Thesis for the degree of Doctor of Philosophy

BIOPHYSICAL STUDIES OF DNA BINDING

BY THE LARGE FILAMENT-FORMING PROTEIN RAD51 AND THE SMALL MINOR-GROOVE BINDER HOECHST 33258

LOUISE H. FORNANDER

Department of Chemistry and Chemical Engineering
Chalmers University of Technology
Gothenburg, Sweden 2015

Biophysical studies of DNA binding – by the large filament-forming protein Rad51 and the small minor-groove binder Hoechst 33258
LOUISE H. FORNANDER
ISBN 978-91-7597-280-0

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie Nr 3961
ISSN 0346-718X

Department of Chemistry and Chemical Engineering
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Front Cover:
Left: Molecular model of three human Rad51 monomers binding to a poly(dT) (12-mer). Model adopted from appended Paper I of this Thesis.
Right: Crystal structure of Hoechst 33258 binding to the oligonucleotide $d(CGCAAATTTGCG)_2$, PDB ID: 264D.

Back Cover:
Photo credit to SVT.

Printed by Chalmers Reproservice
Göteborg, Sweden 2015

BIOPHYSICAL STUDIES OF DNA BINDING BY THE LARGE FILAMENT-FORMING PROTEIN RAD51 AND THE SMALL MINOR-GROOVE BINDER HOECHST 33258

LOUISE H. FORNANDER

Department of Chemistry and Chemical Engineering
Chalmers University of Technology

Abstract

Mechanistic insight into the nature of DNA-binding ligands is crucial for both drug development as well as understanding more complex biological reactions that take place in the cell. In this Thesis, two rather different DNA-binding molecules are considered: 1. the large, filament-forming eukaryotic recombination protein Rad51, which is essential in the strand exchange reaction during homologous recombination, the most accurate repair system of DNA double-strand breaks, and 2. the small, synthetic DNA ligand Hoechst 33258, which is a model drug for DNA minor groove interactions.

The Rad51 filament formation, reflected in the length of short individual Rad51 filament patches on long DNA strands, has been examined by nanofluidics in combination with fluorescence microscopy. Analyses of the dynamics of the Rad51-DNA complex in the nanochannel reveal structural variations that depend on the filament formation conditions; the choice of divalent cations (Mg^{2+} or Ca^{2+}), the DNA substrate (single- or double-stranded), and the Rad51 nucleation concentration affected the macroscopic structure of the filament.

The structural effects that the divalent cations Mg^{2+} and Ca^{2+}, and the accessory protein Swi5-Sfr1 exert on the Rad51-single-stranded (ss) DNA filament at a microscopic level have also been examined by linear dichroism (LD). The naturally unordered bases in ssDNA become preferentially perpendicularly oriented relative to the DNA backbone in presence of Rad51 with Ca^{2+} alone, or Mg^{2+} in combination with the accessory protein Swi5-Sfr1. A preferentially perpendicular base organization is proposed to mechanistically relate to an efficient strand exchange reaction, supposedly due to more critical base matching with the invading double-stranded DNA.

To aid future spectroscopic structural analyses of proteins that contain tyrosine residues, such as Rad51, a combined spectroscopic and *in silico* study of the chromophore in tyrosine has been conducted. It is demonstrated how the spectroscopic properties of tyrosine are sensitively dependent on the polarity of the environment, mainly through the ability to form hydrogen bonds and the rotation of the hydroxyl group.

The last part of the Thesis deals with spectroscopic and thermodynamic studies of the binding of Hoechst 33258 to three different DNA oligonucleotides with AT-tracts of various lengths. The binding at high drug-to-DNA ratio is especially considered, and an important conclusion is that two Hoechst 33258 molecules bind in parallel, a slight distance apart, in the minor groove of an oligonucleotide with 8 consecutive adenines/thymines.

Keywords: DNA, DNA-binding ligands, Rad51, Hoechst 33258, Spectroscopy, Linear Dichroism, Nanofluidics, Isothermal Titration Calorimetry.

List of publications

This thesis is based on the work contained in the following papers:

I Ca^{2+} improves organization of single-stranded DNA bases in human Rad51 filament, explaining stimulatory effect on gene recombination
Louise H. Fornander, Karolin Frykholm, Anna Reymer, Axelle Renodon-Cornière, Masayuki Takahashi and Bengt Nordén*
Nucleic acids research (2012), **40(11)**, 4904-4913.

II Swi5-Sfr1 protein stimulates Rad51-mediated DNA strand exchange reaction through organization of DNA bases in the presynaptic filament
Louise H. Fornander, Axelle Renodon-Cornière, Naoyuki Kuwabara, Kentaro Ito, Yasuhiro Tsutsui, Toshiyuki Shimizu, Hiroshi Iwasaki, Bengt Nordén and Masayuki Takahashi*
Nucleic acids research (2014), **42(4)**, 2358-2365.

III Visualizing the non-homogeneous structure of RAD51 filaments using nanofluidic channels
Louise H. Fornander, Karolin Frykholm, Joachim Fritzsche, Joshua Araya, Philip Nevin, Erik Werner, Ali Çekir, Fredrik Persson, Edwige B. Garcin, Penny Beuning, Bernhard Mehlig, Mauro Modesti and Fredrik Westerlund*
Submitted

IV UV transition moments of tyrosine
Louise H. Fornander, Bobo Feng, Tamás Beke-Somfai and Bengt Nordén*
Journal of Physical Chemistry B (2014), **118(31)**, 9247-9257.

V Minor-groove binding drugs: where is the second Hoechst 33258 Molecule?
Louise H. Fornander, Lisha Wu, Martin Billeter, Per Lincoln and Bengt Nordén*
Journal of Physical Chemistry B (2013), **117(19)**, 5820-5830.

Contribution report

Paper I: The author of this Thesis, Louise H. Fornander, (L.F) prepared and purified the human Rad51 protein, and carried out and analyzed the linear dichroic measurements. A.R-C performed the strand exchange activity measurements and A.R the molecular modeling. L.F wrote the paper.

Paper II: L.F carried out and analyzed the linear dichroic measurements. The proteins were prepared by K.I and N.K, A.R-C performed the strand exchange activity measurements and A.R the molecular modeling. L.F wrote the paper.

Paper III: L.F performed the microscopy measurements. The nanofluidic chips were prepared by J.F, the data analysis was perfomed in collaboration with A.C and E.W and the proteins were prepared by E.G and M.M. L.F wrote the paper.

Paper IV: L.F carried out and analyzed the spectroscopic measurements, QM calculations were performed by T.B-S and B.F. L.F wrote most of the paper.

Paper V: L.F performed and analyzed the CD and ITC measurements, the fitting of the data to theoretical isotherms was performed by P.L. ^{1}H NMR was performed by L.W. L.F wrote the paper.

Publications not included in this thesis

Lifetime heterogeneity of DNA-bound dppz complexes originates from distinct intercalation geometries determined by complex-complex interactions
Johanna Andersson, Louise H. Fornander, Maria Abrahamsson, Eimer Tuite, Pär Nordell and Per Lincoln*
Inorganic Chemistry (2013), **52(2)**, 1151-1159.

Sensing conformational changes in DNA upon ligand binding using QCM-D. Polyamine condensation and Rad51 extension of DNA layers
Lu Sun, Karolin Frykholm, Louise H. Fornander, Sofia Svedhem, Fredrik Westerlund and Björn Åkerman*
Journal of Physical Chemsitry B (2014), **118(41)**, 11895-11904.

Cell surface binding and uptake of arginine- and lysine-rich penetratin peptides in absence and presence of proteoglycans
Helene L. Åmand, Hanna Rydberg, Louise H. Fornander, Per Lincoln, Bengt Nordén and Elin K. Esbjörner*
Biochimica et Biophysica Acta (2012), **1818(11)**, 2669-2678.

The conquest of middle-earth: combining top-down and bottom-up nanofabrication for constructing nanoparticle based devices.
Yuri A. Diaz Fernandez, Tina A. Gschneidtner, Carl Wadell, Louise H. Fornander, Samuel Lara Avila, Christoph Langhammer, Fredrik Westerlund and Kasper Moth-Poulsen*
Nanoscale (2014), **6(24)**, 14605-14616.

Table of contents

1 Thesis introduction

Using light to detect life.

What is life? This is a major question that scientists and mankind have tried to answer and define for a very long time, and yet, there is still not a simple definition that all can agree upon. Numerous definitions have come and gone over the years, but they have all been dismissed in one way or another, this is because *life* and living organisms are extremely diverse, and one single definition that suits all is hard to find.

For me, personally, there are three major factors to describe *life*: *reproduction*, *energy consumption* and *programmed instructions*. Living organisms grow and their whole life form reproduces. They use energy to counteract an otherwise spontaneous entropy decomposition of themselves, and they follow a set of programmed instructions captured in their genetic code, DNA (or RNA). An example of a reproducible non-living organism is crystals; they may break off and reproduce, their formation involves energy consumption, and they form highly ordered structures (low entropy), but they do not possess any genetic code. Thus their organized structure is solely based on the rules of physics, and their structures are the same under similar environmental conditions. However, living matter has its genetic code, which is inherited generation by generation, and by the help of changes in this code, small or large, life forms are able to diversify and adapt to environmental changes.

The genetic code of *life*, DNA, is in the spotlight throughout this Thesis. Literally in the spot*light*. The interactions between *light* and matter (*i.e.* molecules), which is central in the field of spectroscopy, is highly exploited in all appended papers of this Thesis. Spectroscopy is a vast and diverse field, both regarding the type of radiation used (*e.g.* UV, radio waves, X-ray radiation, and even particles such as electrons or neutrons can be used), and the type of interaction studied (*e.g.* absorption, emission or scattering). The diversity of measurement methods provides the field of spectroscopy the possibility to study molecules on both large and small scales (bulk and single molecule), and even the arrangement of single atoms within a molecule can be deduced. In this Thesis, *light*-matter interactions are used in order to deduce the structure of the complexes DNA forms with two particular molecules, Rad51 and Hoechst 33258.

Rad51 and Hoechst 33258 are two DNA-binding molecules that are rather different from one another; Rad51 is a large filament-forming protein that forms a helix around the DNA, while Hoechst 33258 is a small synthetic molecule that binds in the groove of DNA. They differ not only in their molecular sizes, but also in sizes of their DNA binding regions, the type of DNA-binding, and, of course, the effects that the two molecules exert on the DNA. However, both bind to DNA non-covalently. With gained knowledge about their DNA-binding, we might come

one step further when designing future DNA-binding drugs, another extremely diverse field. The DNA molecule has countless opportunities when it comes to treatment of numerous diseases, since the molecule carries all the programmed instructions needed for *life;* that is instructions about development, function and reproduction. By developing drugs that can sequence-specifically target the DNA we can increase the number of approaches on how to handle the event of deficient genetic material. For example, we could increase/reduce the expression of certain genes, be able to remove/exchange a malfunctioning part of the genome, or we could simply use the action of a DNA-binding drug to kill a "bad" cell.

Changes in the genetic code are, as mentioned, vital for any organism since genetic alterations are necessary to the organism in order for it to adapt and evolve. These genetic alterations, called mutations, are permanent changes of the four "letters" of the genetic code, created through various DNA-damaging processes. However, only a minor part of the total number of mutations in an organism results in increased chance for survival and reproduction. Most mutations do not even affect the organism, or if they do it is usually in a negative way, which in worst case could lead to cell death or cancer. The DNA is damaged constantly through various processes, and mutations in the genetic code are usually a result of a DNA lesion that is not repaired correctly by the repair systems of the cell. In humans, the total number of DNA lesions could vary from 10 000 to up to 1 million per day and cell (1–3). Most of these DNA-damages are adequately repaired by our repair systems. One particularly hazardous DNA-damage is a DNA double-strand break (DSB), which is the event where both of the strands in the DNA-helix are broken in close proximity to each other. If such damage is not repaired satisfactory, it may result in major mutations, and most probably cell death. One single DSB could result in arrest of the cell cycle (4), and the rate of programmed cell death, apoptosis, is correlated to the number of DSBs in the genome (5). The cell has two competing processes to handle a DSB, non-homologous end-joining (NHEJ) and homologous recombination (HR) (6), of which the latter will be discussed in detail in this Thesis. HR accurately repairs the broken DNA strand by using another, unbroken, homologous DNA as a template. A central reaction in HR is the strand exchange reaction, in which the recombination protein Rad51 that is studied in this Thesis is essential. Worth to mention is that Rad51 is only found in eukaryotes, but it has orthologs in both bacteria (RecA) and archaea (RadA), demonstrating the importance of accurately healing DSBs throughout all domains of life.

Hoechst 33258 is a synthetic molecule that through hydrophobic interactions and hydrogen bonding with the DNA bases binds non-covalently in the minor groove of DNA. The minor groove is a suitable target when designing sequence-specific

DNA-binding drugs. In the minor groove, a well-fitted drug has the possibility to recognize and interact specifically with a relatively large number of bases, and they are usually small enough to pass the cell membrane and reach the DNA that is positioned in the cell nucleus. Hoechst 33258 has a preference for two of the four "letters" in the DNA (adenines and thymines), but the exact binding mode at high drug concentration is not completely clear. Basic research concerning the binding of this widely used DNA stain may help when designing future DNA-targeting drugs.

In order to comprehend the results of the appended research presented in this Thesis (Papers I-V), I have the ambition to provide the reader with an adequate theoretical foundation in the field of DNA and DNA-binding molecules, as well as an introduction to the techniques used in the Thesis. Opening the theory section is a chapter describing the chemistry of DNA, and also how the DNA molecule behaves as a polymer and as a drug-target (Chapter 2). Thereafter follows a chapter concerning the DNA repair system and HR, where Rad51 in particular will be discussed (Chapter 3). The concept of *light*-matter interactions, the foundation for spectroscopy, is described in the proceeding chapter (Chapter 4), together with the other main experimental techniques used in the Thesis.

The results from the five appended papers are summarized (Chapter 5), and presented as four blocks (5.2-5.5). The first block (5.2) concerns the organization of the DNA bases in the Rad51-DNA filament, and how these are affected by the presence of the divalent cations Ca^{2+} and Mg^{2+}, and accessory protein Swi5-Sfr1. The organization of the DNA bases is also linked to the biological function of the Rad51 filament (Papers I-II). The second block (5.3) deals with how the filament formation conditions (Rad51 concentration, divalent cation present and DNA substrate) affect the macroscopic structure of the filament, and how that structure in turn is mirrored in the behavior of the filament when confined to nanofluidic channels (Paper III). In the third block (5.4), the spectroscopic features of the chromophore in the naturally occurring amino acid tyrosine, found in for example Rad51, are scrutinized. It is also presented how the spectroscopic information about tyrosine could be used for protein structure analysis (Paper IV). Finally, in the fourth and last block (5.5), the binding of Hoechst 33258 to oligonucleotides with AT-rich regions of different lengths is examined at high ligand to DNA ratios (Paper V).

2 Nucleic acids

In this section the focus is on the genetic information carrier, DNA. Its structure, function and physical properties are discussed, followed by a part describing how it can be used as a drug target.

2.1 Structure of DNA

Deoxyribonucleic acid, or DNA, is the molecular information carrier in our cells. It consists of two polymer strands that are intertwined around each other in a right-handed helix. The polymer strands are built up of four different building blocks, the nucleotides, which comprise the molecular information that the DNA molecule carries. Each nucleotide consists of a phosphoric acid that is esterified with a deoxyribose sugar, which in turn is covalently bound to one of four different DNA bases: adenine (A), guanine (G), thymine (T) or cytosine (C) (Figure 1A). A and G are two-ringed bases and are referred to as purines, while T and C are one-ringed pyrimidines. Each nucleotide is attached to its neighbor *via* phosphodiester bonds from the 5' carbon of the first nucleotide to the 3' carbon of the other, resulting in that single-stranded DNA (ssDNA) has an inherent direction. The ssDNA forms, usually, a double-helix with a second DNA strand. The DNA double-helix can exhibit multiple conformations, and the ones best known are the A-, B-, and Z-form, where the B-form is the most common conformation under physiological conditions (Figure 1B). In B-DNA, the helix is right-handed, with the bases centrally located and positioned perpendicular to the phosphate backbone, as steps in a spiral staircase. The DNA bases display a unique hydrogen bonding pattern, resulting in that A pairs with T, and G pairs with C. As a consequence of this, the two DNA strands in the double-helix are complementary to each other, thus carrying the same information. The C-G base pair contains three hydrogen bonds, compared to only two in the A-T base pair, thus base pair sequences rich in C-G are more stable than those rich in A-T.

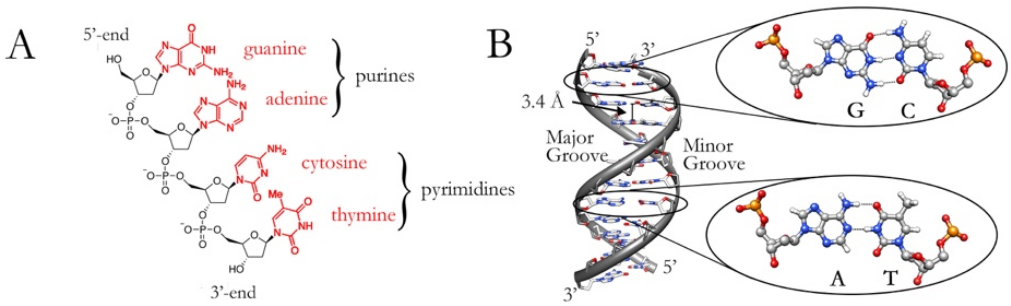

Figure 1. The molecular structure of DNA (A). Deoxyribose and phosphoric acid make up the backbone of the DNA, while the nucleobases adenine (A), guanine (G), thymine (T) and cytosine (C) protrude from the backbone. A B-form DNA is shown in (B) where two polymer chains are linked to each other *via* hydrogen bonds between the DNA bases; A base-pairs with T, and G base-pairs with C. The minor and major grooves of the B-DNA helix are also marked in the figure.

At physiological pH, the phosphate backbone of the DNA carries one negative charge per nucleotide. The repulsive charges between the two strands in the double-stranded DNA (dsDNA) are overcome by stabilizing contributions from the hydrogen bonds previously discussed, and from the hydrophobic interactions between the aromatic DNA bases that are stacked on top of each other in the center of the helix, as well as from the neutralizing contribution of the counterions present in the solvent. In B-DNA, the pitch (numbers of bases to complete a turn) is 10.5, with a helical rise of 3.4 Å. Two grooves run along the B-DNA helix, one on each side of the base pairs in the center. These grooves are called the major and the minor groove and, as the names imply, the major groove is wider than the minor groove (11 nm and 6 nm, respectively), but they hold similar depths (4 nm and 5 nm, respectively) (7). The minor groove, and drugs binding in it, will be discussed in more detail in Chapter 2.4.1.

2.2 Biological function of DNA

Already in 1944 Avery and co-workers demonstrated that DNA was the material for inheritance, and not the proteins as previously thought (8). However, genetic heredity remained poorly understood until 1953 when the classical paper of Watson and Crick was published with the structure of dsDNA (9). Soon thereafter they published another paper that suggested a molecular mechanism for copying and transferring genetic information to the next generation (10). This later lead to the discovery of the central dogma of biology, which explains the general flow of genetic information: from DNA to RNA, and from RNA to proteins (11, 12). The proteins, which are often referred to as the machineries of the cell, are in nature built up of 20 different amino acids, each coded for by a triplet of DNA bases, a codon.

