BIOPHYSICAL STUDIES ON MECHANISMS OF HOMOLOGOUS RECOMBINATIONAL PROTEINS

LE SHIMIN

NATIONAL UNIVERSITY OF SINGAPORE

2015

BIOPHYSICAL STUDIES ON MECHANISMS OF HOMOLOGOUS RECOMBINATIONAL PROTEINS

LE SHIMIN

B. Sci., B.Soc.Sci., Xiamen University, 2010

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

MECHANOBIOLOGY INSTITUTE

NATIONAL UNIVERSITY OF SINGAPORE

2015

DECLEARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis. This thesis has also not been submitted for any degree in any university previously.

LE Shimin
October 16, 2015

Acknowledgement

It's a great pleasure to take this opportunity to express my sincere gratitude to those who had guided, supported and helped me along the way.

First and foremost, I would like to thank my Ph.D. supervisor, Professor Yan Jie, for his enormous support, invaluable guidance, persistent encouragement, and consistent trust throughout my entire Ph.D. period. He discusses science and life with us as friends, settles a stress-free atmosphere in the lab, which make my years of research and study full of relax and happiness. His hard-working, positive attitude and generous offers of opportunities have stimulated me, and helped to establish my confidence and belief on the way to becoming an independent scientist.

I am also grateful to Prof. Hew Choy Leong and Prof. Low Boon Chuan, *et al.* for their appreciation and advices during the MBI admission interview and during my Ph.D period. MBI has provided a wonderful academic environment for my research.

I am also greatly indebted to Dr. Fu Hongxia, Dr. Chen Hu, Dr. Zhang Xinghua, Dr. Cong Peiwen, Dr. Yao Mingxi, and Dr. Artem Yefremov for their help, advices and contributions to our collaborations, and Ms Chen Jin and Ms Zhao Wenwen for their hardworking and help during our collaborations. I am really grateful for the fortune to work with these amazing bunch of colleagues & friends.

I would also like to thank Prof. Michael Cox (University of Wisconsin, Madison), Prof. John Marko (Northwestern University), Prof. K. Muniyappa (Indian Institute of Science), Prof. Juan Alonso (Spanish National Biotechnology Centre), Prof. Blake Wiedenheft (Montana State University), Prof. G.V. Shivashankar (National University of Singapore), Prof. Linda Kenny (National Uni-

versity of Singapore), Prof. Peter Dröge (Nanyang Technological University), and my thesis advisory committees: Prof. Yusuke Toyama (National University of Singapore) and Prof. Wang Zhisong (National University of Singapore), for their simulating advices during my Ph.D. studies.

It is so lucky for me to be supported by and companioned with my lab-mates and friends – Chen Hu, Hongxia, Xinghua, Yingjie, Lin Jie, Peiwen, Yanan, Yuanyuan, Lim Ci Ji, Mingxi, Yuan Xin, Xiaodan, Ricksen, Wenwen, Chen Jin, Rangit, Lee Sin Yi, Wong Wei Juan, Huijuan ..., and my friends outside the lab – Zhao Zhihai, Sun Jichao, Li Jiawei, Wenwei, Ziyu, Zhao Chen, Zhang Bo, Ye Guanqiong, and Li Qiushi, as well as LuJingChenXia.

Lastly and most importantly, I would like to thank my family: my parents and my sisters for their understanding, consideration and support during all these years.

October 16, 2015

Contents

Summary

The integrity of DNA is crucial for the survival of living cells. However, DNA in cells is frequently damaged by both exogenous and endogenous agents. DNA lesions, on the other hand, are also essential steps in cellular process, such as DNA replication, recombination and conjugations. Both the accidental and programmed DNA damages have to be repaired efficiently and accurately; left un-repaired or mis-repaired, DNA damages lead to genome instability, resulting in oncogenous effects or even cell death. Among several major repair pathways, homologous recombination pathway precisely repairs the damaged DNA by searching and using a homologous DNA sequence in the genome as repairing template.

The bacterial RecA nucleoprotein filament formed on ssDNA at the broken DNA ends is the essential player during homologous recombination: it governs the homologous strand search, invasion and exchange. The formation and stability of the filament have to be tightly and precisely regulated by a set of accessory proteins and environmental co-factors; either insufficient or unlimited growth of the filament is lethal. However, the regulatory mechanisms of the filament by these accessory proteins remain elusive, partly due to lack of single-filament resolution studies.

Furthermore, force has been increasingly recognized as an important or even key determinant in diverse biological processes. Forces ubiquitously present on DNA, generated by molecular motors or condensation of topologically constrained DNA. Moreover, increasing evidences have suggested the existence of force on damaged DNA during homologous recombination. However, the potential regulatory role of force on homologous recombination have not been studied yet.

In this thesis work, I aim to understand the molecular mechanisms of dynamics and stability of RecA filament regulated by the accessory proteins and co-factors at single-filament level, and elucidate the role of mechanical force on these processes. I firstly developed a platform for single-ssDNA manipulation using magnetic tweezers, which enables the studies of dynamics and regulation of individual RecA-ssDNA nucleoprotein filaments with high signal-to-noise ratio. Then, I systematically investigated the dynamics and stability of RecA filament regulated by several key accessory proteins including SSB, RecX, RecO, RecR, and force.

I showed that SSB outcompetes RecA binding to ssDNA, inhibiting the nucleation of RecA filament, and de-stablizes pre-formed RecA filament by occupying the vacated ssDNA site at low forces, while RecX promotes ATP-hydrolysis-dependent, step-wise net-depolymerization of RecA filament at low forces. In addition, I showed that RecO tightly folds ssDNA, while RecOR highly extends ssDNA. RecOR stabilizes pre-formed RecA filaments against net de-stablization effect caused by SSB.

Remarkably, I discovered that physiological level forces antagonize the inhibitory effects of SSB and RecX, facilitating repolymerization of partially depolymerized RecA filament in a $3'$-to-$5'$ direction. These findings demonstrated important and potentially broad regulatory functions of force during homologous recombination. Further, theoretical analysis also consistently suggests that force fine-tunes the formation and stability of RecA filament in a biphasic manner. Moreover, the work also suggests the existence of bi-directional polymerization of RecA, in contrast to previously widely accepted $5'$-to-$3'$ unidirectional RecA polymerization.

In summary, the thesis work establishs a framework of molecular mechanisms of dynamic and regulation of RecA filament mediated by accessory proteins and co-factors, and highlights the potential broad regulatory role of force during RecA-dependent homologous recombinational DNA repair.

List of Tables

List of Figures

Chapter 1

Introduction

This thesis is mainly focused on the studies on regulatory mechanisms of bacterial homologous recombination proteins, as well as the potential important role of mechanical force during the process. This chapter is written to introduce the basic background and frameworks on the subjects covered in my thesis studies. To begin with, reviews on DNA damage and homologous recombination pathway are presented to readers in section 1.1, followed by reviews of several major regulatory proteins involved in DNA recombination repair in section 1.2. Further, physiological functions of mechanical force in various cellular processes and its presence during homologous recombination are reviewed in section 1.3. Next, reviews on micro-bio-mechanics of both double-stranded DNA (dsDNA) and single-stranded DNA (ssDNA), as well as the general effects of DNA processing proteins on DNA micro-mechanics are presented to readers in section 1.4. Finally, the objectives and organization of following chapters are presented to readers in section 1.5.

1.1 DNA damage and homologous recombination repair

1.1.1 Genome stability and DNA damages

Living species in a chaotic universe relies on a unique ability of cells to store, retrieve, translate, and maintain the genetic information required to make and

maintain a living organism [1, 2]. The genetic information, termed as genes, is mainly carried in DNA in cells. However, the DNA in cells is frequently damaged by chemicals and radiation from environment as well as by intracellular thermal accidents and molecules [1–6]. Left unrepaired or mis-repaired, DNA lesions caused by damages may lead to oncogenous effects or even cell death [1–6]. Therefore, robust DNA damage repair systems are essential for life survival.

DNA lesions usually include several types of damages on DNA, such as double strands breaks (DSB), single-strand gaps (SSG) and inter-strand crosslinks (ICL) [1, 6–8]. DSB may be induced by certain exogenous agents, such as a variety of chemical compounds (e.g. methyl methanesulfonate and bleomycin) or ionizing radiation, as well as endogenous agents, for instance, replication errors and other metabolites produced in cells [1, 6, 7]. SSGs can appear on one of the newly synthesized daughter strands during semi-conservative DNA replication when errors happened [1, 6, 9]. ICL which blocks DNA replication completely by preventing DNA strand separation is a special class of chemical damage to DNA [6, 8]. In addition, SSG may be converted into a DSB if the gapped ssDNA is cleaved; the DSB may also produce some ssDNA tails; the ICL would be separated into two SSG when incisions take place on the ICL [1, 6, 8, 9].

On the other hand, DNA lesions are not only resulted from damages, but also essential steps in cellular processes. For example, DSB also occurs in DNA during meiotic and mitotic recombination, DNA replication, and restriction endonucleases or topoisomerases actions. SSG also happens during conjugation and natural transformation [1, 6, 9].

Importantly, both the accidental or programmed DNA lesions have to be repaired efficiently and accurately. Several major repairing pathways have been evolved to repair DNA damages and maintain genome stability. For SSG, the other undamaged strand can be used as template for repairing. The major pathways for SSG include base excision repair, nucleotide excision repair and mismatch repair [6]. Furthermore, three major pathways have been evolved to repair the DSB: non-homologous end joining, microhomology-mediated end joining, and homologous recombination [1, 6, 10]. Significantly, homologous recombination precisely repairs the damaged DNA by searching and using a homologous

DNA sequence in the genome as repairing template [1, 6, 10, 11].

1.1.2 Homologous Recombination

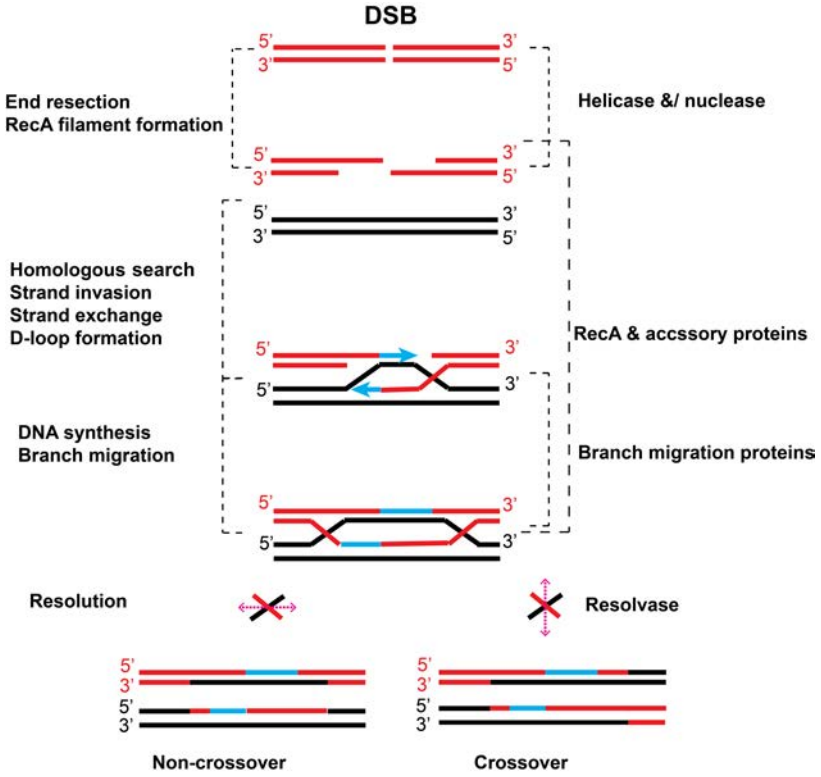

Figure 1.1: **Current model for RecA-dependent DNA homologous recombination**

.

Homologous recombination (HR) is one of the most important pathways to repair DSB and resolve the replication fork collapses, thereby playing essential roles in genome maintenance [3–7, 10, 11]. Decades of intensive studies have come up with several working models of bacterial HR (Figure 1.1) [1, 6, 7, 10, 11], which can be in general divided into several steps, namely, 1). initiation (presynaptic filament formation), 2). strand search and homologous recognition, and 3). D-loop formation, migration, and strand exchange. To begin with, HR is in most cases initiated by creation of an ssDNA region by DNA helicase enzymes. This ssDNA segment is then bound with recombinase protein, RecA in bacteria, which nucleates and polymerizes on the ssDNA, forming an elongated helical nucleoprotein filament, i.e. the presynaptic filament. Secondly, the RecA-ssDNA filament then searches for homologous sequence in an intact dsDNA. Once the

homologous sequence is recognized, the presynaptic RecA nucleoprotein filament invades into the homologous duplex, forming a three-stranded loop, i.e. D-loop. While the D-loop migrates along the duplex, the homologous strand is exchanged. And finally, the D-loop resolved into two repaired DNA duplexes.

However, the mechanisms by which the RecA mediated homologous recombination are still not fully understood. Moreover, during the whole HR process including RecA filament formation, strand search and exchange, and D-loop migration, the RecA filament is complicatedly regulated by an expanding repertoire of accessory proteins and other environmental factors [10, 11]. For instance, the ssDNA generated *in vivo* is in fact firstly covered by ssDNA binding protein (SSB). Therefore, RecA has to replace the SSB from ssDNA in order to form presynaptic RecA filament, which is facilitated by RecOR proteins [10, 11]. On the other hand, some other regulatory proteins, such as RecX, act to destabilize the RecA filament, hence limit the elongation of RecA filament (More details on these major regulatory proteins are introduced in next section). Our understanding of the mechanisms by which the accessory proteins or environmental factors take to regulate the RecA filament, however, is still evolving. One of the main goals of this thesis is to cast light on the mechanisms employed by the accessory proteins and other environmental factors to regulate the RecA filament at a single ssDNA level.

1.2 Homologous recombination repair related proteins

A large family of proteins is known to be involved in the homologous recombination, while new members are still added to the regulatory protein family [10, 11]. Here, I will mainly introduce several major members of the regulatory proteins in the RecA-dependent bacterial homologous recombination.

1.2.1 RecA protein

RecA is the essential player of the bacterial homologous recombination repair. The biochemical properties of RecA have been extensively studied and reviewed previously [6, 10–18]. The *E. coli* RecA (EcRecA), as the prototype protein, is

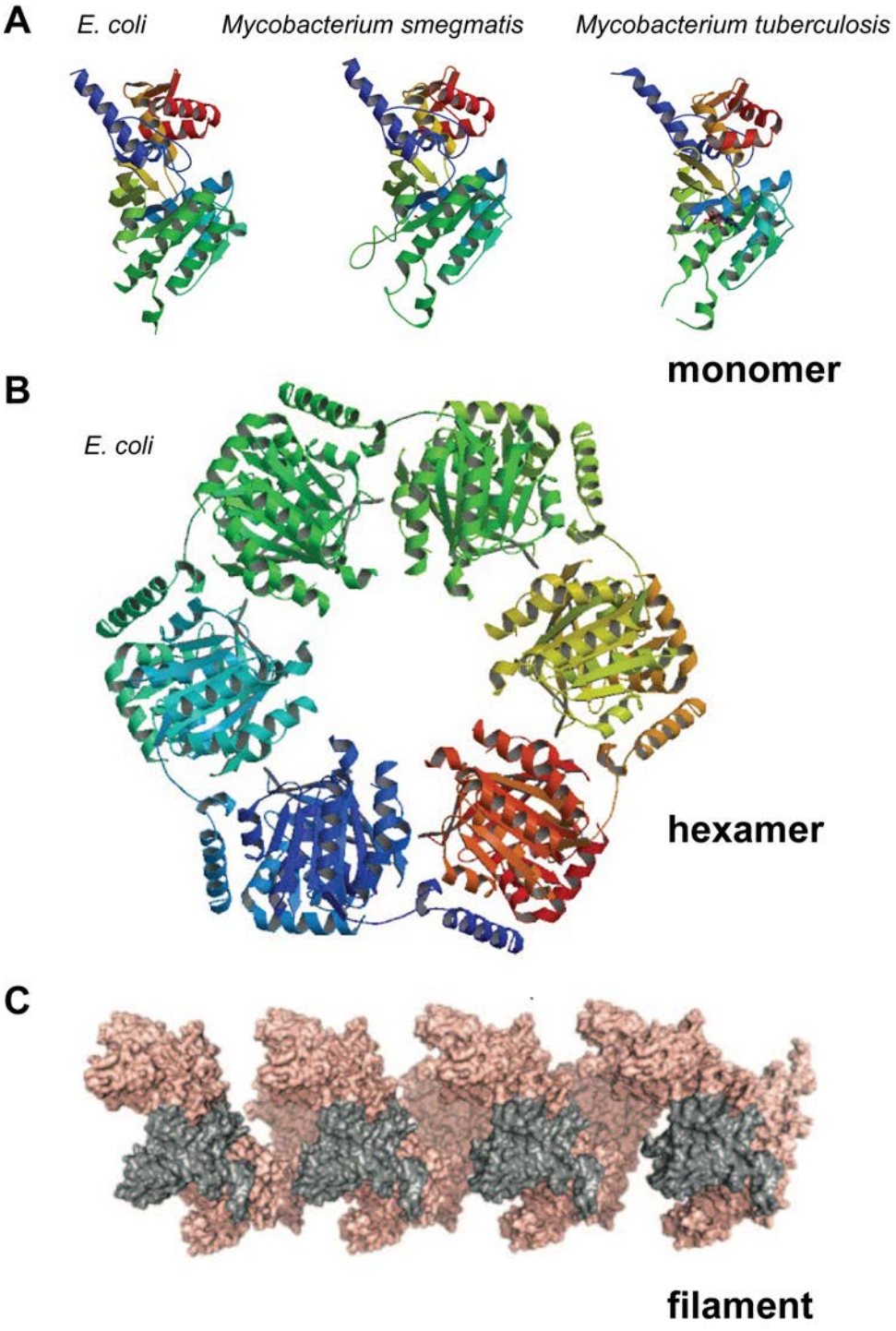

Figure 1.2: **Structures of Bacterial RecA**. (A). RecA monomer. (B). RecA monomer. (C). RecA filament. The figures are adapted from previous publication by Story *et al.* [12, 13]

a 352-residue polypeptide chain with a molecular weight of 37.8 kDa, with an intracellular concentration estimated in micro-molar range [10, 11]. The RecA monomer consists of a large core domain (a motif termed as RecA fold) and

two smaller domains located at the C and N termini [10–13] (Figure 1.2). RecA monomer binds to dsDNA or ssDNA with high positive cooperativity, in the presence of binding co-factors, such as ATP/ ADP / ATPγS as well as Mg^{2+} and Ca^{2+}, forming an extended nucleoprotein filament. Per turn (18 bp or nt) of the RecA filament includes $\sim$6 RecA monomers [10–13]. It is widely believed that the RecA-ssDNA filament polymerization is a unidirectional process from $5'$ to $3'$ on ssDNA, although there is a debate whether the $3'$ to $5'$ reverse polymerization exists [10, 11, 19].

To date, the assembly and disassembly of RecA filaments have been characterized intensively [10, 11, 19–31]. The growth rate of the filament ($5'$ to $3'$) is reported to be in the range of 120 to 1200 subunits min $^{-1}$ [10, 22–24, 30, 31], while the $5'$ dissociation rate is $\sim$70 monomers minute $^{-1}$ (on ssDNA) , and up to 120 monomers minute^{-1}(on dsDNA) [19, 22–24, 30]. The assembly and disassembly of the RecA filament are coupled with the ATP binding and ATP-hydrolysis, respectively, as well as influenced by environmental factors [10, 11, 20, 32]. Furthermore, the dynamics and stability of the RecA filament are also efficiently and tightly regulated by accessory proteins to avoid either insufficient or unlimited formation of the RecA filament [10, 11].

1.2.2 SSB

In order to access the genetic information for DNA replication and recombination, the dsDNA duplex has to be unwound to ssDNA intermediates during the actions [10, 11, 33]. However, such ssDNA intermediates are vulnerable to the nucleolytic and chemical attacks which results in ssDNA lesion or degradation. Therefore, to protect the ssDNA intermediates, a specialized class of SSB has been evolved and conserved in almost all kingdoms of life. Moreover, besides the protection role of SSB proteins, they also act as mediator for other proteins to access the ssDNA [10, 33]. Some prototypes of SSB proteins include the gene 32 protein of bacteriophage T4 [34], the *E. coli* SSB [35], and the eukaryotic equivalent heterotrimeric replication protein A (RPA) protein [36]. A featured DNA binding motif called oligonucleotide/oligosaccharide binding fold (OB fold) has been shown to be shared among all prokaryotic proteins by structural and

functional studies [10, 11, 33]. For example, one OB fold is located in a large amino-terminal domain of the *E. coli* SSB (18.8 kDa) protein [10, 11, 33].

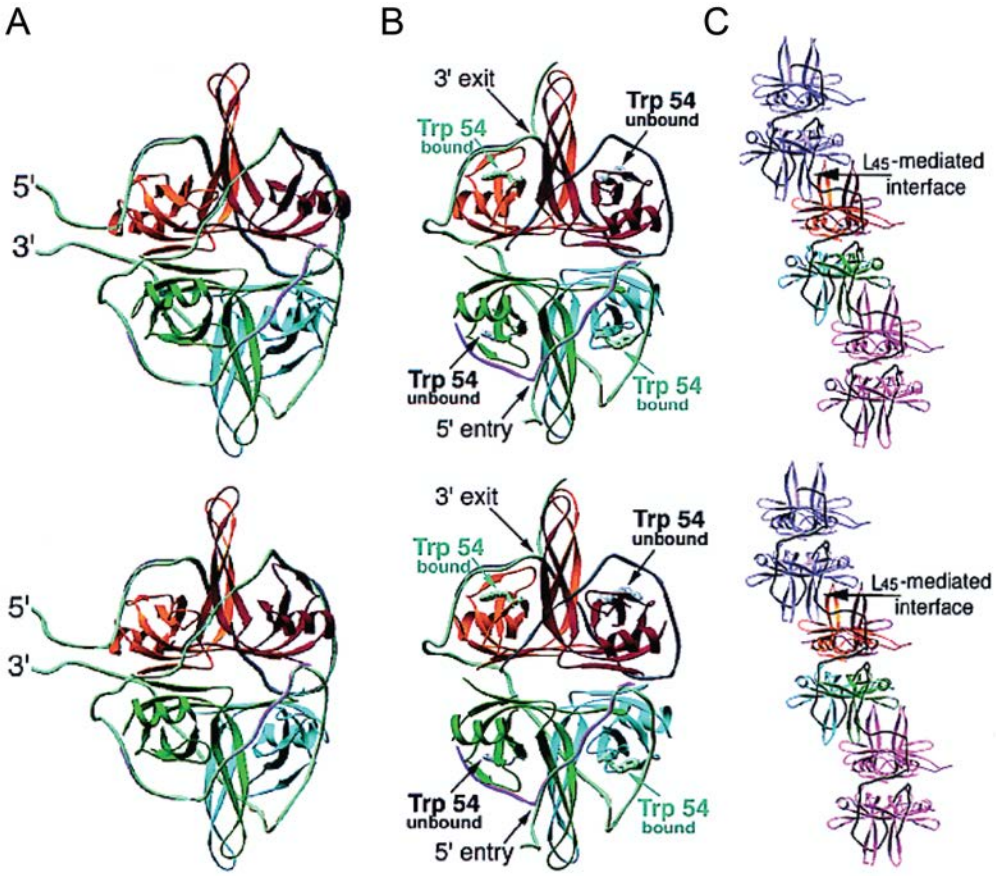

Figure 1.3: **Stereo ribbon diagram construction of the $(SSB)_{65}$ and $(SSB)_{35}$ binding modes**. Models of the $(SSB)_{65}$ (A) and $(SSB)_{35}$ (B) binding modes and the SSB binding to longer ssDNA in $(SSB)_{35}$ modes (C) (stereo ribbon diagram). The figures are adapted from previous publication by Srinivasan Raghunathan *et al.* [37].

SSB proteins bind to ssDNA nonspecifically with multiple modes and high affinity (K_d is usually in nano-molar range), depending on various experimental conditions, and possibly corresponding to various physiological functions [10, 11, 33, 37–41]. For instance, the current understanding of the ssDNA binding of EcSSB mainly involves three modes $(SSB)_{56}$, $(SSB)_{65}$ and $(SSB)_{35}$ where ∼56 / ∼65 or ∼ 35 nt of ssDNA is wrapped around SSB subunits [10, 11, 33, 38–41]. The selection of the SSB binding mode depends on many factors including monovalent salt, pH, magnesium, and protein concentration [10, 11, 33, 37–41]. The less wrapped $(SSB)_{35}$ mode is favored at low salt concentrations at high SSB binding density on ssDNA, while the most wrapped $(SSB)_{65}$ mode is preferred

at higher salt concentrations and lower SSB binding density [10, 11, 33, 37–41] (Figure 1.3). However, all these previously characterized binding modes were obtained for DNA in the absence of mechanical force. It can be anticipated that in the presence of force, the less or even non-wrapping mode would be favoured.

The SSB proteins play complicated and reversal roles in RecA activities. On the one hand, SSB binds to ssDNA with high affinity prior to RecA binding, forming a nucleoprotein array scaffold, which inhibits nucleation and polymerization of RecA filament [10, 11, 33]. This inhibitory effect can be relieved by a set of mediator proteins [10, 11, 33]. On the other hand, in certain conditions, it can act to remove the secondary structures of ssDNA, facilitating the polymerization of RecA along the ssDNA [10, 11, 33]. The regulatory mechanisms of RecA filament by SSB are influenced by multiple factors, including temperature, pH, its own concentration, and other regulatory proteins [10, 11, 33].

1.2.3 RecF, RecO, RecR proteins

As aforementioned, *in vivo*, the ssDNA intermediates produced by DSBs, SSGs or other damages are immediately coated by SSB, protecting the ssDNA from lesion or degradation, or self-association. However, this SSB-coated ssDNA nucleoprotein array in turn leads to barriers against the nucleation and polymerization of RecA onto ssDNA. To overcome this SSB array rampart, two major classes of proteins, including RecO, RecR, and RecF (termed as RecFOR pathway), or RecB, RecC, and RecD proteins (termed as RecBCD pathway) are evolved [10, 11]. These specialized classes of mediator proteins facilitate the loading of RecA onto the SSB-coated ssDNA and the formation of RecA filament. While the RecBCD pathway is mainly responsible for DSB, and RecFOR pathway is for SSG, the RecFOR pathway is also critical for the repair of DSBs and other damages for bacterial species or mutates lacking RecBCD proteins [10, 11].

The most intensively studied prototype of RecFOR mediating pathway is the *E. coli* RecO, RecR and RecF proteins. The 40.5 kDa RecF protein binds to ssDNA with a 1-to-15 nt ratio [10, 11, 42, 43]. It can also bind to dsDNA in the presence of ATP. ATP hydrolysis results in RecF dissociation from DNA [10, 44, 45]. Although RecF was proposed to direct RecA loading to the boundaries of

single strand gaps in duplex [46], it was also reported to reduce the stimulation provided by RecOR [45, 47], resulting in a negative effect on RecA loading in RecFOR pathway. It is possible that the RecF is involving other specific functions in the RecA-mediated HR [10, 47].

RecO protein (27.6 kDa) binds to both ssDNA and dsDNA [48–50], without reported binding or hydrolysis of ATP. It contains three domains: an OB fold, a helical bundle, and a zinc-finger motif moving from the N- to C-terminus [51, 52]. RecO promotes renaturation of complementary DNA strands in an ATP-independent manner [52]. This renaturation effect is enhanced by formation of RecO-SSB complex, and inhibited by formation of RecO-RecR complex [10, 47, 53]. Moreover, since only RecO, rather than RecF or RecR, interacts with SSB, it is believed to play an essential role in the replacement of SSB from ssDNA, making access for RecA loading. RecR protein (22 kDa) has no reported DNA binding or intrinsic enzymatic activities [54]. RecR protein binds to both Reck and RecO *in vitro* [10, 11, 47, 53]. Several lines of evidence have shown that the facilitating effect of RecA loading by RecOR involves the interaction between RecO and the acidic C-terminal region of SSB [10, 55]. Without C-terminal in the SSB-ΔC mutant, both RecA-loading and DNA annealing mediated by RecO/RecOR are inhibited [10, 55].

Up to date, a wealth of knowledge of these proteins has been provided by decades of biochemical, biophysical, and structural studies [10]. However, many aspects of the molecular mechanism(s) of RecFOR interactions with ssDNA, with SSB-coated ssDNA, as well as their positive effects on RecA loading onto ssDNA still remain unclear.

1.2.4 RecX protein

While RecFOR or RecBCD acts to facilitate the formation and/or stabilization of RecA filament, another class of proteins play their roles by inhibiting the formation of RecA filaments or de-stabilizing preformed filaments, since unlimited growth of RecA filament is also lethal *in vivo*. One well-known example of the inhibitory proteins is RecX ($\sim$19 KDa) [10, 56–59] (Figure 1.4). The *recX* gene is often found located on the same coding strand downstream of the *recA* or overlap

with *recA* in some eubacteria [10, 60, 61]. Over-expression of the *recA* in *recX* mutant leads to deadly effects in many eubacteria [56].

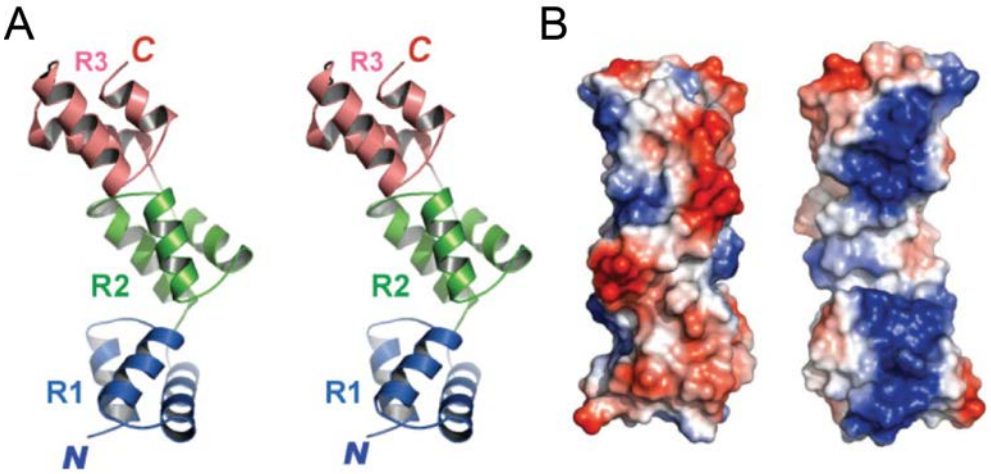

Figure 1.4: **Stereo ribbon diagram construction of RecX protein**. Ribbon draw (A) and Electrostatic plot (B) of RecX protein from *Xanthomonas campestris*. The three helix repeats are coloured in blue (R1), green (R2) and red (R3). The figures are adapted from previous publication by Yang *et al.* [59].

Biochemistry shows that at substoichiometric concentrations, EcRecX promotes net disassembly of EcRecA filaments on circular DNA [62, 63]. It has also been shown that both *E. coli.* and *M. tuberculosis* RecX inhibites RecA-promoted DNA strand exchange and ATP hydrolysis *in vitro* [57, 62, 63]. More recently, crystallographic analysis of RecX protein has shown that RecX is a modular protein with three repeated helix motifs, whose arrangement leads to an elongated and curved shape (Figure 1.4) [59, 64].

1.2.5 Other bacterial regulatory proteins

Besides aforementioned major contributors for the regulation of RecA filament and homologous recombination, there are other members in the regulatory family. For instances, the PsiB protein is believed to interact with RecA and block the formation of RecA filament; the DNA binding protein RdgC was reported to affect both functions of RecA and RecFOR; the UvrD helicase is known to be able to dismantle the RecA filament *in vitro* [10]. Moreover, the network of regulatory proteins is still expanding across different bacterial species.

1.2.6 Homologous recombination homologs in eukaryotic systems

While the eukaryotic cells differ from prokaryotic cells in many aspects, the DNA repair systems are highly conserved–most of the bacterial DNA repair related proteins could find their counter-partners in eukaryotic cells. For instance, the eukaryotic homolog of RecA is Rad51 and DMC1 proteins; the SSB in eukaryotic cell include replication protein A (RPA), RIM1 and so on. In this thesis work, I mainly used the bacterial proteins as model system; readers who are also interested in eukaryotic DNA repair systems are referred to previous researches and reviews [65–67].

1.3 Physiological relevance and functions of force

Mechanical force has been increasingly recognized as a critical physiological factor involved in multiple functions in diverse biological processes [68–70]. *In vivo*, forces can be generated in the ranges from sub-pN to tens of pN, even cumulatively several nN by various cellular machineries [68–70]. For example, the cytoskeletal protein myosin that mediates cell contraction can produce forces of several pN by one myosin [71] and cumulate to nN forces by concerted action of many myosin through stress fiber [68, 70]. In nucleus, individual DNA and RNA polymerases can exert up to ∼30 pN on DNA during actions [72, 73]. During the anaphase of cell division, the mitotic spindle can produce forces to the mitotic chromosomes up to ∼700 pN [74]. In bacteria, the nucleoid is known attached to the cell wall, which may lead to building up tension in DNA due to DNA compaction by nucleoid associated proteins. In general, forces in picoNewton range can be expected based on $\sim k_{\mathrm{B}}T$ interaction energy between proteins and DNA with nm scale of interaction distance. Due to the ubiquitous presence of forces on chromosomal DNA, force may potentially influence and even play crucial role during various regulatory actions in homologous recombinational DNA repair.

Recently, an *in vivo* dynamic imaging experiment has shown that during bacterial homologous recombinational repair of DSB, the two broken ends of DNA remained in close proximity while they were moving over a large distance

during the homologous search process [75]. Similar long-distance homologous searching is also observed in eukaryotic system [76]. These results indicate that the DNA ends are physically tethered during the whole process. As tension of a few pN is anticipated in the DNA and additional larger forces may be involved during this directed active DNA search process, the actions of the DNA damage proteins may be affected by the force.

However, the potential regulation role of force on homologous recombination DNA repair processes has been poorly studied. In this thesis work, I systematically investigated the effects of force on RecA filament regulated by major accessory proteins, including SSB, RecX, and RecOR. These studies reveal the potential universal importance of force in the RecA-dependent homologous recombination.

1.4 Micromechaniscs of double-stranded DNA&single-stranded DNA

The above-introduced DNA repair and recombination, as well as other essential nucleic processes, such as DNA replication, are mainly carried out by enzymes (e.g. RecA for HR), and diverse accessory proteins (e.g. SSB, RecOR for HR). These proteins usually bind to DNA, forming nucleoprotein complexes, or even motor along the DNA during action (for instance, DNA translocases). The cellular machineries should, therefore, be considered both biochemically and biomechanically. Hence, understanding of the micro-bio-mechanical properties of DNA is necessary. Decades of intensive studies, along with development of single-molecular methods, have depicted an insightful picture of the micromechanical properties of DNA, both theoretically and experimentally [77–82]. In this section, I will introduce basic information about the micromechanics of dsDNA and ssDNA, as well as the effects of protein-binding on their properties.

1.4.1 Micromechanics of double stranded DNA

In cells, the dsDNA at normal conditions (termed as B-DNA) is a regular, right-handed double helix, consisting of two chains of nucleic-acid polymers wound

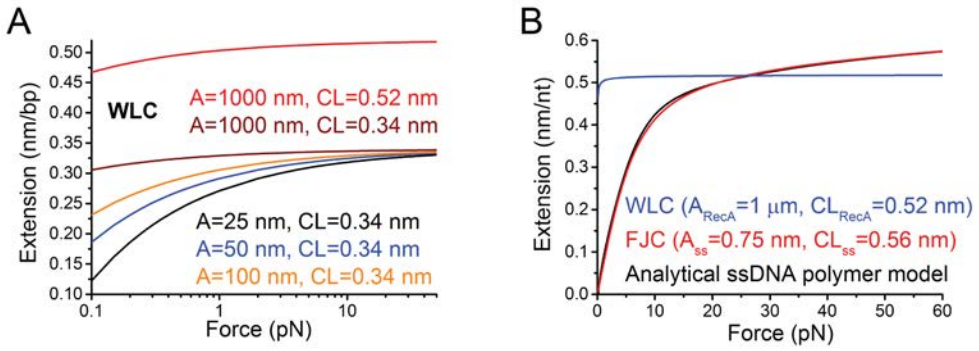

Figure 1.5: **Theoretical force responses of dsDNA and ssDNA as well as nucleoprotein complexes formed on them**. (A). Force extension curves of dsDNA per bp based on Eq. 1.1 with various persistence lengthes (A) and contour lengthes (CL) [78, 79]. (B). Force extension curves of ssDNA per nt based on Eq.1.3 (black) [82] or Eq.1.2(red) [81], as well as that of RecA-ssDNA nucleoprotein filament based on Eq.1.1

around each other. Each of the individual chain contains a series of nucleotides (nt) that are joined together by single covalent bonds. There are four types of nucleotides (adenine (A), thymine (T), guanine (G) and cytosine (C)) which are named after the different bases the nucleotides contain [83, 84]. The sugar-phosphate backbone of the nucleotide chain has a defined directed chemical structure, resulting in a defined direction of the base sequence along a single nucleotide chain (5′-to-3′)[83, 84]. The stability of the dsDNA is accurately regulated by environmental factors, such as, temperature, salt concentration and forces [79]. While the base-paired dsDNA ($> 10 - 30$ bp) is normally stable *in vivo* [79], a few $k_{\mathrm{B}}T$, which can be provided by enzymes proteins, is enough to gradually separate the strands.

dsDNA can be viewed as a semi-flexible polymer, the characteristic relationship between the DNA extension z_{ds}, and the corresponding applied force, f, has been demonstrated by single-molecular experiments [77]. An extensible worm-like-chain (WLC) polymer model [78, 79] fits well with the experimental data at forces <60 pN. The model is described by following Marko-Siggia formula [78, 79]:

$$z_{\mathrm{ds,WLC}}(f) = L(1 - \sqrt{\frac{k_{\mathrm{B}}T}{4Af}} + \frac{f}{f_{\mathrm{s}}}), \tag{1.1}$$

where, $L = N \times 0.34$ nm is the contour length of dsDNA of N base pairs, $A = 50$ nm is the persistence length of B-DNA, describing the bending rigidity

of dsDNA, and $f_s = 1400$ pN is the force constant describing the backbone stretching elasticity.

The value of A ∼50 nm means that an energy of ∼ 1 k_BT is able to bend the DNA of ∼50 nm for ∼1 radian. Despite certain local small inhomogeneity, this single persistence length description for the dsDNA flexibility has been proved extremely useful [78, 79]. As a semi-flexible polymer, dsDNA respondes to stretching by mechanical forces [78, 79] (Figure 1.2 Force responses of dsDNA experiments and theoretical calculation). For forces $< \frac{k_BT}{A} \approx 0.1$ pN, the DNA is coiled due to thermally excited bends which counteract the stretching forces. At this force range, the extension of DNA, the average of end-to-end distance along the force direction, is less than 50% of its B-form contour length. When forces increases from 0.1 pN to 10 pN, the extension increases non-linearly from 50% to 90% of the contour length, due to the suppression of the thermal bending fluctuation by the free energy contributed by stretching forces. When stretching forces continue to increase, deformation of the dsDNA helix takes place with a linear elongation of extension, until the force reaches ∼60 pN.

However, at force slightly above 60 pN, a few pN force increase results in a dramatic and sharp transition with an extension elongation of ∼1.6-1.7 times of B-form DNA contour length. This overstretching transition was originally discovered in experiments using torsion-unconstrained, end-open DNA decades ago [81, 85]. Since then, the nature of this mysterious overstretching transition has been debated and only recently solved [81, 82, 85–97]. Three transitions may involve: i) 'peeling' transition to one peeled ssDNA strand under tension while the other ssDNA strand coils; ii) 'inside-strand separation' transition to two parallel ssDNA strands that share tension (melting bubbles), and ii) 'B-to-S' transition to a novel overstretched base-paired dsDNA, termed as S-DNA [92–97]. The selection between each transition depends on DNA sequence, DNA topology, and environmental factors such as temperature and salt concentration [92–97].

1.4.2 Micromechanics of single stranded DNA

ssDNA intermediates provide gene access for DNA replication, acting as templates for DNA recombination and repair. The unwanted ssDNA must be re-

moved or repaired to maintain the genome stability. As a single nucleic-acid polymer chain, ssDNA, unlike helical structured dsDNA, is a highly flexible polymer with an estimated persistence length of $\sim$ 1 nm [81, 82, 98–100], thereby, a higher force is required to extend the coiled chain. Furthermore, the exposed bases of ssDNA tend to stick to other bases when coiled at low forces, forming secondary structures within the ssDNA strand [82, 98–100].

The force response of ssDNA are much more complicated and less studied. The force extension curve of ssDNA at force <10 pN is highly dependent on salt species and salt concentration. At high salt concentration, the ssDNA at low forces $<$ 10 pN tends to be condensed, which depends on divalent salt concentration more sensitively than on the mono-valent concentration [100]. The force-extension curve of ssDNA in mono-valent salt conditions ($<$ 150 mM NaCl) can be fine fitted by freely joint chain (FJC) model [81]:

$$z_{\rm ss}(f) = Nb_{\rm ss}(\coth(\frac{2Af}{k_{\rm B}T}) - \frac{k_{\rm B}T}{2Af})(1 + \frac{f}{800}) \tag{1.2}$$

where $b_{\rm ss} = 0.56$ nm is the contour length of ssDNA per nt, the $A = 0.75$ nm is the persistence of ssDNA. However, the above equation cannot describe the salt effect on ssDNA. A phenomenological polymer formula[82] has been obtained to describe the monovalent salt effects on DNA:

$$z_{\rm ss}(f) = Nh\frac{a_1 \ln(f/f_1)}{1 + a_3 \exp(-f/f_2 - a_2 - f/f_3)} \tag{1.3}$$

where $h = 0.34$ nm, $a_1 = 0.21$, $a_2 = 0.34$, $f_1 = 0.0037$ pN, $f_2 = 2.9$ pN, and $f_3 = 8000$ pN. The parameter $a_3 = 2.1\ln\frac{(I/0.0025)}{\ln(0.15/0.0025)} - 0.1$ is dependent on the ionic strength I which is the molar concentration of monovalent salt. However, up to date, there is no model fitting well with the force response of ssDNA in the presence of divalent cations, especially at low force ranges ($<$ 10 pN) due to the formation of complex secondary structures.

1.4.3 Effects of DNA-distortion proteins on the force response of dsDNA/ssDNA

DNA binding proteins often distort DNA upon interaction with DNA, resulting in DNA local conformational changes, hence, altering the force responses of DNA by certain level (Figure 1.5). The effects of DNA-distortion proteins on force response of DNA have been calculated theoretically [101–103], and demonstrated experimentally with a great many of proteins [104–113]. Conversely, force can affect the DNA-protein interactions.

Generally, the resulting protein-DNA complex may increase the local rigidity of the DNA (termed as stiffening effect), leading to the elongation of the DNA extension at low force range ($\sim<$ 10 pN) compared to naked DNA, or decrease the local rigidity of DNA (termed as bending/folding effect), resulting in the shortening of the DNA extension at low force range. Force in principle should favour binding of DNA-stiffening proteins while disfavour the DNA-bending/folding proteins.

Since different proteins may have different effects on the force response of DNA, in principle, different protein-DNA binding can be distinguished based on analysis of differential effects on forces responses of DNA. Furthermore, the kinetics and dynamics of protein-DNA interactions can also be revealed by analysing the evolution of DNA extension time trace under forces at a single DNA level.

1.5 Objectives and organization of the thesis

In this thesis, I mainly focus on understanding the mechanisms of the key accessory proteins, including SSB, RecX, RecO, and RecR on regulation of dynamics and stability of RecA filament. Furthermore, I investigate the potential regulatory role of mechanical force on RecA filament in the presence of these accessory proteins, which reveals a potential broad importance of mechanical force during homologous recombinational repair process.

In Chapters 2&3, I describe a new platform for studies of ssDNA-processing proteins on single ssDNA template manipulated by magnetic tweezers. I first detail the single-molecule magnetic tweezers, including the basic apparatus, force

generation, control, calibration and extension measurement methods as well as sample preparations. In addition, I also describe a new disturbance-free rapid solution exchange method, which is critical for studies of dynamics of protein-DNA initial interactions.

In Chapters 4-6, as the main focus of this thesis, I present dynamics and stability of individual RecA-ssDNA nucleoprotein filaments regulated by SSB (Chapter 4), RecX (Chapter 5) and RecO, RecR (Chapter 6), as well as mechanical force. In Chapter 7, I explain the regulatory role of force based on force-dependent free energy cost and binding affinity of RecA to ssDNA, which explains the regulating role of force on these protein accessory dynamics and stability of RecA nucleoprotein filament. , I also present the dynamics and regulation of formation and stability of RecA nucleoprotein filament formed on dsDNA by various environmental factors in Chapter 8.

In addition to studies of dynamics and regulation of RecA nucleoprotein filament formed on both ssDNA and dsDNA, I also present my studies of how mechanical force regulates DNA bending in a single specific protein-DNA (IHF-H$'$) complex (Chapter 9) and the dynamics of CRISPR RNA-guided DNA bending and unwinding (Chapter 10). Finally, I briefly summarize and discuss the work in the thesis in Discussion section (Chapter 11).

Chapter 2

Methods and materials–A platform for studies of ssDNA-processing proteins on individual ssDNA templates

2.1 Chapter Summary

The rapidly developing single-molecule manipulation technologies have enabled studies of molecular interactions at a single-molecule level with nanometer resolution in real time. In this chapter, I introduce the single-ssDNA-manipulation platform to study the dynamic actions of ssDNA processing proteins on individual ssDNA template using magnetic tweezers. This platform consists of several key steps, including 1) generating short single ssDNA tether for detecting the extension with high signal-to-noise ratio from a dsDNA tether, 2) accurate measurement of the force-extension curve of the ssDNA, and 3) anti-drifting technique for stable long time measurement. Accordingly, in this chapter, I first introduced the basics of single molecule manipulation using magnetic tweezers, including force generation, calibration, ssDNA generation and extension measurement with anti-drift control, followed by flow channel, DNA and protein sample preparation. Other experimental and analysis methods used in some of the studies will

be included in the corresponding chapters.

2.2 Single molecule manipulation using magnetic tweezers

As aforementioned, the mechanical properties of DNA (and other molecules) are not only important themselves, but as useful and quantitative approach to study the interactions and competitive regulations of various proteins involved with DNA replication, recombination, repair and organization. Moreover, mechanical force itself is also a proven or potentially critical regulatory factor in these biological processes [68–70, 109, 111–114]. Therefore, the ability to effectively and quantitatively manipulate the DNA at single molecular level is crucial for the advance of our understanding of the bio-mechanical nature of various cellular actions. Thanks to scientists′ decades of efforts, several state-of-art single molecular manipulation techniques have been developed, including atomic force microscopy (AFM), optical tweezers, magnetic tweezers, microfluidics based manipulation, micro-needle manipulation and so on [115]. Among them, the magnetic tweezers are mainly used in this thesis, which is especially useful for short molecule tether manipulation. This chapter will mainly introduce the basic physical principles of magnetic tweezers, while readers may find details about other single-molecular manipulation techniques in recent review [115].

2.2.1 Development of magnetic tweezers

The first demonstration of magnetic tweezers was performed by Crick & Hughes at 1950s to control magnetic particles in cytoplasm of cells [116]. Later, the first single-molecule magnetic tweezers experiment was performed by Smith *et al* at 1990s [77], In this work, they manipulated individual DNA molecules tethered between a microscopic particle and coverslip surface in a flow channel. In 1998, a highly efficient perpendicular design of magnetic tweezers that is suitable for high resolution single-molecule studies was published by Strick *et al.* [117]. Since then, this straightforward method has been extensively used to study the bio-mechanical and bio-chemical properties of DNA/RNA or proteins, as well as

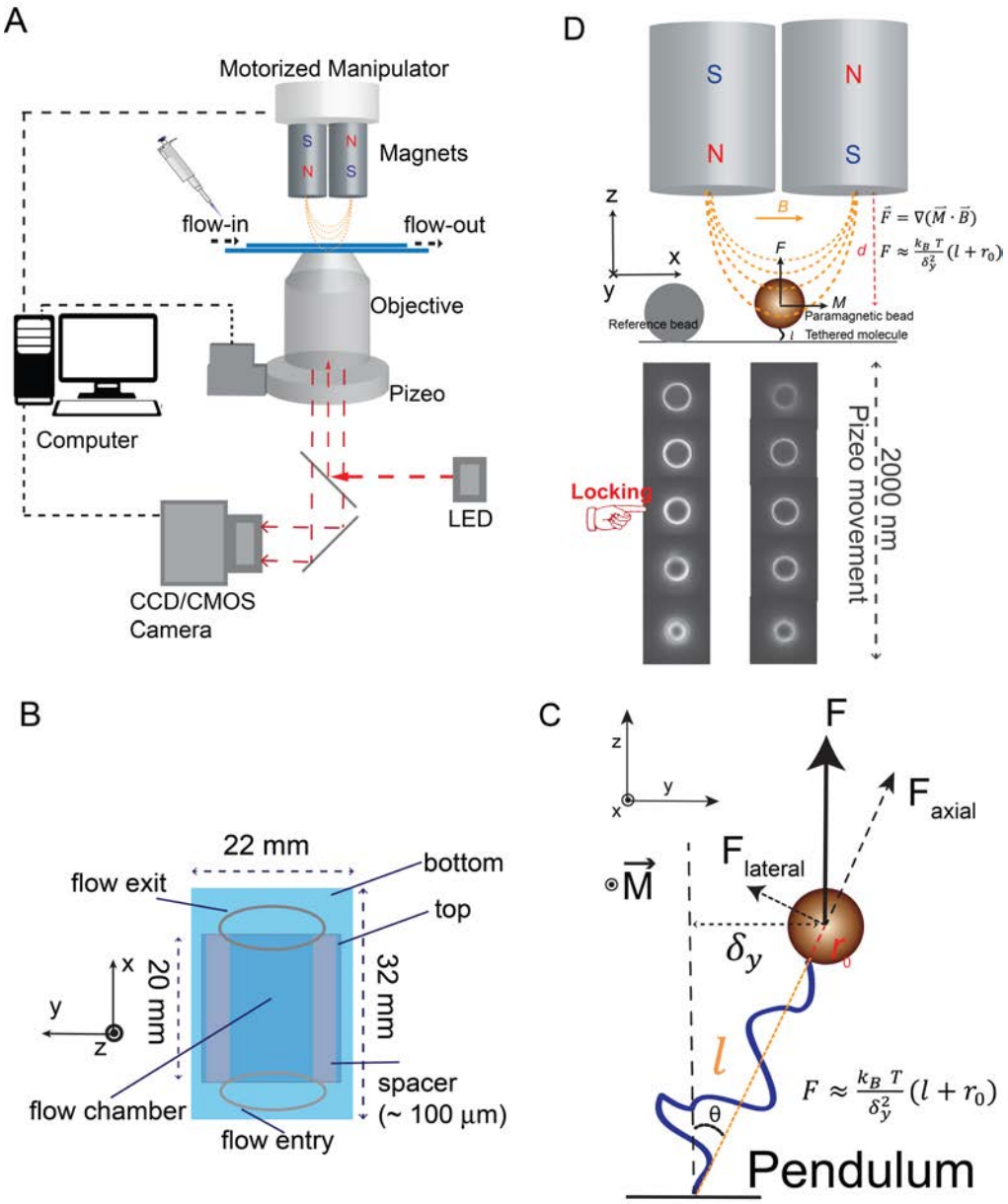

Figure 2.1: **Vertical magnetic tweezers setup**. (A). Basic apparatus of vertical magnetic tweezers. (B). Sketch of a flow channel for vertical magnetic tweezers. (C). Force calibration based on bead fluctuations. (D). Force generation. (E). Bead position determination and anti-drift method.

