Université d’Aix-Marseille

Ecole Doctorale des Science de la Vie et de la Santé
Centre de Recherche en Cancérologie de Marseille
CNRS UMR7258 / INSERM U1068 / Institut Paoli-Calmettes

Thèse présentée pour obtenir le grade universitaire de docteur

Biologie - Spécialité Biochimie structurale

Hongshan ZHANG

A Single Molecule Perspective on DNA Double-Strand Break Repair Mechanisms

Soutenue le 24/07/2017 devant le jury:

Marcelo NOLLMANN CSB - CNRS Rapporteur
Marie-Agnès PETIT L’Institut Micalis - INRA Rapporteur
Erwin PETERMAN Vrije Universiteit Amsterdam Examinateur
James N. STURGIS LISM - CNRS Examinateur
Mauro MODESTI CRCM - CNRS Directeur de thèse
# Table of contents

Abstract 1

Résumé 2

1 Introduction 3
   1.1 DNA: The code of life 3
       1.1.1 DNA structure 3
       1.1.2 DNA-based cellular activity 4
   1.2 DNA double-strand break repair 6
       1.2.1 RAD51-dependent Homologous Recombination 7
       1.2.2 Single-strand Annealing 8
       1.2.3 Classical Non-Homologous End-Joining 9
       1.2.4 Alternative Non-Homologous End-Joining 9
   1.3 Single molecule analysis of DNA repair 10
       1.3.1 Single molecule versus bulk assays 10
       1.3.2 Single DNA molecule manipulation methods 11
       1.3.3 DNA mechanical property 11
       1.3.4 DNA-protein interactions 14
   1.4 Outline of this thesis 16

2 Experimental system: Combination of optical tweezers, fluorescence microscopy and microfluidics 17
   2.1 Experimental setup 18
       2.1.1 Optical tweezers 18
       2.1.2 Fluorescence microscopy 23
       2.1.3 Microfluidics 26
       2.1.4 Combination of optical tweezers, fluorescence microscope and microfluidics 28
Table of contents

2.2 Preparation of reagents ................................................. 29
  2.2.1 Terminally-labeled DNA molecules ............................. 29
  2.2.2 Fluorescent dyes for DNA molecules staining ............... 30
  2.2.3 Fluorescent labeling of proteins ............................... 30
2.3 Experimental procedures for single molecule experiments ....... 32
  2.3.1 Single DNA molecule visualization by Sytox Orange staining .. 32
  2.3.2 Visualization of the interaction of RAD52 with ssDNA using dual-trap optical tweezers ........................................... 34
  2.3.3 Two DNA molecules manipulation and visualization using quadruple-trap optical tweezers ........................................... 36

3 Human RAD52 Captures and Holds DNA Strands, Increases DNA Flexibility and Prevents Melting of Duplex DNA: Implications for DNA Recombination ........................................ 39
  3.1 Introduction .......................................................... 40
  3.2 Results and discussion ............................................... 41
    3.2.1 Experimental approach ........................................ 41
    3.2.2 GFP-RAD52 binding to ssDNA is avid and shortens ssDNA contour length .......................................................... 42
    3.2.3 GFP-RAD52 binding increases flexibility and contour length of dsDNA and prevents melting ........................................ 46
    3.2.4 GFP-RAD52 slides along dsDNA ................................ 50
    3.2.5 DNA bound GFP-RAD52 captures ssDNA in \textit{trans} .......... 53
  3.3 Conclusions and Perspectives ....................................... 55
  3.4 Supplementary ..................................................... 56

4 DNA tethering by the human XRCC4-XLF and XRCC4-Ligase IV complexes ........................................ 63
  4.1 Introduction .......................................................... 64
  4.2 Results and discussion ............................................... 65
    4.2.1 DNA interaction mode of XRCC4-XLF .......................... 65
    4.2.2 DNA tethering by the XRCC4-XLF complexes ................ 65
    4.2.3 Interaction of XRCC4-XLF with DNA ends and tethering ... 71
    4.2.4 DNA tethering by the XRCC4-LigIV complexes ............... 76
  4.3 Conclusions and Perspectives ....................................... 79
  4.4 Supplementary ..................................................... 80
    4.4.1 DNA bridging via XRCC4 or XLF alone ......................... 81
5 DNA tethering by the human MRE11/RAD50/NBS1 complex 82
  5.1 Introduction .............................................. 83
  5.2 Results and discussion .................................. 84
    5.2.1 MRN-DNA interactions are influenced by ATP, AMP-PNP and Mg$^{2+}$............... 84
    5.2.2 DNA tethering by the MRN complex ................. 85
    5.2.3 Interaction of MRN with DNA ends and tethering .... 90
    5.2.4 The MRN complex with RAD50 mutations interaction with DNA molecules ......... 91
  5.3 Conclusions and Perspectives ............................ 92

References 95
Summary 109
List of Publications 111
Fellowship Supports 112
Acknowledgements 113
Abstract

DNA double-strand breaks disrupt the physical continuity of the chromosome and are one of the most severe types of DNA damage. To preserve genome integrity against the potentially deleterious effects of DNA double-strand breaks, human cells have evolved several repair mechanisms including DNA recombinational repair and Non-Homologous End Joining (NHEJ), each catalyzed by specific enzymes. In this thesis we aimed at unraveling the dynamics of protein/DNA transactions involved in DNA double-strand break repair mechanisms at single molecule level. To do this, we combined optical tweezers and microfluidics with wide-field fluorescence microscopy, which allowed us to manipulate individual DNA molecules while directly visualize fluorescently-labeled DNA repair proteins acting on them. We focused the study on three crucial proteins/complexes involved in DNA repair: (i) the human DNA annealing protein RAD52, (ii) the non-homologous end joining human proteins XRCC4 and XLF and the complex XRCC4/Ligase IV, and (iii) the human MRE11/RAD50/NBS1 complex.

Keywords: Single molecule techniques; DNA double-strand breaks; Non-Homologous End Joining; Recombinational repair
Résumé

Les cassures double brin de l’ADN altèrent l’intégrité physique du chromosome et constituent l’un des types les plus sévères de dommages à l’ADN. Pour préserver l’intégrité du génome contre les effets potentiellement néfastes des cassures double brin de l’ADN, les cellules humaines ont développé plusieurs mécanismes de réparation, dont la réparation par recombinaison de l’ADN et la jonction d’extrémités non-homologues (NHEJ), catalysés par des enzymes spécifiques. Pendant ma thèse, nous avons caractérisé la dynamique de certaines des interactions protéines/ADN impliquées dans ces mécanismes au niveau de la molécule unique. Dans ce but, nous avons combiné des pinces optiques et de la micro-fluidique avec de la microscopie de fluorescence à champ large afin de manipuler une ou deux molécules d’ADN individuelles et d’observer directement les protéines de la réparation marquées par fluorescence agissant sur l’ADN. Nous avons concentré notre analyse sur trois protéines/complexes essentiels impliqués dans la réparation de l’ADN: (i) la protéine humaine d’appariement de brin RAD52, (ii) les protéines humaines XRCC4, XLF et le complexe XRCC4/Ligase IV de la NHEJ et (iii) le complexe humain MRE11/RAD50/NBS1.

Mots-clés: Méthodes de molécule unique; Cassures double brin de l’ADN; Jonction d’extrémités non-homologues; Réparation par recombinaison
Chapter 1

Introduction

1.1 DNA: The code of life

As we all know, elephants only give birth to little elephants, dogs to dogs, and so on for every type of living creature, and every species is so similar and yet so different. But why is this so? The underlying reason lies in one molecule called Deoxyribonucleic Acid (DNA), which carries the blueprint of all living organisms. The genetic information stored in the DNA contains the instructions for synthesizing the specific proteins of each living organisms. DNA is passed from adult organisms to their offspring during reproduction, ensuring the development, survival, and propagation of the organism.

1.1.1 DNA structure

DNA was first identified and isolated from leucocytes in 1869 by Swiss biologist and chemist Friedrich Miescher [1]. Then, in the decades following Miescher’s discovery, more and more details about DNA were unraveled. Phoebus Levene discovered the details of DNA composition: four bases (adenine [A], guanine [G], cytosine [C] and thymine [T]), deoxyribose sugars and phosphate groups [2]. Erwin Chargaff found that the DNA bases content varies between species, and within one species, the DNA bases are always present in fixed ratios: the same number of A as T and the same number of C as G [3]. Combining all these evidences with crucial X-ray crystallographic work determined by Rosalind Franklin and Maurice Wilkins, James Watson and Francis Crick proposed the groundbreaking DNA model in 1953: the DNA molecule has the form of a three-dimensional double helix [4].

The double helix of DNA arises from the chemical and structural features of its two polynucleotide chains composed of four types of nucleotide subunits, as shown
1.1 DNA: The code of life

in Figure 1.1. Each nucleotide is composed of a five-carbon sugar, one phosphate group and a nitrogen-containing base. Each carbon of the sugar is numbered as shown in Figure 1.1 A. The phosphate group is attached to the 5’ carbon position. When nucleotides are incorporated into DNA, adjacent nucleotides are linked by a phosphodiester bond: the phosphate group of one nucleotide is attached to 3’ carbon of the sugar of the other nucleotide. Because the two chains are held together by hydrogen bonds between the bases on the different strands, all the bases face inside the double helix, and the negatively charged sugar-phosphate backbones are outside. DNA bases fall into two classes: purines and pyrimidines. The purines are adenine and guanine, and the pyrimidines are cytosine and thymine. Adenine on one chain is always paired with thymine on the other chain and, likewise, guanine is always paired with cytosine. A consequence of these base-pairing requirements is that each strand of a DNA molecule contains a sequence of nucleotides that is complementary to the nucleotide sequence of its partner strand. While the sugar phosphate backbone is regular, the order of bases is irregular and this encodes the information content of DNA. The most common double helical DNA found in nature is B-DNA, which is right-handed with about 10.5 base pairs per turn (Figure 1.1 B). DNA has a diameter of 2 nanometer (nm) and base pairs are separated by 0.34 nm. Base pairing holds the two strands together in an antiparallel manner: this means that one strand has a 3’-5’ orientation while the complementary strand is oriented in a 5’-3’ direction. The two antiparallel strands of DNA twist around each other such that two distinct surfaces are generated: a minor groove that is 1.2 nm wide, and a major groove that is 2.2 nm wide.

1.1.2 DNA-based cellular activity

The structure of DNA is certainly aesthetically pleasing, but is relatively meaningless without an understanding of how it relates to function. The haploid human genome contains approximately 3 billion base pairs of DNA resulting in a length of about 2 meters. Human cells must fit this long DNA into the cell nucleus that has a diameter of less than 10 microns and they achieve this by packing DNA into chromosomes. Most cells in human body (except for female ova and male sperm) are diploid and have 23 pairs of chromosomes. This makes a total of 6 billion base pairs of DNA per cell. Cell is the structural and functional unit of life and many fundamental cellular processes, like DNA replication, gene expression, DNA repair and recombination, are based on genome.
1.1 DNA: The code of life

Figure 1.1 – DNA chemical structure. (A) Complementary base pairs in the DNA double helix. The shape and chemical structure of the bases allow hydrogen bonds to form between A and T and between G and C. (B) The structure of the DNA double helix with two antiparallel strands of DNA twisted around each other. DNA has a diameter of 2 nm and base pairs are separated by 0.34 nm. Two distinct surfaces appear in the DNA double helix: a minor groove, 1.2 nm wide, and a major groove, 2.2 nm wide.

DNA replication

DNA replication and its regulation are fundamental for life. As a cell multiplies and divides, DNA must be accurately duplicated so that genetic information is passed with minimal errors [5]. In essence, DNA replication involves the separation of the two strands of the double helix followed by the polymerization of new complementary strands on the single-stranded templates (Figure 1.2).

Gene expression

Genes in DNA encode protein molecules, which are the "workhorses" of the cell, carrying out crucial functions necessary for life. In the simplest sense, expression of a gene means manufacturing the corresponding protein via two major steps: transcription and translation, as shown Figure 1.2. During transcription, the information in the DNA is copied into a messenger ribonucleic acid (mRNA) molecule. Then, in the second step, the translation process, the mRNA is "read" and its sequence is used as template to orderly assemble the chain of amino acids that forms a protein.
1.2 DNA double-strand break repair

As cells multiply and give rise to new cells, in most cases the genome is accurately copied and information is passed on to the next generation with minimal errors. However, DNA is not inert. It is susceptible to internal and external threats. Both endogenous factors, like byproducts of intracellular metabolism, and exogenous factors including ultraviolet light, ionizing radiations, inflammatory responses, or chemical mutagens can result into DNA damage [6]. DNA damage poses a continuous threat to genomic stability and integrity. To deal with this problem, cells have evolved a variety of DNA repair pathways. This thesis will focus on DNA repair mechanisms, especially the repair of double-strand DNA breaks, which will be described in detail in section 1.2.

1.2 DNA double-strand break repair

DNA serves as the repository of the genetic information in each cell and its integrity and stability is fundamental for proper cellular functioning. DNA, however, suffers attacks from different sources resulting in damage. DNA damage, if not repaired timely and properly, can induce mutagenesis, such as base substitutions, insertions or deletions.

Figure 1.2 – DNA-based cellular activities: DNA replication and gene expression including transcription and translation. DNA replication is the process of precise DNA duplication. Gene expression encompasses the production of proteins via transcription and translation.

DNA repair and recombination

As cells multiply and give rise to new cells, in most cases the genome is accurately copied and information is passed on to the next generation with minimal errors. However, DNA is not inert. It is susceptible to internal and external threats. Both endogenous factors, like byproducts of intracellular metabolism, and exogenous factors including ultraviolet light, ionizing radiations, inflammatory responses, or chemical mutagens can result into DNA damage [6]. DNA damage poses a continuous threat to genomic stability and integrity. To deal with this problem, cells have evolved a variety of DNA repair pathways. This thesis will focus on DNA repair mechanisms, especially the repair of double-strand DNA breaks, which will be described in detail in section 1.2.
Moreover, persistent DNA damage can cause an irreversible state of dormancy of cells, named senescence, or cell suicide, known as apoptosis, or even unregulated cell division, which may increase the probability of tumor formation [7].

DNA damage can be divided into two main categories: single-strand and double-strand lesions. The former ones include lesions affecting only one DNA strand, such as oxidized or alkylated base damage, base loss, DNA adducts, intra-strand cross-links, and single-strand breaks. The latter ones, considered as the most severe type of DNA damage, are lesions that affect both DNA strands, such as inter-strand cross-links and double-strand breaks (DSBs), which directly disrupt the physical integrity of chromosomes. To deal with such potentially deleterious events, cells have therefore evolved a complex network of mechanisms to detect and repair DNA damage and ensure their genomic stability through removal of DNA lesions and reconstruction of the original genomic information.

In human cells, DSBs can be repaired using different strategies (see Figure 1.3): RAD51-dependent homologous recombination (HR), RAD51-independent single-strand annealing (SSA), alternative non-homologous end-joining (A-NHEJ) and classical non-homologous end-joining (C-NHEJ). HR requires intact homologous sequences present on a sister chromatid during the repair process, while the other pathways repair broken DNA without using sister chromatid. Since C-NHEJ does not require a template, it is not restricted to a certain phase of the cell cycle, whereas HR is believed to be only active during S and G2 phases of the cell cycle when the sister chromatid is available.

### 1.2.1 RAD51-dependent Homologous Recombination

RAD51-dependent HR pathway can be divided into three phases: pre-synapsis, synopsis and post synapsis [8]. During pre-synapsis, HR is initiated by detection and resection of a DSB to provide 3’ single-stranded DNA (ssDNA) overhangs. Specific proteins, such as the MRE11-RAD50-NBS1 (MRN) complex and the specialized enzymes CtIP and BLM, are involved in the initiation process [9]. Then, replication protein A (RPA) binds to the newly formed 3’ ssDNA region, eliminating secondary structures on the ssDNA. Because RPA bound to ssDNA forms a kinetic barrier against RAD51 filament assembly, mediators, such as the human breast and ovarian cancer tumor suppressor protein (BRCA2) and RAD51 paralogues, are needed for the formation of RAD51 filament on RPA-coated ssDNA. During synthesis, the RAD51 filament performs homology search and strand invasion, generating "D-loop" configuration where the invading strand guides DNA synthesis. RAD51 filament formation and strand invasion
are crucial features of HR. Finally, the joint intermediates are resolved, restoring two intact chromosomes.

---

**Figure 1.3** – Main DSBs repair pathways in human cells: RAD51-dependent homologous recombination (HR), RAD51-independent single-strand annealing (SSA), alternative non-homologous end-joining (A-NHEJ) and classical non-homologous end-joining (C-NHEJ). The crucial step of HR is strand invasion mediated by RAD51. RAD51 forms a nucleoprotein filament on single-stranded DNA after end resection. SSA and A-NHEJ involve homology and can directly anneal long (SSA) or short (A-NHEJ) complementary sequences (blue and green bars, respectively). A-NHEJ involving microhomology is also termed microhomology-mediated NHEJ (MMEJ). C-NHEJ can directly join broken DNA ends. Some most relevant proteins involved in each pathways are listed in the colored boxes.

### 1.2.2 Single-strand Annealing

The single-strand annealing pathway is often used to repair DSBs occurring between closely repeated sequences that can be as short as 30 nucleotide [10]. Unlike RAD51-dependent homologous recombination that needs identical sequences from sister chromatid, SSA pathway is relatively simple: the DSB ends are processed to form ssDNA tails that anneal together [11].

During the SSA process, after DSB ends resection, single-strand 3’ overhangs are exposed and coated by RPA which prevents the 3’ overhangs from sticking to themselves. Then, RAD52 mediates annealing of complementary sequences of the two RPA-coated single-stranded tails. Subsequently, leftover nonhomologous single-stranded tails are
removed by the endonuclease XPF-ERCC1. The process is finished by filling in the remaining gaps by DNA synthesis and sealing by ligation. Since SSA results in a deletion containing a single copy of repeated sequences, it is considered to be highly mutagenic. SSA does not involve DNA-strand invasion and it is independent of the RAD51 recombinase.

1.2.3 Classical Non-Homologous End-Joining

The general mechanism of C-NHEJ is usually divided into sequential steps [12] which are: (I) recognition of DNA broken ends and assembly of the C-NHEJ complex at the DNA DSB site; (II) processing of the DNA ends for ligation; and (III) ligation of the broken ends and dissolution of the NHEJ complex.

Specifically, the initial step in C-NHEJ is the recognition of DSBs by Ku heterodimer (Ku70/80) which binds to broken ends [13]. The ability of Ku to quickly localize at DSBs is likely due to its extraordinary affinity (dissociation constant of $2 \times 10^9$ M$^{-1}$) for DNA ends, its ability to interact with DSBs in a sequence independent manner, and its abundant concentration (~500,000 Ku molecules/cell) [14, 15, 16, 17]. Once the Ku heterodimer is bound to the DSB ends, it serves as scaffold to recruit the other C-NHEJ factors, including DNA-PK catalytic subunit (DNA-PKcs) [18], Artemis [19], X-ray cross complementing protein 4 (XRCC4) [20, 21], XRCC4-like factor (XLF) [22], and DNA ligase IV [21], to the damage site. Subsequently, the two broken DNA ends are processed by Artemis and bridged by XRCC4 and XLF complex, which can form alternated filaments via interaction of their head domains [23, 24, 25]. The final step is the ligation of the broken ends, which is catalyzed by DNA ligase IV in complex with XRCC4 and is stimulated by XLF [26].

1.2.4 Alternative Non-Homologous End-Joining

In addition to HR, SSA and C-NHEJ, recent studies have demonstrated another DSBs repair process, functioning on simple end-joining principles, but repairing DSBs slower than C-NHEJ [27]. This repair pathway is considered to be an alternative form of NHEJ and is frequently abbreviated as A-NHEJ [28, 29], also termed microhomology-mediated NHEJ (MMEJ) [30].

In normal cells, the C-NHEJ components are more efficient and are sufficient to repair most DSBs, making C-NHEJ dominant with respect to A-NHEJ. When some core factors for C-NHEJ, such as Ku70/80 or XRCC4-ligase IV-XLF, are absent, A-NHEJ is called into action. In comparison to C-NHEJ, A-NHEJ is poorly understood. But these
following features characterize A-NHEJ. First, unlike C-NHEJ, it is independent of Ku70/80 and XRCC4-ligase IV-XLF; in fact, poly(ADP-ribose) polymerase-1 (PARP1) initializes A-NHEJ [31, 32] and other components, including the MRN complex [33, 34, 35], CtIP [36] XRCC1, DNA ligase I, and ligase III [33], also play important roles during this repair process. Second, the junction of A-NHEJ relies on short microhomologies (1 to few base pairs), which can cause chromosome translocations and lead to deletions [37]. Third, A-NHEJ serves as a backup when C-NHEJ is impaired in cells [7].

1.3 Single molecule analysis of DNA repair

In this thesis, I will use a single molecule approach to study certain aspects of DNA DSB repair mechanisms. Different single molecule methods and their advantages and disadvantages will be described in this section.