As mentioned, DNA in its most common conformation consists of two complementary strands base-paired to each other, forming a stable B-form dsDNA. However, the strands of the dsDNA have to be separated in order for the stored genetic information to be transcribed to RNA, or replicated to another set of DNA. In the process of reading the genetic code in the dsDNA, a transcription, or replication, machinery composed of multiple proteins work as an ensemble. They open up the dsDNA helix and stabilizing the short exposed ssDNA strands while a polymerase transcribes the DNA code to a new RNA or DNA strand.

2.3 DNA - a biopolymer

Characterization of a polymer, such as long DNA, is not simply describing the chemistry of its smallest constituents, the nucleotides. Instead the polymer has to be characterized in terms of its molar mass, size and shape. The molar mass depends on the chemistry and the number of monomers, while the size and shape are properties that are also dependent on the solvent and the confinement. The basics of how to model a polymer, and how it behaves in solution and in confinement will be explained in the subsequent chapters (chapter 2.3.1 and chapter 2.3.2, respectively), and for a more comprehensive understanding the reader is referred to (13–16).

2.3.1 Polymer parameters

Depending on the length of the DNA, and if it is single- or double-stranded, which affects the polymer stiffness, different models can explain the polymer behavior in solution and in confinement. Before describing some of the simpler polymer models, parameters used in these will be presented. The parameters are also schematically depicted in Figure 2.

The contour length, L, is the end-to-end distance of the molecule if it is stretched out without any thermal fluctuations. It can be described by the total number monomers, N, multiplied by their height, h. For DNA, it is thus the number of bases or base pairs (bp), multiplied by 0.34 nm (Figure 2A).

The persistence length, P, may be regarded as the distance in contour over which the molecule remains in the same direction. The persistence length of dsDNA is about 50 nm (13, 17) and that of ssDNA is only a couple of nanometers (~2.5 nm) (18) (Figure 2A).

The end-to-end distance, R, is the distance from the start of the polymer to its end. The average root mean square of the end-to-end distance, $\sqrt{\langle R^2 \rangle}$, provides a measure of the spatial extent of the polymer, a useful measure when describing polymers in solution (Figure 2B).

The radius of gyration, R_g, can be used as substitute for R when the polymer exhibits a complex topology. It is a measure of the chain distribution about the average chain center-of-mass position (Figure 2B).

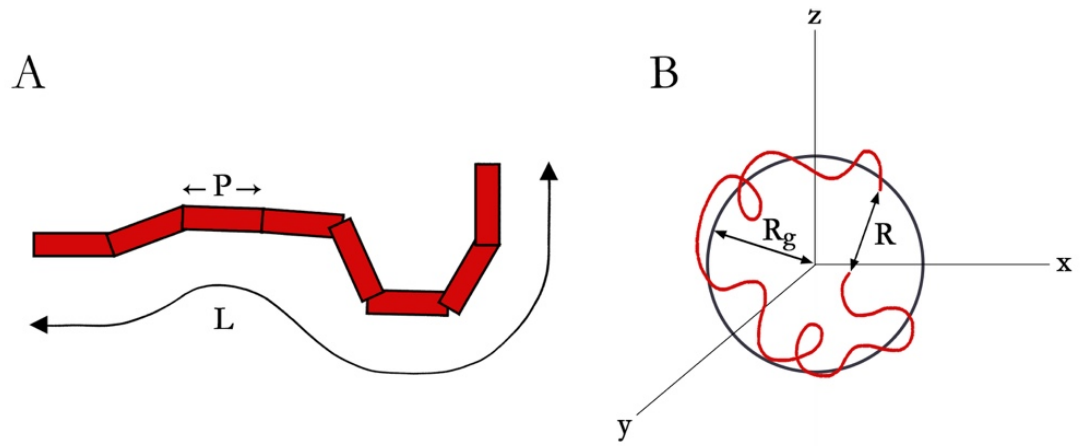

Figure 2. Schematic figures of parameters involved in describing a polymer. The contour length, L, and the persistence length, P, are depicted in (A), while the end-to-end distance, R, and the radius of gyration, R_g, are shown in (B).

2.3.2 Polymer theories

It is essential to understand that the conformation of a polymer in bulk or in confinement is dependent on a balance between entropic and enthalpic effects. For example, the persistence length *per se* is not a fundamental constant, but it is a result of the balance between these two effects. The entropic effects strive to increase the disorganization of the DNA coil, while the enthalpic effects strive to form energetically favorable bonds between segments of the coil. Also, since the conformation of the DNA coil depends on the free energy of the whole system, entropic and enthalpic contributions from molecules in the solvent influence the polymer behavior.

2.3.2.1 Polymers in solution

The simplest of all polymer models in solution is the freely jointed chain model. The freely jointed chain model assumes that there are no interactions between the monomers in the chain, and it also assumes that the monomers are linearly connected with all bonds equally probable. Theoretically, this means that two monomers could take up the same volume. The movement of the polymer in the freely jointed chain model can be approximated by a simple random walk with step length l ($l = 2P$). The root mean square of the end-to-end distance ($\sqrt{\langle R^2 \rangle}$) can be received as a function of the persistence length, P, and the number of nodes that perform a random walk, N ($N = L/2P$):

$$\sqrt{\langle R^2 \rangle} \sim 2P\sqrt{N} \qquad \text{[Eq 1]}$$

This assumption works well when $L \gg P$, (resulting in $\langle R^2 \rangle \approx 2PL$).

While the simple freely jointed chain model describes the polymer as a very flexible strand, the worm like chain (WLC) model takes some more aspects into consideration. In the WLC model, the polymer is treated as a semi-flexible coil, meaning that the polymer is rigid over length scales larger than one monomer unit, and its rigidity is defined by the persistence length. There are two extremes in the WLC model. The first extreme is when the contour length is smaller than the persistence length ($L < P$), which results in that the polymer can be treated as a stiff rod. The second extreme is the other way around, when the contour length is much larger than the persistence length ($L \gg P$), which results in that the polymer behaves as a flexible random coil.

Both the freely jointed chain and the WLC model are rather simple models, and they do not consider several real physical aspects of the polymer. For example, they do not account for the energetic contributions stemming from monomer-monomer and monomer-solvent interactions, and they do not account for the fact that different parts of the polymer cannot occupy the same space simultaneously. These effects are considered in the Flory theory for semi-flexible chains, a theory that describes the most energetically favorable conformation the polymer can occupy in solution. It has also implemented an excluding volume surrounding the polymer, where no interactions can take place. The Flory theory provides the end-to-end distance as:

$$R_F \cong (\omega P)^{1/5} L^{3/5} \qquad \text{[Eq 2]}$$

where ω is the effective width of the polymer.

2.3.2.2 Polymers in confinement

When confining a polymer, an obvious parameter that has to be taken into account is the dimensions of the confining volume, D. Part of the work presented in this Thesis is based upon confining polymers to nanochannels, in which the nanochannel length is described as infinitely long while the width and depth of the channel are restricted. The dimensions of the nanochannel are commonly expressed as he root mean square of a channel cross-section ($D = \sqrt{D_h * D_w}$, where D_h and D_w is the height and width, respectively). The polymers under study are long protein-covered DNA molecules that can be treated as semi-flexible polymers, and the nanochannels used are funnel-shaped (140 x 100-800 nm^2).

Depending on the persistence length of the polymer, P, and the confining volume, D, different polymer theories can describe how the polymer behaves in different parts of a nanofunnel. When the dimensions of the channel are larger than the radius of gyration of the polymer ($R_g < D$), the semiflexible polymer chain will only weakly feel the effects of the confinement and the chain statistics will therefore

be that of a self-avoiding random walk in solution [Eq 2]. When the dimensions of the confinement fall below the radius of gyration, but are still above the persistence length ($P < D < R_g$), the polymer can be modeled as a series of non-interacting "blobs", like pearls on a string. This behavior is described by the theory of de Gennes (19), where each blob is portrayed as an unconfined, self-avoiding blob that behaves as a "Flory coil" with a diameter equal to the channel dimension, D. The end-to-end length of the polymer is given by:

$$R_{de\ Gennes} \cong L\frac{(\omega P)^{1/3}}{D^{2/3}} \qquad \text{[Eq 3]}$$

When the channel dimensions fall below the persistence length ($D < P$), the polymer moves into another specific regime, called Odijk regime (20, 21). Here, the polymer is no longer able to coil up on itself, resulting in very high bending energy while the entropy is low. The movement of the polymer is restricted to small undulations and back-folding on itself is impossible. The movement resembles that of a stretched out polymer with a self-avoiding random walk in 1D. Odijk based his theory on the assumption that the polymer consists of a number of deflections, or undulations, with the length $\lambda = (PD^2)^{1/3}$, and that these undulations have a small angle to the wall. In this regime, the end-to-end length of the polymer along a channel with rectangular cross section, such as that used in Paper III, is given by:

$$R_{Odijk} = L\left[1 - B\left(\left(\frac{D_h}{P}\right)^{2/3} + \left(\frac{D_w}{P}\right)^{2/3}\right)\right] \qquad \text{[Eq 4]}$$

where B is a constant numerically estimated to 0.091 (22).

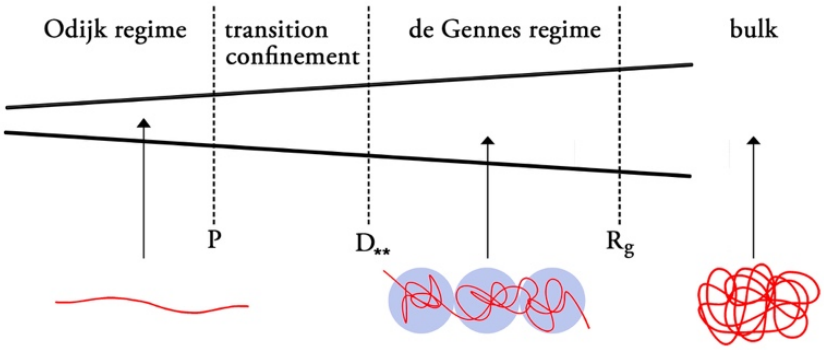

Figure 3. Overview of the different physical regimes a polymer experiences by confinement. When the confining dimensions are large ($D > R_g$), the polymer is not affected by the volume restrictions, and it is said to be in bulk. When the confinement increases, its behavior is determined by the regime in which it resides: as "blobs" in the de Gennes regime ($D_{**} < D < R_g$), as stretched out "blobs" that turns into hairpins in the transition regimes ($P < D < D_{**}$), and more or less stretched out in the Odijk regime (D< P).

Between the de Gennes regime, where the polymer consists of small Flory blobs on a string, and the Odijk regime (Figure 3), where the molecule is stretched out and not able to coil up on itself, there are two other transient regimes, in which the polymer demonstrates slightly different conformations. When increasing the confinement from the de Gennes regimes below a threshold, D_{**}, the polymer forms elongated anisotropic blobs, it is now in the extended de Gennes regime. As the polymer moves into even more confined volumes the bending energy increases. At $P < D < 2P$ the polymer enters the so-called "transition" regime where only hairpins can be formed. Increasing the confinement further causes the polymer to enter the previously discussed Odijk regime ($D < P$). In the appended Paper III, the protein-DNA polymer under study is mostly confined to the Odijk regime, but depending on the protein coverage of the DNA, the polymer moves into the nearby regimes.

2.4 DNA as a drug target

DNA is the genetic code in the cell, and it is therefore a suitable target for therapeutic agents. Cancer for example, which stems from tumor cells proliferating uncontrollably, can be treated or halted by agents that bind to, and occasionally also cleave, the DNA chain. The prolific growth of cancer cells causes them to be more vulnerable to DNA-binding agents than normal cells. Upon binding to DNA, cancer drugs can cause prevention or disruption of the replication process, which in some cases also lead to apoptosis (23). In addition to their antitumor activity, DNA-binding drugs can be used as antimalarial, antibiotic or antifungal agents, and they are also extremely useful for detecting and quantifying DNA both *in vitro* an *in vivo* (24, 25).

A DNA-binding drug has to, besides generating the wanted effects through its binding, conquer several tough obstacles before the actual DNA-binding event can occur in a cell. Firstly, it is beneficial if only selected cells receive the therapeutic effect of the drug, secondly, the drug has to be able to reach the genome located inside the cell nucleus, and, thirdly, the DNA binding should preferably be sequence specific. A sequence specific DNA-binding drug has the advantage of being able to target only certain regions in the genome. If the drug has good sequence specificity it could reduce some of the adverse side effects displayed by DNA-binding agents, such as development of a secondary cancer (26, 27).

A drug can bind either covalently or non-covalently to DNA. However, a covalent binding is usually preceded by non-covalent interactions (28–30). When a ligand binds covalently and irreversibly to the DNA, and if the machineries of the cell cannot repair this defect accurately, the result is usually cell death. For example, the small molecule cisplatin, which is used as a cancer drug, binds covalently to adjacent guanines in the DNA strand, forming an adduct. The adduct elicits the repair machinery and promotes apoptosis (31). The non-covalent binding mode is also exploited by several anticancer drugs, for example by the drugs daunorubicin and doxorubicin, which through interactions with DNA inhibit proper replication of the genome and thereby cause apoptosis (32).

There are two major classes of non-covalent DNA binding, if the electrostatic interactions with the negatively charged DNA backbone are excluded; intercalation and groove binding (Figure 4) (33). Intercalation is usually performed by small, planar and polycyclic aromatic molecules, which can stack between the DNA bases. It causes the DNA to open up, unwind and elongate, events that result in an energy cost. This energy cost is compensated for by the stable π-π interactions between the ring systems of the DNA bases and that of the intercalator. Groove binding is performed by, generally, slightly larger, crescent-shaped molecules that can, depending on their structure, bind either in the minor or major groove of the DNA. The major groove is wider and slightly shallower than the minor groove, and it is therefore usually the target for larger molecules, such as most DNA-binding proteins, while the minor groove, being more narrow, is the target for smaller molecules (7).

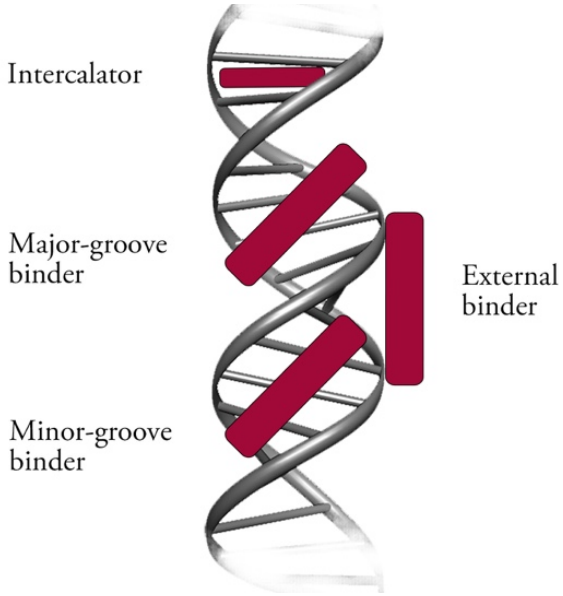

Figure 4. A schematic representation of the main non-covalent binding modes of DNA-binding molecules. The molecule could either bind externally to the phosphate backbone, intercalate between the DNA bases, or sit in one of the grooves (major or minor).

One of the intentions of the field of DNA-binding drugs is to design molecules that can somehow be delivered to the genome of malfunctioning cells where it sequence-specifically binds to the DNA. Intercalators could favor some specific base pairs, but a typical intercalator is not able to "read" longer DNA stretches since its sequence preference is dependent on the energy cost of separating the two adjacent base pairs and unwind the DNA (*e.g.* pyrimidine-purine bases are easier to unstack than purine-pyrimidine) (34). However, groove binders interact with several base pairs simultaneously and they have therefore a larger possibility to act more sequence specific (35, 36). One example is a family of synthetically produced DNA major groove binding molecules called zinc fingers, which design originates from naturally occurring motifs in proteins (37). It is a large family of different binding motifs where a zinc ion coordinates and stabilizes the fold of an amino acid sequence. (Note, the family also includes domains that can bind to RNA or proteins.) What domain the zinc finger recognizes depends on its fold and its amino acid sequence (usually around 9 bp), but several zinc finger domains can also be engineered to sit in an array in order to increase the length of the motif (38). The minor groove is a target for sequence specific DNA-binding molecules as well, and a small synthetic polyamide hairpin is a good example of such (39, 40). The polyamide folds up on itself, like a hairpin, and inserts into the minor groove where it forms specific hydrogen bonds with the DNA bases, and is thus able to recognize all four Watson-Crick base pairs (39–41). The structure and binding specificity of a more classic and typical minor-groove binder, Hoechst 33258, will be discussed in detail in the next chapter (2.4.1).

2.4.1 Minor-groove binding drugs and Hoechst 33258

Minor-groove binding drugs have a typical appearance: they are usually positively charged, crescent-shaped, and consisting of an aromatic ring system that possesses bonds with torsional freedom, thus allowing the molecule to fit snugly into the minor groove without requiring major alteration of the DNA double-helix (42). The minor-groove binding molecule comes in close contact with both the "floor" and the "walls" of the groove, and it can therefore bind rather sequence specifically by utilizing that different DNA sequences exhibit differences in their hydrogen bonding patterns, hydrophobicity and groove widths (43). The DNA minor grooves of AT-sequences are narrower than GC-sequences, thus providing a better fit for the molecule, and the typical minor-groove binder is therefore AT-specific. In addition to just being narrower, AT-rich regions have an over all more negative potential compared to GC-regions, which is favorable for positively charged DNA-binding molecules. In addition, GC-regions also have amino groups protruding from the guanines that sterically hinder the drugs from fitting deeply into the minor groove (43)

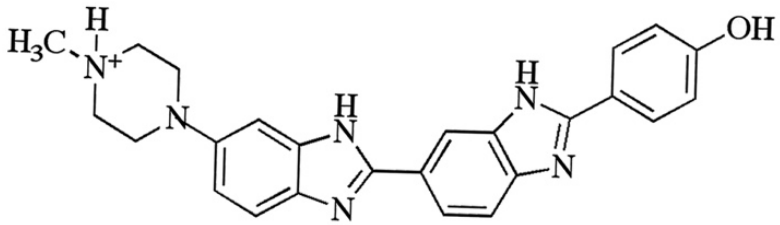

Figure 5. The structure of Hoechst 33258 at neutral pH. The positive charge on the methylpiperazine and its crescent-shaped structure make it a good minor-groove binder.

Hoechst 33258 looks and behaves like a very typical minor-groove binder (Figure 5). Its name originates from the company that invented it, Hoechst AG, and the number announces that it is the 33258th compound made by the company. The bis(benzimidazole) derivate is a typical minor-groove binder, except for the fact that it binds endothermically ($\Delta H° = 4.3$ kcal/mol (44)). The endothermic binding is compensated for by a positive entropy term ($T\Delta S > 0$) so that the total binding free energy is negative (33, 44). The positive entropy is probably due to release of structured water, both positioned in the minor groove and bound to the molecule in the bulk, as well as to the release of condensed positive monovalent counter ions (Na^+) upon ligand binding. The light emission of Hoechst 33258 is quenched in solution, but when the molecule is bound in the minor groove of the DNA, preferably to AT-rich regions, it fluoresces brightly (46–48). Because of its strong binding to DNA and low background fluorescence, Hoechst 33258 is commonly used for detecting DNA in cells and for detecting or quantifying DNA *in vitro* (49, 50). By

footprinting analysis, NMR and crystal structure studies it has been shown that Hoechst 33258 binds to an AT-region of about four base pairs (50–53), but there have been uncertainties regarding the structure of the formed molecular complex. These include: in what direction the positively charged piperazine ring points in the minor groove of several AT-rich sequences, what base pairs the ligand covers in the AT-tract, and what the structure of the Hoechst 33258-DNA complex looks like at high ligand to DNA ratios (54, 55). At high Hoechst 33258 to DNA ratios, Hoechst 33258 could potentially interact with DNA as a defined dimer, like a sandwich dimer that could in principle have sequence recognition power similar to that of the polyamide hairpin discussed in previous chapter (chapter 2.4). The binding of Hoechst 33258 to three different AT-rich oligonucleotides, at high ligand to DNA ratios, have been investigated in Paper V.