DNA-protein, protein-protein interactions by manipulating (stretching/twisting) individual bimolecular tethers [29, 31, 118–121]. Besides the perpendicular design, a transverse design was also developed, which applies forces in the focal plane and the extension is determined by the centroid of bead [122].

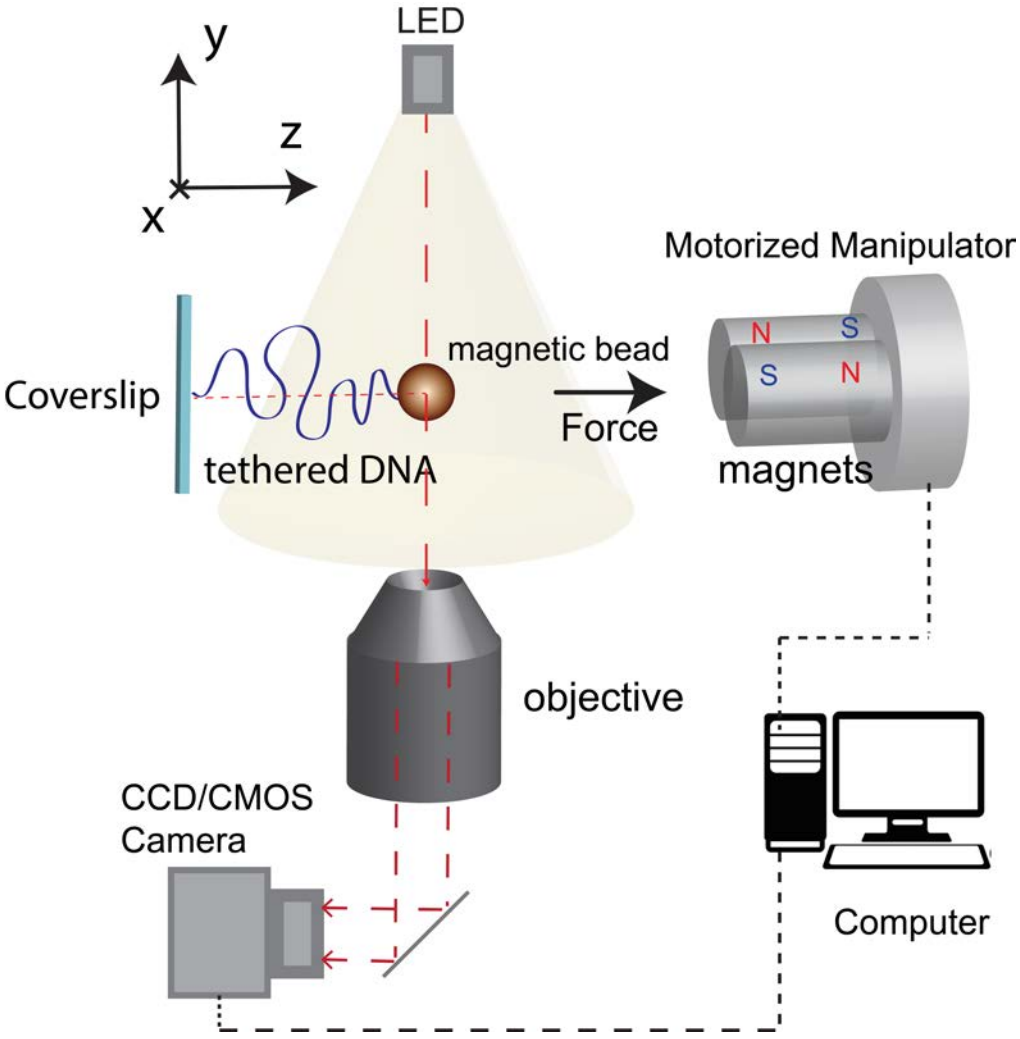

Figure 2.2: **Transverse magnetic tweezers setup**.

2.2.2 Magnetic tweezers apparatus

A basic single-molecule magnetic tweezers apparatus consists of a pair of magnets and magnetic micro-particles (Figure 2.1A&2.2). The magnetic particles, usually, paramagnetic bead (companies) is tethered with one end of single biomolecules whose another end is attached to coverslip surface. The position of the tethered bead is recorded and determined by a computer-controlled microscopic objective with a camera. The force on the paramagnetic bead is tuned by controlling the pair of magnets.

The two designs have their respective strengths. In the vertical design, the tethers are formed on large 2-d coverslip surface, suitable for high-throughput multiplexing experiments [123, 124]. The length of tethers can be shorter than 200 nm, ideal for high signal-to-noise measurements [109, 111, 112, 120]. The

tweezers can be built on a total internal reflection fluorescence microscope, allowing combination with single-molecule spectroscopy technologies such as single-molecule Foster Resonance Energy Transfer (smFRET) [125]. In the transverse design, tethers can be as long as the dimension of the whole view area, ideal of studies of large DNA condensation by proteins [105, 126]. The tethers are stretched in the focal plane, allowing direct observation of fluorescence labeled proteins on DNA. The position of the bead can be determined with nanometer accuracy with long working distance non-contact objective, making it possible to control the temperature of the sample independently from the microscope, convenient for temperature dependent studies [110, 126]. Two different types of in-house built magnetic tweezers: vertical magnetic tweezers and transverse magnetic tweezers were used in the studies [120, 122]. The basic setups of the two types were illustrated in Figure 2.1A&2.2.

In a flow channel for vertical magnetic tweezers, one end of the end-labeled DNA (or other molecules of interest) can be specifically tethered on the specially functionized bottom coverslip surface, while another end of the DNA is attached to the micro-magnetic bead (Figure 2.1A&B). A pair of permanent magnets is placed along the z-direction with its geometric center aligned with the tether, generating a magnetic field gradient (z-direction) perpendicular to the coverslip surface (x-y plane). Hence, the magnets apply a force on the tethered micro-magnetic bead in the direction towards the magnets. The magnitude of the force experienced by the magnetic bead is tuned by controlling the distance between the coverslip surface and the magnets. The illumination from LED light source is placed under the channel in a z-direction through an underneath microscope objective which also record the images of the magnetic bead (as well as the reference bead stuck on the coverslip surface) on its focal plane. The accurate position of microscope objective is controlled by a pizeo, which acts to antagonize the drift of the setup (details can be found in [120]). A CCD/CMOS (short for Charge coupled device/complementary metal oxide semiconductor) camera was used to collect the images obtained and transfer to a computer for further imaging analysis.

In contrast, the design for a transverse magnetic tweezers setup is different.

The end-labeled DNA is tethered between a magnetic micro-bead and the functionized edge of the coverslip (rather than the surface), the pair of magnets is placed in the x-y plane and aligned along the x-direction. Therefore, the force generated by the magnets on the tether is in the x-direction towards the magnets. The illumination is from the top of the channel in z-direction and the images of the moving bead are collected by the bottom microscopy objective.

2.2.3 Force generation

The pair of magnets is placed along the direction (z-direction as in Figure 2.1C&D) perpendicular to focal plane (x-y- plane), with its geometric center aligned with the tether, resulting in a magnetic field ,$\vec{B}$, parallel to focal plane (alone x-direction as in Fig. 1A), and a magnetic field gradient ,$\frac{\mathrm{d}\vec{B}}{\mathrm{d}z}$, perpendicular to the focal plane (i.e., z-direction). Therefore, at a given position in the magnetic field, the tethered micro-magnetic bead is magnetised with a total magnetic dipole moment, $\vec{M}$, along the same direction as $\vec{B}$, resulting a force $F = \bigtriangledown(\vec{M} \cdot \vec{B})$, along the direction towards the magnets (z-direction). In normal magnetic tweezers setup, due to the inherent properties of superparamagnetic bead and large magnetic fields, the magnetisation of the bead reaches a saturated value, M_{max}, which is magnetic field-independent. Hence, the bead experiences a force, $F = M_{\mathrm{max}} \frac{\mathrm{d}B}{\mathrm{d}z}$.

2.2.4 Force control

For a given magnetic bead, i.e., a given M_{max}, the magnitude of the force depends only on $\frac{\mathrm{d}B}{\mathrm{d}z}$, which only dependent on the distance (d) between the focal plane and the magnets, therefore, force control (such as constant force control and force loading-rate control) can be easily achieved in magnetic tweezers by controlling $d(t)$ through a computer-controlled, motorized, position-controlling manipulator. In addition, due to a negligible stiffness of $\sim 10^{-6}$ pN nm^{-1}, force on the tethered bead can be considered as constant even when the molecule moves over μm distance.

2.2.5 Force calibration

For single-protein manipulation experiments, the tethered molecules are usually short ($< \mu$m), the force calibration is accomplished by bead fluctuation analysis at forces < 15 pN, and than extrapolation to larger force ranges up to 200 pN [120].

The basic principle of bead fluctuation analysis based force calibration can be understood as a pendulum model (Figure 2.1C&D). In the schematic representation figure, The movement of the micro bead in the magnetic field is harmonic oscillation. Hence, we have $\alpha = \frac{F_{\text{lateral}}}{(l+r_0)\tan\theta} = \frac{F\sin\theta}{(l+r_0)\tan\theta}$, where α is the stiffness in the lateral direction, r_0 is the radius of the micro bead and the l is the end-to-end distance of the protein (or other tethered polymer). θ in a pendulum is always very small ($\ll l_0$), hence, $\frac{\sin\theta}{\tan\theta} \approx 1$. Therefore, $\alpha \approx \frac{F}{l+r_0}$.

Furthermore, according to the equipartition relation, we have $\frac{1}{2}\alpha\delta_y^2 = \frac{1}{2}k_{\text{B}}T$, where δ_y^2 is the y-directional fluctuation of the micro-bead, the k_{B} is the Boltzmann constant, and T is the temperature in Kelvin scale. Substituting $\alpha \approx \frac{F}{l+r_0}$ into the above equation, we have $F \approx \frac{k_{\text{B}}T}{\delta_y^2}(l + r_0)$. Therefore, by measuring the δ_y^2 and l, the force F can be obtained.

The reason why the above bead fluctuation analysis can only calibrated forces < 15 pN for short protein tethers is explained as follows. To accurately measure δ_y^2 in above equation, the sampling rate, f_{sampling}, has to be faster than the Lorentzian corner frequency, $f_{\text{c}} = \frac{F}{2\pi\gamma(l+r_0)}$. The drag coefficient of the bead $\gamma = 6\pi\eta r_0$, where η is the solution visocity and r_0 is the radius of the spherical bead. The sampling rate in normal setup is in the range ~ 100 Hz. Thereby, a measurable force is up to ~ 100 pN for a 10 μm tether with a 1.4 μm-diameter bead, while the measurable force is less than 15 pN for a 100 nm tether. Therefore, this force calibration method is not able to accurately measure the higher forces for short tethers in the range of hundreds of nm.

The larger force range can be obtained by an extrapolation method detailed in (cites), which utilizes a property that force-distance curves $F(d)$ (typically are exponential decay functions) between any two paramagnetic beads differ only by a constant on logarithm scale (i.e., $\Delta\text{Log}F_{1,2}(d) = \delta$). Therefore, we used long

DNA (48,502 bp λDNA) molecules, which have a $f_c < 50$ Hz at up to 100 pN, to obtain a standard curve $F^*(d)$. In experiments with short tether, we determine the shift δ from the standard curve for each bead at forces < 15 pN through $\mathrm{Log}F(d) = \mathrm{Log}F^*(d) + \delta$. Then, extrapolation to larger forces up to 100 pN can be done. This method is valid for any tethers such as DNA or proteins. Note that, the relative error of forces determined by such extrapolation method is around 10%, mainly caused by the uncertainty in the bead radius [120].

2.2.6 Torque control

In addition, due to $\vec{M}$ is perpendicular to $\vec{F}$ in normal magnetic tweezers setup, an external large torque on the bead-tethered molecule can be easily generated by rotating the pair of magnets. Furthermore, by adjusting the geometry of the magnets, recently developed free-orbiting tweezers, where the $\vec{M}$ is parallel to $\vec{F}$, enable the tethered bead freely rotate in the focal plane.

2.2.7 Scattered vs back-scattered illumination

The tethered bead is usually illuminated by collimated LED (light emitting diode). There are usually two illumination methods: 1) scattered-illumination: the bead is illuminated by light through the gap between the magnets. Well-defined patterns of the concentric rings of he bead are generated due to the interference between light scattered off the bead and unscattered light. 2) backscattered-illumination: the bead is illuminated by the light from the objective; interference between the backscattered light and the reflection light of the top-surface generates a another set of well-defined patterns. These patterned images then are captured on a CCD camera through the objective for further bead position determination. Since the light for back-scattered illumination is from the objective, the gap between the magnets can be much smaller; therefore a much higher force can be generated compared to scattered-illumination. On the other hand, the back-scattered illumination method can only apply for tethers < 1 μm, while the scattered-illumination is suitable for tethers in large length range (e.g. tens of μm).

2.2.8 Bead position measurement

The 3 dimension (x-, y-, z-) position of the bead is determined based on the interference-patterned images: the lateral (x-, y-) position of the bead is determined based on the centroid of the patterned images, while the axial (z-) position of the bead is determined based on patterned images' intensity distribution that is strongly corresponds to the axial (z-) position of the bead to the objective focus. Therefore, an initial library of the patterned interference images of the bead at a series of at defocus planes along the axial direction is necessary for further position determination. This is usually accomplished by recording a set of images of bead' patterned interference during moving the objective along the axial direction using a piezo actuator at the beginning of experiments. During experiments, the real-time, 3 dimensional motions of the bead are determined by comparing the current diffraction pattern with the initial calibration library through correlation function analysis of the power spectrum of Fourier transform of the bead images [120, 127].

2.2.9 Anti-draft method

However, there is always drift during experiments which would largely affect the position determination. To reduce the drift artefacts, in magnetic tweezers experiments, a stuck polybead on the bottom surface is used as a reference bead (Figure 2.1D). By simultaneously tracking ('locking') the reference bead on the surface using an objective pizeo actuator through a low frequency feedback system such that the stuck bead image has the best correlation with a specific image store in the library, the tethered bead therefore is correspondingly tracked with much reduced drift artefacts. Usually, The combination of this active focal plane locking mechanism and the use of the stuck bead as the reference bead for height determination of the moving bead enable highly efficient minimization of mechanical and thermal drift of the setup for long time scale experiments (over hours) with a spatial resolution of $\sim$2 nm, and temporal resolution of $\sim$100 Hz [120].

2.2.10 Molecule extension measurement

The above section has reviewed the bead position determination methods, however, the axial (z-) position determined is in fact the height of the tethered bead in response of force, $H_{\text{bead}}(f)$, rather than the actual molecule extension, $z(f)$. This is because that at given d, in addition to force generated by the magnetic field, there is a torque generated, which would align the bead along the magnetic field. A change in d, therefore change both the force and alignment of the bead due to the torque, which causes rotation of the bead. For short tethers (short DNA or proteins), bead rotations also change the height of bead that cannot be differentiated from real extension change of the tethered molecule. The contribution from bead rotation has to be eliminated in order to obtain real extension change of the molecule.

One possible method is based on a property of the paramagnetic bead that the bead adopts a unique orientation (alignment) that corresponds to a unique height of bead at a given force. Therefore, at a given force, the extension differences of a molecule tether before and after interactions or conformational changes are accurate. Utilizing this property, the force extension curves of a tethered protein or DNA bound with proteins can be measured [109]. For instance, in order to measure the force response of a short single-stranded (ss) DNA, one can first measure the extension differences between the double-stranded (ds) DNA tether and the ssDNA tether generated by this dsDNA at the same set of forces, thereby obtain the $\Delta z_{\text{ss-ds}}(f)$; and then obtain the accurate force response of ssDNA, $z_{\text{ss}}(f) = z_{\text{WLC}}(f) + \Delta z_{\text{ss-ds}}(f)$, by adding back the theoretical dsDNA force-extension curve based on worm-like-chain polymer model. Similarly, in order to measure the force response of a short protein peptide-chain, one can measure measure the extension differences between the folded protein tether and unfolded peptide-chain at the same set of forces, thereby obtain the $\Delta z_{\text{unfold-fold}}(f)$; and then obtain the accurate force response of the chain, $z_{\text{chain}}(f) = z_{\text{fold\_rigidy}}(f) + \Delta z_{\text{unfold-fold}}(f)$ Furthermore, when proteins bind to the tethered molecule, the force response can be similarly obtained by further adding the protein binding-induced extension differences at corresponding force.

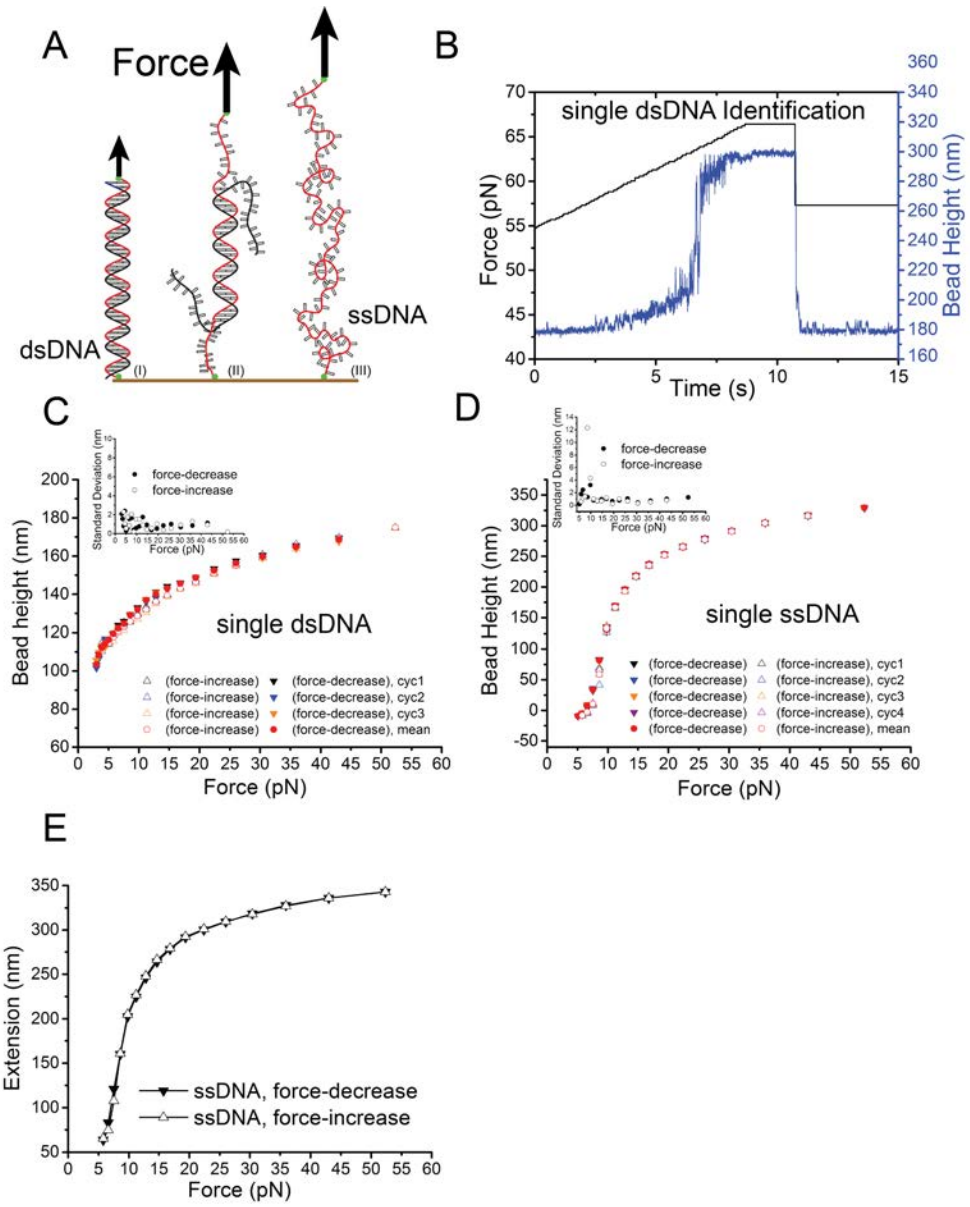

Figure 2.3: **ssDNA generation and force-extension measurement**.

2.3 Flow channel preparation

The flow channel is made with two pieces of coverslips sticking together by parafilms/double-side tape (Figure 2.1B). The bottom piece of coverslips was sinalized with (3-Aminopropyl) triethoxysilane (APTES), and then coated with sulfo-SMCC for Thiol-labeled DNA, or streptavidin for biotin-labeled DNA. During the preparation, 3 μm polybeads are coated on bottom surface as reference beads for experiments. A detailed protocol for channel preparation is included as Appendix A .

2.4 DNA and proteins synthesis

Labelled-DNA synthesis can be divided into three steps: 1.) DNA fragment PCR–a 576 bp DNA construct (as an example) can be generated by PCR from bacteriophage λ-DNA with a pair of forward and reverse primers: 5′[Thiol]ATTATACTCGAGAGCATAAGCAGCGCAACA3′ and 5′ATTATAAG CTTATGACGCAGGCATTATGCT3′ (HindIII cutting site). T_m for PCR this DNA fragment is 55-58 ^{o}C. 2.) HindIII digestion –The DNA is then digested by restriction enzyme Hind III for ∼2 hours, and then purified. 3.) Biotin labelling –The purified DNA is then incubated with 0.1-1 mM Biotin-16-dUTP, 1 mM dATP, 1 mM dGTP, 1 mM dCTP, and 3 μl Vent DNA polymerase in 100 μl reaction volume for $\sim$ 1.5 hours at 65-72 ^{o}C. The DNA product is therefore labeled with thiol at one end, and biotin at another end of the same strand of the DNA [109]. Synthesis of the proteins studied in this thesis is performed following previous protocols indicated in the corresponding chapters.

Chapter 3

A disturbance-free rapid solution exchange method for single molecule manipulation experiments

3.1 Chapter Summary

Single-molecule manipulation technologies have been extensively applied to studies of the structures and interactions of DNA and proteins. An important aspect of such studies is to obtain the dynamics of interactions; however the initial binding is often difficult to be obtained due to large mechanical perturbation during solution introduction. Here, I report a simple disturbance-free rapid solution exchanging method for magnetic tweezers single-molecule manipulation experiments, which is achieved by tethering the molecules inside microwells (typical dimensions - diameter (D): 40-50 μm, height (H): 100 μm; H:D$\sim$2:1). Our simulations and experiments show that the flow speed can be reduced by several orders of magnitude near the bottom of the microwells from that in the flow chamber, effectively eliminating the flow disturbance to molecules tethered in the microwells. We demonstrate a wide scope of applications of this method by measuring the force dependent DNA structural transitions in response to solution

condition change, and polymerization dynamics of RecA on ssDNA /SSB-coated ssDNA/dsDNA of various tether lengths under constant forces *.

3.2 Motivation

Most of previous studies using optical tweezers and magnetic tweezers technologies were conducted after solution flow to avoid mechanical perturbation that impairs high resolution single-molecule measurements [115]. As a result, data during solution exchange often could not be recorded, resulting in loss of important information regarding the initial conformation and binding dynamics in response to solution condition changes.

Several methods have been developed to probe the initial binding for optical tweezers by quickly moving the trapped bead and its attached molecule from one solution condition to another in laminar flow, and the molecular binding is monitored by fluorescence imaging [128, 129]. More recently, additional method has been developed to avoid the influence from the flow drag in the laminar flow by further moving the bead attached molecule from the laminar flow into a flow-free harbor which is connected to the laminar flow through a thin neck [129, 130]. These rapid solution exchange approaches have enabled optical tweezers to probe initial binding dynamics of molecules in response to solution condition changes.

However, the above bead-moving and laminar flow based rapid solution exchange methods cannot be applied to typical magnetic tweezers experiments, as magnetic tweezers apply forces to a paramagnetic bead attached to molecules immobilized on coverslips surfaces. Increasing the solution exchange speed would not solve this problem either, as it would result in large drag force to the bead that can cause large conformational perturbation to the molecule and cause tether breakage. The scope of the applications of magnetic tweezers would be greatly broadened if a rapid disturbance-free solution exchange method could be developed.

In this section, we report a simple method to achieve rapid disturbance-free

*Note that main contents detailed in this chapter are included in *Disturbance-free rapid solution exchange for magnetic tweezers single-molecule studies.* Le S., Yao M. *et al.* Nucleic Acids Reserach, online (2015)

solution exchange for magnetic tweezers by tethering molecules at the bottom of microwells (diameter: 20-60 μm; height: 40-100 μm) in a microwell array. Our results show that complete solution exchange can take place in the order of seconds, making it possible to obtain early binding dynamics upon solution exchange at constant forces.

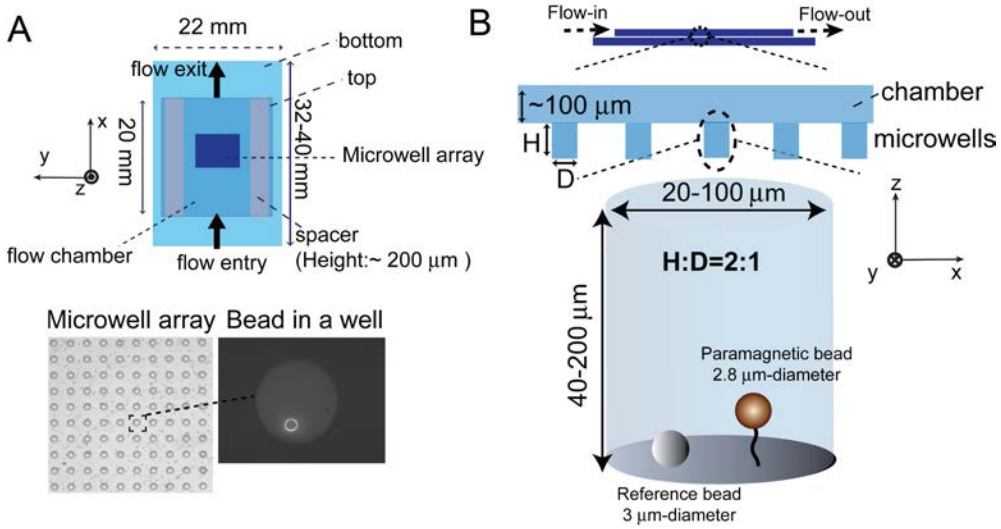

Figure 3.1: **Microwell array and flow channel design**. (A) Top panel: Schematic of flow channel with a middle area covered by a thin microwell array. z- denotes the force direction, x-y plane denotes the focal plane; Bottom panel: A part of microwell array (D: 40 μm; H: 100 μm) covered area imaged using a 20× objective (left panel) and a DNA tethered bead at the bottom of a microwell imaged using a 100× objective (right panel). (B). Sketches of the cross-section of the flow channel (top panel) and a tethered molecule at the bottom of a microwell (bottom panel). A reference bead stuck to the coverslip surface used to eliminate spatial drift is also shown.

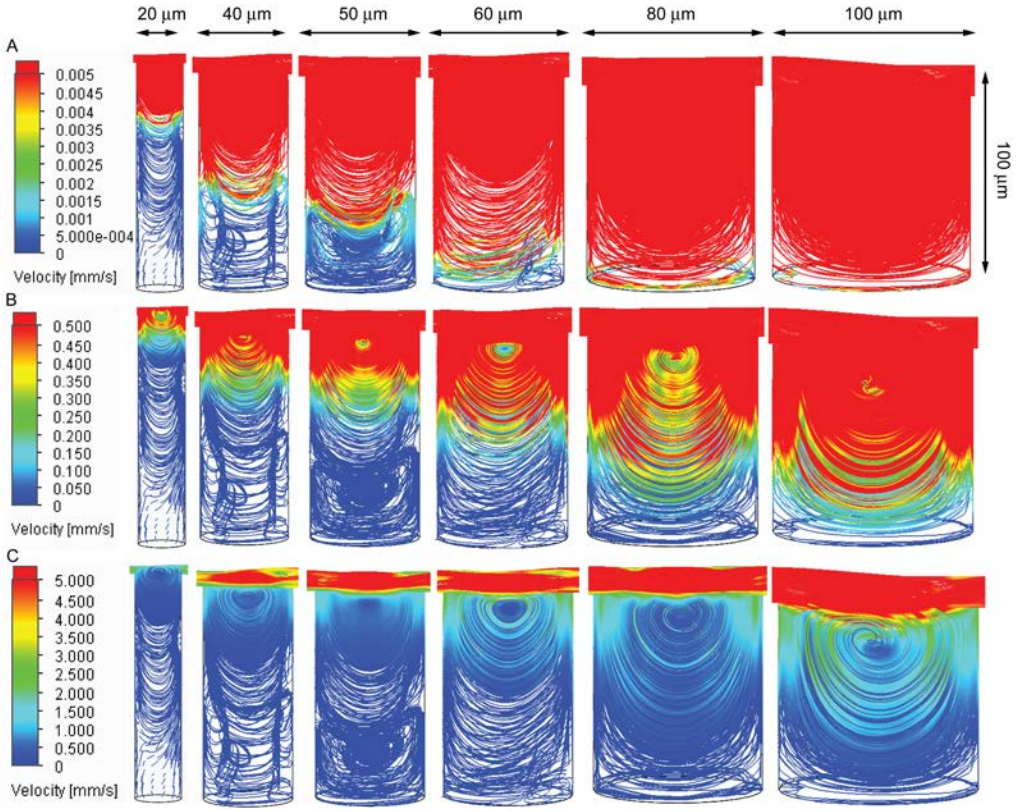

Figure 3.2: **Simulated flow dynamics inside the microwells**. (A-C). Flow dynamics inside microwells with a fixed height of 100 μm and various diameters from 20 μm to 100 μm (from left to right), with different flow speed scale bars. In the simulations, the flow speed in the centre of the channel is set at 100 mul/s ($\sim$16 cm/s). The details of the simulations are described in the methods section. The simulations were performed by my collaborator on this work, Dr. Artem Yefremov (MBI).

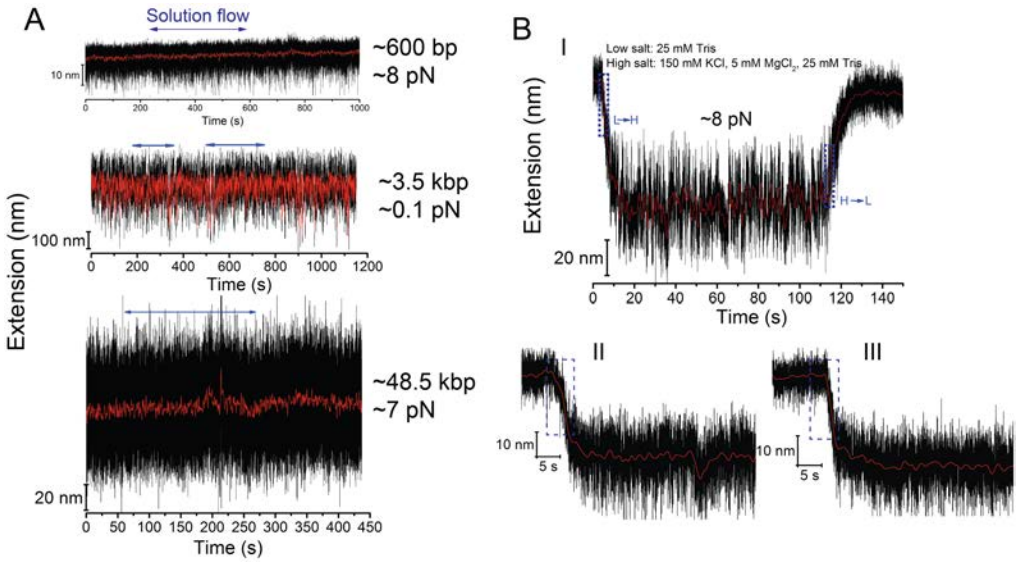

Figure 3.3: **Experimental test of the performance of the disturbance-elimination method**. (A). Representative extension time traces of dsDNA molecules with three different contour lengths (top - 576 bp; middle - ∼3.5 kbp; bottom - ∼48.5 kbp) at the bottom of a microwell (H: 100 μm; D: 40 μm) before, during and after introduction of a flow with speed of ∼10 cm/s. The horizontal blue arrows indicate the period of flow. The results show that the extension fluctuation is not affected by the flow. (B). Representative extension time traces of a 576 nt ssDNA switched between a low salt concentration (20 mM Tris without other salts) and a higher salt concentration (20 mM Tris with 100 mM NaCl and 5 mM $MgCl_2$), which indicates a time scale of several seconds needed for complete solution exchange between fluids in the microwell and in the channel. The blue dash boxes indicate the periods of solution exchange.

3.3 Fabrication of the microwell array

The microwells were fabricated following the procedures described in [131]. Positive mould consisting of 40 μm-diameter, 100 mum-height microwells were constructed using UV light lithography on a 4 inch silicone wafer with SU8 photoresist. The wafer was passivized by silanization and 10:1 PDMS mixture was casted on the SU8 wafer to form a pillar like negative mould. The PDMS mould was mounted on a glass slide with the top of the pillar facing the slide surface. Then UV curable polymer (OFN-134) was added to the side of the mould and was allowed to completely fill the gap between the PDMS mould and the slide. The slide was then submerged in DI water cured under 200 W UV for 8 min. Then the PDMS mould was carefully lifted from the assembly, leaving the cured polymer microwell membrane on the slide for storage.

3.4 Simulations of the flow dynamics in the microwell

To evaluate flow dynamics inside the experimental chamber, we used SOLIDWORKS Flow Simulation package to solve the Navier-Stokes equation. In these calculations we modelled the flow of buffer solution (density 1 g/cm^3, dynamic viscosity 10^-3 Pa s) in the central part of the chamber (dimensions: $5 \times 5 \times 0.2$ mm) containing a single well (e.g. diameter 20 μm, height 40 μm). To estimate the upper bound of the flow velocity at the bottom of the well (where protein/DNA stretching is performed) in the simulation we considered the extreme case when the buffer solution is pumped into the chamber at a very high rate ($\sim$ 16 cm/s). The simulation results are shown in Figure 3.2.

We first searched for dimension of the microwell such that it could simultaneously achieve disturbance-free and rapid exchange between the solution in the microwell and that in the channel. The former requires high height-to-diameter aspect ratio while the latter requires low height of the microwell (Figure 3.1). The flow dynamics in the microwell in the presence of flow with constant velocity in the channel was simulated using Flow Simulation package from SOLIDWORKS (Methods). We found that microwell with a diameter of 20-60 μm and a height

Solution exchange time scale estimation			
Well size (H×D μm)	Boundary speed V_c μm/s	Critical distance H_c μm	Exchange time τ_c s
100×20	2.8	71	25
100×40	4.9	41	8.4
100×50	6.5	31	4.8
100×60	14	14	1
100×80	50	4	0.1
100×100	200	$\sim$1	$\ll$0.1
150×75	5	40	8
200×100	4	50	13

Table 3.1: Solution exchange time scale estimation

of 40-100 μm could meet our need.

As shown in Figure 3.2, for microwell with a diameter of 20 $\sim$m and a height of 40 μm, when a high-speed flow of $\sim$16 cm/sec was applied in the channel, flow vortices were developed at the upper level of the microwell, while the flow speed was drastically reduced by around 10^6-folds (< 0.1 μm/sec) at the bottom level (within 15 μm from the bottom). We expected that the low height of the microwell should allow solution exchange at a time scale of seconds by diffusion facilitated by the vortices. Microwell array with such dimension can be easily produced.

3.5 Experimental validation of disturbance elimination in the microwell

According to the above simulation, the flow velocity is dramatically reduced to a negligible level near the bottom of the microwell in the presence of high-speed flow in the channel. We tested this prediction by monitoring the position of a 2.8-μm-diameter paramagnetic bead attached to one end of a single dsDNA tether before, during, and after a flow of $\sim$ 10 cm/s was applied at constant forces generated by the magnetic field. Figure 3.3 show results obtained from single 576 bp dsDNA tethers, where the flow in channel has negligible perturbation to the bead position, consistent with the simulation results. As such, the designed microwell is suitable for magnetic tweezers single-molecule manipulation experiments with disturbance-free rapid solution exchange.

3.6 Quantification of the solution exchange rate

In order to quantify the exchange rate between solution in the microwell and that in the channel, we measured the characteristic relaxation time of the ssDNA extension change in response to salt concentration change. Being a highly charged flexible polymer, the force-extension curve of ssDNA has been known highly sensitive to salt concentration [99, 100, 109]. At the same force, an ssDNA has a longer extension in lower salt concentration due to electric repelling and shorter extension in higher salt concentration due to electric screening. Utilizing this property, we examined the extension decreasing time course of single 576 nt ssDNA tethers after solution switching between low salt (20 mM Tris without other salts) and higher salt (20 mM Tris with 100 mM NaCl and 5 mM $MgCl_2$) (Figure 3.3).The results show that it took ∼5 seconds (salt concentration increase) to ∼10 seconds (salt concentration decrease) for the ssDNA to be relaxed to a new steady extension level upon solution exchange in the channel, indicating that the microwell is capable of rapid complete solution exchange within several seconds during high speed solution introduction in the channel, consistent with our theoretical estimation (Tabel 3.1). The slightly longer time scale involved in the salt-decrease induced ssDNA extension elongation than ssDNA extension reduction during salt increase can be explained by a longer time needed for decreasing than increasing salt concentration.

In the theoretical estimation, assuming a pure diffusion dependent process, a time scale of $\tau_{\text{diffusion}} \sim \frac{H^2}{(2D_{\text{diffusion}}}$ can be estimated based on the height of microwell (H) and the diffusion coefficient ($D_{\text{diffusion}}$) of the ligand. The diffusion coefficient can be estimated by the ligand size (r of a few nm for typical proteins) through the Einstein relation as $D_{\text{diffusion}} = \frac{k_{\text{B}}T}{6\pi\eta r}$. Further, a characteristic speed can be defined by the ratio of H and $\tau_{\text{diffusion}}$ as:$V_{\text{c}} = \frac{H}{\tau_{\text{diffusion}}}$. If the simulated flow speed V is larger than V_{c}, the mixing is flow dominated which is a fast process. For flow speed comparable or smaller than V_{c}, the mixing is dominated by slow diffusion, which is the rate-limiting step. Using this criterion, we analyzed the flow dynamics of microwells of different dimensions, and identify the height H_{c} at which $V \sim V_{\text{c}}$. The solution exchange time scales for various microwell

dimensions are then estimated by $\tau_c \sim \frac{H_c^2}{2D_{\text{diffusion}}}$, which are included in Tabel 3.1.

We chose the dimension of the microwells such that the critical height H_c is longer than the tether extension (typically shorter than 20 μm) to minimize the flow perturbation. We found a near optimal height-diameter aspect ratio $\sim$ 2:1 of microwells from the simulation, based on which we fabricated microwells accordingly.

3.7 Representative applications

We demonstrated the applications of this disturbance-free, rapid solution exchange method by probing the force dependent DNA structural transitions in response to solution change, and the polymerization dynamics of RecA on ssDNA or SSB-coated ssDNA.

3.7.1 Interconversions between force-dependent DNA structures induced by salt concentration change

Tensile force of 65 pN can induce a so-called DNA overstretching transition for torsion-unconstrained DNA, elongating the contour length of B-DNA by about $\sim$ 1.7 fold [81, 85]. Later studies have revealed three possible overstretched DNA structures [81, 82, 85–97]: i) 'peeling' transition to one peeled ssDNA strand under tension while the other ssDNA strand coils; ii) 'inside-strand separation' transition to two parallel ssDNA strands that share tension (melting bubbles), and ii) 'B-to-S' transition to a novel overstretched base-paired dsDNA, termed as S-DNA. The selection between each transition depends on DNA sequence, DNA topology, and environmental factors such as temperature and salt concentration [92–97]. Therefore, their inter-conversions can be induced by changing these factors while maintaining a constant force applied to the DNA, as demonstrated in a recent study through changing salt concentrations [97]. Surprisingly, in that study, the NaCl concentration decrease induced B-to-S transition took $>$ 40 seconds to finish, which was $>$ 10 times slower than that observed in previously reported force-induced B-to-S transition at a constant NaCl concentration. We

reason that this slow salt decrease induced B-to-S transition was caused by the slow solution exchange (> 5 minutes) that was needed to avoid flow perturbation in that study. To test it, we used the disturbance-free rapid solution exchange method to revisit the dynamics of inter-conversions between B-DNA and S-DNA induced by increasing/decreasing NaCl concentrations under a constant force.

Figure 3.4A shows the extension change of a 600 bp GC-rich torsion-unconstrained, end-closed DNA, being held at a constant force of $\sim$ 60 pN at 23°, induced by decreasing/increasing NaCl concentration between 250 mM and 0 mM in 20 mM Tris (pH 7.4). We observed inter-conversion between the B-DNA at 250 mM with a shorter extension and an overstretched DNA structure at 0 mM NaCl with an elongated extension by $\sim$ 1.6 fold. Due to the use of the end-closed (therefore peeling is pre-excluded) GC-rich DNA, the overstretched form is likely the S-DNA, which could be mixed with some small fraction of internally melted DNA[92–97].

Such DNA structural interconversion induced by decreasing/increasing NaCl concentration was highly reproducible, which took place immediately upon solution exchange and completed within $\sim$ 3 seconds (during salt decrease) and $\sim$ 0.3 seconds (during salt increase). Given that it takes $\sim$ 5 seconds for complete solution exchange in our system, the result indicates that the actual transition dynamics are faster than the observed time scales. The longer time scale involved in the salt-decrease induced DNA elongation transition than that in the reverse transition during salt increase can be explained by a longer time needed for decreasing than increasing salt concentration and/or due to that the critical salt concentration for the transition might not be in the middle of the two concentrations we tested.

3.7.2 The polymerization dynamics of RecA on ssDNA

As introduced in Chapter 1, RecA is a crucial recombinase in bacteria that plays a critical role in homologous recombinational DNA damage repair. It nucleates and polymerizes on ssDNA, forming a rigid extended helical nucleoprotein filament [10, 11]. RecA binds to ssDNA with a high affinity of $\sim$ 100 nM and polymerizes with high cooperativity in a preferential $5'$-to-$3'$ direction [10, 11].

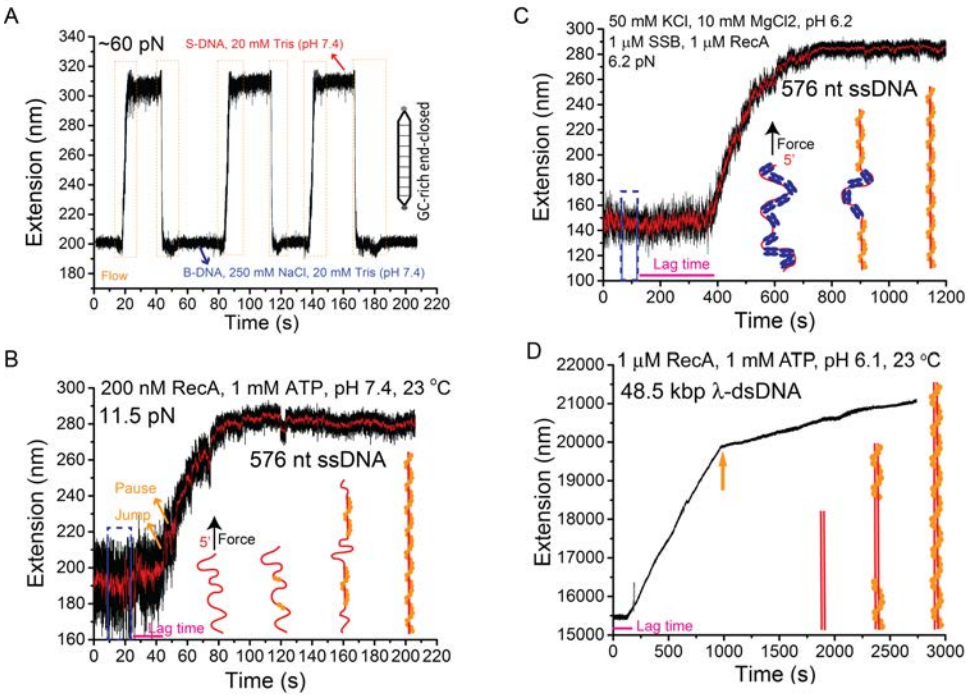

Figure 3.4: **Representative applications of the disturbance-free rapid solution exchange method**. (A). Interconversions of a $\sim$ 600 bp end-closed GC-rich dsDNA between the B-form and the overstretched S-form DNA structures induced by salt concentration change at a constant force of $\sim$ 60 pN. During cycles of switching between 250 mM and 0 mM NaCl in 20 mM Tris (pH 7.4), the DNA extension correspondingly switched between a shorter level (B-DNA) in 250 mM NaCl and a longer level (S-DNA) in 0 mM NaCl. The orange dash boxes indicate the time windows during rapid solution exchange. (B). A typical extension time trace of a 576 nt ssDNA before, during and after 200 nM RecA (with 1 mM ATP, 1x ATP regeneration system, 50 mM NaCl, 10 mM $MgCl_2$, 20 mM Tris pH 7.4) was introduced at a constant force of $\sim$ 11 pN. (C) A typical extension time trace of a 576 nt SSB-coated ssDNA before, during and after solution containing 1 μM RecA and 1 μM SSB (with 1 mM ATP, 1x ATP regeneration system, 50 mM NaCl, 10 mM $MgCl_2$, 20 mM MES pH 6.2) was introduced at a force of $\sim$ 6.5 pN. For B&C, schematics of the RecA polymerization process are shown in figure panels. (D). A representative extension time trace of a 48.5 kbp λ-DNA before, during and after 1 μM RecA (with 1 mM ATP, 1x ATP regeneration system, 50 mM NaCl, 10 mM $MgCl_2$, 20 mM MES pH 6.1) was introduced at a constant force of $\sim$ 10 pN. The solution was introduced within 10 seconds at the beginning of the time trace. The spike at $\sim$ 240 s is due to transient diffusion of a polystyrene bead into the view area that affected the bead imaging. Blue dash boxes in panels B-C indicate the flow periods.

RecA filament can also polymerize on one ssDNA strand inside a double-stranded DNA [10, 11]. Due to a high nucleation barrier and slowed-down polymerization rate inside dsDNA, its polymerization was able to be measured from 2 to 20 monomerssec on micrometer sized DNA [10, 22–24, 30, 31]. In contrast to ds-DNA, due to decreased nucleation barrier and faster polymerization rate, the

dynamics of RecA polymerization on ssDNA has been difficult to be probed in previous single-ssDNA stretching experiments. To our knowledge, there was only one single-ssDNA stretching experiment reporting the polymerization dynamics of RecA, which was achieved using a long ssDNA (7.3 kb) that allowed a longer time of observation [23]. Here we show that the RecA nucleation and polymerization dynamics can be conveniently recorded on much shorter (∼600 bp) ssDNA (therefore better signal-to-noise ratio in extension measurements) using the microwell based disturbance-free rapid solution exchange method.

Figure 3.4B shows a typical time trace of the extension of a 576 nt ssDNA under a constant force of $\sim$ 11.5 pN, before, during and after 200 nM RecA was introduced with a flow speed of $\sim$ 7 mmsec, in the standard RecA reaction buffered solution containing 50 mM NaCl, 10 mM $MgCl_2$, 1 mM ATP, 1x ATP regeneration system, and 20 mM Tris-HCl (pH 7.4), at 23 ^{o}C. The polymerization of RecA began several seconds after the complete solution exchange of RecA solution. We reason that the several seconds lag time (indicated by the magenta line) corresponds to the time for RecA nucleation, as it has been known that RecA polymerization requires a nucleation step [10]. The elongation of the ssDNA extension is due to formation of the rigid RecA nucleoprotein filament (bending persistence of RecA filament is $\sim$ 1000 nm [23]).

The progressive elongation shows an interesting jump-pause-like dynamics of RecA polymerization (indicated by orange arrows). We reason that the jump corresponds to a cooperative polymerization from a newly nucleated RecA patch till the canonical polymerizing end (the $3'$ end) [10] reached the end of ssDNA, or reached the $5'$ end of an existing downstream RecA filament. Each pause can be explained to be the time for formation of a new nucleation site. Overall, the results are consistent with previously reported cooperative nucleation and polymerization of RecA on ssDNA [10].

It took about 60 seconds to reach a steady elongated extension with a total extension increase of $\sim$ 80 nm. It has been known that in RecA-ssDNA nucleoprotein filament, a RecA monomer occupies 3 nt of ssDNA. Therefore, our result of fully polymerization on the 576 nt ssDNA reveals an overall polymerization rate of 3 $\pm$ 1 monomer s^{-1} in 200 nM RecA solution under our solution con-

dition. Given that multiple nucleation sites might exist in the 576 nt ssDNA, the observed overall polymerization rate might be greater than that of a single filament.

3.7.3 Dynamics of RecA loading onto SSB coated ssDNA

In bacteria, newly generated ssDNA fragment is immediately bound by SSB to prevent the formation of secondary structure or ssDNA degradation. The nucleoprotein array formed on ssDNA by SSB on one hand facilitates RecA loading onto ssDNA by removal of secondary structure, on the other hand suppresses the RecA filament formation due to the energy cost to remove the SSB on ssDNA [10]. The dynamics and kinetics of RecA loading onto SSB coated ssDNA therefore are physiologically important. Recently, the RecA loading onto SSB coated ssDNA was probed using combination of optical traps and fluorescent dynamic imaging in the presence of laminar flow, which revealed a bi-directional ($5'$-to-$3'$ and $3'$-to-$5'$) loading processes of RecA onto SSB-coated ssDNA [128]. Here, we demonstrate fluorescence-label-free measurements of the dynamics of RecA loading onto SSB-ssDNA nucleoprotein array using the disturbance-free rapid solution exchange method with magnetic tweezers.