1.3.1 Single molecule versus bulk assays

In conventional ensemble experiments, when researchers study a particular biomolecule, typically, they use solutions where the molecule of interest is at concentrations in the millimolar (mM) to micromolar (µM) range [38]. Thus, a large number of biomolecules are interrogated simultaneously, and averaged-out properties are primarily observed. In bulk biochemical measurements, the averaging over all conformations and species obscures the effects of individual contributions to the observed signal and may lack some transient or intermediate molecular state.

Nevertheless, unravelling these transient and intermediate states as well as the dynamic properties of biomolecules is indispensable for a deep understanding of the biological processes in which they are involved. This promoted the development of single molecule approaches. Single molecule methods allow to study one molecule at a time. Following in real time the progression of a single molecule from reactant to product can reveal intermediate states and unravel detailed mechanisms of a biochemical reaction. Additionally, single molecule methods have excellent sensitivity compared to bulk assays [39]. Thus, biomolecules can be studied at very low concentration (nanomolar (nM), or even picomolar (pM)) that is similar to the concentration of physiological conditions.

For these reasons, single molecule methods have become as a powerful approach to study proteins and DNA interactions in DNA repair. As described in section 1.2, DNA repair involves many processes with multi protein partners continuously assembling and
disassembling at damaged sites, and single molecule analysis can make it possible to monitor and follow each repair step from initial DNA damage recognition to completed DNA repair in real time. A point of caution, however, is that correct interpretation of single molecule experiments requires enough data for statistics, which can be costly and time-consuming.

1.3.2 Single DNA molecule manipulation methods

In addition to the advantages discussed above, certain single molecule techniques provide the unique possibility to directly manipulate single molecules, apply forces and measure tensions, as well as measure the functional and structural responses to mechanical manipulation. Three distinct single molecule methods have been applied to DNA repair research: atomic force microscopy (AFM) [40], magnetic tweezers [41] and optical tweezers [42]. AFM uses a sharp tip mounted at the end of a flexible cantilever as a probe to image single DNA molecules deposited on a flat surface. It can also exert pulling force on a single DNA molecule (attached to the surface) by retraction of the tip tethered to the other DNA end. AFM has a unique ability to apply forces as high as nanonewtons (nN), but it is limited in low force detection (piconewton (pN) level) [40]. Magnetic tweezers use an external, controllable magnetic field to exert force and torque on a single DNA molecule using a superparamagnetic bead attached to one DNA end. It can twist single DNA molecule to study the role of torque and DNA supercoiling and has femtonewton accuracy for force measurement. Another advantage is that multiple DNA molecules can be manipulated and measured simultaneously [41, 43]. Optical tweezers, also known as optical traps, is arguably the most versatile single molecule manipulation method. A single DNA molecule is tethered to one or two microspheres, which are trapped by highly focused laser beams. The displacement of the bead from the trap center results in a restoring force, which can be quantified using for instance back-focal plane interferometry [44]. Optical tweezers provide forces from 0.1 pN to few hundred pN and have a angstrom position resolution and millisecond time resolution [42].

1.3.3 DNA mechanical property

The primary function of DNA is to store and propagate the genetic information, so DNA sequence is often regarded as its most important biochemical property. However, during DNA repair processes, DNA is often manipulated in many ways as a physical object regardless of its specific sequence. During repair, DNA needs to be found, read,
cut, and manipulated during which its mechanical properties may be changed. These changes can be directly measured by single molecule techniques.

**DNA stretching**

In solution, free DNA adopts a random coil conformation which maximizes its entropy. Its physical behavior is essentially similar to that of a twisted rubber hose. Pulling experiments of dsDNA confirmed that B-DNA is an elastic molecule whose force-extension relationship at low force can be well described by the inextensible worm-like chain model (WLC) [45, 46, 47]. This model treats the dsDNA as a flexible chain characterized by the persistence length \( L_P \) that is the distance over which two segments of the chain remain directionally correlated (the persistence length of dsDNA is \( \sim 50 \) nm or \( \sim 150 \) base pairs) [48].

Figure 1.4 – (A) First reported dsDNA force-extension curves. Data were obtained in 150 mM NaCl, 10 mM EDTA, pH 8.0. Reprinted from [49]. (B) Overstretching transition portions of the dsDNA force–extension curves as function of the Na\(^+\) concentration, reprinted from [50]. The overstretching force decreases with salt concentration. The data were obtained by stretching single DNA molecules in solutions with Na\(^+\) at 1000 mM (●), 500 mM (○), 250 mM (▲), 100 mM (△), 50 mM (+), 25 mM (×), 10 mM (▼), and 2.6 mM (▽).

A commonly formula that describes the end to end distance \( x \) of WLC with contour length \( L_0 \) at particular stretching force \( F \) is [45]

\[
\frac{F L_P}{k_B T} = \frac{1}{4} \left( 1 - \frac{x}{L_0} \right)^2 + \frac{x}{L_0} - \frac{1}{4} \tag{1.1}
\]
where $k_B$ is the Boltzmann constant and $T$ is the absolute temperature. This equation describes the entropic elasticity of a WLC, arising from the reduced entropy of the stretched chain, and assumes that DNA is inextensible.

At high force, the inextensible WLC has deviations from dsDNA extension behavior, due to the stretching of DNA beyond its contour length. This region is termed enthalpic regime [51, 46]. This regime can be well described by extensible worm-like chain (eWCL) model by taking into account another parameter, the stretch modulus, which reflects the intrinsic resistance to longitudinal strain [51, 46]. The eWLC model can be expressed as

$$x = L_0 \left[ 1 - \frac{1}{2} \left( \frac{k_B T}{F L_p} \right)^{1/2} + \frac{F}{S} \right]$$

(1.2)

where $S$ is stretch modulus. From this description, two important physical properties of DNA, persistence length and stretch modulus, can be obtained. First measurements of the elasticity of individual DNA molecules were reported in 1992 [52], where relatively modest tensions (<30 pN) were applied. Four years later, first fully extended DNA curves were reported [49], as shown in Figure 1.4 A, where the force-extension curve was fitted with inextensible WLC and eWLC models below 65 pN, indicating that WLC is suitable to describe DNA extension at low forces, while eWLC can well describe the DNA behavior at both low and high forces. When dsDNA is subject to forces of 65 pN or more, it suddenly changes, stretching up to 70% beyond its canonical B-form contour length without significant increase of the tension. This plateau at about 65 pN shows that dsDNA molecule undergoes a structural rearrangement, referred as overstretching transition. Thermodynamic reversibility can be observed when the tension on dsDNA is relaxed. The relaxation curve coincides with the stretching curve below 50 pN. Hysteresis between stretching and relation curves is due to partial melting of dsDNA during stretching.

Above 100 pN forces, dsDNA fully melts into single strand that exhibits the characteristic force-extension behavior of single-stranded DNA. The persistence length of ssDNA is $\sim 0.7$ nm, two orders of magnitude smaller than that of dsDNA, representing just 2 bases and reflecting the high flexibility of ssDNA [49]. The ssDNA elasticity was also well-characterized and was shown to be well-fitted by the extensible freely jointed chain (FJC) model [52] and almost as well using eWLC model [53].

These DNA models are applicable under many experimental conditions (in *vitro*), and perhaps also in physiological conditions (in *vivo*) where DNA experiences low to
intermediate forces in the piconewton range [54]. For example, DNA stretching may happen during cell division when chromosome segregation occurs [55].

**Effects of ions on DNA**

DNA is held together via hydrogen bonds of complementary bases and stacking interactions of adjacent bases while experiences electrostatic repulsion from phosphate backbone. Overall DNA structure is a negatively charged. Ionic strength and the nature of the ions present have major effects on DNA stability. Salt concentration predominantly affects DNA structure through interactions with the negatively charged phosphate backbone [56]. In monovalent salt (e.g. Na$^+$), the measured persistence length is consistent with an electrostatic contribution that varies inversely with the ionic strength, while the elastic stretch modulus is approximately independent from the ionic strength. The overstretching transition force decreases as the salt concentration decrease (Figure 1.4 B), which indicates a weakening of the stability of DNA double helix with reduced salt concentrations [50].

**1.3.4 DNA-protein interactions**

DNA-protein interactions play a crucial role in DNA repair. Each repair pathway requires specific DNA repair proteins. The traditional view of DNA-protein interaction in molecular biology considers DNA as a passive, relatively stable and uniform element and the proteins as active players. However, DNA is not a static object but has its own dynamic behavior and thermal agitation can affect DNA structure, which further influence interactions with proteins. DNA-binding proteins can be divided into two categories. Some proteins, as specific transcription factors, can recognize and bind to specific DNA sequences. And other proteins, which that are not sensitive to DNA sequences, may recognize and interact with particular DNA structures (e.g. ssDNA or DSBs). Several aspects of DNA-protein interactions can be characterized using single molecule techniques and they are described as follows.

**Mechanical analysis of DNA-protein complexes**

Force can be easily applied to DNA-protein complexes to analyze their mechanical response and interrogate their structural properties. Naked DNA force-extension curves (see section 1.3.2) characterize the intrinsic mechanical properties of DNA molecules. However, the elastic properties (contour length, persistence length, and stretch modulus) of DNA may and actually do change upon protein binding. Hysteresis
between DNA stretching and relation curves is a sensitive indicator of protein binding. Moreover, applying force to DNA-protein complexes can be used to monitor some interesting events such as rupture of the complexes, yielding unique insights into the thermodynamics of DNA-protein interactions and unraveling their interaction mechanisms.

**Strength measurement of protein mediated assemblies**

One essential process in DNA repair is protein-mediated DNA bridging, which tethers and holds together two parts of broken DNA and acts as scaffold to recruit other factors for further repair processes. For instance, in C-NHEJ, it is accepted that DNA-PK creates a stable molecular bridge between two DNA ends [57]. And before doing DNA ligation by DNA ligase IV in the final stage, two broken DNA can be bridged by XRCC4/XLF complex [58]. At initial stage of HR and A-NHEJ repair pathways, MRN complex can bridge two DNA ends [59]. Up to now, these bridged structures and their mechanical properties are poorly understood. Knowing the strength of these bridged DNA-protein complexes is one key point to understand their functions. Single molecule manipulation can be used to quantitatively describe these bridging abilities.

**Protein diffusion in real time**

One of the most important questions that remain in DNA repair is how the many diverse different DNA repair proteins, search for specific damage sites, recognize them and decide what corresponding repair pathway to use. The ability of proteins to travel along DNA while seeking out their respective targets has long been recognized as an important mode of interaction that allows many different proteins to efficiently fulfill their biological functions. Single molecule detection permits to directly visualize individual protein diffusing along DNA in real time and to obtain quantitative information regarding the role of protein diffusion in DNA repair mechanisms. In fact, the ability of a protein diffusing along DNA may not only provide a mechanism for target searching, but also allow the coordination with other proteins considering that many proteins are involved in DNA repair. By directly imaging the motion of a protein on DNA, the diffusion coefficient of the protein and the energy of the protein–DNA interaction can be derived. In addition, how protein diffusion behavior is influenced by DNA tension can further be analyzed.
1.4 Outline of this thesis

Single molecule techniques have become widespread and offered unique insights into aspects of DNA repair unattainable with ensemble experiments. Optical tweezers enable to manipulate single DNA-protein complex and measure their mechanical properties with exquisite sensitivity. Fluorescence microscopy has made it feasible to visualize individual proteins dynamic behavior along DNA. In addition, controlled microfluidics laminar flow cell makes it possible to fast control the chemical conditions in single molecule experiments.

Chapter 2 will describe our experimental setup: a combination of optical tweezers, wide-field fluorescence microscopy and laminar flow microfluidics. Using such an apparatus, we can study the dynamic behavior of proteins on single DNA molecules.

Human RAD52 is involved in HR and SSA repair processed and it promotes complementary ssDNA annealing. But does human RAD52 alter the DNA mechanical properties? What is the mechanical behavior of RAD52-ssDNA and RAD52-dsDNA complexes and what are the differences? What is the RAD52 dynamic behavior on DNA? All these questions will be addressed by study DNA-RAD52 interactions at single molecule level in Chapter 3.

Chapter 4 will focus on the human XRCC4, XLF and DNA Ligase IV proteins involved in C-NHEJ. Specially, a powerful way to manipulate two DNA molecules and directly visualize their interactions with these proteins will be described and used to assess the bridging property of human XRCC4/XLF and XRCC4/Ligase IV complexes.

Chapter 5 will discuss the interactions of the human MRE11/RAD50/NBS1 (MRN) complex with DNA. The bridging capability of the MRN complex will be investigated under different buffer conditions.
Chapter 2

Experimental system: Combination of optical tweezers, fluorescence microscopy and microfluidics

Single molecule experiments have unraveled processes that were previously concealed in bulk assays. New insights into the function, structure and interactions of individual molecules have been driven by technological advances in single molecule manipulation and detection. Recent developments of single molecule methods have made it possible to combine optical tweezers, fluorescence microscopy and microfluidics flow system. Such a combination is particularly powerful to study DNA-protein interactions as it allows to manipulate single DNA molecules while observing dynamic behavior of individual proteins on DNA with unprecedented levels of precision and control. This chapter will describe the building up and operation of an experimental system based on optical tweezers, wide-field fluorescence microscopy and microfluidics. Also, several examples will be presented to illustrate the capability of these techniques for the study of DNA-protein interactions.
2.1 Experimental setup

2.1.1 Optical tweezers

Optical tweezers, also known as optical traps, is a technique that uses a highly focused laser beam to create a potential well, in which a micron or sub-micron sized particle, with refractive index greater than that of the surrounding medium, can be stably trapped (e.g. micron-sized polystyrene bead (n=1.58) in water (n=1.33)). The trapped particle can be manipulated by controlling the position of trapping beam. In the plane perpendicular to the propagation direction, the beam intensity has a Gaussian profile for a TEM00 laser mode. The focused laser beam exerts two types of forces on particle: the scattering force, pushing the particle in the direction of the beam propagation, and the gradient force, pointing to the direction of the intensity gradient of the beam and tending to pull the particle towards the beam focus. In general, the scattering force dominates over the gradient force. However, if the laser beam is tightly focused, the intensity gradient around the laser focus becomes very steep, which makes the gradient force surpass the scattering force, leading to stable optical trapping.

For a particle that has a diameter much larger than the wavelength of the trapping beam, conservation of momentum and geometrical (ray) optics can be used to simply describe the principle of optical trapping. Light consists of photons. Thus, if the direction of a photon is altered by a micro-sized bead, by refraction (Figure 2.1 A) and reflection (Figure 2.1 B), a corresponding force is exerted on the bead. On macroscopic objects, this force can be neglected. But on micron or sub-micron sized beads, the effect of these forces can clearly be observed, if a high intensity laser beam is focused by a high numerical aperture microscope objective. In Figure 2.1 C, two rays differ in intensity, resulting from the Gaussian intensity profile, and the momentum change resulting from each ray is proportional to its intensity. The total momentum change generates a net force on the bead that points to the highest intensity position (center of the beam). When the bead is out of focus in the axial direction, the bead is pushed by the reflection of light from its surface, while a restoring force due to refraction pulls the bead toward the focus. To stably trap the bead, the restoring force due to refraction must be sufficient to overcome the scattering force. When the restoring force is big enough, the net force acting on the bead will point to the beam focus, as shown in Figure 2.1 D. Note that due to the scattering force, the equilibrium position of the trapped bead is slightly behind the focal point.
Figure 2.1 – Geometrical optics description of light momentum change and corresponding force on bead. (A) Incident ray $\vec{p}_{in}$ is refracted into $\vec{p}_{out}$ leading to corresponding light momentum change $\Delta\vec{p}$. According to conservation of momentum, the bead gets an opposite momentum exerting a restoring force due to refraction. (B) the bead also experiences a scattering force resulting from reflection. (C), (D) One laser beam is focused by high numerical aperture microscope objective. Two incident rays from the beam, $\vec{p}_{1,in}$ and $\vec{p}_{2,in}$, are refracted by one bead into $\vec{p}_{1,out}$ and $\vec{p}_{2,out}$ resulting in corresponding light momentum changes $\Delta\vec{p}_1$ and $\Delta\vec{p}_2$. (C) shows one bead that is away from the beam center in the lateral direction. Due to refraction, the direction of rays from the beam is altered when they pass through the bead. Because the momentum of the photons is changed, the force exerted on the bead from the high intensity ray (thick) is stronger than that from low intensity ray (thin). The direction of forces is indicated with red arrows and the magnitude is indicated with the size of the arrows. Net force (green arrow) on the bead points to the direction of the highest intensity. (D) one bead out of the focus in axial direction is shown. On one hand, the bead is pushed downward by the reflection of the light from its surface (black dashed arrows: reflection ray; red dashed arrows: scattering force due to reflection). On the other hand, the restoring force due to refraction pulls the bead upwards (black solid arrows: refraction ray; red solid arrows: restoring force). To stably trap the bead, the restoring force must overcome the scattering force. If the restoring force is strong enough, the net force (green arrow) acting on the bead will point upwards.

Optical tweezers can be used to not only trap beads but also to directly measure the force acting on the trapped bead. A bead away from trapping center experiences a restoring force towards the focus. The amplitude of this force $F$ is proportional to the distance, $x$, between the bead and the trap center. This relation can be described by the Hooke’s law and follows the equation $\vec{F} = -k\vec{x}$, where $k$ is a constant, named the trap stiffness. Thus, to measure the external force acting on a trapped bead, the
displacement $x$ and the trap stiffness $k$ should be accurately determined. The position of the trapped bead can be determined with a quadrant photodiode (QPD) located at a conjugate optical plane of the back-focal plane of the condenser lens (a detection objective in our system), as shown in Figure 2.2 A.

To get the trap stiffness $k$, the optical tweezers system need to be calibrated. One calibration procedure is based on the analysis of the power spectral density resulting from the Brownian motion of the trapped bead in a viscous liquid. The power spectral density $X(f)$ has a Lorentzian shape [60],

$$X(f) = \frac{k_B T}{\gamma \pi^2 (f_c^2 + f^2)}$$

where $k_B$ is the Boltzmann constant, $T$ is the absolute temperature, $f_c = k/2\pi\gamma$ is the corner frequency of the trap and $\gamma = 6\pi \eta d$ is the viscous drag acting on the bead. When the viscosity $\eta$ of the medium and the diameter $d$ of the bead are known, the trap stiffness $k$ can be derived from the corner frequency, yielding the conversion factor from displacement to force. In our experiments, the beads with a diameter of 4.26 µm (Spherotech) are normally used, and the viscosity value considered is 0.89 mPa·s, which is the viscosity of water at 25 °C. Figure 2.2 B shows the typical power spectral density of the $X$ direction from our experiments. The corresponding corner frequency (1.344 kHz) is obtained by Lorentzian fitting and thus the trap stiffness (0.3018 pN/nm) is deduced.

In our experimental apparatus, the optical tweezers is based on the 'NanoTracker™ 2' produced by JPK Instruments AG (Germany)[61]. To minimize photodamage on most biological samples [62], a 1064 nm laser (with maximal power 5 W) is used as trapping light source. Two optical traps are generated by splitting of the laser light into two orthogonally polarized beams using a polarizing beam splitter (PBS). The orthogonally polarized beams enable us to independently detect the two traps. One trap is controlled by a piezo-electric mirror for movement in the sample plane ($X$, $Y$) and the steering in $Z$ axis is performed via a movable lens hooked to a linear motor. The other beam passes through an acousto-optic deflector (AOD), which moves the trap in ($X$, $Y$) and another movable lens is used to control $Z$ position of the trap. The laser light is focused into the sample by using a high numerical aperture water immersion objective (60X NA=1.2, Nikon). This ensures a high trap stiffness in all axes and minimal trap degradation when working deep into the sample. The absolute intensity and the power ratio between the two trapping beams can be controlled in a continuous manner. The sample stage consists of a long-distance motorized XY translation stage and a 100 µm × 100 µm × 100 µm piezo stage operated in closed-loop.
2.1 Experimental setup

Figure 2.2 – (A) Back-focal plane interferometry. The configuration for lateral displacement detection is shown. The back-focal plane of the detection objective is projected onto a quadrant photodiode (QPD). When the bead moves by external force away from the trap center, the transmitted laser light is deflected (from red curve to blue curve) resulting in a signal of the QPD. By reading out the differential signals \((I_1+I_3)-(I_2+I_4)/(I_1+I_2+I_3+I_4)\) and \((I_1+I_2)-(I_3+I_4)/(I_1+I_2+I_3+I_4)\), the displacements of \(x\) and \(y\) can be obtained. (B) Typical power spectral density of the \(X\) direction used for calibration. In this experiment, one polystyrene bead with a diameter of 4.26 \(\mu\)m was trapped by 2.5 W trapping beam in water at room temperature. From Lorentzian fit, the corner frequency (1.344 kHz) and the trap stiffness (0.3018 pN/nm) are obtained.
2.1 Experimental setup

The forward scattering laser light is collected using a long working distance water immersion objective (60X NA=1.0, Nikon) and detected by a XYZ detection unit including two QPDs, which are used to independently detect 3-dimensional forces and displacements for the two traps based on polarization.