3 DNA repair

In this section DNA damage and DNA repair are addressed. The focus is on the double-strand break repair system called homologous recombination and its key protein in eukaryotes, Rad51.

3.1 DNA repair

The integrity of DNA, our genetic information carrier, is constantly under attack from internal processes, such as the generation of harmful reactive oxidative species from our natural metabolism, as well as from various external sources, such as chemical agents and UV or ionizing radiation. Every day, more than tens of thousands of DNA lesions occur in a single human cell (1–3), and if these lesions remained unrepaired, they could lead to mutations or cell death (56). Potential DNA damage could be alteration of the chemistry of the DNA molecule, for example alkylation, hydrolysis or oxidation, or it could also be a wrongly incorporated, or missing, base (1, 3). Another possibility is that the phosphate backbone of the double helix is broken; either by a single-strand break (SSB), or by a more lethal double-strand break (DSB). To deal with DNA damage, the cell has three strategies: it might do nothing, it could go through programmed cell death (apoptosis), or it could try to repair the DNA error (57, 58).

Depending on the DNA lesion, the cell type, and where in the cell cycle it occurs, the cell will handle the lesion differently (58, 59). There is a number of different repair pathways and repair machineries that the cell can utilize, yet all pathways have a common strategy: detection of DNA damage, accumulation of DNA repair factors and, finally, reparation of the lesion (60–62). The more subtle DNA lesions, such as alkylation products, oxidative lesions, mismatches or SSBs, are repaired through either the base excision repair pathway, the nucleotide excision repair pathway or the mismatch repair pathway, where on strand is damaged and the other strand works as a template to ensure that the lesion is repaired correctly. Very recently, Lindahl, Modrich and Sancar were rewarded the Nobel Prize in chemistry (year 2015) for their mechanistic research on these particular DNA repair pathways (63). The pathways all work in similar manners, but they utilize slightly different machineries to perform the reparation procedure. However, all pathways remove a section around the damaged base/bases, and the size of the section depends on what pathway is active. The missing region is thereafter replaced with a newly synthesized DNA (64). Some subtle chemical modifications can also be removed by direct chemical reversal.

The most devastating DNA damage is the DSB. Both DNA strands are broken in a DSB, and if they remain unrepaired or inaccurate rejoining of the broken strands takes place, widespread structural rearrangement of the genome occurs, which could have deleterious consequences. There are two major pathways to deal with the lethal

DSBs: homologous recombination (HR) and non-homologous end-joining (NHEJ) (65). The "choice" of when HR or NHEJ is used is still somewhat unclear, but it is shown to depend on the cell type, where in the cell cycle the break happens and how (58, 59, 66). NHEJ directly ligates the two broken ends together, and it functions throughout the cell cycle, even though G1 is preferred (66) (note: G1 is the first growth-phase in the cell cycle, which is composed of four phases: G1 (growth 1) → S (DNA synthesis) → G2 (growth 2) → M (mitosis)). However, by simply joining the broken ends together, insertions, deletions, substitutions or translocations could take place. HR, on the other hand, is a more accurate repair system. It uses a homologue DNA sequence as template for the broken DNA, thus ensuring the correct DNA sequence at the damaged site. HR is relatively slow and acts mainly in the S and G2 phase of the cell cycle (60).

3.1.1 Homologous recombination and strand exchange

In addition to repairing DSBs, HR also plays an important role in meiosis where it mediates the exchange of genetic information between maternal and paternal alleles within gametes, and where it also ensures proper segregation of the homologous chromosome pairs (61, 67). HR is thus vital for maintaining both stability and diversity in the genome. A recombinase protein is the key protein in HR, and over- or under- expression of the recombinase compromises both the integrity of the genome and the survival of the cell. Genome instability, chromosomal damage, increased cancer rate and increased chemical and radio sensitivity have been detected in cells with modified expression of the key protein (68–73). The importance of a properly functioning HR can also be pointed out through its existence throughout all branches of life; bacteria, archaea and eukaryotes have evolutionary conserved recombinases that participate in HR with very similar structure and function.

The central reaction that takes place during HR is an ATP-dependent strand exchange reaction, a reaction where a DNA strand is transferred between two homologous DNA strands. The strand exchange reaction at a DSB site (Figure 6) begins with resection enzymes creating long ssDNA overhangs that are immediately covered by a protein that is called replication protein A (RPA) in eukaryotes. (Note: meiotic recombination between to dsDNA begins with a DSB in one strand.) RPA prevents the ssDNA to back-fold or form secondary structures (74). The RPA-covered ssDNA is in turn a signal for the recombinase protein and its mediators to assemble on the site, displace the RPA and cover the ssDNA (62). The eukaryotic recombinase is called Rad51, and its orthologs in bacteria and archaea are called RecA and RadA, respectively. There is also an eukaryotic recombinase homologue that is only active during meiosis, Dmc1 (75). The recombinase proteins form a right-handed

helical filament around the ssDNA, creating a so-called presynaptic filament. In order to perform the strand exchange reaction, the presynaptic filament initiates a search for a homologous sequence in the vast amount of dsDNA in the cell. This is an extensive task, and it has lately been demonstrated for RecA that the search is not diffusion based but, however, the recognition event is faster within spatial proximity (76). It is suggested that the presynaptic filament creates multiple contacts points with the dsDNA in the search for homology, and thereafter slides along the DNA at these points (77–79). Mediator proteins are also believed to aid the search for homology (78). The initial homology test between the dsDNA and the ssDNA in the presynaptic filament is expected to be very quick and it involves about 8 bp (80, 81). If perfect homology is detected in these few base pairs, the homology test moves into the next phase that requires more bases to match in order to ensure high fidelity (81). When homology with a dsDNA is found, a joint displacement loop, or D-loop, is formed that can be resolved in various ways resulting in either non-crossover or crossover products (Figure 6) (61, 82). The actual strand exchange reaction is believed to involve about 80 bp at the time (83).

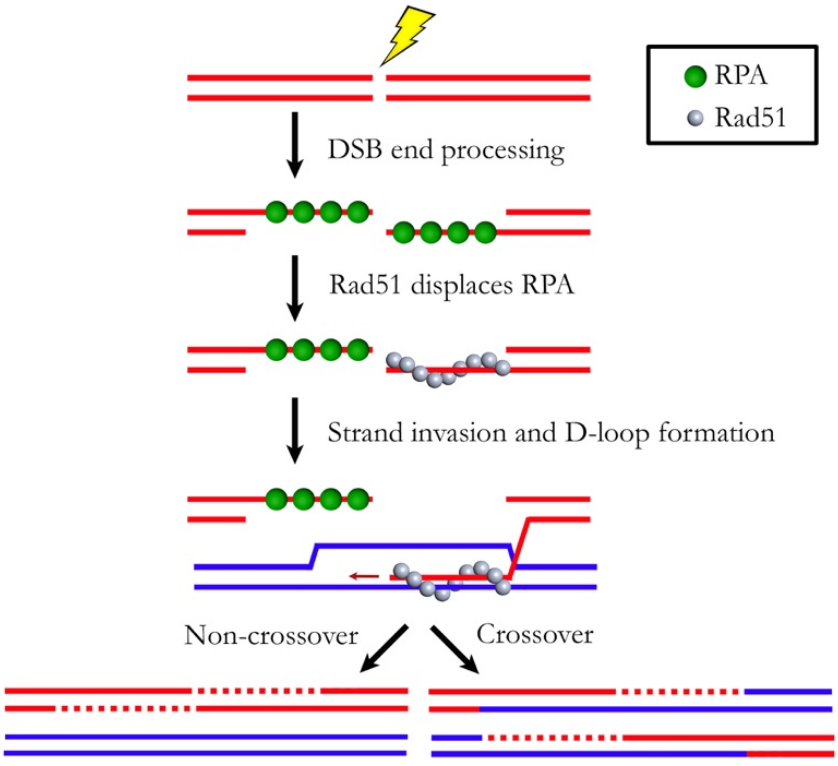

Figure 6. Schematic picture of the strand exchange reaction. At a DSB site, ssDNA overhangs are created by resection enzymes that are quickly covered with RPA. The RPA-ssDNA complex promotes formation of the presynaptic filament, the Rad51-ssDNA filament, which initiates the search for a homologous sequence in the dsDNA. When homology is found, the strand exchange reaction takes place, which results in either a non-crossover or a crossover product.

3.1.2 Rad51 protein

The name "Rad" in the protein Rad51 stems from that its gene, *rad51*, (together with several other members of the *rad51* epistasis group) was identified when mutation or deletion of the gene caused the cells to be deficient in repairing damage induced by ionizing radiation (84). Thereafter it was shown that malfunction of *rad51* resulted in defects in genetic recombination and recombinant repair of DNA damage, thus the gene could be coupled to the HR system (85, 86). Some sequences of the ATP-dependent recombination protein Rad51 are more evolutionary conserved than others and, unsurprisingly, these include the residues that are critical for the recombinase function, such as its DNA binding and ATP hydrolysis (87–89).

The human recombinase Rad51 (HsRad51) is a 37 kDa globular protein. The relative size of the protein to DNA is visualized in Figure 7, where it is demonstrated how four protein monomers bind to three DNA bases each of a ssDNA, forming a right-handed filament with six protein monomers per helical turn (90–92). The pitch and width of the Rad51-DNA filament is dependent on if ATP or its hydrolysis product ADP is bound in the interface between the protein monomers, *i.e.* if the filament is in its active or inactive form, respectively (93). When ATP is bound, the DNA is stretched about 50% compared to normal B-form DNA (91, 94), resulting in that both the pitch and the filament width are approximately 100 Å (91, 95). The stretching of the DNA within the nucleoprotein filaments is mainly achieved by a rise between every third base/bp (~7.8 Å), resulting in that the DNA in a nucleoprotein filament is heterogeneously stretched with triplets that display a more or less normal B-conformation (96, 97). For the filament to dissociate, the bound ATP has to be hydrolyzed to ADP (98). When ADP is bound, the filament exhibits a more compressed, inactive, form that has a shorter pitch and slightly larger width (95). In addition to ATP, the Rad51-DNA filament also requires presence of divalent cations, Mg^{2+} or Ca^{2+}. It has been demonstrated that the presence of Ca^{2+} promotes a more stable filament than that formed with Mg^{2+} (93), probably because Ca^{2+} inhibits the ATP hydrolysis, causing the Rad51-DNA filament to reside in its active form. The structural effects that Ca^{2+} and Mg^{2+} exert on the presynaptic filament have been examined in Papers I, II and III.

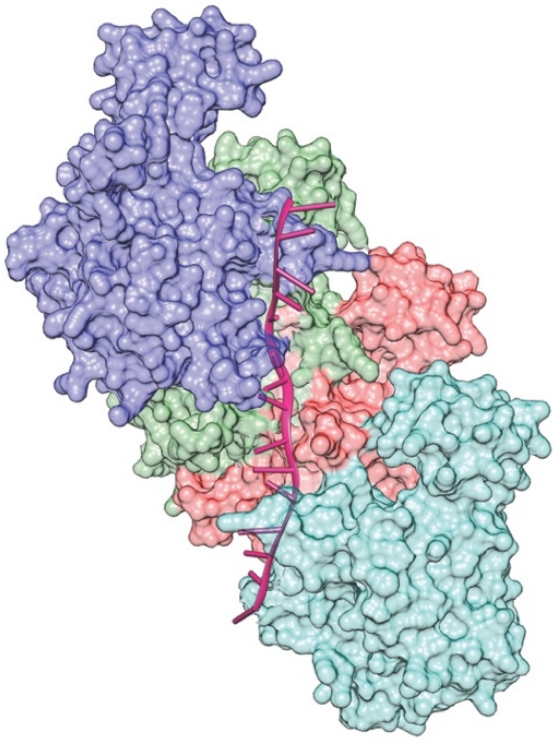

Figure 7. Four HsRad51 monomers bind to a ssDNA in the presence of Mg^{2+} and ATP (molecular model adopted, and modified, from Paper I). It can be visualized how the bases of the DNA strand are structured in triplets.

Rad51 filament formation on DNA is a complex process. *In vivo*, the DNA substrate for Rad51 is ssDNA, while, *in vitro*, the protein is able to form filaments on dsDNA as well (89). As previously mentioned, the structure of the Rad51-DNA filament is dependent on if ATP or ADP is present and the choice of cation (Mg^{2+} or Ca^{2+}). However, the macroscopic structure is also dependent on the concentration of Rad51 monomers during the initial filament formation phase. Filament formation can be divided into two distinct phases, nucleation and polymerization (99, 100), of which the nucleation phase has shown to be dependent on the Rad51 concentration. *In vitro,* it has been demonstrated that low Rad51 concentration results in fewer nucleation sites, which involves approximately 5-6 monomers, compared to when the Rad51 concentration is higher (98). The fewer nucleation sites formed at low Rad51 concentration result in lesser, but probably longer, continuous Rad51 filaments along the DNA strand. These individual Rad51 filaments formed on a DNA strand are usually referred to as Rad51 patches. In Paper III, it is studied how the length of the HsRad51 patches and the naked DNA between the patches are correlated to the structure and dynamics of the macroscopic structure of the HsRad51-DNA filament.

3.1.3 Accessory proteins and Swi5-Sfr1

While Rad51 can perform the strand exchange reaction *in vitro* on its own, there are a number of accessory proteins involved in the process *in vivo*. The recombination process is a balance between accessory proteins that mediate the reaction and those that inhibit it (82, 101). The experiments performed in this Thesis are all performed *in vitro*, thus the competing balance between positive and negative mediators is absent and Rad51 is indeed able to perform the strand exchange reaction by itself. Some of the major accessory proteins involved in the eukaryotic recombination process *in vivo* will be presented here, but the focus is on one particular complex called Swi5-Sfr1.

In the beginning of the recombination process, ssDNA overhangs are created and the eukaryotic protein RPA binds to these with high affinity in order to prevent any secondary structure (102–104). Rad51 is unable to form filaments on the RPA-coated ssDNA on its own, and mediators are therefore required for proper Rad51 filament formation. Rad52 aids the displacement of the RPA and the loading of Rad51 onto ssDNA (105–107). Another mediator, BRCA2, interacts with both ssDNA and Rad51 and it aids the presynaptic filament formation by preventing ATP hydrolysis (108), thus stabilizing the Rad51-ssDNA filament. The mediators Rad55-Rad57 are also involved in aiding the formation of the ssDNA-Rad51 filament (104).

The heterodimer Swi5-Sfr1 (109) has a mechanistic stabilizing role for the Rad51-ssDNA filament. It is a rather large, crescent-shaped complex (110) that fits well in the groove of the helical Rad51-filament (111), thereby stabilizing it (112, 113). When positioned in the groove, Swi5-Sfr1 interacts with up to 10 Rad51 monomers (110), and it can stimulate both Rad51- and Dmc1-mediated strand exchange. *In vivo,* it has been demonstrated for model organisms (fission yeast *Schizosaccharomyces pombe* and mouse) that Swi5-Sfr1 is needed for proper strand exchange (109, 112), but that it also requires Rad22 (the Rad52 homologue in *S. pombe*) for its function and vice versa (112).

When the strand exchange has taken place, the Rad51-dsDNA filament has to dissociate, which in addition to ATP hydrolysis (92, 114–116) requires involvement of accessory proteins *in vivo* (117). For the bacterial ortholog RecA, the disassembly of the filament is coupled the ATP hydrolysis (118), but Rad51 has shown to be bound tighter to its ADP-accommodating product than RecA (98, 119, 120). The key protein responsible for the dissociation of the filament is the translocase Rad54 (117), but Swi5-Sfr1 seem to aid the dissociation as well (121). The structure of the fission yeast Swi5-Sfr1-Rad51-ssDNA complex and the effect that Swi5-Sfr1 has on the strand exchange reaction have been studied in Paper II.

4 Methodology and fundamental concepts

In this section the methods utilized in the Thesis are presented together with some fundamental background concepts. The main field of method exploited throughout the Thesis is spectroscopy of various types, but nanofluidics, thermodynamic measurements and protein expression and purification have also been utilized.

4.1 Spectroscopy

Spectroscopy is the study of interactions between matter and electromagnetic radiation, *i.e.* light. Light is described in classical physics as a wave, but it can also be treated as particles with defined energy quanta. When describing light as a wave, it is comprised of an electric and a magnetic field that harmonically oscillate perpendicular to each other as well as to the direction of propagation. It is characterized by the frequency of the oscillating fields, ν, which in turn is related to the wavelength, λ, and speed of light, c (c = 3 x 10^8 m/s), as:

$$\lambda = c/\nu \qquad \text{[Eq 5]}$$

When light is described as particles, photons that carry discrete energy packages are considered. Interaction between a photon and a molecule results in either scattering or absorption of the photon, and the molecule may also, as a result of absorbing energy, emit another photon. A prerequisite for a molecule to absorb a photon is that the incoming (and absorbed) energy corresponds to the energy difference between two energy states in the molecule (ΔE), according to the Bohr frequency condition:

$$\Delta E = E_{higher} - E_{lower} = h\nu \qquad \text{[Eq 6]}$$

where h is Planck's constant. If the wavelength of the incoming light is in the UV or visible region, such an interaction could cause electronic transitions within a molecule. If the wavelength is longer, thus if the incoming light has lower energy, an interaction could cause vibrational or rotational transitions. Yet, in order for an electronic transition to occur, the electric field component of the light has to constructively interfere with the transient oscillations of the charge density in the molecule as it changes state. Interference gives rise to an induced transient dipole moment, described by a quantum property known as the electric dipole transition moment that determines the probability of the electronic transition i→f to occur as:

$$\mu_{fi} = \int \psi_f^* \hat{\mu} \psi_i \, d\tau \qquad \text{[Eq 7]}$$

where ψ_f^* and ψ_i are the wave functions of the final and initial states, respectively, and $\hat{\mu}$ is the dipole operator.

4.1.1 Absorption spectroscopy

Absorption spectroscopy measures how well an ensemble of molecules absorbs photons of a specific wavelength, λ. The measurement is usually performed on molecules in solution in a spectrophotometer (Figure 8), where the light intensity is measured before (I_0) and after (I) it has moved through the sample, providing a measure of the absorbance as:

$$A(\lambda) = log(I_0(\lambda)/I(\lambda)) \qquad \text{[Eq 8]}$$

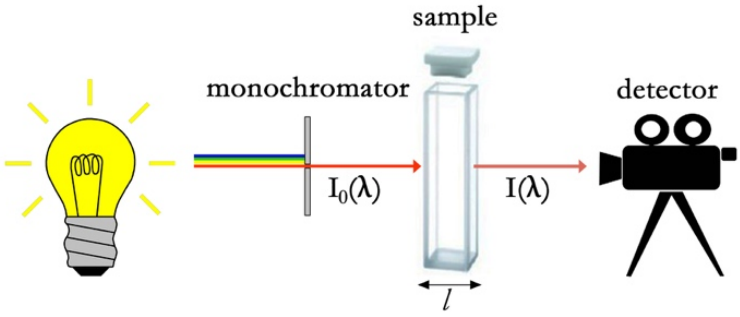

Figure 8. The basics of a spectrophotometer. A light source emits light, a monochromator permits selected wavelengths ($I_0(\lambda)$) to pass to the sample, which is typically in solution in a cuvette, and the transmitted light ($I(\lambda)$) is thereafter detected.