In the experiments, a fully covered SSB nucleoprotein array was formed on ssDNA by incubating the ssDNA tether with 1 ∼M SSB (in standard RecA reaction solution at pH 6.2) at ∼6 pN. Then, a mixture of 1 ∼M SSB and 1 ∼M RecA was introduced. The extension of ssDNA was recorded before, during, and after the introduction of the protein solution (Figure 3.4C). After a lag phase of $\sim$ 200 seconds (indicated by magenta line) following the protein introduction, the extension began to gradually increase without apparent pause-jump-like dynamics observed in the absence of SSB. It took $\sim$ 400 seconds till RecA fully covered the ssDNA. From the extension time traces, a loading rate of ~ 0.5 monomer s^{-1} was estimated, which is consistent with recent reported speed of 0.1 - 1 monomer s^{-1} by fluorescent dynamic imaging using optical tweezers [128].

3.7.4 The polymerization dynamics of RecA on long double-stranded λ-DNA

To demonstrate that the disturbance-free rapid solution exchange method is also suitable for longer tethers, we used RecA polymerization dynamics on long λ-DNA (48,502 bp, $\sim$ 16 μm) as an example. As shown in Figure 3.4D, after RecA solution (1 $\sim$M) was introduced, there was a lag phase of over 200 seconds before RecA polymerization started. This long lag phase likely corresponds to nucleation of RecA inside dsDNA. After the nucleation step, RecA began to polymerize on the dsDNA, indicated by progressive extension increase of the dsDNA with a rapid growth phase followed by a slowly growing phase after a transition point (indicated by orange arrow in Figure 3.4D). Together, these results demonstrate that this method is capable to probe dynamics of DNA and DNA-protein interactions over a wide range of length scales.

Chapter 4

ATP-hydrolysis and Force mediated competitive regulation of RecA-ssDNA nucleoprotein filament by SSB

4.1 Chapter summary

As introduced in Chapter 1, the bacterial RecA filament formed on ssDNA is the essential factor of homologous recombinational DNA damage repair. While plenty of our understanding of the RecA filament is based on direct binding of RecA to ssDNA, the SSB protein actually binds to ssDNA prior to RecA *in vivo*, and forms a SSB-ssDNA nucleoprotein array, which plays complicated roles in the RecA filament regulation. However, the molecular mechanisms of the regulation by SSB are still elusive. In this chapter, based on the distinctly different force-responses of individual single short ssDNA (576-nt) associated with RecA or SSB proteins, we investigate the mechanisms by which SSB regulates the RecA filament formation and stability, as well as the affecting factors of the regulation mechanisms. We find that the SSB protein inhibits RecA nucleation and polymerization by outcompeting even higher concentrations of RecA, forming SSB-ssDNA nucleoprotein array. Furthermore, in the presence of SSB, the

pre-formed RecA filaments go through de-polymerization process in an ATP hydrolysis dependent manner at low forces (∼3 pN). The net de-polymerization of RecA filament in the SSB and ATP hydrolysis dependent manner is likely resulted from that SSB occupies the vacated naked ssDNA due to the RecA disassociation at 5′ end of the DNA during ATP hydrolysis. Interestingly, at higher forces (>∼15 pN), re-polymerization of RecA filament occurs, driving away the SSB on the ssDNA and forming a fully re-polymerized RecA filament. Importantly, since the 5′-to-3′ direction polymerization has been inhibited by the occupied SSB array on the ssDNA, the re-polymerization of RecA filament has to take place at a 3′-to-5′ direction from the remained partial RecA filament. These force-dependent re-polymerization, therefore, provides evidence of the bi-directional (5′-to-3′, and 3′-to-5′) polymerization of RecA on ssDNA, which is contradictory to previous well-known uni-directional (5′-to-3′) polymerization of RecA filament. Moreover, the study depicts a physical picture of the competitive regulation of the SSB on the scale of entire nucleoprotein SSB array on the RecA filament formation and stability under mechanical forces, which has broad biological implications, especially competitive molecular regulations. *

4.2 Introduction

The *E.coli* RecA nucleates and polymerizes on the ssDNA, forming the RecA nucleoprotein filament which is essential for the homologous recombination, and DNA repair [10, 11]. The formation of the RecA filament can in general divided into two kinetically distinct steps: a slow initial nucleation step, and a faster directional, highly cooperative polymerization step. Both the nucleation and polymerization of RecA onto ssDNA requires ATP or its analogs, such as dATP, ATPγS; ATP-hydrolysis leads to disassociation of the RecA from ssDNA [10, 11]. Therefore, the stability of the dynamic RecA filament results from the competition of polymerization and de-polymerization of RecA on ssDNA.

The formation and stability of active (formed with ATP) RecA nucleoprotein

*Note that main contents detailed in this chapter have been published in *Force and ATP hydrolysis dependent regulation of RecA nucleoprotein filament by single-stranded DNA binding protein*, Fu, H., Le S. *et al.*, Nucleic Acids Research 41, 924-932 (2013).

filament have been intensively studied for decades based on bulk biochemical and biophysical approaches [15, 16, 132]. Recently, advancement of single-molecule manipulation techniques, such as optical tweezers, magnetic tweezers and single-molecule FRET experiments [23, 24, 31, 133, 134], have enhanced our understanding of the formation and stability of the RecA nucleoprotein filament, by providing the direct measurements of kinetics of RecA nucleation, polymerization, and de-polymerization on ssDNA/dsDNA with high resolutions.

However, *in vivo*, the ssDNA produced during the DNA replication, recombination, and damage is immediately bound with SSB proteins, forming an SSB-ssDNA nucleoprotein array, which protects the ssDNA intermediates from degradation and secondary structure [10, 11, 33, 135, 136]. This SSB-ssDNA nucleoprotein array, at the same time, leads to the energy barrier for the formation of the RecA-ssDNA nucleoprotein filament. The gaps still persist in understanding of how the SSB protein and other recombination accessory proteins affect the formation and stability of RecA nucleoprotein filament.

SSB binds to ssDNA nonspecifically with high affinity and multiple binding modes [10, 11, 33, 40, 41, 137, 138]. Up to now, the binding modes of SSB to ssDNA are mainly divided into two groups: 1) more-wrapped modes $(SSB)_{56}$ and $(SSB)_{56}$, where each SSB tetramer wrap $\sim$ 56 or 65 nt of ssDNA, and 2) the less-wrapped mode $(SSB)_{35}$, where each SSB tetramer binds $\sim$ 35 nt of ssDNA. Many factors including salt, pH, and protein concentration have been demonstrated to affect the selection of SSB binding modes [10, 11, 33, 40, 41, 137, 138]. According to previous reports, the less-wrapped $(SSB)_{35}$ mode is favored at low salt concentrations at high SSB binding density on ssDNA. However, these SSB binding modes were studied in the absence of force on ssDNA, which likely regulates the binding modes of SSB on tension mediated ssDNA template [10, 11, 33, 40, 41, 137, 138].

While RecA and SSB do not interact with each other, they actively competes the ssDNA binding sites. Although effects of SSB and RecA competition on ssDNA have been extensively studied [10, 11, 135, 136], how the SSB competes with RecA on ssDNA and its regulation on the formation of RecA nucleoprotein filaments have not yet been studied at single-DNA molecule resolution with

the presence of mechanical forces which can potentially affect the competition. Moreover, although individual SSB tetramer has been demonstrated can sliding on ssDNA [139], since *in vivo* the SSB tetramers are densely packed on the ssDNA as a nucleoprotein array where events of individual tetramers sliding along ssDNA is unlikely or rare during homologous recombination, the properties of the tightly coated SSB tetramers array and its impact on the formation of RecA nucleoprotein filaments has not yet been studied at an entire SSB-ssDNA nucleoprotein array scale at single molecular level.

In this chapter, to address the aforementioned unknowns, we investigate dynamics of ssDNA and pre-formed RecA nucleoprotein filaments on ssDNA stretched under mechanical forces, in the presence of free SSB protein in solution. The dynamics reveals the competitive regulation of SSB nucleoprotein array on the formation and stability of RecA nucleoprotein filament.

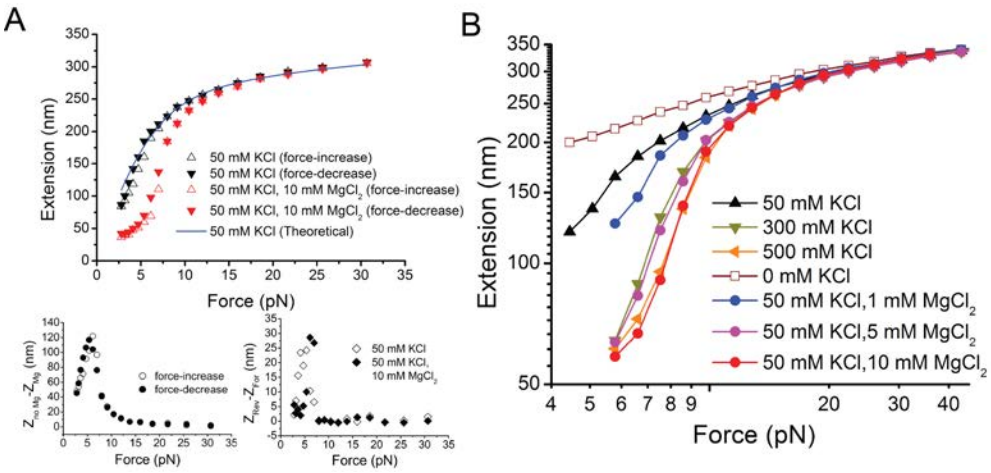

Figure 4.1: **Force Responses of ssDNA**. (A). Force extension curves of a 576-nt ssDNA in solutions containing 50 mM KCl, with (red) or without (black) 10 mM $MgCl_2$, 20 mM Tris-pH7.4, at 23°C. The hollow symbols denote force-increasing scans and the solid symbols denote force-decreasing scans. Blue line is the theoretical force response of ssDNA based on Eq. 1.3. (B). Left panel: extension difference of the ssDNA in the presence or absence of $MgCl_2$; Right panel: extension difference during force-decreasing scans and force-increasing scans. (C). Force extension curves of a 576-nt ssDNA in various salt concentrations. Only force-decreasing curves are plotted in figure for simplicity; Data are plotted in Log-scale for better presentation of extension differences at low force <15 pN.

4.3 Results

4.3.1 Force-responses of ssDNA in various solution conditions

The studies are largely based on the differences of the ssDNA force responses bound with SSB, RecA or other proteins, resulting from the conformational rearrangement due to the protein-ssDNA complex formation. Therefore, the single ssDNA tether used has to be short enough to suppress thermal noise from the longitude conformational fluctuation of the flexible ssDNA, in order to clearly observe the signals. However, the ssDNA on the other hand has to be long enough for the formation of stable RecA filament and SSB-ssDNA nucleoprotein array. As introduced in the Methods and Materials in Chapter 2, a 576 nucleotides ssDNA which is short enough to increase the signal-to-noise ratio is constructed in our experiments. Moreover, the 576-nt DNA tether can form a SSB-ssDNA nucleoprotein array with ∼8-16 SSB tetramers (corresponding to ∼ 16 SSB proteins in the $(SSB)_{35}$ binding mode or ∼ 8 SSB proteins in the $((SSB)_{65}$ binding mode) [10, 40, 137] and is much longer than the minimal length of the RecA nucleation

site (<12 nt), and the minimal length of stable RecA filament (∼35nt) [15, 20].

Figure 4.1 A shows the force-extension curves of a single ssDNA tether recorded during force-decreasing scans (solid down tri-angles) and force-increasing scans (open up tri-angles) at ∼23 °C in buffered solutions containing 20 mM Tris (pH 7.4), 50 mM KCl (black symbols), and 20 mM Tris (pH 7.4), 50 mM KCl and 10 mM $MgCl_2$ (red symbols). Secondary structures within the coiled ssDNA can potentially form at low force, leading to hysteresis in the ssDNA extension between the force-decreasing and force-increasing scans. The ssDNA was hold at each force for 5 seconds (unless specifically stated). The force decreasing-increasing cycle scans were repeated for ≥3 times. The extension at each force was averaged for data recorded for the repeated measurements (the standard deviations were smaller or comparable with the symbols size and were not shown for simplicity). We quantified the hysteresis level of 576-nt ssDNA by plotting the extension difference at each force during the force-decreasing scan and force-increasing scan (Figure 4.1 A, bottom panel). The secondary structures formed with a peak hysteresis value of ∼ 30 nm at ∼ 7 pN.

Furthermore, While the force-extension curve of ssDNA at 50 mM KCl agrees well with theoretical curves based on Eq.1.3 [82], the one with 10 mM $MgCl_2$ differs from the theoretical values at low force range, likely due to ssDNA condensation. To gain more insights on the effects of salt, we further quantified the effects of various concentrations of KCl (0 mM, 50 mM, 300 mM, and 500 mM) and $MgCl_2$ (0 mM, 1mM, 5 mM, and 10 mM) on the force response of the 576-nt ssDNA tether. As Figure 4.1 B shows, the extension of ssDNA monotonically decreases as KCl concentration is increased (triangles indicated by different colors). At 50 mM KCl concentration, the increase of $MgCl_2$ concentration also resulted in monotonic decrease in the extension of ssDNA (solid circles indicated by different colors. Moreover, ssDNA is condensated at high KCl concentration or high magnesium concentration, which depends on magnesium concentration more sensitive than on the KCl concentration. The more sensitive dependence on magnesium concentration than on monovalent cation concentration is consistent with previous observation that ssDNA tethers tend to collapse at forces lower than 10 pN in a high NaCl concentration (> 150 mM) and in high magnesium

concentration ($\sim$ 10 mM) [98, 99].

In summary, the force-response of the 576-nt ssDNA and the effects of monovalent and magnesium concentration on it are in consistence with previous studies of ssDNA force responses [98, 99]. In addition, hereafter, the following force-extension curves plotted will only show the data obtained during the fore-decreasing scans unless specifically stated to minimize the additional effects of secondary structure formation on ssDNA extension.

4.3.2 Distinct force-responses of RecA-ssDNA nucleoprotein and SSB-ssDNA nucleoprotein array corresponds to their distinct conformations

Having quantified the force-responses of the naked short ssDNA in various solutions, we further investigate the effects of RecA and SSB proteins on the force-response of ssDNA when these proteins form complex with ssDNA, by comparing the force-extension curves of ssDNA in the presence and absence of RecA or SSB.

Next, we measured the force-responses of ssDNA in the presence of 1 pM to 1 μM SSB proteins in various concentrations of KCl (0, 50, 300, 500 mM)and $MgCl_2$ (0, 5, 10 mM) in the force range of $\sim 3-40$ pN(Figure 4.2). We emphasize that since no free ssDNA was present in solution, SSB is always in excess to ssDNA. In the absent of KCl and $MgCl_2$ (Figure 4.2 A), the force-extension curves of SSB bound ssDNA with $>$ 10 pM SSB in solution are always shorter than that of naked ssDNA, indicating wrapping of ssDNA by SSB. In the presence of increasing concentration of KCl or $MgCl_2$ (Figure 4.2 B-F), the extension changes of ssDNA can be roughly divided into three ranges. First, at forces below $\sim$5-7 pN, and $>$ 100 pM-nM SSB, the extensions of SSB-ssDNA array are longer than that of naked ssDNA, which is likely caused by removal of secondary structures within ssDNA and repulsive interactions between adjacent units of SSB bound to ssDNA, although the ssDNA is wrapped by SSB. Second, in contrast, in the force range from lower limits of 5-7 pN to up-limits of 15-20 pN, the extensions of SSB-ssDNA nucleoprotein array are shorter than naked ssDNA, which can be attributed to wrapping effect of SSB tetramers on ssDNA [40, 41, 137]. Third, for forces >15-20 pN, the extensions of ssDNA bound by SSB overlap (or are

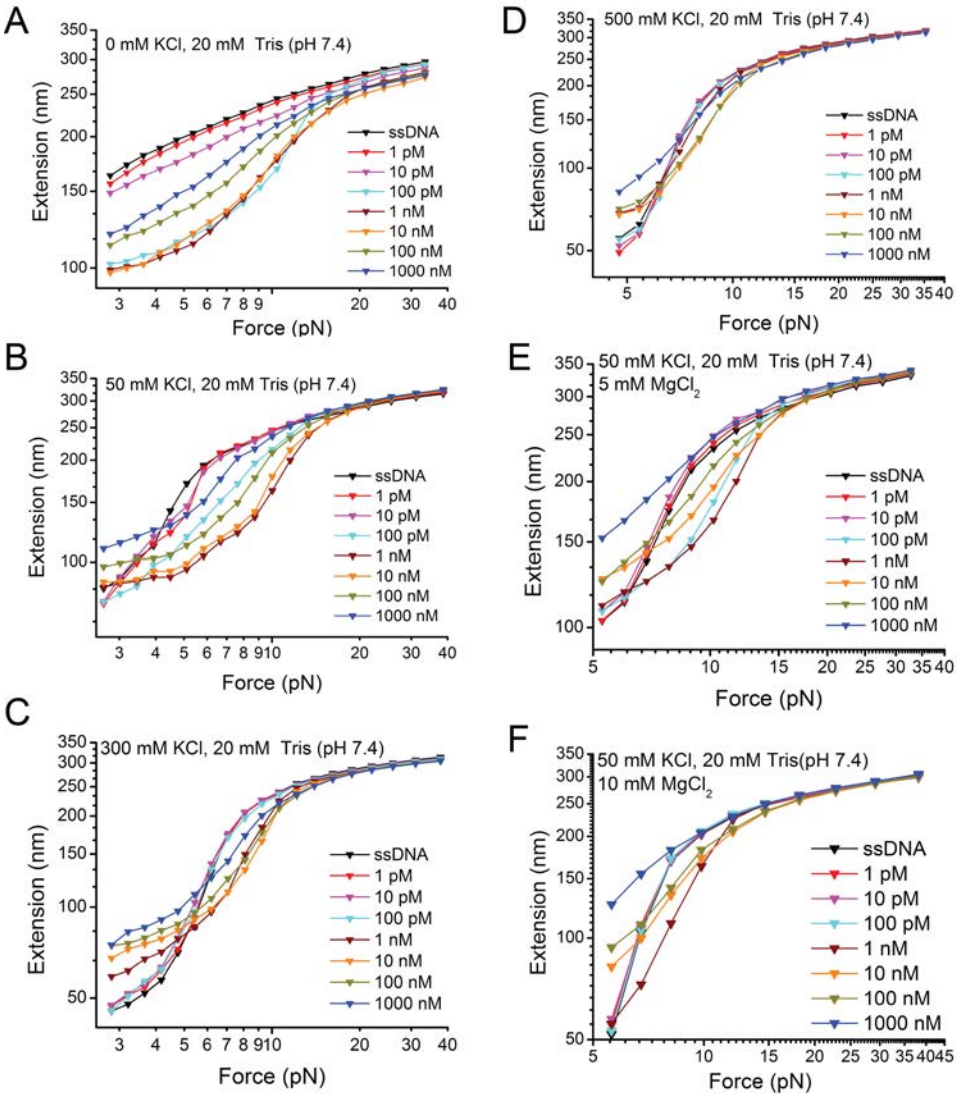

Figure 4.2: **Force Responses of SSB-ssDNA nucleoprotein array in different salt and SSB concentrations**. (A-D). Force extension curves of individual 576-nt ssDNAs coated with SSB in 0, 50, 300, 500 mM KCl (panel A-D, respectively) in the presence of SSB concentration ranging from 1 pM to 1μM (indicated by colours). (E-F). Force extension curves of individual 576-nt ssDNAs coated with SSB in 50 mM KCl with 5 mM (E) or 10 mM (F) $MgCl_2$, in the presence of SSB concentration ranging from 1 pM to 1μM (indicated by colours).

very slightly shorter than) that of naked ssDNA for all the SSB concentrations tested. In this force range, either dissociation of SSB proteins from the ssDNA or binding of SSB on the ssDNA in a non-wrapping mode can explain the overlapped extensions. To distinguish the two possibilities, we performed 'force-jumping cycles' of the ssDNA extension. As Figure 4.3 shows, the forces on the naked ssDNA (orange), the same ssDNA in 10 nM SSB (red), and the same ssDNA in 1 μM SSB (blue) were jumped from a high force ($\sim$41 pN, black) to a series

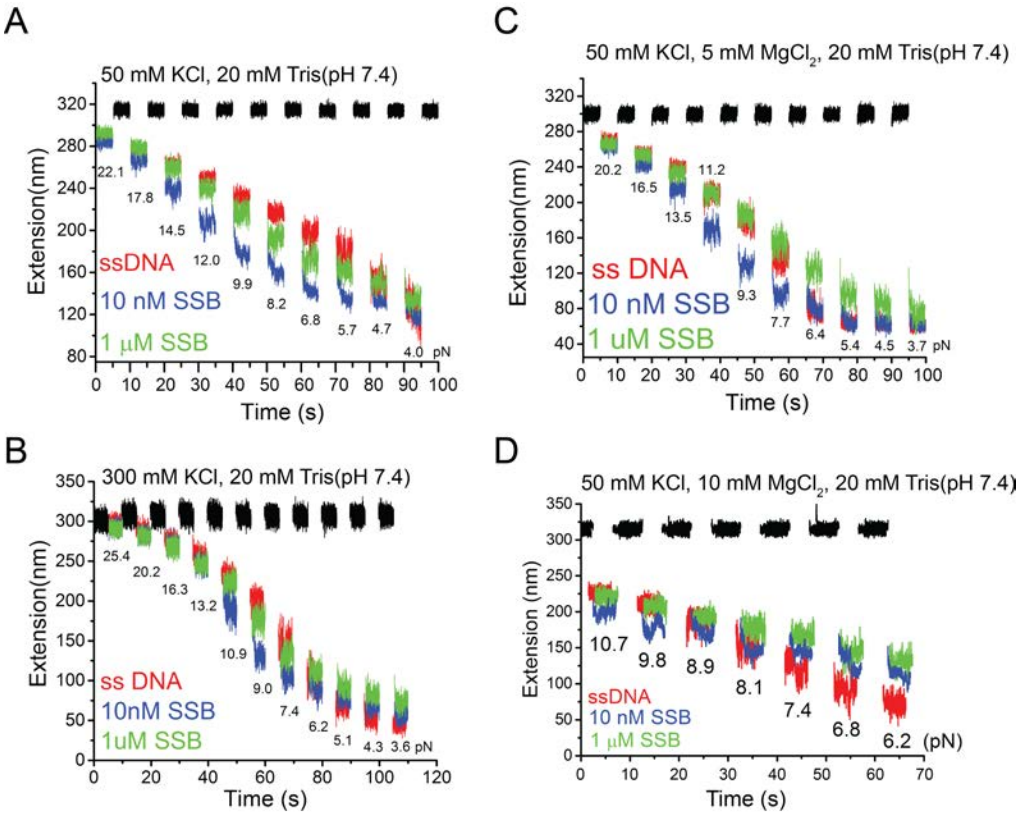

Figure 4.3: **Extension time traces of SSB-ssDNA nucleoprotein array in different salt and SSB concentrations during force-jumping**. (A-D). Extension time traces of SSB-ssDNA nucleoprotein array in different salt and SSB concentrations during force-jumping in the presence of SSB concentrations of 10 nM and 1μM (indicated by colours) in different salt concentrations.

of lower forces (colors). At each force, the DNA was held for 5 seconds. Since the force-jumping process takes only $\sim$ 0.1 second, we can monitor the initial dynamics of the extension when force was jumped from the high force (where the extension is almost overlapped with naked ssDNA). In the presence of 10 nM SSB, immediately after the force jump, the ssDNA extension is longer at $<$8 pN and shorter at $>$8 pN than that of naked ssDNA, consistent with the results in Figure 3.2B. Considering the quick jumping of $\sim$ 0.1 second, this result highly suggests that SSB is still bound with ssDNA in a non-wrapping mode at high force before force-jumping. Otherwise, immediately after the force-jumping, the initial ssDNA extension should be close to naked ssDNA extension. Similarly, in the presence of 1 μM SSB, at lower forces immediately after the force-jumping, the ssDNA extension is consistent with the force-extension curve in the force-scan in Figure 3.2B, again suggesting that the SSB is still associated with ssDNA at

high force. Interestingly, at low forces (6-9 pN in Figure 3.2C), the extensions of ssDNA in 10 nM SSB dropped from an extension similar to that in 1 μM SSB, to the shorter extension during the 5 seconds recording. The dynamics of the extensions may suggest that the SSB-ssDNA nucleoprotein array in the 10 nM SSB reorganized from a non-wrapping mode to a wrapping mode when jumped from high force to low force. Therefore, these results demonstrate that force on ssDNA can switch the SSB-ssDNA nucleoprotein array from more-wrapped binding mode(s) to less-wrapped or non-wrapped binding mode(s).

Interestingly, we also found that SSB wrapping effects reach the maximum at SSB concentration of 100-10 nM (for different salt concentration), as revealed by the maximum extension shortening effects in the medium force range ($\sim$ 5-20 pN). Below the maximum wrapping-effect SSB concentration, the wrapping effects of the SSB-ssDNA nucleoprotein array increases as SSB concentration increases, likely due to more SSB binds to ssDNA; In contrast, Above the maximum wrapping-effect SSB concentration, the wrapping effects of SSB-ssDNA nucleoprotein array decreases as SSB concentration increases, which likely suggests that with higher SSB concentration, SSB-ssDNA nucleoprotein array switches from more-wrapped binding modes to less-wrapped (or non-wrapped) binding modes to incorporates more SSB onto ssDNA. Particularly, with 1 μM SSB at 5-10 mM MgC_2 concentration, the force-extension curves of the SSB-ssDNA nucleoprotein arrays seems exist in non-wrapped binding mode. These results demonstrate that the SSB binding modes on SSB-ssDNA nucleoprotein array are regulated by SSB concentration. Particularly, at high SSB concentrations ($>$ 100 nM), which is likely the *in vivo* SSB concentration, with the presence of force, it is likely in a less-wrapping or non-wrapping state.

On the other hand, RecA nucleoprotein filament is expected to have distinctively different force-responses due to its high rigidity. Figure 4.4 A shows extension dynamics of a 576-nt ssDNA at different forces in the presence of 1 μM RecA in standard solution. At $\sim$ 9-10 pN, RecA progressively polymerization was observed, indicated by dramatic step-wise increase of extension by $\sim$150 nm. ssDNA secondary structures are likely formed at lower forces, which impedes the polymerization of RecA-ssDNA nucleoprotein filament. After fully

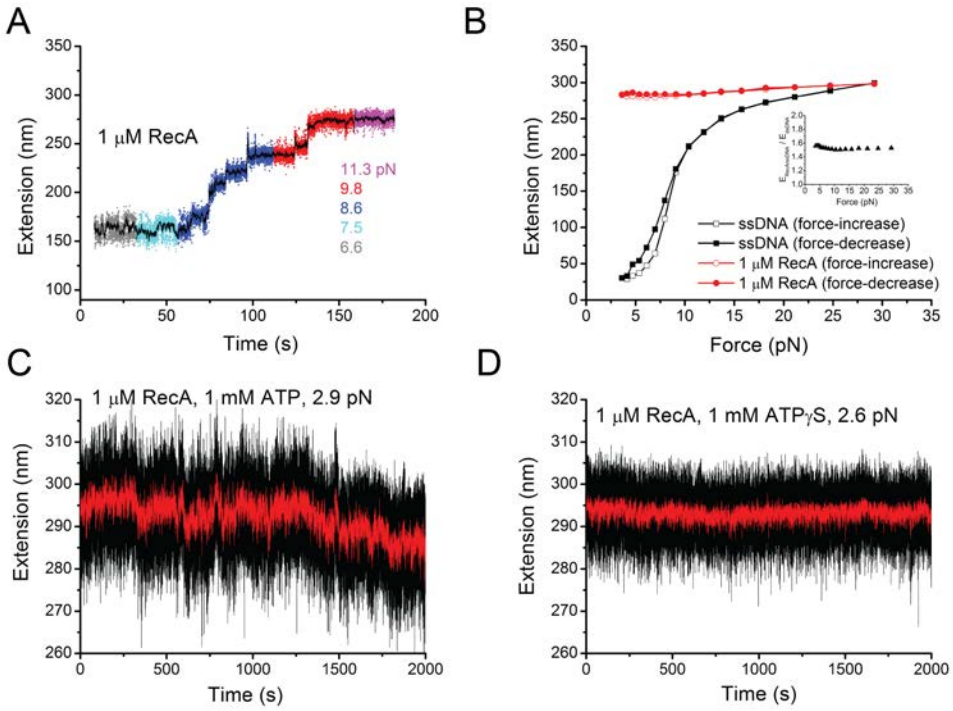

Figure 4.4: **Extension dynamics and force-extension curves of RecA-ssDNA nucleoprotein filament**. (A). Extension dynamics of a 576-nt ssDNA in the presence of 1 μM RecA in standard reaction solution at different forces (indicated by colours). (B) the force-extension curves of the naked ssDNA (squares) and the RecA nucleoprotein filament (circles) formed on the same ssDNA. (C-D) Extension dynamics of fully polymerised RecA nucleoprotein filaments formed on 576-nt ssDNA with ATP (C) or ATPγS as binding co-factor, in standard reaction solution 1 μM free RecA. The red lines are 100-points FFT smooth of the raw extension (black).

polymerization of RecA-ssDNA nucleoprotein filament, we measured the force-extension curves of it in the standard RecA reaction solution in the presence of free RecA with a concentration of 1 μM and 1 mM ATP, 1 $\times$ATP regeneration system (circles in Figure 4.4B). The concentration of 1 μM is over saturated for the fully polymerization of RecA filament on the ssDNA tether, since there is no free ssDNA presented in solution. In another word, the RecA is always in excess to ssDNA binding sites.

RecA nucleoprotein filament is known a stiff structure with a bending persistence length larger than 1 μm [21]. Consistent with that, the force-extension curve of the ssDNA polymerized by RecA is always longer than that of the naked ssDNA before introduction of RecA at forces up to $\sim$30 pN(Figure 4.4B). Moreover, the extension of RecA filament is about 150% of the B-form dsDNA of an equal number of base pairs (Figure 4.4B, inset), which is in accord with previous

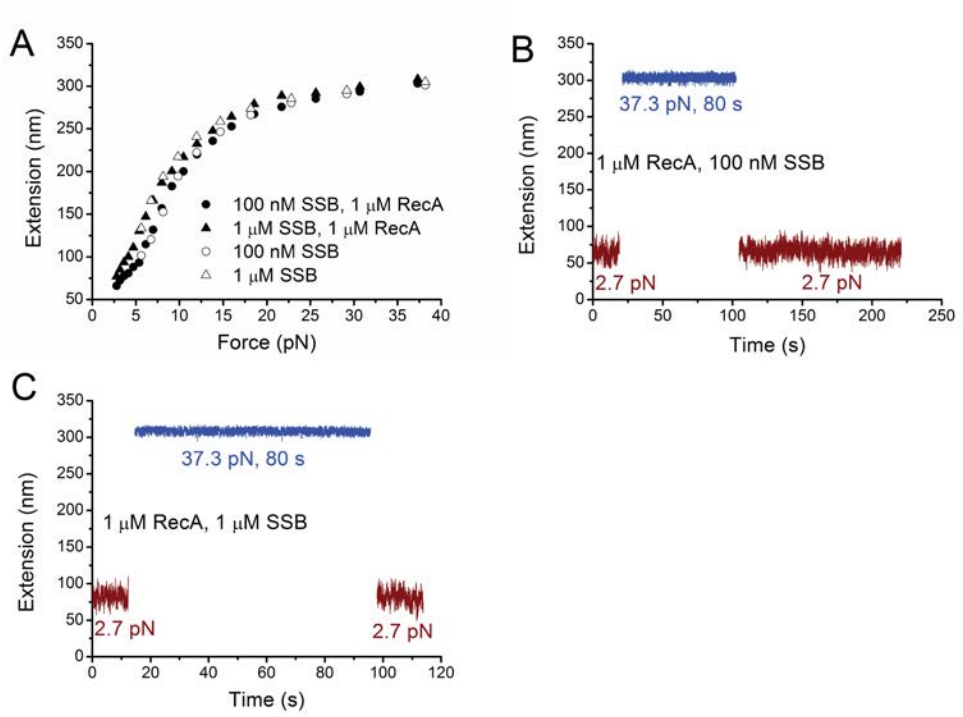

Figure 4.5: **SSB inhibits RecA nucleation and polymerization on ssDNA**. (A). force extension curves of a 576-nt ssDNA in the presence of mixture of 100 nM (solid circles) or 1 μM (solid tri-angles) SSB and 1 μM RecA in standard reaction solution, compared with those with SSB only (hollow symbols). (B-C) Extension dynamics of the ssDNA in the presence of mixture of SSB and RecA during jumping between a higher force of $\sim$37 pN and lower force of $\sim$3 pN.

studies. Additionally, in the presence of 1 μM free RecA in solution, the RecA-ssDNA nucleoprotein filament is stable at low force ($\sim$ 3 pN) over long time (> 2000 sec as shown in Figure 4.4C&D) with ATP (panel C) or ATPγS (panel D) as binding co-factor, the latter is less dynamics due to lacking of ATP-hydrolysis.

Overall, we show differential force responses (extensions) of ssDNA bound with RecA and SSB, which depend on the applied forces and concentrations of the accessory proteins. Again, we emphasize that since no free ssDNA was present in solution, both RecA and SSB proteins are always in excess to ssDNA. Hence, the force and SSB concentration dependence of ssDNA extension likely suggests the concentration dependent switch of binding modes with different binding density of SSB on ssDNA. Besides the known wrapping binding modes of SSB, such as $(SSB)_{35}$, $(SSB)_{56}$, $(SSB)_{65}$ [10], there is non-wrapping binding mode of SSB when ssDNA is stretched under larger enough force.

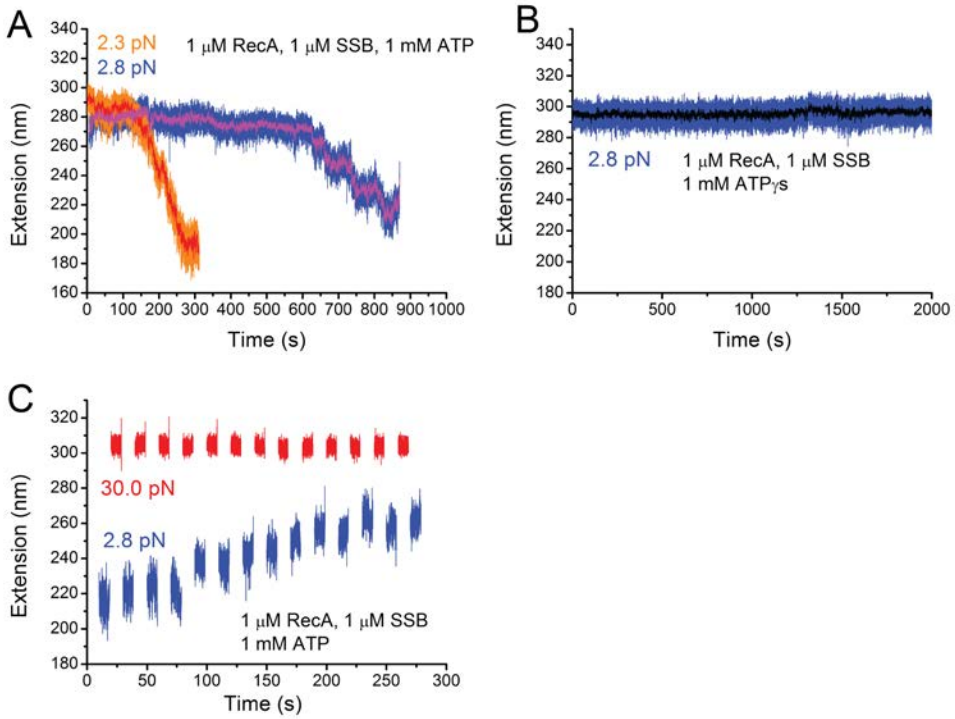

Figure 4.6: **SSB, ATP hydrolysis and force-dependent stability of preformed RecA nucleoprotein filaments**. (A). Typical extension dynamics of de-polymerization of preformed RecA nucleoprotein filaments (ATP as binding co-factor) at low forces of ∼3 pN, in mixture of 1 μM SSB and 1 μM RecA, 1mM ATP in standard reaction solution. (B) Typical extension dynamics of preformed RecA nucleoprotein filament with ATPγS as binding co-factor, in mixture of 1 μM SSB and 1 μM RecA, 1 mM ATPγS in standard reaction solution. (C). force assisted re-polymerization of partially de-polymerized RecA nucleoprotein filament during cycles of force-jumping between a higher force of ∼30 pN and a lower force of ∼3 pN.

4.3.3 SSB inhibits nucleation and polymerization of RecA on ssDNA by outcompeting RecA on binding to ssDNA binding sites

Having characterized the force responses of RecA-ssDNA nucleoprotein filament and SSB-ssDNA nucleoprotein array, next, we investigated how SSB and RecA proteins compete with each other on ssDNA. The physiological concentration of SSB is in the range of a few hundred nM [140]. Protein mixtures of 1 μM RecA together with either 100 nM or 1 μM SSB in standard RecA reaction solution with 1 mM ATP, 1 ×ATP regeneration system were introduced to a naked ssDNA tether.

Interestingly, Figure 4.5 shows that in the presence of 100 nM or 1 μM SSB

and 1 μM RecA, the RecA filament cannot form at any forces. Moreover, The resulting force-extension curves are almost identical to those in the presence of only SSB without RecA (Figure 4.5A). Together, these findings indicate that SSB outcompetes RecA for binding sites on ssDNA. In addition, since force is known to facilitate the formation of RecA nucleoprotein filament, to determine whether RecA polymerization can take place in the assistance of large force over a longer time scale, the tether was held at ∼37 pN for over 80 seconds, but RecA polymerization did not occur at the high force in our time scale (Figure 4.5B&C). In a word, these results demonstrate that the SSB completely inhibits the nucleation and polymerization of RecA on ssDNA by outcompeting the RecA on binding to ssDNA and form SSB-ssDNA nucleoprotein array in our solution conditions. Note that, we also demonstrated that in a lower pH (pH 6.2) solution, the RecA is able to nucleate and polymerize on the SSB-ssDNA nucleoprotein array (Figure 3.4C or 6.4C), which may be resulted from the dimerization of RecA at low pH [128, 141].

4.3.4 Dynamics of de-polymerization and re-polymerization of pre-formed RecA nucleoprotein filament regulated by ATP-hydrolysis, SSB and force

Having demonstrated that SSB-ssDNA nucleoprotein array formed during competitive binding of SSB and RecA to ssDNA inhibits the nucleation and polymerization of RecA nucleoprotein filament, in this section, we further to investigate that how the SSB affect the stability of pre-formed RecA nucleoprotein filament (Figure 4.6).

To begin with, a fully polymerized RecA-ssDNA nucleoprotein filament was formed by holding an ssDNA tether in the presence of 1 μM RecA and 1 mM ATP, 1 × ATP regeneration system at ∼30 pN. Then, a mixture of 1 μM RecA, 1 μM SSB, and 1 mM ATP, 1 × ATP regeneration system was introduced. Subsequently, the force was decreased to ∼2.8 pN (blue in Figure 4.6A). An initial lag phase of ∼ 600 seconds was observed. Within the lag phase, the filament extension fluctuated around a constant average of ∼ 290 nm, indicating that the RecA nucleoprotein filament stays in a nearly steady state. Remarkably, dynamic de-

polymerization of RecA filament indicated by a rapid extension decreasing phase where the extension of ssDNA tether decreased ∼70 nm within ∼200 seconds was observed after the lag phase. Since RecA nucleoprotein filament in the same solution without SSB protein is stable at such forces, the de-polymerization of RecA filament is dependent on the presence of SSB. Repeating experiments on different ssDNA tethers on the SSB dependent RecA filament de-polymerization were performed (for instance, orange in Figure 4.6A). Similarly, a rapid extension decreasing phase following a lag phase was observed. However, the two steps might be highly stochastic, suggested by that both the duration of the lag phase and the speed of de-polymerization in the extension decreasing phase were significantly different from each other. In addition, similar SSB-dependent de-polymerization of pre-formed RecA nucleoprotein filament was also observed at the lower SSB concentration of 100 nM [109].

Next, we examine the role of ATP hydrolysis in the SSB dependent RecA nucleoprotein filament de-polymerization. Similar experiments with 1 mM ATPγS ascertaining the role of ATP-hydrolysis in the formation of RecA nucleoprotein filament were performed. A mixture of 1 μM RecA, 1 mM ATPγS, and 1 μM SSB was subsequently introduced to a pre-formed fully polymerized RecA-ssDNA nucleoprotein filament tether formed with ATPγS. Then, the extension of the filament was monitored for $>$ 2000 seconds at a force of $\sim$ 2.8 pN (Figure 4.6B). Over the time, no extension decrease was observed, indicating that the de-polymerization of RecA nucleoprotein filament at low force also depends on ATP hydrolysis besides the presence of SSB.

Furthermore, an interesting question is that whether a higher force can lead to the re-polymerization of the partially de-polymerized RecA filament, since force is known to play a positive role on the formation of RecA nucleoprotein filament. To test it, the partially de-polymerized RecA nucleoprotein filament (which is the same ssDNA tether as shown in blue in Figure 4.6A), cycles of force-jumping between a high force ($\sim$ 30 pN) and a low force ($\sim$ 2.8 pN) were repeated multiple times (Figure 4.6C). The tether was held for 5 seconds at each force. Clearly, the ssDNA at low force became longer after each high-force-holding in most of the cycles, demonstrating that a partially de-polymerized RecA filament

at low force are able to re-polymerized at higher forces. In summary, the data in this section demonstrate that a pre-formed RecA nucleoprotein filament de-polymerizes at low force in a SSB and ATP-hydrolysis dependent manner; the partially de-polymerized RecA filament can re-polymerize assisted by high force, even in the presence of SSB.

4.4 Discussion

In this chapter, we studied the regulatory effects of the formation and stability of RecA-ssDNA nucleoprotein filament by SSB, ATP hydrolysis and mechanical force. Based on the data, we proposed a model for this force, ATP-hydrolysis and SSB dependent regulation of RecA nucleoprotein filament (Figure 4.7). In this model, at pH 7.4 and 24 ^{o}C, in solution with >100 nM SSB concentrations and 1 μM RecA, SSB tetramers outcompete RecA binding to ssDNA, and form SSB-ssDNA nucleoprotein array. This tightly packed nucleoprotein array inhibits the nucleation and polymerization of RecA on ssDNA, by imposing an energy barrier. At the same environmental condition (pH 7.4 and 24 ^{o}C, with >100 nM SSB concentrations and 1 μM RecA in solution), a pre-formed RecA nucleoprotein filament de-polymerizes in an ATP-hydrolysis and SSB dependent manner at low force of $\sim$3 pN. Moreover, the partially de-polymerized RecA filament re-polymerizes at high force >$\sim$20 pN, which is in physiological range. Hence, the study provides important mechanistic insights how RecA nucleation and polymerization might take place and the regulation of the stability of RecA nucleoprotein filament by SSB and mechanical force.

Firstly, we show that > 100 nM SSB outcompetes RecA for binding sites on ssDNA, inhibiting the formation of RecA nucleoprotein filament. The initial step of the formation of RecA nucleoprotein filament involves the formation of a stable RecA nucleation site which comprises $\sim$ 3-5 RecA monomers, corresponding to $\sim$ 9-18 nt ssDNA [10]. Since there is no evidence for direct interaction between RecA and SSB, a reasonable speculation is that the high concentration SSB proteins bind to ssDNA much faster than RecA and occupy the binding site, therefore, depletes the naked ssDNA for RecA nucleation. Secondly, a pre-formed RecA nucleoprotein filament de-polymerized at low forces in the presence of >100 nM SSB, 1 μM RecA and ATP-hydrolysis. Previous studies demonstrated that ATP in RecA nucleoprotein filament hydrolyzes primarily as a wave in the 5$'$-to-3$'$ on ssDNA [19]. Since the RecA filament bound with ADP is unstable, the directional ATP-hydrolysis causes the dissociation of RecA from ssDNA from the 5$'$ end, thereby vacating areas of naked ssDNA for SSB binding [19].

The binding of SSB at the naked area from 5′ end blocks re-polymerization of RecA at low force, therefore leading to net de-polymerization of RecA filament. Consistently, without ATP-hydrolysis or without SSB, the RecA nucleoprotein filament is stable at low force.

On the other hand, interpretations of the re-polymerization of RecA filament assisted by high force seem to be complicated. Three possibilities are proposed and discussed: The first possibility is that RecA is able to invade and nucleate into the SSB-ssDNA nucleoprotein array at high force, leading to new polymerization in the 5′-to-3′ direction. However, this possibility has been disproved by our results that SSB-ssDNA nucleoprotein array completely inhibit the nucleation of RecA onto ssDNA at the both low and high force ranges in our experimental conditions. Therefore, the re-polymerization of RecA filament requires a pre-existing nucleation site (i.e., the partially de-polymerized RecA filament). Consistently, an earlier study reported that a pre-formed RecA nucleation cluster is able to displace a single SSB bound to poly-dT ssDNA [24]. The second possibility is that initially RecA filament de-polymerizes from the 3' end, hence the re-polymerization of RecA filament still takes place in the 5′-to-3′ direction. However, extensive previous studies have demonstrated that RecA filament de-polymerizes occurs at the 5′ end corresponding to 5′-to-3′ directional ATP-hydrolysis wave [10, 19].

The most likely possibility is that RecA filament de-polymerizes primarily occurs at the 5′ end, whereas the re-polymerization takes place from the 5′ end of the remaining partially RecA filament in a direction of 3′ to 5′. Although this third interpretation is conflicted the widely accepted unidirectional polymerization of RecA in a 5′ to 3′ direction, there is no experiments disproved the existence of 3′ to 5′ polymerization. While only 5′ to 3′ polymerization of RecA has been observed in most previous experiments, it is likely that the orthodox 5′ to 3′ direction polymerization occurs much faster than the reverse 3′ to 5′ polymerization so that 3′ to 5′ polymerization was not able to be detected. Consistently, in our experimental condition, the 5′-to-3′ polymerization has been blocked by the SSB-ssDNA array at the 5′ end area of ssDNA, the slower 3′ to 5′ polymerization therefore takes place. In addition, this interpretation is also

consistent with two previous experiments based on single-molecule fluorescence resonance energy transfer (FRET) assay of the RecA dynamics at the 5′ end and the measurements of the lifetime of synaptic intermediates during homologous searching [24, 142]. Moreover, a contemporaneous microscopy imaging study using fluorescence labeled RecA and SSB demonstrated that RecA is able to slowly nucleate and polymerize on SSB bound long ssDNA from both 5′ to 3′ and 3′-to-5′ directions at pH 6.5 [128]. In that study, higher pH (pH 7.5 or 8.0) highly repressed the nucleation and polymerization. These observations are also consistent with our results. In our study (pH 7.4), the re-polymerization of RecA filament is assisted by high force. Studies have demonstrated that force can facilitate the formation of RecA filament and slow down dissociation of RecA from the filament in the absence of SSB.

Although it is still unclear about the mechanisms of RecA polymerization facilitated by large force, it can be quantitatively while preliminarily understood based on mechanical differences between the RecA nucleoprotein filaments and SSB tetramers nucleoprotein array. The ssDNA within the RecA nucleoprotein filament is stretched and elongated compared to both naked ssDNA and SSB nucleoprotein filament, thereby the work done by force pre-positions the ssDNA for RecA polymerization. The polymerization energy of RecA combines with the energy supplied by force, which may exceed the energy needed to displace SSB tetramers from ssDNA. Moreover, force on ssDNA reduces the dissociation rate of RecA from ssDNA which also adds to stabilize the RecA filament at high force.

To sum up, our results in this chapter elucidates how formation (nucleation, polymerization) and stability of RecA nucleoprotein filaments are regulated by SSB and ATP-hydrolysis and force. Our work underscores in the importance of SSB nucleoprotein array and mechanical force on the competitive regulatory mechanisms.

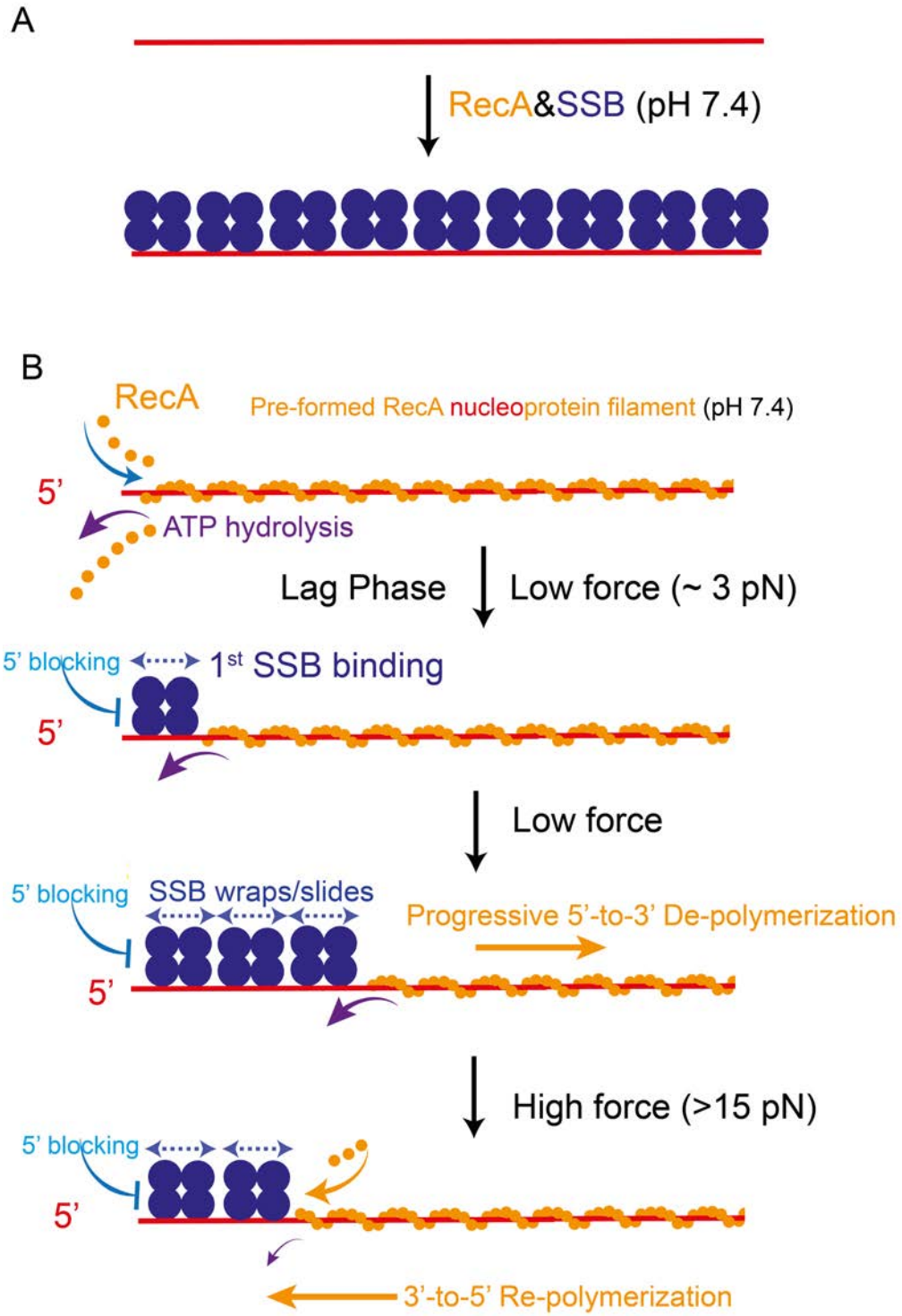

Figure 4.7: **Sketch of SSB, ATP-hydrolysis, and force dependent dynamics of RecA-ssDNA nucleoprotein filament**. (A). More than 100 nM SSB outcompetes RecA (1μM) at pH 7.4, resulting in the formation of a pure SSB-ssDNA nucleoprotein array, which inhibits nucleation and polymerization of RecA nucleoprotein filament. (B) In >100 nM SSB and 1 μM RecA, net depolymerization of a pre-formed RecA nucleoprotein filament occurs at low force in an ATP hydrolysis- and SSB-dependent manner. Increasing force results in re-polymerization of RecA nucleoprotein filament likely with the direction from $3'$ to $5'$ of the ssDNA. Note that the geometric objects indicating ssDNA, RecA and SSB do not represent their real conformations.