We measured the movement range of the two traps in sample plane. For each trap, the ideal movement range should have minimal difference of trap stiffness. We thus measured the trap stiffness at intervals of 10 \( \mu \text{m} \) along \( X \) and \( Y \) axis in the entire field of view for each trap and determined a 30 \( \mu \text{m} \times 40 \mu \text{m} \) common movement range with minimal stiffness variations and big enough for our purpose of DNA manipulation.

Besides two traps, our optical tweezers can generate multiple independent traps thanks to the AOD. This is based on the so-called time-sharing multiplexing approach that relies on quickly moving a beam from one trap position to the next one where it resides in a certain dwell time [63]. In an AOD, a crystalline material is bonded to a piezoelectric transducer that converts a high frequency (MHz) oscillating voltage into an acoustic wave. The wave propagates through the crystal with the same frequency as the applied voltage and forms a standing wave pattern of compressions and expansions parallel to the transducer, leading to periodic changes in the index of refraction of the crystal [64]. So, when a beam passes through the AOD, it is deflected and a fast change in the frequency of the applied voltage provides a high switching rate (up to MHz) of beam positioning that can be used to create hundreds of simultaneous stable traps using one beam. Nevertheless, stable optical trapping needs enough dwell time on each trap and uninterrupted manipulation requires that the bead should not escape from the trap when the laser is scanning the other positions. In our experiments, after optimization, the dwell time was set to 60 \( \mu \text{s} \). An example of trap multiplexing with 27 parallel traps is shown in Figure 2.3.

![Figure 2.3](image-url)  
*Figure 2.3 – Example of trap multiplexing. Beads with a diameter of 1.87 \( \mu \text{m} \) are simultaneously trapped by 27 traps in solution and arranged to form the CRCM logo. Note that some traps have more than one bead. Scale bar: 5 \( \mu \text{m} \).*
2.1.2 Fluorescence microscopy

A fluorescence microscopy is an optical microscope that uses fluorescence to generate an image. Its basic function is to irradiate a fluorescently-labeled specimen with a specific band of wavelengths, and then to spectrally separate the emitted fluorescence from the excitation light. The excitation laser beam is reflected from a specific dichroic mirror, and then passes through the microscope objective to illuminate the specimen with intense light. The emitted light from the fluorescent sample gathered by the same objective passes through the dichroic mirror, and is subsequently filtered by an emission filter, which blocks unwanted excitation wavelengths. The fluorophore emission spectral profile is often a mirror image (or nearly so) of the excitation curve, but shifted to longer wavelengths. This well-documented phenomenon is known as Stokes’ shift. In order to achieve maximum fluorescence intensity, a fluorophore is usually excited at wavelengths near or at the peak of its excitation spectrum, and the widest possible range of emission wavelengths, including the emission peak, is selected for detection. Also, coupling different fluorescent probes to different molecules allows us to distinguish them using multi-color excitation and detection.

![Figure 2.4](image)

Figure 2.4 – Schematic of wide-field fluorescence microscopy applied to DNA-protein interaction studies. One DNA molecule is tethered by two optical trapped beads in a fluorescent protein solution. The entire field of view is illuminated by a collimated excitation laser (light green). All fluorescent proteins (dark green dots) either in surrounding medium or on the DNA molecule are excited and emit fluorescence.

In general, fluorescence microscopy refers to any microscope that uses fluorescence to generate an image, whether it is a simple setup like a wide-field (epi-fluorescence) microscope, or a complicated design such as a confocal microscope, which has better resolution. For wide-field microscopy, the entire field is illuminated, and the whole image of the sample is captured at once with a camera. However, because the whole field is illuminated, all fluorescent probes within the field are excited and emit fluorescence (see
Figure 2.4), which results in a high background noise and therefore it is not suitable for single molecule imaging in a solution with a high concentration of fluorescently-labeled proteins.

In our experiments, we use an inverted microscope (Nikon Eclipse Ti-U) around which the whole setup is assembled. A blue light-emitting diode (LED) is used to illuminate the whole field to facilitate optical trapping and sample manipulation. The trans-illumination image is recorded by a charge-coupled device (CCD) camera. The excitation lasers source is a commercial laser combiner (Andor, Belfast, Northern Ireland) which includes three lasers (488 nm, 532 nm and 640 nm wavelength) coupled to an optical fiber. Before the excitation light coupling into microscope, a quarter-wave plate is used to generate a circular polarization light to avoid the influence of the linear polarization on sample. Three different excitation lasers allow us to acquire 3-color images. Since the optical tweezers is integrated in the microscope, the influence of the optical trapping light cannot be neglected during fluorescent imaging. Typically, the power of the trapping laser is in the order of Watts, leading to a density of energy in focus of 100 MW/cm$^2$, while epi-fluorescence excitation lasers use power in the range of milliwatts (100 W/cm$^2$ in focus). In contrast, the intensity of typical fluorescence signals is in the order of femtowatts [65]. To detect such weak fluorescence signals, we choose optical dichroic mirrors and filters. A dichroic mirror (DLHS NIR 1064 nm, Qioptiq), DM 2 in Figure 2.8, is used to couple the 1064 nm light to the microscope. It reflects IR and transmits the excitation light, the blue LED light and the fluorescence signals. Another dichroic mirror (ZT405/488/532/640, Chroma), DM 3 in Figure 2.8, is used as main dichroic and reflects the excitation light and transmits the fluorescent signals. A filter (FF01-842/SP, Semrock) (Filter 2 in Figure 2.8) is used to suppress the infrared trapping light in the detection pathway. The microscope detection system mainly includes the CCD for transmission illumination, an Optosplit III (Cairn Research Ltd, England) and an electron-multiplying charge-coupled device (EMCCD) (iXon DU-897; Andor, Belfast, Northern Ireland) for fluorescence detection. The Optosplit III is a 3-way image splitter device (Figure 2.5) for dividing an image into one, two or three separate, spectral components which can be aligned side by side on a camera chip, enabling a single EMCCD to record images simultaneously in one, two or three different colors (Figure 2.6). In Optosplit III, we can combine different dichroic mirrors and filters. Specifically, we use an emission filter FF01-510/20 (Semrock) for the detection of GFP or Sytox Orange. we use the emission filter FF01-572/28 (Semrock) for the detection of Alexa fluor 555. The emission filter FF01-605/15 (Semrock) is used for the detection of mCherry. The emission filter FF01-670/30
2.1 Experimental setup

(Semrock) is used for detection of Atto 647N. The Dicroic mirror 540dclp (Chroma) is used to simultaneously record fluorescent signals GFP and Alexa fluor 555 or GFP and mCheery. The dicroic mirror 625dclp (Chroma) is used to simultaneously record GFP and Atto 647N. The combination of these dichroic mirrors and filters is based on the experimental requirements and constraints. The entire microscope detection system is enclosed in a black box to reduce the background noise.

![Figure 2.5](image)

**Figure 2.5** – Schematic of the Optosplit III. The image size is selected by an aperture. Fluorescence signals (yellow) are divided into three components (blue, green and red) by different dichroic mirrors. Emission filters are used to select the desired detection bands, and reflecting mirrors are used to align the images recorded by the EMCCD.

![Figure 2.6](image)

**Figure 2.6** – Three color wide-field fluorescence images of fluorescent beads simultaneously recorded. (A) Fluorescent beads excited with the 488 nm laser. (B) Fluorescent beads excited with 532 nm laser. (C) Fluorescent beads excited with 640 nm laser. (D) Merge of the three images. Scale bar: 5 µm.
2.1 Experimental setup

2.1.3 Microfluidics

Single molecule studies of DNA-protein interactions require precise and fast control of the chemical environment. Also, wide-field microscope has significant background noise caused by free fluorophores floating in surrounding solution. To facilitate exchange of different chemical conditions and reduce background noise, a microfluidics system based on a multi-channel laminar flow cell is used. In the laminar flow cell, the mixing between different solutions only occurs within a few microns around the buffer interface [66]. In our system, flow cells and the corresponding holder are custom-made. Figure 2.7 A shows the schematic of a flow cell with five channels. Three inlet channels 500 µm-wide and 100 µm-deep combine to form one main channel 2 mm-wide and 100 µm-deep. The fourth and fifth inlet channels (500×100 µm) join the main channel in the middle of the flow cell. In general, we use channel 1 as beads channel and channel 2 as DNA channel. The components in channel 3, 4, and 5 are different according to different experimental requirements. Polyethylene tubes are used to connect the inlets to solution reservoirs and the output to a trash container. The flow cell is mounted on the wide-field inverted microscope sample stage using a holder and allows rapid switching of buffer conditions simply by moving microscope sample stage such that the optically trapped DNA/DNA-protein molecules are brought into the different channels. Figure 2.7 B shows our 5-channel flow cell filled with red and blue dyes to illustrate that the mixing of solutions is negligible when the flow is on and very slow when the flow is off.

The flow in channels is generated and controlled by a pressure control unit (MFCSTM-EX: Extended Flow Control, FLUIDGENT, France) which allows to apply stable flow with a short response time (100ms) and a stabilization time of 1 s. The input pressures can be independently applied to different reservoirs. To understand the relationship between the pressure and the flow velocity, we trapped one bead with a diameter of 4.26 µm near the center of the main channel. By releasing the bead and measuring its displacement versus time at different pressures, we find that the flow velocity has a linear relationship with the input pressure (Figure 2.7 C). We also measured the force acting on a trapped bead at different flow pressures and find again a linear relationship (Figure 2.7 C). Figure 2.7 D shows that the force acting on the bead increases as the flow velocity increases. In fact, the force acting on particles in a fluid cell is determined by the Stokes' drag force $\vec{F}_d$, described as $\vec{F}_d = -6\pi\eta r \vec{v}$, where $r$ is the radius of the particle, $\eta$ the viscosity of solution and $\vec{v}$ the flow velocity. Nevertheless, as shown in Figure 2.7 D, the measured force has a deviation from the drag force. The deviation most probably arises from the flow velocity measurement.
2.1 Experimental setup

Figure 2.7 – (A) Schematics of a 5-channel flow cell with five inlets and one outlet. Channel 1 is the beads channel and channel 2 is DNA channel. Channel 3 and channel 4 are buffer channels and channel 5 is loaded with protein solution. Normally, the imaging is done with flow off in buffer channel, as shown the position of the blue rectangle. (B) Our 5-channel flow cell filled with red and blue dyes. When all channels are flowing, the solutions in the 5 channels do not mix. Next, channels 1 and 2 are closed, and after 30 s all channels are closed. 180 s after stopping the flow there is no mixing deep into channel 4 in the imaging area. Scale bar: 2 cm. (C) The flow velocity is linearly dependent on the applied pressure, and the force acting on the bead also has a linear relationship with the applied pressure. The measurement is done near the center of the main channel. (D) Relationship between the flow velocity and the force acting on a trapped bead. The forces measured from the trap is shown in black while the red squares are the values expected from the Stokes’ drag force.
2.1.4 Combination of optical tweezers, fluorescence microscope and microfluidics

Figure 2.8 – Simplified diagram of the experimental setup including trapping laser steering, 3D force/position detection, microfluidics flow, excitation laser and microscope detection systems. Trapping beams, excitation beams, LED light and fluorescence signals are shown in red, light green, blue, and yellow, respectively. DM indicates dichroic mirrors, PBS polarizing beam splitter, QPD quadrant photodiode, and LED indicates light-emitting diode.

As we illustrated, our experimental setup combines optical tweezers, wide-field fluorescence microscopy and a microfluidics flow system. The basics of these techniques
2.2 Preparation of reagents

To attach single DNA molecules to optically trapped polystyrene beads, the DNA need to be modified. Also, since DNA is invisible, fluorescent dyes are widely used to stain and visualize individual DNA molecules. Moreover, if the proteins are not naturally fluorescent, the proteins of interest need to be labeled with specific fluorescent probes for fluorescence detection. This section will describe the preparation of the DNA, the dye used for staining and the protein fluorescent labeling.

2.2.1 Terminally-labeled DNA molecules

The DNA molecules used in our experiments are bacteriophage lambda DNA (Roche), which are linear DNA (48,502 base pair long) with two complementary 12-nucleotides single-strand 5’ overhangs. In our experiments, streptavidin-coated polystyrene beads are used for trapping. Because biotin binds to streptavidin with an extremely high affinity, fast on-rate, and high specificity, the DNA constructs for trapping experiments are modified by attaching multiple biotins to both DNA ends. For double-strand DNA experiment, normal nucleotides (dGTP and dTTP) and biotinylated nucleotides (biotin-11-dATP and biotin-16-dCTP) are incorporated into the 3’ ends of lambda DNA using Klenow DNA polymerase exo−, as shown in Figure 2.9 A. Besides the dsDNA construct, we can also prepare DNA constructs for single-strand DNA experiments using the force-induced DNA melting approach [67]. To this end, biotins should be attached to the 3’ and 5’ end of the same DNA strand. Here, we use different strategies to prepare both long ssDNA constructs and short ssDNA constructs, as shown in Figure 2.9 B and C. For long ssDNA constructs, firstly, one biotinylated oligonucleotide (5’-agg tcg ccg ccc ttt ttt tTt TtT-3’, where uppercase bases are those modified with biotin) and one unlabeled oligonucleotide (5’-ggg cgg cga cct gga caa-3’) are annealed and ligated to the lambda DNA to create the 5’-labeled end. Secondly, another biotinylated oligonucleotide (5’-TtT tTt ttt aga gta ctg tac gat cta gca tca atc ttg tcc-3’) is annealed and ligated to the 3’ end of the lambda DNA to produce the 3’-labeled end. Finally, long ssDNA molecules (48,517 nt, contour length 16.5 µm) are produced within flow cell by force-induced DNA melting using optical tweezers. This long ssDNA construct has been used for the study of RAD52 and ssDNA interactions in Chapter 3.
2.2 Preparation of reagents

Besides the long ssDNA constructs, we can also prepare short ssDNA constructs. In this case, first, dGTP, dTTP, biotin-11-dATP and biotin-16-dCTP are incorporated into the 3’ ends of lambda DNA using Klenow DNA polymerase exo−. Then the restriction enzyme ApaI is used to cut the DNA at the GGGCC∧C site, generating a long DNA segment and a short one. Subsequently, one biotinylated oligonucleotide (5’-cTcTcTcTcTcTcTcTcTc-3’) is annealed and ligated to the 5’ end of DNA. This way, one long DNA construct (contour length 13 µm) containing 38,435 bp with a 22 nucleotides overhang at 5’-end DNA and one short DNA construct (contour length 3.5 µm) of 10,109 bp with a 22 nucleotides overhang at 5’-end are generated. Finally, ssDNA can be produced in flow cell by force-induced melting.

2.2.2 Fluorescent dyes for DNA molecules staining

Fluorescent intercalating dyes, such as YOYO-1, Sytox Orange and TOTO-1, are often used to directly visualize dsDNA. These dyes exhibit a large enhancement of fluorescence upon intercalation [68] and their background fluorescence is extremely low, which make them ideal probes to visualize dsDNA in single-molecule experiments [69]. These dyes can also be used to monitor some enzymatic reactions on DNA. In fact, they report on the length of dsDNA and thus can be used to infer whether dsDNA is processed by enzyme (for instance opened) [70, 71]. Nevertheless, the intercalating dyes can change the structural properties of dsDNA, making it longer (increasing its contour length), while the bending rigidity (the persistence length) of DNA remains almost unaltered [72, 73]. Some studies [74] showed that the DNA-binding affinity of these dyes is mainly governed by a tension-dependent dissociation rate. The rate can be tuned over a range of seven orders of magnitude by changing DNA tension, intercalating species and ionic strength. Of note, if the intensity of the excitation laser and the dye concentration are very high, intercalating dye can cause DNA breaks easily.

2.2.3 Fluorescent labeling of proteins

Ideal fluorophores used to label proteins should be chemically stable, have minimal interference on the folding and biological functions of the protein, have high extinction coefficient and quantum yield, and also have sufficient photostability to be imaged for the duration of the experiment [75]. In multiple color experiments, the fluorophores should have minimal crosstalk in their excitation and emission spectra. In general, there are two main strategies to label proteins [76]. The first strategy relies on synthetic
2.2 Preparation of reagents

Figure 2.9 – DNA constructs preparation for trapping experiments. (A) dsDNA labeling procedure. Unmodified nucleotides (dGTP and dTTP) and biotinylated nucleotides (biotin-11-dATP and biotin-16-dCTP) are incorporated into the 3' ends of lambda DNA using Klenow DNA polymerase exo−. (B) Long ssDNA labeling procedure. One biotinylated oligonucleotide (red) and one unlabeled oligonucleotide (grey) are annealed and ligated to the lambda DNA to generate a 5'-labeled end. Then another biotinylated oligonucleotide (green) is annealed and ligated to create the 3'-labeled end. Finally, long ssDNA molecules are produced by force-induced DNA melting. (C) Short ssDNA labeling procedure. After polymerization (the same procedure as dsDNA labeling), the DNA is digested with the restriction enzyme ApaI. Then, one biotinylated oligonucleotide (red) is annealed and ligated to the 5' end of DNA. Finally, two labeled DNA segments are generated and ssDNA can be produced by force-induced DNA melting.

Fluorescent organic compounds (such as ATTO dyes, Cy dyes, and Alexa fluors) and fluorescent labeling of proteins is accomplished using a reactive derivative of the organic dyes that covalently and selectively binds to a functional group contained in the protein of interest. The excitation and emission spectra of the organic dyes cover the visible spectrum and extend into the infrared. The second strategy uses fluorescent proteins,
such as jellyfish green fluorescent protein (GFP), GFP variants (eGFP (enhanced-GFP), YFP (yellow fluorescent protein), and RFP (red fluorescent protein), to label the protein of interest via the generation of chimeric fusion protein. This strategy ensures that exactly one fluorescent protein is attached to one target protein, which can be very convenient to quantify the stoichiometry and oligomerization state of the protein of interest [76]. Nevertheless, fluorescent organic compounds and fused fluorescent proteins may alter the conformation of the target protein and potentially interfere with its functions [77]. It is therefore necessary to carefully verify the functionality of the fluorescently-labeled proteins.

In this thesis, one intercalating dye (Sytox Orange), two fluorescent organic compounds (Alexa Fluor 555 and Atto 647N) and two fluorescent proteins (eGFP and mCherry) have been used for experiments.

2.3 Experimental procedures for single molecule experiments

Using our experimental setup, we can not only manipulate individual DNA molecules but also visualize them in real time. In this section, the experimental procedures for catching single DNA molecules, stretching them, as well as visualizing single DNA-protein complexes will be introduced.

2.3.1 Single DNA molecule visualization by Sytox Orange staining

At the beginning of experiments, the different components (polystyrene beads, DNA, buffer, and dye solution (Sytox Orange)) are prepared and placed into different reservoirs. Flow pressure will then flush these components into separate laminar flow channels, as indicated in Figure 2.10. The experiment is conducted as follows:

(1) Two trapping beams are used to trap two beads in the beads channel. Normally, beads are diluted 1:200 in stock buffer 100 mM NaCl, 20 mM Tris HCl pH 7.5 and 20 mM EDTA. The size and shape of the trapped beads are visually checked to avoid the situation where more than one bead is trapped simultaneously in one trap or that contaminations are attached to the bead.

(2) The trapped beads are then moved into the DNA channel where DNA molecules are moving following the flow direction. The DNA is diluted in the same buffer used for beads. The dilution rate depends on the concentration of DNA constructs. We usually
2.3 Experimental procedures for single molecule experiments

**Figure 2.10** – Schematics of the laminar flow cell showing the procedure for single DNA molecule staining by an intercalating dye. Four channels of 5-channel flow cell are used: beads channel, DNA channel, buffer channel and dye solution channel. The arrows indicate flow direction. (1) Two beads are optically trapped in the beads channel. (2) The trapped beads are moved into the DNA channel where one single DNA molecule is caught between the beads. (3) In buffer channel, the DNA molecule is manipulated and stretched. (4) The DNA molecule is stained by the dye and visualized directly in the dye channel.

use 1:1000 in stock buffer. To catching DNA, the left bead is kept fix and the right bead is moved back and forth around the expected position of the end of flow-stretched DNA. When the left bead starts moving following the movements of the right one, it means that at least one DNA molecule has been caught between the two beads.

(3) After catching DNA, the two beads are moved into the channel containing only buffer (25 mM KCl, 20 mM Tris HCl pH 7.5). Here, in the absence of flow, the DNA force-extension curve is measured to ensure that only one DNA molecule is tethered between the two beads. In fact, the force-extension curves of a single DNA molecule and of multiple DNA molecules are different. Figure 2.11 A shows the typical force-extension curve of one DNA molecule with the plateau at about 65 pN. If there are more than one molecule, the plateau will occur at much higher forces.

(4) To reduce spatial fluctuations, the DNA is stretched at 5 pN tension and then moved into the dye channel containing 50 nM Sytox Orange in 100 mM NaCl, 20 mM Tris HCl pH 7.5 and 20 mM EDTA. When the DNA is illuminated by the 532 nm excitation laser, the fluorescence image of Sytox Orange-stained DNA can be observed (Figure 2.11 B). The bright fluorescent signal of the dye upon intercalation and the
low background of free dye allow the visualization of DNA directly in the dye solution channel.

