACCTACACGACATCATTTTATATCAA-3'
dsDNA dN-50p	31.157	(Forward) 5'-Biotin-TCGAGAGGGTAAACCACAATTATCGCCTACCCAAAACTATTTTATATCAA-3' (Reverse) 5'-TTGATATAAAATAGTTTTGGGTAGGCGATAATTGTGGTTTACCCTCTCGA-3'
ssDNA dT-50	15.541	5'-Biotin-TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT-3'
ssDNA dA-50	15.992	5'-Biotin-AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA-3'

L: hexaethylene glycol flexible linker. **Biotin TEG**: Biotin linked to a flexible triethylene glycol spacer. **Biotin**: Biotin linked to a standard C6 spacer. **dN**: mixed base composition. **dT**: deoxythymine. **dA**: deoxyadenine.

Author contributions

FE, OD and FP conceived and planned the project. FP conducted all SPR experiments, protein purification and mathematical modelling. AES and TB designed and generated the short DNA oligos for SPR experiments. JA contributed to the conceptualisation of the thermodynamic impact of polymerisation. FE, OD and FP wrote the manuscript with input from all contributing authors.

Acknowledgements

FE and OD are supported by Wellcome Trust Senior Research Fellowships in Basic Biomedical Science (101009/Z/13/Z and 207537/Z/17/Z, respectively). JA is supported by National Science Foundation grant (DMS 1454739). TB is thankful for the support by Biotechnology and Biological Sciences Research Council (BB/J001694/2). FP is a recipient of the Systems Biology Doctoral Training Centre Scholarship, funded by Engineering and Physical Sciences Research Council. We thank Marcus Bridge for assistance with SPR, Nicola Trendel for assistance with ABC-SMC, David Staunton for the SEC-MALS analysis of the RAD51 mutant, and Robert Rambo for the DNA oligo SAXS analysis. We also acknowledge the use of the University of Oxford Advanced Research Computing (ARC) facility in carrying out this work (http://dx.doi.org/10.5281/zenodo.22558).

Figure legends

Figure 1. **Kinetics of RAD51 binding to ssDNA of varying length.**

A and B. Biotinylated ssDNA oligos of indicated lengths were separately immobilised onto SPR CM5 chips *via* biotin-streptavidin interaction. WT RAD51 was injected at the indicated concentrations to measure the association and dissociation kinetics. The estimated number of maximum RAD51 molecules per oligo was calculated using the structurally determined RAD51 : nucleotide/base pair stoichiometry of 1:3 (Short et al., 2016). The dotted and solid curves respectively show the binding of RAD51 to DNA oligos measured by SPR and the ODE model fits. **C**. Kinetic representation of the ordinary differential equation (ODE) model describing RAD51

polymerisation on ssDNA, consisting of five parameters: $k_{ad.f}$ (adsorption forward rate), $k_{ad.r}$ (adsorption reverse rate), $k_{el.f}$ (elongation forward rate), $k_{el.r}$ (elongation reverse rate), k_s (stabilisation rate). Dotted arrows describe monomer binding and unbinding, solid arrows indicate dimer binding. The ODE model was calibrated by simultaneously fitting the SPR curves of ssDNA dN-8, dN-14, dN-17 and dN-50 using the best fit ABC-SMC particle. Mean values ± 1 SEM of 3 ODE model fits (n=3).

Figure 2. **Kinetics of RAD51 binding to dsDNA of varying length.**

A and B. Biotinylated dsDNA molecules of indicated lengths were separately immobilised onto SPR CM5 chips *via* biotin-streptavidin interaction, and WT RAD51 was injected at the indicated concentrations to measure association and dissociation kinetics. The estimated number of maximum RAD51 molecules per oligo was calculated using the structurally determined RAD51 : nucleotide/base pair stoichiometry of 1:3 (Short et al., 2016). The dotted and solid curves respectively show the binding of RAD51 to DNA oligos measured by SPR and the ODE model fits. **C**. Kinetic representation of the ordinary differential equation (ODE) model describing RAD51 polymerisation on dsDNA, consisting of four parameters: $k_{ad.f}$ (adsorption forward rate), $k_{ad.r}$ (adsorption reverse rate), $k_{el.f}$ (elongation forward rate), k_s (stabilisation rate). Dotted arrows describe monomer binding and unbinding, solid arrows indicate dimer binding. The ODE model was calibrated by simultaneously fitting the SPR curves of dsDNA dN-5p, dN-8p, dN-11p and dN-50p using the best fit ABC-SMC particle. Mean values ± 1 SEM of 3 ODE model fits (n=3).

Figure 3. **Kinetics of RAD51 binding to DNA of varying flexibility.**

A. A depiction of the Bend-to-Capture (BTC) mechanism. RAD51 generates two sequential, non-covalent interactions (a RAD51 protomer-protomer interaction and a RAD51-DNA interaction, or *vice versa*) faster on flexible DNA (depicted as ssDNA) compared to rigid DNA (depicted as dsDNA). The free RAD51 to be incorporated into the growing polymer is shown with an asterisk. **B**.

Biotinylated DNA molecules of varying flexibility were separately immobilised onto SPR CM5 chips *via* biotin-streptavidin interaction, and WT RAD51 was injected at the indicated concentrations to measure association and dissociation kinetics. The expected entropic penalties upon RAD51 binding to respective DNA are indicated in green boxes. **C**. WT RAD51 SPR curves for respective DNA oligos (dotted lines) and ODE model fits (solid lines; fitted $k_{el.f}$ values are highlighted in the corresponding colour). **D and E**. Bar plots of fitted $k_{el.f}$ values (D) and k_s values (E) for each SPR curves. In panel D, all parameters except $k_{el.f}$ were fixed to the mean values as identified in Fig. 1C and 2C, and $k_{el.f}$ was fitted using lsqcurvefit (MATLAB). In panel E, all parameters except k_s and $k_{el.f}$ were fixed to the mean values as identified in Fig. 1C and 2C, and k_s and $k_{el.f}$ were fitted using lsqcurvefit (MATLAB). Mean ± 1 SD of 3 ODE model fits (n=3). Asterisks indicate unpaired, two-tailed t-tests: * $p < 0.05$; ** $p < 0.01$. ns = non-significant.

Figure 4. **Kinetics of monomeric RAD51 F86E binding to DNA of varying flexibility.**

A. A depiction of the Entropy-Enthalpy Compensation (EEC) mechanism. The enthalpic contribution of WT RAD51 polymerisation enables RAD51 to overcome the entropic penalty to binding flexible DNA (here depicted as ssDNA). F86E RAD51 has a lower RAD51 protomer-protomer affinity and a consequent lower enthalpic contribution to polymerisation, and thus cannot overcome the entropic penalty for binding to flexible DNA. **B**. Biotinylated DNA molecules of varying flexibility were separately immobilised onto SPR CM5 chips *via* biotin-streptavidin interaction, and F86E RAD51 was injected at the indicated concentrations to measure association and dissociation kinetics. The expected entropic penalties upon F86E RAD51 binding to respective DNA are indicated in green boxes. **C.** F86E RAD51 SPR curves (dotted lines) and corresponding model fits using the ssDNA model (ssDNA dT-50, dN-50, dA-50) and the dsDNA model (dsDNA dN-50p). RAD51 lifetimes ($1/k_s$) are shown in the corresponding colours. A single forward rate ($k_{el.f}$) and a single reverse rate (k_s) were fitted to the experimental data using lsqcurvefit (MATLAB). **D and E.** Bar plots of fitted $k_{el.f}$ values (D) and k_s values (E) for each SPR curves. Mean ± 1 standard

deviation of 3 ODE model fits (n=3). Asterisks indicate unpaired, two-tailed t-tests: * $p < 0.05$; ** $p < 0.01$, *** $p < 0.001$. ns = non-significant.