Chapter 5

Antagonizing effect of machanical force on the inhibitory actions of RecX on RecA nucleoprotein filaments stability in *M. tuberculosis*

5.1 Chapter summary

RecA nucleoprotein filament formed on ssDNA produced during bacterial recombinational DNA repair requires tightly regulated assembly and disassembly, which involves a variety of mediating proteins. Among them, the RecX protein plays a crucial inhibitory role in the formation and stability of RecA nucleoprotein filaments. Furthermore, RecA nucleoprotein filaments are likely subject to tension, as during homologous searching, the two broken DNA ends are tethered. However, the interplay between RecX and force on formation and stability of RecA nucleoprotein filaments has not yet been investigated. In this chapter, using the new platform for studies of ssDNA-processing proteins on single ssDNA by magnetic tweezers, we found that *M. tuberculosis* (Mt) RecX catalyzes stepwise net de-polymerization of preformed MtRecA filament in the presence of

ATP-hydrolysis at low forces (< 7 pN), which can be reasonably explained by a $3'$-capping model previously proposed. However, interestingly, larger forces applied on ssDNA can antagonize the inhibitory actions of MtRecX on RecA nucleoprotein filament; with the assistance of larger force, a partially de-polymerized MtRecA filament could re-polymerize in the presence of MtRecX, which cannot be explained by any previous models. As the forces that can antagonize the inhibitory actions of RecX are in a physiological range; our findings highlight a broad potential mechanosensitive regulation during homologous recombination.*

5.2 Introduction

DNA damages, which severely affect the stability and integrity of the genome, have to be repaired efficiently *in vivo*. In eubacteria, the RecA protein plays an essential role in homologous recominational DNA repair, by forming a right handed RecA filament on single stranded DNA (ssDNA) to promote the homologous pairing and exchange of DNA strands in the presence of ATP or ATP-analogues and other co-factors [10, 11]. The stability of the RecA filament is dynamically regulated by polymerization and de-polymerization in the presence of ATP-hydrolysis [10, 11].

As introduced in Chapter 1, various proteins are involved in the regulation of the polymerization and de-polymerization of RecA to avoid either insufficient or unlimited formation of the RecA filament [10, 11]. Among the mediating proteins, RecX protein strongly inhibits RecA filament nucleation and polymerization [10], however, much less is understood about the regulatory mechanisms of RecX on RecA nucleoprotein filament.

E.coli (Ec) RecX was shown to promote EcRecA nucleoprotein filaments disassembly from circular DNA in an ATP-hydrolysis dependent manner at substoichiometric concentrations. A $3'$ capping model where EcRecX blocks the growing end ($3'$ end) of the EcRecA filament, resulting in net EcRecA disassembly were proposed to explained the observations [62]. Later it was showed that higher con-

*Note that main contents detailed in this chapter have been published in *Mechanical force antagonizes the inhibitory effects of RecX on RecA filaments formation in M. tuberculosis.* Le S. *et al.*, Nucleic Acids Research 42 (19): 11992-11999 (2014).

centrations of EcRecX resulted in faster de-polymerization of EcRecA filament [64]. Besides, EcRecX is found to bind within the major helical groove in the monomer-monomer interface along the length of the active RecA-ssDNA filament by electron microscopic and X-ray crystallographic studies [58, 64, 143]. Therefore, an additional internal-nicking mechanism where RecX can generate nicks inside the EcRecA filament, resulting in more de-polymerization ends, was suggested [64]. In addition, another additional model that RecX may facilitate RecA filament end de-polymerization can also explain the observation of RecX concentration dependent RecA filament de-polymerization [144]. While aforementioned studies have been focused on RecX promoted RecA de-polymerization, much fewer attentions have been paid to the antagonizing factors of RecX inhibitory effects.

As introduced in Chapter 1, recent *in vivo* studies show that DNA during humongous recombinational repair is likely subject to tension [75, 76]. Furthermore, considering the markedly different force responses between the soft ssDNA and the rigid RecA nucleoprotein filament [21, 81, 109], one can intuitively anticipate the potential regulatory effect on RecA filament stability. However, it has yet remained poorly investigated of the potential regulatory role of force and its interplay with other cellular factors on the regulation of RecA filament.

In this Chapter, aiming to understand the effect of force on the RecX-mediated regulation of RecA filament stability, I directly observed and quantified RecX-mediated dynamics of individual MtRecA filaments regulated by mechanical forces within physiological ranges, using the new platform introduced in Chapter 2.

5.3 Results

5.3.1 Force-response of MtRecA nucleoprotein filament formed on ssDNA

As introduced in Chapter 1, ssDNA is a very flexible polymer, whose bending persistence is estimated to be $\sim$1 nm [81]. In distinct contract, the RecA nucleoprotein filament formed on ssDNA is very rigid with a bending persistence

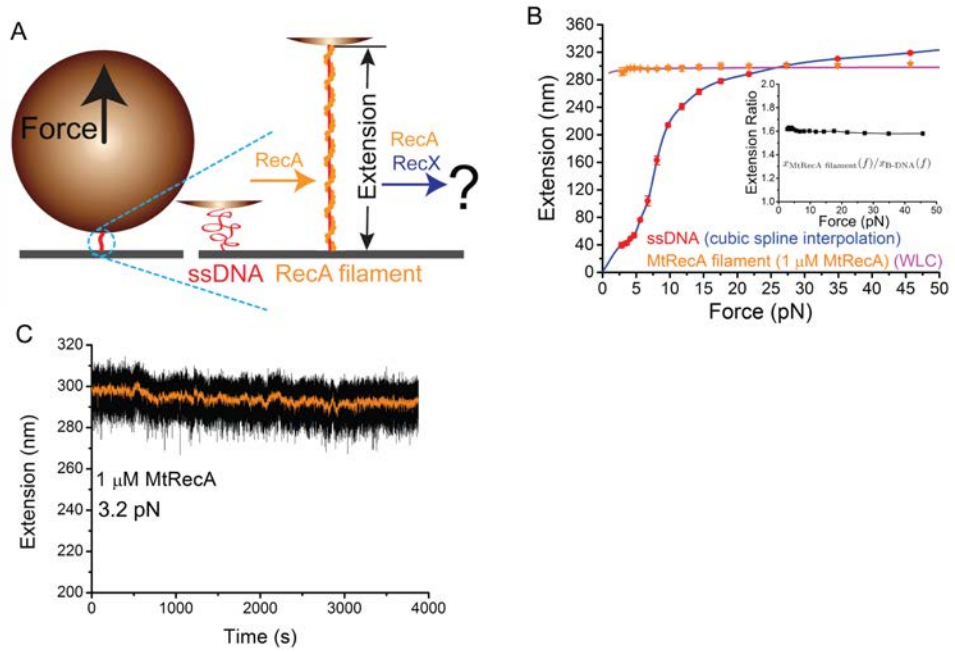

Figure 5.1: **Distinct Force Responses of ssDNA and MtRecA-ssDNA nucleoprotein filament**. (A).Schematic of the experiment. A short ssDNA tethered between a cover glass surface and a paramagnetic bead is subject to forces applied to the bead in the upward direction (left panel). Middle and right panels show a coiled ssDNA (red), and a rigid, extended RecA filament (orange). The effects of MtRecX were explored by the changes in extension of preformed MtRecA filament after induction of MtRecX. (B). Typical force-extension curves of a naked 576 nt ssDNA (red circles: experimental data; blue line: cubic spline interpolation of data) in the standard assay buffer and a fully formed MtRecA filament (orange tri-angles: experimental data; magenta line: WLC fitting) on the same ssDNA with 1 μM MtRecA in solution. Inset shows the mea- sured extension of fully formed MtRecA filament divided by the theoretical extension of a B-form dsDNA of equal number of base pairs (576-bp). The error bars are standard deviations (s.d.) obtained from repeating measure- ments (>3 times) of the same DNA tether under each condition. (C). Long time trace of the extension of a preformed MtRecA filament (the same filament as in (B)) recorded at ~3.2 pN with 1 μM MtRecA in solution.

length of $\sim$ 1000 nm [21]. This distinctive difference of bending rigidity leads to distinctive force responses, which can be used as a quantitative tool for investigating the dynamics and stability of the RecA nucleoprotein filaments regulated by mediating proteins in single-DNA stretching experiments as sketched in Figure 5.1A.

Figure 5.1B shows force-extension curves of a naked ssDNA (576-nt) and MtRecA nucleoprotein filament formed on this ssDNA in our standard RecA reaction solution: 20 mM Tris (pH 7.4), 50 mM KCl, 10 mM $MgCl_2$, 1 mM ATP and 1×ATP regenerating system at 23 °C. The extension and error bar at each

force were averaged mean value and standard deviations (s.d.) from multiple (>3) cycles of force-increasing and force-decreasing scans. The DNA was held for 5-seconds at each force during the scans.

Due to the distinct micro-mechanical properties of ssDNA and MtRecA filament [21, 81], the extension of MtRecA nucleoprotein filament is longer than that of the naked ssDNA for forces $<\sim$ 30 pN, with larger difference at smaller forces. At forces $>$ 30 pN, the ssDNA is much more stretched and exceeds the extension of MtRecA filament. In addition, the extension of MtRecA filament is $\sim$ 1.5-1.6 times of that of dsDNA with the same bases (Figure 4.2B inset). Furthermore, MtRecA nucleoprotein filaments are stable in the presence of 1 μM MtRecA in solution over a wide force range: no significant net de-polymerization of MtRecA filament was observed over $>$ 1000 sec in forces from 1-90 pN (Figure 5.1C is a typical example of the extension time course of MtRecA filament at $\sim$ 3 pN).

5.3.2 MtRecX catalyzes net stepwise de-polymerization of MtRecA filament in an ATP-hydrolysis dependent manner at low forces

To investigate how MtRecX regulates the dynamics of individual MtRecA nucleoprotein filament, we introduced mixtures of 1 μM RecA and 80-1000 nM RecX in standard reaction solution to pre-formed MtRecA filament at low forces $\sim$3 pN (Figure 5.2). At low concentrations of MtRecX (80 nM), MtRecX catalyzes a slow progressive net de-polymerization of MtRecA filament, indicated by net DNA extension reductions with an speed of 0.038 (mean) $\pm$ 0.019 (s.d.) nm/s (Figure 5.2A). Occasional MtRecA re-assembly events (arrows, Figure 2A) were also observed in addition to the dominant de-polymerization process. Besides, the disassembly process, although noisy, appeared stepwise overall.

As shown in Figure 5.2B&C, similar stepwise MtRecX catalyzed de-polymerizations of MtRecA filaments were observed at higher concentrations (400 nM or 1 μM) of MtRecX with faster overall de-polymerization speeds (0.15 $\pm$ 0.075 nm/s for 400 nM and 0.27$\pm$0.31 nm/s for 1 μM). In addition, obvious large-step re-assembly event was absent.

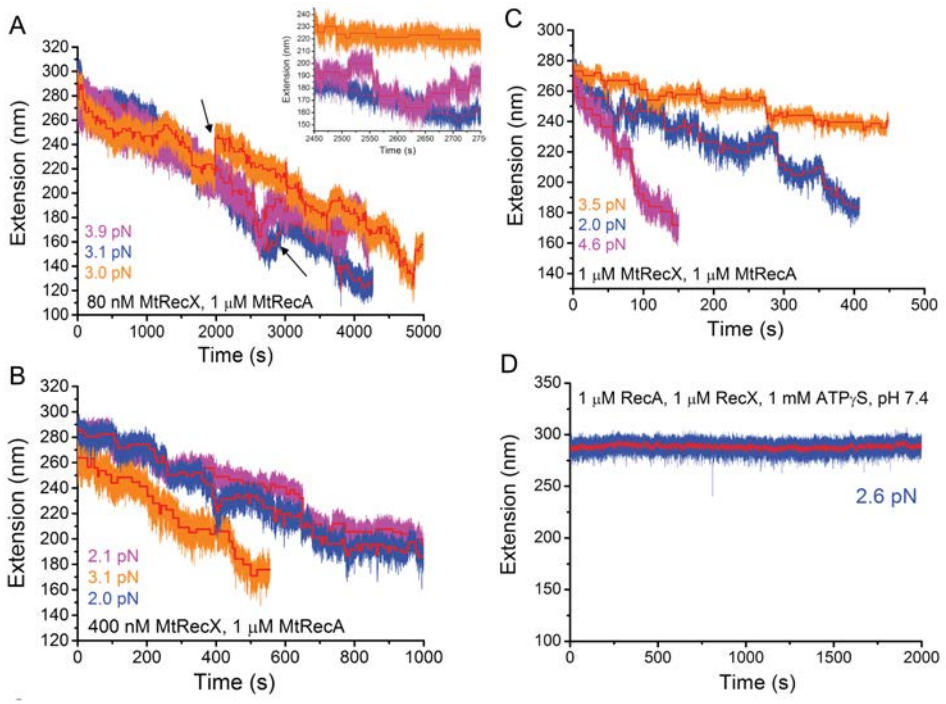

Figure 5.2: **De-polymerization of preformed MtRecA filaments at different MtRecX concentrations**. (A). Time traces of the extension obtained on three independent preformed MtRecA filaments formed on three ssDNA tethers (indicated by different colors) after addition of 80 nM MtRecX at forces of 2-4 pN. Inset shows steps in zoom-in time traces. (B-C) show extension time traces of three independent preformed MtRecA filaments at 400 nM MtRecX (B) and 1 μM MtRecX (C) at forces of 2-4 pN. The red lines in (A-C) show stepwise de-polymerization and re-polymerization obtained from a step finding algorithm. (D) A long extension time trace of a preformed MtRecA in 1 μM MtRecX, 1 μM MtRecA, 1 mM ATPγS (other conditions remained the same), where no net de-polymerization of MtRecA filament occurs over the experimental time scale.

Importantly, when we replaced ATP with its non-hydrolysable homologue, ATPγS, while other conditions remained unchanged, MtRecA nucleoprotein filament was stable in the presence of MtRecX over a long period of time (>2000 seconds) (Figure 5.2D). This result indicates that MtRecX mediated de-polymerization of MtRecA filament is also dependent on ATP hydrolysis. In addition, a lower pH (6.1) significantly slowed down (0.053 $\pm$ 0.015 nm/s) the overall MtRecX-mediated de-polymerization speed of MtRecA filament compared to that at a higher pH (7.4), (0.27 $\pm$ 0.31 nm/s). It can be expalined by the fact that lower pH stabilizes RecA filaments, likely due to dimerization of RecA at lower pH [128].

5.3.3 Kinetics of MtRecX mediated net de-polymerization of preformed MtRecA filaments

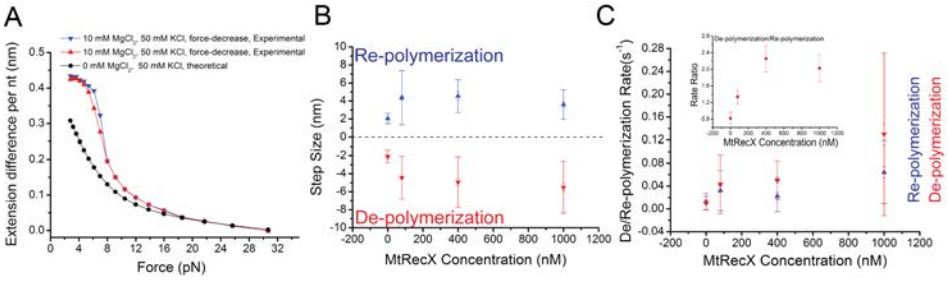

Figure 5.3: **Step sizes and rates of de-polymerization or re-polymerization of MtRecA filaments in different concentrations of MtRecX.**. (A) Extension difference of MtRecA filament and ssDNA. The black circles are estimated by theoretical models, using naked ssDNA force response with 50 mM KCl and MtRecA filament force response based on the WLC model with a bending persistence length of 1 μm and an estimated contour length of 0.51 nm/nt. The blue tri-angles are data measured in experiments in 50 mM KCl and 10 mM $MgCl_2$. The data are averaged from multiple experiments ($N = 3$). In the force range of 2-4 pN, extension change is $\sim$0.3 nm/nt based on theoretical estimation and $\sim$0.4 nm/nt based on experimental data. The difference is likely due to the effect of magnesium, which is not considered in the theoretical formula of naked ssDNA force-response. (B) Step sizes of de-polymerization (red down-triangles) and re-polymerization (blue up-triangles). (C) Rates of de-polymerization (red down-triangles) and re-polymerization (blue up-triangles). The error bars with wider and narrower caps indicate standard deviations (s.d.) and standard errors (s.e.) obtained from multiple (>3) independent experiments, respectively. Insets in (C) show the ratio of the rates of de-polymerization and re-polymerization, error bars were obtained from the standard deviations and standard errors of the rates of re-polymerization and de-polymerization through error propagation.

In addition, in multiple, independent experiments using the same MtRecX concentrations, different speeds of de-polymerizations were observed (Figure 5.2A-C). Though there are several possible explanations for this variation, we speculate that it is due to the stochastic nature of the de-polymerization and re-polymerization processes from a single de-polymerizing end, which is supported by kinetics simulations

We analyzed the de-polymerization and re-polymerization steps in each time trace. A step finding algorithm was employed to extract negative steps (de-polymerization) and positive steps (re-polymerization) from the data. Detection of such stepwise signal was automated by detecting abrupt decreases or increases in extension using a method similar to that developed by Cui et al.

[118] (Methods–Step finding algorithm). Red lines in the Figure5.2A-C show the stepwise time traces identified by this algorithm. Note as this method cannot detect steps smaller than the noise level, the numbers of potential steps as well as the transition rates are likely underestimated.

In the absence of MtRecX, ∼2 nm steps were detected for ATP-dependent MtRecA dissociation and re-association, which are larger than the expected monomer dissociation and re-association steps. Note a RecA monomer is known to associate three consecutive nucleotides of ssDNA, corresponding to ∼ 0.9-1.2 nm ssDNA extension changes at 2-4 pN per MtRecA monomer dissociation or re-association (Figure 5.3A). Such small steps are below the noise level (∼ 2 nm for smoothed data) of the extension fluctuation; therefore they cannot be detected. The averages, standard deviations, and standard errors of the step sizes and the rates of de-polymerization and re-polymerization are shown in Figure 5.3B&C.

In the presence of MtRecX in the range of 80 nM to 1 ∼M and force range of 2-4 pN, the average step sizes for both de-polymerization and re-polymerization are around 4 nm, roughly corresponds to release of 9 nt of ssDNA, or equivalently simultaneous dissociation of three RecA monomers (i.e., half helical turn of RecA filament) (Figure 5.3A). Although the causes of this step wise de-polymerization is unclear, it may be related to cooperative ATP hydrolysis in the RecA filament reported [19]. Over the range of MtRecX concentrations, the step sizes do not depend on the MtRecX concentration. In contrast, the kinetics of the de-polymerization is dependent on the MtRecX concentration. From 80 nM to 1 μM MtRecX, the average de-polymerization rate increases by ∼3-folds, and the ratio of de-polymerization rate over the re-polymerization rate increases by $>$ 2-folds (Figure 5.3., inset). These results suggest that MtRecX facilitates the rate of de-polymerization in addition to $3'$ capping [62]. The large standard deviations suggest a highly stochastic process of de-polymerization.

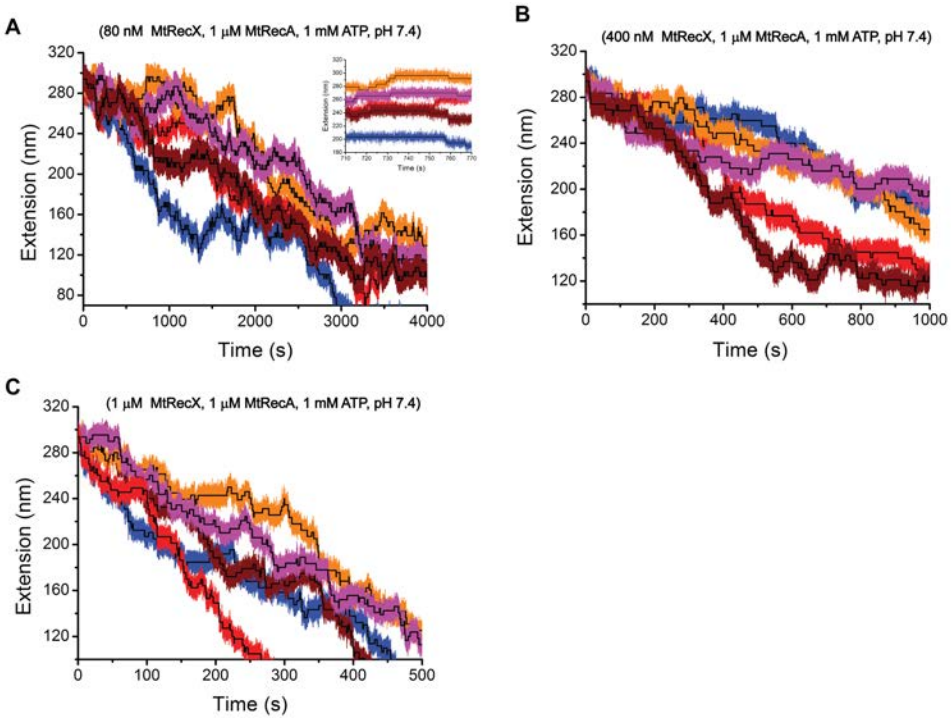

Figure 5.4: **Simulated extension evolutions of MtRecA filaments in the presence of MtRecX**.(A-C). The simulated extension time traces of MtRecA filaments using the kinetics simulation algorithm based on the average values of rates and step sizes of de-polymerization or re-polymerization in Figure S4. Black lines, simulated extension time traces. Gaussian noises with a standard deviation of 4 nm (close to the global standard deviation of the raw time traces in our experiments) are added to each line to mimic the experimental time traces. Five independent simulations at each condition were plotted.

5.3.4 Large variations in MtRecA de-polymerization speeds can be explained by stochastic de-polymerization and re-polymerization kinetics

In multiple, independent experiments using the same MtRecX concentrations, we observed different speeds of de-polymerizations (Figure 5.2). Though there are several possible explanations for this variation, we reason it is likely due to the stochastic nature of the de-polymerization and re-polymerization processes. Each extension time trace can be understood by a one-dimensional random walk process, with an average de-polymerization rate of k_{off} and step size l_{off}, as well as a re-polymerization rate of k_{on} and step size of l_{on}. These parameters were estimated in experiments for each MtRecX concentration (Figure 5.3). Based on the averages of these parameters, we simulated the extension evolution of a pre-formed MtRecX filament using kinetics simulation algorithm (Methods–kinetics

simulation).

Using the averaged values of the kinetic and step size parameters estimated, five independent simulated time traces were generated for each corresponding MtRecX concentration over similar time scales (Figure 5.4). These stepwise time traces were superimposed with Gaussian noise using a standard deviation of $\sim$ 4 nm, which is similar to the noise level of the raw time traces obtained in our experiments. We found that the simulated time traces were consistent with the corresponding experimental time traces. The simulated and experimental traces exhibit similar overall net de-polymerization speeds and large variations from one experiment to another. Hence, the simulation results support our hypothesis that the variations of the extension time traces for multiple, independent experiments using the same MtRecX concentration can be explained by the stochastic nature of MtRecA de-polymerization and re-polymerization steps of a single filament. RecA polymerization is intrinsically stochastic, as it involves diffusion of free RecA proteins to the filament. De-polymerization is facilitated by ATP hydrolysis mainly taking place at the $5'$ end. ATP turnover is a stochastic process, and spontaneous dissociation of RecA after ATP hydrolysis should also be a stochastic process. All these may potentially contribute to the stochastic nature of MtRecA de-polymerization process as observed in experiments and simulation. These results do not exclude the possibility that MtRecX may create limited nicks inside the MtRecA filament.

5.3.5 $5'$-to-$3'$ polymerization of MtRecA filament revealed by its re-polymerization on MtSSB bound ssDNA assisted by force

$3'$-to-$5'$ reverse polymerization of RecA in *E.coli.* has been demonstrated by several groups including my previous study detailed in Chapter 4 [24, 109, 128, 142]. Here, the possible existence of the $3'$-to-$5'$ reverse polymerization of RecA in *M. tuberculosis* was examined. In reaction solution containing 20 mM Tris (pH 7.4), 50 mM KCl, 10 mM $MgCl_2$, 1 mM ATP, at 23°C (the same environmental condition as in main text), an MtSSB protein array was formed on ssDNA in the MtSSB concentration from 0.25 nM to 250 nM. In this concentration range,

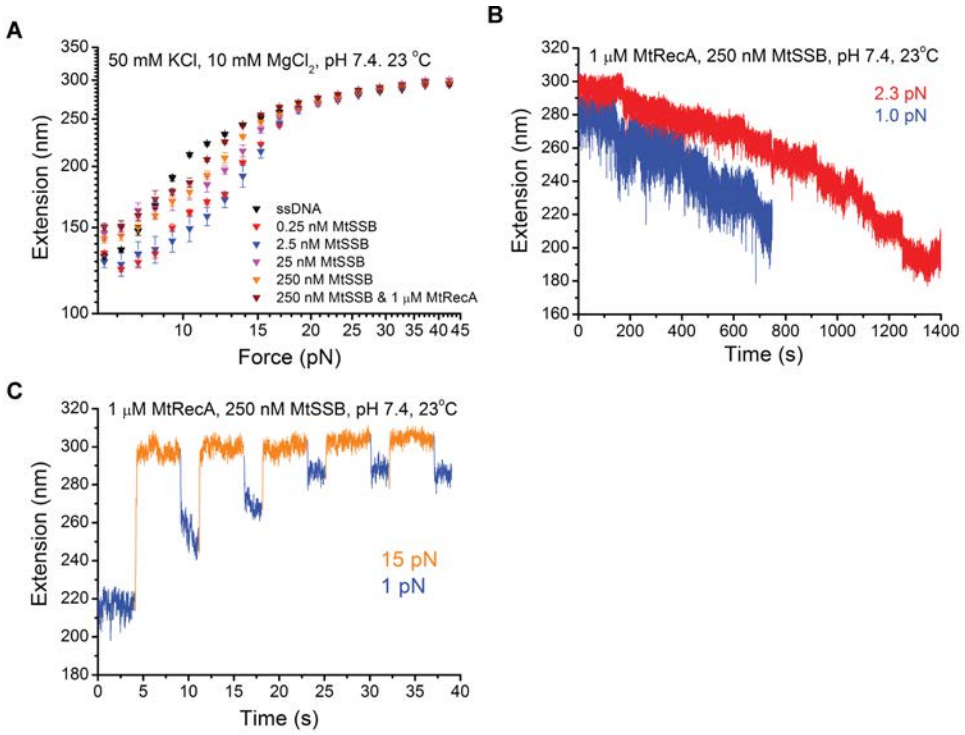

Figure 5.5: **Force assisted re-polymerization of MtRecA filament in the presence MtSSB**. (A) Force extension curves of ssDNA with different concentrations of MtSSB (0.25 nM & 250 nM), or mixture of MtSSB and MtRecA, indicated by different colors. Symbols and error bars represent the average values and standard deviations of multiple (>3) repeating force scans at each condition. (B) De-polymerization of two preformed MtRecA filaments after a mixture of 250 nM MtSSB and 1 μM MtRecA was introduced at low forces <3 pN. (C) Extension evolution of a partially de-polymerized MtRecA filament during cycles of force-jumping between a low force of $\sim$1 pN and a high force of $\sim$15 pN.

MtSSB binding resulted in slight extension reduction in a moderate force range of 5-15 pN, which is expected from ssDNA wrapping around MtSSB tetramers. A mixture of 250 nM MtSSB and 1 μM MtRecA was then introduced in the same buffer solution. The force-extension curve remained nearly unchanged, indicating that MtRecA could not nucleate and polymerize on the MtSSB coated ssDNA under this solution condition (Figure 5.5A).

When an MtRecA filament was pre-formed on ssDNA followed by introduction of the mixture of 250 nM MtSSB and 1 μM MtRecA, net de-polymerization of the MtRecA filament was observed, indicated by progressive extension reduction at a few pN forces (Figure 5.5B). This can be explained by binding of MtSSB to vacated ssDNA at the 5′ end of the MtRecA filament due to ATP-hydrolysis mediated MtRecA disassociation, resulting in net de-polymerization.

Such MtSSB dependent de-polymerization of preformed MtRecA filament should lead to a partition of the ssDNA into MtSSB array at the 5′ side of the ssDNA and the remaining MtRecA filament at the 3′ side of ssDNA.

Before the MtRecA filament was completely de-polymerized, upon switching to high forces, re-polymerization of the MtRecA filament was observed, revealed by extension re-elongation at low forces. One example is shown in Figure 5.5C - switching force between a low force (1 pN) and a high force (15 pN) on a partially de-polymerized MtRecA filament resulted in progressive elongation of the ssDNA extension at 1 pN. As this experiment began with a pre-formed fully polymerized MtRecA filament, there was no space left at the 3′ end for the canonical 5′-to-3′ directional polymerization. The only space on ssDNA available for MtRecA re-polymerization was at the 5′ side which was however occupied by MtSSB. Therefore, we conclude that the force-assisted re-polymerization should take place from the 5′ end of remained MtRecA filament in a reversed 3′-to-5′ direction.

5.3.6 Force facilitates re-polymerization of MtRecA filament, antagonizing the inhibitory effects of MtRecX

We have directly monitored and quantified the dynamics of MtRecX mediated de-polymerization of individual MtRecA filaments at low forces (∼2-3 pN). These results are overall consistent with that previously reported bulk-biochemistry experiments with RecX from across several bacterial species [62, 64, 144]. Next, we explore the potential regulatory roles of force on the MtRecX mediated dynamics of MtRecA filament.

To investigate the regulatory role of force, we monitored the extension evolutions of partially de-polymerized MtRecA filaments in the presence of mixture of MtRecX and MtRecA, during jumps between lower forces (∼2.0 pN) and higher forces (>17 pN) (Figure 5.2A-B). At each force, the DNA was held for 5 seconds. Interestingly, the extensions of the DNA returned to that when it is coated with fully polymeriezed MtRecA filament DNA after several cycles of force jumping, indicating that the partially de-polymerized MtRecA filaments re-polymerized at higher forces in the presence of MtRecX with all concentrations we tested (80

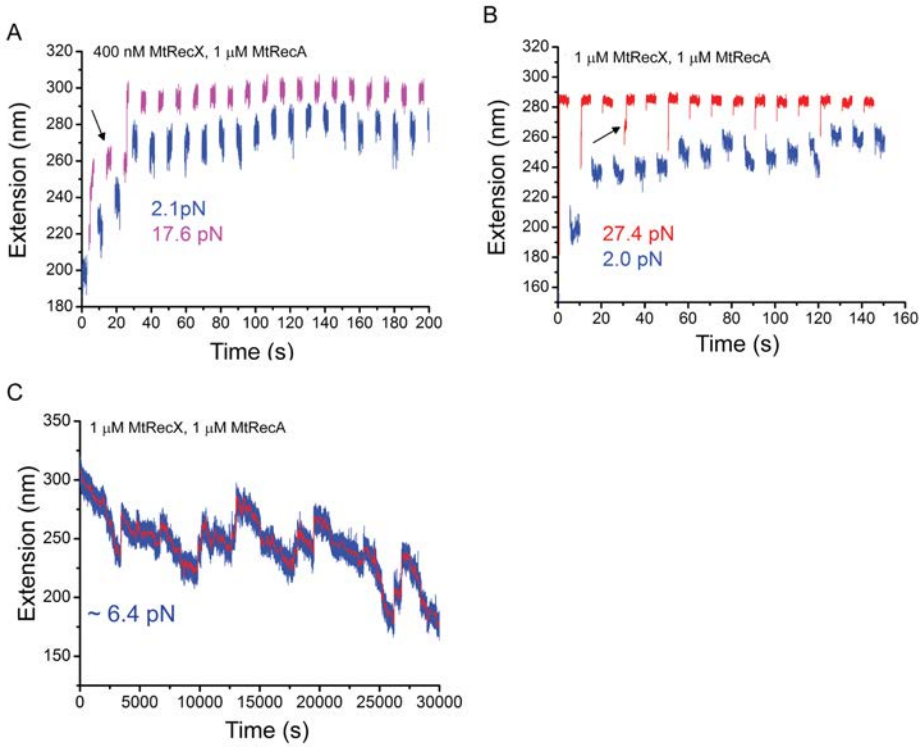

Figure 5.6: **Force assisted re-polymerization of MtRecA filament in the presence of MtRecX.**. (A) Extension evolution of a partially de- polymerized MtRecA filament (the same tether as that indicated by blue color in Figure 5.2B) during quick jumps between two forces of ∼17.6 pN and ∼2.0 pN with 400 nM MtRecX, 1 ∼M MtRecA and 1 mM ATP. (B) Extension evolution of a partially de-polymerized MtRecA filament (the same tether as that indicated by blue in Figure 5.2C) during quick jumps between two forces of ∼27.4 pN and ∼2.0 pN with 1 μM MtRecX, 1 μM MtRecA and 1 mM ATP. In both A and B the partially de-polymerized MtRecA filaments re-polymerized at higher forces (>17 pN) after several force-jump cycles, indicated by the elongated extension when force was jump back to lower forces. Black arrows indicate 'locked' conformation of MtRecA filament with shorter extension. (C) A long extension time trace obtained at ∼7 pN with 1 μM MtRecX, 1 μM MtRecA and 1 mM ATP, showing nearly balanced de-polymerization and re-polymerization $> 30,000$ s.

nM -1 μM) (cites).

Furthermore, we also determined the critical force to be ∼7 pN for the balance of de-polymerization and re-polymerization of the MtRecA filament is nearly reached (Figure 5.2C). In addition, MtRecX also induces EcRecA filament de-polymerization at low forces ∼3 pN, which is also antagonized by higher forces applied on the filament (Figure 5.7A&B). These results suggest that the regulatory effect of force on the RecX mediated dynamics of RecA nucleoprotein filament is likely conserved across bacterial species.

Additionally, we found that the partially de-polymerized MtRecA nucleopro-

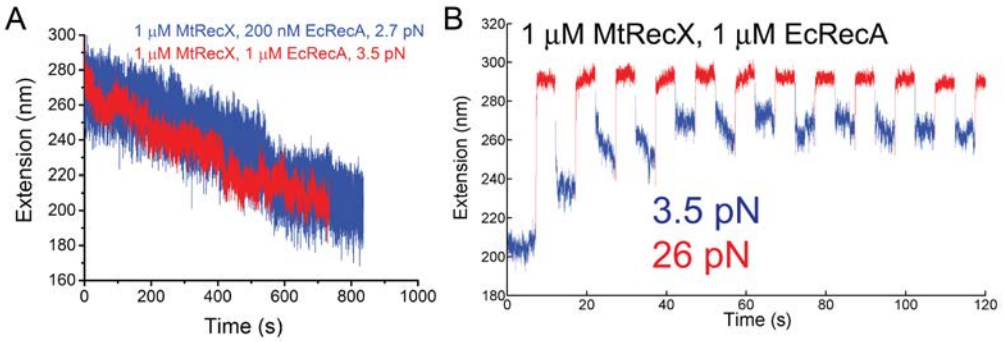

Figure 5.7: **Force dependence of MtRecX mediated EcRecA filament dynamics**. (A). MtRecX (1 μM) promotes net de-polymerization of EcRecA filaments (200 nM and 1 μM tested) at low forces of $\sim$3 pN. (B). Higher forces assist re-polymerization of partially de-polymerized EcRecA filament in the presence of MtRecX.

tein filaments were sometimes mechanically stably 'locked' at a short-extension state at low forces. These locked states could withstand higher forces for several seconds (arrows in Figure 5.2A-B) and then unlocked with sudden extension jumps. While it is unclear about the nature of the locked conformations, we speculated that it might be the bridging between the vacated naked ssDNA region and the remained partial MtRecA filament region through MtRecA secondary binding sites [10].

5.4 Discussion

In this chapter, we investigate the MtRecX-mediated dynamics of MtRecA nucleoprotein filament, and the regulatory role of force on the dynamics. We show that MtRecX catalyzes ATP-hydrolysis dependent, stepwise net de-polymerization of MtRecA filaments at low forces (<7 pN) at single RecA filaments level. These results are consistent with previous ensemble biochemical experiments reporting RecX-mediated net RecA disassembly [62, 64, 144]. Furthermore, we discovered that higher forces can antagonize the inhibitory actions of MtRecX on MtRecA filaments and facilitating re-polymerization of the MtRecX-induced partially de-polymerized MtRecA filament (Figure 5.8). Importantly, this force-facilitated re-polymerization of MtRecA filament in the presence of MtRecX has never been previously reported, and it cannot be explained by previous proposed models. Hence, besides directly probing the dynamics of individual RecA nucleoprotein

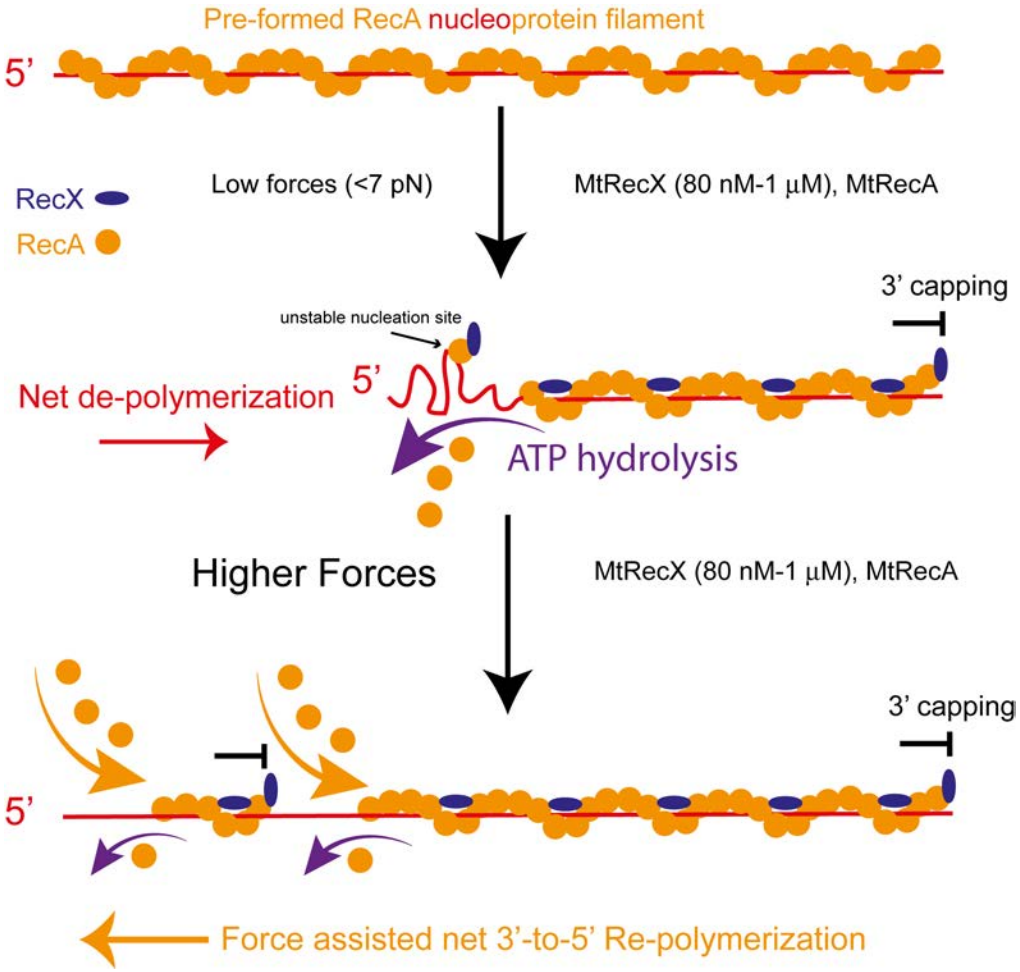

Figure 5.8: **Mechanistic model of the effects of force on MtRecX dependent MtRecA filament dynamics**. A pre-formed MtRecA flament (orange) on ssDNA (red) is capped by MtRecX (blue) at its 3′ end (based on the 3′ capping model proposed by Dress et.al (11)), and are bound with MtRecX at the filament grooves (based on electronic microscopy reconstruction (14)). At low forces (<7 pN according to our measurement), MtRecA dissociates from the ssDNA at the 5′ end that requires ATP hydrolysis, resulting in net de-polymerization. Any potential new nucleation sites formed on the vacated ssDNA are not stable since MtRecX caps them at the 3′ end. At higher forces (>7 pN), due to the stabilizing effect of force on RecA filament, a partially de-polymerized MtRecA filament may re-polymerize from the 5′ end of the remained filament and/or from the new nucleation sites in a force assisted 3′-to-5′ reverse direction.

filaments, these results have uncoverd important new insights regarding how microenvironment such as mechanical force may regulate the actions of RecX-mediated RecA filaments. Previously, a 3′ growing-end capping by RecX model was proposed to explain EcRecX-mediated disassembly of EcRecA filament [62]. In this model, RecX capped the 3′ growing-end of RecA filament, therefore inhibiting polymerization of RecA filament. In addition, the capping of RecX at 3′ end also inhibits formation of stable nucleation of RecA on vacated ssDNA due to

ATP-hydrolysis induced RecA disassembly. The combination of the two effects of 3′-end capping is able to explain the net de-polymerization of RecA filament at the 5′ end induced by RecX from *E. coli.* [64], Neisseria Gonorrhoeae [144] and *M. tuberculosis.* Together, these results suggest that RecX 3′ capping mechanism is likely a universal dominating factor that regulates dynamics of RecA nucleoprotein filament by RecX proteins across bacterial species.

The stochastic variations of the dynamics of RecA filament de-polymerization/re-polymerization processes can also be explained by the 3′ capping mechanism. In our experimental assays, for each RecA filament, it began with a fully polymerized RecA filament, presumably capped by a RecX at the 3′ end accordingly, resulting in the 5′ end, the only end, for RecA de-polymerization due to ATP hydrolysis. During the net de-polymerization process, the position of the 5′ end is governed by ATP-hydrolysis induced RecA de-polymerization and RecA reverse re-polymerization from the 5′ end. The balance of the two competitive processes determined the dynamics of RecA filament. Since both ATP-hydrolysis and spontaneous RecA association/dissociation are stochastic [19], the movement of the de-polymerizing 5′ end can be understood as a one-dimensional random walk along the track of ssDNA template. Consistently, the kinetics simulation predicted extension evolutions of a pre-formed MtRecX filament well mimic that observed in experiments.

Importantly, we found that higher forces (> 7 pN) can antagonize the inhibitory effect of MtRecX on MtRecA filament, and resulting re-polymerization of a partially de-polymerized MtRecA filament in the presence of MtRecX (80 nM to 1 μM). This new finding is unexpected and interesting, since it seems to suggest that the dominating 3′ capping mechanism at low forces cannot apply to the situation when sufficiently large force is applied on MtRecA filament. Mechanosensitive factors have to be considered in order to explain this force-assisted re-polymerization effect. While there are several possibilities, a likely explanation is a potential force-facilitated reversed (3′-to-5′) re-polymerization of partially de-polymerized MtRecA filament.

In this force-assisted reverse polymerization of RecA filament scenario, although MtRecX caps the 3′ growing end of RecA filament, the 5′ end of the

filament can still polymerize in a reverse (3′-to-5′) direction, which is assisted by force and outcompetes de-polymerization rate at 5′ end (see sketch in Figure 5.8). While a canonical 5′-to-3′ polymerization of RecA filament has been widely known, our previous studies (detailed in Chapter 3), as well as some other studies [24, 109, 128, 142], have demonstrated the existence of reversed 3′-to-5′ polymerization of RecA filament. This reverse polymerization of RecA filament is likely promoted by force, and starts from the end of a partially de-polymerized filament and/or from potential new RecA nucleation sites whose 3′ end is capped by RecX on vacated ssDNA. Besides, we note that if the 3′ capping capability of MtRecX is reduced by mechanical force the canonical 5′-to-3′ polymerization might also take place from new nucleation sites on vacated ssDNA.

To sum up, together with previous results obtained from EcRecX and NgRecX [62, 64, 144], our results advanced our understanding of the functions of RecX across bacterial species. Particularly, we highlight the importance of mechanical force, which antagonizes the inhibitory effects of MtRecX on the stability of MtRecA filament. Furthermore, due to the ubiquitous presence of force in vivo, and as the effect of force for RecA filament regulation is likely resulted from the distinct micromechanics between naked ssDNA and RecA filament, we underscore a need to further explore the mechanosenstive regulation of homologous recombination reaction and other cellular processes.

5.5 Methods and Materials

Experimental setup and materials–The experiments were performed on the single-ssDNA manipulation platform detailed in Chapter 2. MtRecX and MtRecA proteins were purified as described [57, 145]. The standard assay buffered solution containing 20 mM Tris (pH 7.4), 50 mM KCl, 10 mM $MgCl_2$, 1 mM ATP and 1×ATP regenerating system (note that 1 mM ATP and 1×ATP regenerating system were included to maintain 1mM ATP in solution. In the main text and figures, it is referred as 1 mM ATP for simplicity). All experiments were done at 23 °C.

Step finding algorithm–A step finding algorithm similar to that developed by

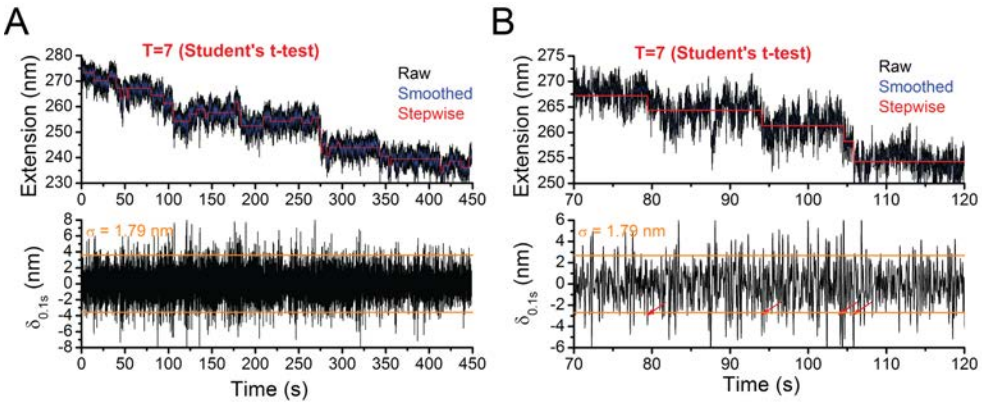

Figure 5.9: **Steps detected from the extension time trace of MtRecA filament in the presence of MtRecX**. (A). The raw data (black) were smoothed by the Savitzky-Golay method within 0.2 sec time window (blue). Trial steps are detected when $\delta_{0.1s}(t) \geq 2\sigma = 3.58$ (orange lines). Only steps with the Student′s T-value greater than 7 were accepted as de-polymerization or re-polymerization steps, and the corresponding stepwise time trace was plotted in red. (B) show the zoom-in of (A) from 70-120 second.

Cui et al. [118] was used to estimate the steps sizes and kinetics of MtRecA filament dynamics. First, the raw data were smoothed by the Savitzky-Golay method within 0.2 sec time window. Then a local extension difference $\delta_{0.1s}(t)$, which is the extension measured at time t minus the extension measured at $t - 0.1s$, is calculated throughout the whole time trace, and a global standard deviation, σ, is obtained. Trial steps are detected when $\delta_{0.1s}(t)$ exceeds 2σ. Typical noise levels in our experiments are in the range of $2\sigma \sim$ 2-4 nm for partially de-polymerized MtRecA filaments depending on the amount of vacated flexible ssDNA during de-polymerization. Only steps with the Student's T-value greater than 7 were accepted as de-polymerization or re-polymerization steps, and the corresponding step size and the time point of the accepted stepwise extension changes were recorded. A typical step finding process was plotted as Figure 5.9.

Kinetics simulation–In the Kinetics simulation, a small time step $\Delta t = 0.01$ s (corresponding to the experimental temporal resolution 100 Hz) was chosen, within which the probabilities for de-polymerization, re-polymerization, and neither de-polymerization nor re-polymerization to occur are $p_{\text{off}}(\Delta t) = k_{\text{off}}\Delta t$, $p_{\text{on}}(\Delta t) = k_{\text{on}}\Delta t$, and $p_{\text{none}}(\Delta t) = 1 - (p_{\text{off}}(\Delta t) + p_{\text{on}}(\Delta t))$, respectively. A uniformly distributed random number $0 < r < 1$ was generated, which is compared

to the above probabilities: de-polymerization is selected if $0 < r < p_{\text{off}}(\Delta t)$; re-polymerization is selected if $p_{\text{off}}(\Delta t) < r < p_{\text{off}}(\Delta t) + p_{\text{on}}(\Delta t)$; and neither de-polymerization nor re-polymerization is selected if $p_{\text{off}}(\Delta t) + p_{\text{on}}(\Delta t) < r < 1$.