Figure 2.11 – (A) Force-extension curve (black) and force-relaxation curve (red) of a single lambda DNA molecule. The blue curve is the eWLC fit curve from which the contour length (16.2 µm) and persistence length (43 nm) are obtained. (B) Fluorescent image of a single DNA molecule stained with Sytox Orange. The image is taken directly in the Sytox Orange channel. The two bright balls are the trapped beads, and one DNA intercalated by Sytox Orange is tethered between them. Scale bar: 2 µm.

2.3.2 Visualization of the interaction of RAD52 with ssDNA using dual-trap optical tweezers

In this section, we will describe the procedure to obtain single ssDNA molecules and visualize ssDNA-RAD52 complexes (Figure 2.12):

1. Two beads are trapped using two trapping beams in the beads channel.
2. The two beads are moved into the DNA channel and a single DNA molecule is caught between them.
3. After DNA capture, the beads are moved into a channel containing only a melting buffer (20 mM Tris HCl pH 7.5). The melting buffer has no salt to facilitate DNA melting [67]. Force-induced melting is employed to produce ssDNA: the DNA between two beads is stretched above 70 pN (normally at 90 pN) for at least 20 s. The DNA will melt due to the high tension acting on it. Using this force-induced melting method, single ssDNA can be produced within the flow cell and the outcome can be monitored by recording DNA force-extension curves (Figure 2.13 A).
2.3 Experimental procedures for single molecule experiments

Figure 2.12 – Schematics of the laminar flow cell showing the procedure for single ssDNA-protein complex visualization. Five channels are used: beads channel, DNA channel, melting buffer channel, buffer channel and protein channel. Arrows indicate flow direction. (1) Two beads are optically trapped in the beads channel. (2) The trapped beads are moved into the DNA channel where one single DNA molecule is caught between the beads. (3) In the melting buffer channel, the ssDNA molecule is produced by force-induced melting. (4) The ssDNA molecule is incubated in the protein channel. (5) The ssDNA-protein complex is moved back to the buffer channel and imaged.

(4) ssDNA is held with a tension of 10 pN and moved into the protein channel containing 500 pM of human eGFP-RAD52 in 50 mM KCl, 20 mM Tris HCl pH 7.5, 1 mM DTT, 1 mM MgCl$_2$, 0.5mM EDTA. Here, ssDNA is incubated in a controllable incubation time according to different experimental requirements. If we want to have only few individual protein patches on DNA, the incubation time is very short (e.g. several seconds). If we want to have a DNA highly coated with proteins, the incubation time is on the order of minutes.

(5) After incubation in the protein channel, single ssDNA-RAD52 complexes are moved into the buffer channel with no free protein. Fluorescent image with low background can be obtained under 488 nm laser excitation, as shown in Figure 2.13 B, which demonstrates the advantage of using a laminar flow cell for rapid exchange of buffer conditions and for drastically reducing the fluorescence background. At this
stage, ssDNA-protein interactions can be observed and studied in real time. The detailed study of DNA and RAD52 interactions will be presented in Chapter 3.

Figure 2.13 – (A) ssDNA production by force-induced melting. Initially, the force-extension curve of dsDNA is obtained when the DNA is stretched. The DNA will melt when the tension acting on it is kept above 70 pN. After about 20 seconds, when DNA is relaxed, ssDNA curve (red) will be obtained. The red curve is the proof of ssDNA generation. The magenta and blue curves are eWLC fits to dsDNA and ssDNA from which the contour lengths (12.9 µm and 20.1 µm, respectively) are obtained. Here, short ssDNA construct (contour length 13 µm) is used (described in section 2.2.1). (B) Fluorescent image of a single ssDNA-eGFP-RAD52 complex. In this example, only one protein patch is bound to ssDNA. The incubation time was 60 s and the fluorescence image was taken in buffer channel with no free protein. Scale bar: 2 µm.

2.3.3 Two DNA molecules manipulation and visualization using quadruple-trap optical tweezers

As described in section 2.1.1, our optical tweezers system can generate multiple independent traps controlled in 3 dimensions. Figure 2.14 illustrates the procedure used for manipulating and visualizing two DNA molecules by employing four optical traps (quadruple-trap):

1. Four beads are trapped by four traps.
2. Trapped beads are moved into the DNA channel and two DNA molecules are caught.
(3) Both DNA molecules are inspected by force-extension to ensure that only one molecule is tethered between each bead pair. As an example, Figure 2.15 A shows two DNA molecules stained with Sytox Orange. Then they are kept at 5 pN tension and incubated in the protein channel.

(4) After incubation in the protein channel, DNA-protein complexes are moved into the buffer channel. The two DNA molecules can be independently manipulated by controlling the corresponding beads. For example, one DNA molecule can be wrapped around the other one by moving one trap in 3 dimensions. Figure 2.15 B shows a fluorescence image of two MRN-coated DNA molecules wrapped together (one DNA tethered between the top beads and the other held by the bottom beads). The detailed study of DNA-MRN interactions will be presented in Chapter 5. This clearly illustrates the capability of our experimental system to manipulate and visualize multiple DNA molecules and study their interactions with proteins at single molecule level.
2.3 Experimental procedures for single molecule experiments

Figure 2.15 – (A) Fluorescent image of two DNA molecules stained with Sytox Orange (50 nM Sytox Orange). (B) Fluorescent image of two MRN-coated DNA molecules wrapped together. One DNA molecule is tethered between the two top beads and the other one is caught by the two bottom beads. The image was taken in the buffer channel without free MRN protein after incubating the DNA for 5 minutes in the protein channel containing 5 nM GFP-MRN in 20 mM Hepes pH 7.3, 100 µl/ml acetylated BSA, 1 mM DTT, 0.25 mM MgCl$_2$, 1 mM ATP. Scale bar: 2 µm.
Chapter 3

Human RAD52 Captures and Holds DNA Strands, Increases DNA Flexibility and Prevents Melting of Duplex DNA: Implications for DNA Recombination

Human RAD52 promotes annealing of complementary single-stranded DNA (ssDNA). In-depth knowledge of RAD52-DNA interaction is required to understand how its activity is integrated in DNA repair processes. Here, we visualize individual fluorescent RAD52 complexes interacting with single DNA molecules. The interaction with ssDNA is rapid, static and tight, where ssDNA appears to wrap around RAD52 complexes that promote intra-molecular bridging. With double-stranded DNA (dsDNA), interaction is slower, weaker and often diffusive. Interestingly, force-spectroscopy experiments show that RAD52 alters the mechanics dsDNA by enhancing DNA flexibility and increasing DNA contour length, suggesting intercalation. RAD52 binding changes the nature of the overstretching transition of dsDNA and prevents DNA melting, which is advantageous for strand clamping during or after annealing. DNA-bound RAD52 is efficient at capturing ssDNA in trans. Together, these effects may help key steps in DNA repair such as second-end capture during homologous recombination or strand annealing during RAD51-independent recombination reactions.
3.1 Introduction

Human RAD52 belongs to a ubiquitous class of proteins that helps to overcome the thermodynamic barrier required to anneal complementary DNA strands under biological conditions [78, 79]. In mammalian cells, RAD52 is important for repair of DNA double-strand breaks (DSBs) by the mutagenic RAD51-independent single-strand annealing pathway (SSA) [80, 81, 82]. However, unlike its yeast ortholog, vertebrate RAD52 does not seem to be crucial for recombinational repair of DSBs via RAD51 [83, 84]. The reason for this is that the RAD52 function to facilitate the loading of RAD51 on RPA-coated ssDNA appears to have been taken over by breast cancer susceptibility protein 2 (BRCA2) (see commentary in reference [85] and references therein). However, the finding that RAD52 deficiencies are synthetically lethal with BRCA2 deficiencies suggests a functional redundancy between RAD52 and BRCA2 [86, 87]. These reports, together with the recent discovery that RAD52 is implicated in promoting DNA synthesis after replication stress [88, 89, 90] and in modulating antibody class-switch recombination [91], are fueling a regain of interest in studying the function of RAD52, for both fundamental and therapeutic purposes [92].

RAD52 forms ring-shaped structures in vitro [93]. In solution, full-length RAD52 forms stable heptameric rings with a large central channel, a structural organization reminiscent of hexameric DNA helicases [94]. However, unlike hexameric DNA helicases, there is no evidence indicating that DNA passes through the channel of the RAD52 ring. In contrast, it has been proposed that ssDNA wraps around the outer surface of the RAD52 ring, interacting with an exposed positively charged groove [95]. Annealing of complementary ssDNA might then involve ssDNA wrapping and dynamic interactions between multiple RAD52 rings [96, 97].

Here, we directly visualize and quantify the interaction of fluorescently labeled human RAD52 with individual ssDNA and dsDNA molecules using a single-molecule approach that combines optical trapping with microfluidics and fluorescence microscopy. We report intrinsic properties of RAD52-DNA interactions including binding stoichiometry, diffusivity and effect on DNA mechanics and discuss the implications of our findings on the biological roles of RAD52.
3.2 Results and discussion

3.2.1 Experimental approach

After expression in bacteria, human RAD52 tagged at the N-terminus with enhanced green fluorescent protein (GFP-RAD52) was purified (Figure 3.1 A) and tested in strand-annealing kinetic assays (details of experimental procedures can be found in Supplemental Experimental Procedures). We found no appreciable differences in activity compared to the untagged variant (Figure 3.1 B and C), in agreement with previous studies where GFP-RAD52 fully rescued the synthetic lethality of BRCA2 RAD52 double deficient cells [86, 87].

To study the interaction of GFP-RAD52 with individual DNA molecules we used a combination of optical trapping, fluorescence microscopy and microfluidics [98, 99, 53]. Using two independent optical traps, individual DNA molecules could be manipulated while simultaneously detecting the tension on the DNA. dsDNA molecules were biotinylated on the 3’ ends of both top and bottom DNA strands [67] and were tethered to optically trapped streptavidin-coated polystyrene microspheres (Figure 3.2

Figure 3.1 – Functionality of GFP-RAD52. (A) SDS-PAGE and Coomassie staining analysis of purified wild-type RAD52 and N-terminally tagged GFP-RAD52. Lane 1: molecular size marker. Lanes 2 and 3: 0.5 µg of RAD52 and GFP-RAD52, respectively. (B) Kinetics of single-strand annealing reactions mediated by RAD52, eGFP-RAD52 or in the absence of protein using a FAM-labeled 50-mer oligonucleotide and its unlabeled complementary strand as described in Experimental Procedures. Progressive annealing was detected by taking samples at different time points. After fractionation of the samples by native PAGE, fluorescence signal was collected using a CCD camera. (C) Quantification of annealing activities detected as in (B). The extent of annealing is assessed as the percentage of dsDNA formed over time for RAD52 (black dataset), GFP-RAD52 (red dataset), and spontaneous annealing in the absence of protein (grey dataset). Error bars: standard error of the mean (SEM) from four independent experiments.
A). Single ssDNA tethers (Figure 3.2 C) were generated by biotinylation of both the 5’ and 3’ ends of the top strand of a dsDNA molecule and subsequent detachment of the bottom strand by force-induced melting [67]. After incubation of the dsDNA and ssDNA constructs in a GFP-RAD52 solution, the DNA tethers were brought into a buffer channel where DNA-bound proteins could be visualized in the absence of fluid flow and fluorescent proteins in solution (Figure 3.2 B and 3.2 D).

**Figure 3.2 – Experimental assay.** (A) Schematic of a dsDNA molecule (purple) tethered between two optically trapped micrometer-sized polystyrene beads (grey) with GFP-RAD52 complexes (green) bound to the dsDNA molecule. By manipulating the position of the beads the extension of the DNA molecule can be controlled while the tension in the molecule is monitored. At the same time, the proteins can be directly visualized with single-fluorophore resolution using widefield fluorescence microscopy. (B) Typical fluorescence signal image emitted by GFP-RAD52 complexes bound to dsDNA such as schematically shown in (A). Scale bar: 2 µm. (C) Schematic, such as shown in (A), but for GFP-RAD52 complexes bound to ssDNA complex. (D) Typical fluorescence signal image emitted by GFP-RAD52 complexes bound to ssDNA such as shown schematically in (C). Scale bar: 2 µm.

### 3.2.2 GFP-RAD52 binding to ssDNA is avid and shortens ssDNA contour length

The affinity of RAD52 for ssDNA has been reported to be much higher than for dsDNA [100]. To quantitatively assess the affinity for ssDNA, single ssDNA molecules were (unless otherwise indicated) incubated at a tension of 10 pN in a buffer containing 1 nM of GFP-RAD52, 20 mM Tris HCl pH 7.6, 100 mM KCl and either 1 mM Ca\(^{2+}\), 1 mM Mg\(^{2+}\) or no divalent cations. Protein complexes formation on the DNA was assessed from the fluorescence intensity (under constant continuous excitation with 500 ms exposure time/frame) immediately after transfer of the construct to a protein-free
environment. We observed discrete fluorescent patches along the ssDNA constructs, with each patch corresponding to an oligomeric DNA-bound GFP-RAD52 complex. In the presence of Ca\textsuperscript{2+}, loading of GFP-RAD52 is remarkably fast: after 5 s incubation in 100 pM GFP-RAD52, significant amounts of fluorescent protein were bound to the ssDNA (Figure 3.3 A left panel). Under these conditions, the observed rate of patch formation was 35 ± 3 oligomers s\textsuperscript{-1} nM\textsuperscript{-1} (N = 43). In the presence of Mg\textsuperscript{2+}, a much lower rate of (5 ± 1)\times10\textsuperscript{-2} oligomers s\textsuperscript{-1} nM\textsuperscript{-1} (N = 11) was observed and loading of similar amounts of protein on ssDNA required a much longer incubation (100 s) at 10-fold higher GFP-RAD52 concentration (Figure 3.3 A middle panel). When no divalent cation was present, binding of GFP-RAD52 was even less efficient and the patch formation rate was (1.2 ± 0.5)\times10\textsuperscript{-2} oligomers s\textsuperscript{-1} nM\textsuperscript{-1} (N = 25) (Figure 3.3 A right panel). Interactions appeared independent of DNA sequence (Figure 3.7 A), most patches detected on ssDNA appeared static (see analysis of protein diffusion below), and within the (limited) observation time of our experiments (about 2 min) we did not observe dissociation of the GFP-RAD52 patches from the DNA, which implies a dissociation rate smaller than 0.008 s\textsuperscript{-1}. We thus show that GFP-RAD52 interacts efficiently and statically with ssDNA in a cation-dependent manner.
3.2 Results and discussion

Figure 3.3 – Binding of GFP-RAD52 to ssDNA. (A) Fluorescence images (top panels) and kymographs (bottom panels) of GFP-RAD52 on ssDNA in the presence of the indicated divalent cations. ssDNA molecules were held at 10 pN tension and incubated in a buffer containing 100 pM GFP-RAD52 for 5 s in the presence of CaCl$_2$ or 1 nM GFP-RAD52 for 100 s in the presence of MgCl$_2$ or in the absence of divalent cation. Scale bars: 2 µm and 5 s.

(continued)
Figure 3.3 (continued) – (B) Size distribution of GFP-RAD52 oligomers bound to ssDNA, measured in CaCl\(_2\) at 100 pM with an average of 11 ± 1 monomers (mean ± SEM, N = 238). Dashed blue lines indicate multiples of 7 monomers. Similar distributions were obtained for all the conditions tested. (C) Bar plot showing how the average patch size varies with varying divalent cations at 1 nM and 100 pM GFP-RAD52. At 100 pM, no binding was detected within our incubation times in the presence of Mg\(^{2+}\) or in the absence of divalent cation. Error bars: statistical errors in the number of counts. (D) Force-extension curves during successive extension and retraction of GFP-RAD52-ssDNA complexes formed by incubation of the ssDNA with 5 nM GFP-RAD52 in 30 mM KCl. Red trace shows the corresponding (calculated) contour length during the extension trace. (E) Top panel: kymograph of the fluorescence signal corresponding to (D). Bottom panel: enlargement of the events indicated by (1), (2), and (3) showing clear ruptures of protein-protein bridges. Scale bars: 5 \(\mu\)m and 10 s.

Next, we determined the stoichiometry of the DNA-bound complexes by quantifying the fluorescence intensity of the complexes and normalizing to that of an individual GFP (Figure 3.8). Stoichiometry distributions were typically very broad, ranging from 1 to several tens of GFP-RAD52 monomers per complex (Figure 3.3 B). We found no evidence for the strict heptameric structure reported previously (blue dotted lines in Figure 3.3 B) [94]. Heptamers, the dominant species in solution, may thus rearrange into different oligomeric complexes when the protein interacts with ssDNA. Nevertheless, further experiments are needed to directly observe this putative rearrangement and assess to what extent the deviation from the 7 \(\cdot\) n distribution expected for heptamers and multiples of heptamers is caused by quenching of the eGFP fluorescence by homo-FRET or because of a dark, non-fluorescent fraction in the GFP-RAD52 preparations. In the presence of Ca\(^{2+}\), the average patch size was significantly larger than in the presence of Mg\(^{2+}\) or without divalent cations (Figure 3.3 C). Applying tension to the ssDNA substrate did not have a significant effect on the average number of patches (Figure 3.7 B) and their size (Figure 3.7 C). To further address how RAD52 interacts with ssDNA we performed an experiment where a ssDNA construct was repeatedly incubated in the GFP-RAD52 channel, each incubation lasting 2 s. The position of the fluorescent patches was recorded and photobleached afterward. Subsequently, the recorded positions of the successive incubations were compared (Figure 3.7 D, representative example out of 8 experiments). The probability of detecting a fluorescent patch at the same position during successive incubations is in the order of 14 ± 4 % (SEM) indicating that the rates of initial patch formation and of patch growth are
of the same order unlike RAD51, which exhibits rate-limiting nucleation and a fast polymerization rate [101].

We also studied the impact of GFP-RAD52 binding on the mechanical properties of ssDNA. To this end, we incubated ssDNA constructs in a buffer containing GFP-RAD52 at very low tension, allowing different segments of the ssDNA to interact with each other. Next, the constructs were brought into a protein-free buffer channel, where force-extension and force-relaxation curves were measured. In these curves, two clear effects were observed. In the extension curves (Figure 3.3 D), rupture events were observed, where a large, abrupt increase in DNA extension was observed without an increase in force (orange arrows). The average increase in contour length for a single rupture event was $0.35 \pm 0.02 \mu m$ (mean $\pm$ SEM, Figure 3.7 E). The relaxation curves, on the other hand, showed no such ruptures and appeared smooth but the ssDNA was significantly shorter than naked ssDNA. This shortening persists up to forces above 80 pN (Figure 3.3 D). These rupture events could be attributed to tension-induced rupture of protein-protein bridges, as can be observed in the corresponding fluorescence kymographs (Figure 3.3 E) or to the rupture of short stretches of dsDNA that were formed through RAD52-mediated annealing of partially complementary DNA segments. In addition, the shortening of the GFP-RAD52-ssDNA construct with respect to naked ssDNA can be attributed either to very strong (and thus unbroken) protein-protein bridges or to ssDNA that is wrapped around the protein complexes. Hence, this behavior is consistent with the proposal that ssDNA wrapping and ring-ring contacts might be involved in RAD52-promoted strand annealing [96, 97].

### 3.2.3 GFP-RAD52 binding increases flexibility and contour length of dsDNA and prevents melting

Although biochemical studies have mostly focused on the binding of RAD52 to ssDNA, there is also evidence for RAD52 interacting with dsDNA [102, 100], despite the biological relevance of such interaction remains subject of debate. To detect dsDNA binding we worked at 20 nM GFP-RAD52 concentration, as reference the RAD52 concentration in yeast is circa 1 nM [103]. Interestingly, RAD52 has a secondary DNA binding site that is important for dsDNA binding, likely regulated by phosphorylation, and required to introduce positive supercoiling in dsDNA upon RAD52 binding [104, 105]. Here, we used our single-molecule approach to directly observe the binding of GFP-RAD52 to dsDNA. As for the interaction with ssDNA, binding was dependent on the divalent cation present: GFP-RAD52 binds to dsDNA more readily in the
presence of Ca$^{2+}$ ((average patch formation rate $(26 \pm 4) \times 10^{-3}$ oligomers s$^{-1}$ nM$^{-1}$ (N = 66)), than in the presence of Mg$^{2+}$ ((3.9 $\pm$ 0.3) $\times$ 10$^{-3}$ oligomers s$^{-1}$ nM$^{-1}$ (N = 101)) or in the absence of divalent cations ((2.9 $\pm$ 0.5) $\times$ 10$^{-3}$ oligomers s$^{-1}$ nM$^{-1}$ (N = 38)) (Figure 3.4 A). From these numbers, it is clear that in the presence of divalent ions the affinity for dsDNA is 1 to 3 orders of magnitude lower compared to the ssDNA experiments. Again, the interaction of GFP-RAD52 with dsDNA appeared not to depend on DNA sequence (Figure 3.7 A) and dissociation was slower than photobleaching, as found for the interaction with ssDNA. Unlike GFP-RAD52 binding to ssDNA, which predominantly involved static complexes, binding to dsDNA involved both static and diffusive complexes (Figure 3.4 A), depending on the cation and applied tension. Also, when measured in Ca$^{2+}$ and at 50 pN, the average patch size on dsDNA did not depend on incubation time (Figure 3.9 A), indicating that patch growth is not cooperative.