Table 1. Persistence lengths (L_p) ± 95% confidence intervals are estimated from fitted Kuhn lengths for the ssDNA dN-50, dA-50, dT-50 and the dsDNA dN-50p SAXS plots. All data manipulation and model fitting were done using SasView. It is likely that the relatively low dsDNA dN-50p L_p of 7.34 nm compared to the average dsDNAp L_p of 30 - 55 nm (Baumann et al., 1997; Brunet et al., 2015) is due to the low GC content of the dN-50p (36%).

Figure S1. Experimental setup.

A. Depiction of the experimental setting. A biotinylated DNA oligo is immobilised onto a surface plasmon resonance (SPR) CM5 chip *via* biotin-streptavidin interaction. RAD51 protein is injected over the DNA-coated SPR matrix to measure polymerisation kinetics. Throughout this period, association and dissociation of RAD51 take place simultaneously. Following protein injection (stop), running buffer is injected to measure dissociation kinetics. **B**. WT RAD51 SPR curves for ssDNA dN-50 and dsDNA dN-50p. RAD51 was injected at 150 nM in the presence of 2.5 mM ATP (pH 7.5) and 10 mM $CaCl_2$. Curves were normalised to the R_{max}, whereby the normalised response value of X (vertical axis) indicates the average number of RAD51 molecule bound to a single DNA oligo.

Figure S2. **Evaluation of models describing RAD51 binding to ssDNA and dsDNA**

A. ABC-SMC simulation for the ssDNA model. Heat maps indicate particle frequency and describe pair-wise correlations between model parameters ($k_{ad.f}$, $k_{ad.r}$, $k_{el.f}$, $k_{el.r}$, and k_s). Each vertical and horizontal axis of heat maps is labelled with the colour of the corresponding parameter ($k_{ad.f}$ - purple, $k_{ad.r}$ - brown, $k_{el.f}$ - orange, $k_{el.r}$ - black, k_s - green). A high particle frequency corresponds to a good model fit to the data for the specified parameter pair. $k_{ad.r}$ and $k_{ad.f}$ are correlated. The

probability densities for individual parameters are presented along the top left diagonal and were estimated using a Gaussian kernel. Particles were initialised *via* uniform priors with the following lower and upper bounds: $0.0001 < k_{ad.f} < 0.01$, $0.001 < k_{ad.r} < 20$, $1 < k_{el.f} < 1000$, $0.0001 < k_{el.r} < 0.01$, $0.00001 < k_s < 0.005$. **B.** Adaptive threshold for the ABC-SMC simulation. The plateau at 10 iterations means that the model has reached a best fit to the data. The simulation was carried out using in-house MATLAB ABC-SMC scripts with 10 iterations and 2000 particles. **C**. ssDNA_dN-X data fit as in Fig. 1B, except assuming that RAD51 polymerisation occurs only *via* the addition of RAD51 monomers. **D**. ssDNA_dN-X data fit as in Fig. 1B, except in the absence of the elongation step ($k_{el.r} = k_s$). In panels C and D, parameter values are derived using the best fit ABC-SMC particles (described in Fig. 1). **E**. Sum-of-squares residuals (SSRs) for the three ssDNA models, described in Fig. 1C, S2C and S2D. The chosen model (shown in Fig. 1C), circled in red, best describes the experimental data. **F**. Experimental data of ssDNA dN-50 from Fig. 1 (dotted lines) and predicted polymers with defined length on each single DNA oligo using the model from Fig. 1 (solid lines). Predicted polymer curves were calculated by multiplying the concentration of a specific polymer by its length (i.e. [n-mer](t) * n). **G and H**. As in F, except ssDNA dN-17 (G) and ssDNA dN-14 (H) experimental data and predicted polymers are shown.

Figure S3. **Evaluation of models describing RAD51 binding to dsDNA**

A. ABC-SMC simulation for the dsDNA model. Heat maps indicate particle frequency and describe pair-wise correlations between model parameters ($k_{ad.f}$, $k_{ad.r}$, $k_{el.f}$, and k_s). Each vertical and horizontal axis of heat maps is labelled with the colour of the corresponding parameter ($k_{ad.f}$ - purple, $k_{ad.r}$ - brown, $k_{el.f}$ - orange, k_s - green). A high particle frequency corresponds to a good model fit to the data for the specified parameter pair. $k_{ad.f}$, $k_{ad.r}$, $k_{el.f}$, and k_s are well determined. The probability densities for individual parameters are presented along the top left diagonal and were estimated using a Gaussian kernel. Particles were initialised *via* uniform priors with the following lower and upper bounds: $0.0001 < k_{ad.f} < 0.01$, $0.001 < k_{ad.r} < 20$, $0.1 < k_{el.f} < 10$, $0.00001 < k_s <$

0.001. **B.** Adaptive threshold for the ABC-SMC simulation. The plateau at 10 iterations means the model has reached a best fit to the data. **C**. dsDNA_dN-Xp data fit as in Fig. 2, except assuming polymerisation only *via* RAD51 monomer addition. **D**. dsDNA_dN-Xp data fit as in Fig. 2, except in the presence of the elongation step (i.e. $k_{el.r} \neq k_s$). In panel F and G, parameter values are derived using the best fit ABC-SMC particles (described in Fig. 2). **E**. Sum-of-squares residuals (SSRs) for the three dsDNA models, described in Fig. 2C, S3C and S3D. The chosen model (shown in Fig. 2C), circled in red, describes the experimental data equally well as the model in S3D but requires only four parameters ($k_{ad.f}$, $k_{ad.r}$, $k_{el.f}$, k_s). **F**. Experimental data for the dsDNA dN-50p from Fig. 2 (dotted lines) and predicted polymers with defined length on each single DNA oligo using the dsDNA model from Fig. 2 (solid lines). Predicted polymer curves were calculated by multiplying the concentration of a specific polymer by its length (i.e. [n-mer](t) * n). **G and H**. As in F, except dN-11p (G) and dN-8p (H) experimental data and predicted polymers are shown.

Figure S4. The flexibility measurements of DNA oligos used in this study.

A-D. The flexibilities of ssDNA dT-50 (A), dN-50 (B), dA-50 (C) and dsDNA dN-50p (D) were assessed by small-angle x-ray scattering (SAXS). The left panels show the absolute scattering intensity (cm^{-1}) measured as a function of the scattering vector ($Å^{-1}$). The flexible cylinder model was fit to the data (fitting range: 0.0037 - 0.27 $Å^{-1}$) to predict the kuhn lengths (persistence length = kuhn length / 2). The right panels show residuals of the flexible cylinder model fits.

Figure S5. **Evaluation of models describing RAD51 binding to DNA of varying flexibility.**

A-D. RAD51 SPR curves (dotted lines) for the ssDNA dT-50, dN-50 and dA-50 were fitted by varying a single parameter of the ssDNA model using lsqcurvefit (MATLAB). Each panel shows the ODE model based on a varying adsorption forward rate constant ($k_{ad.f}$) (**A**), adsorption reverse rate constant ($k_{ad.r}$) **(B),** elongation reverse rate constant ($k_{el.r}$) **(C),** and the stabilisation rate constant (k_s) (**D**). A higher or lower $k_{ad.f}$, or $k_{ad.r}$, but not $k_{el.r}$, or k_s, can generate accurate model fits.

Figure S6. The characterisations of F86E RAD51 used in this study.

A. Purified WT and F86E RAD51 on an SDS-PAGE gel are visualized by coomassie blue staining. **B.** Size-exclusion multi-angle light scattering (SEC-MALS) curves for F86E RAD51 for the five elution concentrations as indicated. Dotted lines represent the molar mass derived values for each elution peak, solid lines represent light scattering. **C.** Calculated molecular weights of the five F86E RAD51 elution peaks. Data were analysed using the ASTRA 6.1 chromatography software for SEC-MALS. RSE: relative standard error. **D and E.** F86E RAD51 displays reduced affinity for ssDNA and dsDNA. F86E RAD51 SPR binding curves for the ssDNA dT-50 (D) and the dsDNA dN-50p (E) at the colour-coded concentrations are shown.

References

Baumann, C.G., Smith, S.B., Bloomfield, V.A., and Bustamante, C. (1997). Ionic effects on the elasticity of single DNA molecules. Proceedings of the National Academy of Sciences *94*, 6185-6190.