The measured absorption (A in [Eq 8]) is directly proportional to an experimentally determinable parameter called the extinction coefficient ($\varepsilon(\lambda)$, typically given in $M^{-1}cm^{-1}$), the concentration of absorbing molecules (c), and the path length (l), as demonstrated by the Beer-Lambert law:

$$A = \varepsilon(\lambda)cl \qquad \text{[Eq 9]}$$

By the use of the Beer-Lambert law, absorption spectroscopy provides a straightforward way to determine sample concentrations as well as to identify unknown samples. This relatively simple, yet powerful, technique can also be used to characterize molecular properties and to follow or quantify chemical reactions. Noteworthy, ε depends not only on the wavelength and the molecule, but also on the molecular environment.

4.1.2 Polarized light spectroscopy

Light consists of a magnetic and an electric wave oscillating perpendicular to each other and to the direction of propagation. For "common" white light, or non-polarized light, all oscillating directions are equally probable. This chapter will discuss two techniques used in this Thesis involving polarized light, *i.e.* light where

only certain oscillating direction(s) are allowed and the polarization is defined by the direction of the electric field. The two techniques are linear dichroism (LD) and circular dichroism (CD), and they will be presented briefly below. For more details about the techniques and their implementations, reference (122) is recommended.

4.1.2.1 Linear dichroism spectroscopy

Linear dichroism, or LD, is defined as the difference in absorption of light that is polarized parallel (∥) and perpendicular (⊥) to an orientation axis:

$$LD = A_{\parallel} - A_{\perp} \qquad \text{[Eq 10]}$$

LD can be applied to systems that are either intrinsically oriented or oriented during the experiment. Biomolecules that display intrinsically oriented features involve for example several components of the cell, such as DNA and membranes. However, they still have to be macroscopically oriented for an LD measurement to be possible.

The probability of an electronic transition to occur within a molecule depends on the angle, Θ, between the transition dipole moment of the molecule and the electric field of the light as:

$$A \propto \left|\mu_{fi}\right|^2 cos^2\Theta \qquad \text{[Eq 11]}$$

As the above equation states, there is maximum absorption probability if the light polarization and the induced transition moment are parallel to each other ($\Theta = 0°$), and, conversely, no absorption takes place if the angle is perpendicular ($\Theta = 90°$). In order to utilize LD, an isotropic sample has to be oriented. This can be done for smaller molecules by incorporating them into stretchable polymer films, which causes the small molecules to align in the stretching direction (Figure 9A). Orientation for larger molecules in solution can be achieved by shear flow in a couette cell (Figure 9B). The cell consists of two cylinders, one smaller positioned inside a larger, with a narrow gap between them where the sample solution is kept. One of the cylinder rotates, creating shear flow in the solution that causes the molecules to align in the direction of the flow. The two mentioned orientation methods have been used in this Thesis. It is also possible to align the sample by electric or magnetic fields.

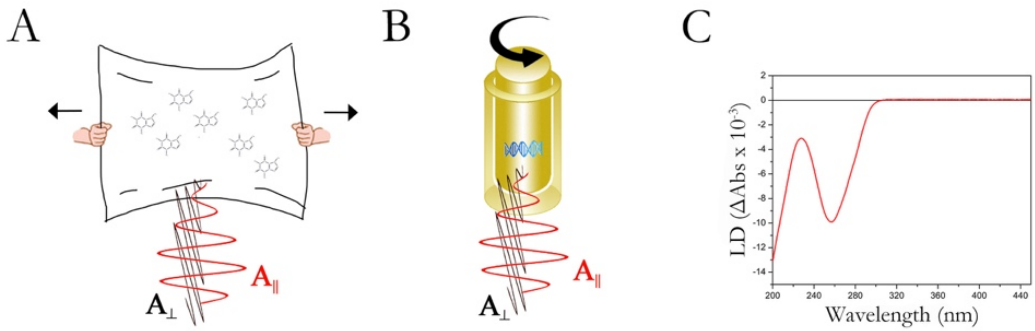

Figure 9. (A) and (B) demonstrate two examples of techniques to align molecules for LD measurements. Smaller molecules can be aligned by incorporating them into stretchable polymer films (A). The small molecules align in the stretching direction of the film. Longer polymers, such as DNA, can be aligned in a couette cell (B). The couette cell has two coaxial cylinders with a small gap in between (approx. 0.5 mm). One of the cylinders rotates, creating a laminar flow in which sample is oriented. (C) shows an example of an LD spectrum of DNA The negative LD signal at 260 nm corresponds to the transition moments of the DNA bases that are perpendicularly aligned relative to the orientation axis ($A_\perp > A_\parallel$).

By normalizing the LD signal with respect to the isotropic absorption (A_{iso}) the reduced LD (LD^r) is obtained, a quantity that has the advantage of being independent of path length and concentration:

$$LD^r = \frac{LD}{A_{iso}} \qquad \text{[Eq 12]}$$

If the macroscopic orientation, S, of the molecule is known, the direction of the transition moment relative to the orientation axis, Θ, can be deduced, or vice versa:

$$LD^r = \frac{3}{2} S(3cos^2\Theta - 1) \qquad \text{[Eq 13]}$$

The above equation can be applied when there are no overlapping electronic transitions. S is 1 for a perfectly orientated sample and 0 for an unoriented sample. In the special case when there is a uniaxial orientation distribution of the molecules following equation is valid:

$$A_{iso} = \frac{1}{3}(A_\parallel + 2A_\perp) \qquad \text{[Eq 14]}$$

LD spectroscopy is a powerful technique for analyzing the structure and functionality of organized systems. DNA is a regular biomolecule that has UV-absorbing transition moments oriented in the plane of the DNA bases stemming from $\pi \rightarrow \pi^*$ transitions. Thus, when the DNA in a sample is macroscopically oriented, tentatively by shear flow as shown in Figure 9B, the bases will be organized preferentially perpendicular to the orientation axis resulting in a negative LD ($A_\perp > A_\parallel$) in the region where the bases absorbs (around 260 nm) (Figure 9C). If a

hypothetic small and planar ligand with its transition moment oriented along its long-axis (as a typical DNA ligand) is externally bound to DNA in a regular fashion, the LD signal arising from its binding could provide information about the binding mode. A negative LD signal reveals that the transition moment of the ligand is parallel to DNA bases, such as it is upon intercalation, while a positive LD signal demonstrates that the transition moment is preferentially parallel to the DNA axis, a groove binding of the ligand may then be more possible.

However, LD spectroscopy is not as straightforward when many identical chromophores have different orientations, as they do in for example a protein. If the absorptions of several chromophores are identical, the LD of the sample will be the average of the overlapping LD signals of the individual chromophores, and the angle Θ in [Eq 13] will thus be the average direction for all transition moments. However, if the absorptions of the different chromophores are dissimilar, the LD^r will vary with wavelength and the angles for the individual transition moments could be deduced. For example, if we assume two similar chromophores (denoted 1 and 2) in an oriented protein [Eq 13] will become:

$$LD^r = \frac{3}{2}S\left[\frac{(3cos^2\theta_1 - 1)A_1(\lambda) + (3cos^2\Theta_2 - 1)A_2(\lambda)}{A_1(\lambda) + A_2(\lambda)}\right] \quad \text{[Eq 15]}$$

If the absorption curves A_1 and A_2 are not the same at different wavelengths, say if the absorptions of the chromophores are shifted due to environment, and the orientations of the two chromophores are different, *i.e.* $\Theta_1 \neq \Theta_2$, the LD^r ratio will vary with wavelength. Knowledge on how the environment may cause shifts of the absorption bands and/or changes in the directions of the transition moments could be used to disentangle protein structures containing many identical chromophores. This is the incentive behind the study of the transition moments of the chromophore of tyrosine conducted in Paper IV. The LD technique was used thoroughly in Papers I, II and IV.

4.1.2.2 Circular dichroism spectroscopy

Circular dichroism, or CD, is defined as the difference in absorption of left and right circularly polarized light:

$$CD = A_l - A_r \quad \text{[Eq 16]}$$

As opposed to LD, where the electric field is oscillating in the same plane, the magnitude is constant in CD, while the electric field vector traces out either a left-handed or a right-handed helix about the direction of propagation. CD is used to study chiral molecules since they absorb the orthogonal forms of circularly polarized

light differently, producing a CD signal. Achiral molecules, on the other hand, do not demonstrate any absorption difference of circularly polarized light and have, hence, zero CD signal. CD is particularly useful for studying several biological systems, such as DNA or proteins; two systems that are both intrinsically chiral. Interestingly, an achiral molecule can display an induced CD signal upon incorporation into a chiral environment, for example an achiral ligand binding to the chiral DNA molecule. Also, when two CD-producing chromophores are close to each other in space, shifts or splits in the CD spectrum may arise depending on their relative orientations and distances to each other that result in a specific CD "signature". This feature is exploited in Paper V, where the binding of Hoechst 33258 to DNA has been examined.

4.2 Fluorescence spectroscopy

The event of absorption has been discussed the previous chapter. However, after the molecule has absorbed a photon of appropriate energy (predicted by [Eq 6]), the energy gained can disperse in various ways. The possible transition routes between the electronic states of a molecule can be pictured in a Jabłonski diagram (Figure 10).

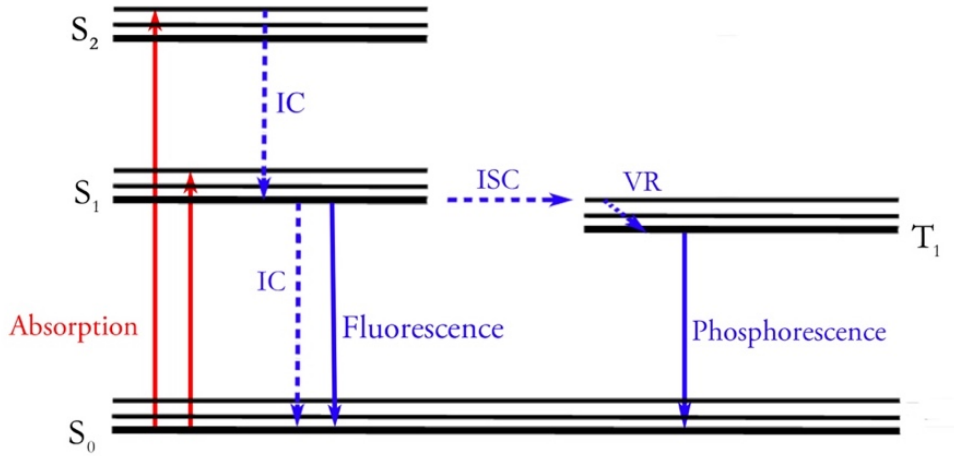

Figure 10. A Jabłonksi diagram demonstrates, highly simplified, the pathways a molecule can take upon excitation. The solid red and blue arrows show absorption and emission involving a photon, respectively, while the dotted arrows are non-radiative processes. IC is internal conversion between two states of the same spin, ISC is intersystem crossing between two states of different spins, and VR is vibrational relaxation. VR is a rapid event that puts the molecule in its lowest vibrational level in an electronic state, preceding the other events (IC, ISC, fluorescence and phosphorescence)

A molecule is usually found in its lowest electronic singlet state, S_0, the so-called ground state. Also, most of the molecules are in their lowest vibrational level (three vibrational levels are displayed in Figure 10). When a molecule absorbs a photon it is

excited to a higher electronic state with the same spin (S_1 or S_2 in Figure 10). From here, the molecule can return to its ground state in various ways. As in all electronic states, it moves quickly *via* vibrational relaxation (VR) to the lowest vibrational level. Thereafter, the energy can dissipate by a radiative or a non-radiative pathway back to S_0. Depending on the molecule and its environment (*i.e.* on what time-scales the competing decays happen), non-radiative internal conversion (IC) to a lower singlet state can take place, or a radiative process called fluorescence may happen. There is also an additional route for the excited molecule which is less common; the excited molecule could change spin of one electron and undergo non-radiative intersystem crossing (ISC) to a triplet state, and from here, after the common vibrational relaxation, the molecule may return to a lower singlet *via* a slow spin-forbidden radiative process called phosphorescence.

4.2.1 Fluorescence microscopy

Light microscopy is used to obtain a magnified image of a sample, and the image stems from light that has been transmitted through or reflected from the sample (Figure 11A). Fluorescence microscopy is very similar to light microscopy, the major difference being that it uses light emitted by the sample *via* fluorescence to generate the image. The sample is illuminated by a light source (*e.g.* mercury, laser or LED lamp) with a certain wavelength that excites fluorophores within the sample, which in turn emit photons of a slightly longer wavelength. The wavelength shift between the absorbed and the emitted light is called Stoke's shift, and it has to be large enough so that it is possible to separate the emitted light from the exciting light. When the emitted light is filtered out, only objects that fluoresce will be detected, which results in an extremely good contrast to the black non-fluorescent background. There are several different fluorescence microscopes, but the simplest one, which is also the basis for other more advance fluorescence microscopes, is the epifluorescence microscope. Here, both the excitation and emission radiation travel through the same light path, *i.e.* the objective. An emission and an excitation light filter together with a beam-splitting dichroic mirror perform the filtering process in epifluorescence microscopes.

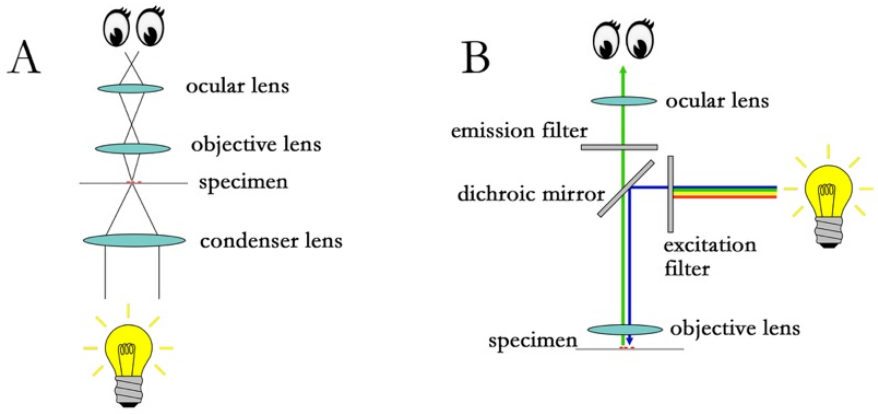

Figure 11. The basics of light microscope (A) and fluorescence microscope (B) are displayed schematically. In a light microscope, the specimen is illuminated with visible light, which is focused by a condenser on the sample, the transmitted light is collected by the objective lens that focuses a real image of the specimen inside the microscope, and the ocular lens then magnifies the image. The fluorescence microscope has the same basic principle, but the light is not passed through the sample, instead the sample itself gets excited and emits light. The filtering process for an epifluorescent microscope involves two filters and a dichroic mirror that work in concert to separate the exciting light (here blue) and the emitted light (here green) from each other.

The smallest distance two objects can be separated and still be detected as two independent units is given by the resolution limit, d. In 1873 the diffraction-limited resolution was proposed by Ernst Abbe, which was a couple of years later refined by Lord Rayleigh, stating that it is impossible to resolve two elements closer than about half of the wavelength of the light. A more exact limit (in the x,y-plane) is given by Abbe's criterion (123):

$$d = \frac{\lambda}{2n \sin \alpha} \qquad \text{[Eq 17]}$$

where $n \sin \alpha$ in the denominator is called the numerical aperture (NA). n is the refractive index of the immersion medium between the sample and the objective (*e.g.* oil=1.52, water=1.33, air=1.00), and α is half the aperture angle under which the point source is observed.

Due to the diffraction limit, basic light spectroscopy has visualization limitations; two objects have to be separated more than 100-200 nm in order to be detected as two distinct units. However, the Nobel Prize in Chemistry 2014 was awarded to Betzig, Moerner and Hell for their "development of super-resolved fluorescence microscopy", which allows capturing of images with a higher resolution than Abbe's limit (124). The basic procedure of the two super-resolved fluorescence techniques, called STEM and STORM, is to collect multiple images containing a small number of fluorophores at distances larger than the diffraction limit and thereafter pile the images together into a super-resolved image.

4.2.2 Fluorescence microscopy combined with nanofluidics

Fluorescence microscopy has provided the ability to receive plenty of information regarding reactions and structures of both ensembles of cells (tissues) as well as organelles within a single cell. It has also provided the ability to study single molecules. (Even without the use of the newly developed super-resolution methods mentioned in the previous chapter.) Long single molecules in solution, for example DNA, can be stretched out by confining them to nanochannels. Obviously, the degree of stretching of the molecule can be manipulated since it is correlated to the confining volume. The molecule is readily studied by fluorescence microscopy by attaching, covalently or non-covalently, fluorescent probes. Confining polymers to nanochannels allows them to be studied at lower forces than, for example, magnetic or optical tweezers (125). This opens up the possibility to detect characteristics that are only present at low degree of stretching. Several studies of DNA have been conducted by combining these two techniques, revealing information about the physical properties of the DNA molecule (16, 126). Additionally, the techniques can also provide new ways to quickly and crudely sequence DNA (127, 128).

The nanochannel chips used in the nanofluidic setup in this Thesis are made out of fused silica and the shape of the channels can be customized depending on the study. In Paper III, DNA coated with HsRad51 was studied by confining the nucleoprotein filaments to funnel-shaped nanochannels, which enabled the same individual filament to be examined at different confinements (Figure 12). A drawback for nanochannels when studying proteins, and positively charged DNA-binding proteins in particular, is the non-specific binding of positively charged, or hydrophobic, residues to the negatively charged silica walls. When solely negatively charged DNA is under study, the wall repels the molecule and it does not stick. The hurdle with non-specific sticking of the protein to the silica surface is overcome by passivating the walls with a lipid bilayer (129), which is spontaneously formed when flushing the chip with lipid vesicles at high pressure.

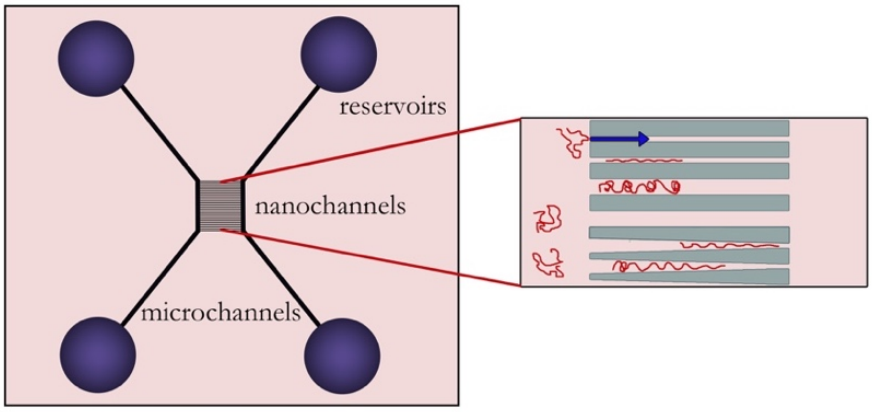

Figure 12. Schematic figure of the nanofluidic chip used in Paper III. Two reservoirs are connected to each other *via* a microchannel, and an array of nanochannels connects the two microchannel systems. The nanochannels can be constructed with various dimensions and designs (different widths and funnel-shaped design depicted in the cartoon), depending on what feature and what polymer that is under study.