Stoichiometry distributions of complexes bound to dsDNA were even broader than for ssDNA (Figure 3.4 B), and also for dsDNA we did not discern a clear 7 · n distribution reminiscent of heptameric complexes. Again, a dependence of the stoichiometry distributions on the divalent cation present was observed, as well as a significant dependence on the tension applied to the construct (Figure 3.4 C). At lower DNA tensions, GFP-RAD52 oligomers were larger than at higher tensions. This observation is in contrast to the interaction with ssDNA, which did not show dramatic tension dependence. Most likely, this is due to structural changes of bare dsDNA upon application of tension, causing disruption of base pairing resulting in force-induced melting or other structural states that are more ssDNA-like where RAD52 would preferentially bind with a stoichiometry reminiscent of that observed for ssDNA [106]. For ssDNA, such structural transitions do not occur. On both templates, the size of these complexes depends on the presence of Ca$^{2+}$ and Mg$^{2+}$. While the difference in the presence or absence of divalent cation can be explained by the shielding effect of their positive charges on the negatively charged phosphodiester backbone of DNA, the difference between Ca$^{2+}$ and Mg$^{2+}$ cannot be directly explained. Nevertheless, it is interesting to note here that Ca$^{2+}$ greatly favors association kinetics of GFP-RAD52 with DNA. In light of previous studies that showed an important role of Ca$^{2+}$ in the control of homologous recombination in human by affecting the ATPase activity of RAD51 [107, 108], our findings suggest that Ca$^{2+}$ could have a much wider impact on DNA recombination transactions in human cells not only by stimulating RAD51-mediated strand exchange but also RAD52-mediated strand annealing.
Figure 3.4 – Binding of GFP-RAD52 to dsDNA. (A) Fluorescence images (top panels) and kymographs (bottom panels) of GFP-RAD52 on dsDNA in the presence of the indicated divalent cations and dsDNA template tensions. The dsDNA molecule was incubated in a channel containing 20 nM GFP-RAD52. Incubation times were 20 s for CaCl$_2$ at 5 pN, 30 s for CaCl$_2$ at 50 pN, 50 s for MgCl$_2$ at both forces and for no divalent cation at 50 pN and 300 s for no divalent cation at 5 pN. As on ssDNA, most binding is observed in the presence of Ca$^{2+}$, slightly less binding is observed in presence of Mg$^{2+}$ and the lowest affinity is observed without divalent cation. GFP-RAD52 shows a dynamic behavior at low force, while it binds in a more static fashion at higher forces. Scale bars: 2 µm and 5 s.

(continued)
Figure 3.4 (continued) – (B) Size distribution (N = 242) of GFP-RAD52 oligomers bound to dsDNA, average of 34 ± 3 monomers (mean ± SEM), measured at 20 nM GFP-RAD52, 50 pN tension, in the presence of CaCl$_2$. Dashed blue lines indicate multiples of 7 monomers. Similar distributions were obtained for all the conditions tested. (C) Relation between average patch size and DNA tension for the cationic conditions studied. As dsDNA tension is increased the average patch size decreases by 4-fold in the 5-50 pN range. In addition, cationic conditions slightly influence the patch size. Error bars: SEM. (D) Mechanical properties of GFP-RAD52-dsDNA complexes determined by force-relaxation experiments (red curve). Up to 30 pN, the curve is well described by the eWLC model (dark grey). Compared to bare dsDNA (blue and light-grey), a significant decrease in persistence length and a slight increase in contour length are observed (see Table 3.1). The inset shows the fluorescence image of the dsDNA, covered by more than 4500 GFP-RAD52 proteins, recorded before the stretching cycle. Scale bar: 2 µm.

Next, we investigated the effect of GFP-RAD52-binding on the mechanical properties of dsDNA using dsDNA constructs with a high coverage of GFP-RAD52. We observe smooth force-extension and relaxation curves (Figure 3.4 D, Figure 3.9 B and C) that did not show evidence for protein-protein bridges or DNA wrapping, in contrast to our observation for ssDNA. Individual force-extension curves show substantial variations (Figure 3.9 B), likely due to DNA-bound protein complexes sticking nonspecifically to the microspheres. Relaxation curves, however, were reproducible and smooth (Figure 3.9 C). For forces below 30 pN, these curves can be accurately described by the extensible-Worm-Like-Chain (eWLC) model [109] (Figure 3.4 D). From the fit parameters in the presence and absence of GFP-RAD52 (Table 3.1), we deduce that GFP-RAD52 binding results in a (9 ± 1)-fold decrease of the persistence length and a (1.21 ± 0.03)-fold increase in contour length, while the stretch modulus is not affected. These properties suggest a possible binding mechanism of RAD52 to dsDNA through intercalation and opening of the double helix, which might be crucial for ATP-independent homology recognition and strand exchange by RAD52 [110, 111]. In the buffer conditions used, at forces above 30 pN, extension-relaxation cycles of bare dsDNA typically show a saw-tooth like overstretching transition with a large hysteresis between the extension and relaxation curve (Figure 3.9 D), signature of the force-induced melting of the DNA strands [112]. For dsDNA coated with GFP-RAD52 the behavior is different: the curves remain smooth and the hysteresis between extension and relaxation curves is much smaller, indicating that force-induced melting of the DNA strands no longer occurs. The formation of ssDNA is thus prevented by GFP-RAD52, providing evidence for strand annealing and clamping activity for...
### 3.2 Results and discussion

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Stretching direction</th>
<th>Number of experiments</th>
<th>$L_p$ (nm)</th>
<th>$L_c$ (µm)</th>
<th>$S$ (pN)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GFP-RAD52-dsDNA Relaxation</td>
<td></td>
<td>8</td>
<td>5.5 ± 0.9</td>
<td>19.4 ± 0.4</td>
<td>$(21 ± 7) \times 10^2$</td>
</tr>
<tr>
<td>Naked dsDNA Extension</td>
<td></td>
<td>20</td>
<td>45 ± 2</td>
<td>15.96 ± 0.02</td>
<td>$(18 ± 1) \times 10^2$</td>
</tr>
<tr>
<td>Naked dsDNA Relaxation</td>
<td></td>
<td>18</td>
<td>47 ± 2</td>
<td>16.00 ± 0.02</td>
<td>$(17 ± 1) \times 10^2$</td>
</tr>
</tbody>
</table>

$L_p = \text{persistence length}; L_c = \text{contour length}; S = \text{stretch modulus.}$ GFP-RAD52-dsDNA complexes and naked dsDNA were measured in a buffer containing CaCl$_2$. Errors: SEM.

**Table 3.1 – GFP-RAD52-dsDNA complexes as an extensible Worm-Like-Chain.**

RAD52 reminiscent of the activity proposed for the bacteriophage λ Red/β ortholog, which is thought to clamp DNA strands together to secure homology recognition [113]. RAD52 DNA-strand clamping might be an important property during second-end capture, for holding together annealed DNA repeats to allow processing of ssDNA flaps and DNA repair synthesis during the various types of homologous recombination after D-loop formation. Moreover, as was discovered in yeast, RAD52 could be part of a complex that tightly tethers the two ends of broken chromosomes allowing them to withstand the pulling forces of the mitotic spindle [114, 115].

#### 3.2.4 GFP-RAD52 slides along dsDNA

When examining the binding of GFP-RAD52 to dsDNA (Figure 3.4 A) we observed that, depending on the buffer and the tension applied on the construct, a fraction of the fluorescent patches moved along the DNA in a diffusive manner. To quantify this diffusive motion, we used custom-written tracking software to determine the trajectory of each individual fluorescent patch over time and used mean squared displacement (MSD) analysis to determine the one-dimensional diffusion coefficient of each patch (Figure 3.10 A-D). Then, we applied a threshold of 583 nm$^2$/s (the minimal detectable diffusion coefficient in our experimental conditions, see Supplemental Experimental Procedures) to determine whether the complex was static or diffusive. For each tension and buffer condition studied, we generated the distribution of above-threshold diffusion coefficients (Figure 3.5 A) and used two parameters to quantify the diffusive behavior: the average diffusion coefficient (calculated based on only the diffusive particles) and the diffusive fraction (the fraction of particles that diffused). Diffusion was most prominent at forces below 15 pN, where almost 100% of the particles were mobile. Both the average diffusion coefficient (Figure 3.5 B) and the diffusive fraction (Figure
3.5 C) decreased with increasing tension. The transition between static and diffusive behavior is reversible: when the force is increased, particles switch from a diffusive to a static state, and when the force is decreased again, particles may switch back to the diffusive mode (Figure 3.5 D). Yet, no dependence of the diffusion coefficient on the nature of the divalent cation was observed (Figure 3.5 E), and the diffusive fraction also did not change accordingly (Figure 3.10 E). Finally, our data shows no strong correlation between the diffusion coefficient and the size of the fluorescent patch (Figure 3.5 F), indicating that diffusion is not limited by the drag force acting on the protein complex. A quantitative analysis revealed that RAD52 complexes also diffuse on ssDNA, albeit with smaller diffusion coefficients and diffusive fractions than on dsDNA and in a manner independent of the tension applied on the ssDNA construct (Figure 3.10 F and 3.10 G). In total, 25 ± 3 % of the complexes bound to ssDNA showed diffusive behavior and the average diffusion coefficient was roughly 50-fold smaller for particles diffusing on ssDNA than on dsDNA (Figure 3.5 E). From these data we conclude that GFP-RAD52 can interact with DNA in either static or diffusive binding modes. The diffusive binding mode is predominantly observed on dsDNA and is slowed down when tension is applied on the dsDNA, which favors immobilization of RAD52 complexes by intercalation in the double helix (Figure 3.5 G top panel). This diffusive binding mechanism suggests a role for RAD52 in a diffusive search mechanism for localizing DNA structural intermediates such as ssDNA-dsDNA interfaces.
3.2 Results and discussion

**Figure 3.5** – GFP-RAD52 can diffuse along DNA. (A) Histogram (N = 77) of the diffusion coefficients of the diffusive GFP-RAD52 complexes along dsDNA, measured in 20 nM GFP-RAD52 at 5 pN and in the presence of CaCl$_2$. (B) Relation between average diffusion coefficient and dsDNA tension. A clear 8-fold decrease is observed as the tension on the dsDNA molecule increases. Error bars: SEM. (C) The fraction of complexes that are mobile decreases with increasing tension. At low force, virtually all complexes show diffusive behavior, while at high tension, only a small fraction of complexes is mobile. Error bars: statistical errors in the number of counts. (D) Kymograph recorded during successive extension-relaxation cycles of a GFP-RAD52-dsDNA complex showing a clear force-dependence of the diffusion: at low force, most complexes diffuse. When the force is increased, complexes switch to a static binding mode. When the force is decreased, complexes may switch back to diffusive behavior. Intensity is scaled logarithmically. (E) Bar plot of average diffusion coefficient for different divalent cations, forces and DNA substrates. Diffusion is fastest in the presence of Ca$^{2+}$, slower in presence of Mg$^{2+}$, and slowest in the absence of divalent cations. Error bars: SEM.

(continued)
3.2 Results and discussion

Figure 3.5 (continued) – (F) Relation between diffusion coefficient and patch size, measured on dsDNA with a tension of 5 pN in the presence of Ca$^{2+}$. Grey dataset shows all individual data points, red dataset shows the average diffusion coefficient of 20 consecutive data points of increasing complex size. Error bars: SEM. (G) Schematic summarizing the interaction of RAD52 with DNA. GFP-RAD52 interacts with dsDNA in a diffusive mode at low tension (1-5 pN) (top). As the applied tension is increased, diffusion halts as GFP-RAD52 complexes intercalate the double helix. At high tension (> 50 pN), GFP-RAD52 tightly clamps the DNA strands. The process is reversible, as GFP-RAD52 complexes resume diffusion when the applied tension is brought back to 5 pN. GFP-RAD52 binding to ssDNA is rapid, stable and static, consisting of a combination of wrapping and bridging modes (bottom).

3.2.5 DNA bound GFP-RAD52 captures ssDNA in trans

Given the involvement of RAD52 in second-end capture during homologous recombination [116, 117, 118], we explored its ability to capture DNA in trans. GFP-RAD52 was first bound to dsDNA or to ssDNA and the constructs were subsequently exposed to a solution of 60-mer ssDNA oligonucleotides fluorescently end-labeled with Atto647N. GFP-RAD52 bound to dsDNA exhibits a remarkably efficient ability to capture the ssDNA oligonucleotide in trans (Figure 3.6 A). All out of 27 individual GFP-RAD52 fluorescent patches (green) observed on 6 independent dsDNA molecules held at 50 pN, captured at least one ssDNA oligo (red). Under the same conditions, no Atto647N signal was detected on constructs that were not incubated with GFP-RAD52 (Figure 3.6 B). Next, we performed similar experiments on ssDNA constructs held at 5 pN (Figure 3.6 C). Again, analysis of 6 independent molecules pre-incubated with GFP-RAD52 showed detection of Atto647N signal co-localizing with the GFP-RAD52 patches. Control experiments with ssDNA constructs not pre-incubated with GFP-RAD52 revealed that the oligonucleotides bind to naked ssDNA to a much lesser extent than in the presence of GFP-RAD52 (Figure 3.6 D). We conclude that DNA bound GFP-RAD52 is efficient at capturing ssDNA from solution, reminiscent of its role in second-end capture.
Figure 3.6 – GFP-RAD52 captures ssDNA in trans. (A) Two representative experiments showing ssDNA capture by GFP-RAD52 bound to dsDNA (N = 6). A dsDNA construct held at 50 pN was first incubated in a channel containing 50 nM GFP-RAD52 in the presence of Ca\(^{2+}\) for 30 s and then moved into a protein-free buffer channel to detect the binding positions of GFP-RAD52 (green). Next, the construct was incubated in a channel containing 10 nM Atto647N labeled ssDNA Oligo for 30 s, and subsequently moved to the observation channel to detect the binding positions of the oligo (red). Both signals were merged to detect where colocalization (yellow) has occurred. (B) Two representative control experiments as in (A) but without GFP-RAD52.

(continued)
Figure 3.6 (continued) – (C) Two representative experiments showing ssDNA capture by GFP-RAD52 bound to ssDNA (N = 6). A ssDNA construct held at 5 pN was first incubated in a channel containing 5 nM GFP-RAD52 in the presence of Ca$^{2+}$ for 30 s and then moved into a protein-free buffer channel to detect the binding positions of GFP-RAD52 (green). Next, the construct was incubated in a channel containing 10 nM Atto647N labeled ssDNA Oligo for 30 s, and subsequently moved to the observation channel to detect the binding of the oligo (red). Both signals were merged to detect where colocalization (yellow) has occurred. (D) Two representative control experiments as in (C) but without GFP-RAD52. Scale bars: 2 µm.

3.3 Conclusions and Perspectives

We have provided a quantitative assessment of the interaction of human RAD52 with DNA, suggesting properties important for its physiological roles as summarized in Figure 3.5 G. While interacting tightly with ssDNA through a combination of wrapping and bridging, RAD52 complexes bound to dsDNA profoundly affect dsDNA mechanics, and can diffuse in a tension-dependent way along dsDNA. The substantial decrease in persistence length and slight increase in contour length observed upon RAD52 binding indicate that RAD52 binding increases dsDNA flexibility probably by destabilizing and intercalating into duplex DNA. Our findings suggest that the way by which RAD52 promotes strand exchange in vitro is not by a strand-invasion mechanism like the RAD51 or RecA nucleoprotein filament, but rather results from the ability of RAD52 to change dsDNA structure, intercalating in the helix to make the bases available for pairing. Consistent with our model, it was previously observed that increasing the fractional A.T content of DNA increases the yields of in vitro strand exchange reactions by RAD52, likely because it would be easier for RAD52 to intercalate in A.T regions [110, 119]. Also, the overstretching behavior of dsDNA is profoundly altered in the presence of RAD52, suggesting that RAD52 prevents force-induced melting and thus providing evidence for strand-clamping activity. The methodology and findings reported here can now be used in future experiments to extend this analysis by studying how RAD52 interacts with RPA, a pivotal ssDNA binding protein. Indeed, RAD52 will need to deal with RPA-coated ssDNA in the physiological context and its direct physical and functional interaction with RPA appears to be essential for homologous recombination in yeast and mammalian cells, especially when long ssDNA substrates need to be processed [120, 121, 122]. Further, the activities of BRCA2 could be similarly analyzed and directly compared to the ones of RAD52 to explore possible
3.4 Supplementary

Protein purification

Plasmid pET28a-eGFP-polyHis-hRAD52 for the expression of N-terminally tagged GFP-RAD52 was constructed by modification of pET28a-polyHis-hRAD52 [124]. RAD52 and GFP-RAD52 expression was performed in Rosetta(DE3)pLysS cells (Novagen) grown in LB medium containing kanamycin at 25 µg/ml and chloramphenicol at 34 µg/ml. Two liters of the same medium was inoculated with 20 ml of a saturated overnight pre-culture and incubated at 37 °C with shaking. Expression was induced with 0.5 mM IPTG when the optical density at 600 nm reached 0.5. After overnight growth at 16 °C, the cells were harvested by centrifugation (4,500 × g, 15 min, 4 °C), resuspended with 15 ml PBS and stored at -20 °C. The cell paste was thawed and mixed with one volume of lysis buffer (40 mM Tris HCl, pH 7.5, 1600 mM NaCl, 4 mM 2-mercaptoethanol, 20 mM Imidazole, 10% glycerol) supplemented with 1 tablet of protease inhibitors without EDTA (Pierce), 1 mM PMSF, 10 mg lysozyme and 1 ml of 10% Triton X-100. After resuspension the lysate was treated by sonication to reduce the viscosity. The lysate was clarified by centrifugation (25,000 × g, 60 min, 4 °C). The supernatant was collected and passed through a 5 ml HisTrap FF Crude column (GE Healthcare) equilibrated with 20 mM Tris HCl, pH 7.5, 800 mM NaCl, 2 mM 2-mercaptoethanol, 10 mM Imidazole and 10% glycerol. The column was washed extensively with 20 column volumes of the same buffer and proteins were eluted with 25 ml of the same buffer but containing 300 mM NaCl. The eluate was dialyzed against 2 liters of 20 mM Tris HCl, pH 8.0, 50 mM KCl, 1 mM EDTA, 1 mM DTT, 10% glycerol, overnight at 4 °C and applied to a 5 ml Heparin column (GE Healthcare) equilibrated with the dialysis buffer, washed by raising the salt to 100 mM and eluted with 0.1-1 M KCl gradient. The peaks of RAD52 and GFP-RAD52 eluting around 300 mM KCl (Figure 3.1 A), were collected, divided in aliquots and flash frozen in liquid nitrogen before storage at -80 °C.

Annealing assay

Two complementary 50-mer oligonucleotides, oligo 1 (5’-TAA ATG CCA ATG CTG CTG ATA CGT ACT CGG ACT GAT TCG GAA CTG TAA CG-3’), and oligo 2
(5’-CGT TAC AGT TCC GAA TCA GTC CGA GTA CGT ATC AGC AGC ATT GGC ATT TA-3’) were purchased from Eurogentec. Oligo 1 was 5’-end labeled with FAM. FAM-labeled oligo 1 (final concentration 10 nM), complementary oligo 2 (10 nM), RAD52 protein (final concentration 100 nM) or GFP-RAD52 (final concentration 100 nM) were added in reaction buffer (final volume 90 µl) containing 0.25 mM DTT, 0.25 mM EDTA, 5 mM Tris HCl PH 8.0, 40 mM KCL, 2.5% Glycerol, 50 µg/mL BSA, 2.5 mM MgCl2, kept at room temperature. At the time indicated, 10 µl of reaction mixture were taken out and quenched by addition of 2 µl of stop buffer containing 5 µM unlabeled oligo 1, 5% SDS, 250 mM EDTA and 2.5 mg/ml proteinase K for 15 min at room temperature. DNA products were fractionated by native polyacrylamide gel electrophoresis using a Tris-Borate buffer system. Fluorescent signal (Figure 3.1 B) was detected and quantified using a ChemiDoc MP imaging system (BIO-RAD).

**Experimental conditions for optical trapping experiments**

Catching of the beads and the DNA were performed in PBS buffer, consisting of 10 mM phosphate and 150 mM sodium chloride at pH 7.3-7.5. DNA melting for generation of ssDNA templates was performed in 20 mM Tris HCl pH 7.6. Buffer conditions in the protein channel and imaging channel were 20 mM Tris HCl pH 7.6, 100 mM KCl and either 1 mM MgCl2, 1 mM CaCl2 or no divalent cation.