Chapter 6

Dynamics and Regulation of RecA-ssDNA filament by SSB, RecO, and RecR

6.1 Chapter summary

SSB, RecO and RecR are crucial mediator proteins regulating RecA-ssDNA nucleoprotein filament in eubacteria. However, the ssDNA binding properties of each individual protein and the heteroprotein complexes (RecOR, RecO-SSB, RecOR-SSB), as well as their influence on RecA filament formation and stability are still elusive. In this work, we addressed these questions by single-ssDNA manipulation by the new platform using magnetic tweezers. We show that RecO tightly folds ssDNA, RecOR highly extends ssDNA, and SSB/RecO-SSB/RecOR-SSB complexes wrap ssDNA at different levels. Under condition that RecA is not able to nucleate and polymerize on SSB coated ssDNA, neither RecO nor RecOR can facilitate RecA nucleation. However, under the same condition RecOR rather than RecO alone stabilizes a pre-formed RecA filament against net de-polymerization caused by the presence of SSB. Further, mechanical forces and solution conditions regulate the conformations of these nucleoprotein complexes sensitively. Based on these findings, we proposed a model that highlights the potential role of mechanical properties of the nucleoprotein complexes formed by

these mediator proteins on RecA nucleation, polymerization, and stability.*

6.2 Introduction

As introduced in Chapter 1, in bacteria, the essential player of the HR is the nucleoprotein filament formed by RecA on ssDNA [10, 11]. The formation and stability of RecA filament is elaborately regulated by environmental factors such as temperature, pH, salt concentration, and mechanical force, as well as by a set of mediator proteins such as SSB, RecO, and RecR that are targets of this research [10, 11].

In vivo, the ssDNA intermediates produced during DNA damage repair are first coated by SSB, protecting the ssDNA from degradation, or formation of secondary structures. However, this SSB-coated ssDNA nucleoprotein array imposes a significant physical barrier for the nucleation and polymerization of RecA onto ssDNA [10, 11, 33]. Further, as shown in Chapter 3, SSB destabilizes pre-formed RecA filaments by occupying the vacated ssDNA sites from ATP-hydrolysis dependent RecA dissociation at the $5'$ end [109]. A specialized class of mediator proteins involved in the RecFOR pathway is evolved to overcome the strong inhibitory effect of SSB on RecA filament formation [10, 11, 47]. Among them, RecO and RecR work together to facilitate formation of RecA filament on SSB-coated ssDNA, while RecF has been proposed to be involved in other functions in RecA-mediated HR [10, 11, 47].

E.coli RecO protein (27.6 kDa) promotes renaturation of complementary DNA strands in an ATP-independent manner [10, 11, 146]. This function is enhanced by formation of RecO-SSB complex while is inhibited by formation of RecO-RecR complex [10, 146]. RecO interacts with both ssDNA and SSB while RecR (22 kDa) has so far no reported ssDNA or SSB binding activities [10, 146]. Therefore, it is believed that in the RecFOR pathway RecO plays an essential role in the replacement of SSB from ssDNA, making access for RecA loading, while RecR plays its roles through direct interaction with RecO [10, 146].

*Note that main contents detailed in this chapter are included in *Dynamics and Regulation of RecA-ssDNA filament by SSB, RecO, and RecR.* Le S. *et al.*, manuscript in preparation (2015).

Although a wealth of knowledge of these proteins have been obtained from decades of imaging and bulk biochemical biophysical studies, their ssDNA binding properties and concerted interplay to regulate RecA filament nucleation, polymerization, and stability on individual single ssDNA templates remain largely unexplored. In this work, using *E.coli* as a model system, we systematically studied the conformations of nucleoprotein complexes formed by SSB, RecO, RecOR, and SSB-RecOR, as well as their effects on RecA filament dynamics and stability, on single ssDNA molecules manipulated by magnetic tweezers.

6.3 Results

6.3.1 RecO induces folding of ssDNA, while RecOR highly extends ssDNA

The ssDNA binding property of RecO and the conformations of the resulting RecO-ssDNA complex still remain unclear, which is addressed in this section. We investigated the effects of a wide concentration range (1 nM-1 μM) of RecO on the force-extension curves of a short ssDNA tether (576-nt) in solutions containing 50 mM NaCl, 20 mM Tris-pH 7.4, with/without 10 mM $MgCl_2$. We note that EcRecA and EcSSB are purchased, EcRecO and EcRecR are expressed and purified by our collaborator Dr. Korolev Lab (Saint Louis University).

In 50 mM NaCl, the force extension curves of ssDNA recorded with $\sim$ 300 nM RecO do not have significant hysteresis and are nearly overlapped with that of naked ssDNA, indicating low level of RecO binding to ssDNA (Figure 6.1A). When concentration of RecO was increased to 600 nM or higher, significant extension reductions and hysteresis between force-decrease and force-increase scans were observed, suggesting that ssDNA is highly folded by high concentrations of RecO ($\sim$600 nM). Furthermore, the folded RecO-ssDNA complex is stable under moderate force range ($<$10 pN). The unfolding trajectories under higher constant forces exhibit irregular extension increase pattern, indicating that the folded complexes are likely lacking of a regular nucleoprotein structure.

$MgCl_2$ (a concentration of $\sim$ 1-10 mM) is believed to be presented during RecA-mediated homologous recombination *in vivo*. Therefore, we next studied

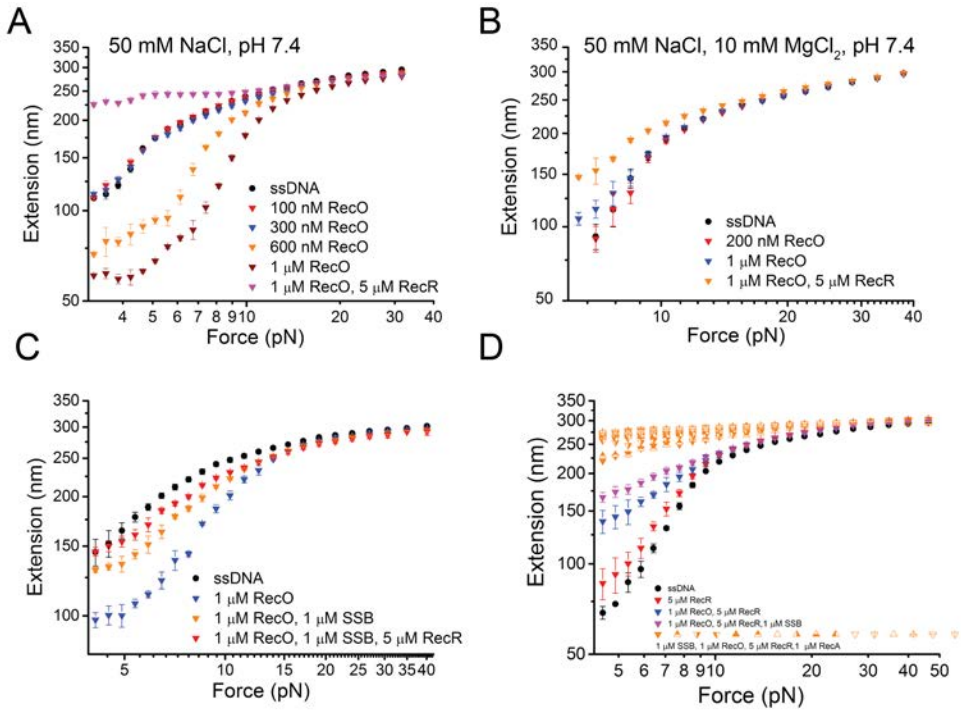

Figure 6.1: **Force responses of ssDNA bound with RecO, RecOR, RecOR-SSB**. (A-B). Force extension curves of 576-nt ssDNA bound with RecO, or RecOR in solutions containing 50 mM KCl, without (A) or with (B) 10 mM $MgCl_2$, with RecO concentration various from 100 nM to 1 μM, and 5 μM RecR at pH 7.4, 23°C. (C). Force-extension curves of ssDNA bound with RecO, RecO-SSB, or RecOR-SSB with 50 mM KCl. (D). Force-extension curves of ssDNA bound with RecO, RecO-SSB, or RecOR-SSB with 50 mM KCl and 10 mM $MgCl_2$. The orange data show the force-extension curves after 1 μM RecA was introduced to RecOR-SSB bound ssDNA. In these experiments, RecO was first introduced to ssDNA tethers.

how RecO binds to ssDNA in the presence of 50 mM NaCl, and 10 mM $MgCl_2$. For concentrations of RecO <300 nM, no obvious effects of RecO on ssDNA was observed in our force range; at a concentration of 1 μM, slightly extension changes at low forces (<8 pN) was observed (Figure 6.1B). However, it does not necessarily suggest that RecO-ssDNA complex presents a non-folding conformation, since high concentration of $MgCl_2$ (10 mM) is known able to condense ssDNA at low force [109]. It is likely that the folding level of RecO-ssDNA complex is similar to that of ssDNA in the presence of high $MgCl_2$.

RecR is known to form dimer and bind to RecO with a monomer ratio of 2:1, while has no reported binding to ssDNA [10]. To ensure sufficient RecR for formation of RecOR complex, the concentration of RecR in our experiments was kept as 5 μM when the RecO concentration was 1 μM. In 50 mM NaCl,

when a mixture of RecR and RecO was introduced to the RecO-ssDNA folded complex or naked ssDNA, the force-extension curve of ssDNA became much more stiffening than the naked ssDNA, presumably due to the formation of much extended structure of the RecOR-ssDNA complex (Figure 6.1A). In the presence of 10 mM $MgCl_2$, the RecOR-ssDNA complex is also more extended than ssDNA (Figure 6.1B). These results indicate that RecR relieves the folded RecO-ssDNA complex to an extended RecOR-ssDNA nucleoprotein complex.

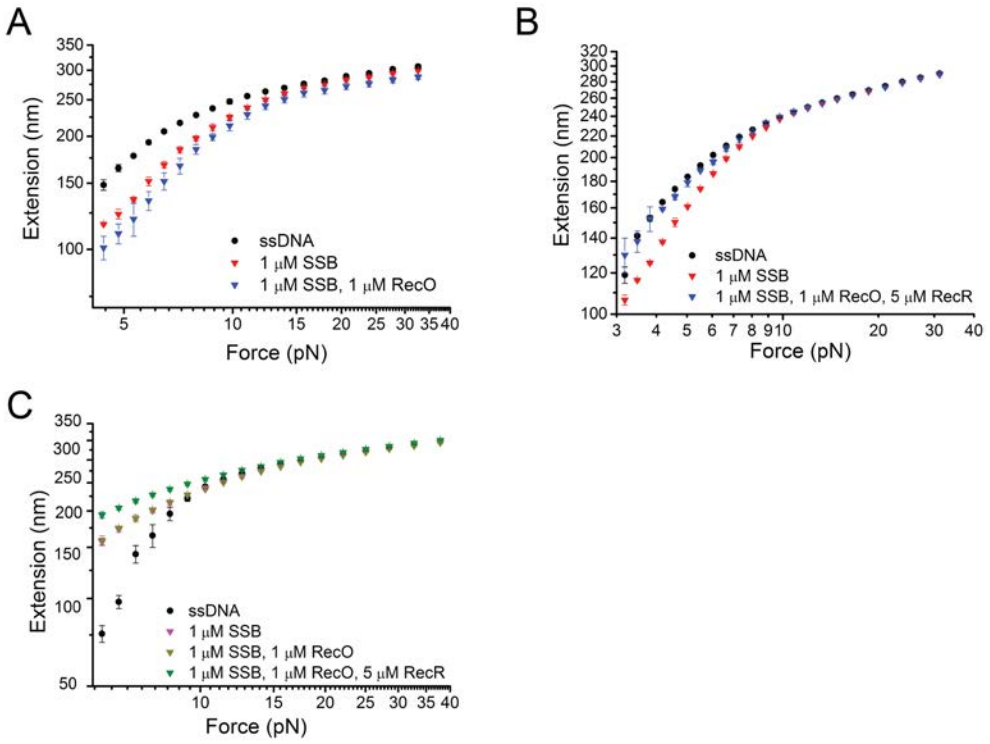

Figure 6.2: **Force responses of ssDNA bound with SSB, SSB-RecO, SSB-RecOR**. (A-B). Force extension curves of 576-nt ssDNA bound with SSB, SSB-RecO, or SSB-RecOR in solutions containing 50 mM KCl, without (A &B) or with (C) 10 mM $MgCl_2$. In these experiments, SSB was first introduced to the naked ssDNA tethers.

6.3.2 Neither RecO or RecOR complex removes SSB from ssDNA

Next, we examined how RecO and SSB interact (corporately or competitively) with each other upon binding to ssDNA. We first introduced 1 μM RecO to ssDNA tether in 50 mM NaCl solution in the absence of $MgCl_2$ to form a RecO-ssDNA folded complex (Figure 6.1C). Then a mixture of 1 μM RecO and 1 μM SSB was introduced to this RecO bound ssDNA tether, the force response of

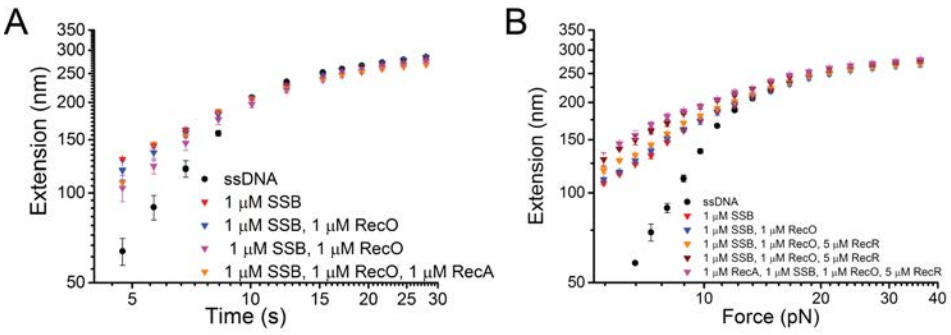

Figure 6.3: **Neither RecO, nor RecOR is able to facilitate RecA polymerization on SSB coated ssDNA at pH 7.4, 23 °C**. (A-B). Force extension curves of SSB coated ssDNA after mixture of SSB, RecO, and RecA (A) or SSB, RecOR, and RecA was introduced.

ssDNA became much less folded (Figure 6.1C) (Note that the force-response in the RecO and SSB mixture is still more folded compared to that of SSB-ssDNA complex), which likely suggests formation of an SSB-RecO-ssDNA complex. On the other hand, when 1 μM SSB was introduced in the first place to the naked ssDNA in the same buffered solution, forming a SSB-ssDNA array in the less wraping mode (Figure 6.2A). Then the mixture of 1 μM RecO and 1 μM SSB was introduced to this SSB pre-bound ssDNA tether, the force response of ssDNA became slightly more folded then that of SSB along, which is consistent with the formation of a SSB-RecO-ssDNA complex. Together, this SSB-RecO-ssDNA complex (or RecO-SSB-ssDNA complex) is mildly more folded than SSB-ssDNA complex and much more unfolded then RecO-ssDNA complex.

When 10 mM $MgCl_2$ is presented in the solution, the SSB coated ssDNA is overall more stiffened than naked ssDNA in the force range, which suggests that SSB binds to ssDNA in a less wrapping mode [109]. After the introduction of a mixture of 1 μM RecO and 1 μM SSB, the force-response of ssDNA remains almost overlapped with the one coated with SSB in the presence of 1 μM SSB alone (Figure 6.2B). This result suggests that either RecO is not able to bind to SSB-ssDNA complex in the presence of 10 mM $MgCl_2$, or RecO binds to SSB-ssDNA and forms a RecO-SSB-ssDNA complex which has similar the force response as SSB-ssDNA complex.

To sum up, RecO (1 μM) alone fold the ssDNA which reduce the force response of ssDNA, SSB (1 μM) alone binds to ssDNA in a less- or non-wrapping

mode which extended the force response of ssDNA, SSB relieves some level of the folding effects of RecO when forming the RecO-SSB-ssDNA complex. RecR (5 μM) further relieve the folding effects and even extended the architecture by forming the RecOR-ssDNA nucleoprotein array or RecOR-SSB-ssDNA nucleoprotein array.

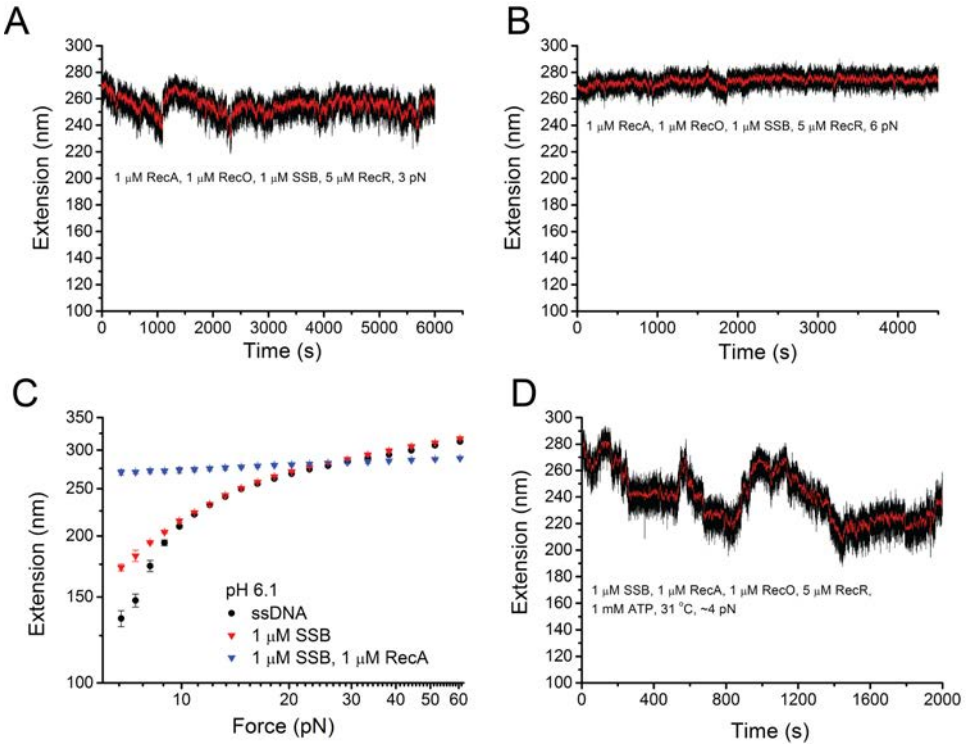

Figure 6.4: **RecOR stabilizes pre-formed RecA filament in the presence of SSB**. (A-B). Extension time traces of a pre-formed RecA filament in solutions containing 1 μM RecA, 1 μM SSB, 1 μM RecO, 5 μM RecR with 50 mM KCl,10 mM $MgCl_2$, 20 mM Tris-pH7.4, at 23°C, at 3 pN (A) or 6 pN (B). (C). Force extension curves of ssDNA bound with SSB (red) or RecA (blue) at pH 6.1. (D). Extension time traces of SSB-coated ssDNA after introduction of mixture of 1 μM RecA, 1 μM SSB, 1 μM RecO, 5 μM RecR at 31 °C. We note that the solution introduction was performed at a force of $\sim$ 20 pN for $\sim$ 1 min, during which RecA polymerizes on the SSB coated ssDNA.

6.3.3 Neither RecO alone or RecOR complex is able to assists RecA loading on SSB-bound ssDNA at pH 7.4 and 23 °C

Having examined the effects of SSB, RecO and RecOR on ssDNA alone or combined, we further investigated how RecA load onto SSB coated ssDNA in the presence of these mediator proteins. We first tested whether RecO alone or RecOR are able to facilitate RecA loading onto SSB coated ssDNA. A fully

coated SSB-ssDNA tether was formed with 1 μM SSB in standard RecA reaction solution, followed by sequential introduction of a mixture of 1 μM SSB and 1 μM RecO, and a mixture of 1 μM SSB, 1 μM RecO, and 1 μM RecA (Figure 6.2B). Note that 1 mM ATP and 1x ATP regeneration system were presented when RecA was included in the protein mixture. The resulting force responses were overlapped with that with the SSB alone; suggesting that RecO alone is not able to promote the loading of RecA onto SSB bound ssDNA at pH 7.4 and 22 ^{o}C.

Next, we repeated the experiments with additional 5 μM RecR in the protein mixture. The resulting force-response of ssDNA in the mixture of SSB, RecO, RecR and RecA overlapped with that in the mixture of SSB, RecO, and RecR, suggesting that RecOR is also not able to facilitate loading of RecA onto SSB bound ssDNA at this environmental condition (pH 7.4, 23 ^{o}C) (Figure 6.3A&B).

In contrast, in the event of the RecO (RecOR) is pre-bound with ssDNA, followed by the mixture of SSB, RecO, and RecR, and the mixture of SSB, RecO, RecR, and RecA, the RecA polymerized on ssDNA and formed a stable RecA filament even in the presence of SSB (Figure 6.1D). These results suggest that the barrier step is the rearrangement of RecOR and SSB on the ssDNA, consistent with previous biochemical studies [10].

6.3.4 The inhibitory effect of SSB on RecA filament formation is antagnized by lower pH or higher temperature

RecA filament has been reported to be more stable at low pH or high temperature in physiological range. Hence, we further tested whether these favorable environments are able to reverse the inhibitory effect of SSB on RecA filament formation with/without presence of RecOR complex. To begin with, we formed a SSB-ssDNA nucleoprotein array with 1 μM SSB in 50 mM NaCl, 10 mM $MgCl_2$, at pH 6.1 and 23^{o}C. Then, a mixture of SSB and RecA at the same buffered solution was introduced. Interestingly, the extension of ssDNA elongated immediately after introduction of the protein mixture, and reached an extended steady state, suggesting the formation of a stable RecA filament. This result indicates that RecA are able to nucleate and polymerize on SSB bound ssDNA at pH 6.1,

even without the assistance of RecOR, which may related to the dimerization of RecA at low pH [141]. Next, the effect of higher temperature was examined. An SSB-ssDNA nucleoprotein array was preformed with 1 μM SSB in 50 mM NaCl, 10 mM $MgCl_2$, at pH 7.4, 23 ^{o}C. Then the temperature was increased to 31 ^{o}C, followed by introduction of a mixture of SSB, RecO, RecR, and RecA. The resulting 50% elongated extension (compared to B-form dsDNA) indicating the formation of RecA filament at 31 ^{o}C (Figure 6.3E). The highly dynamic extension time trace at low force may suggest the competition of the inhibitory effect of SSB, the stabilizing effect of RecOR and the polymerization-impulsion of the remained RecA filament.

6.3.5 RecOR stabilizes the pre-formed RecA filament in the presence of SSB

As we previous reported, a pre-formed RecA filament would de-polymerize due to binding of SSB to the vacated ssDNA during ATP-hydrolysis mediated RecA disassociation at low forces ($\sim$3 pN). An interesting question is that what is the role of RecOR on a pre-formed RecA filament when SSB is presented. Hence, we preformed a RecA filament in the presence of 1 μM RecA at pH 7.4 and 23 ^{o}C, and then introduced the protein mixture of RecA, RecO, RecR, and SSB into the tethered RecA filament. No net de-polymerization of RecA filament was observed within our measuring time (>6000 s), indicating that the RecOR complex stabilizes the RecA filament in the presence of SSB (Figure 6.4A-C). Together with previous sections, these results suggest that RecOR is able to facilitate polymerization rather than nucleation of RecA on SSB-coated ssDNA at pH 7.4 and 23 ^{o}C, while lower pH or higher temperature in physiological range promotes the nucleation of RecA on SSB-coated ssDNA.

6.4 Discussion

In this work, we studied the ssDNA binding properties of RecO and the hetero-protein complexes (RecOR, RecO-SSB, RecOR-SSB), as well as their influence on RecA filament formation and stability. We show that RecO tightly folds ssDNA,

RecOR highly extends ssDNA, and SSB/RecO-SSB/RecOR-SSB complexes wrap ssDNA at different levels. Under conditions that RecA is not able to nucleate and polymerize on SSB coated ssDNA, neither RecO nor RecOR can facilitate RecA nucleation. However, under the same condition RecOR rather than RecO alone stabilizes a pre-formed RecA filament against net de-polymerization caused by the presence of SSB. Further, mechanical forces and solution conditions regulate the conformations of these nucleoprotein complexes sensitively. Based on these findings, we proposed a model that highlights the potential role of mechanical properties of the nucleoprotein complexes formed by these mediator proteins on RecA nucleation, polymerization, and stability.

We showed that RecO binds to ssDNA, resulting in highly folded RecO-ssDNA nucleoprotein structure. RecO is known to mediate DNA strand-annealing, which is likely mediated by ssDNA folding by RecO. RecO-SSB-ssDNA nucleoprotein complex is also relatively more folded compared to SSB-ssDNA nucleoprotein array. *in vivo*, SSB first binds to ssDNA and facilitates DNA annealing by removing/preventing the secondary structures on ssDNA, although adding energy cost to access ssDNA. RecO later binds to SSB-ssDNA array and forms RecO-SSB-ssDNA complex, on which RecO and SSB rearrange their binding modes on ssDNA and accomplish DNA annealing together. RecOR, on the other hand, relieves the folding effect of RecO and extends ssDNA into an filamentous RecOR-ssDNA or RecOR-SSB-ssDNA nucleoprotein structures. RecOR is known to facilitate RecA loading onto SSB-coated ssDNA in RecF pathway. Together, distinctive micro-mechanical properties of folded RecO-ssDNA and extended RecORssDNA conformations likely guided the selection of DNA strand annealing or RecA-dependent homologous recombination.

Although RecOR has been shown to facilitate RecA-loading onto SSB-coated ssDNA under certain conditions (<pH 6.8) [128], we showed that under conditions that RecA is not able to nucleate and polymerize on SSB coated ssDNA (pH 7.4, 23 ^{o}C), neither RecO nor RecOR can facilitate RecA nucleation and polymerization. At the same environmental conditions, however, RecOR, rather than RecO alone, can stabilize pre-formed RecA nucleoprotein filament, antagonizing the de-polymerization effect of SSB. These results likely suggest that at

these conditions, RecOR is able to stabilize the existing RecA nucleation sites, rather than facilitating formation of new nucleation sites. From the existing nucleation sites, RecA therefore can polymerize in both directions. In addition, at higher temperature or lower pH within physiological range, RecA is able to spontaneously nucleate and polymerize on SSB-coated ssDNA, RecOR at these conditions, while is not essential, can enhance the polymerization speed of RecA [128].

Chapter 7

Theoretical analysis of force effect on dynamics and stability of RecA-ssDNA filament

7.1 Chapter summary

In previous chapters, we have investigated the dynamics and regulation of RecA nucleoprotein filament formed on ssDNA by various mediating proteins and environmental factors. Interestingly, an mechanosensitive regulation of RecA nucleoprotein filament is identified and is likely universal exist during RecA-mediated homologous recombinational DNA repair as well as other nucleic acids reactions. However, the underlining mechanisms of these mechanosensitive regulations are still elusive. In this chapter, we theoretically analyze the effects of force on protein-DNA binding kinetics and conformational free energies of naked ssDNA and RecA nucleoprotein filament. Both the theoretical analysis and experimental results consistently suggest that mechanical force is capable to sensitively regulate formation and stability of RecA filament, which provides a possible mechanism for the observation in previous chapters.

7.2 Force-free equilibrium binding and unbinding of protein

We first reviewed equilibrium binding/unbinding of protein based on equilibrium statistics physics. Assuming a solution of total volume of V contains N molecules (proteins), each protein occupies a unit volume of ν. The number density and molar density of the protein are defined as $c_{\mathrm{N}} = \frac{N}{V}$, $c_{\mathrm{A}} = \frac{N/N_{\mathrm{A}}}{V} = \frac{c}{N_{\mathrm{A}}}$, where $N_{\mathrm{A}} \approx 6.02 \times 10^{23}$ is the Avogadro's number. The total volume can be discretised into M lattices by $M = \frac{V}{\nu}$. The lattice occupancy fraction (also termed as volume fraction) is therefore defined as $\theta = \frac{N}{M} = \frac{N\nu}{V} = c\nu$. Based on the number of available states (Ω), the entropy is defined as $S = k_{\mathrm{B}}\ln\Omega = -k_{\mathrm{B}}N\ln\theta$ for diluted solution (i.e., $\theta \ll 1$). Hence, each protein carries an entropy of $-k_{\mathrm{B}}\ln\theta$.

When a protein is unbound 'off' or bound 'on' with a substrate (DNA), the free energy of the system is: $G_{\text{off}} = -TS_{\mathrm{N}} = Nk_{\mathrm{B}}T\ln\theta$, $G_{\text{on}} = -\varepsilon - TS_{\text{N-1}} = -\varepsilon - (N-1)k_{\mathrm{B}}T\ln\theta$ Hence, the free energy cost of binding for the protein is :

$$\Delta G_{\text{on-off}} = G_{\text{on}} - G_{\text{off}} = -\varepsilon - k_{\mathrm{B}}T\ln\theta = -\varepsilon - k_{\mathrm{B}}T\ln(c\nu) \tag{7.1}$$

Next, we denote the probabilities of the bound and unbound states as p_{on} and p_{off}, respectively. At equilibrium, p_{on} and p_{off} should follow the Boltzmann distribution:

$$p_{\text{on}} = \frac{e^{-\beta\Delta G_{\text{on-off}}}}{1 + e^{-\beta\Delta G_{\text{on-off}}}} \tag{7.2a}$$

$$p_{\text{off}} = \frac{1}{1 + e^{-\beta\Delta G_{\text{on-off}}}} \tag{7.2b}$$

where $\beta = \frac{1}{k_{\mathrm{B}}T}$. Since for a protein, $\Delta G_{\text{on-off}} = -\varepsilon - k_{\mathrm{B}}\ln\theta$, the two terms control the probability of binding: $\varepsilon > 0$ leads to a negative energy cost, which favours protein-substrate binding; θ is always < 1, hence $k_{\mathrm{B}}\ln\theta$ always < 0, leads to a positive energy cost, thereby unfair protein-substrate binding, which is purely an entropic effect. In addition, the more diluted the protein concentration, the smaller the θ, thereby the higher the energy cost for binding.

Further, the ratio of binding and unbinding probabilities is:

$$\frac{p_{\text{on}}}{p_{\text{off}}} = e^{-\beta \Delta G_{\text{on-off}}} = \frac{\theta}{e^{-\beta\varepsilon}} = \frac{\theta}{\theta^*} = \frac{c\nu}{\theta^*} = \frac{c}{k_{\text{D}}} \tag{7.3}$$

where $\theta^* = e^{-\beta\varepsilon}$ is the critical occupancy fraction where $p_{\text{on}} = p_{\text{off}}$. $k_{\text{D}} = \frac{\theta^*}{\nu}$ is often referred as dissociation constant (with the same unit as c). The binding energy ε therefore can be calculated from k_{D} through:

$$\varepsilon = -k_{\text{B}}T\ln\theta^* = -k_{\text{B}}T\ln(k_{\text{D}}\nu) \tag{7.4}$$

7.3 Effects of force on equilibrium binding and unbinding of proteins

Now we consider the situation that an external force F is applied to the substrate (DNA). The F introduces a force-dependent conformational free energy to the bound state and unbound state:

$$\Phi_{\text{i}}(F) = -\int_0^F x_{\text{i}}(f')df', \text{i} = \text{on,off.} \tag{7.5}$$

The total free energies of the two states therefore are:

$$g^F_{\text{off}} = \Phi_{\text{off}}(F) \tag{7.6a}$$

$$g^F_{\text{on}} = \Phi_{\text{on}}(F) + \mu \tag{7.6b}$$

where $\mu = -\varepsilon - k_{\text{B}}T\ln\theta = -\varepsilon - k_{\text{B}}T\ln(c\nu)$ is the binding energy at zero force, here we assume that force does not affect ε.

At equilibrium, the binding and unbinding probabilities in the presence of force, p^F_{on}, and p^F_{off} are determined by Boltzmann distribution:

$$p^F_{\text{off}} = \frac{1}{Z^F} e^{-\frac{g^F_{\text{off}}}{k_{\text{B}}T}} \tag{7.7a}$$

$$p_{\text{on}}^F = \frac{1}{Z^F} e^{-\frac{g_{\text{on}}^F}{k_{\text{B}}T}} \tag{7.7b}$$

where $Z^F = e^{-\frac{g_{\text{off}}^F}{k_{\text{B}}T}} + e^{-\frac{g_{\text{on}}^F}{k_{\text{B}}T}}$ is the partition function.

Further, the ratio of binding and unbinding probabilities becomes:

$$\frac{p_{\text{on}}^F}{p_{\text{off}}^F} = \frac{e^{-\beta g_{\text{on}}^F}}{e^{-\beta g_{\text{off}}^F}} = \frac{\theta e^{\beta\varepsilon} e^{-\beta\Phi_{\text{on}}(F)}}{e^{-\beta\Phi_{\text{off}}(F)}} = \frac{\theta}{e^{-\beta\varepsilon}} e^{-\beta(\Phi_{\text{on}}(F)-\Phi_{\text{off}}(F))} = \frac{\theta}{e^{-\beta\varepsilon}} e^{-\beta\Delta\Phi_{\text{on-off}}(F)} \tag{7.8}$$

The critical occupancy fraction θ^* becomes $\theta^* = e^{-\beta\varepsilon} e^{\beta\Delta\Phi_{\text{on-off}}(F)}$. Therefore, we can re-define a force-dependent dissociation constant $k_{\text{D}}(F)$ as:

$$k_{\text{D}}(F) = \frac{\theta^*}{\nu} = \nu^{\text{-}1} e^{-\beta\varepsilon} e^{\beta\Delta\Phi_{\text{on-off}}(F)} = k_{\text{D}}^0 e^{\beta\Delta\Phi_{\text{on-off}}(F)} \tag{7.9}$$

7.4 Effect of force on de-polymerization and re-polymerization of RecA filament

The stability of RecA filament results from the competition between RecA polymerization and de-polymerization. When a tensile force F is applied to the ssDNA, the probability ratio of RecA polymerization and de-polymerization depends on the total free energy cost of polymerization, $\Delta G(F,\xi)$, by the Boltzmann distribution: $\frac{p_{\text{on}}}{p_{\text{off}}} = e^{\frac{\Delta G(F,\xi)}{k_{\text{B}}T}}$, where a parameter vector ξ describes the solution conditions such as temperature, salt concentration and pH. $\Delta G(F,\xi)$ consists of a force-independent term, $\Delta G_0(\xi)$, from the physical interaction between RecA and ssDNA, and a force-dependent term, $\Delta\Phi(F,\xi)$, which is the conformational free energy difference between RecA filament and ssDNA:$\Delta\Phi(F,\xi) = \Phi_{\text{RecA}}(F,\xi) - \Phi_{\text{ssDNA}}(F,\xi)$. Therefore, at forces when $\Delta\Phi(F,\xi) < 0$, RecA filament formation is promoted by the decreased conformational free energy during polymerization.

The force dependent conformational free energies of RecA filament and ssDNA, $\Phi_{\text{RecA}}(F,\xi)$ and $\Phi_{\text{ssDNA}}(F,\xi)$, respectively, can be calculated from their respective force-extension curves $x_{\text{RecA}}(F,\xi)$ and $x_{\text{ssDNA}}(f,\xi)$, through relations:

$$\Phi_{\mathrm{RecA}}(F,\xi) = \int_0^F x_{\mathrm{RecA}}(f',\xi)df' \tag{7.10a}$$

$$\Phi_{\mathrm{ssDNA}}(F,\xi) = \int_0^F x_{\mathrm{ssDNA}}(f',\xi)df' \tag{7.10b}$$

$x_{\mathrm{RecA}}(F,\xi)$ and $x_{\mathrm{ssDNA}}(f,\xi)$ were directly measured in our buffered reaction solution conditions. The calculated $\Delta\Phi(F,\xi)_{\text{RecA-ssDNA}}$ per RecA monomer is negative up to 90 pN, force in this range facilitates RecA polymerization by reducing the free energy cost for polymerization and is optimized in force range of 20-25 pN (Figure 7.1A). Further, we calculated the effect of force on binding affinity of RecA on ssDNA. Based on $k_{\mathrm{D}}(F) = k_{\mathrm{D}}^0 e^{\beta\Delta\Phi_{\text{RecA-ssDNA}}(F)}$, we found that force effectively reduce the k_{D} in the force range up to ~90 pN (Figure 7.1B). The

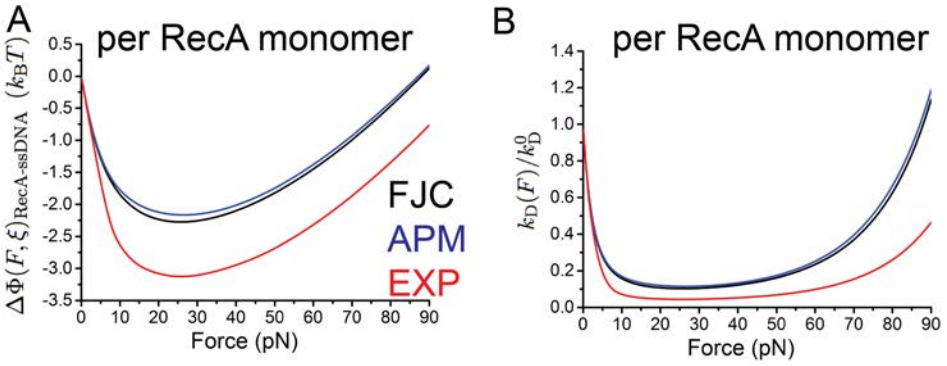

Figure 7.1: **Force dependent free energy difference and dissociation constant**. (A). Force dependent free energy difference between RecA nucleoprotein filament and ssDNA per RecA monomer. (B). Force dependent dissociation constant of RecA binding to ssDNA per RecA monomer. Black and blue lines are calculated based on worm-like chain mode (Eq. 1.2 with A = 1000 nm and L = 0.52 nm) for RecA, and free-joint chain model (Eq. 1.2) or an analytical polymer model (Eq.1.3) for ssDNA, respectively (Figure 1.5). Red lines are calculated based on experimental data (Figure 5.1).

equilibrium between RecA polymerization and de-polymerization depends on the rates of polymerization and de-polymerization, k_{on} and k_{off}, respectively, through $\frac{p_{\mathrm{on}}}{p_{\mathrm{off}}} = \frac{k_{\mathrm{on}}}{k_{\mathrm{off}}}$. Therefore, the effect of force on shifting the equilibrium must be through either increasing k_{on}, or decreasing k_{off}, or both of the RecA filament.To investigate the effects of force on the on-rate, we recorded the extension change of the ssDNA right after 200 nM MtRecA was introduced at different forces in 50 mM KCl, 10 mM $MgCl_2$, 20 mM Tris (pH 7.4), with 1 mM ATP, at 23

°C. Similarly, the force-dependent off-rate was measured for a fully polymerized MtRecA filament after removal of free RecA in the channel at different forces. Our results show that force facilitates polymerization while it suppresses depolymerization, with maximal effects around 20 pN (Figure 4 in reference [112]).

Chapter 8

Dynamics and stability of RecA nucleoprotein filament formed on dsDNA

8.1 Chapter summary

In previous chapters, we mainly focus on the regulation of RecA nucleoprotein filament formed on ssDNA. RecA filament formed on dsDNA has also been proposed to be an important functional state during homologous recombinational DNA repair *in vivo*. The nucleation, polymerization and de-polymerization of RecA-dsDNA nucleoprotein filament are dynamically regulated by various physiological factors. Nevertheless, up to date, there is still lacking a comprehensive understanding of the regulatory mechanisms employed by these factors. Here, in this chapter, using the single-DNA manipulation platform by magnetic tweezers, we systematically investigate dynamics of individual EcRecA filaments polymerization and de-polymerization regulated by a set of physiological factors including temperature, pH, DNA topology and mechanical forces. We identified the optimal conditions that favors spontaneous RecA nucleation and polymerization of RecA filament, as well as conditions that maintain RecA filament stability. Furthermore, nano-meter spatial resolution examination of the striking dynamic saw-tooth pattern fluctuation of DNA extension uncovers dynamic competition between RecA filament stochastic catastrophic de-polymerization and slow re-

polymerization. In addition, our data also demonstrated that the S-DNA (a stretched base-paired form of DNA) is not, although it was previous proposed to be, a nucleation substrate for RecA filament. Overall, our studies resolved several previous single-molecule studies that reported seemly contradictory and inconsistent results on RecA nucleation, polymerization and stability. Furthermore, our findings also cast lights on the molecular nature of RecA-dsDNA filament formation and stability *in vivo*. *

8.2 Introduction

As introduced and discussed in previous chapters, the bacterial RecA filament formed on ssDNA, which is dynamically regulated by various factors, is critical in homologous recombinational DNA repair. In each RecA monomer, there are two distinct (primary and secondary) DNA binding sites [10, 11]. While RecA-ssDNA filament formation is through the primary binding sites, the secondary binding sites of RecA filament weekly interact with dsDNA during homologous search and strand exchange [10, 11, 147].

Studies have reported that RecA can aslo nucleate and polymerize on ds-DNA besides its primary form of RecA-ssDNA nucleoprotein filament [10, 11, 18, 22, 29, 30, 147–150]. Importantly, the RecA-dsDNA nucleoprotein filament possibly exists as a functional state during DNA strand exchange [147]. Hence, a comprehensive understanding of the properties of RecA-dsDNA filament in physiologically relevant conditions, and the regulatory factors, ultimately the nature of dynamics of RecA-dsDNA nucleoprotein filament may provide insights for homologous recombination.

Previously, plenty of single-molecule manipulation experiments studying the polymerization/de-polymerization dynamics of RecA nucleoprotein filament formed on dsDNA [18, 22, 29, 30, 149]. While some of them demonstrated that RecA fails to nucleate and polymerize on dsDNA at low force (several pN), and a pre-formed RecA-dsDNA filament is unstable, leading to a net de-polymerization

*Note that main contents detailed in this chapter have been published in *Dynamics and Regulation of RecA polymerization and de-polymerization on double-stranded DNA*. Fu H., Le S. *et al.*, PLoS ONE 8(6): e66712 (2013).

[149], while some others have shown spontaneous polymerization of RecA on ds-DNA at low mechanical forces, leading to a stable RecA-dsDNA nucleoprotein filament [18, 22, 29, 30]. Furthermore, in studies reporting net RecA polymerization at low force, while some reported spontaneous nucleation at low force [22, 30], an initial large force (∼65 pN) for DNA overstretching transition was often required to promote the initial RecA nucleation [18, 29, 149]. Up to now, it is still unclear about the causes of the above contradictions.

In addition, the nature of RecA polymerization assisted by DNA overstretching is also unclear. Since the torsion-unconstraint DNA overstretching transition at $\sim$ 65 pN involves two DNA structural transitions: strand-pealing transition which produces two ssDNAs, and B-to-S transitions resulting in a base-paired extended DNA structure now normally termed as the S-DNA. [92–97] While the ssDNA is a known binding substrate for RecA nucleation, the elongated S-DNA has also been proposed for RecA nucleation and polymerization[18, 89, 92]. Hence, an interesting question is whether one or both of the structural transitions assist the nucleation of RecA.

To understand the nature of dynamics of RecA-ssDNA filament, firstly, it is important to resolve the conflicting observations. A possible cause of the contradictions might be the different experimental conditions employed in those previously studies, since microenvironmental factors, such as pH, temperature, salt concentration and force affect various cellular processes. Therefore, we first systematically examined this premise by investigating the dynamic competition between RecA-dsDNA filament polymerization and de-polymerization regulated by temperature, pH, salt, tensile force, and dsDNA topology. We indentified the optimal conditions that favour spontaneous RecA nucleation and polymerization of RecA filament, as well as conditions that maintain RecA filament stability, and resolved the contradictory observations. Moreover, previous experiments were performed with long DNA molecules (usually >10 μm), which leads to missing the detailed dynamics of RecA-dsDNA filament. Hence, we then monitored the detailed real-time dynamics of the competition between RecA filament polymerization and de-polymerization using much shorter DNA (∼200 nm) which enables nano-meter scale resolution. Furthermore, based on the different dynam-

ics of RecA filament formed on differently designed dsDNA, we proposed a RecA filament end-capping model, which explains the molecular natures of formation and stability of RecA-dsDNA nucleoprotein filament.

8.3 Results

8.3.1 Balance between RecA polymerization and de-polymerization regulated by temperature and pH within physiological range

To begin with, in this section, we investigated the effects of temperature, pH and force on dynamics of RecA-dsDNA filament, since these factors were different in several previous studies [18, 22, 29, 30, 149, 150]. At 24^{o}C in 20 mM Tris (pH 7.4), 1 μM RecA, 50 mM KCl, 10 mM $MgCl_2$, 1 mM ATP and 1$\sim$ ATP regeneration system, no RecA polymerizes on dsDNA over a wide force ranges up to 48.9 pN, indicated by the constant DNA extensions at corresponding forces, without extension elongation at corresponding DNA holding time (Figure 8.1A). To test whether at this reaction condition, DNA overstretching transition can assist the nucleation and polymerization of RecA, we applied a large force of $\sim$72.8pN, at which, DNA went through overstretching transition, as indicated by the initial extension elongation of $\sim$1.7-fold (red in the figure panel of Fig. 1A). With the assist of DNA overstretching at this force, RecA starts to polymerize along DNA, indicated by the shortening of DNA extension afterwards. The observed DNA extension shortening is due to extension difference of a RecA-dsDNA filament and an overstretched DNA, with contour length of $\sim$ 1.5 times [18, 22, 149] and $\sim$ 1.7 times [81, 85] compared to that of B-DNA, respectively. Roughly, the extension difference between them is $\sim$ 0.07 nm/bp ($(1.7 - 1.5) \times 0.34$ nm/bp). Therefore, during holding DNA for $\sim$ 80 sec at $\sim$ 72.8 pN, $\sim$ 4,286 bp of dsDNA is covered by RecA filaments with DNA extension reduced $\sim$300 nm.

To test whether the partially polymerized RecA-dsDNA filament is able to continue polymerization at low force, we decreased the force to $\sim$6.2 pN. The initial extension after force drop is $\sim$777 nm longer than B-DNA, which is consistent with that RecA is partially polymerized on DNA at high force. An estimation

of ∼4,571 bp RecA-dsDNA filaments agrees well with that estimated at higher force (∼10 % relative error).

However, the partial RecA-dsDNA filament went through progressive de-polymerization and returned to B-DNA at this force, indicated by the corresponding DNA extension shortening. These observation of unstable RecA-dsDNA filament at low forces are fully consistent with that observed by Feinstein et al. [149] in similar reaction solutions.

We repeated the experiments at 37 °C with other conditions remain unchanged. After the initial force-assisted nucleation of RecA at ∼ 58.8 pN (the onset force of DNA overstretching transition at this temperature), progressive RecA polymerization continues when force was decreased to ∼ 6.2 pN (Figure 8.1B). The DNA extension elongated to about 21 μm (∼1.3 times of the B-DNA) in the first 600 second, followed by a slower growing to nearly steady state in the following 1,200 seconds. In addition, we note that at this temperature, spontaneous nucleation and polymerization of RecA on dsDNA can also tack place without facilitation of force-induced DNA overstretching transition as shown in inset of Figure 8.1B. This higher temperature assisted spontaneous RecA filament polymerization is consistent with the study by Shivashankar *et al.* [22]. Additionally, in experiments, same DNA tether can be re-used by utilizing the fact that RecA-dsDNA filaments quickly and fully de-polymerize and return to B-DNA when ATP is replaced with ADP in the absence of free RecA in solution (Fig. 8.1C).

Next, we repeated the experiments with a low pH (pH 6.1) while remaining the temperature at 24°C. As shown in Figure 8.1D, during introduction of the proteins, spontaneous RecA nucleation and polymerization took place at low force. The polymerized RecA filament then rapidly de-polymerized when pH was changed to 7.4 (Figure 8.1E).

Together the results depicted in Figure 8.1A-E, in this section, we demonstrated that the dynamics of RecA-dsDNA filaments is sensitively governed by temperature and pH. At low forces, RecA can polymerize at higher temperature and/or lower pH values and the resulting RecA-dsDNA filament is stable; On the other hand, RecA is not able to nucleate and polymerized at lower temperature

and higher pH, where a pre-existing RecA-dsDNA filament is unstable and goes through net de-polymerization until return to B-DNA.

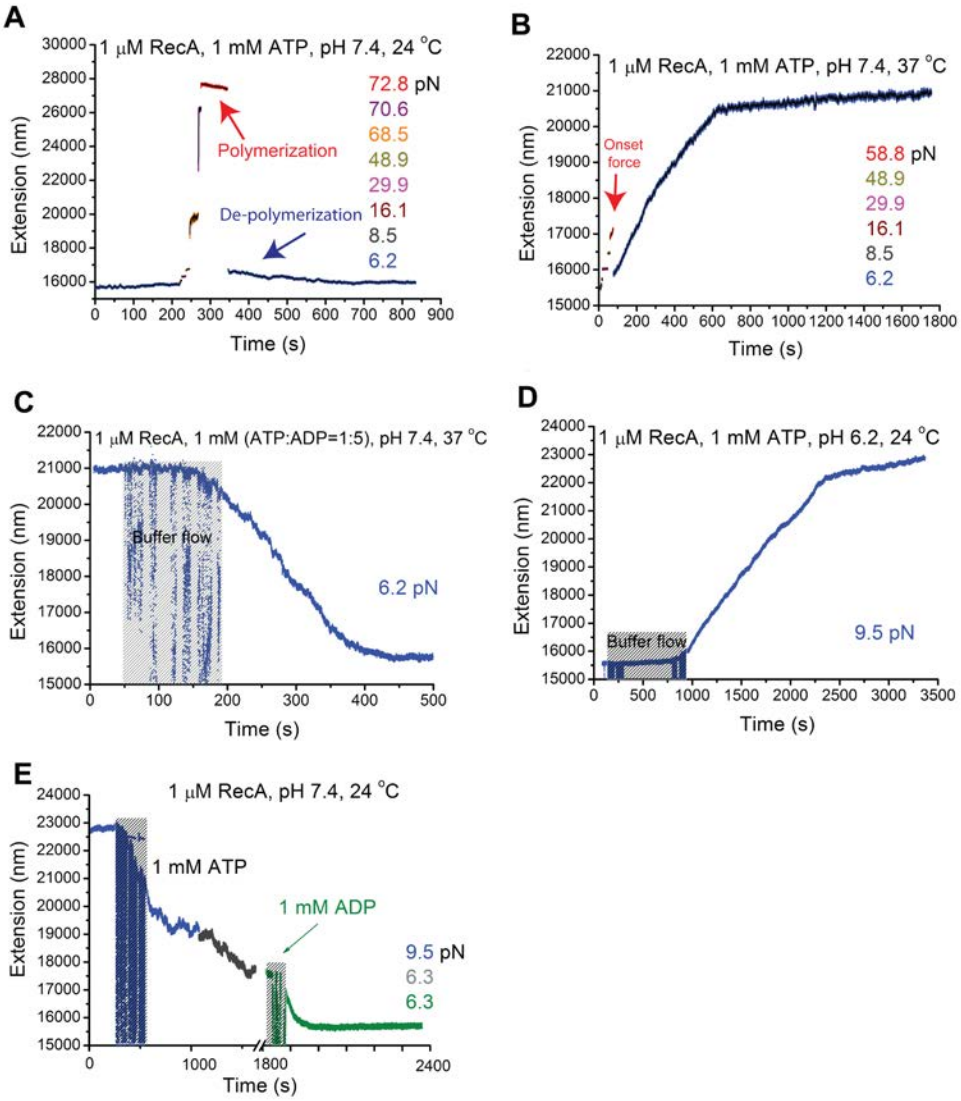

Figure 8.1: **Effects of temperature and pH on the formation and stability of RecA filament**. (A) Time trace of RecA polymerization and de-polymerization in a λ-DNA in 1 μM RecA, 50 mM KCl, 10 mM $MgCl_2$, 1 mM ATP, 1× ATP regeneration system, pH 7.4, and 24° C, at different forces indicated by different colors. Progressive polymerization was observed at ∼72.8 pN after DNA overstretching indicated by shortening in DNA extension (red allow), while de-polymerization was observed when force was decreased to ∼ 6.2 pN (blue arrow). (B) Following the complete de-polymerization in (A), time trace was obtained on the same DNA at 37 °C (other conditions remained unchanged). Progressive polymerization was observed at ∼6.2 pN (blue data points) after initiation with DNA overstretching transition by ∼58.8 pN for a short time duration (red arrow). (C) Time trace of de-polymerization of the RecA nucleoprotein filament formed in (B) after introduction of 1 mM mixture of ATP and ADP (ATP : ADP = 1 : 5) (other conditions remained unchanged) at ∼6.2 pN. (D) Time trace of spontaneous RecA polymerization on a different λ-DNA at pH 6.2, 24 °C, and 9.5 pN without initiation by DNA overstretching. (E) Time trace of de-polymerization of RecA nucleoprotein filament formed in (D) after pH was changed to 7.4 with 1 mM ATP (blue and dark grey data at different forces) and with 1 mM ADP (green data). The noisy data in (C-E) in the shadowed areas were recorded during buffer exchanging.