In order to detect the sample in the nanochannels it has to be fluorescently labeled. Commonly, DNA is detected by adding the intercalating dye YOYO to the sample, which fluoresces when bound to DNA. In the appended Paper III, HsRad51-covered DNA is studied, and instead of labelling the DNA, fluorescently marked protein is used. To not affect the protein-DNA or protein-protein interactions, an amino acid on the surface of HsRad51 (cysteine 319) was chosen to be labeled with a commercially available fluorophore (Alexa-555).

4.3 Thermodynamics

The field of thermodynamics is based upon three fundamental laws. These laws will be presented in this chapter, together with some equations that have been exploited in the Thesis. The first law states that the energy for an isolated system is constant, thus the energy can be transformed but never destroyed. The law can be formulated as, for a closed system, that the change in internal energy (ΔU) is the combination of added heat (Q) and work (W) exerted on it:

$$\Delta U = Q + W \qquad \text{[Eq 18]}$$

The first law does not account for the fact that natural processes have a preferred direction of progress. This, however, is implemented in the second law of thermodynamics, where also the quantity entropy (S) is introduced. Entropy is the measure of how many ways a system may be arranged, and it is popularly known as the measure of disorder, or chaos. The second law states that the entropy of an isolated

system can only increase (and tend to do so to its maximum value). For a reversible process in a closed system, S is defined as:

$$\mathrm{d}S = \frac{dQ}{T} \qquad \text{[Eq 19]}$$

where T is the temperature of the system. For a process to be spontaneous, it has to fulfill the following criteria, called the Clausius inequality:

$$\mathrm{d}S - \frac{dQ}{T} \geq 0 \qquad \text{[Eq 20]}$$

Another useful thermodynamic quantity is enthalpy, H, which is defined as the internal energy plus the product of pressure and volume ($H = U + pV$). At constant pressure, and no additional work ($W = 0$ in [Eq 18]), dQ can be substituted by dH. Then, Eq 20, which states if a reaction is spontaneous or not, can be presented in terms of Gibbs free energy, G. Gibbs free energy is a thermodynamic quantity that determines if a reaction is spontaneous ($\Delta G < 0$) or not ($\Delta G > 0$). The following equation, defining Gibbs free energy, is valid at fixed temperature:

$$\Delta G = \Delta H - T\Delta S \qquad \text{[Eq 21]}$$

The changes in Gibbs free energy also demonstrates that a process with unfavorable enthalpy change ($\Delta H > 0$) could still be spontaneous if the entropy change is high enough ($T\Delta S > \Delta H$). For a system in equilibrium and at standard conditions, the standard reaction Gibbs energy, $\Delta G°$, can be utilized, which is defined as the difference in standard molar Gibbs energies between the reactants and products. $\Delta G°$ is commonly presented in relation to the equilibrium constant, K, of the system:

$$\Delta G° = \Delta H° - T\Delta S° = -RTlnK \qquad \text{[Eq 22]}$$

where R is the gas constant.

Finally, to close the circle, the last of the three laws of thermodynamics states that the entropy of a perfect crystal at the temperature absolute zero, is zero.

4.3.1 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) provides thermodynamic information about molecules interacting with each other, normally a macromolecule and a smaller ligand. It measures the changes in heat that is directly coupled to the enthalpy changes of the system (ΔH). The binding affinity, K_d, can also be determined by using Eq 22, and the binding stoichiometry, n, may also be received.

A typical experimental setup consists of two identical cells, one reference cell containing buffer or water, and one sample cell in which, usually, the macromolecule is kept, for example DNA (Figure 13A). When a ligand is titrated in small aliquots into the sample cell, the heat that arises from all events in the sample cell, mainly the interactions between the two molecules, is recorded. The working principle is to keep the reference and the sample cells at thermal equilibrium at a fixed temperature, and in order to do so, thermal power (μcal/s) has to be added to or retrieved from the sample cell during the experiment. The raw data is presented as a series of peaks showing the thermal power *versus* time (Figure 13B, upper panel), and the integrated area of each peak normalized to the sample concentrations provides the heat evolved per mole injectant (Figure 13B, lower panel).

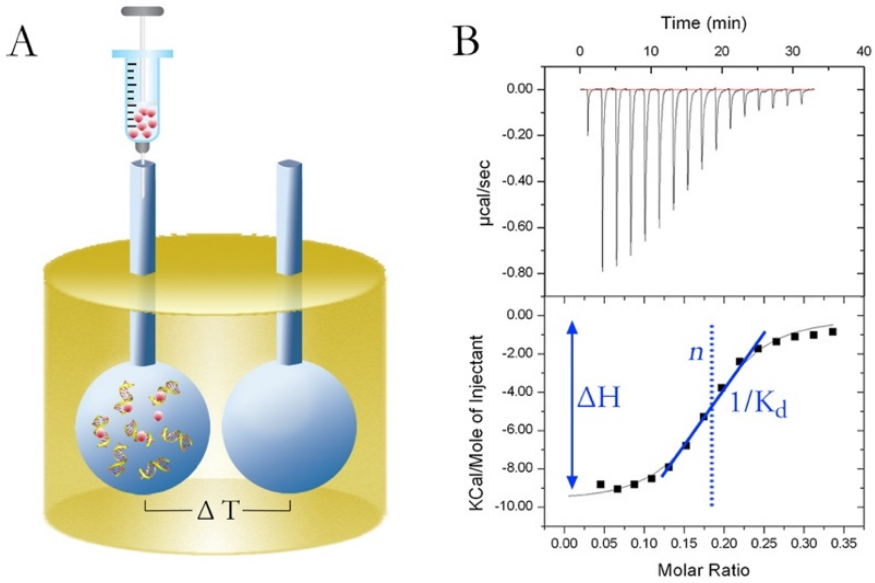

Figure 13. Illustration of the fundamental concepts of the ITC technique. Two identical cells, one sample cell and one reference cell, are kept in an adiabatic environment at a constant temperature (A). The reference cell holds the same solvent as in the sample cell (but without sample), while the sample cell usually holds the macromolecule (here DNA). A concentrated injectant (here red spheres) is titrated in small aliquots into the sample cell. The power applied to the sample cell in order to keep ΔT between the two cells minimal is measured and presented as a thermogram as a series of peaks (B, upper panel). The peaks are negative if the interaction is exothermic, and positive if it is endothermic. Integration of the peak area and correction for the sample concentrations gives the binding isotherm that presents the heat evolved per mole injectant for each injection (B, lower panel), from which the enthalpy (ΔH), the dissociation constant (K_d) and the binding stoichiometry (n) can be obtained. The example shown in (B) is data from a titration of ethidium bromide (458 μM) into DNA of mixed sequence (274 μM base pairs calf-thymus DNA).

In Figure 13B, the binding isotherm of a titration of ethidium bromide into calf thymus DNA is shown. This system is more or less ideal; the ligand has a simple DNA binding mode (intercalation) at low ligand to DNA ratios, it binds with high affinity and the binding is highly exothermic resulting in easily detected signals. The concentrations of the molecules under study should be chosen so that all injected agents are bound in the first few injections, resulting in a plateau in the beginning of the binding isotherm (Figure 13B, lower panel). The binding sites on the DNA are gradually occupied for each injection of the ligand, causing the curve to change in a sigmoidal manner, finishing off by only detecting the heat from the ligand being diluted in the buffer. The heat of dilution is subtracted from the binding isotherm, and the corrected ITC binding isotherm can provide a complete thermodynamic profile. In this rather simple example; the initial plateau gives the ΔH of the system, the slope of the curve provides the binding constant, K_d, and the center of the sigmoidal curve, when half of the ligand has bound to the macromolecule, denotes the binding stoichiometry, n (Figure 13B). The changes in Gibbs free energy, ΔG, and the entropy, ΔS, can also be calculated (from equation [Eq 22]), and if the titration is performed at several different temperatures, the heat capacity, C_p, can be deduced as well. Simple binding modes can easily be fitted to the binding isotherms mentioned, while more complex interactions, such as cooperative or anti-cooperative behavior, requires a more intricate analysis.

The ITC technique was exploited when analyzing the DNA-binding ligand Hoechst 33258 in Paper V. The received thermodynamic profile of the binding event of Hoechst 33258 to DNA was rather complex, and a more intricate analysis than that explained above had to be conducted.

4.4 Protein expression and purification

Human Rad51 (HsRad51) used in Paper I was expressed in *E. coli* BL21(DE3) cells containing pLysE plasmids, a bacterial strain that is commonly used for expression of proteins. Figure 14 shows a simplified scheme of the expression and purification procedure of the HsRad51 protein. The vector plasmid, pT7 carrying the gene encoding for HsRad51, was transformed into calcium competent cells through heat shock. The cells were grown at room temperature and the protein expression was induced by addition of IPTG at OD_{600}=0.5. The cells were harvested after approximately 18 hours. Normally, 4 liter of media produced around 15 g cell mass that usually resulted in approximately 7.5 mg pure HsRad51.

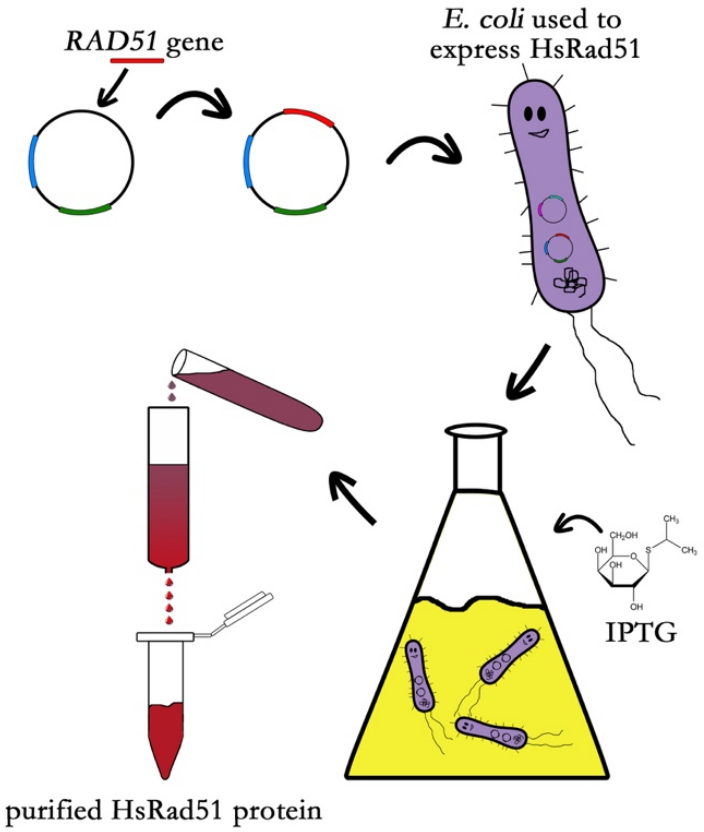

Figure 14. Schematic, and simplified, figure of the expression and purification of the HsRad51 protein. An expression vector, *i.e.* a plasmid containing a T7 promoter upstream of the *RAD51* gene as well as a gene for antibiotic resistance, is inserted into a host cell, in this case *E. coli* BL21(DE3) cells containing pLysE plasmids. By antibiotic selection, only the bacteria containing the expression plasmid are grown and the protein expression is thereafter induced by addition of IPTG. IPTG is a non-hydrolysable analogue of allolactose that displaces the lac operon repressor, thus induces the HsRad51 expression. The *E. coli* cells are thereafter lysed, and the HsRad51 protein is purified in several steps.

The HsRad51 protein was purified according to (57), and the concentration of the pure HsRad51 was thereafter determined by a colorimetric assay where the protein is colored through a chemical reaction and the absorbance at a particular wavelength is proportional to the protein concentration. A protein of known concentration, here bovine serum albumin, was used as a reference.

5 Results

In this section the main results from Papers I-V are presented and commented on. For a complete overview of the results and the experimental procedures, the reader is referred to the appended papers.

5.1 Aim

The major part of this Thesis focuses on increasing the understanding of the binding of the filament-forming protein Rad51 to DNA. The main questions that I focus on are:

- How does the choice of divalent cation, Mg^{2+} or Ca^{2+}, and the presence of the accessory protein Swi5-Sfr1 affect the structure of the Rad51-ssDNA filament? (Papers I and II)
- Does the choice of divalent cation (Mg^{2+}/Ca^{2}), DNA substrate (ssDNA/dsDNA) and Rad51 protein concentration affect the macroscopic filament structure? (Paper III)
- Is it possible to gain more structural information about Rad51, or any other protein, through polarized light spectroscopy techniques by better knowledge of the electronic transitions involved? (Paper IV)

In addition, the DNA minor-groove binder Hoechst 33258 has also been studied, and the question I address is:

- How does Hoechst 33258 bind to AT-rich DNA sequences at high drug to DNA ratios? (Paper V)

5.2 Microscopic structure of the Rad51-DNA filament

In this section, the first question above is addressed. I correlate the choice of Mg^{2+} or Ca^{2+} as divalent cation, and the presence or absence of the accessory protein Swi5-Sfr1 to the structure of the Rad51-ssDNA filament. In particular, the organization of the DNA bases within the filament is analyzed. The Mg^{2+}/Ca^{2+} study is performed on human Rad51 (HsRad51), while the study with Swi5-Sfr1 is performed on the protein homologue from fission yeast, *S. pombe* (SpRad51). A summary of the main results is presented, and the complete studies are found in Papers I and II.

5.2.1 LD spectroscopy on Rad51 filaments formed on poly(dεA)

The formation of the presynaptic filament, *i.e.* the Rad51-ssDNA filament that is present before the strand exchange reaction takes place, requires the presence of divalent cations (Ca^{2+}/Mg^{2+}) and ATP (or an ATP analogue). There are also several

accessory proteins involved in the formation and dissociation processes *in vivo* (130). To obtain detailed structural information about how the divalent cations (Mg^{2+}/Ca^{2+}) and the accessory protein Swi5-Sfr1 affect the structure of the presynaptic filament, LD spectroscopy has been exploited. However, in order to utilize LD spectroscopy for this purpose, the absorption profiles of the chromophores within the complex have to be individually distinguishable. Unfortunately, there are several components of the Rad51-DNA complex that all absorb in the same region and overlap each other (Figure 15). First, DNA has four different bases (adenine, thymine, cytosine and guanine) with absorption peaks around 260-280 nm, and there is also an additional adenine in the bound ATP or ADP molecule. Furthermore, there are several tyrosines and phenylalanines in the the amino acid sequence of each Rad51 monomer that also absorb in this region (10 tyrosines and 12 phenylalanines for HsRad51).

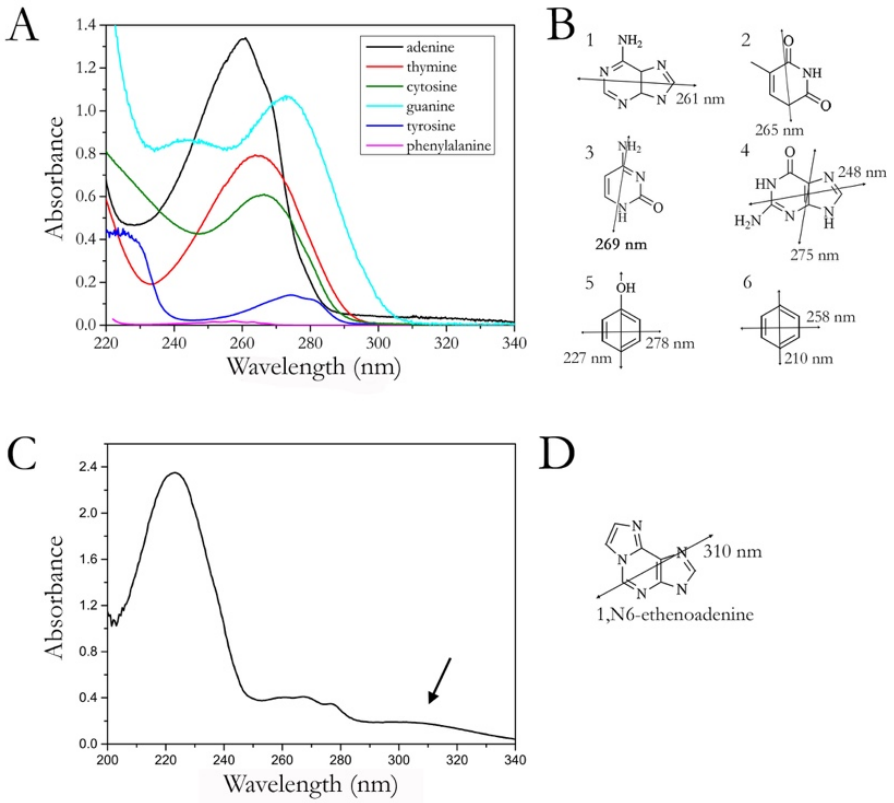

Figure 15. In (A), absorbance spectra of the chromophores found in the Rad51-DNA filament are shown with their absorbance intensities relative to each other. Adenines (in DNA and ATP/ADP), thymines, cytosines, and guanines (131), as well as phenylalanines and tyrosines (132) all absorb in the same region (240-280 nm). The chromophores and their corresponding transition moments are shown in (B): (1) adenine, (2) thymine, (3) cytosine, (4) guanine, (5) tyrosine and (6) phenylalanine. In (C), the absorbance spectrum of poly(dεA) is shown. There is a strong absorption peak around 225 nm, but it has also absorption above 300 nm (marked with an arrow), a region where no other chromophore in a Rad51-poly(dεA) filament absorbs. In (D), the molecular structure of 1,N6-ethenoadenine, the base in poly(dεA), and its transition moment at 310 nm are shown.

In Papers I and II, the issue of overlapping absorbance spectra in the Rad51-ssDNA filaments was addressed by using a poly(dA) analogue called poly(dεA) (dεA = 1,N6-ethenoadenine). Poly(dεA) absorbs strongly at 225 nm and has an additional absorption around 310 nm (Figure 15C and D), a wavelength where no other chromophore in the Rad51-poly(dεA) complex absorbs: thus any LD signal at this wavelength stems solely from organization of the poly(dεA) bases.

5.2.2 *Effect of varying the cation*

By using poly(dεA) it was possible to show that the choice of cation (Mg^{2+} or Ca^{2+}) has structural effects on the presynaptic filament. The bases were preferentially perpendicularly aligned relative the DNA axis when Ca^{2+} was present (negative LD at 310 nm), whereas they were basically unordered in the presence of Mg^{2+} (Figure 16). To exclude the trivial explanation that the reason for the unoriented bases in the HsRad51-ssDNA filament with Mg^{2+} was incomplete protein coverage, HsRad51 was added in surplus. The protein addition resulted in no changes in the LD signal, verifying that the DNA was covered to maximum extent. The LD signal was also confirmed to be stable during a period of six hours, thus dismissing depolymerization as an explanation for the weaker signal in presence of Mg^{2+}.