**Quantification of fluorescence intensities**

To obtain the stoichiometry of DNA-bound RAD52 complexes, we applied a step-fitting algorithm [125] to the photobleaching traces of all individual fluorescent patches and determined the average intensity of a single GFP from a Lorentzian fit to the histogram of step intensities (Figure 3.8). From this, the total number of GFP molecules in each patch could be inferred from the initial fluorescence intensity of each patch.

**Quantification of protein diffusion**

The diffusion of GFP-RAD52 complexes along dsDNA and ssDNA was quantified by tracking the proteins for a large number of frames (on average 29 ± 2 s). The corresponding diffusion coefficient is calculated using 1-dimensional mean square displacement (MSD) analysis [53]. Because the pixel size of camera is 130 nm, we estimate that the minimal diffusion we could observe would be if the particle moves a distance of 1 pixel during the average interaction time. Therefore, we estimate the
minimum detectable diffusion of RAD52 interactions using our system to be 583 nm$^2$/s. We thus define that a particle is static if its diffusion coefficient is lower than this value and that it is diffusive when its diffusion coefficient is higher than this value.

**Experiments of ssDNA oligonucleotide capture in *trans* by RAD52**

The sequence of the 60-mer fluorescent oligonucleotide was Atto647N-5’-ACA GCC AGA CCC GGA CGC TGA CGC TCG ACC GTG AAA TCA CGC TGC CAT CCT CCG GTA CCG-3’. Experiments were performed in a flowcell with six channels. DNA molecules with GFP-RAD52 bound were exposed for 30 s to a 10 nM solution of the oligonucleotide in 20 mM Tris pH 7.6, 100 mM KCl and 1 mM CaCl$_2$ and visualized in the same buffer in the absence of oligonucleotides in solution.
Figure 3.7 – GFP-RAD52 binding to ssDNA. (A) Relative position along the DNA of all detected GFP-RAD52 complexes on ssDNA (red) and on dsDNA (blue). Because the orientation of the DNA molecule between the two beads is unknown in the experiments, the distance to the closest bead was used as a measure for the position. Error bars indicate Poisson statistics. (B) Relation between the average number of patches and the force on ssDNA at the indicated GFP-RAD52 concentrations and given cationic conditions. (C) Relation between the average patch size and the force on ssDNA at the indicated GFP-RAD52 concentrations and given cationic conditions. In contrast to what was observed on dsDNA (Figure 3.4 (C)), GFP-RAD52 patch size on ssDNA is independent of force. Error bars indicate SEM. (D) For the ssDNA molecule shown in the figure, the probability of a new RAD52 patch to bind to a location of a previously bound one is \(7.5 \pm 3\%\) (SEM) on average over the 5 subsequent 2 s-long dipping. Globally, we observed 21 co-localizing patches over the 144 detected patches on 8 different ssDNA constructs with an overall average patch grow probability of about \(14 \pm 4\%\) (SEM). Scale bar: \(5 \mu\text{m}\). (E) Histogram of contour length changes observed in the rupture events for ssDNA. Data were generated by applying a step-fitting algorithm [125] to the traces such as the red curve shown in Figure 3.3 (D). On average, the contour length changes in step of \(0.35 \pm 0.02 \mu\text{m}\) (SEM).
Figure 3.8 – Calibration of intensity and bleaching time of individual GFP-RAD52 proteins. (A) Example of fluorescence intensity trace of an individual GFP-RAD52 complex as a function of time. Step values were obtained from such traces using a step fitting algorithm [125]. (B) Histogram of step intensities. Lorentzian fit was used to determine the most likely value giving 345 ± 10. (C) Histogram of photobleaching times from the dataset shown in (B). The bleaching time obtained from an exponential fit is 2.17 ± 0.03 s. Error bars are statistical errors in the number of counts.
Figure 3.9 – GFP-RAD52 binding to dsDNA, related to Figure 3.4. (A) Average patch size as a function of the incubation time, shown for the data measured on dsDNA at a GFP-RAD52 concentration of 20 nM, in the presence of CaCl$_2$ and with a DNA tension of 50 pN. For the other conditions, similar results were obtained. Data display no dependence of the patch size on the incubation time, showing no evidence for a mechanism of cooperative patch growth. Error bars indicate SEM. (B) Force-extension curves of 11 different GFP-RAD52-dsDNA complexes shows irregular force-extension behavior, caused by nonspecific sticking of DNA-bound RAD52 to the beads. This made it impossible to obtain reliable eWLC fits to this data. Nevertheless, the data do not show the typical plateau at around 65 pN indicating the melting of the duplex DNA. Instead the curves show a continuous raise of the force during extension. (C) Force-relaxation curves of 8 different GFP-RAD52-dsDNA complexes. These curves are more regular and reliable eWLC fits can be performed on them. Fit parameters are reported in Table 3.1. (D) Example of force-extension data recorded during a stretching-relaxation cycle in the absence of GFP-RAD52. Blue curve shows extension curve and red curve relaxation data. Up to 30 pN, the curves are well described by a fit to the eWLC model. Fits are shown in light grey (extension data) and dark grey (relaxation data). Fit parameters for extension and relaxation are reported in Table 3.1.
Figure 3.10 – Analysis of GFP-RAD52 diffusion, related to Figure 3.5. (A) Typical example of a kymograph of diffusive GFP-RAD52 particles. Data were recorded at 20 nM GFP-RAD52 in the presence of CaCl$_2$ on dsDNA with a tension of 5 pN. Scale bars: 2 µm, 10 s. (B) Typical example of a kymograph of static GFP-RAD52 particles. Data were recorded at 20 nM GFP-RAD52 in the presence of MgCl$_2$ on dsDNA with a tension of 50 pN. Scale bars: 2 µm, 10 s. (C) Displacement from the initial position for a mobile particle (red dataset, corresponding to the particle indicated by the red dot in (A)) and static particle (black dataset, corresponding to the particle indicated by the black dot in (B)). (D) Mean-squared displacement analysis of the particles in (C). Linear fits to the first 5 points of the MSD reveal a diffusion coefficient of (4.4×10$^4$) nm$^2$/s (red dataset) and 91 nm$^2$/s (black dataset). Error bars indicate SEM. (E) GFP-RAD52 diffusive fraction as a function of cationic conditions for dsDNA at different forces and for ssDNA. Error bars indicate Poisson statistics. (F) Effect of applied tension on ssDNA constructs on GFP-RAD52 complexes diffusion coefficients in CaCl$_2$ or MgCl$_2$. 
Chapter 4

DNA tethering by the human XRCC4-XLF and XRCC4-Ligase IV complexes

Classical Non-Homologous End Joining (C-NHEJ) is predominant DNA double-strand break repair mechanisms in human cells. During C-NHEJ, the core components have been identified, but it is still difficult to assess the specific roles of these components and the dynamic process of bridging and holding two DNA fragments. Traditional bulk assays have found that XRCC4 and XLF proteins can form a complex and bridge DNA. DNA Ligase IV (LigIV), existing as a complex with XRCC4, is involved in last step during C-NHEJ, ligating the broken ends together and stimulated by XLF. Here, by combining optical tweezers, fluorescence microscope and microfluidics, the interaction of XRCC4, XLF, XRCC4-XLF complex, XRCC4-LigIV complex with DNA are studied at single molecule level in real time. The bridging of two DNA molecules via XRCC4, XLF, XRCC4-XLF complex can be directly observed and corresponding bridging strength is measured. Our results suggest that XRCC4–XLF complexes form mobile sleeve-like structures around DNA that can connect the broken DNA very rapidly and hold them together. Moreover, our experiments illustrate that XRCC4-LigIV complex also has DNA bridging ability, in addition to its ligation activity. The DNA bridges formed by XRCC4-LigIV are immobile.
4.1 Introduction

Classical Non-Homologous End Joining (C-NHEJ) is predominant DNA repair mechanism for DNA double-strand breaks in human cells. It operates by direct religation of the two broken DNA ends, but may be error-prone due to degradation of broken ends [126]. Many proteins, including Ku70/80 [13], DNA-PKcs [18], XRCC4 [20], XLF [22] and LigIV [21], are involved in C-NHEJ. The general mechanism of C-NHEJ can be divided into individual and sequential steps. Two major steps have been identified: (i) the Ku heterodimer (Ku 70 and Ku 80 subunits) first recognizes DNA ends of the broken DNA molecules and protects them from degradation, followed by recruiting DNA-dependent protein kinase catalytic subunit (DNA-PKcs) which promotes end synapsis [126] and (ii) a complex of XRCC4-LigIV ligates the broken ends, stimulated by XLF [127]. Recent biochemical studies have shown that XRCC4 and XLF are also present at the early stage of C-NHEJ, performing a LigIV-independent function [128][129]. In particular, when XRCC4 is present in large excess over LigIV [130], it associates with XLF to form complexes promoting LigIV-independent broken end bridging [131][132].

XRCC4 and XLF are structurally related. Both protein can form homo-dimers characterized by N-terminal heads and alpha-helical tails [133, 134]. Crystal structures and biochemical analysis of XRCC4–XLF complexes revealed that they form alternating helical filaments capable of bridging DNA [58, 135, 136]. LigIV forms a complex with XRCC4 and catalyzes the final ligation step during C-NHEJ in a manner stimulated by XLF. How two DNA broken ends are reconnected together is the center question in C-NHEJ mechanism. Therefore, the study of XRCC4-XLF and XRCC4-LigIV interacting with two DNA fragments is particular intriguing.

Here, we use a single-molecule approach combining optical tweezers, fluorescence microscopy and microfluidics to study the XRCC4, XLF and LigIV interactions with DNA at single molecule level in real time. Moreover, to study XRCC4-XLF and XRCC4-LigIV interactions with two DNA fragments, we extend our research approach by using quadruple-trap optical tweezers that allows us to independently manipulate two DNA molecules. Furthermore, we can precisely generate two DNA molecules with free ends, and thus directly observe XRCC4-XLF complexes and two DNA ends interactions. Our results suggest that XRCC4–XLF complexes form mobile sleeve-like structures around DNA that can reconnect the broken ends very rapidly and hold them together. Moreover, our experiments show that XRCC4-LigIV complex also can
4.2 Results and discussion

4.2.1 DNA interaction mode of XRCC4-XLF

When XRCC4 and XLF are mixed in solution, they can form XRCC4-XLF protein complexes [137]. In our experiments, we use a single cysteine mutant of XRCC4 labeled with the fluorophore Alexa 555 (XRCC4-Alexa555) and XLF fused with eGFP (eGFP-XLF) at 1:1 ratio. Figure 4.1 A and B shows the fluorescence colocalization image of DNA molecule in XRCC4-Alexa555-eGFP-XLF protein channel and the fluorescence colocalization image of DNA coated with XRCC4-Alexa555-eGFP-XLF in buffer channel, respectively. The image in buffer channel clearly shows several individual protein patches binding to DNA while the image in protein channel has high background. We notice that when a DNA molecule is moved to buffer channel from protein channel, the fluorescence signal on DNA decreased fast, indicating that XRCC4-XLF proteins dissociate rapidly and thus are bound transiently to the DNA.

In order to test whether XRCC4-XLF complexes affect the mechanical properties of DNA upon binding to it. DNA force-extension curves in absence and presence of XRCC4-XLF are recorded. Figure 4.1 C shows that the force-extension curve of DNA is not altered by the presence of XRCC4-XLF, indicating that the binding of XRCC4-XLF does not influence the elastic properties of DNA even at high density. If DNA can wrap around XRCC4-XLF complexes, DNA contour length should become shorter. But the contour length of DNA does not change upon XRCC4-XLF binding to it. The results suggest that it is the XRCC4-XLF complexes that wrap around DNA, not vice versa.

4.2.2 DNA tethering by the XRCC4-XLF complexes

DNA bridging via XRCC4-XLF in protein channel

Bulk assays have shown that XRCC4-XLF complexes play an important role during C-NHEJ by bridging and holding DNA fragments together [58, 135, 136]. To study this bridging activity in a well-controlled way, quadruple-trap optical tweezers is used to tether two DNA molecules simultaneously. We have developed three different DNA
4.2 Results and discussion

**Figure 4.1** – (A) Fluorescence colocalization image of DNA molecule in XRCC4-Alexa555-eGFP-XLF protein channel. Protein concentration: 50 nM XRCC4-Alexa555, 50 nM eGFP-XLF. Protein buffer: 25 mM KCl, 20 mM Tris HCl pH 7.5, 1 mM EDTA, 1 mM DTT. (B) Fluorescence colocalization image of DNA coated with XRCC4-Alexa555-eGFP-XLF in buffer channel. The buffer is the same as protein buffer. Incubation time in protein channel: 60 s. Green: eGFP-XLF. Red: XRCC4-Alexa555. Scale bars: 5 µm. (C) Force-extension curve (red) of DNA highly coated with XRCC4-XLF and force-extension curve (black) of naked DNA. The red curve is measured in protein channel containing 100 nM XRCC4-Alexa555 and 100 nM eGFP-XLF in 25 mM KCl, 20 mM Tris HCl pH 7.5, 1 mM EDTA, 1 mM DTT. When DNA is incubated in protein channel, it is completely relaxed and there is no force acting on it. The incubation time is 120 s.
Figure 4.2 – Three different methods of manipulating two DNA molecules and DNA bridging via XRCC4-XLF in protein channel. (A) Pairing manner. Top DNA molecule (black) is slightly stretched and touch the bottom one (red) in the presence of flow. The arrow shows the flow direction. (B) Crossing manner. One DNA molecule (black) is put on the top of the other one (red), making them cross each other. (C) Wrapping manner. The detailed process is shown in Figure 4.3. (A1), (B1), and (C1) are typical examples for each manipulation manner in XRCC4-XLF protein channel. Black dashed line indicates one DNA molecule tethered by two beads (orange circles). Red dashed line is another DNA molecule caught by the other two beads (blue circles). (A2), (B2) are the bridged DNA molecules in protein channel and related to (A1), (B1). (C2) Unwrapped DNA molecules in protein channel, showing DNA bridging. Protein concentration: 50 nM XRCC4-Alexa555 and 50 nM eGFP-XLF. Protein buffer: 25 mM KCl, 20 mM Tris HCl pH 7.5, 1 mM EDTA, 1 mM DTT. The images are excited with 488 nm laser. Scale bars: 5 μm.
4.2 Results and discussion

manipulation methods: pairing, crossing and wrapping (Figure 4.2) and can visualize DNA bridging via XRCC4-XLF in protein channel.

For the pairing manner, two DNA molecules can be slightly stretched and touch each other in the presence of a gentle flow, as indicated in Figure 4.2 A. Since DNA is invisible, this manner requires that fluorophores bind to DNA and make it visible. Figure 4.2 A1 is one typical example that shows pairing of two DNA molecules by XRCC4-XLF complexes visualized in protein channel. After 60s incubation, the two DNA molecules are bridged via XRCC4-XLF (Figure 4.2 A2).

In another configuration, we put one DNA molecule on the other, so that the two molecules cross each other, as shown in Figure 4.2 B. As an example, Figure 4.2 B1 presents two crossed DNA molecules in XRCC4-XLF protein channel. When we try to separate the two crossed molecules after 60s incubation, we find that the two DNA molecules are connected with each other due to XRCC4-XLF (Figure 4.2 B2).

The third way of DNA manipulation used in DNA bridging experiments is the wrapping manner (Figure 4.2 C). The detailed wrapping and unwrapping process is described in Figure 4.3. Two DNA molecules are caught by four beads (Figure 4.3 A). Then one molecule is wrapped around the other one by controlling one trapped bead (Figure 4.3 B). Finally, the wrapped construct is unwrapped by moving the same bead in the opposite direction (Figure 4.3 C). Figure 4.2 C1 shows the wrapped DNA molecules coated with XRCC4-XLF in protein channel. After unwrapping, the two DNA molecules are bridged to each other via XRCC4-XLF (Figure 4.2 C2).

Real time observation of DNA bridging via XRCC4-XLF in buffer channel

DNA bridging by XRCC4-XLF can be visualized in protein channel, but the background is too high. In order to observe individual XRCC4-XLF bridges, we conducted the DNA bridging in buffer channel. As shown in Figure 4.4, after 120 s incubation in protein channel, two DNA molecules are moved into buffer channel where several isolated XRCC4-XLF patches can be observed (Figure 4.4 A). The two complexes are wrapped and incubated 120 s in this configuration (Figure 4.4 B). Then the wrapped construct is unwrapped. Figure 4.4 C shows that the unwrapped DNA molecules are bridged by XRCC4-XLF and the proteins XRCC4 and XLF are located simultaneously at the bridged junction.

Sliding of the XRCC4-XLF tethers

Next, we investigate the dynamic behavior of the XRCC4-XLF junctions. To this end, we have tried to control one DNA moving along the other by manipulating two
trapped beads and monitor the movement of bridges. We find that the DNA junction bridged by XRCC4-XLF can slide back and forth smoothly along DNA, as shown in Figure 4.5. The movement distance of the bridged junction is almost the whole contour length of lambda DNA, which is more than 40,000 bp. At some points, this sliding is stalled, probably when the bridge runs into a stationary XLF and/or XRCC4 complex or a DNA end. Over 113 bridged experiments, there are 79 events showing sliding of bridged junction. The junction bridged by XRCC4 or XLF alone can also slide along DNA (Supplementary Figure 4.16). This mobility indicates that the bridges consist of one or a few of sleeves. A single sleeve could surround two DNA molecules, allowing them to slide independently within the sleeve. The property of two DNA fragments sliding to each other may be important during NHEJ mechanism to regulate broken ends.

**Strength of the XRCC4-XLF tethers**

Another interesting question is how strong the bridged junctions are. To study this, the position of four beads are rearranged for easily force measurement, as shown in Figure
4.2 Results and discussion

Figure 4.4 – DNA bridging via XRCC4-XLF in buffer channel using wrapping manner. (A) Fluorescence colocalization image of two DNA molecules coated with XRCC4-XLF in buffer channel. (B) Wrapped construct. (C) Unwrapped construct, showing DNA have been connected by XRCC4-XLF. At the bridged junction (yellow), XRCC4 and XLF are located simultaneously. In this experiment, DNA molecules are first incubated 120 s in protein channel containing 50 nM XRCC4-Alexa555 and 50 nM eGFP-XLF in 25 mM KCl, 20 mM Tris HCl pH 7.5, 1 mM EDTA, 1 mM DTT. Then they are moved into buffer channel containing the same buffer with protein buffer. Green: eGFP-XLF. Red: XRCC4-Alexa555. Scale bar: 5 µm.

4.6 B. The strength of bridged junction is measured by moving the right bead along x direction (Figure 4.6 C). In this configuration, the total force $F$ should equal the sum of $f$ and the projections of $f_1$ and $f_2$ on $x$ axis. If the angle $\theta$ between $f$ and $f_1$ is quite close to $90^\circ$, total force $F$ will approximately equal to $f$. In this case, the strength of bridged junction can be measured by calibrating the left optical trapped bead and moving the right one. Figure 4.6 D shows the bridged junction is broken due to stretch and therefore leading to the bridged two DNA molecules separation. The corresponding
4.2 Results and discussion

The strength of bridged junction is measured, which is about 10 pN, as shown in Figure 4.7 A. The averaged bridging strength is $30.3 \pm 17.6$ pN over 38 measurements whose strength distribution is presented in Figure 4.7 B. 12 over 38 experiments, two DNA molecules are still intact after force measurement. In other cases, DNA breaks due to stretch. Together, these experiments indicate that XRCC4–XLF complexes readily form stable DNA–DNA bridges.

![Image](image_url)

**Figure 4.5** – Sliding of XRCC4-XLF tethers. (A) Bridged molecules after unwrapping. The bridged junction is indicated by arrow. (B), (C), (D), (E) The junction can slide smoothly back and forth by controlling the movement of one DNA (black). Green: eGFP-XLF. Red: XRCC4-Alexa555. Scale bar: 5 µm.

### 4.2.3 Interaction of XRCC4-XLF with DNA ends and tethering

**Interaction of XRCC4-XLF with one free DNA end and tethering**

Our experiments using two end-occluded DNA molecules have demonstrated that two DNA can be stably bridged by XRCC4-XLF and bridged junction can slide along DNA. However, it raises the question of how XRCC4–XLF complexes interact with double-strand DNA broken ends. To this end, first, we study the interaction of XRCC4–XLF with a free DNA end. The end can be generated by manually stretching one DNA. Here, a 5-channel laminar flow cell is used (Beads channel, DNA channel, Sytox Orange channel, protein channel, and buffer channel). This experiment is performed in a slight buffer flow, such that, after generation of a free end, the end can be extended in the direction of flow to prevent the DNA from bridging onto itself. While for the force measurement, the flow is off. Figure 4.8 A shows that two DNA molecules are bridged by XRCC4-XLF in buffer channel. The bridged junction can slide along DNA (Figure 4.8 B, C, D, E, F, G). If one DNA is overstretched, it may break and generate a free
Figure 4.6 – Strength measurement of the XRCC4-XLF tethers. (A) The unwrapped DNA molecules are bridged by XRCC4-XLF. (B) The position of four beads are rearranged for force measurement. (C) The strength of bridged junction is measured by moving the right bead along x direction. The force analysis at this configuration is shown (orange arrows). (D) The bridged junction is broken and two DNA molecules are separated. Green: eGFP-XLF. Red: XRCC4-Alexa555. Scale bar: 5 µm.