8.3.2 Nano-meter scale, detailed dynamics of RecA-dsDNA filament polymerization and de-polymerization

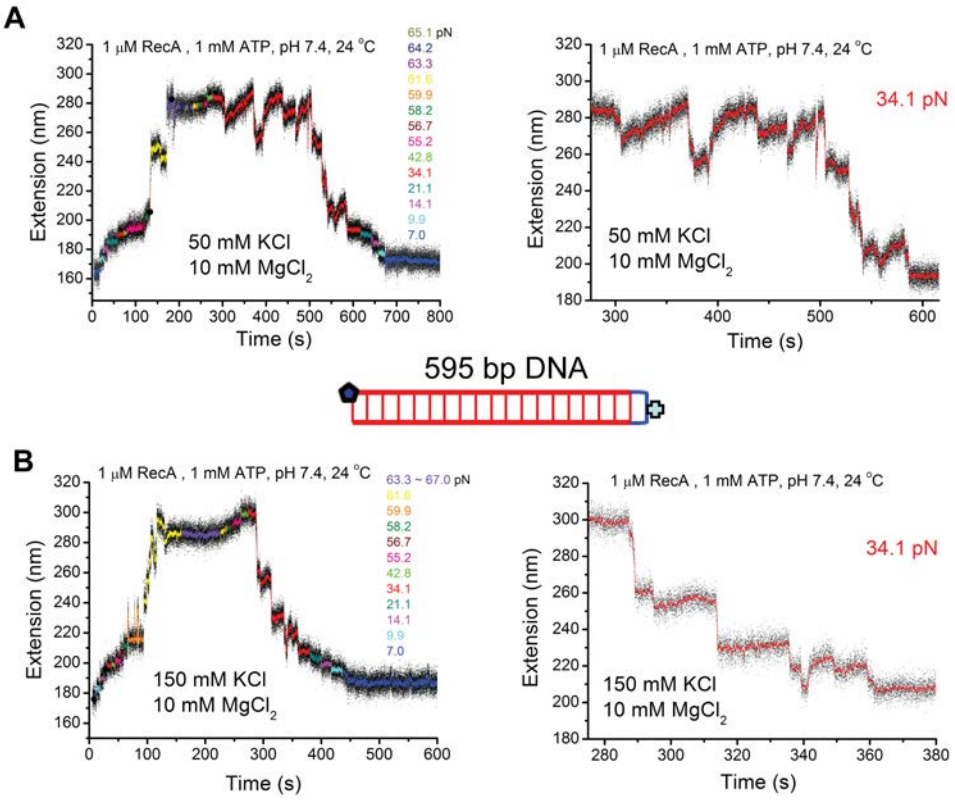

Figure 8.2: **Detailed dynamics of short RecA filaments in different KCl concentrations**. (A-B) Time traces of polymerization and de-polymerization of RecA nucleoprotein filament on a 595 bp dsDNA in 1 μM RecA 1 mM ATP, 1× ATP regeneration system, 24 °C, pH 7.4, with 50 mM KCl (A) first then 150 mM KCl (B) next. Data in the left panels were recorded at different forces indicated by different colors. Right panels show dynamics of the competition between polymerization and de-polymerization under a constant force of ∼34.1 pN. Inset shows the sketch of the 595 bp DNA containing one closed end and one open end.

In previous section, we have systematically shown that temperature and pH governed the balance of polymerization and de-polymerization dynamics of RecA-dsDNA filament. However, due to the use of long DNA (∼16 μm) template in the experiments, the detailed competition dynamics were buried in the large noise due to DNA longitude fluctuation. Hence, in following sections, we examine the nano-meter scale dynamics of competitive dynamics of RecA-dsDNA filament polymerization and de-polymerization using much shorter DNA (∼600 bp). We note that, for short dsDNA tethers, we only focus on extension changes

at constant forces, which reveals the dynamics of RecA-dsDNA nucleoprotein filament.

As shown in Figure 8.2A-B, we first examined the detailed dynamics of the RecA nucleoprotein filament formed on the 595 bp dsDNA, whose end is topologically closed to avoid potential strand-dissociation due to DNA overstretching, in reaction solutions containing 10 mM $MgCl_2$, 20 mM Tris-HCl (pH 7.4), 1 uM RecA, 1mM ATP, 1×ATP regeneration system, and 50 mM (A) or 150 mM (B) KCl, at 24 °C. At this temperature, RecA polymerization was initiated with assistance of DNA overstretching transition, and indicated by the extension decrease at $\sim$ 60 pN. After the DNA is nearly fully polymerized with RecA, we decreased the forces to lower value and monitored the extension dynamics over long time.

Figure 8.2A, left panel shows the time course of the experiment from overstretching assisted polymerization to de-polymerization at lower force values. RecA was fully polymerized at force greater than 63.3 pN. The right panel shows the extension dynamics at a constant force of $\sim$34.1 pN over $\sim$ 380 seconds.

Interestingly, with 50 mM KCl, we observed a dynamic saw-tooth pattern of the DNA extension with stochastic abrupt extension drops (step sizes up to $\sim$30 nm) and a followed slow extension elongation at the decreased constant forces (Figure 8.2A). We reason that the saw-tooth dynamic pattern is due to that at this force range, RecA de-polymerization processes (indicated by the abrupt drops of extension) only slightly out-competed RecA re-polymerization processes (indicated by slow extension elongation). Consistently, it took a longer time of $>$ 300 seconds for the RecA-dsDNA filament de-polymerize for $\sim$ 100 nm, from $\sim$ 280 nm to $\sim$ 190 nm, compared to lower forces, where RecA fully de-polymerized from DNA which thereby returned to B-DNA.

Since dsDNA stability is sensitive to salt concentration, we hypothesize that the saw-tooth dynamics pattern due to the balance of RecA polymerization/de-polymerization may also sensitive to salt concentration. To test it, we repeat the experiments with higher KCl concentration of 150 mM while other conditions remain the same (Figure 8.2B). Consistent with the hypothesis, stochastic abrupt extension drops due to dynamic RecA de-polymerization were still ob-

served, however, the followed slow re-polymerization processes were much less often compared to that with 50 mM KCl. Moreover, a plateau of DNA extension was often observed between two continuous RecA de-polymerization events. The overall net de-polymerization process was much faster (it took ∼100 seconds for ∼ 100 nm extension drop) compared to that with lower salt concentration, which also demonstrated a predominated RecA de-polymerization dynamics at higher salt concentrations.

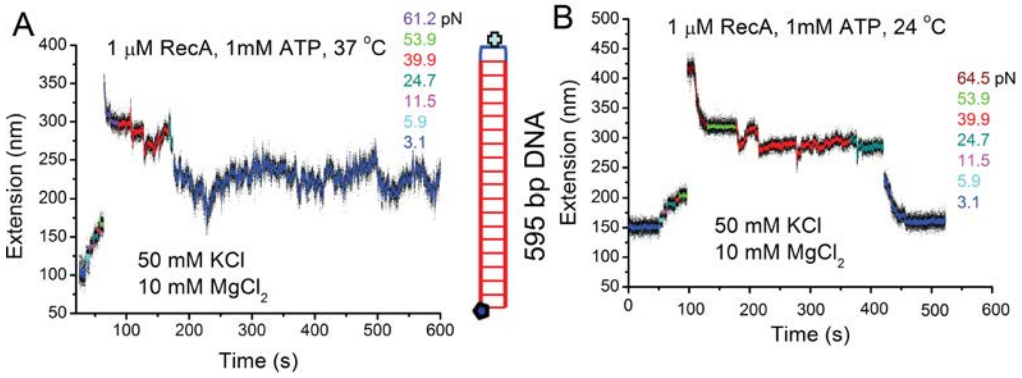

Figure 8.3: **Effects of temperature on short RecA filaments**. (A-B) Time traces of polymerization and de-polymerization of RecA nucleoprotein filament on another 595 bp dsDNA tether in 1 μM RecA 1 mM ATP, 1× ATP regeneration system, 50 mM KCl, 10 mM $MgCl_2$, pH 7.4, at 37 °C (A) first then 24 °C (B) next. Different colors indicate data at different forces. The results revealed stable RecA nucleoprotein filament at 37 °C and unstable RecA nucleoprotein filament at 24 °C at low force, consistent with results obtained on large λ-DNA (Figure 8.2A-B).

Further, we examine the detailed dynamics of RecA-dsDNA nucleoprotein filament affected by temperature. At a high temperature of 37 °C with 50 mM KCl, 10 mM $MgCl_2$ (Figure 8.3A), the detailed saw-tooth dynamics were also observed at ∼ 40 pN. However, no net de-polymerization of RecA filament was observed over the experimental time scale, which is consistent with that higher temperature facilitates polymerization of RecA-dsDNA filament therefore shifts the balance and maintain the stability of RecA-dsDNA filament. Next, we further reduced the force to ∼ 3 pN, the saw-tooth fluctuation of RecA-dsDNA filament sustained with an extension of 1.5 times that of B-DNA, indicating a fully polymerized RecA-dsDNA with dynamic de-polymerization and re-polymerization.

Next, we repeat the experiment at 24°C (with other conditions unchanged) on the same DNA tether(Figure 8.3). The resulting saw-tooth DNA extension

dynamics at ∼ 40 pN and net RecA de-polymerization at low forces are consistent with findings shown in Figure 8.1A and Figure 8.2A. In addition, the pH effects on the dynamics of RecA-dsDNA filament is also re-examined with short DNA tether and consistent observations were obtained.

In summary, in this section, by examining the detailed dynamic of RecA-dsDNA filament at nano-meter resolution, we found a saw-tooth polymerization/de-polymerization dynamics pattern of RecA-dsDNA filament. These results indicate that RecA-dsDNA filament de-polymerization process involves highly stochastic, abrupt, stepwise events with extension reduction step size of ∼5-40 nm, corresponding to ∼ 10-80 bp DNA at per de-polymerization event. In other word, each RecA de-polymerization event is a highly cooperative, stochastic process involving ∼ 10- 80 bp DNA patches covered by 3-20 RecA monomers. In contrast, RecA re-polymerization on dsDNA seems to be much slower with no clear steps determinable by our experimental approach.

8.3.3 The elongated S-DNA is not the binding template for RecA nucleoprotein filament

As introduced previously, DNA overstretching transition at ∼ 65 pN is usually utilized to assist initial RecA nucleation and/or polymerization afterwards on dsDNA. However, the facilitation mechanism is still unclear. Since the DNA overstretching transition at ∼ 65 pN of torsion-unconstrained DNA in fact involves two different structural changes: 1) 'strand-peeling' by which one strand is peeled-off and coiled while another strand remains under force; 2) 'B-to-S' transition, which produce a elongated, base-paired dsDNA, now normally termed as S-DNA [92–97]. Note that, another overstretching form termed as 'melting bubble', or 'inside-strand-separation' of dsDNA where the two strands of ssDNA parallelly shared the force is not considered in our experiments since the designed DNA ensures it is not the case in our experimental condition [92–96]. Further, in physiological conditions at forces of ∼65 pN, strand-peeling and B-to-S transitions may co-exsist, whose selection is tuned by base-pair stability difference due to salt-concentration, temperature, and GC percentage [92–96]. B-to-S transition out-competes the other transition under conditions of higher base-pair

stability [92–95]. While ssDNA is known binding template for RecA filament, the elongated S-DNA is also proposed to be a suitable template for RecA filament [18, 89, 92]. Hence, the question remaining unclear is that whether S-DNA is the binding template for RecA nucleation and polymerization.

To test this hypothesis, we examined the potential RecA binding on pure S-DNA after B-to-S transition at $\sim$ 65 pN (Figure 8.4). The two DNA are specially designed to ensure B-to-S transition by using GC-rich handles or closed-ends, both of which prevents strand-peeling and resulted in pure S-DNA, which lacks extension hysteresis between force increasing scan and force decreasing scan (Figure 8.4A) [92–96].

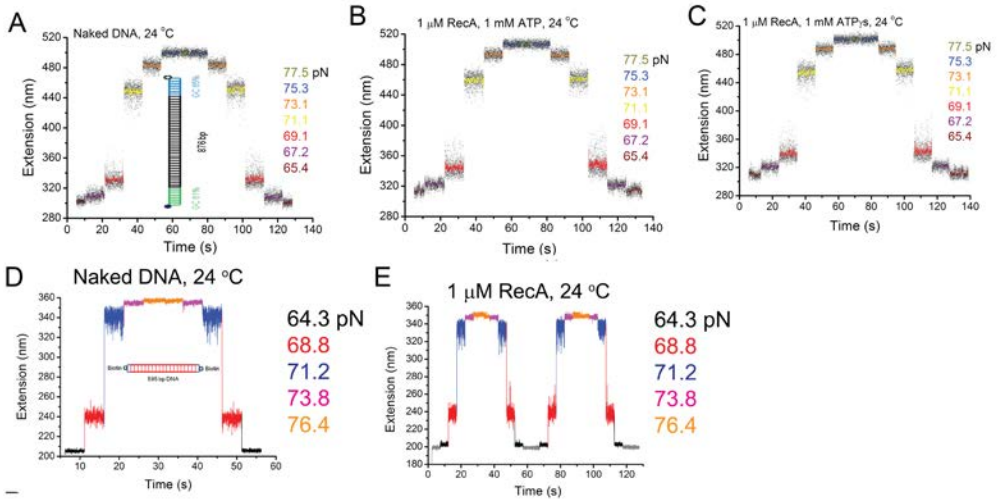

Figure 8.4: **The S-DNA produced in the B-to-S transition does not promote RecA nucleoprotein filament formation**. (A) Time traces of the extension of an 876 bp DNA tether with two GC-rich ends recorded during B-to-S transition in force-increase scan and the reverse S-to-B transition in the subsequent force-decrease scan in 50 mM KCl, 10 mM $MgCl_2$, pH 7.4, at 24 °C. The transitions are completely reversible, indicated by the same extensions recorded at the same forces during the force-increase and force-decrease scans. Inset on the top shows a sketch of the 876 bp DNA; (B-C) Results identical to those in (A) were obtained when the same experiments were repeated on the same DNA tether in 1 μM RecA and 1 mM ATP, 1× ATP regeneration system (B) or in 1 μM RecA and 1 mM ATPγS (C), indicating no RecA filaments formed on the S-DNA produced by the B-to-S transition within the experimental time scale.

Next, we introduced 1 μM RecA with ATP (B) or ATPγS (C) to the DNA, and monitored the extension dynamics of the DNA during and after B-to-S transition (Figure 8.4B&C). The resulting extension in the presence of RecA were identical with that in the absence of RecA, indicating that RecA did not nucle-

ated and polymerized during and after B-to-S transition. Since RecA was always start to nucleate and polymerize within seconds during DNA overtretching in identical solutions as shown in previous sections, it is safe to conclude that S-DNA is not a preferable binding template for RecA nucleoprotein filament. In other words, the often used DNA overstretching facilitated RecA nucleation is mainly through RecA binding to ssDNA produced during the transition.

8.3.4 The molecular nature of RecA nucleoprotein filament on dsDNA

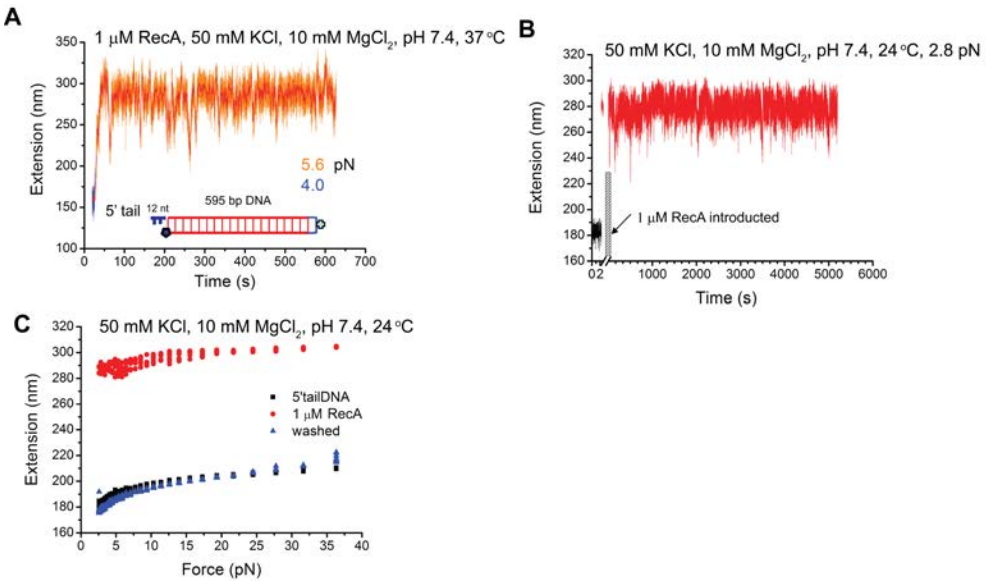

Figure 8.5: **Effects of ssDNA 5′ overhangs on the formation of RecA nucleoprotein filaments**. (A) Time traces of the extension of one end-closed 595 bp DNA with a 12 nt 5′ ssDNA tail in 1 μM RecA, 1 mM ATP, 1×ATP regeneration system , 50 mM KCl, 10 mM $MgCl_2$, pH 7.4, at 37 °C. A stable RecA nucleoprotein filament was formed after a spontaneous RecA polymerization at ∼5 pN. (B) Time traces of a DNA with the same structure as the DNA in (A) at 24 °C (other conditions remained unchanged). A stable RecA nucleoprotein filament was also observed for over 5000 second, after a spontaneous RecA polymerization at ∼3 pN. The shadowed area represents the process of solution flow with RecA.(C). Force responses of a 595 bp one-end closed DNA with a 12 nt 5′ ssDNA (Black), and the same DNA with RecA filament formed (Red). The extension of the DNA formed with RecA filament is about 50% longer than that of naked DNA before RecA was introduced. After remove the RecA by exchanging to pure buffered solution, the RecA filament de-polymerized and resulting DNA extension (blue) overlaps with naked DNA extension.

A prediction from the foregoing conclusion is that a pre-existing ssDNA-overhang would facilitate spontaneous nucleation polymerization of RecA at low forces

under conditions favoring net RecA polymerization. In addition, it has been widely accepted that RecA polymerize on DNA predominately in 5′-to-3′ direction. Therefore, we tested it by monitoring the extension dynamics of a short dsDNA (595-bp) containing a 12-nt 5′ ssDNA-overhang in standard reaction solution at 37^{o}C. Note that the selection of 12 nt is because that a minimal biding site for initial RecA nucleation has been demonstrated to ∼9 nt [19].

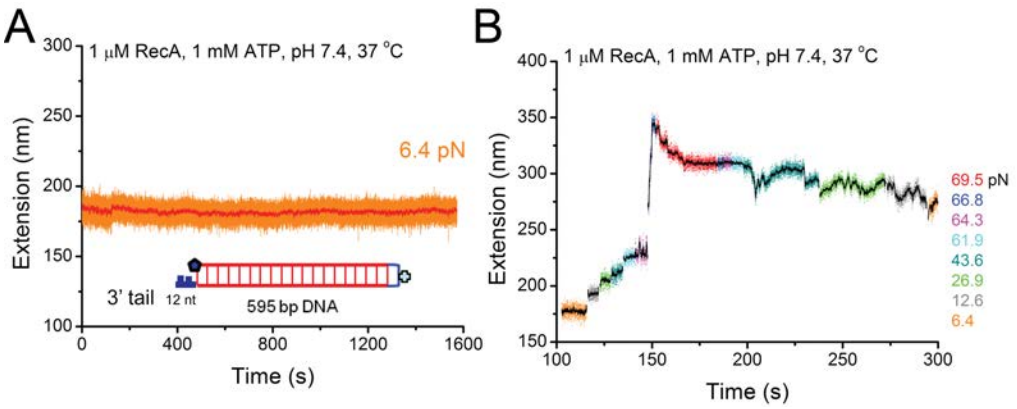

Figure 8.6: **Time traces of the extension of a 595 bp DNA with a 12 nt 3′ ssDNA tail and another end sealed** (A) Time traces of the extension of the DNA in 1 μM RecA, 50 mM KCl, 10 mM $MgCl_2$, 1 mM ATP, 1×ATP regeneration system, pH 7.4, and 37 ^{o}C. Within the experimental time scale of 1600 seconds, the DNA extension remained at the B-DNA extension, indicating that 3′ ssDNA overhang did not promote RecA filament formation at low force. (B) Time trace of the same DNA in the same solution and temperature condition, when the force was subsequently increased to $>$ 60 pN where DNA overstretching transition occurred, the RecA polymerization immediately started and the resulting RecA filament was stable. When the force was reduced to 6.4 pN, the DNA extension was still ∼120 nm longer than the B-DNA before RecA polymerization, indicating a stable RecA filament at low force.

As shown in Figure 8.5A, at ∼5.6 pN, when RecA was introduced to the DNA with 5′-ssDNA handle, RecA spontaneously and rapidly polymerized on the DNA, resulting in a steady fully polymerized RecA filament within ∼40 seconds, indicated by ∼120 nm longer (∼50% to B-DNA) extension compared to that before protein introduction. Further, the fully polymerized RecA-dsDNA filament also underwent large rapid dynamics while the overall extension remained that of a fully polymerized RecA filament over a long time (∼ 600 seconds tested).

In addition, we repeat similar experiments with the DNA construct with a 12nt 3′-ssDNA-overhang at the same solution conditions. No spontaneous RecA nucleation and polymerization was observed over 1500 seconds tested at low

force of $\sim$ 6.4 pN (Fig. S7A). Besides, the much less dynamic fluctuation of the DNA extension ($\sim$ 4 nm standard variance) compared to that with RecA-dsDNA filament (18 nm standard variance) also consistently suggests that RecA did not polymerize/de-polymerize on this DNA structure at this condition. Future, we facilitated the nucleation and polymerization of RecA on this DNA by increasing force to $>$ 60 pN, consistently, RecA started to polymerize and results in stable RecA filament on dsDNA (Figure 8.6). The results further support that a 5′-to-3′ direction polymerization of RecA along ssDNA is pre-dominated.

Surprisingly, rapid spontaneous RecA polymerizationwith on DNA containing 12-nt 5′-ssDNA at 24^{o}C in the same solution condition was observed (Figure 8.5B). The resulting fully polymerized RecA filament ($\sim$ 110 nm longer ($\sim$ 50%) than B-DNA (Figure 8.5B&C) also underwent rapid large dynamics while maintaining an overall constant averaged extension over $\sim$5000 seconds tested (Figure 8.5B). In addition, after washing away RecA and ATP in solution, bound RecA de-polymerized and the DNA returned to B-DNA (Figure 8.5C).

This result is interesting because that RecA was not able to nucleate and polymerize spontaneously on blunted-ends DNA at the same solution condition as shown in Figures 8.1A, 8.2A, and 8.3B. More importantly, at the same condition, a pre-existing fully polymerized RecA filament on blunted-ends DNA is unstable with a step-wise net de-polymerization. These inconsistencies of the dynamics of recA filament formed on ssDNA-tailed-DNA and blunted-ends DNA highlight the importance of the ssDNA tail on the stability of RecA filament. The tail role can be understood as following: RecA nucleates and polymerized on the 5′ ssDNA tail, forming a stable RecA-ssDNA filament at the 5′ end region; the filament is able to outcompetes the re-annealing of DNA pairing a pre-formed stable RecA-ssDNA filament on the 5′ end outcompetes the re-annealing of DNA pairing by hydrogen bonds to a complementary sequence, thereby continuously invading into dsDNA region along 5′ to 3′ direction and forming a stable RecA-dsDNA.

8.4 Discussion

In this chapter, we systematically investigated how physiological factors such as temperature, pH, force and ionic strength, as well as the topology of DNA govern the RecA filament polymerization/de-polymerization dynamics on individual dsDNA at nano-meter scales. We showed that the balance of the competition between polymerization and de-polymerization of RecA filament is delicately regulated by these factors over wide physiological ranges, which reconciled a large set of previous results with apparently discrepancy. We summarized the effects of these factors in Table 8.1 and Figure 8.7. Furthermore, the molecular nature of dynamics and stability of RecA filament is uncovered based on the findings, which will be detailed below.

We note that there are still many other regulatory factors on the dynamics of RecA-dsDNA nucleoprotein filaments, such as the nucleotide types (ADP, ATP, ATPγS, etc) and concentrations, divalent salts ($CaCl_2$, $MgCl_2$), torques on DNA [28, 29, 147]. Also, RecA from different bacterial species may have different kinetics and stability when they form filaments on dsDNA. These additional factors will be further investigated in our future studies.

We found that the ssDNA generated due to ssDNA stand-separation during DNA overstretching transition is the binding template for RecA nucleation, therefore, alleviates the energy barrier. This is further supported by the finding that RecA can spontaneously nucleate and polymerize onto dsDNA with a $5'$ ssDNA overhang. Further, these findings suggest that RecA-dsDNA nucleoprotein filament formation is likely started from an initial nucleation on one ssDNA strand of the dsDNA and then invaded via this ssDNA strand into dsDNA and polymerizes along the same stran in a $5'$-to-$3'$ direction, while another strand interacting with the RecA on the leading strand through secondary binding sites on RecA, resulting in a polymerized RecA nucleoprotein filament on dsDNA. This culminates in formation ssDNA-RecA-ssDNA co-filament, consistent with the model proposed by Pugh and Cox [147]. Note that in this study, we still refer this filament as RecA-dsDNA filament.

Moreover, the dynamics and stability of a RecA-dsDNA nucleoprotein fila-

Summary of major results of RecA nucleation, polymerization on dsDNA and its stability		
Experimental Conditions	DNA structures	RecA nucleation, polymerization, and stability
24 °C, pH 7.4	DNA with blunt ends: 48,502 bp λ-DNA; or 595 bp one-end-sealed DNA	1.Nucleation requires force-induced DNA strand-peeling transition; 2. Polymerization requires high force (> 40 pN); 3. Pre-formed RecA filament is unstable at forces of several pN;
24 °C, pH 7.4	DNA with 5′ssDNA 12-nt tail	1. Spontaneous nucleation and polymerization without assistance of DNA strand-peeling; 2. Pre-formed RecA filament is stable at forces of several pN;
24 °C, pH 7.4	876 bp DNA with two GC rich handles; or 600 bp GC rich end-closed DNA	Nucleation and polymerization do not occur during DNA B-to-S transition;
37 °C, pH 7.4	DNA with blunt ends: 48,502 bp λ-DNA; or 595 bp one-end-sealed DNA	1.Nucleation requires requires force-induced DNA strand-peeling transition; 2. Progressive polymerization occurs at forces of several pN; 3. Pre-formed RecA filament is stable forces of several pN;
37 °C, pH 7.4	DNA with 5′ ssDNA 12-nt tail	1. Spontaneous nucleation and polymerization without assistance of DNA strand-peeling; 2. Pre-formed RecA filament is stable at forces of several pN;
24 °C, pH 6.2	DNA with blunt ends: 48,502 bp λ-DNA; or 595 bp one-end-sealed DNA	1.Spontaneous nucleation and polymerization at forces of several pN; 2. Pre-formed RecA filament is stable at forces of several pN;

Table 8.1: Summary of major results of RecA nucleation, polymerization on dsDNA and its stability

ment is in fact governed by three main factors: 1) nucleation and polymerization of RecA on the leading ssDNA strand, 2) de-polymerization of RecA on the leading ssDNA strand due to ATP hydrolysis, 3) base-pair stability of dsDNA/ re-annealing of base paired strands. The environmental factors, including pH, temperature, force and salt concentration, as well as DNA topology are in factor selectively tuning one or all off these three pre-dominate factors. For instance, lower pH facilitates dimerization of RecA, and hence assists nucleation and polymerization of RecA on the leading ssDNA strand. Higher temperature and higher force lower the stability of dsDNA, at the same time facilitate polymerization of RecA along the leading ssDNA strand and reduce de-polymerization of RecA from the leading ssDNA strand. A 5′-end ssDNA tail ensures RecA nucleation and polymerization along the tailed ssDNA strand, furthermore, due to lacking a complementary ssDNA strand to base pair the tail, the resulting RecA-dsDNA filament is stablly capped by a RecA-ssDNA filament at the tail region. In addition, RecA cannot nucleate and polymerize on elongated base-paired S-DNA is consistently due to the increased stability of dsDNA in S form state.

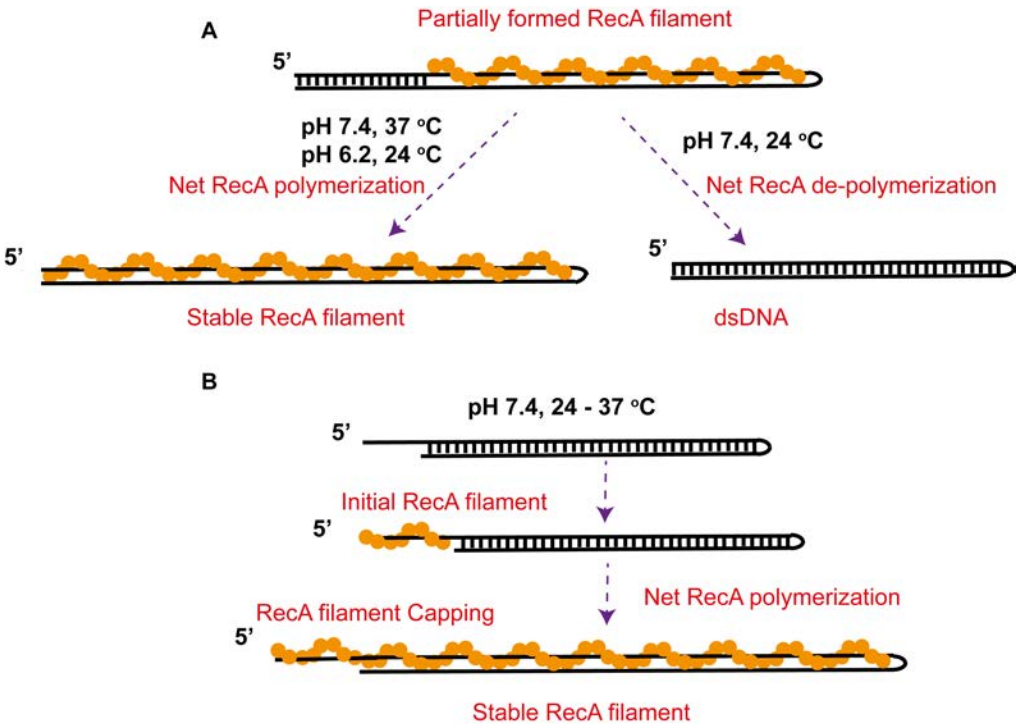

Figure 8.7: **Mechanistic models of the stability of RecA filaments formed on dsDNA**. (A) On dsDNA with blunt ends, at 37 °C and pH 7.4 or 24 °C and pH 6.2, polymerization of a partially formed RecA filament outcompetes DNA re-hybridization, leading to a net extension of the RecA filament into a stable fully coated RecA filament. In contrast, at 24 °C and pH 7.4, DNA re-hybridization outcompetes RecA polymerization, leading to a net de-polymerization of the RecA filament into a stable B-form DNA. (B) On dsDNA with a 5′ ssDNA overhang that provides sites to initial RecA nucleation and polymerization, invasion of the RecA filament formed on the ssDNA overhang into the dsDNA region occurred in 24-37 °C and pH 7.4, leading to stable fully coated RecA filament explained by the end-capping mechanism discussed in the text.

In summary, our findings based on systematic examination of the dynamics and stability of RecA-dsDNA nucleoprotein filament regulated by a set of environmental physiological factors, including pH, force, temperature, DNA overhangs, not only reconcile previous conflicting results from separated studies, but also advanced our understanding of the molecular nature of dynamics and stability of RecA nucleoprotein filament formed on dsDNA in various physiological microenvironments, which further provide new insights to RecA activities during homologous recombination *in vivo*.

Chapter 9

Other studies

I–Mechanosensitive regulation of single specific protein-DNA (IHF-H′) complex

9.1 Chapter summary

Force has been recognized as one of the critical determinants in many crucial biological processes. In previous chapters, we have demonstrated its potential critical role in dynamics and regulation of RecA nucleoprotein filament during DNA repair and recombination. Due to the ubiquitous presence of force in vivo, mechanical force also likely affects gene regulation. To test this possibility, we systematically examined how force may affect the binding dynamics of *E. coli* integration host factor (IHF), an important transcription factor, to its specific binding sequence at a single protein-DNA complex level in real time. We found that, pico-Newton force drastically re-defined the stability of this specific protein-DNA complex. These results indicate that force is likely one of the critical physiological factors for transcription regulation. In addition, we demonstrated the first time that single short DNA manipulation using magnetic tweezers can resolve single site specific protein-DNA binding dynamics and can be further

used to quantify DNA bending angle by a single protein. Overall, this study has broad implications for site-specific DNA distorting proteins regarding their mechanosensitivity. *

9.2 Introduction

DNA in cells is the main bank storing the genetic information that can be accessed by gene transcription. Gene transcription is tightly governed by various protein-DNA interactions. Particularly, site-specific DNA-protein binding plays crucial role in the regulation of gene transcription. For instance, to initiate gene transcription, RNA polymerase is recruited to some specific promoter sequences where a variety of transcription factors binds to as well to regulate transcription level [151, 152]. An estimation of ∼400 transcription factors that bind to DNA in *E. coli* has been reported previously [152, 153]. Furthermore, Site-specific DNA binding by proteins is also key factors in many other cellular processes such as DNA replications, site-specific recombination [154, 155]. Hence, detailed information of these site-specific DNA binding by proteins as well as their regulatory factors will enhance our understanding of their cellular functions. It has been well known that various cellular proteins and environmental factors including pH, temperature and salt osmolarity sensitively regulate the activities of these site-specific DNA-protein complexes [152]. For decades, these microenvironmental factors have been extensively studied by biochemical, biophysical approaches [152].

However, a gap still waits to be bridged due to lack of investigation on the potential regulatory roles of force in site-specific DNA-protein interactions during DNA transactions. As introduced in Chapter 1, mechanical force has now been established to be a key dominates in diverse cellular processes. Force *in vivo* can be generated by various cellular machineries. Particularly, in nucleus, forces up to ~ 30 pN on DNA can be exerted by DNA/RNA polymerases [72, 73]. In addition, due to the attachment of the nucleoid and the cell wall [6], tension in DNA is

*Note that main contents detailed in this chapter have been published in *Mechanosensing of DNA bending in a single specific protein-DNA complex.* Le S. *et al.*, Scientific Reports 3-3508 (2013).

likely built up during DNA packaging. Furthermore, forces of pN are expected based on $\sim k_{\mathrm{B}}T$ range interaction energy in nm scale interaction distance between protein and DNA. Considering the ubiquitous presence of forces on chromosomal DNA, force is expected to influence the site-specific DNA-protein interactions. To test it, we systematically probed the effects of force on a specific DNA binding by *E. coli* integration host factor (IHF), an important gene transcription factor [152–154].

Discovered decades ago as an essential factor in site-specific recombination of phage λ gene into *E. coli* genome [154], IHF is now known to also regulate gene transcriptions in *E. coli* [156] and *S. typhimurium* [157]. The capability of IHF to sharply bend several specific sequences plays major roles for its functions [158, 159]. By inducing sharp bending of these sequences, it bridges remote sequences, thereby promotes long-distance interactions. A so-called H′ sequence (34-bp) has the highest binding affinity to IHF with a dissociation constant reported to be in a range of 2-20 nM based on ensemble biochemical measurement [155, 160], and < 1 nM reported by stop flow measurements [161]. Furthermore, the sharp bending angle of H′ by IHF binding has also been determined with a value of $\sim 160^{\mathrm{o}}$ (X-ray crystallization [158]), or $> 120^{\mathrm{o}}$ (atomic force microscopy (AFM) imaging [162–164]).

IHF-H′ specific binding has been intensively studied by various approaches including isothermal titration calorimetry (ITC) [165], tethered particle motion [166], electrophoretic mobility shift assays (EMSAs) and footprinting [154–156], fluorescence resonance energy transfer (FRET) [161], stop-flow fluorimetry [159] and temperature jump [167]. These studies have revealed that both specific and non-specific H′ binding modes may occur; the selection of the two modes depends on protein concentration and salt concentration: high protein concentration and low salt concentration favor non-specific binding mode, while low protein concentration or high salt concentration favors another [165]. Additionally, intermediate binding modes are also been proposed based on recent time-resolved FRET [161] and stop-flow fluorimetry [159] experiments.

However, these experiments were performed in the absent of force applied on DNA; hence, the potential role of physiological force on regulation of specific

IHF-H′ nucleoprotein interaction remains unexplored. To address this remained unknown, we systematically investigated the effects of force on dynamics and stability of IHF-H′ interactions in different temperature, salt and IHF concentrations at a single H′ DNA level. Additionally, we determined the value of H′ DNA bending angle induced by IHF binding in physiological solution.

9.3 Results

9.3.1 Force-sensitive two-state conformational fluctuation of IHF-H′ complex

The experimental principle is that when an IHF protein binds to and bends a H′ sequence in DNA tether under force, the DNA extension would reduced by certain level Figure 9.1A. Due to thermal noise from longitude fluctuation of DNA, the DNA has to be short enough to increase the signal-to-noise, in order to monitor the extension changes induced by IHF binding. In the experiments, based on theoretical calculation, we designed a 455 bp DNA (contour length of $\sim$ 150 nm) with a single H′ sequence in the middle, while other parts of the DNA contain no known high-affinity IHF-binding sequences. Parallel experiments on the same length DNA lacking the H′ sequence were performed as a negative control. Dynamics of DNA extension were monitored by magnetic tweezers [120] introduced in Chapter 2.

Figure 9.1B shows a set of representative DNA extension time traces in the presence of 10 nM IHF in solutions containing 50 mM KCl, 2.5 mM $MgCl_2$, 10 mM Tris (pH 7.4) at 21 $\pm$ 1 ^{o}C with forces applied on DNA in the range of 0.5-1 pN. We observed clear two-state fluctuations of DNA extension between a shorter extension (refer as bent DNA), and longer extension (refer as unbent DNA). The raw extension dynamics can be well digitized (red) using a noise-beating step-finding algorithm (see methods). The two-state extension fluctuations were neither observed before introduction of IHF nor with the control DNA lacking the H′ sequence in the presence of IHF, we therefore, conclude that these two-state fluctuations are both IHF and H′ sequence specific. Additionally, as another negative control, we measured the force extension curves of a much longer DNA

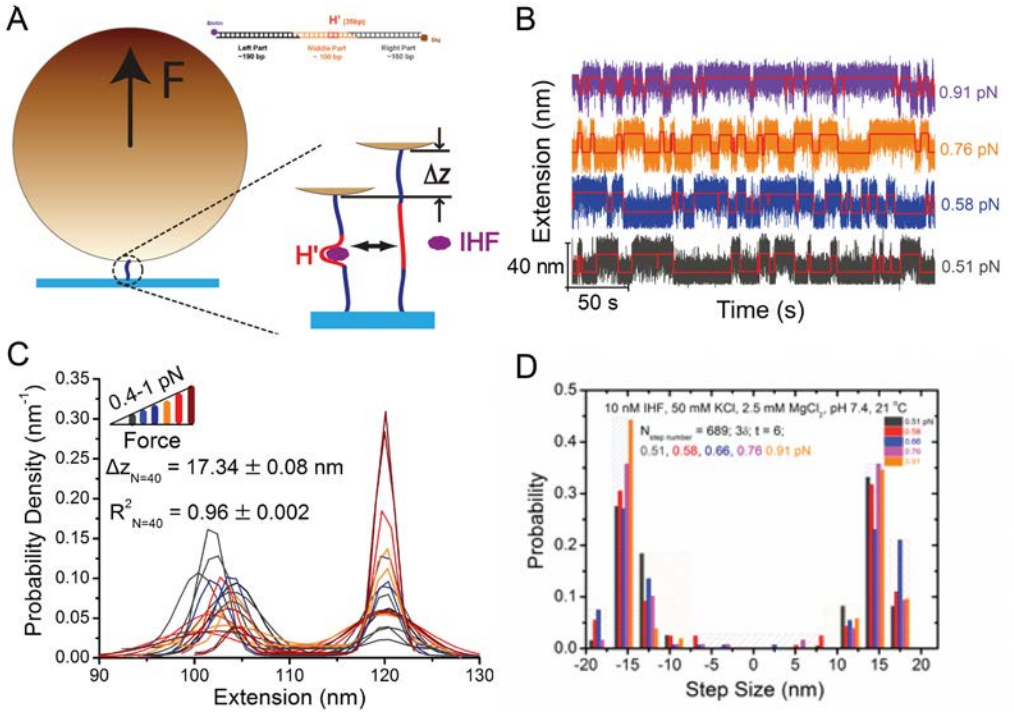

Figure 9.1: **Two-state fluctuation of DNA-IHF interaction**. (A). A sketch of the principle of the measurement. The H′ DNA bending by IHF binding reduces the DNA extension, which is detected by the resulting change of the height of the bead. (B). Dynamic fluctuations between two extensions in 10 nM IHF, 50 mM KCl, 2.5 mM $MgCl_2$, 10 mM Tris (pH 7.4), and 21 °C at different forces (0.5-1 pN) indicated by different colors. The red line represents extension steps detected a noise-beating step-finding algorithm (introduced in Chapter 4). (C). The probability densities of the two extension states, which were produced by the double Gaussian fitting of the normalized histogram of smoothed data at different forces (0.4-1 pN, represented by different colors) using a bin size of 1 nm. A statistics of 40 distributions gave a step size of 17.34 ± 0.08 (mean ± s.d.) nm and $R^2 = 0.96 \pm 0.022$ (mean ± s.d.). Note, only 13 fitting lines from 5 DNA molecules were shown in figure for clarity. (D). Distribution of extension increasedecrease steps detected by noise-beating step-finding algorithm.

(48,502 bp λ-DNA, ∼16 mm in contour length) in the absence and presence of 10 nM IHF protein at the same solution condition, the resulting extension curves are indistinguishable from each other, indicating that non-specific binding of IHF to other DNA sequence has no observable effect in this condition (Figure S3 in Ref. [111]). We note that there are four known high affinity consensus sits on λ-DNA [154, 160], whose contribution to force response of the vast DNA is negligible [101].

We then analyzed the normalized histogram of apparent bimodal extension distributions by fitting with double Gaussian distribution (Figure 9.1C). Calculating the peak-to-peak distance of the fitted histograms from >40 extension

distributions at the force range, the extension difference of these two conformational states of DNA (bent and unbent, respectively) is determined to be 17.34 ± 0.08 nm (mean ± s.d.). Remarkably, sub-pN force change dramatically switched the balance of the two conformational state of DNA, suggesting that the IHF-H′ specific interaction is fined-tuned by pN scale forces. We note that, in addition to the obvious two-state fluctuation, Detailed extension analysis show several minor sub-steps, which might related to transient intermediated binding states of IHF-H′ complex (Figure 9.1D).

9.3.2 Force-sensitive stability of IHF-H′ interaction regulated by temperature, KCl concentration and IHF concentration

Further, the probability of bent-DNA state, P_{bent}, is the function of force and is related to free energy difference between the two conformational states through Boltzmann distribution:

$$P_{\mathrm{bent}} = \frac{e^{\frac{\Delta G - \Delta z \times f}{k_{\mathrm{B}}T}}}{1 + e^{\frac{\Delta G - \Delta z \times f}{k_{\mathrm{B}}T}}} \tag{9.1}$$

In above equation, P_{bent} can be calculated based on the relative area ratio of the bimodal extension distribution, Δz is the extension difference between the two conformational state, f is applied force, ΔG is the zero-force Gibbs free energy difference between the two conformational state of DNA. The intrinsic stability of IHF-H′ interaction can be described by a protein concentration independent quantity, the dissociation constant, K_{d}. The lower the K_{d}, the stronger the interaction. K_{d} is related to ΔG through: $K_{\mathrm{d}} = C \times e^{-\frac{\Delta G}{k_{\mathrm{B}}T}}$. Therefore, Eq.9.1 becomes:

$$P_{\mathrm{bent}} = \frac{\frac{C}{K_{\mathrm{d}}} \times e^{\frac{\Delta G - \Delta z \times f}{k_{\mathrm{B}}T}}}{1 + \frac{C}{K_{\mathrm{d}}} \times e^{\frac{\Delta G - \Delta z \times f}{k_{\mathrm{B}}T}}} \tag{9.2}$$

The above equation explicitly shows that P_{bent} depends on force, dissociation constant, and protein concentrations. C and f were experimental parameters, and P_{bent} and Δz were measured; hence the free parameter K_{d} can be determined by fitting the experimental data to Eq.9.2. We note that, in our fitting the

narrowly distributed Δz in our narrow force range is set to be a constant (the average value) for simplicity.

Based on above equations, we then analyzed the effects of force, temperature, salt concentration and protein concentration on the stability of IHF-H′ interaction through calculating the value change of K_{d} (Figure 9.2). First, temperature increases from 21 ± 1 °C (dark gray) to 31 ± 1 °C (red) in 50 mM KCl, resulting in an increase of K_{d} from 0.44 ± 0.06 nM to 1.54 ± 0.08 nM, indicating that higher temperature destabilizes the bent DNA state.

Next, we examined the effect of salt concentration on the stability of IHF-H′ complex. However, we found that the two-state conformational fluctuation was absence at an increased KCl concentration of 200 mM at 21 °C (while other conditions remain unchanged. This is likely due to decreased binding affinity at increased salt concentration. To test it, we further increase the concentration of IHF to 100, 500 and 1000 nM. Consistent with the hypothesis, the two-state fluctuations were observed at higher protein concentration. Further, we calculated the force-dependent Pbent (symbols, Figure 9.2) and K_{d} from fitting based on Eq. 9.2 (lines, Figure 9.2). The three K_{d} values estimated at 200 mM KCl with 100, 500 and 1000 nM IHF protein were overall consistent with a value of 28.5 ± 3.9 nM (mean ± s.d.). Additionally, Pbent at 10 nM IHF, 200 mM KCl in the force range of 0.5-1 pN therefore, can be predicted based on the estimated Kd, to be <5%. This low probability consistently explained why on obvious two-state fluctuation was observed in 10 nM IHF at 200 mM KCl.

Interestingly, when the free IHF in solution was removed, the two-state fluctuation was still sustained for a long time (over hundreds of seconds, Figure 9.3). This is consistent with previously suggested possible unbent intermediate conformational states [159, 161]. Further, this result also consistent with the much slow-downed dissociation of proteins from DNA when no free protein in solution [168].

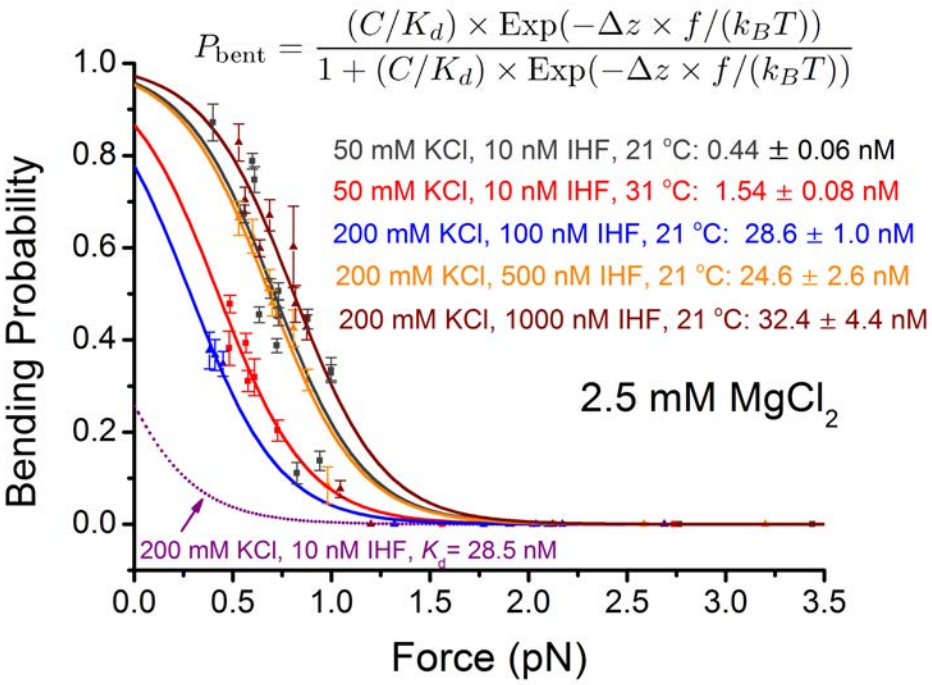

Figure 9.2: **Stability of IHF-H′ complex affected by temperature, KCl concentration and IHF concentration**. The bending probability as a function of force in 2.5 mM $MgCl_2$, 10 mM Tris (pH 7.4), at 10 nM IHF, 50 mM KCl, 21 °C (dark gray symbols), 10 nM IHF, 50 mM KCl, 31 °C (red symbols), 100 nM IHF, 200 mM KCl, 21 °C (blue symbols), 500 nM IHF, 200 mM KCl, 21 °C (orange symbols), and 1000 nM IHF, 200 mM KCl, 21 °C (wine symbols). Data for each solution condition were obtained from multiple (≥3) independent DNA molecules. Error bars for each data points (symbols) were standard deviations from multiple (≥3) repeating measurements for the same DNA molecules. The bending probabilities were calculated by the relative area of the two species in the bimodal extension distribution. Under each solution condition, data obtained were fitted by the two-state model (Eq. 9.2) to obtain K_{d} and the standard error ofK_{d} (fitting error), which are indicated in figure panels by corresponding colors. The goodness of the fitting (R^2) are 0.84, 0.98, 0.94, 0.93, 0.99, respectively. The purple dot line is the theoretical calculation of the bending probability in 10 nM IHF, 200 mM KCl, 2.5 mM $MgCl_2$, based on the two-state model with an averaged K_{d} of 28.5 nM.