Baumann, P., Benson, F.E., and West, S.C. (1996). Human Rad51 Protein Promotes ATP-Dependent Homologous Pairing and Strand Transfer Reactions In Vitro. Cell *87*, 757-766.

Benson, F.E., Stasiak, A., and West, S.C. (1994). Purification and characterization of the human Rad51 protein, an analogue of E. coli RecA. The EMBO journal *13*, 5764-5771.

Brunet, A., Tardin, C., Salome, L., Rousseau, P., Destainville, N., and Manghi, M. (2015). Dependence of DNA Persistence Length on Ionic Strength of Solutions with Monovalent and Divalent Salts : A Joint Theory − Experiment Study. Macromolecules *48*, 3641-3652.

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

Buisson, R., Dion-Côté, A.-M., Coulombe, Y., Launay, H., Cai, H., Stasiak, A.Z., Stasiak, A., Xia, B., and Masson, J.-Y. (2010). Cooperation of breast cancer proteins PALB2 and piccolo BRCA2 in stimulating homologous recombination. Nature structural & molecular biology *17*, 1247-1254.

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

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

Chen, R., and Wold, M.S. (2014). Replication protein A: Single-stranded DNA's first responder. BioEssays *36*, 1156-1161.

Chi, Q., Wang, G., and Jiang, J. (2013). The persistence length and length per base of single-stranded DNA obtained from fluorescence correlation spectroscopy measurements using mean field theory. Physica A *392*, 1072-1079.

Esashi, F., Galkin, V.E., Yu, X., Egelman, E.H., and West, S.C. (2007). Stabilization of RAD51 nucleoprotein filaments by the C-terminal region of BRCA2. Nature Structural and Molecular Biology *14*, 468-474.

Fan, J., and Pavletich, N.P. (2012). Structure and conformational change of a replication protein A heterotrimer bound to ssDNA. Genes & Development *26*, 2337-2347.

Harris, L.A., Hogg, J.S., Tapia, J.J., Sekar, J.A., Gupta, S., Korsunsky, I., Arora, A., Barua, D., Sheehan, R.P., and Faeder, J.R. (2016). BioNetGen 2.2: advances in rule-based modeling. Bioinformatics *32*, 3366-3368.

Hilario, J., Amitani, I., Baskin, R.J., and Kowalczykowski, S.C. (2009). Direct imaging of human Rad51 nucleoprotein dynamics on individual DNA molecules. Proceedings of the National Academy of Sciences *106*, 361-368.

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

Kang, J., Jung, J., and Keun, S. (2014). Flexibility of single-stranded DNA measured by single-molecule FRET. Biophysical Chemistry *195*, 49-52.

Ma, C., Kwon, Y., Sung, P., and Greene, E.C. (2017a). Human RAD52 interactions with replication protein A and the RAD51 presynaptic complex. Journal of Biological Chemistry *292*, 11702-11713.

Ma, C.J., Gibb, B., Kwon, Y., Sung, P., and Greene, E.C. (2017b). Protein dynamics of human RPA and RAD51 on ssDNA during assembly and disassembly of the RAD51 filament. Nucleic Acids Research *45*, 749-761.

Miné, J., Disseau, L., Takahashi, M., Cappello, G., Dutreix, M., and Viovy, J.L. (2007). Real-time measurements of the nucleation, growth and dissociation of single Rad51-DNA nucleoprotein filaments. Nucleic Acids Research *35*, 7171-7187.

Miyazaki, T., Bressan, D.A., Shinohara, M., Haber, J.E., and Shinohara, A. (2004). In vivo assembly and disassembly of Rad51 and Rad52 complexes during double-strand break repair. EMBO Journal *23*, 939-949.

Murphy, M.C., Rasnik, I., Cheng, W., Lohman, T.M., and Ha, T. (2004). Probing Single-Stranded DNA Conformational Flexibility Using Fluorescence Spectroscopy. Biophysical Journal *86*, 2530-2537.

Myszka, D.G. (1999). Improving biosensor analysis. Journal of molecular recognition : JMR *12*, 279-284.

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

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

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

Short, J.M., Liu, Y., Chen, S., Soni, N., Madhusudhan, M.S., Shivji, M.K.K., and Venkitaraman, A.R. (2016). High-resolution structure of the presynaptic RAD51 filament on single-stranded DNA by electron cryo-microscopy. Nucleic Acids Research *44*, 9017-9030.

Sim, A.Y.L., Lipfert, J., Herschlag, D., and Doniach, S. (2012). Salt dependence of the radius of gyration and flexibility of single-stranded DNA in solution probed by small-angle x-ray scattering. Physical Review E *86*, 1-5.

Subramanyam, S., Ismail, M., Bhattacharya, I., and Spies, M. (2016). Tyrosine phosphorylation stimulates activity of human RAD51 recombinase through altered nucleoprotein filament dynamics. Proceedings of the National Academy of Sciences.

Taylor, M.R.G., Ma, C.J., Greene, E.C., Krejci, L., Boulton, S.J., Ma, C.J., Carzaniga, R., Takaki, T., and Collinson, L.M. (2016). A Polar and Nucleotide-Dependent Mechanism of Action for RAD51 Paralogs in RAD51 Filament Article A Polar and Nucleotide-Dependent Mechanism of Action for RAD51 Paralogs in RAD51 Filament Remodeling. Molecular Cell *64*, 926-939.

Toni, T., Welch, D., Strelkowa, N., Ipsen, A., and Stumpf, M.P. (2009). Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. J R Soc Interface *6*, 187-202.

Tubbs, A., and Nussenzweig, A. (2017). Endogenous DNA Damage as a Source of Genomic Instability in Cancer. Cell *168*, 644-656.

Van der Heijden, T., Seidel, R., Modesti, M., Kanaar, R., Wyman, C., and Dekker, C. (2007). Real-time assembly and disassembly of human RAD51 filaments on individual DNA molecules. Nucleic Acids Research *35*, 5646-5657.

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

Yu, D.S., Sonoda, E., Takeda, S., Huang, C.L.H., Pellegrini, L., Blundell, T.L., and Venkitaraman, A.R. (2003). Dynamic control of Rad51 recombinase by self-association and interaction with BRCA2. Molecular Cell *12*, 1029-1041.

Zhao, W., Vaithiyalingam, S., Filippo, J.S., Chazin, W.J., Wiese, C., Zhao, W., Vaithiyalingam, S., Filippo, J.S., Maranon, D.G., and Jimenez-sainz, J. (2015). Promotion of BRCA2-Dependent Homologous Recombination by DSS1 via RPA Targeting and DNA Mimicry Article Promotion of BRCA2-Dependent Homologous Recombination by DSS1 via RPA Targeting and DNA Mimicry. Molecular Cell *59*, 176-187.

Zhou, Y., Caron, P., Toulouse, F., and Paull, T.T. (2014). Quantitation of DNA double-strand break resection intermediates in human cells. Nucleic Acids Research *42*.

Zou, Y.U.E., Liu, Y., Wu, X., and Shell, S.M. (2006). Functions of Human Replication Protein A (RPA): From DNA Replication to DNA Damage and Stress Responses. Journal of Cellular Physiology *273*, 267-273.