Figure 4.7 – (A) The strength measurement of junction bridged by XRCC4-XLF related to Figure 4.6. (B) The bridging junction strength distribution of two ends-occluded DNA molecules over 38 measurements. The averaged bridging strength is 30.3±17.6 pN. The concentration of XRCC4 and XLF in protein channel is 50 nM. Incubation time in protein channel: 120 s. Incubation time in buffer channel: 120 s. Error: Standard deviation.
Figure 4.8 – DNA free end sliding mediated by XRCC4-XLF. (A) Two unwrapped DNA molecules are bridged via XRCC4-XLF. The image is fluorescence colocalization image of XRCC4 (red) and XLF (green). Both XRCC4 and XLF are presented at the bridged junction (yellow). (B), (C), (D), (E), (F), (G) show the sliding of the bridged junction. (H) One DNA (red line) is manually broken by stretching the left bottom bead. The free end is indicated by arrow. (I), (J), (K), (L), (M) show that the free end of DNA gradually slides to bridged junction by moving right bottom bead, as indicated with arrows. Images from (B) to (M) are eGFP-XLF illumination. Protein concentration in protein channel: 50 nM XRCC4-Alexa555 and 50 nM eGFP-XLF. Protein buffer: 25 mM KCl, 20 mM Tris HCl pH 7.5, 1 mM EDTA, 1 mM DTT. Incubation time in protein channel and in buffer channel: 120 s. Green: eGFP-XLF. Red: XRCC4-Alexa555. Scale bar: 5 µm.
end (Figure 4.8 H). Interestingly, this free end can slide smoothly to bridged junction (Figure 4.8 I, J, K, L) and it remains stably connected with the junction (Figure 4.8 M).

In order to determine the construct configuration at this situation, it is moved to a Sytox Orange channel. Figure 4.9 A is the construct stained with Sytox Orange. It clearly shows the DNA end is stably connected with the junction via XRCC4-XLF. The strength of bridged junction is also measured (Figure 4.9 B, C), which can be up to 80 pN (Figure 4.9 E). Note that the curve of force measurement is different from the one in Figure 4.7 A, because the construct is fully stained with Sytox Orange. We also find that Sytox Orange can hinder the diffusion of XRCC4-XLF and inhibit the sliding of bridged junction, which can be explained by intercalation of Sytox Orange into DNA double helix.

**Interaction of XRCC4-XLF with two blunt DNA ends and tethering**

Furthermore, we investigate XRCC4-XLF interaction with two blunt DNA ends. DNA molecule with blunt end can be precisely produced using restriction enzyme SfoI. The enzyme can recognize and cut lambda DNA at GGC^GCC site, generating a 45,679 bp DNA with a blunt end (Figure 4.10). The idea is to use SfoI to digest bridged DNA molecules and generate two DNA blunt ends for study. In this experiment, a 5-channel laminar flow cell is used (Beads channel, DNA channel, buffer channel, protein channel, SfoI enzyme channel). This experimental process is also performed in a gentle flow except for force measurement.

As before, we get two wrapped and bridged DNA molecules (Figure 4.11 A). Then it is moved into in the SfoI channel. The best working temperature for SfoI is at 37°C. But it is still quite efficient at room temperature. In short time, the two DNA molecules are digested (Figure 4.11 B and C). After SfoI treatment, it produces a hybrid DNA complex consisting of two fragments from different molecules. This hybrid construct is held by XRCC4-XLF (Figure 4.11 D). We stretch the construct (Figure 4.11 E) and find that it can be stretched up to 46 µm and the lengths of the two fragments are almost the same (Figure 4.11 F). The strength of the bridged junction is very high, which can up to 240 pN (Figure 4.11 G).

Another interesting result is presented in Figure 4.12. In this case, the bridged junction connects two DNA fragments, slides as a sleeve to DNA ends (Figure 4.12 D, E, F) and holds two ends for a while. Finally, one DNA end slides out of the bridge when the hybrid construct is continuously stretched. The bridged junction strength
4.2 Results and discussion

Figure 4.9 – DNA-XRCC4-XLF complex stained with Sytox Orange, related to Figure 4.8. (A) The DNA-XRCC4-XLF complex stained with Sytox Orange in Sytox Orange channel. The fluorescence colocalization image of XLF (green) and Sytox Orange (magenta) shows that DNA end is located at the bridged junction. (B) The position of the construct is changed for strength measurement (C). (D) One DNA (red line) end detaches from right bead due to stretch. It is still connected with another one (black) via the bridged junction. (E) The recorded stretching force. Bridged junction is indicated by the arrow. Sytox Orange channel contains 50 nM Sytox Orange, 25 mM KCl, 20 mM Tris HCl pH 7.5, 1 mM EDTA, 1 mM DTT. Green: eGFP-XLF, Magenta: Sytox Orange, Scale bar: 5 µm.

Figure 4.10 – (A) Restriction enzyme SfoI cutting position on lambda DNA. (B) A blunt DNA end produced by SfoI.

measured in this case is less than 5 pN. This behavior was observed in 3 out of 27 independent experiments.
The bridging strength distribution of two DNA molecules with free ends over 27 measurements is shown in Figure 4.13. The averaged bridging strength is 59.2±58.4 pN, which is higher than the one shown in Figure 4.7. It can be explained by two reasons. One is that force measuring method are different. For the force measurement of end-occluded DNA molecules, the real force should be higher than measured value according to the force analysis (Figure 4.6). Another reason is the protein concentration is different. The former one is 50 nM and the later one is 100 nM. During the same incubation time, the bridging strength may depend on the concentration of protein. More protein coated on DNA will result in higher bridging strength.

Taken together, these experiments suggest that after DNA double-strand breaks formation, XRCC4–XLF complexes are able to bridge and hold the broken DNA fragments together such that their ends remain close to each other, while the mobility of the bridged junction ensures that the ends can be accessible for further processing by other repair factors, e.g. LigIV.

4.2.4 DNA tethering by the XRCC4-LigIV complexes

During C-NHEJ, XRCC4, XLF and LigIV can form a complex and work together. At last, LigIV, existing as a complex with XRCC4, performs ligation of the broken ends. The ligation process is stimulated by XLF. We have demonstrated that XRCC4-XLF can bridge DNA and hold two broken ends together. It is intriguing to study the bridging capability of LigIV. Because LigIV always exists as a complex with XRCC4, we use XRCC4-LigIV complexes in our experiment.

A 4-channel laminar flow cell is used (Beads channel, DNA channel, buffer channel, and protein channel). Two DNA molecules are caught and moved into protein channel and incubated 120 s. Then DNA-XRCC4-LigIV complexes are moved into buffer channel and conducted by DNA wrapping. After 120 s incubation, the complexes are unwrapped. But in most cases we did not observe DNA bridging phenomenon (2 out of 9 events).

However, when the experimental strategy is modified, the bridging efficiency is greatly increased. After 120 s incubation in protein channel, two DNA-XRCC4-LigIV complexes are brought to the buffer channel and wrapped. Then the wrapped construct is put back to protein channel and incubated 120 s. Subsequently, the wrapped construct is unwrapped in buffer channel. By doing so, we observe DNA bridging via XRCC4-XLF in all 15 experiments. Figure 4.14 is one typical experiment that shows DNA bridging via XRCC4-LigIV. We find that in most cases, the bridged junction is immobile (Figure 4.14 E, F). Only 3 out of 15 measurements show small movement of
4.2 Results and discussion

Figure 4.11 – Interaction of XRCC4-XLF with two blunt DNA ends and strong tethering. (A) The fluorescence image of XLF shows two unwrapped DNA molecules are bridged. (B) One DNA molecule (red line) is cut by SfoI in SfoI channel (1 in 250 1X Cutsmart buffer). (C) The other DNA molecule (black line) is also cut by SfoI. The hybrid construct consists of two DNA fragments from two different DNA molecules. (D) The fluorescence colocalization image of XRCC4 (red) and XLF (green) shows that both XRCC4 and XLF are located at the bridged junction. (E) The position of the construct is rearranged. (F) The hybrid DNA complex is stretched for force measurement. The left DNA is stretched 22.8 µm and the right one is stretched to 23.2 µm. (G) Corresponding force-extension curve. Protein concentration: 50 nM XRCC4-Alexa555 and 50 nM eGFP-XLF. Protein buffer: 50 nM eGFP-XLF, 25 mM KCl, 20 mM Tris HCl pH 7.5, 1 mM EDTA, 1 mM DTT. Green: eGFP-XLF. Red: XRCC4-Alexa555. Scale bar: 5 µm.

We also find that some bridges are very strong. DNA molecules break one by one during force measurement (Figure 4.14 H, J). Figure 4.15 A is the strength of the bridged junction related to the construct in Figure 4.14 G. The averaged bridging strength is 63.8±42.1 pN (Figure 4.15 B). A small fraction broke at very high forces, which can be up to 180 pN. Together, these experiments show that XRCC4-LigaseIV complex can bridge DNA molecules, in addition of its ligation capability. Although XRCC4 alone can bridge DNA, the bridged junction by XRCC4 can slide smoothly in...
4.2 Results and discussion

**Figure 4.12** – Interaction of XRCC4-XLF with two blunt DNA ends and weak tethering. (A) The fluorescence colocalization image of XRCC4 (red) and XLF (green) shows two unwrapped DNA molecules are bridged. (B) The fluorescence image of XLF (green) shows one DNA molecule (red line) is cut by SfoI in SfoI channel. (C) The other DNA molecule (black line) is also cut by SfoI. (D) The hybrid construct consists of two DNA fragments from two DNA molecules. (E) and (F) Hybrid DNA complex is stretched. The left DNA is stretched 13.3 µm and the right one is stretched to 11.5 µm. (G) Two molecules are separated. Green: eGFP-XLF. Red: XRCC4-Alexa555. Scale bar: 5 µm.

most cases. However, in the presence of LigIV, the bridged junction is stalled, resulting in DNA fragments immobility.

Two DNA-protein complexes are wrapped in buffer channel after 120 s incubation (Figure 4.14 A, B). The wrapped complex is put back to protein channel and incubated 120 s again, allowing more protein to bind to DNA. Wrapped DNA-protein complex (Figure 4.14 C) is unwrapped in buffer channel (Figure 4.14 D). The sliding of bridged junction is not observed by moving top two beads, as indicated in Figure 4.14 E and F.
4.3 Conclusions and Perspectives

In this study, we have demonstrated that the combination of quadruple-trap optical tweezers, fluorescence microscopy and microfluidics is a powerful tool to study the bridging activity of XRCC4-XLF and XRCC4-Ligase IV complexes. Our experiments show that XRCC4-XLF complex can form strong and mobile DNA bridges capable of holding the DNA fragments. These complexes can bridge the DNA fragments at any location, preventing the fragments from moving apart. The mobility of the bridged junction could allow them to slide along the DNA to the broken ends accessible for further processing. Moreover, our experiments suggest that XRCC4-LigIV complex also has DNA bridging ability, and this bridging behavior is quite different from the bridging via XRCC4-XLF complex. The bridges formed by XRCC4-LigIV lead to DNA bridged fragments immobility. In future, more experiments should be conducted to address: 1) the bridging ability of Ligase IV; 2) how XRCC4, XLF, LigIV are regulated together; 3) how other factors, such as Ku and DNA-PKcs, come into play in C-NHEJ.

Figure 4.13 – The bridging junction strength distribution of two DNA molecules with free ends over 27 measurements. The averaged bridging strength is 59.2±58.4 pN. The concentration of XRCC4 and XLF in protein channel is 100 nM. Incubation time in protein channel: 120 s. Error: Standard deviation.
4.4 Supplementary

Protein labeling and purification

All detailed information of XRCC4 and XLF protein labeling and purification has been described in [98]. Recombinant proteins are expressed and purified from Escherichia coli cells. In XRCC4-XLF experiment, XRCC4 fluorescent labeling are performed with...
maleimide Alexa Fluor 555 at position 93 or position 218 and XLF fluorescent labeling is performed with eGFP. In XRCC4-LigIV experiment, XRCC4 fluorescent labeling is performed with fused mCherry and LigIV fluorescent labeling was performed with fused eGFP. For all protein samples used, protein activities are similar to wild-type activities. This has been confirmed by biochemical DNA bridging assays.

4.4.1 DNA bridging via XRCC4 or XLF alone

DNA molecules can also be bridged via XRCC4 (5 out of 5 tests) or XLF (4 out of 4 tests) alone. The bridged junctions can also slide along DNA, for XRCC4 (all 5 tests) and for XLF (2 out of 4 tests).

Figure 4.15 – (A) Strength measurement of bridges, related to the construct in Figure 4.14 (J). (B) The bridging junction strength distribution over 15 measurements. The averaged bridging strength is 63.8±42.1 pN. Error: Standard deviation.

Figure 4.16 – DNA molecules can be bridged by XRCC4 or XLF alone. (A) DNA molecules are bridged by XRCC4 by pairing manner. (B) DNA molecules are bridged by XRCC4 in a wrapping manner. (C) DNA molecules are bridged by XRCC4 in a pairing manner. Image was taken in protein channel. (D) DNA molecules are bridged by XRCC4 in a wrapping manner. Green: eGFP-XLF. Red: XRCC4-Alexa555. Scale bar: 5 µm.
Chapter 5

DNA tethering by the human MRE11/RAD50/NBS1 complex

In mammalian cells, the MRE11/RAD50/NBS1 (MRN) complex initiates and coordinates broken DNA repair and signaling, promoting homology-directed DNA double-strand break repair. Its interaction with DNA ends is critical for the recruitment of DNA-processing enzymes, activation of ATM protein kinase and broken ends tethering. The mechanism of DNA bridging promoted by the MRN complex, is still not very clear. Here we have directly visualized MRN binding to duplex DNA molecules at different buffer conditions using dual-trap optical tweezers and find that the affinity of MRN for DNA is strongly affected by ATP, non-hydrolysable ATP analog (AMP-PNP) and Mg$^{2+}$. Furthermore, quadruple-trap optical tweezers are used to study the bridging of two DNA molecules by the MRN complex. Our results demonstrate that the MRN bridging ability depends on protein-protein interactions, and that the MRN complex can slide onto DNA end and hold two ends together. Interestingly, we find that mutant MRN with the truncation of the hook and coiled coil domains of RAD50 can also bridge DNA molecules, suggesting that the hook and the coiled coil of RAD50, indispensable for DNA repair, are not required for DNA bridging.
5.1 Introduction

The human MRE11/RAD50/NBS1 (MRN) complex plays critical roles in DNA damage signaling [138], recruitment of DNA-processing enzymes [139, 140], activation of ATM protein kinase [140] and broken DNA end tethering [141]. The complex consists of a globular DNA-binding domain in which a MRE11 dimer and two NBS1 molecules associate with two RAD50 molecules. In this complex two 50 nm long coiled coils of RAD50 extend out and contain Zinc hook domains at their apices [142, 143, 144]. The coiled coil of RAD50 are flexible [145] and their apices can self-associate [144]. The flexibility of the coiled coil allows these apices to adopt an orientation favorable for interaction [146].

These structural analyses of the complex, as well as genetic data related to the roles of the complex in DSB repair, suggest that one of the biological function of the MRN complex on DSB repair is primarily attributed to its ability to bridge two DNA molecules in trans via the hook and coiled-coil domains of RAD50 [147, 146, 141, 145]. Atomic force microscopy (AFM) studies have revealed that the MRN complex exists as two conformations: the intracomplex and the intercomplex [146]. The intracomplex model proposes that the MRN complex functions in a ring-like structure in which dimerization at the hook and globular domain complete the dimeric assembly [59, 148, 149]. However, DNA binding by MRN globular domain can induce RAD50 opening, leading to intercomplex associations via the RAD50 apices [146]. The intercomplex model posits that zinc-mediated dimerization is required for DNA bridging and thus promotion of DNA repair [144]. However, recent researches determined the structure of the human RAD50 hook and coiled-coil domains and found that the predominant form of MRN complex is the intracomplex, demonstrating that hook and a interface within the coiled-coil domains of RAD50 coordinately promote intracomplex assembly and defining the intracomplex as the functional form of MRN complex [150].

In our experiments, we have directly visualized MRN binding to single DNA molecules under different buffer conditions using dual-trap optical tweezers and find that the affinity of MRN for DNA is enhanced by ATP and non-hydrolysable ATP analog (AMP-PNP). Furthermore, DNA tethering via MRN is investigated by manipulating two DNA molecules using quadruple-trap optical tweezers, demonstrating MRN bridging ability. Our results show that this bridging ability depends on protein-protein interactions and the bridging strength has no significant difference with or without ATP or AMP-PNP. However, ATP and AMP-PNP can greatly enhance DNA bridging rate. In addition, the study of MRN and DNA ends interactions directly prove that
5.2 Results and discussion

5.2.1 MRN-DNA interactions are influenced by ATP, AMP-PNP and Mg$^{2+}$

Recent structural analysis of MRE11/RAD50 proteins illuminate the physical interactions between MRE11 and RAD50 in both the nucleotide-free and ATP-bound states, and reveal that ATP binding results in extensive conformational changes at the base of the RAD50 coiled coil controlling the nuclease activity of the complex [151, 152, 153]. In our experiments, combining dual-trap optical tweezers, wide-field microscope and controlled laminar flow microfluidics allows us to directly observe the interactions of MRN complex with single DNA molecules in the absence or in the presence of ATP, AMP-PNP or Mg$^{2+}$. eGFP is fused to C-terminal of MRE11. Linear lambda phage DNA molecules are biotinylated on the 3' ends of two DNA strands and are tethered at 10 pN by optically trapped streptavidin-coated polystyrene microspheres with a diameter of 4.26 $\mu$m. After 120 s incubation in eGFP-MRN protein channel, the DNA construct is brought into a buffer channel where DNA-MRN complex can be visualized in the absence of flow and fluorescent proteins in solution. Figure 5.1 shows the fluorescent images of MRN complexes bound to DNA in the buffer channel in different conditions. In the absence of Mg$^{2+}$, isolated fluorescent patches bound to DNA can be observed either in the absence of ATP/AMP-PNP (Figure 5.1 A) or in the presence of ATP (Figure 5.1 B) or AMP-PNP (Figure 5.1 C). However, increasing the amount of Mg$^{2+}$ in protein buffer results in less binding of MRN (Figure 5.1 D, E, F). When it contains 1 mM Mg$^{2+}$, no MRN complexes attached to DNA (Figure 5.1 G, H, I). In the presence of ATP or AMP-PNP, more patches can be observed on DNA. The statistics of the numbers of MRN patches from 70 independent DNA molecules are shown in Figure 5.2. This experiment demonstrates that the affinity of MRN for DNA is enhanced by ATP and AMP-PNP while it is inhibited by Mg$^{2+}$. Our results are consistent with previous crystal structure analysis and biochemical assay.
which illustrated that ATP-driven conformational changes in RAD50 can promote DNA binding [59].

<table>
<thead>
<tr>
<th></th>
<th>-ATP/-AMPPNP</th>
<th>1 mM ATP</th>
<th>1mM AMPPNP</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 mM MgCl₂</td>
<td>(A)</td>
<td>(B)</td>
<td>(C)</td>
</tr>
<tr>
<td>0.5 mM MgCl₂</td>
<td>(D)</td>
<td>(E)</td>
<td>(F)</td>
</tr>
<tr>
<td>1 mM MgCl₂</td>
<td>(G)</td>
<td>(H)</td>
<td>(I)</td>
</tr>
</tbody>
</table>

**Figure 5.1** — Fluorescent images of DNA-MRN complex in buffer channel in different conditions, showing the MRN binding to DNA is affected by ATP, AMP-PNP and Mg²⁺. DNA molecule tension is kept at 10 pN. GFP-MRN concentration: 5 nM. The basic buffer includes 20 mM Hepes pH 7.3, 1 mM DTT and 100 μg/ml acetylated BSA. Other components are indicated in the table. Scale bar: 5 μm.

### 5.2.2 DNA tethering by the MRN complex

**Real time observation of DNA bridging via MRN complex**

To examine the dynamics of DNA bridging by MRN, we use quadruple-trap optical tweezers to manipulate DNA molecules and observe DNA bridging by fluorescently labeled MRN complexes in real time. In our experiments, DNA molecules can be bridged by MRN in different configurations: pairing and wrapping. For the pairing configuration, two DNA molecules are moved close to each other by manipulating four beads in the presence of gentle flow in protein channel (See Figure 5.3 A and B). After 120 s incubation in protein channel, DNA-MRN complexes are brought into the buffer channel in the absence of flow and bridges between two DNA molecules can be clearly visualized in real time (See Figure 5.3 C). This bridging occurred in protein channel in the absence of ATP and AMP-PNP. Our experiment shows that the bridging rate in pairing configuration is up to 70% in absence of ATP and AMP-PNP (6 out of 9 tests). For the wrapping configuration, two DNA molecules are incubated 120 s in protein channel and moved into buffer channel where DNA molecules are wrapped (See Figure 5.4 A and B). After 120 s incubation in wrapping configuration in buffer channel,
5.2 Results and discussion

![Graphs showing patch counts with different concentrations of MgCl2 and ATP/AMP-PNP.](A) In the absence of MgCl2, the distribution of MRN complex patches on DNA without or with ATP, AMP-PNP. The binding of MRN is inhibited by increasing the concentration of MgCl2 either in the absence of ATP/AMP-PNP (B) or in the presence of ATP (C) or AMP-PNP (D).