9.3.3 Theoretical prediction of the effect of bending angle and force on the extension of short DNA

X-ray crystallization structure of the specific IHF-H′ complex shows that an IHF hetero-dimer bent the 34-bp H′DNA over an angle of ∼160° [158], while the AFM images of H′-IHF complex estimated the bending angle to be >120° [162–164]. However, the two methods determined the angle based on static IHF-H′ complex conformations. Moreover, certain specific solution conditions are usually required

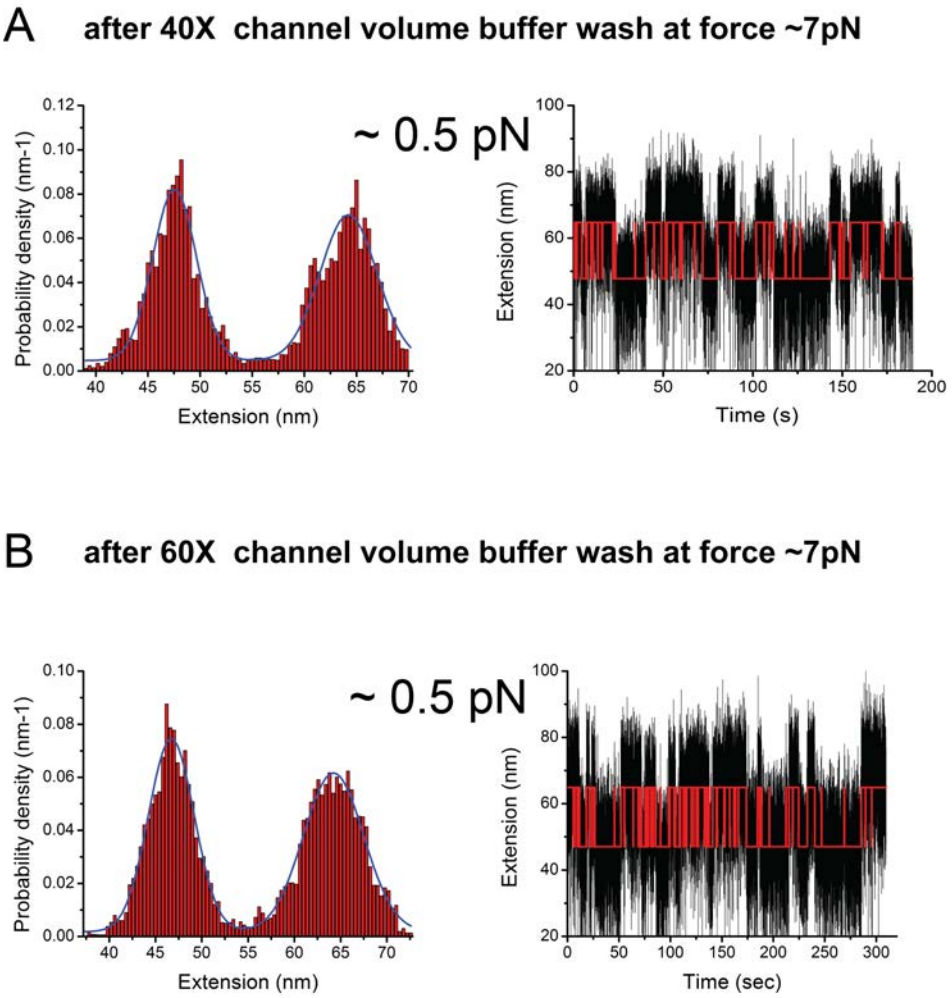

Figure 9.3: **Two-state fluctuation of IHF-H′ complex in the absence of free IHF in solution**. (A&B). The two-state extension fluctuation of an IHF-H′ complex in IHF-free solution after removal of free IHF by 40X or 60X channel volume solution flow.

for both methods. Here, we show that the bending angle can be estimated based on the dynamic two-state conformational fluctuations, where the IHF-H′ complex is in physiological solution conditions.

As a semi-flexible polymer, DNA can be well described by WLC model with a DNA bending persistence length, A $\sim$ 50 nm [78, 79]. Based on WLC model, DNA is a chain of N segments with a segment length of $b = L/NA$, where L is the contour length of the DNA. The bending energy cost of one segment (in units of $k_{\mathrm{B}}T$) is $E_i(\hat{t}_i, \hat{t}_{i+1}) = \frac{1}{2}\frac{A}{b}(\hat{t}_i - \vec{t}_{i+1})^2$, where $\hat{t}_i$ and $\hat{t}_{i+1}$ denote the tangent vector of the i_{th} and $(i+1)_{\mathrm{th}}$ segment respectively. Therefore, The total bending energy cost is $E = \sum_{i=1}^{N-1} E_i(\hat{t}_i, \hat{t}_{i+1})$, i.e., the sum of all the bending energies

cost of the segments.

The ∼150 nm DNA used in our experiments contains one H′ site for IHF specific binding. To estimate the effect of the bending of H′ site when associated with an IHF protein on the overall DNA extension reduction, the DNA is segmentized with $b = 1$ nm, and the IHF associated H′ site is modeled as a point-like site (located at the middle of the DNA) with a preferred bending angle described by a parameter γ: $\theta = \cos^{-1}\gamma$. Therefore, while the rest of DNA remaining in naked DNA state, the bending energy cost of kink site is modified as $E_{\text{kink}}(\hat{t}, \hat{t}') = \frac{a}{2}(\hat{t} \cdot \hat{t}' - \gamma)^2$ [101, 103, 148], where $\hat{t}$ and $\hat{t}'$ denote the two tangent vectors of the two successive kink-involved segments; a is the dimensionless parameter used to describe the deformability of the kink. A large value of $a = 50$ is chosen to ensure the high rigidity of the kink in the calculation. Note that the resulting force dependent extension reduction profiles were not sensitive to the value of a as shown in Figure 9.4D).

Next, we calculated the DNA force-extension curves with a set of values of θ $[0-180^o]$ at the kinked H′ site using transfer-matrix method [101, 103, 148](Figure 9.4A), and compared these curves with that of naked DNA lacking the kink defect. Four representative resulting force-dependent extension reduction Δz are plotted in Figure 9.4B, Note that the values of θ $[0 - 180^o]$ are converted from by $\theta = \cos^{-1}\gamma$ with γ in the range of $[-1, 0]$. The extension reduction Δz in our experiments is ∼17.34 nm in the force range of ∼0.4-1 pN where we observed the two-state DNA conformational fluctuations. By Comparing the force and extension reduction with the theoretical calculations, we can estimate a preferred bending angle of 140-180° for H′ sequence associated with an IHF protein. By an approximate analytical formula derived by Kulić *et al* [169].: $\Delta z = 4(1 - cos\frac{\theta}{4})\sqrt{\frac{k_{\text{B}}TA}{f}}$, we can also estimate the bending angle at the same range Figure 9.4C. We note that this approximate formula is only valid when forces $\gg \frac{k_{\text{B}}T}{A} \sim 0.08$ pN where the entropic DNA conformational fluctuation, which is not considered in the formula, is suppressed. The bending angle estimated based on force dependent DNA extension dynamics agrees with that observed in previous AFM imaging and X-ray crystallization experiments [158, 162–164].

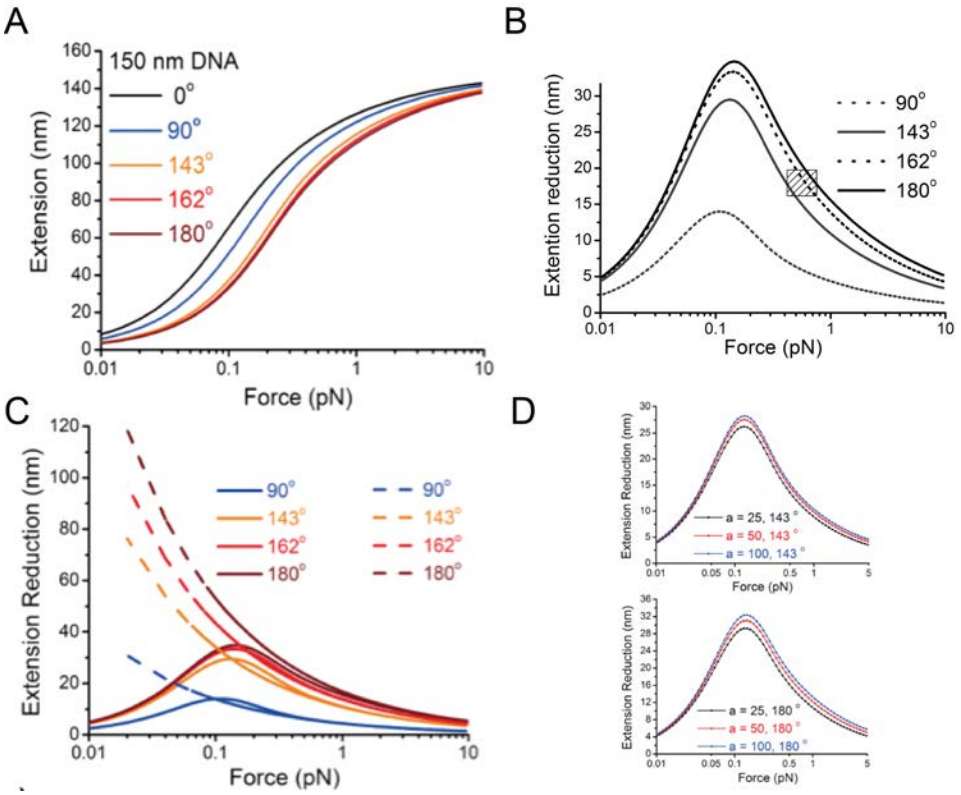

Figure 9.4: **Theoretical prediction of the effect of bending angle and force on the extension of short DNA**. (A-B). Theoretical prediction of the extension (A) and the extension reduction (Δz) (B) of a DNA with a contour length of 150 nm induced by a kink placed at the middle as a function of force and bending angles of 90 ^{o} (gray dotted line), 143 ^{o} (gray solid line), 162 ^{o} (black dotted line), 180 ^{o} (black solid line). The shadow area represents the rough force range and Δz range measured in experiments. (C). The force and bending angle dependent extension reduction calculated by an approximate analytical formula derived by Kulić *et al* [169] (dash lines), compared with data in panel B (solid lines). (D). Representative curves of the force and bending angle dependent extension reduction calculated with different value of $a = 25, 50, 100$ for $\theta =$ 143 or 180^o.

9.4 Discussion

In this chapter, we demonstrated that dynamics and stability of single specific DNA-protein interactions can be probed by single short DNA manipulation. Previously, the DNA used in single-DNA stretching experiments was usually $\sim$ microns or above [115, 170, 171]. To detect effects of protein-DNA interactions in these experiments, the protein density on DNA has to be high enough. For instance, an obvious force-extension curves shift can only be observed with at least $\sim$ one protein per 100 nm distance for proteins which bent DNA by $\sim$ 90^o [101]. This is mainly due to the large conformational fluctuation of DNA

with long contour length, which lower the signal-to-noise ratio. Hence, those previous single-DNA stretching experiments were mainly focus on non-specific DNA-protein interactions where high protein concentrations were used to ensure sufficient protein binding density.

On the other hand, specific DNA-protein interactions were previously often studied by biochemistry methods [154–156, 159, 165, 167]. Comparing to these methods, single short DNA manipulation has unique advantages to directly probe the dynamics of single protein-DNA interactions. Next, compared to other single-molecule fluorescence methods such as FRET [161], single short DNA manipulation is capable to monitor the interaction dynamics over longer time and distance scales while does not requires fluorescence labelling. Thirdly, compared to the tethered particle motion methods [166], single short DNA manipulation has unique advantages of well controlled force on DNA, which enables the study of potential role of force on the protein-DNA interactions. Overall, we demonstrated that single short DNA manipulation can be a powerful approach for studies of single protein-DNA specific interactions regulated by environmental factors including mechanical forces. This is a good complimentary to other biophysical and biochemistry methods.

The *in vivo* IHF concentration is estimated to be 12-55 μM [172] during different state, which is much higher than that required for specific interaction functions. Therefore, non-specific IHF-DNA interaction may also occurs *in vivo*. However, for non-specific IHF-DNA binding, each protein occupies less base pairs of DNA ($\sim$10 bp) [165], and induces much smaller bending ($<30^{o}$) of DNA [105], which is distinctively different to that of site-specific IHF-H$'$ interaction.

Despite that it might have difference between *in vitro* and *in vivo* dynamics and stability of protein-DNA interactions due to much-complicated environments *in vivo*, our finding of ultra force sensititivity of the IHF-H$'$ specific interaction demonstrated potential importance of of force in gene transcription regulation. Forces may be generated on DNA both actively by cellular machinaries or passively due to DNA packaging and other DNA distortion interactions [173]. Hence, the force sensitive interactions of IHF-H$'$ probably as well as other DNA binding transcription factors is likely physiologically important.

We showed that P_{bent} is over 95% at zero force in 50 mM KCl, 2.5 mM $MgCl_2$ and 21 °C, which agrees well with the reported stable IHF-H′ complex in previous experiments without force [155, 159, 164, 166]. A slight increase of force to ∼1 pN switches IHF induced DNA bending to be minority. Such ultra sensitivity of force-dependence in fact can be expected based on the large DNA conformational changes induced by protein binding: the large conformational change of DNA (∼17 nm in extension in our experiments) implies that perturbing this protein-DNA interaction only require a force of $\sim \frac{k_{\text{B}}T}{17\text{nm}} \approx 0.2$ pN which is consistent with the sub-pN force sensitivity discovered in experiments. Furthermore, as large DNA conformational changes are involved in many site-specific transcription factors and other proteins, and play key roles for their functions, such as sharp DNA bending or DNA looping [168, 173, 174], our findings generally suggest that force in pN physiological ranges on DNA may potentially regulate a broad class of protein-DNA interactions *in vivo*.

While the lower-extension state of DNA is induced by sharply bending of H′ sequence bound with an IHF protein, the longer-extension state might have two alternative possibilities: the IHF disassociates from DNA, or an intermediate state of IHF-H′ complex in which the H′ is in a non-bent conformation. The consistent values of K_{d} determined in our experiments with 10-1000 nM IHF protein in solution and previous ensemble biochemical experiments [155, 159] suggest that the observed two states are likely dominated by association and dissociation of an IHF to the H′ sequence. However, in the absence of free IHF in solution, the bound IHF on H′ can remain associated for a long time, fluctuating between bent and unbent states. This is consistent with previously suggested possible unbent intermediate conformational states [159, 161]. Further, this result also consistent with the much slow-downed dissociation of proteins from DNA when no free protein in solution [168].

Our results also show that the distortion level of a DNA (bending angle) induced by single site-specific protein binding can be estimated by analyzing protein binding/unbinding induced DNA extension fluctuations based polymer physics. The bending angle of H′ sequence induced by an IHF binding estimated by this approach (140-180°) is within 15% difference with that determined by

X-ray crystallization (∼160°) [155]. While the resolution is lower, this approach has a unique advantage that it is not limited to specific solution conditions and can probe the dynamics of bending angle.

Chapter 10

Other studies II–Mechanics and dynamics of CRISPR RNA-guided DNA bending and unwinding

10.1 Chapter summary

Clustered regularly interspaced short palindromic repeats (CRISPRs) are the molecular memory of an adaptive immune system that protects bacteria and archaea from invading viruses and plasmids. In *Escherichia coli.*, short CRISPR-derived RNA (crRNA) in a multi-subunit surveillance complex called Cascade (CRISPR-associated complex for antiviral defense) contains a 32-nt guide sequence that is essential for recognition of an invading DNA target. Here we use single-molecule manipulation and FRET methods to measure the mechanics and dynamics of DNA conformational changes upon target binding by Cascade. We show that hybridization of the Cascade-crRNA with complementary double-stranded DNA results in target bending of >100 degrees, target unwinding of 34 degree per bp, and stabilization of the displaced strand in an extended conformation. Interactions between Cascade and a DNA target are irreversible within our experimental time scale (>1000 seconds). Together these results provide new

insights into crRNA-guided strand invasion and R-loop formation. *

10.2 Introduction

Prokaryotes have evolved nucleic acid based adaptive immune systems that detect and destroy invading foreign nucleic acids [175–184]. These defense systems rely on clustered regularly interspaced short palindromic repeats (CRISPRs) and a diverse set of CRISPR-associated (cas) genes. Upon viral or plasmid challenge, short fragments of foreign nucleic acids, termed protospacers, are selectively integrated into the host CRISPR loci [176, 177, 185]. The selection of foreign protospacers for integration into the host CRISPR is performed by two conserved Cas proteins (Cas1 and Cas2) and relies on the detection of a short di- or tri-nucleotide motif called a protospacer adjacent motif (PAM) [186, 187]. The CRISPR locus is transcribed and the primary transcript is processed into short CRISPR RNAs (crRNAs) that each contains a unique guide sequence derived from a previously encountered foreign genetic element. The crRNAs are bound by Cas proteins to form crRNA-guided surveillance complexes. Some crRNA-guided surveillance systems identify foreign targets through recognition of the PAM sequence prior to complementary base pairing of the crRNA-guide to the DNA target [188]. Bound protospacer targets are then degraded by dedicated nucleases [175–184].

The crRNA-guided search and target binding steps are crucial in CRISPR-mediated adaptive immunity. Structures of the crRNA-guided surveillance complex from Escherichia coli have been determined using electron microscopy, and X-ray crystallography [189, 190]. The cryo-EM and X-ray structures reveal a conformational change in Cascade upon DNA binding, and atomic force microscopy (AFM) experiments indicate that Cascade bends the target DNA in a broad range of angles [191, 192]. Recent studies have revealed that crRNA-guided dsDNA target recognition results in a crRNA-ssDNA hybrid and a displaced ssDNA R-loop [189, 191–193]. Direct evidence of R-loop formation on single DNA molecules has been demonstrated by Szczelkun *et al* [119], in which they quantified the dynam-

*Note that main contents detailed in this chapter are included in *Mechanics and dynamics of CRISPR RNA-guided DNA bending and unwinding.* Le S. *et al.*, to be submitted (2015).

ics of torque-dependent R-loop formation by both Cascade and Cas9 complexes. They show that the stability of the R-loop increases as torque decreases. These results are consistent with DNA unwinding during R-loop formation; however, important information regarding how the DNA bending is related to R-loop formation and the dynamics of DNA bending remains unaddressed.

Here we measure the dynamics of DNA bending and its correlation with R-loop formation, the level of DNA unwinding during R-loop formation, and the micromechanics of the R-loop structure. In contrast to the previous single-molecule DNA stretching study where the DNA linking number (the number of times the two DNA strands wind around each other) was constrained, we used torsion-unconstrained DNA, which allows simultaneous detection of DNA extension dynamics and twist angle fluctuations during Cascade invasion. We show that Cascade rapidly (∼0.4 sec) reduces the extension of the DNA target by 15-20 nm at ∼ 1 pN force, which corresponds to bending of the DNA target by more than 100°. By measuring rotation angles between the two ends of a DNA tether, we observed protospacer DNA unwinding by 34° per base pair (bp).

10.3 Results

10.3.1 Cascade induced DNA bending

A broad range of DNA bending by Cascade has been reported using AFM imaging (Figure 10.1) [191, 192]; however, DNA binding has not been quantified in solution by analyzing DNA bending dynamics. DNA bending reduces the end-to-end distance of DNA (i.e., DNA extension) when DNA is held under weak forces [101, 103, 111, 148]. To measure the dynamics of bending, we stretched a single 509-bp fragment of DNA using magnetic tweezers and then added Cascade (Figure 10.2A and Methods). The addition of Cascade resulted in a rapid decrease in the extension by 15-20 nm at ∼1 pN of force, corresponding to Cascade induced DNA bending of over 100° (Figure 10.2B&C) [101, 103, 111, 148]. The decrease in DNA extension was only observed when Cascade was added to a DNA target that contains a protospacer sequence and a PAM (Data not shown). Cascade-induced DNA bending takes 0.43 ±0.21 seconds to complete after the

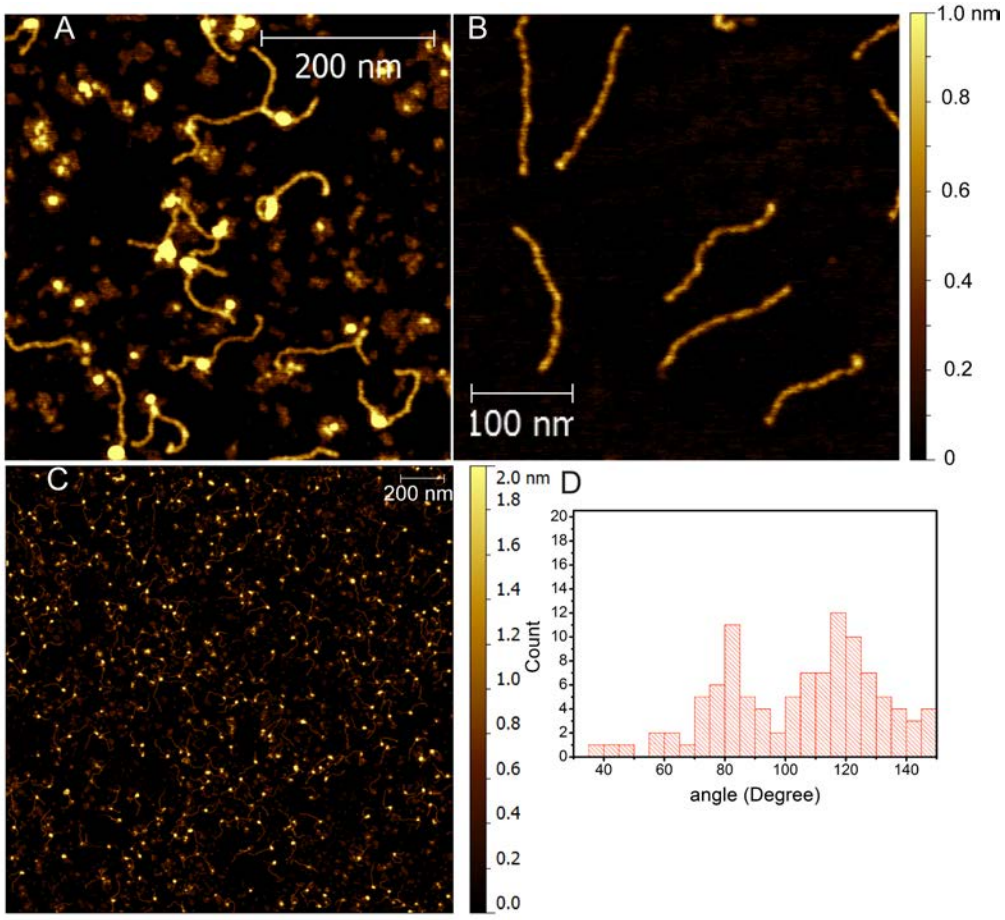

Figure 10.1: **AFM imaging of cascade induced DNA bending**. (A). Cascade-DNA complex. (B). DNA without Cascade. (C). a large area view of AFM images of Cascade-DNA complex. (D). Bending angle distribution of Cascade-DNA complex.

initiation of DNA bending (Methods - relaxation time scales). After bending, unbound Cascade was washed out of the chamber and the bent molecules were monitored for >1000 seconds. We monitored six single molecules for more than 1000 seconds and none of these ever returned to the relaxed state. These results suggest that the DNA-bound complex is highly-stable over long periods of time.

10.3.2 Cascade induced DNA unwinding

To directly probe the DNA unwinding dynamics caused by Cascade binding, we measured DNA backbone rotation in real time, using free orbiting magnetic tweezers (Figure 10.3A&B). Three independent experiments at three different forces, revealed a total DNA unwinding (i.e. negative DNA backbone rotation) angle of 1098.8 $\pm$39.6^{o} (mean $\pm$ s.d.) that took 38.7 $\pm$21.7 seconds to complete

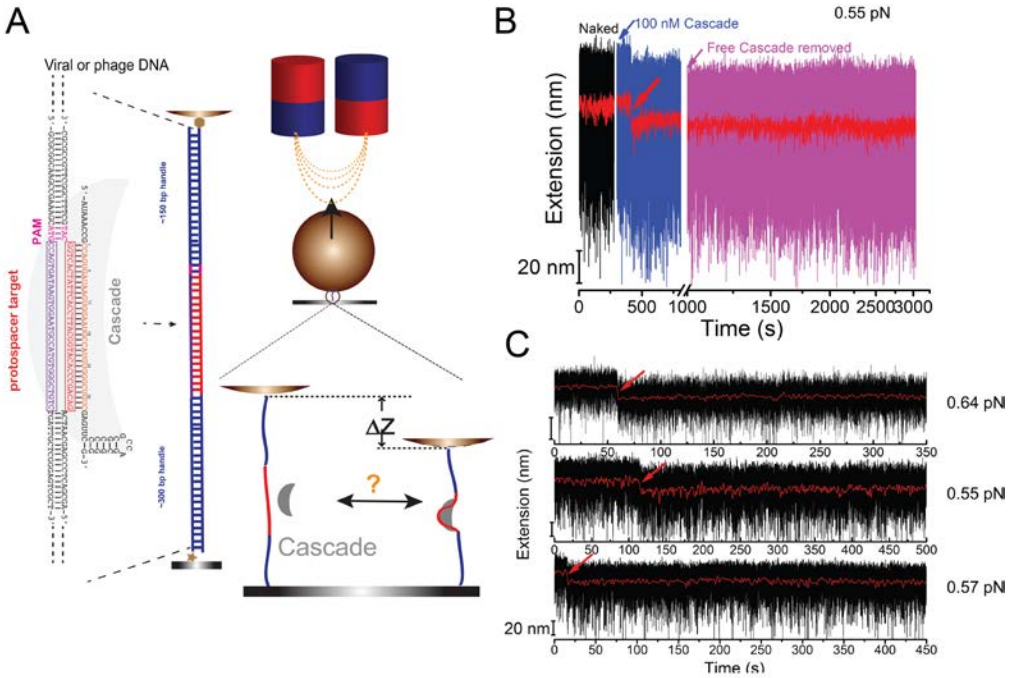

Figure 10.2: **Cascade induced DNA bending**. (A). A schematic of Cascade (gray), with the 32-nt crRNA-guide sequence (orange) base paired to a complementary target sequence (red). The displaced strand of the target is shown in purple. A DNA target, containing a protospacer (32-bp) and PAM (3-bp), is tethered between a 1-μm-diameter paramagnetic bead and coverslip surfaces. Extension changes due to Cascade binding, Δz, are traced. (B). A representative DNA extension time trace over the course of an experiment. (C). DNA extension time traces reveal a rapid stepwise extension drop of 15-20 nm (indicated by arrows) at $\sim$ 1 pN, seconds to minutes after introduction of Cascade (100 nM).

(Figure 10.3C). The slower rate of apparent DNA unwinding as compared to DNA bending is likely caused by a slower bead rotation relaxation (see Discussion). The total unwinding angle corresponds to unwinding of 34.3 ± 1.2^o per bp, which was estimated by dividing the unwinding angle by 32-bp of the protospacer DNA. This value is close to the twist angle per bp in B-form DNA ($\sim 34^o$ per bp) [83].

The unwinding angle corresponds to complete unlinking (i.e. linking number = 0) of the two strands over the length of the 32-bp protospacer, which is consistent with the -2.81 turns of shift in the twist-extension curve observed in the recently reported supercoiling assay [119]. In other words, the two ssDNA strands are positioned in such a way that they are topologically invariant to two parallel strands . These results reveal that Cascade is rigidly linked to both strands of the DNA target, such that the Cascade induced unwinding of the protospacer

DNA could generate sufficient torsion stress in the DNA to drive the observed bead rotation.

Target binding by Cascade bends the DNA and produces an R-loop [119, 189]. However, the location and conformational flexibility of the displaced DNA strand has not been established. To determine the conformational flexibility of the displaced strand we used the free orbiting magnetic tweezers to measure the rotational fluctuation of the DNA tether before and after unwinding. The overall twist stiffness of the DNA tether calculated from the rotational fluctuations did not show significant change upon Cascade binding (Figure 10.4). This result confirms that a majority of the displaced strand is tightly bound to the Cascade complex, thereby restricting the rotational freedom of the two DNA strands in the Cascade complex.

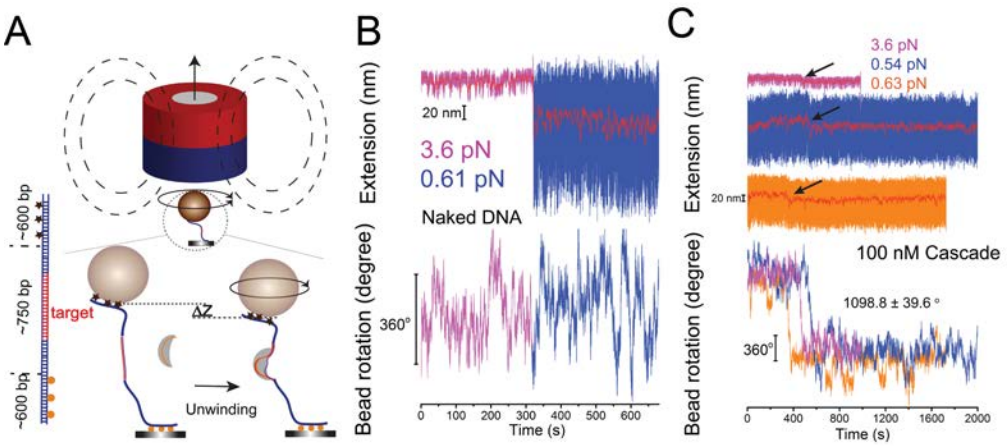

Figure 10.3: **Cascade induced DNA unwinding**. (A). Schematics of experiments using free orbiting magnetic tweezers. A 750-bp DNA (containing the same protospacer and PAM sequence used in Fig.1) between two ∼600-bp DNA handles tethered between the coverslip surface and streptavidin coated paramagnetic bead. DNA unwinding is traced by the bead rotation. (B). Extension and backbone rotation dynamics of naked DNA at two different forces, both fluctuate around constant values. (C). Three representative DNA extension and DNA backbone rotation time traces after the introduction of 100 nM Cascade. In each time trace, a single step DNA extension drop and DNA unwinding (1098.8 ± 39.6°) are observed after the introduction of Cascade.

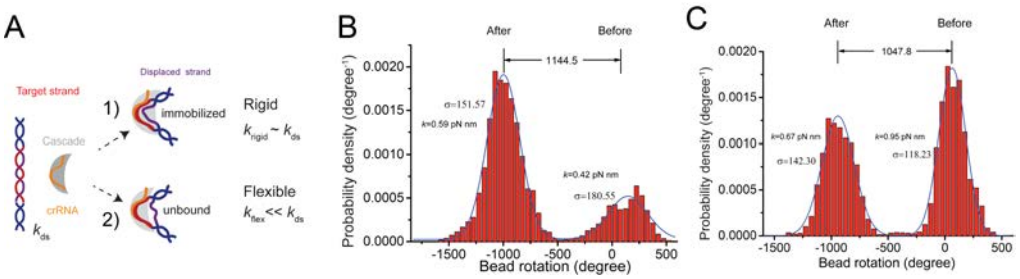

Figure 10.4: **Conformational state of the displaced strand of the DNA target**. (A). Schematic representation of: 1) an immobilized displaced strand (purple) in a rigid Cascade-protospacer DNA complex, and 2) an unbound displaced strand resulting in a torsional flexible joint between the two DNA handles. (B-C). Two typical distributions of DNA angular fluctuation before and after Cascade induced DNA unwinding. Blue lines are double-Gaussian fitting of the distribution histogram with $R^2 > 0.98$. The corresponding two-peak angular distance, standard deviation (σ) and the twist stiffness $k = \frac{(k_B T)}{\sigma^2}$ are indicated in figures.

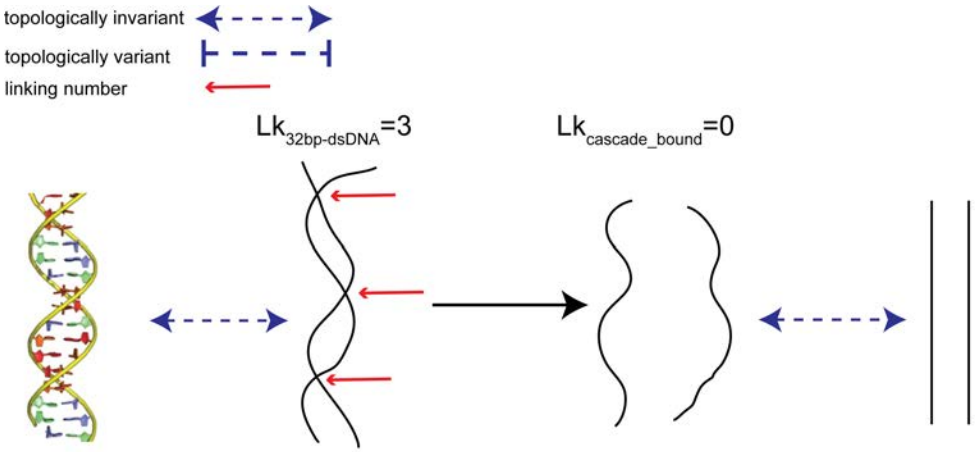

Figure 10.5: **Schematics of the topology of protospacer DNA before and after Cascade binding**. The 32-bp protospacer B-DNA has a linking number of 3 before Cascade binding (left panel). After R-loop formation, the DNA unwound by ~ 3 turns resulting in a linking number around zero, indicating that the two DNA strands are unlinked (right panel).

10.4 Discussion

Efficient target recognition by Cascade and recruitment of the trans-acting CRISPR-associated nuclease 3 (Cas3) are critical to crRNA-guided adaptive immunity in *E. coli*. Target recognition relies on R-loop formation, which involves formation of the hybrid of crRNA-ssDNA with a displaced strand. In this work we directly probed the dynamics of R-loop formation by monitoring DNA deformation during Cascade binding to a single target molecule. We show that R-loop

formation involves sharp DNA bending ($>100^o$) and unwinding of DNA backbone by $\sim 34^o$ per bp, where the displaced strand is stretched and rotationally restricted. This binding mechanism results in a rotation of the displaced strand relative to the complementary strand during crRNA-ssDNA hybridization.

After DNA bending and unwinding induced by Cascade binding, spontaneous bending or twist relaxation was never observed in our experimental time scale ($\sim$ 1000 s), even after removal of free Cascade from solution. This result indicates that the Cascade-DNA complex in the bent and unwound conformation is a highly stable structure. Such a highly stable complex is consistent with the 'locking' mechanism recently proposed by Szczelkun et al. [119] where a stable R-loop structure may provide a robust template for recruitment of Cas3.

The apparent bead rotation resulted from DNA unwinding took ~ 30 seconds to complete, which is much longer than the < 1 second time involved in the bending step. However, the observed differences in deformation time scales do not imply that Cascade induced DNA bending and unwinding are two independent processes. This is because the translational and rotational bead motions caused by the corresponding DNA bending and unwinding are associated with very different relaxation time scales. It only takes 10-3 seconds for the bead to diffuse to a new equilibrium position 20 nm away during DNA bending, while it takes ~ 10 seconds for the bead to relax to a new equilibrium angular position during DNA unwinding (Supplementary Information: Text S3 - relaxation time scales). Therefore, a likely scenario is that R-loop formation is synchronized with DNA bending, which takes a time of $\sim$0.4 seconds, followed by a much slower rotational relaxation of the bead to the new equilibrium angular position. However, our result does not exclude the alternative possibility that DNA bending occurs prior to R-loop formation, which may actually have a benefit to destabilize DNA duplex to facilitate strand invasion. Indeed, a pre-destabilized DNA target generated by applying negative torque promotes Cascade binding [119] .

We did not observe Cascade induced DNA deformations on control DNA lacking the target sequence. This result indicates that non-protospacer binding, should it occur, would not cause a stable association that could be detected using our single-DNA assay. Overall, the results are consistent with previous biochem-

ical and imaging experiments that reported transient unstable non-protospacer sequence binding [194, 195].

The crystal structure of Cascade bound to an ssDNA target reveals that the target strand forms a bent ribbon-like hybrid that is stretched and underwound. Our results suggest the displaced strand in the R-loop structure is also stabilized by Cascade in a stretched and conformationally restricted in a way that results in unlinking of the two DNA strands from the original B-DNA duplex. This unique organization of the crRNA-ssDNA hybrid and the displaced strand in the R-loop structure may form the physical basis for recruitment of downstream factors such as Cas3.

10.5 Materials and Methods

DNA synthesis –Two dsDNA targets (509-bp and 2033-bp) containing a 32-bp protospacer and a 3-bp PAM (underlined) (5′-GACAGCCCACATGGC ATTC-CACTT ATCACTGGCAT-3′) were generated by PCR amplification. The 509-bp DNA was generated with a thiol labeled forward primer and a biotin labeled reverse primer using lambda phage DNA as template. The 2033-bp PCR product was produced using Mega-PCR (184) with a multiple dig labeled forward primer and a multiple biotin labeled reverse primer. A control 576-bp DNA containing 3-bp PAM sequences (5′-CAT-3′), but no protospacer sequence was generated as previously described [109, 112].

Protein expression and purification–The protein subunits of the *E. coli* Cascade complex were co-expressed with a synthetic CRISPR containing four identical J3-spacers targeting the J-gene in λ-phage [196]. Cascade was purified using previously described methods [189, 197, 198].

Magnetic tweezers and measurements–A vertical magnetic tweezers setup with a sampling rate of 100 Hz and spatial resolution of $\sim$2 nm [120] was used to stretch the 509 bp DNA tethered between coverslips and a 1-μm-diameter paramagnetic bead, with a maximum force of $\sim$10 pN. A free-orbiting magnetic tweezers setup was used to stretch the multiple dig and biotin labeled 2,033-bp DNA tether. In these experiments, one end of each DNA strand is fixed to a coverslip while

the other end is attached to a paramagnetic bead. The free orbiting magnetic tweezers apply a magnetic field in the same direction of force [199], which allows unconstrained bead rotation around an axis along the magnetic field, allowing measurement of DNA winding or unwinding by monitoring the rotation of the bead. The trade-off of this system is that it can only apply forces up to 4 pN, and causes larger extension noise due to increased rotational fluctuation of the bead compared to traditional magnetic tweezers. The single-molecule stretching experiments were performed at 23°C in solution containing 50 mM KCl, 2.5 mM $MgCl_2$, and 20 mM Tris pH 7.4, in the absence or presence of Cascade. During solution exchange (∼10 sec), the DNA tethers were held at higher forces (∼4 pN for free-orbiting or ∼10 pN for traditional magnetic tweezers) to avoid bead-surface interaction due to the drag force of the solution flow. The force was immediately reduced to ∼1-pN after solution exchange. Detailed methods for the single-molecule stretching experiments can be found in Chapter 2.

relaxation time scales–The relaxation time (t_{rot}) for the bead to relax to a new equilibrium rotational position, is dependent on the twist stiffness of DNA, k, and the rotational friction coefficient of the bead, υ, through relation $t_{\mathrm{rot}} = \Upsilon/k$. υ depends on the radius of the bead, r_{bead}, and the bead rotational radius, r_{circle} through relation $\upsilon \approx \frac{8\pi\eta r_{\mathrm{bead}}^3}{1-(1/8)D^3} + \frac{6\pi\eta r_{\mathrm{bead}}^2 r_{\mathrm{circle}}}{1-(9/16)D+(1/8)D^3}$, where the η is the viscosity of the fluid, which is $\sim 10^{-3}\ kg\ m^{-1}\ s^{-1}$ for aqueous buffer. $D = r_{\mathrm{bead}}/S$, where S is the distance of bead to surface. In the experiments, $r_{\mathrm{bead}} \approx 500$ nm, $r_{\mathrm{circle}} \approx 600$ nm, $1/2 < D < 1$, t_{rot} is therefore estimated to be $\sim 11 - 13$ s. The estimated t_{rot} is in the same scale of the measured bead rotation time due to Cascade induced DNA unwinding; therefore, the actual Cascade induced DNA unwinding process can be much faster than the observed bead rotation time.

In contrast, for DNA bending signal, the translational relaxation time (t_{trans}) for the bead to relax to a new equilibrium position,$t_{\mathrm{trans}} \approx \frac{\Delta z^2}{2D}$, where the diffusion constant $D = \frac{k_{\mathrm{B}}T}{6\pi\eta r_{\mathrm{bead}}}$, and $\Delta z$is the measured extension change. For $\Delta z = 20$ nm, $t_{\mathrm{trans}} \approx 10^{-3}$ s, which is over two orders of magnitude smaller than the observed DNA bending time (∼0.4 s). Therefore, the observed DNA bending time reflects the real time involved in the Cascade induced DNA bending.

Cascade induced DNA topology change –The right-handed DNA double helix

has a helical pitch of $h = 3.6$ nm, containing 10.5 bp with two DNA strands winding over each other once per helical turn. The number of the times the two strands wind each other is defined as the linking number and a torsion-unconstrained DNA with N base pairs and a contour length L would have a relaxed linking number, $Lk_{\text{32bp-dsDNA}}$, calculated by: total twist angle/360°, or $N/10.5$, or L/h. For the 32 bp of the protospacer DNA, the relaxed linking number is: $Lk_{\text{32bp-dsDNA}} = 32/10.5 \sim 3$, as shown in supplementary Figure 10.5.

After Cascade invasion, the DNA tether unwound by total ∼1098°, indicating a linking number change of $\Delta Lk = -1098/360 \sim -3$, which results in a new linking number of $Lk_{\text{cascade-bound}} \sim 0$ for the Cascade bound protospacer DNA segment. In other words, in the R-loop structure, the crRNA bound ssDNA strand and the displaced ssDNA strand are in a conformation that is topologically invariant to two parallel strands (i.e, by smooth continuous deformation the two strands can become parallel to each other without any entanglement between the two strands).

The relative locations of the two strands in the R-loop should be spatially immobilized inside the complex. This is because any significant relative motion would decrease the local torsion stiffness and result in increased level of rotational fluctuation of the DNA tether, which was not observed in our experiments. The crRNA is also immobilized in the complex, since it is hybridized with complementary ssDNA strand. Together, our results suggest that the whole Cascade-protospacer DNA complex is a rigid body with fixed relative positions of the crRNA-ssDNA hybrid and the displaced ssDNA strand, in which the two ssDNA strands has a topology that is invariant to two parallel ssDNA strands. This suggests a unique spatial organization of the displaced strand relative to the complementary strand, whose detailed structure has yet to be determined.

Chapter 11

Conclusion and Discussion

In this thesis work, I developed a new platform for single ssDNA manipulation using state-of-art magnetic-tweezers, combined with a disturbance-free rapid solution exchange method (Chapter 2). With the platform, I systematically investigated several key accessory proteins including SSB, RecX, RecO, RecR on regulation of the formation and stability of RecA nucleoprotein filament, and the regulatory function of force on these protein-mediated dynamics and stability of RecA nucleoprotein filament (Chapters-3-5).

I investigated the salt concentration, protein concentration and force dependent DNA binding properties of SSB and identified a force-dependent non-wrapping binding mode of SSB. I showed that SSB outcompetes RecA binding to ssDNA, inhibiting the nucleation and polymerization of RecA filament, and de-stablizes pre-formed RecA filament by occupying the vacated ssDNA site due to RecA dissociation due to ATP-hydrolysis at low forces (several pN). Remarkably, I demonstrated that higher forces (> 15 pN) facilitate repolymerization of partially depolymerized RecA filament in a $3'$-to-$5'$ reverse direction, which not only highlights a potentially important regulatory role of force, but also point out the existence of bi-directional polymerization of RecA [24, 128, 142], in contrast to previously widely accepted $5'$-to-$3'$ unidirectional RecA polymerization [10].

I showed RecX promotes ATP-hydrolysis-dependent, step-wise net-depolymerization of RecA filament at low forces. Moreover, I demonstrated that the inhibitory effect of RecX is antagonized by higher forces, which cannot be explained by previous models of RecX-mediated RecA filament activities, and again empha-

sizes mechanosensitive regulation of RecA filament. Further, I showed that RecO tightly folds ssDNA, which is likely related to its ssDNA strand annealing function; RecOR highly extends ssDNA, which is likely related to its role of promoting RecA filament formation; RecOR stabilizes pre-formed RecA filaments against net de-stablization effect caused by SSB, while at the same condition, RecOR cannot facilitate RecA nucleation on SSB coated ssDNA.

To understand the underlining mechanisms of the mechanosensitive regulations on RecA filament, I theoretically analyzed the effects of force on conformational free energies of naked ssDNA and RecA nucleoprotein filament, which shows that force can increase the binding affinity of RecA to ssDNA over 10 times, consistently suggesting that formation and stability of RecA filament can be fine tuned by mechanical force.

In addition, I showed that the dynamics and stability of RecA filament formed on dsDNA is a competitive balance of polymerization/depolymerization of RecA on the one of the ssDNA strand (the leading ssDNA strand) and re-annealing of the complementary strand. As a result, RecA filament is stable on dsDNA with a $5'$ ssDNA overhang which acts as a 'cap' of the filament. Moreover, I demonstrated that RecA can not bind to S-DNA, a previously proposed binding substrate of RecA.

These studies establish a framework of molecular mechanisms of dynamic and regulation of RecA nucleoprotein filament mediated by accessory proteins and co-factors, and highlight the potentially broad regulatory role of force over physiological range in RecA-dependent homologous recombinational DNA repair, as well as other nucleic acid-protein interactions. Future efforts will be directed to studies of force regulatory role on other proteins involved in DNA replication, repair and recombination in both prokaryotic and eukaryotic systems.

Besides the studies on dynamics and regulations of RecA nucleoprotein filaments, I also studied site-specific single protein-DNA bindings. In Chapter 8, I systematically examined how force may affect the binding dynamics of *E. coli* IHF, a well-known DNA bending protein, to its specific binding sequence at a single protein-DNA complex level in real time. I showed that the stability of the specific IHF-DNA complex is fine-tuned by sub-pN forces. Furthermore, I

demonstrated for the first time that single short DNA manipulation using magnetic tweezers can resolve single site specific DNA binding dynamics and can be further used to quantify DNA bending angle by a single protein. Overall, this study has broad implications for site-specific DNA distorting proteins regarding their mechanosensitivity.

In Chapter 9, I show that hybridization of the Cascade-crRNA with complementary double-stranded DNA results in target bending of >100 degrees, target unwinding of 34 degree per bp, and stabilization of the displaced strand in an extended conformation. Interactions between Cascade and a DNA target are irreversible within our experimental time scale (>1000 seconds). Together these results provide new insights into crRNA-guided strand invasion and R-loop formation.

Bibliography

[1] Alberts, B, Wilson, J, & Hunt, T. (2008) *Molecular biology of the cell.* (Garland Science, New York), 5th edition.

[2] Nasheuer, H. P. (2010) *Genome stability and human diseases*, Subcellular biochemistry. (Springer, Dordrecht) Vol. 50.

[3] Friedberg, E. C, Elledge, S. J, Lehmann, A, Lindahl, T, & Muzi-Falconi, M. (2013) *DNA repair, mutagenesis, and other responses to DNA damage: a subject collection from Cold Spring Harbor perspectives in biology*, Cold Spring Harbor perspectives in biology.

[4] Generoso, W. M, Shelby, M. D, & De Serres, F. J. (1980) *DNA repair and mutagenesis in eukaryotes.* (Plenum Press, New York) Vol. 15.

[5] Kreuzer, K. N. (2013) Dna damage responses in prokaryotes: regulating gene expression, modulating growth patterns, and manipulating replication forks. *Cold Spring Harb Perspect Biol* **5**, a012674.

[6] Higgins, N. (2005) *The Bacterial Chromosome*, American Society Mic Series. (ASM Press).

[7] Khanna, K. K & Jackson, S. P. (2001) Dna double-strand breaks: signaling, repair and the cancer connection. *Nat Genet* **27**, 247–54.

[8] Clauson, C, Schärer, O. D, & Niedernhofer, L. (2013) Advances in understanding the complex mechanisms of dna interstrand cross-link repair. *Cold Spring Harb Perspect Med* **3**, a012732.

[9] Travers, A & Muskhelishvili, G. (2005) Bacterial chromatin. *Current opinion in genetics & development* **15**, 507–514.

[10] Cox, M. M. (2007) Regulation of bacterial reca protein function. *Crit Rev Biochem Mol Biol* **42**, 41–63.

[11] Kowalczykowski, S. C, Dixon, D. A, Eggleston, A. K, Lauder, S. D, & Rehrauer, W. M. (1994) Biochemistry of homologous recombination in escherichia coli. *Microbiological reviews* **58**, 401–465.

[12] Story, R. M & Steitz, T. A. (1992) Structure of the reca protein-adp complex. *Nature* **355**, 374–6.

[13] Story, R. M, Weber, I. T, & Steitz, T. A. (1992) The structure of the e. coli reca protein monomer and polymer. *Nature* **355**, 318–25.

[14] Pugh, B. F & Cox, M. M. (1988) General mechanism for reca protein binding to duplex dna. *J Mol Biol* **203**, 479–93.

[15] Bianco, P. R & Weinstock, G. M. (1996) Interaction of the reca protein of escherichia coli with single-stranded oligodeoxyribonucleotides. *Nucleic acids research* **24**, 4933–4939.

[16] Flory, J, Tsang, S. S, & Muniyappa, K. (1984) Isolation and visualization of active presynaptic filaments of reca protein and single-stranded dna. *Proc Natl Acad Sci U S A* **81**, 7026–30.

[17] Cox, M. M. (2007) Motoring along with the bacterial reca protein. *Nat Rev Mol Cell Biol* **8**, 127–38.

[18] Leger, J. F, Robert, J, Bourdieu, L, Chatenay, D, & Marko, J. F. (1998) Reca binding to a single double-stranded dna molecule: a possible role of dna conformational fluctuations. *Proc Natl Acad Sci U S A* **95**, 12295–9.

[19] Cox, J. M, Tsodikov, O. V, & Cox, M. M. (2005) Organized unidirectional waves of atp hydrolysis within a reca filament. *PLoS Biol* **3**, e52.