Figure 1

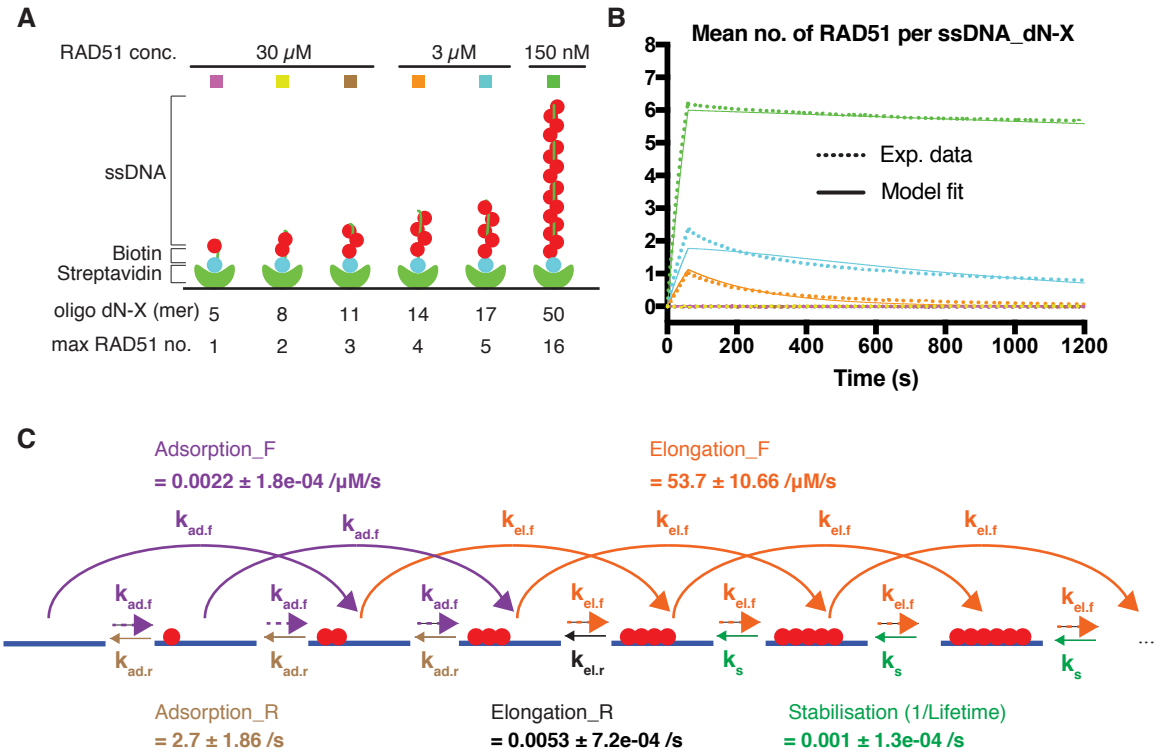

A
RAD51 conc.
30 μM
3 μM
150 nM
ssDNA
Biotin
Streptavidin
oligo dN-X (mer) 5 8 11 14 17 50
max RAD51 no. 1 2 3 4 5 16
B
Mean no. of RAD51 per ssDNA_dN-X
Exp. data
Model fit
Time (s)
C
Adsorption_F
= 0.0022 ± 1.8e-04 /μM/s
Elongation_F
= 53.7 ± 10.66 /μM/s
Adsorption_R
= 2.7 ± 1.86 /s
Elongation_R
= 0.0053 ± 7.2e-04 /s
Stabilisation (1/Lifetime)
= 0.001 ± 1.3e-04 /s

Figure 2

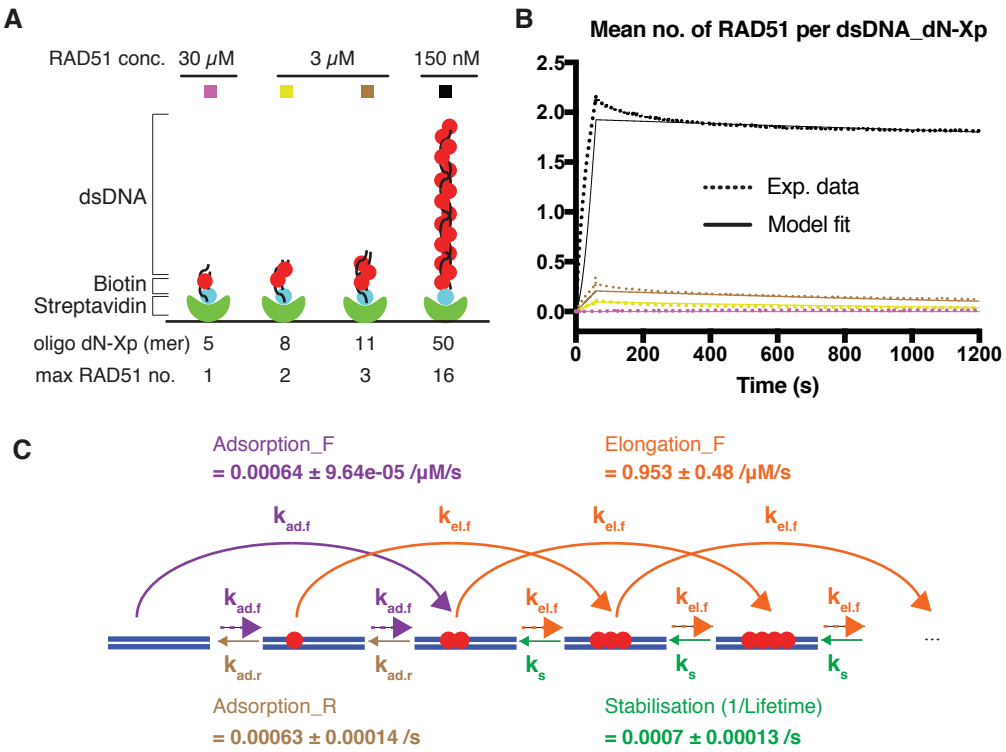

A
RAD51 conc.
30 μM
3 μM
150 nM
dsDNA
Biotin
Streptavidin
oligo dN-Xp (mer) 5 8 11 50
max RAD51 no. 1 2 3 16
B
Mean no. of RAD51 per dsDNA_dN-Xp
Exp. data
Model fit
Time (s)
C
Adsorption_F
= 0.00064 ± 9.64e-05 /μM/s
Elongation_F
= 0.953 ± 0.48 /μM/s
Adsorption_R
= 0.00063 ± 0.00014 /s
Stabilisation (1/Lifetime)
= 0.0007 ± 0.00013 /s