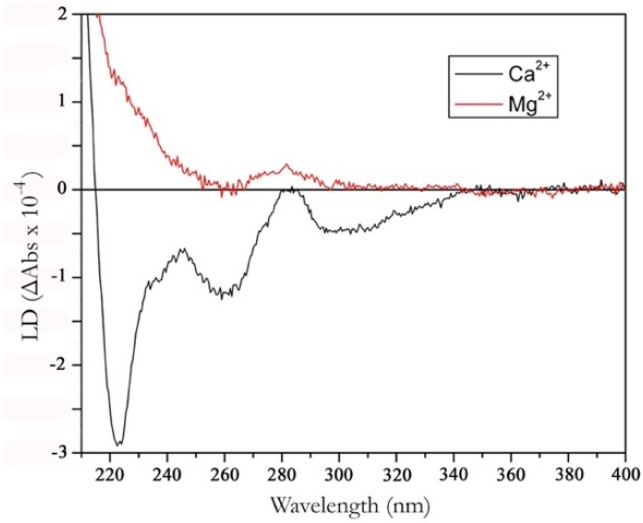

Figure 16. The LD signals arising from HsRad51 filaments formed on poly(dεA) in the presence of Ca^{2+} (black) and Mg^{2+} (red).

To determine the reason for the increased alignment of the DNA bases in the HsRad51-ssDNA filament in presence of Ca^{2+} was challenging. In Paper I it is argued, based on molecular modeling, that it is probably the somewhat larger polarizability and size of Ca^{2+} compared to Mg^{2+} that is indirectly responsible for this effect. Molecular modeling showed that presence of cations regulated the conformation of a

HsRad51 segment called the L2-loop. When the counterion was absent, the L2-loop was in the structure of a random coil. However, when a cation was bound to the filament, the L2-loop transformed to an α-helix that interacted with the bound DNA (Figure 17). In the presence of Ca^{2+}, the α-helical L2-loop is probably more stable, which in turn enables more stable interactions between the HsRad51 monomers and the DNA strand, resulting in ordering of the DNA bases.

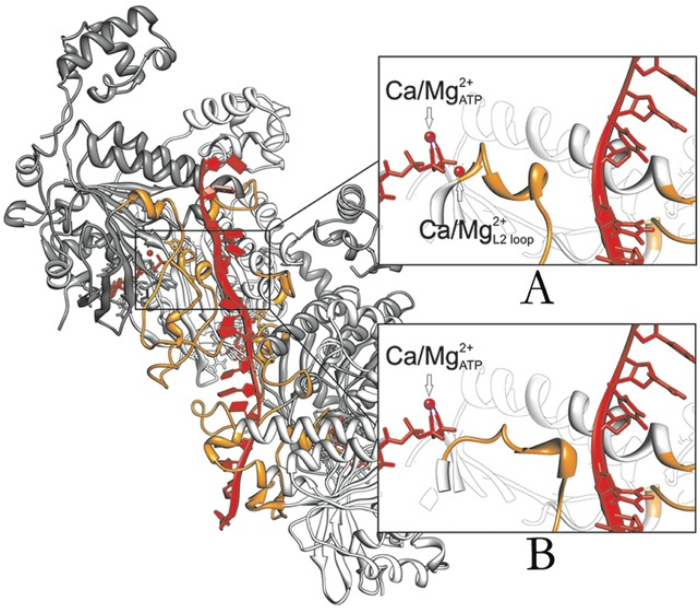

Figure 17. Model structure of the HsRad51-ssDNA filament adopted from Paper I, with two (A) or one (B) divalent cation bound. There are two binding pockets for the divalent ions, one that coordinates the oxygen atoms of the phosphate in ATP (labeled 'ATP'), and one that coordinates the alpha-helical region of the L2 loop (labeled 'L2 loop'). When Mg^{2+} or Ca^{2+} is bound in the 'L2 loop'-pocket, the cation affects the DNA indirectly by altering the structure of the L2 loop.

5.2.3 Swi5-Sfr1

The effect that the accessory protein Swi5-Sfr1 exerts on the presynaptic filament, and in particular the orientation of DNA bases, was also studied by LD spectroscopy (Figure 18). Swi5-Sfr1 binds in the groove of the right-handed Rad51 filament that is formed around DNA, and one Swi5-Sfr1 heterodimer covers about 10 Rad51 monomers (110). By LD spectroscopy, again with poly(dεA) as DNA substrate, it was possible to deduce that the addition of fission yeast Swi5-Sfr1 causes the otherwise less well ordered bases in the SpRad51-ssDNA filament formed in presence of Mg^{2+} (Figure 16) to become highly perpendicularly organized relative to the filament axis (Figure 18A). The addition of Swi5-Sfr1 to the presynaptic filament with Ca^{2+} also caused a certain improvement in the perpendicular organization of the DNA bases, however the increase was not as large as that for Mg^{2+}. By using an N-terminal deleted

Sfr1 (Swi5-Sfr1C), thus eliminating the part of the heterodimer that has affinity for DNA (110), it was also possible to demonstrate that the organization of the DNA bases is related to Swi5-Sfr1 interacting with the nucleoprotein filament, and not simply by interaction with the DNA directly.

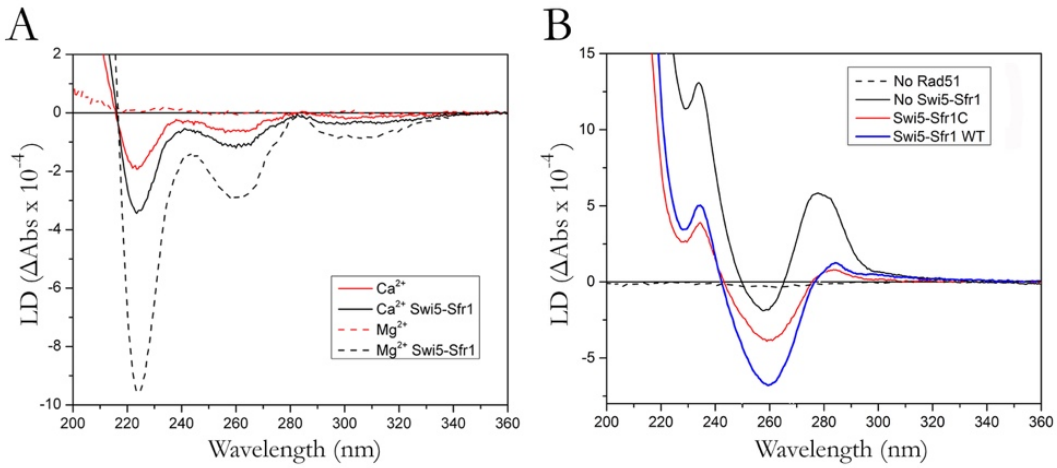

Figure 18. In (A), LD spectra of SpRad51 filaments formed on poly(dεA) with Ca^{2+} (solid lines) or Mg^{2+} (dashed lines), in the presence (black) or absence (red) of Swi5-Sfr1. Perpendicular alignment of the ssDNA bases in the filament is reflected in the amplitude of the negative LD signal around 310 nm. In (B), the binding of the N-terminal deleted heterodimer, Swi5-Sfr1C (red), onto poly(dT) demonstrates similar structural effects on the SpRad51-poly(dT) filament as the wild type protein (blue). The LD signal of the SpRad51-poly(dT) filament without Swi5-Sfr1 is also shown (black), as well as the non-existing signal of poly(dT) and Swi5-Sfr1 alone, without SpRad51 (dashed).

5.2.4 Biological implication of DNA base alignment in the filament

Previous section (chapter 5.2.3) considered how the cation affects the alignment of the DNA bases in the presynaptic filament, and a legitimate hypothesis is that the base alignment can be coupled to the biological function of the filament, *i.e.* the ability to perform the strand exchange reaction. Therefore, the efficiency of the strand exchange reaction displayed by the different complexes was analyzed (according to (133)). Basically, the extent of transferred DNA between a ssDNA (58-mer) that has been preincubated with SpRad51 and a homologous dsDNA (32 bp) was measured in presence of Ca^{2+} or Mg^{2+} and increasing amounts of Swi5-Sfr1 (Figure 19). The result from the strand exchange assay correlated well to the LD measurements: good perpendicular DNA base alignment corresponded to high strand exchange efficiency. The SpRad51-DNA filaments formed with Ca^{2+} showed more than twice the strand exchange efficiency as the filaments formed with Mg^{2+}, and the addition of Swi5-Sfr1 caused an increase in the strand exchange efficiency that was greater in the presence of Mg^{2+} than in the presence of Ca^{2+}.

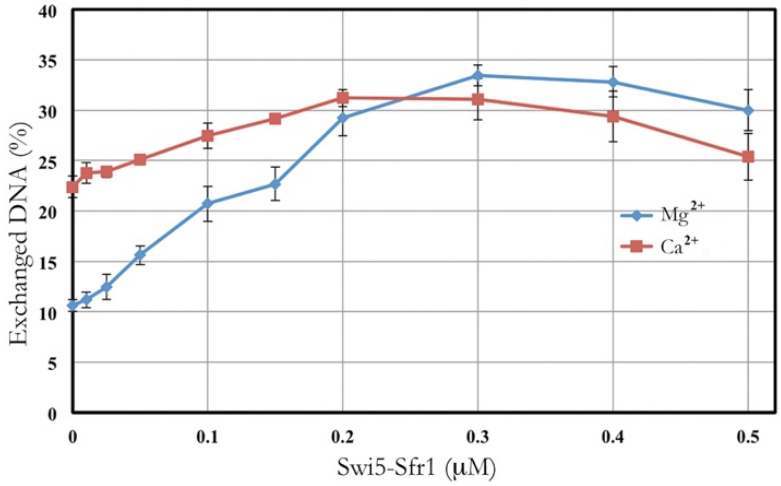

Figure 19. The percentage of exchanged DNA for SpRad51 filaments in presence of Ca^{2+}(red) or Mg^{2+}(blue), and increasing concentration of the accessory protein Swi5-Sfr1. The SpRad51 filaments were formed on ssDNA (58-mer) and subsequently incubated with homologous dsDNA (23 bp).

It is still not known which of cations is the main metal ion cofactor for Rad51 *in vivo*, Ca^{2+} or Mg^{2+}, or both. The most common divalent cation in the cell is Mg^{2+}, which is present in millimolar concentrations (134), while cellular concentrations of Ca^{2+} are in the nanomolar range (135, 136). Even though the Ca^{2+} concentration is increased when the cell is under stress, the level is still far below the concentration of Mg^{2+}, and it thus seems most likely that Mg^{2+} is the main metal ion cofactor in the Rad51 filament. In addition to the large concentration differences in the cell, the presynaptic filament with Mg^{2+} bound seems to be more responsive to external changes, such as addition of Swi5-Sfr1. The eukaryotic Rad51 works in concert with a large number of accessory proteins, thus it may be beneficial if it can be manipulated rather easily and exhibit some kind of response to its accessory proteins.

5.3 Macroscopic structure of the Rad51-DNA filament

In this section, the second question posed in chapter 5.1 is addressed. The macroscopic structures of single HsRad51-DNA filaments in solution have been monitored by fluorescence spectroscopy in combination with nanofluidics, and distinct macroscopic characteristics were revealed. These characteristics could be correlated to the choice of cation, DNA substrate and Rad51 concentration. A summary is presented, and the complete study is found in Paper III.

5.3.1 Transient bundles and stationary kinks

The filament formation of Rad51 on DNA is a complex process, as described in chapter 3.1.2. Essentially, Rad51 forms short "patches" on the DNA, and there may be gaps of naked DNA of various lengths between two adjacent patches. When HsRad51-DNA filaments were confined to nanochannels, certain characteristics were detected that could not be seen for filaments formed by its bacterial homologue RecA, a protein that covers the DNA homogenously. When confined to nanochannels of 650 x 140 nm^2 dimensions, the homogenous RecA-DNA filaments undulated smoothly between the walls of the nanochannel, while the heterogeneous HsRad51-DNA filaments displayed two distinct features: stationary kinks, which are bends in the filament that resides throughout the imaging time (6 seconds) and usually longer, and transient bundles, which arise when a filament curls up on itself locally, resulting in high intensity areas that form and disappear continuously. Interestingly, one individual HsRad51 filament did usually not display both features; it exhibited either kinks or bundles.

The heterogeneous DNA-coverage of HsRad51, which resulted in the presence of stationary kinks and transient bundles, made the analysis of the HsRad51-DNA filaments in nanochannels far from trivial, and standard polymer theories could not be applied. Paper III is therefore mostly based upon qualitative analysis. Typically, captured images of a polymer confined to a nanochannel are presented in a figure called a kymograph, which is a time-trace of the emission intensity of the polymer with the time on the y-axis (Figure 20A). Sudden changes in the polymer over time can be detected in a kymograph, and transient bundles in the HsRad51-DNA filaments were thus readily detected in such "average" kymograph. Identification of stationary kinks required creation of several kymographs, one for each pixel line along the nanochannel (line-by-line kymograph, Figure 20B), or they could also be detected by superimposition of the images (superimposed trajectories, Figure 20C). By visualizing the Rad51-filaments in the nanochannel by average kymographs in combination with liny-by-line kymographs or superimposed trajectories, it was possible to detect filaments that displayed transient bundles or stationary kinks for HsRad51-DNA filaments. Interestingly, these characteristics were not detected for RecA-DNA filaments (Figure 20).

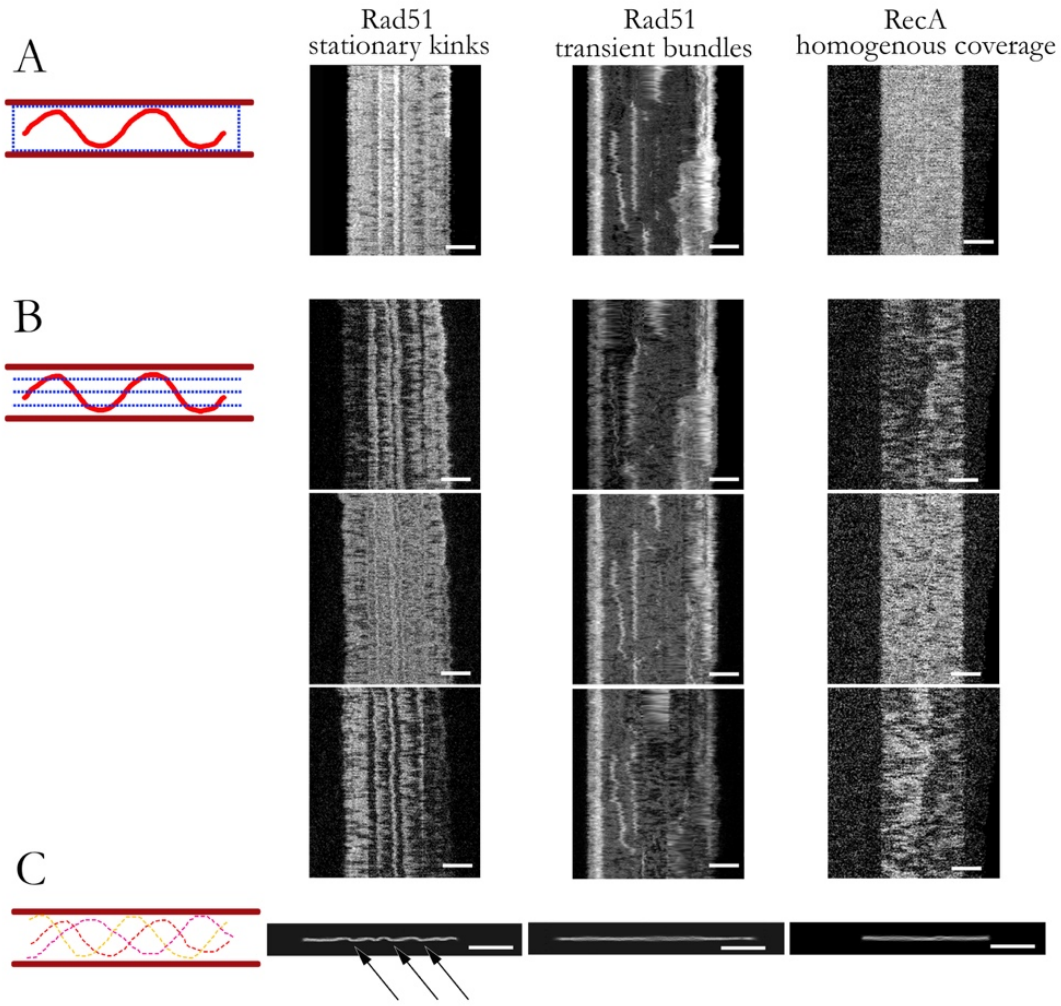

Figure 20. In (A), average kymographs are shown for three filaments confined to 650 x 140 nm^2: a HsRad51-dsDNA filament that displays stationary kinks (left), a HsRad51-dsDNA filament that displays transient bundles (center) and a RecA-dsDNA filament that displays regular undulations (right). In (B) and (C), their corresponding line-by-line kymographs and superimposed trajectories are shown. Bends in the superimposed trajectory, reflecting in stationary kinks in the filament, are denoted with arrows. The scale bars are 5 μM.

In Paper III, it is concluded that the reason for the two different dynamic features displayed by the HsRad51-DNA filaments was the presence of naked DNA between the HsRad51 patches. Non-covered DNA is very flexible, exhibiting a persistence length of around 2.5 nm or 50 nm depending on if it is singe- or double-stranded, while protein covered DNA is much stiffer, with a persistence length of several hundred nanometers. If the naked DNA was very short between the patches, the filament displayed stationary kinks, while longer stretches of naked DNA correlated to formation of transient bundles (Figure 21). In comparison, when the DNA was completely covered, such the case for RecA (Figure 20), the filament undulated smoothly in the nanochannel.

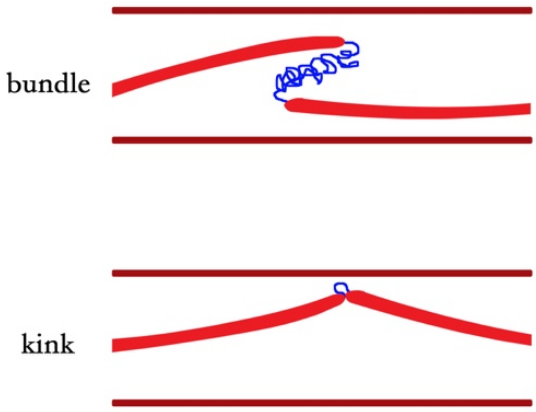

Figure 21. Schematic picture of the reasons for the bundles and kinks displayed by HsRad51-DNA filaments in the nanochannels. HsRad51 covers the DNA (blue strand) in patches (red areas). When the naked DNA stretch is sufficiently long, the patches may stack on top of each other, giving rise to a bundle. A stationary kink is probably due to a strong interaction between two adjacent HsRad51 patches, which is seemingly benefited by short stretches of naked DNA between the protein patches.

The existence of long naked DNA between the HsRad51 patches was verified by moving filaments that displayed transient bundles into a stronger confinement (a possibility and advantage of using funnel-shaped nanochannels), where it became highly stretched out and large regions of DNA that was not covered by fluorescent HsRad51 could be detected. To experimentally verify the presence of short stretches of naked DNA was harder, since these gaps were too small to be detected by the fluorescence microscope. However, the absence of detectable naked DNA regions in these filaments demonstrated that they did not contain longer stretches of naked DNA.

5.3.2 Manipulating the number of kinks

The Rad51 filament formation depends on the rate of nucleation on the DNA and the rate of the subsequent polymerization. Factors that were demonstrated to affect the HsRad51 filament formation were the DNA substrate (ssDNA or dsDNA), the choice of cation (Mg^{2+} or Ca^{2+}) and the concentration of HsRad51 monomers during the nucleation event. By modifying these factors, the number and lengths of the HsRad51 patches formed on DNA were varied, which was reflected in the filament characteristics when confined to nanochannels.