Figure 5.2 – Statistics of MRN complex patches attached on DNA from 70 independent experiments. Each dot represents one molecule. (A) In the absence of Mg2+, the distribution of MRN complex patches on DNA without or with ATP, AMP-PNP. The binding of MRN is inhibited by increasing the concentration of Mg2+ either in the absence of ATP/AMP-PNP (B) or in the presence of ATP (C) or AMP-PNP (D).

![Graphs showing patch counts with different concentrations of MgCl2 and ATP/AMP-PNP.](A) In the absence of MgCl2, the distribution of MRN complex patches on DNA without or with ATP, AMP-PNP. The binding of MRN is inhibited by increasing the concentration of MgCl2 either in the absence of ATP/AMP-PNP (B) or in the presence of ATP (C) or AMP-PNP (D).

the construct is unwrapped and DNA bridging can be observed (Figure 5.4 C). The bridging shown in Figure 5.4 occurred in buffer channel in the presence of AMP-PNP. The bridging rate in the wrapping configuration over 63 independent experiments in all different buffer conditions is about 70%.

**Sliding of the MRN tethers**

Next, we investigate the dynamic behavior of the MRN junctions. To study this, we controlled one molecule moving along the other by manipulating two trapped beads. We find that the MRN junctions can be forced to smoothly slide back and forth along
5.2 Results and discussion

Figure 5.3 – An example of two DNA molecules bridging via MRN in pairing manner. (A) Two DNA molecules tethered by four beads in protein channel in presence of gentle flow. (B) Two DNA molecules touch to each other by controlling the beads in protein channel in the presence of gentle flow. (C) Two bridged DNA in buffer channel in the absence of flow. (D) The bridged junction slides to the right side of the bottom DNA. (E) The bridged junction slides to the left side of the bottom DNA. In this experiment, the protein buffer contains 20 mM Hepes pH 7.3, 1 mM DTT and 100 µg/ml acetylated BSA. MRN protein concentration: 5 nM. The flow direction in protein channel is shown by wide arrow. Bridged junction is indicated by the small arrow. Scale bar: 5 µm.

Figure 5.4 – An example of two DNA molecules bridging via MRN in wrapping manner. (A) Two DNA-MRN complexes tethered by four beads in buffer channel after 120 s incubation in protein channel. (B) Wrapped DNA-MRN complexes in buffer channel. (C) Unwrapped DNA-MRN complexes in buffer channel. It shows that DNA are bridged by MRN complex. (D) The bridged junction slides to the right side of the bottom DNA. (E) The bridged junction slides to the left side of the bottom DNA. In this experiment, the protein buffer contains 20 mM Hepes pH 7.3, 1 mM DTT, 100 µg/ml acetylated BSA, 0.25 mM Mg$^{2+}$ and 1 mM AMP-PNP. MRN protein concentration: 5 nM. Bridged junction is indicated by small arrow. Scale bar: 5 µm.
DNA both in pairing configuration (Figure 5.3 C, D, E) and in wrapping configurations (Figure 5.4). The sliding distance is more than 40,000 bp.

**Strength of the MRN tethers**

Another intriguing question about these bridges is how strong they are. To address this issue, we determine the strength of these junctions by measuring the rupture forces. Figure 5.5 is one typical process of such rupture force measurement. Two DNA-MRN complexes are wrapped in buffer channel after 120 s incubation in protein channel (Figure 5.5 A and B). After 120 s incubation in wrapped configuration, DNA-MRN complexes are unwrapped and two DNA molecules are bridged by MRN complex (Figure 5.5 C). To measure the strength of the bridge, the positions of four beads are rearranged. We calibrate the left bead and stretch the right one. The corresponding force is shown in Figure 5.5 F. Since the strength is not strong, two DNA molecules are separated at last (Figure 5.5 E). We measured the strength of bridges in 3 different buffer conditions: in the absence of ATP/AMP-PNP/Mg\(^{2+}\) (15 measurements), in the presence of 1 mM ATP and 0.25 mM Mg\(^{2+}\) (11 measurements), and in the presence of 1 mM AMP-PNP and 0.25 mM Mg\(^{2+}\) (22 measurements). The average strength is 18.0 ± 12.6 pN, 19.3 ± 12.6 pN and 17.7 ± 8.9 pN, respectively, as shown in Figure 5.6 A. Our results show that the strength of bridges has no significant difference in different buffer conditions. We also compared the bridging rate in those conditions, as indicated in Figure 5.6 B where it shows that the bridging rate is about 50% in the absence of ATP/AMP-PMP/Mg\(^{2+}\) (over 42 experiments), while it can be up to 90% in the presence of 1 mM ATP and 0.25 mM Mg\(^{2+}\) (over 10 experiments) and 85% in the presence of 1 mM AMP-PNP and 0.25 mM Mg\(^{2+}\) (over 20 experiments). Our experiments demonstrate that ATP and AMP-PNP do not enhance the strength of bridges while they can stimulate MRN complex to bridge DNA.

**DNA bridging via MRN depends on protein-protein interactions**

The intercomplex model suggests that DNA binding by MRN globular domain can induce the opening of coiled coils of RAD50, leading to zinc-mediated connection with another opened coiled coils of RAD50 from another MRN complex for DNA bridging [146, 144]. The intracomplex model proposes that the predominant form of MRN complex is intracomplex which is the functional form for DNA repair and oligomerization [150]. Both for intracomplex and for intercomplex model, protein-protein interactions are required for DNA bridging and thus DNA repair. To test the importance of protein-protein interactions for DNA bridging via MRN complex, we use
Figure 5.5 – Typical force measurement of the bridged junction. (A) Two DNA-MRN complexes in buffer channel. (B) Wrapped DNA-MRN complexes in buffer channel. (C) Unwrapped DNA-MRN complexes in buffer channel showing DNA bridging. (D) The positions of four beads are rearranged for the force measurement during which the left bead is calibrated for recording force and the right is stretched. (E) Two DNA molecules are separated due to the force measurement. (F) The recorded force curve during force measurement. In this experiment, the protein buffer contains 20 mM Hepes pH 7.3, 1 mM DTT, 100 µg/ml acetylated BSA, 0.25 mM Mg$^{2+}$ and 1 mM ATP. MRN protein concentration: 5 nM. Scale bar: 5 µm.

one DNA molecule coated with MRN complex and one bared DNA molecule, as shown in Figure 5.7 A. These two molecules are wrapped and the MRN complex is located at the wrapped junction (Figure 5.7 B). After 120 s incubation, wrapped molecules are unwrapped and no bridging occurs, as indicated in Figure 5.7 C. Even the incubation time is increased to 300 s, there is still no DNA bridging. We repeat the experiments 6 times in the presence of 1 mM AMP-PNP and 0.25 mM Mg$^{2+}$ and 2 times in the
5.2 Results and discussion

Figure 5.6 – (A) Comparison of force distributions in 3 different conditions. In the absence of ATP/AMP-PNP/Mg$^{2+}$, the averaged bridging strength is $18.0 \pm 12.6$ pN over 15 measurements. In the presence of 1 mM ATP and 0.25 mM Mg$^{2+}$, the averaged bridging strength is $19.3 \pm 12.6$ pN over 11 measurements. In the presence of 1 mM AMP-PNP and 0.25 mM Mg$^{2+}$, the averaged bridging strength is $17.7 \pm 8.9$ pN over 22 measurements. (B) The comparison of bridging rate in 3 different conditions. In the absence of ATP/AMP-PNP/Mg$^{2+}$, in the presence of 1 mM ATP and 0.25 mM Mg$^{2+}$, and in the presence of 1 mM AMP-PNP and 0.25 mM Mg$^{2+}$, the bridging rate is about 50% (over 42 experiments), 90% (over 10 experiments) and 85% (over 20 experiments), respectively. In these experiments, basic protein buffer contains 20 mM Hepes pH 7.3, 1 mM DTT, 100 $\mu$g/ml acetylated BSA and MRN protein concentration is 5 nM.

absence of ATP/AMP-PNP/Mg$^{2+}$. All of them show that no DNA bridging occurs if there is no protein-protein interactions from two DNA molecules.

5.2.3 Interaction of MRN with DNA ends and tethering

Since the MRN complex, a sensor of DNA damage, is involved in DNA double-strand break repair, we want to know how the MRN complex interacts with broken ends. To study this interaction under precise control, we introduce DNA blunt ends using SfoI restriction enzyme which can recognize and cut lambda DNA at GGC$^\sim$GCC site, producing a 45,679 bp-long DNA with blunt end. In this experiment, a laminar flow cell with 5 channels is used (Beads channel, DNA channel, buffer channel (the same buffer as protein buffer), protein channel (5 nM MRN complex, 20 mM Hepes PH 7.3, 1 mM DTT, 100 $\mu$g/ml acetylated BSA), and SfoI enzyme channel (1 in 250 1X Cutsmart buffer)). This experiment is conducted in slight buffer flow, so that, after generation of free end, the DNA end can be extended in the direction of the flow. First, two bare DNA molecules tethered by four beads are brought into SfoI enzyme channel and cut
Figure 5.7 – Typical example of DNA bridging via MRN depends on protein-protein interactions. (A) One DNA molecule (top) coated with MRN and one bared DNA (bottom) in buffer channel. (B) Wrapped construct in buffer channel. (C) No bridging occurs after unwrapping. In this experiment, the protein buffer contains 20 mM Hepes pH 7.3, 1 mM DTT, 100 µg/ml acetylated BSA, 0.25 mM Mg\(^{2+}\) and 1 mM AMP-PNP. MRN protein concentration: 5 nM. Scale bar: 5 µm.

in short time. Because DNA molecules are invisible in SfoI channel, they are inspected to ensure that they have been cut. Subsequently both cut DNA molecules with blunt ends are moved into MRN protein channel. Figure 5.9 shows that two DNA molecules with blunt ends bound with MRN complex in protein channel. We manipulate the two DNA with blunt end to touch each other (Figure 5.9 A). One end of DNA can bridge to another DNA via MRN complex (Figure 5.9 B) and the bridged junction can slide along the DNA to its end (Figure 5.9 C, D, E, F). This end-end bridged hybrid DNA complex can be relaxed and stretched by controlling beads, suggesting that MRN complex can stably hold two ends together (Figure 5.9 G, H). Finally, this complex can be separated by stretching it. The strength of bridged junction is less than 20 pN. This experiment is performed in the absence of ATP, APM-PNP and Mg\(^{2+}\) and repeat 3 times. Our results directly demonstrate that MRN complex can tether two blunt DNA ends independent of ATP and AMP-PNP.

5.2.4 The MRN complex with RAD50 mutations interaction with DNA molecules

Previous models suggest that MRN complex facilitates and coordinates DSBs repair by tethering two DNA molecules via hook and coiled-coil domains of RAD50 ([154, 144, 146]). Our experiments have directly shown that DNA can be bridged by MRN complex.
To further verify whether the DNA bridging via MRN is mediated by the hook and coiled coil domains of RAD50, we truncate the hook and coiled coil domains of RAD50 and use the truncated MRN\textsuperscript{\Delta CC} complex to test DNA bridging. Surprisingly, we find that the MRN\textsuperscript{\Delta CC} complex can bridge DNA. Figure 5.10 A is a typical example of DNA bridging via MRN\textsuperscript{\Delta CC} complex in the absence of ATP, AMP-PNP and Mg\textsuperscript{2+}. 2 out of 5 independent experiments (40 \%) show DNA bridging in this case. The averaged bridging force is 24.5±6.6 pN. Comparing with wide type MRN complex in the same buffer condition, the bridging rate is smaller and strength of bridges is similar within error range. Our results provide evidence that the hook and coiled coil of RAD50 may not be necessary for DNA bridging. Biochemical studies suggest that the DNA-binding head bridges variable ends via MRE11 protein, which interacts with ssDNA, dsDNA, and forked DNA structures [155, 156]. Also, X-Ray structure analysis of MRE11-DNA complex suggests that dimeric MRE11 can align and tether two DNA ends for short range DNA synapsis [154]. Our results of two blunt DNA tethering via MRN\textsuperscript{\Delta CC} complex could be attributed to MRE11 protein instead of the RAD50 apices.

5.3 Conclusions and Perspectives

In this study we have directly visualized MRN binding to single DNA molecule under different buffer conditions using dual-trap optical tweezers and find that MRN and DNA affinity is enhanced by ATP and AMP-PNP and inhibited by high concentration
Figure 5.9 – Two blunt DNA molecules tethering by MRN complex. Related to Figure. (A) Two DNA molecules with blunt ends are manipulated to close to each other. (B) One end of DNA molecule (black) bridges to another DNA (red) via MRN complex. (C), (D), (E), (F) This end can smoothly slide along the DNA (red). (G), (H) Two blunt ends are held by MRN complex stably. (I) Two DNA molecules are separated due to stretching. This experiment is performed in a slight buffer flow in protein channel. The protein buffer contains 20 mM Hepes pH 7.3, 1 mM DTT, 100 µg/ml acetylated BSA. MRN protein concentration: 5 nM. The arrow indicates the flow direction. Scale bar: 5 µm.
of Mg$^{2+}$. Furthermore, DNA tethering via MRN are investigated by manipulating two DNA molecules using quadruple-trap optical tweezers, demonstrating that MRN complex can form stable, mobile DNA bridges capable of holding DNA molecules. MRN can bridge DNA molecules at any location. The mobility of the bridged junction allows them to slide along DNA to its broken ends accessible for further processing, which is directly verified by blunt DNA ends tethering experiments. Our results show that this bridging ability depend on protein-protein interactions and the bridging strength has no significant difference with or without ATP or AMP-PNP. Remarkably, we find that MRN$\Delta$CC complex can also bridge DNA molecules, suggesting that the hook and coiled coil domains of RAD50 indispensable for DNA repair may not be necessary for DNA bridging. In the future, more experiments should be conducted about MRN$\Delta$CC complex in different conditions to test the bridging function of MRN$\Delta$CC complex in the presence of ATP or AMP-PNP.
References


[38] Ignacio Tinoco et al. “Biological mechanisms, one molecule at a time Biological mechanisms, one molecule at a time”. In: (2011), pp. 1205–1231.


References


[87] B H Lok et al. “RAD52 inactivation is synthetically lethal with deficiencies in BRCA1 and PALB2 in addition to BRCA2 through RAD51-mediated homologous recombination.” In: Oncogene 32.30 (2013), pp. 3552–8.


[90] Sotirios K. Sotiriou et al. “Mammalian RAD52 Functions in Break-Induced Replication Repair of Collapsed DNA Replication Forks”. In: Molecular Cell 64.6 (2016), pp. 1127–1134.

[91] Hong Zan et al. “Rad52 competes with Ku70/Ku86 for binding to S-region DSB ends to modulate antibody class-switch DNA recombination”. In: Nature Communications 8 (Feb. 2017), p. 14244.


References


107

References


Summary

The primary objective of this thesis was to study DNA double-strand break repair mechanisms at the single molecule level by combining optical tweezers, wide-field fluorescence microscopy and microfluidics in order to manipulate single DNA molecules while observing fluorescently-labeled protein interacting in real time with DNA.

First, we studied the interaction of human RAD52 with ssDNA and dsDNA. We found that the interaction with ssDNA is rapid, static, and tight. ssDNA appears to wrap around RAD52 complexes, which promotes intra-molecular DNA connections. On the contrary, dsDNA interaction is slower, weaker and often diffusive. Force-spectroscopy experiments showed that RAD52 alters the mechanics of dsDNA by enhancing DNA flexibility and increasing DNA contour length, suggesting intercalation. RAD52 binding also changes the nature of the overstretching transition of dsDNA preventing DNA melting, which is advantageous for strand clamping during or after annealing. Furthermore, we directly demonstrated that DNA-bound RAD52 is efficient at capturing ssDNA in trans. Together, these effects may help key steps in DNA repair, such as second-end capture during homologous recombination or strand annealing during RAD51-independent recombination reactions.

Furthermore, using quadruple-trap optical tweezers, we developed innovative methods to manipulate two DNA molecules and investigate the DNA bridging activity of the XRCC4-XLF, XRCC4-LigIV and MRN complexes. We confirmed that XRCC4-XLF can form strong and mobile DNA bridges capable of holding DNA fragments together. The mobility of the bridged junctions allows them to slide along the DNA, which is advantageous for end processing and subsequent ligation. We found that also the XRCC4-LigIV complex has a DNA bridging ability. The bridges formed by XRCC4-LigIV are mostly immobile and resist very high force. We also demonstrated that the MRN complex can form DNA bridges and its bridging ability is greatly enhanced in the presence of ATP. Remarkably, we found that MRN complexes lacking the hook and coiled coil domains of RAD50 are able to bridge DNA molecules. This suggests that the hook and the coiled coil of RAD50, while indispensable for DNA repair,
are not required for DNA bridging, which challenges previous models that invoked a predominant role of the hook and coiled coil domains for efficient DNA bridging.
List of Publications


Fellowship Supports

Financement 1: Financement d’un Etablissement d’enseignement supérieur
Type de contrat : Contrat doctoral
Durée : du 1 décembre 2013 au 30 novembre 2016
Employeur : AMU - Contrat Collège - Origine des fonds : MESER

Financement 2: Financement d’une association ou fondation
Type de contrat : Contrat de travail à Durée Déterminée CDD
Durée : du 1 décembre 2016 au 31 mai 2017
Employeur : AMU - Contrat Collège - Origine des fonds : La Ligue Nationale contre le Cancer

Financement 3: Partiellement par LASERLAB-EUROPE (grant agreement no. 284464, EC’s Seventh Framework Programme).

Financement 4: Partiellement par le CNRS

Financement 5: Partiellement par l’INSERM
Acknowledgements

When I start this part, I realize that my doctoral term is about to end. I don’t know how to express my feeling, because a lot of memories are in my mind at this moment and there are many people to whom I would like to say thanks.

There is no doubt that, first of all, I am eternally grateful to my advisor, Mauro, for giving me the opportunity to pursue my Ph.D. in this very exciting interdisciplinary field. Mauro, I admire your optimistic spirit and persistent attitude to research. You always guide me to think deeply and enlighten me on how to deal with new projects and how to better express an experiment and its results. When I felt frustrated by non-productive experiments, you encouraged me to believe in myself and gave me effective advice. Besides, your humor and enthusiasm to life are very impressive to me. Whenever and wherever, you can create an optimistic and pleasant atmosphere around you, infecting with me.

Also, I would like to thank my co-advisor, Sophie, for generous supports to my work. I admire your patience and keen mind on experiment. You are able to point out the valuable essence from massive data and offer relative suggestions for next step.

A special thanks to Davide for his unconditional supports and guidance on experiments. We spent so many unforgettable days in dark room building up experimental system, discussing experiments, sharing exciting moments together. You inspired me to consider questions from different perspectives and to analyze data from different manners, which play an important role in analyzing my experimental results in this thesis.

This thesis would not complete without continuous help and supports from my wonderful cooperation partners. Erwin and Gijs, you are great scientists. Thank you very much for giving me chances to go to your lab doing RAD52 project and learning experimental skills. Ineke, it was great to work together with you on the RAD52 project. You have amazing working efficiency. Thanks to your efforts, we had a nice paper that is an essential part of this thesis. Andrea, thank you for useful discussions.
and helps when I encountered experimental problems. For the MRN project, Tanya and Logan, thank you for vast efforts on sample preparation and great suggestions.

Besides the scientific work, I am also thankful to my nice friends in Mauro’s group. Satish, thank you for your patient explanation and countless helps on molecular biology. Edwige, thank you for always being there when I needed help for anything. Abi, Stéphanie, Quynh, Seungyun, thank you for sharing so many great and unforgettable moments in our crowned office, during coffee breaks, and in cafeteria. Many thanks to the friends in CRCM who always hold a special place in my heart. Arato, and Hiro thank you for spending fun time together. We have many fantastic adventures in Marseille and my life here becomes more interesting and exciting. Pierre-Marie, and Eunice, thank you for inviting me many times to your home having delicious food and for helping me solve various problems. Luisa, Jean-Hugues, Pierre-Henri, Bertrand, Sarah, Samar, Sylvain, thank you for useful discussions and having pleasant time together. Also, a big thanks to former and present people in Mosaic team in Institut Fresnel: Naveen, Wei, Xueqin, Cesar, Haitham, Alberto, Carolina, for playing games and having great time in Fresnel. Especially to Hervé, thank you for fruitful discussions and useful advices on my presentations.

Finally, I would like to thank my wife Nan, my parents Xiaolin and Wenlan, and my brother Hongchuan, for providing selfless and unwavering backup in my life. This thesis is also for you.