Figure 3

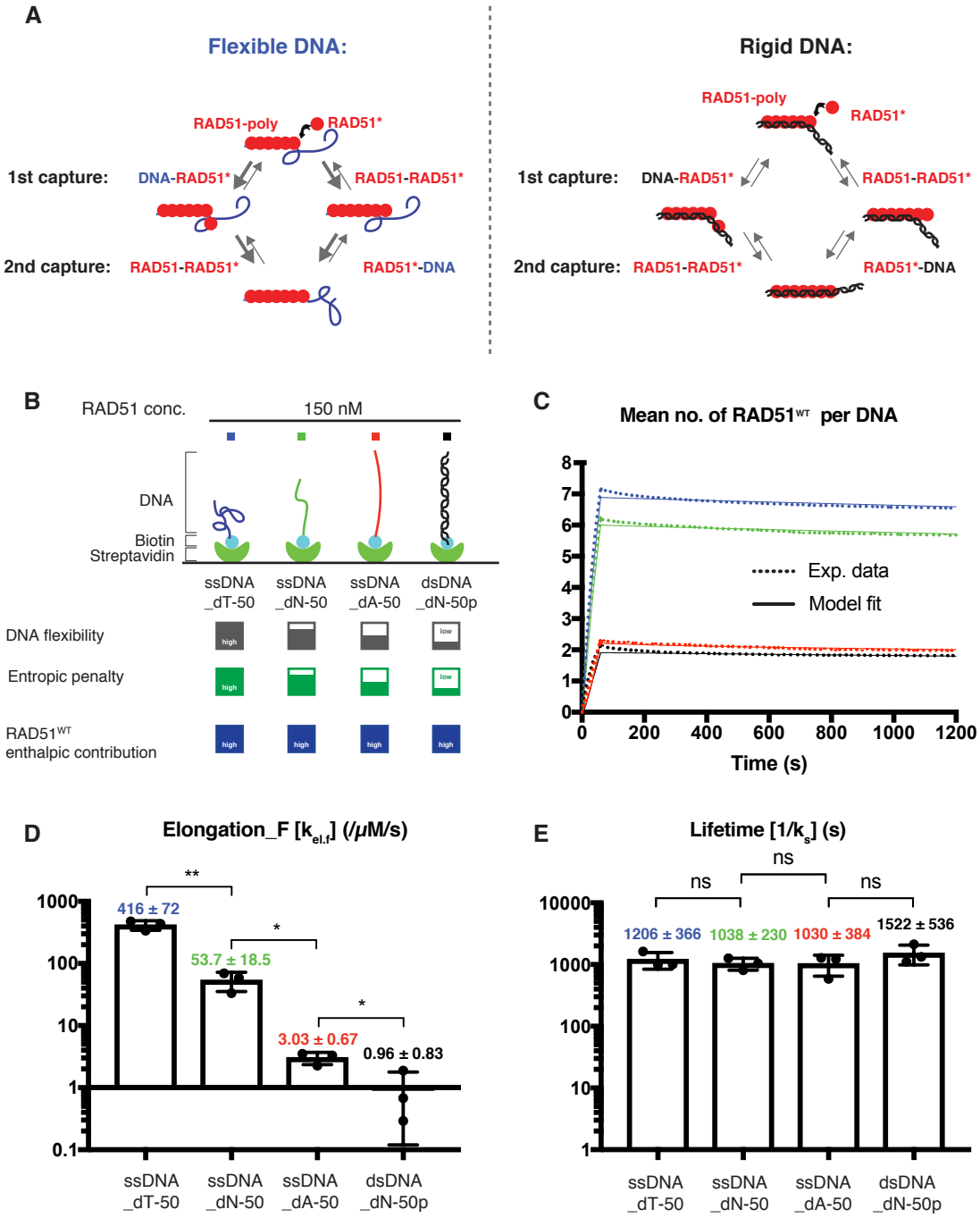

A
Flexible DNA:
RAD51-poly
RAD51*
1st capture:
DNA-RAD51*
RAD51-RAD51*
2nd capture:
RAD51-RAD51*
RAD51*-DNA
Rigid DNA:
RAD51-poly
RAD51*
1st capture:
DNA-RAD51*
RAD51-RAD51*
2nd capture:
RAD51-RAD51*
RAD51*-DNA
B
RAD51 conc.
150 nM
DNA
Biotin
Streptavidin
ssDNA _dT-50
ssDNA _dN-50
ssDNA _dA-50
dsDNA _dN-50p
DNA flexibility
Entropic penalty
RAD51WT enthalpic contribution
high
low
C
Mean no. of RAD51WT per DNA
Exp. data
Model fit
Time (s)
D
Elongation_F [kel.f] (/μM/s)
416 ± 72
53.7 ± 18.5
3.03 ± 0.67
0.96 ± 0.83
ssDNA _dT-50
ssDNA _dN-50
ssDNA _dA-50
dsDNA _dN-50p
E
Lifetime [1/ks] (s)
ns
1206 ± 366
1038 ± 230
1030 ± 384
1522 ± 536
ssDNA _dT-50
ssDNA _dN-50
ssDNA _dA-50
dsDNA _dN-50p

Figure 4

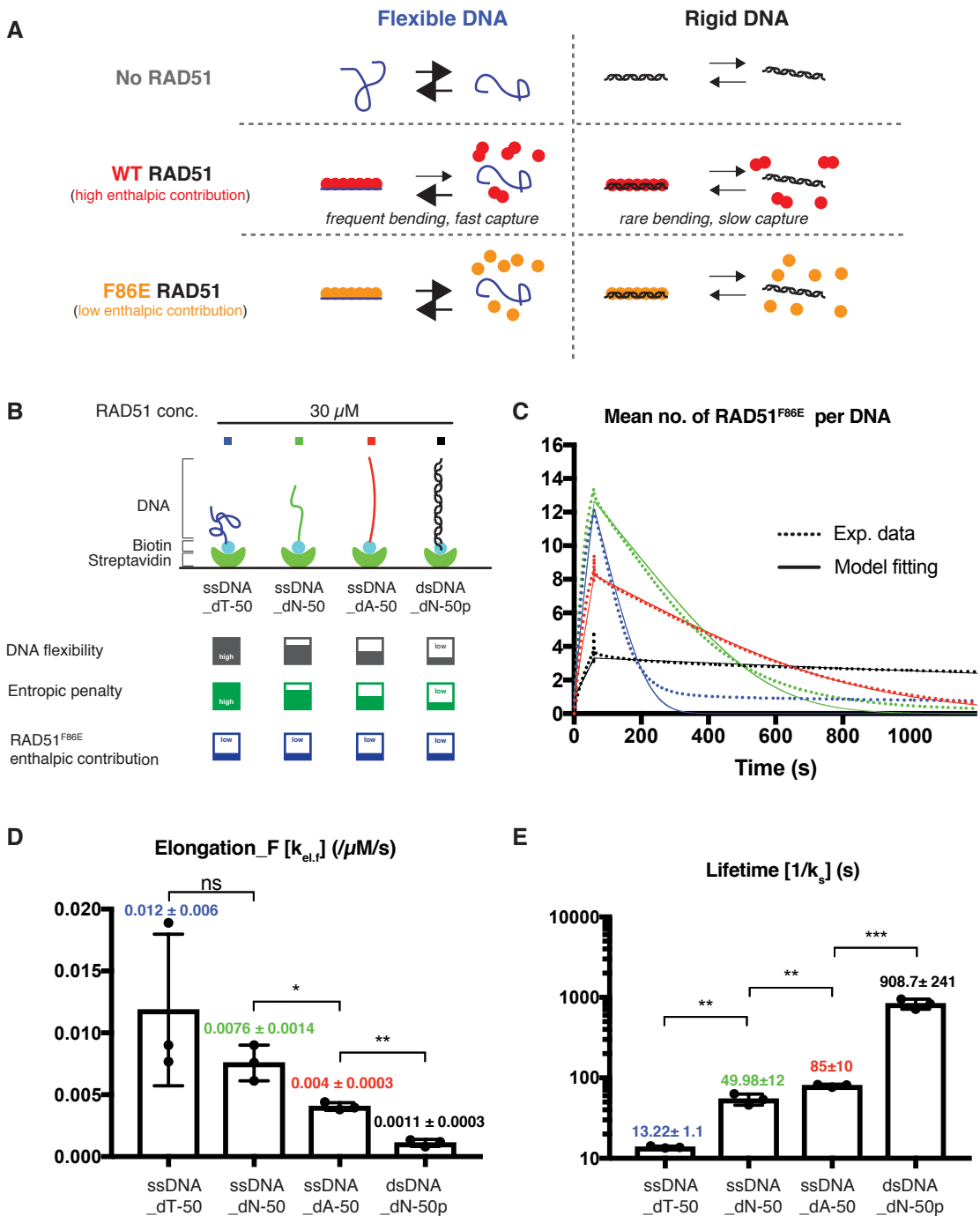

A
Flexible DNA
Rigid DNA
No RAD51
WT RAD51
(high enthalpic contribution)
frequent bending, fast capture
rare bending, slow capture
F86E RAD51
(low enthalpic contribution)
B
RAD51 conc.
30 μM
DNA
Biotin
Streptavidin
ssDNA _dT-50
ssDNA _dN-50
ssDNA _dA-50
dsDNA _dN-50p
DNA flexibility
Entropic penalty
RAD51F86E enthalpic contribution
high
low
C
Mean no. of RAD51F86E per DNA
Exp. data
Model fitting
Time (s)
D
Elongation_F [kel.f] (/μM/s)
ns
0.012 ± 0.006
0.0076 ± 0.0014
0.004 ± 0.0003
0.0011 ± 0.0003
E
Lifetime [1/ks] (s)
13.22± 1.1
49.98±12
85±10
908.7± 241

Table 1. SAXS-derived persistence length measurement of DNA oligos.