When Mg^{2+} was present, it was possible to detect distinct changes in the filament characteristics, especially in the number of kinks, when varying the nucleation conditions (Figure 22A). By decreasing the HsRad51 nucleation concentration (from ratio 1:1 to 1:600 HsRad51 monomers to DNA bp) the filaments displayed fewer

kinks. Additionally, the changes were even more pronounced when ssDNA was used as substrate instead of dsDNA. Of all the assessed conditions, the HsRad51 filaments formed on ssDNA with Mg^{2+} and low HsRad51 nucleation concentration (1:600) behaved most similarly to the homogenously covered RecA-DNA (Figure 21C, right column). Surprisingly, when Ca^{2+} was used as a counterion, the results were the opposite (Figure 23B). Filaments formed on dsDNA seemed to demonstrate a smoother movement than those formed on ssDNA, and when the nucleation concentration was decreased from 1:1 to 1:600 a higher number of stationary bends were detected. However, the changes for Ca^{2+} were not as evident as for Mg^{2+}, and the results therefore have to be interpreted with care. To conclude, the balance between nucleation and polymerization, which determines the number and length of the protein patches along the DNA, is shifted between the ratios 1:1 and 1:600 HsRad51 monomers to DNA bases/bp for Mg^{2+}, but not for Ca^{2+}.

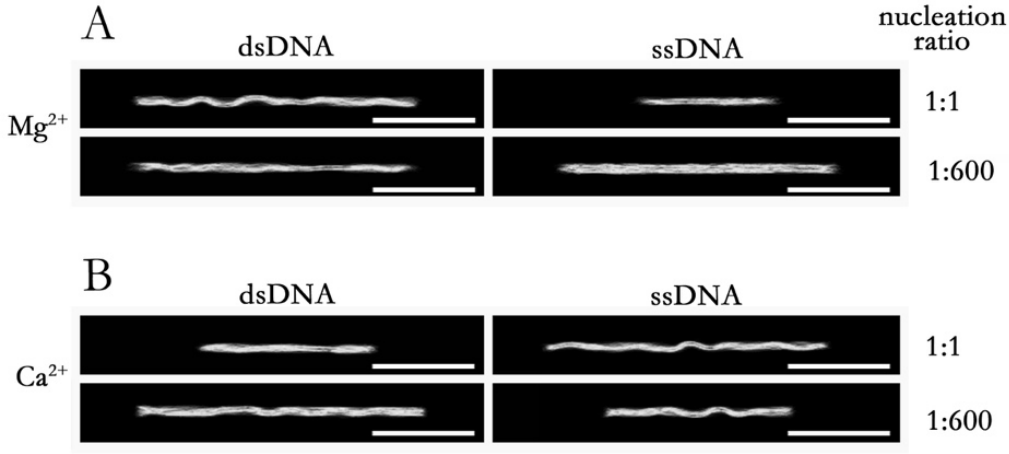

Figure 22. Representative HsRad51 filaments formed in eight different conditions (at least 20 filaments were monitored per condition). In (A), Mg^{2+} is used as a cation, and the DNA is either dsDNA (left column) or ssDNA (right column. The nucleation ratio is high (1:1 HsRad51 to DNA bp) or low (1:600 HsRad51 to DNA bp). Similar conditioned are displayed in (B), but with Ca^{2+} as cation. Noteworthy, the ssDNA is not uniform in length since the production method (rolling circle amplification) does not allow that parameter to be strictly controlled. The scale bars are 5 μM.

5.4 UV transition moments of tyrosine

In this section, the third question posed in chapter 5.1 is addressed. I intend to increase the amount of structural information about proteins, Rad51 in particular, that can be deduced from polarized light spectroscopy measurements. I demonstrate how the first two electronic transition moments of the chromophore in the amino acid tyrosine, *para*-cresol, are affected by environmental factors, and thereafter a structural

analysis of the HsRad51-dsDNA filament with the newly gained knowledge is conducted. The complete study, consisting of both spectroscopic measurements and quantum mechanical calculations, is found in Paper IV.

5.4.1 Shifts in the L_a and L_b transitions

The two lowest lying π→π* transitions of *para*-cresol, L_a and L_b, give rise to an absorbance spectrum with two peaks in the UV-region. The transition moment L_a is oriented parallel to the symmetry axis of the molecule (z-axis) and absorbs strongly at 225 nm, while L_b is almost perpendicular to the symmetry axis but still in the plane of the aromatic ring (*i.e.* y-polarized) and absorbs at 280 nm (Figure 23).

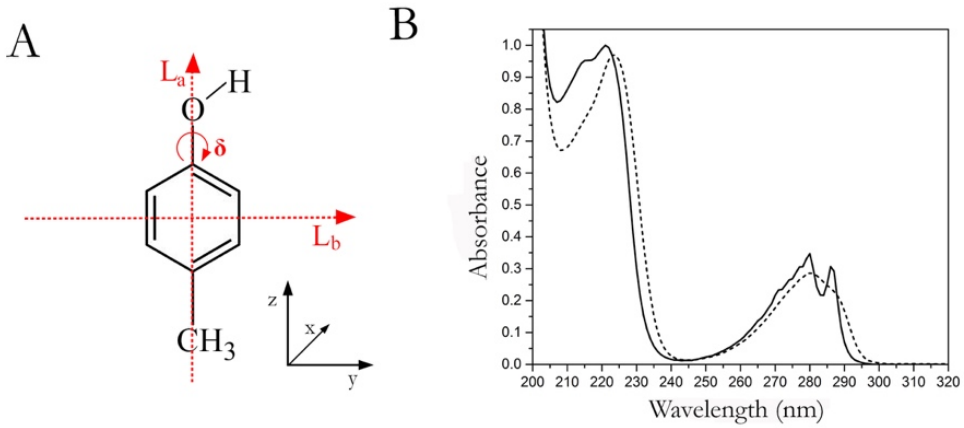

Figure 23. In (A), the molecular structure of *para*-cresol with the directions of its first two electronic transitions moments (L_a and L_b) shown as dotted red arrows. L_a and L_b are directed almost parallel to the z- and y-axis, respectively. δ denotes the angle of rotation of the hydroxyl group around the C-O bond. In (B), the absorbance spectra of *para*-cresol in nonpolar cyclohexane (solid) or polar methanol (dotted). The prominent peak around 225 nm reveals a shift of approximately 3 nm between the two solvents.

The isotropic absorbance spectrum of *para*-cresol demonstrated a 3 nm red-shift when the solvent was changed from nonpolar cyclohexane to polar methanol (Figure 23B). Also, the clear vibronic structure seen in cyclohexane was smoothed out in the polar environment. The reason for the smoothening of the spectrum is that *para*-cresol can reside in multiple transient configurations in the polar environment, and all these absorbance spectra are averaged, forming the smooth absorption peak. The absorbance shift is slightly harder to explain. When examining the spectral shift by TDDFT calculations (method found in Paper IV), it was evident that the shift was in fact not due to bulk polarity itself, but rather due to the hydrogen bonds that *para*-cresol formed with solvent molecules. When *para*-cresol acted as a hydrogen bond donor, the calculated shift was 3 nm to longer wavelengths, and when the

molecule acted as an acceptor the shift was of equal magnitude but to shorter wavelengths. Thus, the reason for the experimentally detected 3 nm red-shift is probably that the molecule act as a hydrogen bond donor in methanol.

5.4.2 Directions of the L_a and L_b transitions

Through LD, the directions of the transition moments of *para*-cresol were resolved. *Para*-cresol, being a small molecule, was aligned in two different stretched films; polyvinylalcohol (PVA), which has polar character, and polyethylene (PE), which is nonpolar. From the obtained LD spectra (Figure 24A and B), the long- and short-axis polarizations were calculated (Figure 24C and D) (method described in Paper IV). In both PVA and PE, the long-axis polarizations (A_z) did not contribute significantly to the absorption above 240 nm, indicating that the L_b transition is purely polarized along the y-axis. The L_a transition on the other hand demonstrated a slight perpendicular contribution (A_y), seen in Figure 24C and D as the peak around 220 nm in the solid line spectra. This perpendicular contribution, which was more prominent in nonpolar solvents, is believed to arise from intensity mixing with stronger transitions found at shorter wavelengths, in accordance with the perturbation theory that explains how weaker transitions are more easily affected by stronger close-lying transitions. Just below 200 nm are B_a and B_b, which are two strong mixed transitions (having both parallel and perpendicular character) that may perturbate the L_a transition. An explanation for that the perturbation was only detected in nonpolar environments is that B_b shifts slightly to the red in nonpolar solutions (137), and comes therefore closer to the L_a transition.

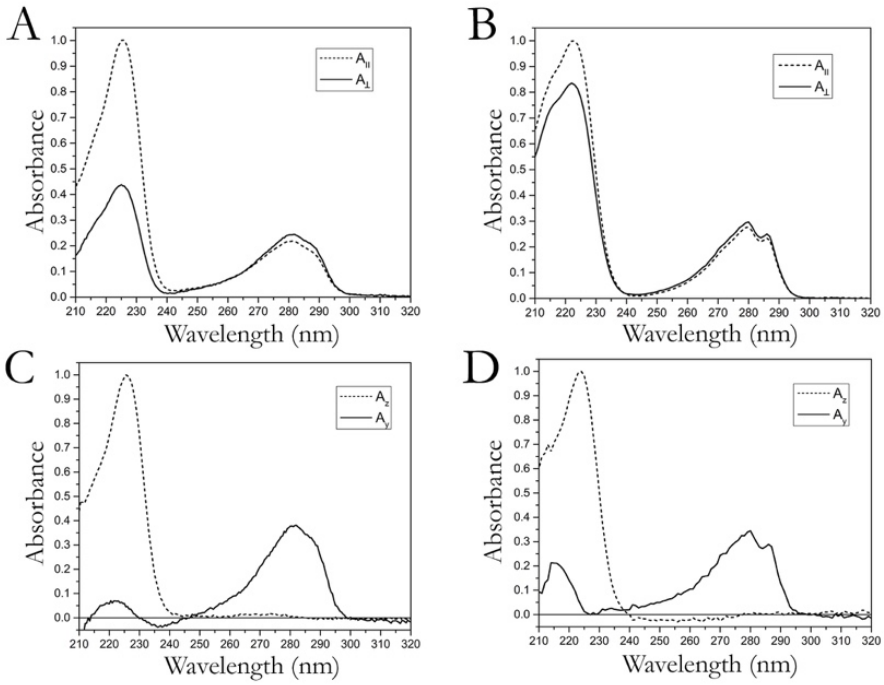

Figure 24. Polarized absorbance spectra of *para*-cresol in stretched films of PVA (A) and PE (B). The dotted lines are the absorptions parallel ($A_{\parallel}$) to the orientation axis, and the solid lines are the absorptions perpendicular ($A_{\perp}$)to the orientation axis. From the absorbance spectra in (A) and (B), the long- (A_z, dotted) and short-axis (A_y, solid) polarizations of *para*-cresol were calculated, and these are shown in (C) for PVA and (D) for PE.

TDDFT calculations of the directions and magnitudes of the transition moments agreed well with the LD results for *para*-cresol in the stretched films. The transition moment directions were not sensitively dependent on the environment and they may therefore be used without any correction to interpret the orientation of the tyrosine ring using the LD signal. By contrast, the transition energies were shown to depend on environmental factors. The most interesting feature found from the TDDFT calculations was that the energy gap for the $S_0 \rightarrow S_1$ transition changed slightly when the hydroxyl group was rotated out-of-plane (δ in Figure 24A), reflected in a predicted blue shift of the absorption peak. In solution, this will not matter since the molecule will be in its lowest energy state with the hydroxyl group in the same place as the aromatic ring. However, when *para*-cresol is in a tyrosine residue in a protein, hydrogen bonding to its neighboring amino acids or to water molecules trapped in polar pockets could cause an out-of-plane rotation of the hydroxyl group. The hydrogen bond will, in itself, cause an absorbance shift as explained previously (red or blue depending on if *para*-cresol acts as a donor or acceptor), but the possible rotation of the hydrogen group will also contribute to the shift.

5.4.3 Analyzing the tyrosines within a protein

Due to its large extinction coefficient, high quantum yield and fluorescence properties, that all depend on the environment, tryptophan is routinely used as an intrinsic marker in spectroscopic studies of protein structure. However, with the better sensitivity of the instruments of today, tyrosine could, and should, be exploited to a higher extent. The tyrosine residue is more polar than tryptophan, and is therefore expected to be more exposed to the solvent in a folded protein, a feature that is beneficial when studying effects of environmental changes. In addition, the wavelength shift of the absorption peak of tyrosine is more distinct than that of tryptophan when the surrounding environment changes. In Paper IV, a first attempt to analyze a protein structure (HsRad51-dsDNA filament), using the newly gained knowledge on how hydrogen bonding and rotation of the hydroxyl group of a tyrosine residue cause spectral shifts, was performed.

An LD study, where tyrosines were site-specifically mutated to phenylalanines, was revisited (138). This so-called SSLD method permits, in principle, the LD signal for each individual tyrosine to be deduced. For the HsRad51-dsDNA filament, eight resolved SSLD spectra of separate tyrosine residues in the HsRad51 protein have been analyzed, and the residues that demonstrated the most distinctive absorbance shifts in the L_b transition were investigated more closely. In general, a detected red-shift is most probably due to hydrogen bonding with *para*-cresol acting as a hydrogen bond donor, while a blue-shift, could be due to one of the following reasons: hydrogen bonding as an acceptor, being exposed to the bulk where hydrogen bonds are formed only briefly, or out-of-plane rotation of the hydroxyl group. The reasons for the detected spectral shifts of the tyrosine residues in the HsRad51-dsDNA filament were more or less obvious, depending on their positions within the protein and the number of possible interactions to consider. The most intriguing result of the analysis is however that of residue Tyr232. This residue is positioned in a region of the protein that has yet to be resolved by crystal structure determination. It was earlier proposed that Tyr232 intercalates between two DNA bases, contributing in coordinating the incoming dsDNA (138). However, the SSLD signal of Tyr 232 is only slightly red-shifted, and in the event of intercalation the red-shift is expected to be larger. Even more important, the SSLD is positive, a counterargument for intercalation which should have been associated with negative LD. The fact that there exists an SSLD signal demonstrates that Tyr232 is in fact oriented, but the slight red-shift could be better explained by Tyr232 acting as a hydrogen bond donor. A suggestion is that Tyr232 is somehow aligned with the DNA bases, forming a hydrogen bond with the phosphate backbone of DNA.

In conclusion, the revealed spectral shifts of *para*-cresol add an interesting new dimension to LD spectroscopy when analyzing systems with tyrosine residues that have dissimilar orientations and/or are positioned in dissimilar environments.

5.5 Minor-groove binding of Hoechst 33258

I will now switch the focus from Rad51 to a much smaller DNA-binding molecule, the minor-groove binding molecule Hoechst 33258. The last question posed in chapter 5.1, how Hoechst 33258 binds to AT-rich DNA at high concentrations, is addressed in this chapter. The whole study, which has been conducted by both spectroscopic (CD) and thermodynamic (ITC) techniques, can be found in Paper V.

5.5.1 Position of a second Hoechst 33258

Hoechst 33258 (Figure 25A) binds strongly in the minor groove of AT-rich regions, as discussed in chapter 2.4.1. Even though the molecule is a routinely used dye for staining DNA, its binding at high concentrations is not yet completely understood. Hoechst 33258 is a model compound for minor-groove binders, and additional knowledge of its binding could possibly aid the development of future DNA-targeting drugs. It is known that the molecule binds strongly to AT-rich sequences of at least 4 bp, however there are controversies about how additional Hoechst 33258 molecules bind to such sequences. Suggestions vary from electrostatic binding to the backbone, to dimers and even larger aggregates (50, 54). In Paper V, the binding of Hoechst 33258 to three dodecamers, 5'-CGCGAATTCGCG-3 (A_2T_2), 5'-CGCAAATTTGCG-3' (A_3T_3) and 5'-CGAAAATTTTCG-3' (A_4T_4), were examined in order to characterize the binding of the molecule.

Examination of the binding mode of Hoechst 33258 was conducted by combining, in a global data analysis, the information obtained from CD and ITC measurements when Hoechst 33258 was titrated into the three different oligonucleotides (A_2T_2, A_3T_3 and A_4T_4). SVD analysis (method described in Paper V) revealed the different components of the CD spectra, *i.e.* the number of binding modes. For A_2T_2 and A_3T_3 there were two component detected, most probably arising from the first dye binding in the minor groove and the second binding externally. For A_4T_4, on the other hand, two similar, non-cooperative binding modes were detected before similar external binding took place (Figure 25B and C). The ITC measurements did also confirm external binding after ratio 1:1 for A_2T_2 and A_3T_3, and after ratio 2:1 for A_4T_4, by displaying a slight exothermic signal that was similar for all oligonucleotides (Figure 25D-F).

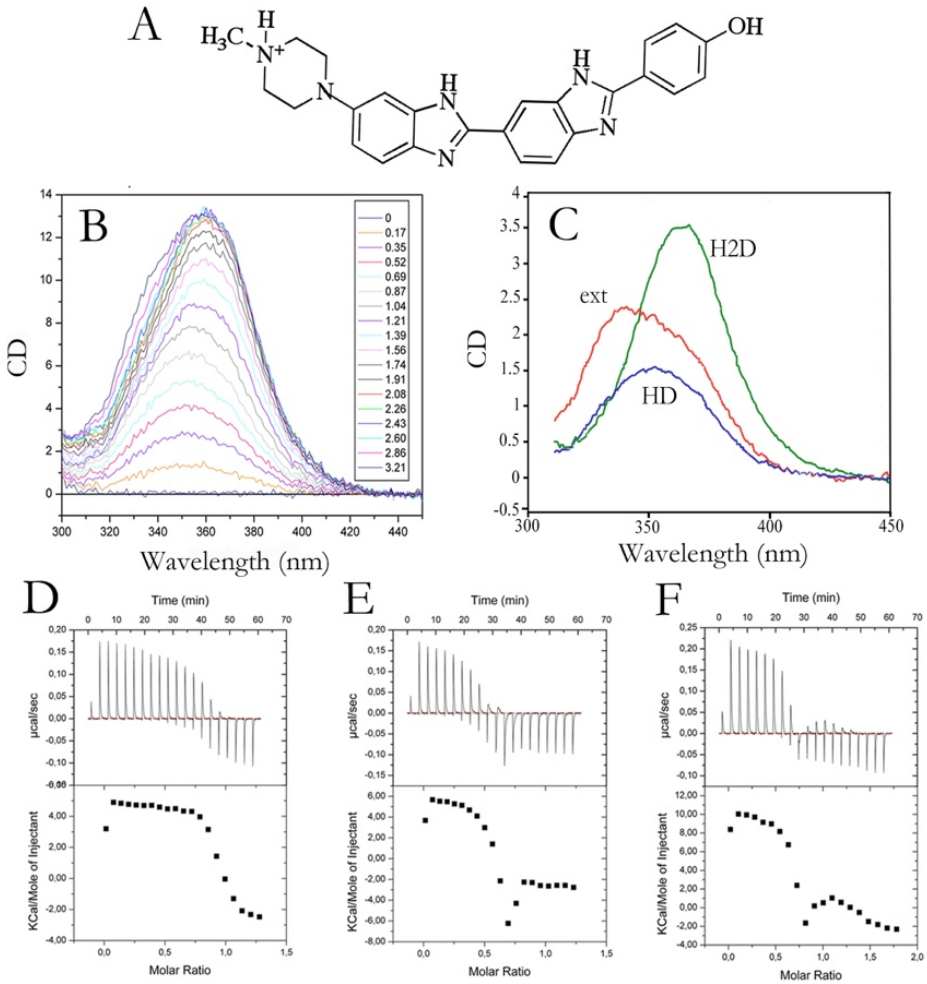

Figure 25. In (A), the molecular structure of Hoechst 33258 at neutral pH is shown. In (B), the raw CD spectra of Hoechst 33258 being titrated into A_4T_4 and (C) shows the three resolved components received from SVD analysis of the CD titration, here denoted "HD", "H2D" and "ext" (C). Also shown are the ITC thermograms and its corresponding binding isotherms for Hoechst 33258 being titrated into A_2T_2 (D), A_3T_3 (E) and A_4T_4 (F).