[20] Brenner, S. L, Mitchell, R. S, Morrical, S. W, Neuendorf, S. K, Schutte, B. C, & Cox, M. M. (1987) reca protein-promoted atp hydrolysis occurs throughout reca nucleoprotein filaments. *The Journal of biological chemistry* **262**, 4011–4016.

[21] Hegner, M, Smith, S. B, & Bustamante, C. (1999) Polymerization and mechanical properties of single reca-dna filaments. *Proc Natl Acad Sci U S A* **96**, 10109–14.

[22] Shivashankar, G. V, Feingold, M, Krichevsky, O, & Libchaber, A. (1999) Reca polymerization on double-stranded dna by using single-molecule manipulation: the role of atp hydrolysis. *Proc Natl Acad Sci U S A* **96**, 7916–21.

[23] van Loenhout, M. T. J, van der Heijden, T, Kanaar, R, Wyman, C, & Dekker, C. (2009) Dynamics of reca filaments on single-stranded dna. *Nucleic Acids Res* **37**, 4089–99.

[24] Joo, C, McKinney, S. A, Nakamura, M, Rasnik, I, Myong, S, & Ha, T. (2006) Real-time observation of reca filament dynamics with single monomer resolution. *Cell* **126**, 515–27.

[25] Shan, Q & Cox, M. M. (1996) Reca protein dynamics in the interior of reca nucleoprotein filaments. *J Mol Biol* **257**, 756–74.

[26] Shan, Q & Cox, M. M. (1997) Reca filament dynamics during dna strand exchange reactions. *J Biol Chem* **272**, 11063–73.

[27] Ramreddy, T, Sen, S, Rao, B. J, & Krishnamoorthy, G. (2003) Dna dynamics in reca-dna filaments: Atp hydrolysis-related flexibility in dna. *Biochemistry* **42**, 12085–94.

[28] Fulconis, R, Bancaud, A, Allemand, J.-F, Croquette, V, Dutreix, M, & Viovy, J.-L. (2004) Twisting and untwisting a single dna molecule covered by reca protein. *Biophys J* **87**, 2552–63.

[29] van der Heijden, T, van Noort, J, van Leest, H, Kanaar, R, Wyman, C, Dekker, N. H, Dekker, N, & Dekker, C. (2005) Torque-limited reca polymerization on dsdna. *Nucleic Acids Res* **33**, 2099–105.

[30] Galletto, R, Amitani, I, Baskin, R. J, & Kowalczykowski, S. C. (2006) Direct observation of individual reca filaments assembling on single dna molecules. *Nature* **443**, 875–8.

[31] van der Heijden, T, Modesti, M, Hage, S, Kanaar, R, Wyman, C, & Dekker, C. (2008) Homologous recombination in real time: Dna strand exchange by reca. *Mol Cell* **30**, 530–8.

[32] Cox, M. M. (1994) Why does reca protein hydrolyse atp? *Trends Biochem Sci* **19**, 217–22.

[33] Shereda, R. D, Kozlov, A. G, Lohman, T. M, Cox, M. M, & Keck, J. L. (2008) Ssb as an organizer/mobilizer of genome maintenance complexes. *Crit Rev Biochem Mol Biol* **43**, 289–318.

[34] Alberts, B. M & Frey, L. (1970) T4 bacteriophage gene 32: a structural protein in the replication and recombination of dna. *Nature* **227**, 1313–8.

[35] Sigal, N, Delius, H, Kornberg, T, Gefter, M. L, & Alberts, B. (1972) A dna-unwinding protein isolated from escherichia coli: its interaction with dna and with dna polymerases. *Proc Natl Acad Sci U S A* **69**, 3537–41.

[36] Erdile, L. F, Heyer, W. D, Kolodner, R, & Kelly, T. J. (1991) Characterization of a cdna encoding the 70-kda single-stranded dna-binding subunit of human replication protein a and the role of the protein in dna replication. *The Journal of biological chemistry* **266**, 12090–12098.

[37] Raghunathan, S, Kozlov, A. G, Lohman, T. M, & Waksman, G. (2000) Structure of the dna binding domain of e. coli ssb bound to ssdna. *Nat Struct Biol* **7**, 648–52.

[38] Urbanke, C & Schaper, A. (1990) Kinetics of binding of single-stranded dna binding protein from escherichia coli to single-stranded nuclei acids. *Biochemistry* **29**, 1744–9.

[39] Chrysogelos, S & Griffith, J. (1982) Escherichia coli single-strand binding protein organizes single-stranded dna in nucleosome-like units. *Proc Natl Acad Sci U S A* **79**, 5803–7.

[40] Bujalowski, W, Overman, L. B, & Lohman, T. M. (1988) Binding mode transitions of escherichia coli single strand binding protein-single-stranded dna complexes. cation, anion, ph, and binding density effects. *The Journal of biological chemistry* **263**, 4629–4640.

[41] Lohman, T. M & Overman, L. B. (1985) Two binding modes in escherichia coli single strand binding protein-single stranded dna complexes. modulation by nacl concentration. *J Biol Chem* **260**, 3594–603.

[42] Blanar, M. A, Sandler, S. J, Armengod, M. E, Ream, L. W, & Clark, A. J. (1984) Molecular analysis of the recf gene of escherichia coli. *Proc Natl Acad Sci U S A* **81**, 4622–6.

[43] Madiraju, M. V & Clark, A. J. (1991) Effect of recf protein on reactions catalyzed by reca protein. *Nucleic Acids Res* **19**, 6295–300.

[44] Madiraju, M. V & Clark, A. J. (1992) Evidence for atp binding and double-stranded dna binding by escherichia coli recf protein. *J Bacteriol* **174**, 7705–10.

[45] Webb, B. L, Cox, M. M, & Inman, R. B. (1999) Atp hydrolysis and dna binding by the escherichia coli recf protein. *J Biol Chem* **274**, 15367–74.

[46] Morimatsu, K & Kowalczykowski, S. C. (2003) Recfor proteins load reca protein onto gapped dna to accelerate dna strand exchange: a universal step of recombinational repair. *Mol Cell* **11**, 1337–47.

[47] Sakai, A & Cox, M. M. (2009) Recfor and recor as distinct reca loading pathways. *J Biol Chem* **284**, 3264–72.

[48] Kolodner, R, Fishel, R. A, & Howard, M. (1985) Genetic recombination of bacterial plasmid dna: effect of recf pathway mutations on plasmid recombination in escherichia coli. *J Bacteriol* **163**, 1060–6.

[49] Luisi-DeLuca, C & Kolodner, R. (1994) Purification and characterization of the escherichia coli reco protein. renaturation of complementary single-stranded dna molecules catalyzed by the reco protein. *J Mol Biol* **236**, 124–38.

[50] Umezu, K, Chi, N. W, & Kolodner, R. D. (1993) Biochemical interaction of the escherichia coli recf, reco, and recr proteins with reca protein and single-stranded dna binding protein. *Proc Natl Acad Sci U S A* **90**, 3875–9.

[51] Leiros, I, Timmins, J, Hall, D. R, & McSweeney, S. (2005) Crystal structure and dna-binding analysis of reco from deinococcus radiodurans. *EMBO J* **24**, 906–18.

[52] Ryzhikov, M, Koroleva, O, Postnov, D, Tran, A, & Korolev, S. (2011) Mechanism of reco recruitment to dna by single-stranded dna binding protein. *Nucleic Acids Res* **39**, 6305–14.

[53] Gupta, R, Ryzhikov, M, Koroleva, O, Unciuleac, M, Shuman, S, Korolev, S, & Glickman, M. S. (2013) A dual role for mycobacterial reco in reca-dependent homologous recombination and reca-independent single-strand annealing. *Nucleic Acids Res* **41**, 2284–95.

[54] Alonso, J. C, Stiege, A. C, Dobrinski, B, & Lurz, R. (1993) Purification and properties of the recr protein from bacillus subtilis 168. *J Biol Chem* **268**, 1424–9.

[55] Ryzhikov, M & Korolev, S. (2012) Structural studies of ssb interaction with reco. *Methods Mol Biol* **922**, 123–31.

[56] Pagès, V, Koffel-Schwartz, N, & Fuchs, R. P. P. (2003) recx, a new sos gene that is co-transcribed with the reca gene in escherichia coli. *DNA Repair (Amst)* **2**, 273–84.

[57] Venkatesh, R, Ganesh, N, Guhan, N, Reddy, M. S, Chandrasekhar, T, & Muniyappa, K. (2002) Recx protein abrogates atp hydrolysis and strand exchange promoted by reca: insights into negative regulation of homologous recombination. *Proc Natl Acad Sci U S A* **99**, 12091–6.

[58] Stohl, E. A, Brockman, J. P, Burkle, K. L, Morimatsu, K, Kowalczykowski, S. C, & Seifert, H. S. (2003) Escherichia coli recx inhibits reca recombinase and coprotease activities in vitro and in vivo. *J Biol Chem* **278**, 2278–85.

[59] Yang, C.-Y, Chin, K.-H, Yang, M.-T, Wang, A. H.-J, & Chou, S.-H. (2009) Crystal structure of recx: a potent regulatory protein of reca from xanthomonas campestris. *Proteins* **74**, 530–7.

[60] Yang, M. K, Chou, M. E, & Yang, Y. C. (2001) Molecular characterization and expression of the recx gene of xanthomonas campestris pv. citri. *Curr Microbiol* **42**, 257–63.

[61] Vierling, S, Weber, T, Wohlleben, W, & Muth, G. (2000) Transcriptional and mutational analyses of the streptomyces lividans recx gene and its interference with reca activity. *J Bacteriol* **182**, 4005–11.

[62] Drees, J. C, Lusetti, S. L, Chitteni-Pattu, S, Inman, R. B, & Cox, M. M. (2004) A reca filament capping mechanism for recx protein. *Mol Cell* **15**, 789–98.

[63] Drees, J. C, Lusetti, S. L, & Cox, M. M. (2004) Inhibition of reca protein by the escherichia coli recx protein: modulation by the reca c terminus and filament functional state. *J Biol Chem* **279**, 52991–7.

[64] Ragone, S, Maman, J. D, Furnham, N, & Pellegrini, L. (2008) Structural basis for inhibition of homologous recombination by the recx protein. *EMBO J* **27**, 2259–69.

[65] Candelli, A, Modesti, M, Peterman, E. J. G, & Wuite, G. J. L. (2013) Single-molecule views on homologous recombination. *Quarterly Reviews of Biophysics* **46**, 323–348.

[66] Ohnishi, T, Mori, E, & Takahashi, A. (2009) Dna double-strand breaks: their production, recognition, and repair in eukaryotes. *Mutat Res* **669**, 8–12.

[67] Bardwell, A. J, Bardwell, L, Wang, Z, Siede, W, Reagan, M. S, Tomkinson, A. E, Friedberg, A. S, Pittenger, C, Feaver, W. J, & Svejstrup, J. (1994) Recent insights on dna repair. the mechanism of damaged nucleotide excision in eukaryotes and its relationship to other cellular processes. *Ann N Y Acad Sci* **726**, 281–91.

[68] Vogel, V & Sheetz, M. (2006) Local force and geometry sensing regulate cell functions. *Nat Rev Mol Cell Biol* **7**, 265–75.

[69] Shivashankar, G. V. (2011) Mechanosignaling to the cell nucleus and gene regulation. *Annu Rev Biophys* **40**, 361–78.

[70] Iskratsch, T, Wolfenson, H, & Sheetz, M. P. (2014) Appreciating force and shape—the rise of mechanotransduction in cell biology. *Nat Rev Mol Cell Biol* **15**, 825–33.

[71] Finer, J. T, Simmons, R. M, & Spudich, J. A. (1994) Single myosin molecule mechanics: piconewton forces and nanometre steps. *Nature* **368**, 113–119.

[72] Wang, M. D, Schnitzer, M. J, Yin, H, Landick, R, Gelles, J, & Block, S. M. (1998) Force and velocity measured for single molecules of rna polymerase. *Science* **282**, 902–7.

[73] Wuite, G. J, Smith, S. B, Young, M, Keller, D, & Bustamante, C. (2000) Single-molecule studies of the effect of template tension on t7 dna polymerase activity. *Nature* **404**, 103–6.

[74] Nicklas, R. B. (1988) The forces that move chromosomes in mitosis. *Annu Rev Biophys Biophys Chem* **17**, 431–49.

[75] Lesterlin, C, Ball, G, Schermelleh, L, & Sherratt, D. J. (2014) Reca bundles mediate homology pairing between distant sisters during dna break repair. *Nature* **506**, 249–53.

[76] Cho, N. W, Dilley, R. L, Lampson, M. A, & Greenberg, R. A. (year?) Interchromosomal homology searches drive directional alt telomere movement and synapsis. *Cell* **159**, 108–121.

[77] Smith, S. B, Finzi, L, & Bustamante, C. (1992) Direct mechanical measurements of the elasticity of single dna molecules by using magnetic beads. *Science* **258**, 1122–6.

[78] Bustamante, C, Marko, J. F, Siggia, E. D, & Smith, S. (1994) Entropic elasticity of lambda-phage dna. *Science* **265**, 1599–600.

[79] Marko, J. F & Siggia, E. D. (1995) Stretching dna. *Macromolecules* **28**, 8759–8770.

[80] Marko & Siggia. (1995) Statistical mechanics of supercoiled dna. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* **52**, 2912–2938.

[81] Smith, S. B, Cui, Y, & Bustamante, C. (1996) Overstretching b-dna: the elastic response of individual double-stranded and single-stranded dna molecules. *Science* **271**, 795–9.

[82] Cocco, S, Yan, J, Léger, J.-F, Chatenay, D, & Marko, J. F. (2004) Overstretching and force-driven strand separation of double-helix dna. *Phys Rev E Stat Nonlin Soft Matter Phys* **70**, 011910.

[83] WATSON, J. D & CRICK, F. H. C. (1953) Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature* **171**, 737–738.

[84] Crick, F. H. C & Watson, J. D. (1954) The complementary structure of deoxyribonucleic acid. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **223**, 80–96.

[85] Cluzel, P, Lebrun, A, Heller, C, Lavery, R, Viovy, J. L, Chatenay, D, & Caron, F. (1996) Dna: an extensible molecule. *Science* **271**, 792–4.

[86] Williams, M. C, Wenner, J. R, Rouzina, I, & Bloomfield, V. A. (2001) Effect of ph on the overstretching transition of double-stranded dna: evidence of force-induced dna melting. *Biophys J* **80**, 874–81.

[87] Williams, M. C, Wenner, J. R, Rouzina, I, & Bloomfield, V. A. (2001) Entropy and heat capacity of dna melting from temperature dependence of single molecule stretching. *Biophys J* **80**, 1932–9.

[88] Wenner, J. R, Williams, M. C, Rouzina, I, & Bloomfield, V. A. (2002) Salt dependence of the elasticity and overstretching transition of single dna molecules. *Biophys J* **82**, 3160–9.

[89] Bosaeus, N, El-Sagheer, A. H, Brown, T, Smith, S. B, Akerman, B, Bustamante, C, & Nordén, B. (2012) Tension induces a base-paired overstretched dna conformation. *Proc Natl Acad Sci U S A* **109**, 15179–84.

[90] Sheinin, M. Y, Forth, S, Marko, J. F, & Wang, M. D. (2011) Underwound dna under tension: structure, elasticity, and sequence-dependent behaviors. *Phys Rev Lett* **107**, 108102.

[91] van Mameren, J, Gross, P, Farge, G, Hooijman, P, Modesti, M, Falkenberg, M, Wuite, G. J. L, & Peterman, E. J. G. (2009) Unraveling the structure of dna during overstretching by using multicolor, single-molecule fluorescence imaging. *Proc Natl Acad Sci U S A* **106**, 18231–6.

[92] Fu, H, Chen, H, Marko, J. F, & Yan, J. (2010) Two distinct overstretched dna states. *Nucleic Acids Res* **38**, 5594–600.

[93] Fu, H, Chen, H, Zhang, X, Qu, Y, Marko, J. F, & Yan, J. (2011) Transition dynamics and selection of the distinct s-dna and strand unpeeling modes of double helix overstretching. *Nucleic Acids Res* **39**, 3473–81.

[94] Zhang, X, Chen, H, Fu, H, Doyle, P. S, & Yan, J. (2012) Two distinct overstretched dna structures revealed by single-molecule thermodynamics measurements. *Proc Natl Acad Sci U S A* **109**, 8103–8.

[95] Zhang, X, Chen, H, Le, S, Rouzina, I, Doyle, P. S, & Yan, J. (2013) Revealing the competition between peeled ssdna, melting bubbles, and s-dna during dna overstretching by single-molecule calorimetry. *Proc Natl Acad Sci U S A* **110**, 3865–70.

[96] King, G. A, Gross, P, Bockelmann, U, Modesti, M, Wuite, G. J. L, & Peterman, E. J. G. (2013) Revealing the competition between peeled ssdna, melting bubbles, and s-dna during dna overstretching using fluorescence microscopy. *Proceedings of the National Academy of Sciences* **110**, 3859–3864.

[97] Zhang, X, Qu, Y, Chen, H, Rouzina, I, Zhang, S, Doyle, P. S, & Yan, J. (2014) Interconversion between three overstretched dna structures. *J Am Chem Soc* **136**, 16073–80.

[98] Saleh, O. A, McIntosh, D. B, Pincus, P, & Ribeck, N. (2009) Nonlinear low-force elasticity of single-stranded dna molecules. *Phys. Rev. Lett.* **102**, 068301.

[99] McIntosh, D. B & Saleh, O. A. (2011) Salt species-dependent electrostatic effects on ssdna elasticity. *Macromolecules* **44**, 2328–2333.

[100] McIntosh, D & Saleh, O. A. (2010) Elastic behavior of ssdna in salty solutions. *Biophysical Journal* **98**, 593a –.

[101] Yan, J & Marko, J. F. (2003) Effects of dna-distorting proteins on dna elastic response. *Phys Rev E Stat Nonlin Soft Matter Phys* **68**, 011905.

[102] Chen, H & Yan, J. (2008) Effects of kink and flexible hinge defects on mechanical responses of short double-stranded dna molecules. *Phys Rev E Stat Nonlin Soft Matter Phys* **77**, 041907.

[103] Yan, J, Kawamura, R, & Marko, J. F. (2005) Statistics of loop formation along double helix dnas. *Phys Rev E Stat Nonlin Soft Matter Phys* **71**, 061905.

[104] Liu, Y, Chen, H, Kenney, L. J, & Yan, J. (2010) A divalent switch drives h-ns/dna-binding conformations between stiffening and bridging modes. *Genes Dev* **24**, 339–44.

[105] Lin, J, Chen, H, Dröge, P, & Yan, J. (2012) Physical organization of dna by multiple non-specific dna-binding modes of integration host factor (ihf). *PLoS One* **7**, e49885.

[106] Laurens, N, Driessen, R. P. C, Heller, I, Vorselen, D, Noom, M. C, Hol, F. J. H, White, M. F, Dame, R. T, & Wuite, G. J. L. (2012) Alba shapes the archaeal genome using a delicate balance of bridging and stiffening the dna. *Nat Commun* **3**, 1328.

[107] van Mameren, J, Modesti, M, Kanaar, R, Wyman, C, Peterman, E. J. G, & Wuite, G. J. L. (2009) Counting rad51 proteins disassembling from nucleoprotein filaments under tension. *Nature* **457**, 745–8.

[108] Laurens, N, Rusling, D. A, Pernstich, C, Brouwer, I, Halford, S. E, & Wuite, G. J. L. (2012) Dna looping by foki: the impact of twisting and bending rigidity on protein-induced looping dynamics. *Nucleic Acids Res* **40**, 4988–97.

[109] Fu, H, Le, S, Chen, H, Muniyappa, K, & Yan, J. (2013) Force and atp hydrolysis dependent regulation of reca nucleoprotein filament by single-stranded dna binding protein. *Nucleic Acids Res* **41**, 924–32.

[110] Fu, H, Le, S, Muniyappa, K, & Yan, J. (2013) Dynamics and regulation of reca polymerization and de-polymerization on double-stranded dna. *PLoS One* **8**, e66712.

[111] Le, S, Chen, H, Cong, P, Lin, J, Droge, P, & Yan, J. (2013) Mechanosensing of dna bending in a single specific protein-dna complex. *Sci. Rep.* **3**.

[112] Le, S, Chen, H, Zhang, X, Chen, J, Patil, K. N, Muniyappa, K, & Yan, J. (2014) Mechanical force antagonizes the inhibitory effects of recx on reca filament formation in mycobacterium tuberculosis. *Nucleic Acids Res* **42**, 11992–9.

[113] Chen, J, Le, S, Basu, A, Chazin, W. J, & Yan, J. (2015) Mechanochemical regulations of rpa's binding to ssdna. *Sci Rep* **5**, 9296.

[114] Vogel, V & Sheetz, M. P. (2009) Cell fate regulation by coupling mechanical cycles to biochemical signaling pathways. *Curr Opin Cell Biol* **21**, 38–46.

[115] Neuman, K. C & Nagy, A. (2008) Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat Methods* **5**, 491–505.

[116] Crick, F & Hughes, A. (1950) The physical properties of cytoplasm: A study by means of the magnetic particle method part i. experimental. *Experimental Cell Research* **1**, 37 – 80.

[117] Strick, T. R, Allemand, J. F, Bensimon, D, & Croquette, V. (1998) Behavior of supercoiled dna. *Biophys J* **74**, 2016–28.

[118] Cui, Y, Petrushenko, Z. M, & Rybenkov, V. V. (2008) Mukb acts as a macromolecular clamp in dna condensation. *Nat Struct Mol Biol* **15**, 411–8.

[119] Szczelkun, M. D, Tikhomirova, M. S, Sinkunas, T, Gasiunas, G, Karvelis, T, Pschera, P, Siksnys, V, & Seidel, R. (2014) Direct observation of r-loop formation by single rna-guided cas9 and cascade effector complexes. *Proc Natl Acad Sci U S A* **111**, 9798–803.

[120] Chen, H, Fu, H, Zhu, X, Cong, P, Nakamura, F, & Yan, J. (2011) Improved high-force magnetic tweezers for stretching and refolding of proteins and short dna. *Biophys J* **100**, 517–23.

[121] Chen, H, Zhu, X, Cong, P, Sheetz, M. P, Nakamura, F, & Yan, J. (2011) Differential mechanical stability of filamin a rod segments. *Biophys J* **101**, 1231–7.

[122] Yan, J, Skoko, D, & Marko, J. F. (2004) Near-field-magnetic-tweezer manipulation of single dna molecules. *Phys Rev E Stat Nonlin Soft Matter Phys* **70**, 011905.

[123] De Vlaminck, I & Dekker, C. (2012) Recent advances in magnetic tweezers. *Annu Rev Biophys* **41**, 453–72.

[124] De Vlaminck, I, Henighan, T, van Loenhout, M. T. J, Burnham, D. R, & Dekker, C. (2012) Magnetic forces and dna mechanics in multiplexed magnetic tweezers. *PLoS One* **7**, e41432.

[125] Ha, T, Enderle, T, Ogletree, D. F, Chemla, D. S, Selvin, P. R, & Weiss, S. (1996) Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc Natl Acad Sci U S A* **93**, 6264–8.

[126] Qu, Y, Lim, C. J, Whang, Y. R, Liu, J, & Yan, J. (2013) Mechanism of dna organization by mycobacterium tuberculosis protein lsr2. *Nucleic Acids Res* **41**, 5263–72.

[127] Gosse, C & Croquette, V. (2002) Magnetic tweezers: micromanipulation and force measurement at the molecular level. *Biophys J* **82**, 3314–29.

[128] Bell, J. C, Plank, J. L, Dombrowski, C. C, & Kowalczykowski, S. C. (2012) Direct imaging of reca nucleation and growth on single molecules of ssb-coated ssdna. *Nature* **491**, 274–8.

[129] Candelli, A, Wuite, G. J. L, & Peterman, E. J. G. (2011) Combining optical trapping, fluorescence microscopy and micro-fluidics for single molecule studies of dna-protein interactions. *Phys Chem Chem Phys* **13**, 7263–72.

[130] Omar, M. A, Miskovsky, P, & Bánó, G. (2014) Proof-of-principle for simple microshelter-assisted buffer exchange in laser tweezers: interaction of hypericin with single cells. *Lab Chip* **14**, 1579–84.

[131] Masters, T, Engl, W, Weng, Z. L, Arasi, B, Gauthier, N, & Viasnoff, V. (2012) Easy fabrication of thin membranes with through holes. application to protein patterning. *PLoS One* **7**, e44261.

[132] Cox, J. M, Abbott, S. N, Chitteni-Pattu, S, Inman, R. B, & Cox, M. M. (2006) Complementation of one reca protein point mutation by another. evidence for trans catalysis of atp hydrolysis. *J Biol Chem* **281**, 12968–75.

[133] Ragunathan, K, Joo, C, & Ha, T. (2011) Real-time observation of strand exchange reaction with high spatiotemporal resolution. *Structure* **19**, 1064–1073.

[134] Forget, A. L & Kowalczykowski, S. C. (2012) Single-molecule imaging of dna pairing by reca reveals a three-dimensional homology search. *Nature* **482**, 423–7.

[135] Kowalczykowski, S. C, Clow, J, Somani, R, & Varghese, A. (1987) Effects of the escherichia coli ssb protein on the binding of escherichia coli reca protein to single-stranded dna. demonstration of competitive binding and the lack of a specific protein-protein interaction. *J Mol Biol* **193**, 81–95.

[136] Kowalczykowski, S. C & Krupp, R. A. (1987) Effects of escherichia coli ssb protein on the single-stranded dna-dependent atpase activity of escherichia coli reca protein. evidence that ssb protein facilitates the binding of reca protein to regions of secondary structure within single-stranded dna. *J Mol Biol* **193**, 97–113.

[137] Bujalowski, W & Lohman, T. M. (1986) Escherichia coli single-strand binding protein forms multiple, distinct complexes with single-stranded dna. *Biochemistry* **25**, 7799–7802.

[138] Roy, R, Kozlov, A. G, Lohman, T. M, & Ha, T. (2007) Dynamic structural rearrangements between dna binding modes of e. coli ssb protein. *J Mol Biol* **369**, 1244–57.

[139] Zhou, R, Kozlov, A. G, Roy, R, Zhang, J, Korolev, S, Lohman, T. M, & Ha, T. (2011) Ssb functions as a sliding platform that migrates on dna via reptation. *Cell* **146**, 222–232.

[140] Meyer, R. R & Laine, P. S. (1990) The single-stranded dna-binding protein of escherichia coli. *Microbiological Reviews* **54**, 342–380.

[141] McEntee, K, Weinstock, G. M, & Lehman, I. R. (1981) Binding of the reca protein of escherichia coli to single- and double-stranded dna. *Journal of Biological Chemistry* **256**, 8835–8844.

[142] Mani, A, Braslavsky, I, Arbel-Goren, R, & Stavans, J. (2010) Caught in the act: the lifetime of synaptic intermediates during the search for homology on dna. *Nucleic Acids Res* **38**, 2036–43.

[143] VanLoock, M. S, Yu, X, Yang, S, Galkin, V. E, Huang, H, Rajan, S. S, Anderson, W. F, Stohl, E. A, Seifert, H. S, & Egelman, E. H. (2003) Complexes of reca with lexa and recx differentiate between active and inactive reca nucleoprotein filaments. *J Mol Biol* **333**, 345–54.

[144] Gruenig, M. C, Stohl, E. A, Chitteni-Pattu, S, Seifert, H. S, & Cox, M. M. (2010) Less is more: Neisseria gonorrhoeae recx protein stimulates recombination by inhibiting reca. *J Biol Chem* **285**, 37188–97.

[145] Kumar, R. A, Vaze, M. B, Chandra, N. R, Vijayan, M, & Muniyappa, K. (1996) Functional characterization of the precursor and spliced forms of reca protein of mycobacterium tuberculosis. *Biochemistry* **35**, 1793–802.

[146] Kantake, N, Madiraju, M. V. V. M, Sugiyama, T, & Kowalczykowski, S. C. (2002) Escherichia coli reco protein anneals ssdna complexed with its cognate ssdna-binding protein: A common step in genetic recombination. *Proc Natl Acad Sci U S A* **99**, 15327–32.

[147] Pugh, B. F & Cox, M. M. (1988) General mechanism for reca protein binding to duplex dna. *Journal of molecular biology* **203**, 479–493.

[148] Chen, Z, Yang, H, & Pavletich, N. P. (2008) Mechanism of homologous recombination from the reca-ssdna/dsdna structures. *Nature* **453**, 489–4.

[149] Feinstein, E, Danilowicz, C, Conover, A, Gunaratne, R, Kleckner, N, & Prentiss, M. (2011) Single-molecule studies of the stringency factors and rates governing the polymerization of reca on double-stranded dna. *Nucleic Acids Res* **39**, 3781–91.

[150] Sagi, D, Tlusty, T, & Stavans, J. (2006) High fidelity of reca-catalyzed recombination: a watchdog of genetic diversity. *Nucleic Acids Res* **34**, 5021–31.

[151] Kadonaga, J. T. (2004) Regulation of rna polymerase ii transcription by sequence-specific dna binding factors. *Cell* **116**, 247–57.

[152] Hughes, T. R. (2011) *A handbook of transcription factors*, Subcellular biochemistry. (Springer, Dordrecht) Vol. 52.

[153] Ishihama, A. (2000) Functional modulation of escherichia coli rna polymerase. *Annu Rev Microbiol* **54**, 499–518.

[154] Nash, H. A & Robertson, C. A. (1981) Purification and properties of the escherichia coli protein factor required for lambda integrative recombination. *J Biol Chem* **256**, 9246–53.

[155] Yang, S. W & Nash, H. A. (1995) Comparison of protein binding to dna in vivo and in vitro: defining an effective intracellular target. *EMBO J* **14**, 6292–300.

[156] Arfin, S. M, Long, A. D, Ito, E. T, Tolleri, L, Riehle, M. M, Paegle, E. S, & Hatfield, G. W. (2000) Global gene expression profiling in escherichia coli k12. the effects of integration host factor. *J Biol Chem* **275**, 29672–84.

[157] Mangan, M. W, Lucchini, S, Danino, V, Cróinín, T. O, Hinton, J. C. D, & Dorman, C. J. (2006) The integration host factor (ihf) integrates stationary-phase and virulence gene expression in salmonella enterica serovar typhimurium. *Mol Microbiol* **59**, 1831–47.

[158] Rice, P. A, Yang, S, Mizuuchi, K, & Nash, H. A. (1996) Crystal structure of an ihf-dna complex: a protein-induced dna u-turn. *Cell* **87**, 1295–306.

[159] Sugimura, S & Crothers, D. M. (2006) Stepwise binding and bending of dna by escherichia coli integration host factor. *Proc Natl Acad Sci U S A* **103**, 18510–4.

[160] Craig, N. L & Nash, H. A. (1984) E. coli integration host factor binds to specific sites in {DNA}. *Cell* **39**, 707 – 716.

[161] Kuznetsov, S. V, Sugimura, S, Vivas, P, Crothers, D. M, & Ansari, A. (2006) Direct observation of dna bending/unbending kinetics in complex with dna-bending protein ihf. *Proc Natl Acad Sci U S A* **103**, 18515–20.

[162] Seong, G. H, Kobatake, E, Miura, K, Nakazawa, A, & Aizawa, M. (2002) Direct atomic force microscopy visualization of integration host factor-induced dna bending structure of the promoter regulatory region on the pseudomonas tol plasmid. *Biochem Biophys Res Commun* **291**, 361–6.

[163] Dame, R. T, van Mameren, J, Luijsterburg, M. S, Mysiak, M. E, Janićijević, A, Pazdzior, G, van der Vliet, P. C, Wyman, C, & Wuite, G. J. L. (2005) Analysis of scanning force microscopy images of protein-induced dna bending using simulations. *Nucleic Acids Res* **33**, e68.

[164] Bao, Q, Chen, H, Liu, Y, Yan, J, Dröge, P, & Davey, C. A. (2007) A divalent metal-mediated switch controlling protein-induced dna bending. *J Mol Biol* **367**, 731–40.

[165] Holbrook, J. A, Tsodikov, O. V, Saecker, R. M, & Record, Jr, M. T. (2001) Specific and non-specific interactions of integration host factor with dna: thermodynamic evidence for disruption of multiple ihf surface salt-bridges coupled to dna binding. *J Mol Biol* **310**, 379–401.

[166] Zurla, C, Samuely, T, Bertoni, G, Valle, F, Dietler, G, Finzi, L, & Dunlap, D. D. (2007) Integration host factor alters laci-induced dna looping. *Biophys Chem* **128**, 245–52.

[167] Vivas, P, Velmurugu, Y, Kuznetsov, S. V, Rice, P. A, & Ansari, A. (2012) Mapping the transition state for dna bending by ihf. *J Mol Biol* **418**, 300–15.

[168] Browning, D. F, Grainger, D. C, & Busby, S. J. (2010) Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. *Curr Opin Microbiol* **13**, 773–80.

[169] Kulić, I. M, Mohrbach, H, Thaokar, R, & Schiessel, H. (2007) Equation of state of looped dna. *Phys Rev E Stat Nonlin Soft Matter Phys* **75**, 011913.

[170] Revyakin, A, Ebright, R. H, & Strick, T. R. (2005) Single-molecule dna nanomanipulation: improved resolution through use of shorter dna fragments. *Nat Methods* **2**, 127–38.

[171] Revyakin, A, Liu, C, Ebright, R. H, & Strick, T. R. (2006) Abortive initiation and productive initiation by rna polymerase involve dna scrunching. *Science* **314**, 1139–43.

[172] Ali Azam, T, Iwata, A, Nishimura, A, Ueda, S, & Ishihama, A. (1999) Growth phase-dependent variation in protein composition of the escherichia coli nucleoid. *J Bacteriol* **181**, 6361–70.

[173] Stavans, J & Oppenheim, A. (2006) Dna-protein interactions and bacterial chromosome architecture. *Phys Biol* **3**, R1–10.

[174] Matthews, K. S. (1992) Dna looping. *Microbiol Rev* **56**, 123–36.

[175] Wiedenheft, B, Sternberg, S. H, & Doudna, J. A. (2012) Rna-guided genetic silencing systems in bacteria and archaea. *Nature* **482**, 331–8.

[176] Wiedenheft, B. (2013) In defense of phage: viral suppressors of crispr-mediated adaptive immunity in bacteria. *RNA Biol* **10**, 886–90.

[177] Sorek, R, Lawrence, C. M, & Wiedenheft, B. (2013) Crispr-mediated adaptive immune systems in bacteria and archaea. *Annu Rev Biochem* **82**, 237–66.

[178] Barrangou, R. (2013) Crispr-cas systems and rna-guided interference. *Wiley Interdiscip Rev RNA* **4**, 267–78.

[179] Reeks, J, Naismith, J. H, & White, M. F. (2013) Crispr interference: a structural perspective. *Biochem J* **453**, 155–66.

[180] Barrangou, R & Marraffini, L. A. (2014) Crispr-cas systems: Prokaryotes upgrade to adaptive immunity. *Mol Cell* **54**, 234–44.

[181] Vestergaard, G, Garrett, R. A, & Shah, S. A. (2014) Crispr adaptive immune systems of archaea. *RNA Biol* **11**, 156–67.

[182] van der Oost, J, Westra, E. R, Jackson, R. N, & Wiedenheft, B. (2014) Unravelling the structural and mechanistic basis of crispr-cas systems. *Nat Rev Microbiol* **12**, 479–92.

[183] Makarova, K. S, Haft, D. H, Barrangou, R, Brouns, S. J. J, Charpentier, E, Horvath, P, Moineau, S, Mojica, F. J. M, Wolf, Y. I, Yakunin, A. F, van der Oost, J, & Koonin, E. V. (2011) Evolution and classification of the crispr-cas systems. *Nat Rev Microbiol* **9**, 467–77.

[184] Bondy-Denomy, J & Davidson, A. R. (2014) To acquire or resist: the complex biological effects of crispr–cas systems. *Trends in Microbiology* **22**, 218 – 225.

[185] Ishino, Y, Shinagawa, H, Makino, K, Amemura, M, & Nakata, A. (1987) Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in escherichia coli, and identification of the gene product. *J Bacteriol* **169**, 5429–33.

[186] Yosef, I, Goren, M. G, & Qimron, U. (2012) Proteins and dna elements essential for the crispr adaptation process in escherichia coli. *Nucleic Acids Res* **40**, 5569–76.

[187] Nuñez, J. K, Kranzusch, P. J, Noeske, J, Wright, A. V, Davies, C. W, & Doudna, J. A. (2014) Cas1-cas2 complex formation mediates spacer acquisition during crispr-cas adaptive immunity. *Nat Struct Mol Biol* **21**, 528–34.

[188] Sternberg, S. H, Redding, S, Jinek, M, Greene, E. C, & Doudna, J. A. (2014) Dna interrogation by the crispr rna-guided endonuclease cas9. *Nature* **507**, 62–7.

[189] Jore, M. M, Lundgren, M, van Duijn, E, Bultema, J. B, Westra, E. R, Waghmare, S. P, Wiedenheft, B, Pul, U, Wurm, R, Wagner, R, Beijer, M. R, Barendregt, A, Zhou, K, Snijders, A. P. L, Dickman, M. J, Doudna, J. A, Boekema, E. J, Heck, A. J. R, van der Oost, J, & Brouns, S. J. J. (2011) Structural basis for crispr rna-guided dna recognition by cascade. *Nat Struct Mol Biol* **18**, 529–36.

[190] Hochstrasser, M. L, Taylor, D. W, Bhat, P, Guegler, C. K, Sternberg, S. H, Nogales, E, & Doudna, J. A. (2014) Casa mediates cas3-catalyzed target degradation during crispr rna-guided interference. *Proc Natl Acad Sci U S A* **111**, 6618–23.

[191] Westra, E. R, van Erp, P. B. G, Künne, T, Wong, S. P, Staals, R. H. J, Seegers, C. L. C, Bollen, S, Jore, M. M, Semenova, E, Severinov, K, de Vos, W. M, Dame, R. T, de Vries, R, Brouns, S. J. J, & van der Oost, J. (2012) Crispr immunity relies on the consecutive binding and degradation of negatively supercoiled invader dna by cascade and cas3. *Mol Cell* **46**, 595–605.

[192] Westra, E. R, Nilges, B, van Erp, P. B. G, van der Oost, J, Dame, R. T, & Brouns, S. J. J. (2012) Cascade-mediated binding and bending of negatively supercoiled dna. *RNA Biol* **9**, 1134–8.

[193] Wiedenheft, B, Lander, G. C, Zhou, K, Jore, M. M, Brouns, S. J. J, van der Oost, J, Doudna, J. A, & Nogales, E. (2011) Structures of the rna-guided surveillance complex from a bacterial immune system. *Nature* **477**, 486–9.

[194] Westra, E. R, Semenova, E, Datsenko, K. A, Jackson, R. N, Wiedenheft, B, Severinov, K, & Brouns, S. J. J. (2013) Type i-e crispr-cas systems discriminate target from non-target dna through base pairing-independent pam recognition. *PLoS Genet* **9**, e1003742.

[195] Cho, S. W, Kim, S, Kim, Y, Kweon, J, Kim, H. S, Bae, S, & Kim, J.-S. (2014) Analysis of off-target effects of crispr/cas-derived rna-guided endonucleases and nickases. *Genome Res* **24**, 132–41.

[196] Westra, E. R, Pul, U, Heidrich, N, Jore, M. M, Lundgren, M, Stratmann, T, Wurm, R, Raine, A, Mescher, M, Van Heereveld, L, Mastop, M, Wagner, E. G. H, Schnetz, K, Van Der Oost, J, Wagner, R, & Brouns, S. J. J. (2010) H-ns-mediated repression of crispr-based immunity in escherichia coli k12 can be relieved by the transcription activator leuo. *Mol Microbiol* **77**, 1380–93.

[197] Jackson, R. N, Golden, S. M, van Erp, P. B. G, Carter, J, Westra, E. R, Brouns, S. J. J, van der Oost, J, Terwilliger, T. C, Read, R. J, & Wiedenheft, B. (2014) Structural biology. crystal structure of the crispr rna-guided surveillance complex from escherichia coli. *Science* **345**, 1473–9.

[198] Brouns, S. J. J, Jore, M. M, Lundgren, M, Westra, E. R, Slijkhuis, R. J. H, Snijders, A. P. L, Dickman, M. J, Makarova, K. S, Koonin, E. V, & van der Oost, J. (2008) Small crispr rnas guide antiviral defense in prokaryotes. *Science* **321**, 960–4.

[199] Lipfert, J, Wiggin, M, Kerssemakers, J. W. J, Pedaci, F, & Dekker, N. H. (2011) Freely orbiting magnetic tweezers to directly monitor changes in the twist of nucleic acids. *Nat Commun* **2**, 439.

Appendix A

Protocols of flow channel preparation, DNA synthesis and ssDNA extension measurement

1. Coverslip cleaning and surface functionalization
   1.1. Place bottom coverslips (#1.5, 22 mm × 32 mm) and top coverslips (#1.5, 20 mm × 20 mm) into coverglass staining jars (each jar can hold 7 pieces of coverslips, the volume is ∼ 20 mL); Rinse the coverslips in the jars by distilled (DI) water 2-5 times.
   1.2. Add ∼20 mL of 5%-40% detergent solution into each jar, and then place in ultrasonic cleaning bath for 30 min; Rinse with DI water for >10 times to remove the detergent.
   1.3. Add ∼ 20 mL of acetone (CAUTION, toxic/flammable) into each jar and place in ultrasonic cleaning bath for 5 min; Pour the waste acetone into the waste bottle for acetone; Rinse the jars with DI water for >10 times to remove the acetone; Perform this step in fume hood for flammable chemicals.
   1.4. Dry the coverslips in the jars in oven (∼150 ^{o}C; CAUTION, hot), or by N_2 gas; Store the dried top coverslips in dry cabinet.
   1.5. Use plasma (O_2 gas) to clean the coverslips in the jar for 10 min; During the 10 min, prepare 20 mL of 1% (3-Aminopropyl)triethoxysilane (APTES) solution in Methanol (CAUTION, toxic/flammable); Immediately after the plasma cleaning, add the 1% APTES solution into the jars and incubate for 1 hour; Pour the waste into waste bottle specific for 1% APTES Methanol; Perform the methanol-related steps in fume

hood for flammable chemicals); Rinse the jars for > 10 times with DI water, and then dry by oven (∼150 ^{o}C; CAUTION, hot); Store the APTES-coated bottom coverslips in dry cabinet if not in-use for up to two weeks.

2. Assemble the flow channel

   2.1. Prepare two pieces of spacer–parafilm or double-side tape (∼ 4 mm × 20 mm) for each channel; Place the two pieces of spacer on a bottom coverslip along the long-edge; Place a top coverslip on the spacer, forming a flow cell in between (∼ 10 mm × 20 mm area).

   2.2. If parafilm is used as spacer, place the flow channel on a heater (60 - 120 ^{o}C; CAUTION, hot) for ∼ 5 -10 sec while gently press the sides of top coverslip to stick the two coverslip together by parafilm. The resulting flow channel has a height of ∼ 100 μm, thereby the volume of the channel is ∼ 20 μL.

   2.3. Seal the long edge of the channel with silicone glue to avoid leakage; Use silicone glue to make a small sink-like structure at each open edge of the flow channel, which serve as entry and exit of solution. The entry and exit can also be made by other ways, e.g., by adhering small plastic rings using wax.

3. Tether DNA onto the bottom surface of the flow channel

   3.1. Prepare amino-coated polystyrene beads (diameter: 3.00 μm) in DI water; Vortex the bead solution and then flow into channels; Incubate the bead solution in channels for ∼30 minutes - 2 hours depending on the bead concentration; Remove unstuck beads by washing with DI water for 200 μL. Adjust the incubation time to achieve surface density of 1-5 bead per 50 mm × 50 mm area. The channels deposited with the reference beads can be stored up to 3 days.

   3.2. Dilute sulfo-SMCC (sulfosuccinimidyl 4-(N-maleimidomethyl)cyclohexane-1-carboxylate) powder into 1X PBS (phosphate buffered saline) solution (0.1-1 mg/mL); Vortex the solution and then flow into the channel; Incubate the SMCC solution in channel for 30 min; Remove SMCC solution by washing with large amount (1 mL, ∼ 50 times of

the channel volume) of 1X PBS solution.

3.3. DNA tethering: Dilute thiol-biotin labeled DNA into 1X PBS, with a resulting DNA concentration of ∼0.3 nM; Gently pipette-mix the solution, flow the DNA solution into the SMCC-coated channel and incubate for 30 min; Gently wash away free DNA by 200 μL of blocking solution that contains 1X PBS with 2-4% BSA, and 0.01% 2-Mercaptoethanol.

3.4. Block the channel surfaces by incubating the channel in blocking solution for 2 hours; after this step, the channel is ready for experiments; The prepared channel can be kept at 4 ^{o}C for ∼1 day.

4. Labeled-DNA synthesis

4.1. A 576 bp DNA construct is generated by PCR from bacteriophage λ-DNA with primers: 5′-[Thiol]ATTATACTCGAGAGCATAAGCAGCG CAACA-3′ and 5′-ATTATAAGCTTATGACGCAGGCATTATGCT-3′ (underlined sequences: HindIII cutting site). The DNA is then digested by Hind III restriction enzyme for 2 hours.

4.2. Purified DNA product is incubated with 0.1-1 mM Biotin-16-dUTP, 1 mM dATP, 1 mM dGTP, 1 mM dCTP, and 3 μL Vent DNA polymerase in 100 μL reaction volume for 1.5 hours at 65-72 ^{o}C; The resulting DNA is labeled with thiol at one end, and biotin at the other end on the same strand of the DNA; Store the labeled DNA in -20 ^{o}C.

5. Identify single dsDNA tether

5.1. Tether formation: Flow in properly diluted paramagnetic beads in standard RecA reaction solution (50 mM NaCl, 10 mM $MgCl_2$, 20 mM Tris-pH 7.4) into a channel; Incubate for 10 min to allow beads to bind to biotin-labeled DNA molecules immobilized on the SMCC-coated surface through the thiol labeled end; Gently wash away untethered beads using 200 μL standard reaction solution.

5.2. Mount the channel onto the microscope stage; Search for beads on bottom surface using 100X oil immersion objective.

5.3. Select a reference bead on surface and a moving tethered bead; Build the initial image libraries of both the reference bead and the tethered

bead at different defocus planes.

5.4. Determine whether the tether is a single dsDNA molecule by applying 65 pN forces. A single dsDNA molecule is determined if the tether undergoes the characteristic DNA overstretching transition. Repeat the process until a single dsDNA thether is found.

5.5. Record the x-,y-,z- positions of the dsDNA tethered bead in a force range of 1 to 50 pN. At each force, the bead is held for 5 sec to obtain the average values of the positions. The z-position is denoted by $H_{\mathrm{dsDNA}}(f)$.

6. Generate ssDNA tether

6.1. The identified dsDNA tether is converted to ssDNA tether by applying $>$ 65 pN forces in low salt concentration ($<$ 50 mM NaCl, 0 mM $MgCl_2$) through the force-induced strand peeling transition.

6.2. Record the x-,y-,z- positions of the ssDNA tethered bead in a force range up to 70 pN. At each force, the bead is held for 5 sec to obtain the average values of the bead positions. The z-position is denoted by $H_{\mathrm{ssDNA}}(f)$.

7. Extension measurement

7.1. The force-extension curve of the ssDNA is recovered through the equation:$z_{\mathrm{ssDNA}}(f) = \Delta z_{\text{ss-ds}}(f) + z_{\mathrm{ds,WLC}}(f)$, where the extension difference, $\Delta z_{\text{ss-ds}}(f) = H_{\mathrm{ssDNA}}(f) - H_{\mathrm{dsDNA}}(f)$, $z_{\mathrm{ds,WLC}}(f)$ is the theoretical force-extension curve of dsDNA based on the worm-like-chain (WLC) polymer model of dsDNA with a persistence length of 50 nm.

7.2. After proteins, such as SSB, RecA, etc, bind to the ssDNA, the force-extension curves of the resulting nucleoprotein complexes are similarly obtained by: $z_{\mathrm{ssDNA,p}}(f) = \Delta z_{\text{ss,p-ds}}(f) + z_{\mathrm{ds,WLC}}(f)$.

PUBLICATION LIST

1. **S. Le***, Ruchuan Liu, Chwee Teck Lim, and Jie Yan, *Uncovering mechanosensing mechanisms at the single protein level using magnetic tweezers.* Methods, online (2015).

2. **S. Le, M.Yao**, J. Chen, A. K. Efremov, S. Azimi, and J. Yan, *Disturbance-free rapid solution exchange for magnetic tweezers single-molecule studies.* Nucleic Acids Res. online (2015).

3. **J. Chen, S. Le** , A. Basu, W. J. Chazin, and J. Yan, *Mechanochemical regulations of RPA's binding to ssDNA.* Sci Rep.19,5:9296 (2015).

4. **S. Le** , H. Chen, X. Zhang, J. Chen, K. Patil, K. Muniyappa, and J. Yan, *Mechanical force antagonizes the inhibitory effects of RecX on RecA filaments formation in M. tuberculosis.* Nucleic Acids Res. 42 (19): 11992-11999. (2014).

5. **H. Fu, S. Le** , H. Chen, K. Muniyappa, and J. Yan, *Force and ATP hydrolysis dependent regulation of RecA nucleoprotein filament by single-stranded DNA binding protein.* Nucleic Acids Res. 41, 924 (2013).

6. **H. Fu, S. Le**, K. Muniyappa, and J. Yan, *Dynamics and Regulation of RecA polymerization and de-polymerization on double-stranded DNA.* PLoS ONE 8(6): e66712 (2013).

7. **S. Le**, H. Chen, P. Cong, J. Lin, P. Dröge, and J. Yan, *Mechanosensing of DNA bending in a single specific protein-DNA complex.* Sci Rep. 3-3508 (2013).

8. X. Zhang, H. Chen, S. Le, I. Rouzina, P. Doyle and J. Yan, *Revealing the competition between peeled ssDNA, melting bubbles, and S-DNA during DNA overstretching by single-molecule calorimetry.* Proc. Natl. Acad. Sci. U.S.A.10.1073 (2013)

*Bold: First or Joint First Author.

9. **S. Le** , S. Golden, C. Lim, B. Wiedenheft, J. Yan, *CRISPR RNA-guided DNA deformation.* Nucleic Acids Res. under revision (2015).

10. **S. Le**, J. Chen, H. Chen, H. Fu, and J. Yan, *Dynamics and regulation of single-stranded DNA processing proteins studied by single-molecule manipulation using magnetic tweezers.* JoVE, under revsion (2015).

11. **S. Le** , S. Korolev, and J. Yan, *Competitive Regulation of SSB, RecO and RecR on the dynamics, formation and stability of RecA-ssDNA filament.* Manuscript in preparation (2014).