Oligo Name	kuhn length (nm)	persistence length (nm)	persistence length error (95% CI) (nm)
ssDNA dT-50	2.6702	1.3351	0.200185
ssDNA dN-50	4.5177	2.25885	0.048095
ssDNA dA-50	9.3488	4.6744	0.13126
dsDNA dN-50p	14.674	7.337	0.07379

Figure S1

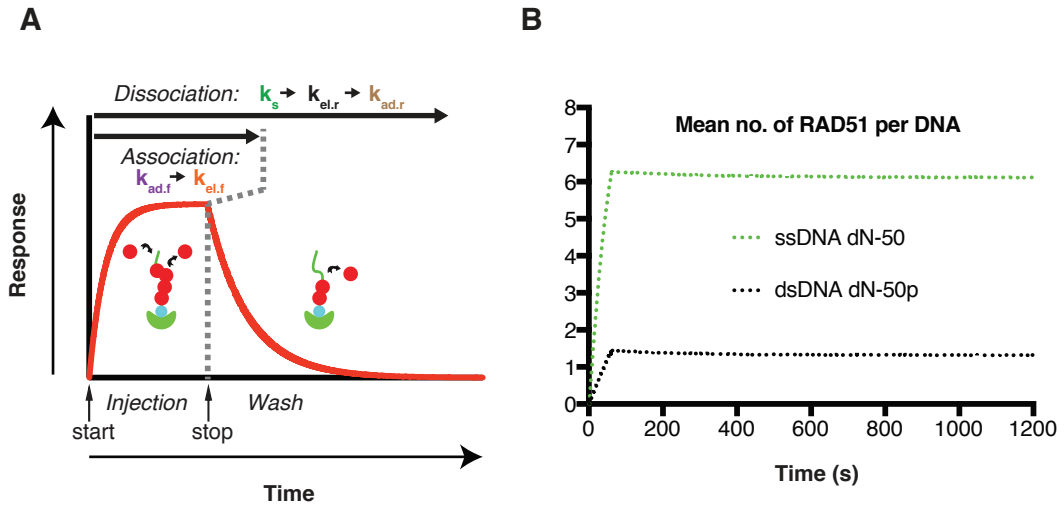

A
Dissociation: k_s → $k_{el.r}$ → $k_{ad.r}$
Association: $k_{ad.f}$ → $k_{el.f}$
Response
Injection
Wash
start
stop
Time
B
Mean no. of RAD51 per DNA
ssDNA dN-50
dsDNA dN-50p
0
1
2
3
4
5
6
7
8
0
200
400
600
800
1000
1200
Time (s)

Figure S2

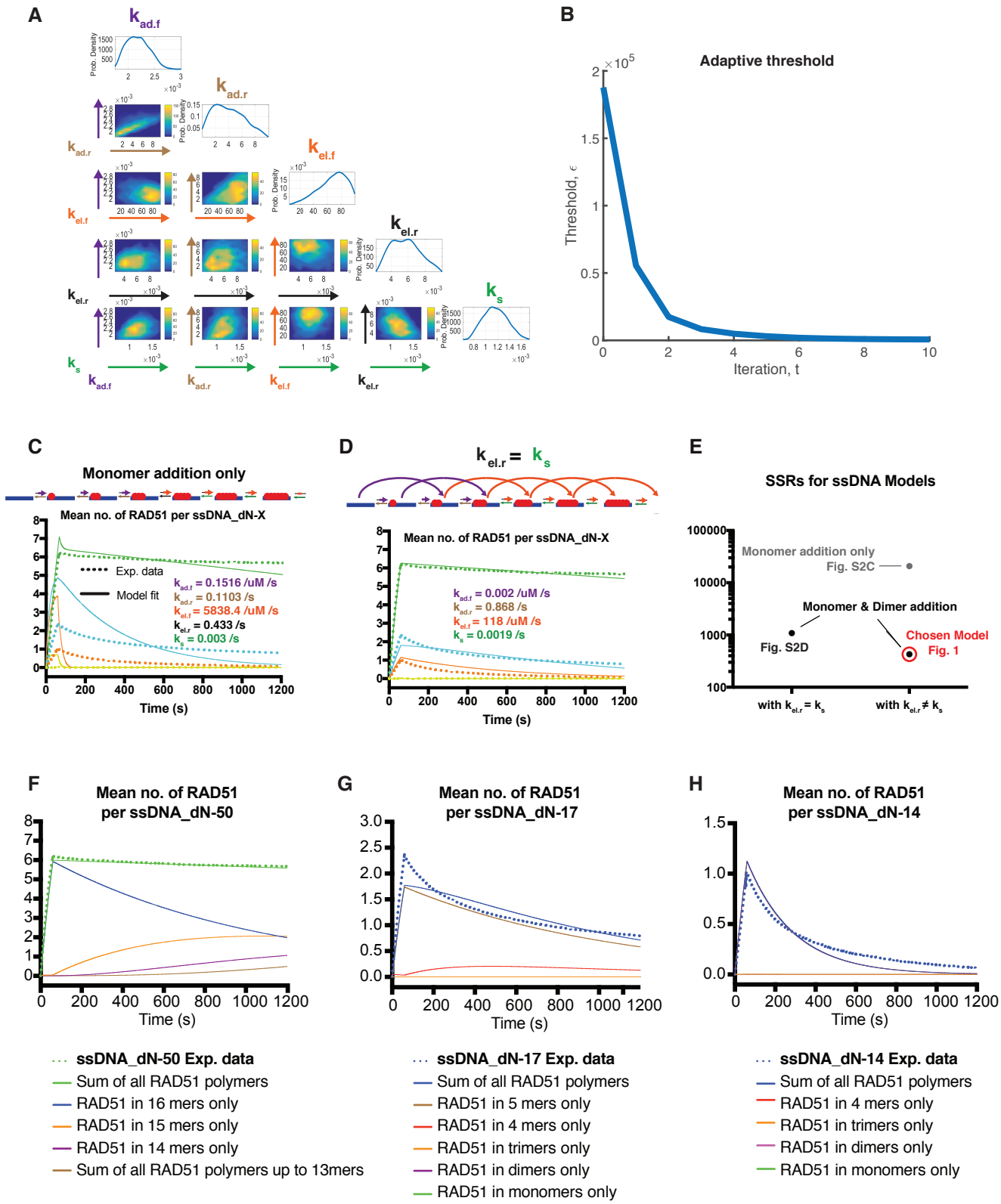

A
B
Adaptive threshold
Threshold, ϵ
Iteration, t
C
Monomer addition only
Mean no. of RAD51 per ssDNA_dN-X
Exp. data
Model fit
Time (s)
D
Mean no. of RAD51 per ssDNA_dN-X
Time (s)
E
SSRs for ssDNA Models
Monomer addition only
Fig. S2C
Monomer & Dimer addition
Fig. S2D
Chosen Model
Fig. 1
F
Mean no. of RAD51 per ssDNA_dN-50
Time (s)
G
Mean no. of RAD51 per ssDNA_dN-17
Time (s)
H
Mean no. of RAD51 per ssDNA_dN-14
Time (s)
ssDNA_dN-50 Exp. data
Sum of all RAD51 polymers
RAD51 in 16 mers only
RAD51 in 15 mers only
RAD51 in 14 mers only
Sum of all RAD51 polymers up to 13mers
ssDNA_dN-17 Exp. data
Sum of all RAD51 polymers
RAD51 in 5 mers only
RAD51 in 4 mers only
RAD51 in trimers only
RAD51 in dimers only
RAD51 in monomers only
ssDNA_dN-14 Exp. data
Sum of all RAD51 polymers
RAD51 in 4 mers only
RAD51 in trimers only
RAD51 in dimers only
RAD51 in monomers only

Figure S3

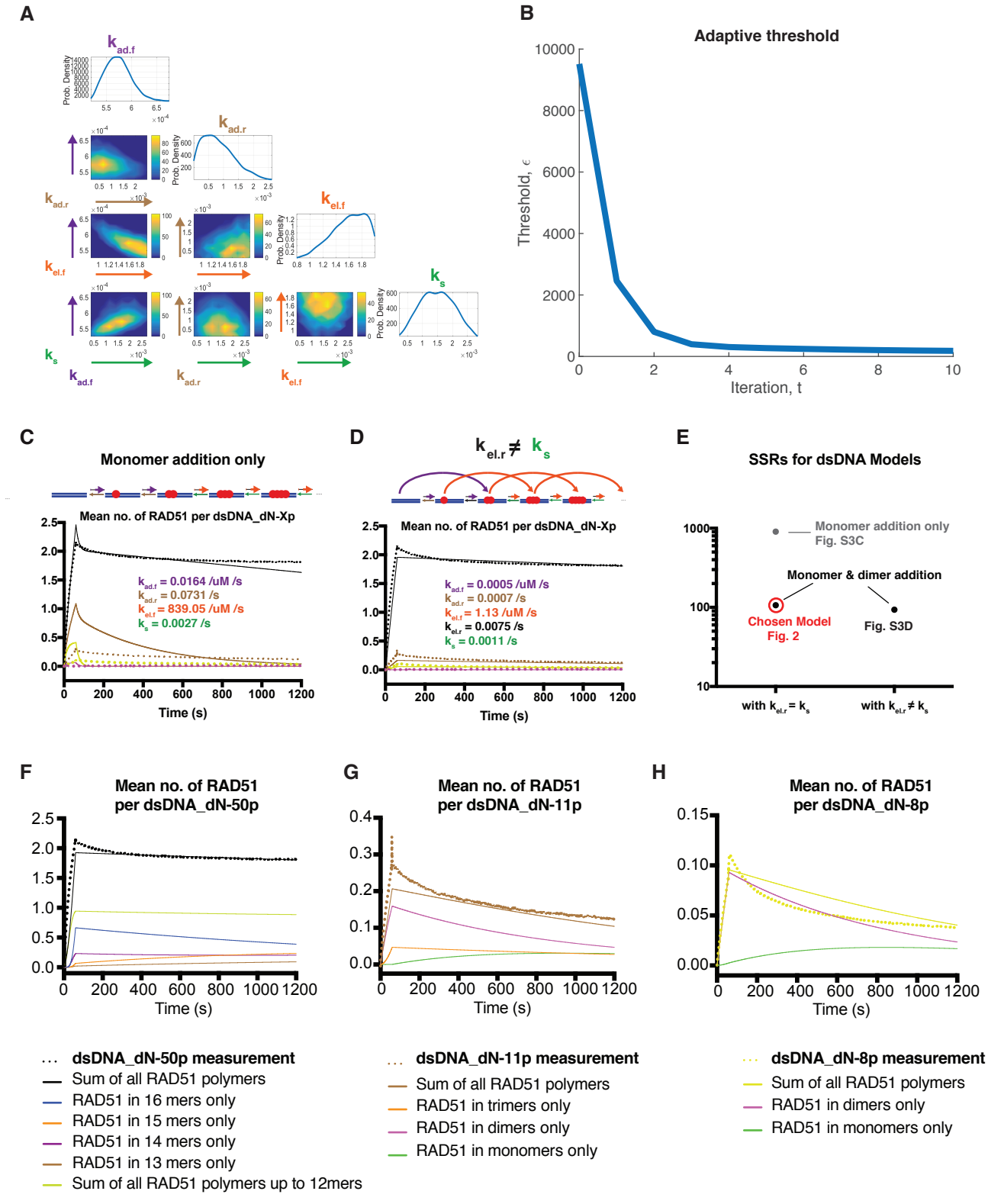

A
B
Adaptive threshold
Threshold, ϵ
Iteration, t
C
Monomer addition only
Mean no. of RAD51 per dsDNA_dN-Xp
k_ad.f = 0.0164 /uM /s
k_ad.r = 0.0731 /s
k_el.f = 839.05 /uM /s
k_s = 0.0027 /s
Time (s)
D
k_el.r ≠ k_s
Mean no. of RAD51 per dsDNA_dN-Xp
k_ad.f = 0.0005 /uM /s
k_ad.r = 0.0007 /s
k_el.f = 1.13 /uM /s
k_el.r = 0.0075 /s
k_s = 0.0011 /s
Time (s)
E
SSRs for dsDNA Models
Monomer addition only
Fig. S3C
Monomer & dimer addition
Chosen Model
Fig. 2
Fig. S3D
with k_el.r = k_s
with k_el.r ≠ k_s
F
Mean no. of RAD51
per dsDNA_dN-50p
Time (s)
G
Mean no. of RAD51
per dsDNA_dN-11p
Time (s)
H
Mean no. of RAD51
per dsDNA_dN-8p
Time (s)
dsDNA_dN-50p measurement
Sum of all RAD51 polymers
RAD51 in 16 mers only
RAD51 in 15 mers only
RAD51 in 14 mers only
RAD51 in 13 mers only
Sum of all RAD51 polymers up to 12mers
dsDNA_dN-11p measurement
Sum of all RAD51 polymers
RAD51 in trimers only
RAD51 in dimers only
RAD51 in monomers only
dsDNA_dN-8p measurement
Sum of all RAD51 polymers
RAD51 in dimers only
RAD51 in monomers only

Figure S4

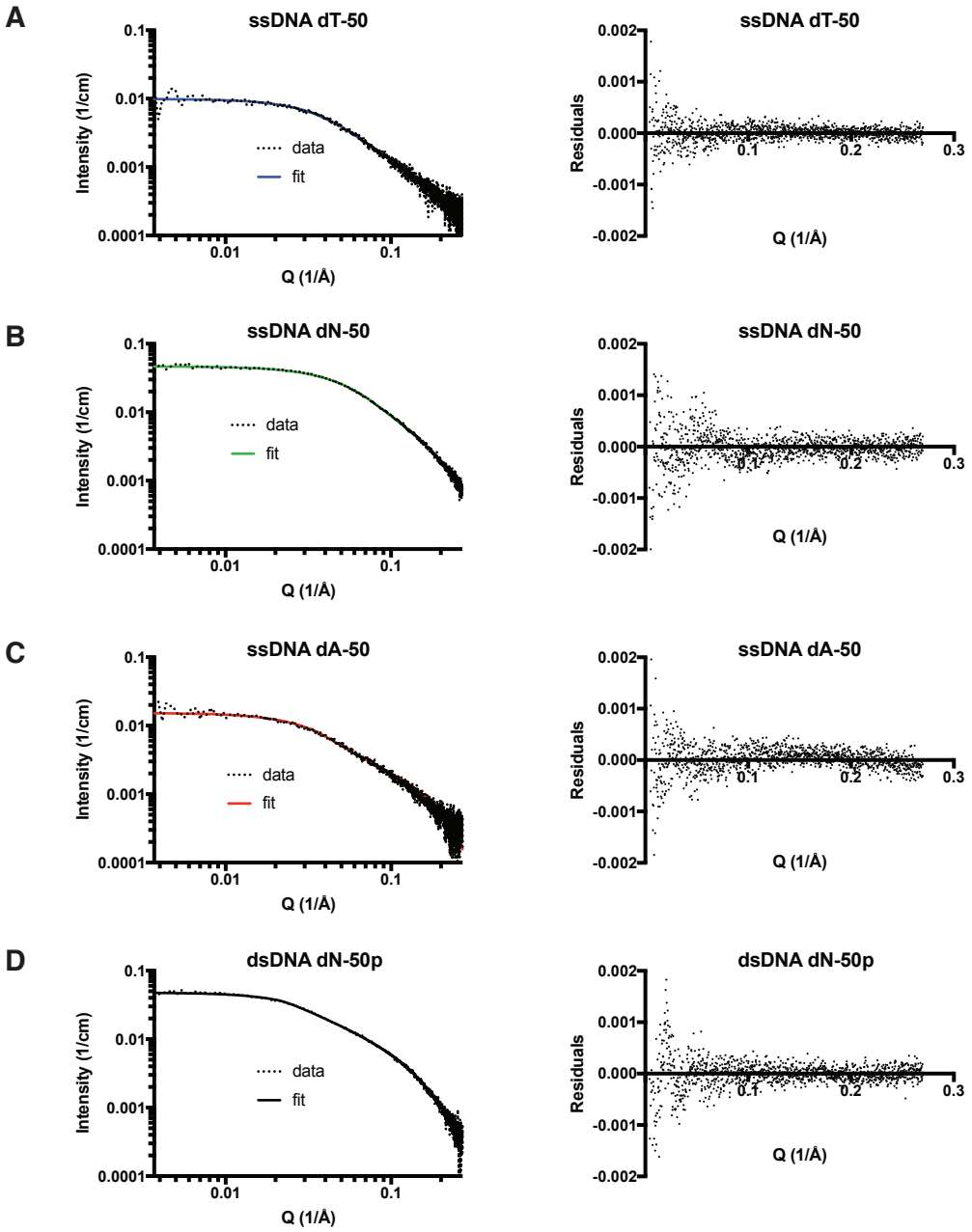

A
ssDNA dT-50
Intensity (1/cm)
0.1
0.01
0.001
0.0001
data
fit
0.01
0.1
Q (1/Å)
ssDNA dT-50
Residuals
0.002
0.001
0.000
-0.001
-0.002
0.1
0.2
0.3
Q (1/Å)
B
ssDNA dN-50
Intensity (1/cm)
data
fit
Q (1/Å)
ssDNA dN-50
Residuals
Q (1/Å)
C
ssDNA dA-50
Intensity (1/cm)
data
fit
Q (1/Å)
ssDNA dA-50
Residuals
Q (1/Å)
D
dsDNA dN-50p
Intensity (1/cm)
data
fit
Q (1/Å)
dsDNA dN-50p
Residuals
Q (1/Å)

Figure S5

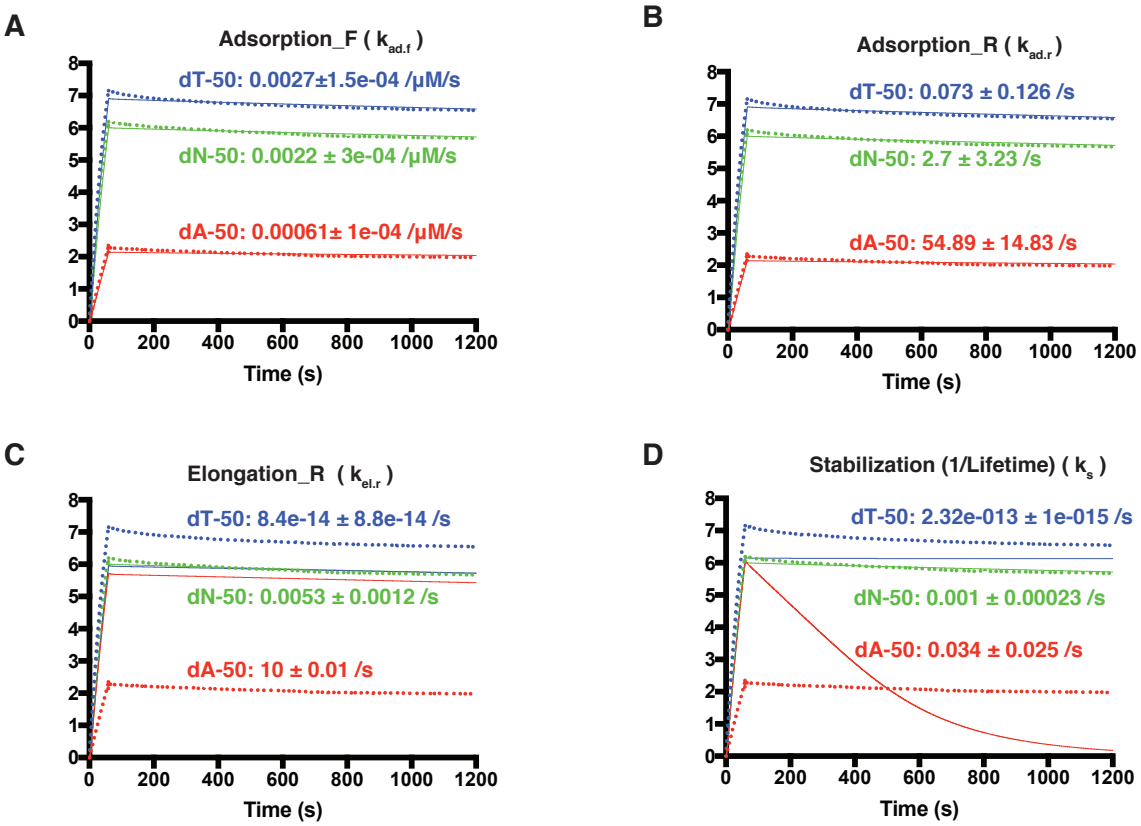
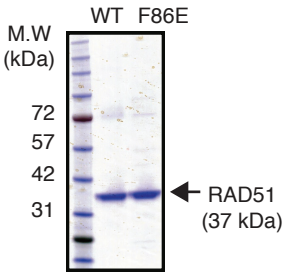
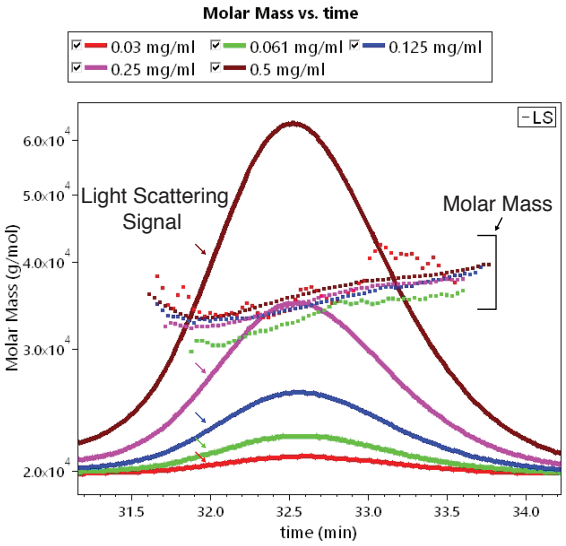

A
Adsorption_F ($k_{ad.f}$)
dT-50: 0.0027±1.5e-04 /μM/s
dN-50: 0.0022 ± 3e-04 /μM/s
dA-50: 0.00061± 1e-04 /μM/s
Time (s)
B
Adsorption_R ($k_{ad.r}$)
dT-50: 0.073 ± 0.126 /s
dN-50: 2.7 ± 3.23 /s
dA-50: 54.89 ± 14.83 /s
Time (s)
C
Elongation_R ($k_{el.r}$)
dT-50: 8.4e-14 ± 8.8e-14 /s
dN-50: 0.0053 ± 0.0012 /s
dA-50: 10 ± 0.01 /s
Time (s)
D
Stabilization (1/Lifetime) (k_s)
dT-50: 2.32e-013 ± 1e-015 /s
dN-50: 0.001 ± 0.00023 /s
dA-50: 0.034 ± 0.025 /s
Time (s)

Figure S6

A



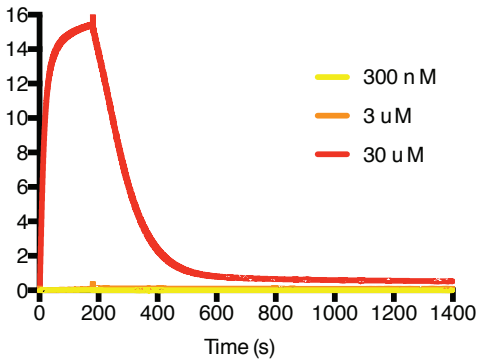
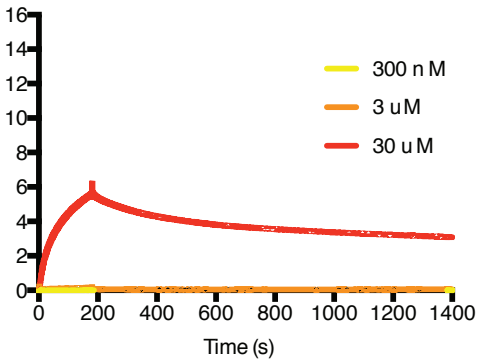
B



C

[F86E] (mg/ml)	Mw (kDa)	RSE
0.03	36.2	4.6%
0.061	33.3	1.1%
0.125	34.9	0.8%
0.25	34.9	0.2%
0.5	35.9	0.4%
Average	35	

D



E