Four feasible binding modes for the second Hoechst 33258 to A_4T_4 were considered. Two Hoechst 33258 molecules could sit in the minor groove in the form of a sandwich dimer (Figure 26A), or a partially overlapping dimer (Figure 26B), or with a slight separation between them (Figure 26C). The fourth alternative is external binding for the second Hoechst 33258 molecule (Figure 26D), which was already ruled out by both the CD and ITC measurements. By examination of the CD spectra (Figure 25B), it was possible to rule out the sandwich dimer and the overlapping dimer too, since these complexes would have resulted in a blue-shift or a CD with two opposite peaks, respectively. The remaining model, with binding of two Hoechst 33258 molecules in the minor groove but a slight distance apart would result in a slight red-shift in the CD, which we also detected.

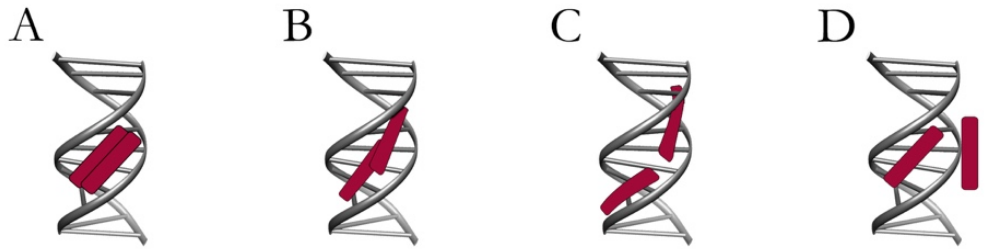

Figure 26. Schematic pictures of the four feasible models for the binding of two Hoechst 33258 to the A_4T_4 duplex. The two molecules could be positioned on top of each other in the minor groove (A), they could be accommodated in the minor groove slightly overlapping each other (B) or a small distance apart (C), or the second Hoechst 33258 could bind externally to the DNA backbone (D). All models except (C) are ruled out based on the CD and ITC analysis.

5.5.2 Exothermic peak

Through ITC measurements it was revealed that the endothermic binding of Hoechst 33258 to A_3T_3 and A_4T_4 demonstrated exothermic peaks around ratio 0.75:1 Hoechst 33258 to oligonucleotide (Figure 25E and F). The peak of A_4T_4 was analyzed more closely, and it could not be explained by dye aggregation, the existence of different binding sites, nor the presence of different DNA populations (examined by ^{1}H NMR), the latter being a hypothesis suggested for a similar system (139, 140). A decent fit to the data in our global analysis could only be obtained when two DNA populations with different thermodynamic profiles were considered (red line in Figure 27A). Both DNA populations (HD and HD* in Figure 27B) result in the same Hoechst 33258-DNA complex, but they have different binding affinity and enthalpy. A potential structural explanation for the exothermic contribution is that the first bound Hoechst 33258 rearrange slightly in the groove when a second molecule enters, and this rearrangement exhibit an exothermic contribution.

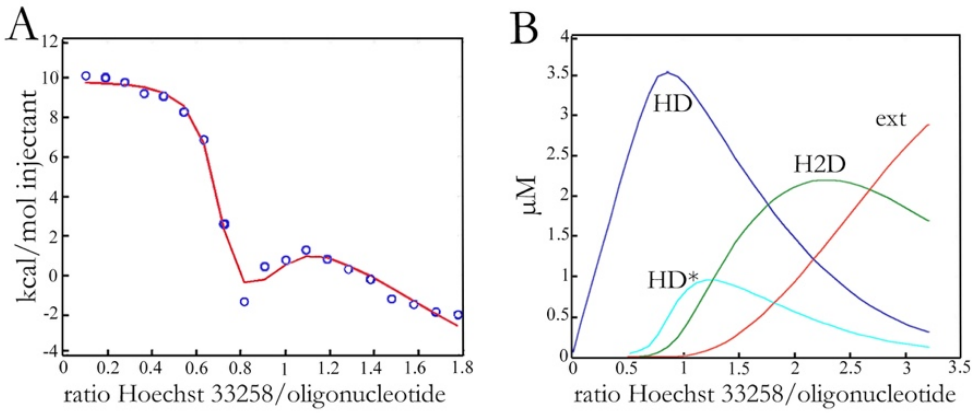

Figure 27. Illustrations of the best fits for the measured signals to the theoretical binding isotherm for Hoechst 33258 binding to A_4T_4. In (A), the theoretical binding isotherm (red) is fitted to the ITC signal (blue). In (B), the concentration profiles for the first binding ("HD" and "HD*", two binding modes that have the same structural complex, but different binding enthalpies), second binding ("H2D") and external binding ("ext") are shown.

To conclude, it is suggested Paper V that the two Hoechst 33258 molecules are positioned in the minor groove of A_4T_4 a slight distance apart. Thus, the two molecules do not form a dimer in the groove. Additional knowledge from the study is that direct comparison between ITC and CD data has to be done with care, since CD measures the equilibrated complex while ITC measures all the ongoing events in the sample cell.

6 Concluding remarks

In this section, a short summary of the research conducted in the appended papers of the Thesis is presented, followed by concluding remarks about the field of DNA-binding molecules.

The work presented in this Thesis concerns the binding to DNA, at high ligand coverage, by two rather different molecules: the cooperatively binding protein Rad51, which performs the strand exchange reaction in a DNA repair system called homologous recombination, and the small minor-groove binding molecule Hoechst 33258, which is routinely used for fluorescent staining of DNA. The work has mainly been conducted through spectroscopic techniques, in some cases in combination with measurements of the biological activity or thermodynamic profiles.

In summary, the prerequisites for an efficient strand exchange reaction performed by the Rad51 protein have been analyzed in terms of the coplanar orientation of the bases in the Rad51-ssDNA filament, which was considered in Papers I and II (for HsRad51 in Paper I and SpRad51 in Paper II). In Paper III, the macroscopic structure of the HsRad51-DNA filament was studied, and distinct features that the filaments displayed when confined to nanochannels were correlated to the conditions during the filament formation. With the wish to provide additional information regarding the Rad51-DNA structure, and also potentially other protein structures, at atomistic level by polarized light spectroscopy, the polarization properties and environment dependence of the L_a and L_b transitions in the chromophore of tyrosine have been studied in Paper IV. Finally, the binding of the model minor-groove binder Hoechst 33258 to oligonucleotides with A/T-regions of different lengths, at high ligand coverage, was considered in Paper V. The most important findings presented in this Thesis are the following:

- Ca^{2+} promotes a nearly coplanar alignment of the DNA bases, perpendicular to the filament axis, in the presynaptic filament (HsRad51-ssDNA in presence of ATP), while the bases are virtually unordered in presence of Mg^{2+} (Paper I).
- Addition of the accessory protein Swi5-Sfr1 causes an increased perpendicular organization of the DNA bases in the presynaptic filament, both in presence of Mg^{2+} and Ca^{2+}. The increase is largest when adding Swi5-Sfr1 to a filament formed in presence of Mg^{2+} (Paper II).
- A coplanar, preferentially perpendicular alignment of the DNA bases in the presynaptic filament is proposed to be a beneficial structural factor for increased strand exchange activity, supported by the observation that increased organization correlates with higher strand exchange efficiency (Papers I and II).

- The number of HsRad51 patches formed on long DNA is determined by the HsRad51 nucleation concentration, the type of cation present (Ca^{2+} or Mg^{2+}) and the DNA substrate (ss or ds) (Paper III).
- The length of the naked DNA stretches between the HsRad51 patches determines the physical appearance of the HsRad51-DNA filament (*i.e.* transient bundles or stationary kinks) in nanochannels (Paper III).
- The chromophore in tyrosine, *para*-cresol, has its first two low-lying electronic transitions, L_a and L_b, polarized parallel and perpendicular to the symmetry axis, respectively. L_b is purely perpendicularly polarized, while the parallel transition moment, L_a, displays a perpendicular character as well in non-polar solvents (Paper IV).
- The shifts of the absorption bands of the L_a and L_b transitions of *para*-cresol depend on the formation of stable hydrogen bonds, if *para*-cresol is acting as a hydrogen bond donor or acceptor, and on the rotation of its hydroxyl group. (Paper IV).
- The newly obtained knowledge about the absorbance shifts of tyrosine can be exploited for structural analysis of tyrosine-containing proteins by polarized light spectroscopy, as demonstrated for the HsRad51-dsDNA filament where a tyrosine was shown to not intercalate between the DNA bases, as previously proposed (Paper IV).
- In a short oligonucleotide that contains a region of 4 adenines and 4 thymines (3'-CGAAAATTTTCG-5'), two Hoechst 33258 molecules bind in a parallel manner in the minor groove with a small distance between them (Paper V).

The field of DNA-binding drugs is vast, and it is still expanding. The linkage between Rad51 and Hoechst 33258 may not be completely obvious, but both are highly connected to DNA and cancer therapy. The route of several cancer chemotherapies is to cause an excessive amount of damage to the DNA in the cancer cell, in order for it to die directly or make it unable to replicate. However, the response of the cell to such DNA damage is to use its DNA repair systems, and the efficacy of the cancer drug is therefore dependent on the capacity of the reparation machineries. By instead selectively inhibiting the repair systems of the cancer cells, or by inhibiting them in combination with DNA-damaging agents, they can be killed with higher specificity and with less side-effects (23). In order to do so, an increased understanding of the repair machineries in the cell is required, since the different systems may interplay with each other. In addition, we also need deeper knowledge of the binding and sequence-specificity of small DNA-binding molecules that can damage the DNA one way or another.

7 Acknowledgements

I would like to send my warm regards to the following people for their direct or indirect support:

My supervisor **Bengt Nordén** for his support, intelligent mind, and funny ideas. Thank you for always encouraging me, and also for letting me have the freedom to perform all my "side activities", such as chemistry outreach.

My co-supervisor **Per Lincoln** for his scientific support, and for helping me find inspiration, and solutions, when it felt impossible to draw any scientific conclusion at all.

My extra, unofficial, supervisor **Fredrik Westerlund** for his never-ending support. Even though you have your own group, you always seem to have time for me. Thank you for letting me work with you and your wonderful group members the last years.

Masayuki Takahashi for your scientific support, you are the King of Knowledge in my field of research (recombination proteins). Thank you for our collaborations and your valuable input.

Collaborators and **co-authors**, you made the research fun, and I have learnt a lot from you all. An extra thanks to **Karolin Frykholm** and **Anna Reymer**, you two helped me get started with my PhD studies, and now you help me finish it by proof-reading this Thesis.

Per Thorén and **Molecular Frontiers** for collaborating with me and the rest of Untamed Science, in order to produce inspirational chemistry videos, both in Sweden and in Singapore. It was definitely a lot of fun, and it turned out pretty good too.

Colleagues at Chalmers. All of you. Really, everyone. Thank you for fikas, laughter, moral and scientific support, and keeping me company while climbing, running and doing yoga.

Colleagues at Universeum. I truly value your enthusiasm. Thank you for great collaborations during our rewarding excursions to schools. I know we were there to inspire and teach the students science, but I think I was the one who got inspired the most.

My **friends** outside of Chalmers for just being awesome. I am happy to have you all in my life, and for you letting me be a part of yours.

My **family** for your endless support. The support over the last year has been especially valuable as our schedule at home has been rather full. Thank you also for being the best friends my daughter can have.

Jonas & Evelyn. I cannot even explain my gratitude for having you two crazy, happy and lovable people in my life. You are my everything.

8 References

1. Lindahl,T. (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709–715.

2. Hoeijmakers,J.H.J. (2009) DNA Damage, Aging, and Cancer. *N. Engl. J. Med.*, **361**, 1475–1485.

3. Bernstein,C., Prasad,A.R., Nfonsam,V. and Bernstein,H. (2013) DNA Damage , DNA Repair and Cancer. In Chen,C. (ed), *New Research Directions in DNA Repair*. InTech, pp. 413–466.

4. Huang,L.C., Clarkin,K.C. and Wahl,G.M. (1996) Sensitivity and selectivity of the DNA damage sensor responsible for activating p53-dependent G1 arrest. *Proc. Natl. Acad. Sci. U. S. A.*, **93**, 4827–4832.

5. Lips,J. and Kaina,B. (2001) DNA double-strand breaks trigger apoptosis in p53-deficient fibroblasts. *Carcinogenesis*, **22**, 579–585.

6. Rich,T., Allen,R.L. and Wyllie, a H. (2000) Defying death after DNA damage. *Nature*, **407**, 777–783.

7. Stofer,E. and Lavery,R. (1994) Measuring the geometry of DNA grooves. *Biopolymers*, **34**, 337–346.

8. McCarty,M. and Avery,O.T. (1946) Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types : Ii. Effect of Desoxyribonuclease on the Biological Activity of the Transforming Substance. *J. Exp. Med.*, **83**, 89–96.

9. Watson,J.D. and Crick,F.H.C. (1953) Molecular structure of nucleic acids. *Nature*, **171**, 737–738.

10. Watson,J.D. and Crick,F.H.C. (1953) Genetical implications of the structure of deoxyribonucleic acid. *Nature*, **171**, 964–967.

11. Crick,F.H.C. (1956) On protein synthesis. *Symp. Soc. Exp. Biol.*, **XII**, 139–163.

12. Crick,F.H.C. (1970) Cendral Dogma of Molecular Biology. *Nature*, **227**, 561–563.

13. Reisner,W., Pedersen,J.N. and Austin,R.H. (2012) DNA confinement in nanochannels: physics and biological applications. *Reports Prog. Phys.*, **75**, 106601.

14. Bloomfield,V.A., Crothers,D.M. and Ignacio Tinoco,J. (2000) Nucleic Acids - Structure, Properties and Functions. University Science Books, Sausalito, CA.

15. Atkins,P. and Paula,J. de (2002) Atkins' Physical Chemistry. 7th ed. Oxford University Press, New York.

16. Persson,F. and Tegenfeldt,J.O. (2010) DNA in nanochannels--directly visualizing genomic information. *Chem. Soc. Rev.*, **39**, 985–999.

17. Baumann,C.G., Smith,S.B., Bloomfield,V. a and Bustamante,C. (1997) Ionic effects on the elasticity of single DNA molecules. *Proc. Natl. Acad. Sci. U. S. A.*, **94**, 6185–6190.

18. Murphy,M.C., Rasnik,I., Cheng,W., Lohman,T.M. and Ha,T. (2004) Probing single-stranded DNA conformational flexibility using fluorescence spectroscopy. *Biophys. J.*, **86**, 2530–2537.

19. Daoud,M. and De Gennes,P.G. (1977) Statistics of macromolecular solutions trapped in small pores. *J. Phys.*, **38**, 85–93.

20. Odijk,T. (1983) On the Statistics and Dynamics of Confined or Entangled Stiff Polymers. *Macromolecules*, **1344**, 1340–1344.

21. Odijk,T. (1984) Similarity Applied to the Statistics of Confined Stiff Polymers. *Macromolecules*, **17**, 502–503.

22. Burkhardt,T.W., Yang,Y. and Gompper,G. (2010) Fluctuations of a long, semiflexible polymer in a narrow channel. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, **82**, 1–9.

23. Helleday,T., Petermann,E., Lundin,C., Hodgson,B. and Sharma,R. a (2008) DNA repair pathways as targets for cancer therapy. *Nat. Rev. Cancer*, **8**, 193–204.

24. Palchaudhuri,R. and Hergenrother,P.J. (2007) DNA as a target for anticancer compounds: methods to determine the mode of binding and the mechanism of action. *Curr. Opin. Biotechnol.*, **18**, 497–503.

25. Ehsanian,R., Van Waes,C. and Feller,S.M. (2011) Beyond DNA binding - a review of the potential mechanisms mediating quinacrine's therapeutic activities in parasitic infections, inflammation, and cancers. *Cell Commun. Signal.*, **9**, 1–18.

26. Gurova,K. (2010) New hopes from old drugs: revisiting DNA-binding small. *Futur. Oncol.*, **5**, 1–28.

27. Ebbinghaus,S.W. (2003) Site-Selective DNA Binding Drugs. *Science (80-.).*, **10**, 895–897.

28. Norden,B. (1978) Rearrangement of a platinum (II) complex in DNA from intercalation outer-sphere position to non-intercalation coordination. *FEBS Lett.*, **94**, 204–206.

29. Nordén,B. (1978) Structural evidence on DNA carcinogen interactions: N-acetoxy-N-2acetylaminofluorene binding to DNA. *Biophys. Chem.*, **8**, 385–391.

30. Tjerneld,F., Nordén,B. and Ljunggren,B. (1979) Interaction between DNA and 8-methoxypsoralen studied by linear dichroism. *Photochem. Photobiol.*, **29**, 1115–1118.

31. Pinto,A.L. and Lippard,S.J. (1985) Binding of the anitumor drug cis-diamminedichloroplatinum(II) (cicplatin) to DNA. *Biochim. Biophys. Acta*, **780**, 167–180.

32. Gewirtz,D.A. (1999) A critical evaluation of the mechanisms of action proposed for the antitumor effects of the anthracycline antibiotics adriamycin and daunorubicin. *Biochem. Pharmacol.*, **57**, 727–741.

33. Strekowski,L. and Wilson,B. (2007) Noncovalent interactions with DNA: an overview. *Mutat. Res.*, **623**, 3–13.

34. Chaires,J.B. (2006) A thermodynamic signature for drug-DNA binding mode. *Arch. Biochem. Biophys.*, **453**, 26–31.

35. Joubert,A., Sun,X.-W., Johansson,E., Bailly,C., Mann,J. and Neidle,S. (2003) Sequence-Selective Targeting of Long Stretches of the DNA Minor Groove by a Novel Dimeric Bis-benzimidazole. *Biochemistry*, **42**, 5984–5992.

36. Bailly,C. and Chaires,J.B. (1998) Sequence-Specific DNA Minor Groove Binders. Design and Synthesis of Netropsin and Distamycin Analogues. *Bioconjug. Chem.*, **9**, 513–538.

37. Laity,J.H., Lee,B.M. and Wright,P.E. (2001) Zinc finger proteins: New insights into structural and functional diversity. *Curr. Opin. Struct. Biol.*, **11**, 39–46.

38. Proudfoot,C., McPherson,A.L., Kolb,A.F. and Stark,W.M. (2011) Zinc finger recombinases with adaptable DNA sequence specificity. *PLoS One*, **6**, 1–9.

39. Dervan,P.B. (2001) Molecular recognition of DNA by small molecules. *Bioorganic Med. Chem.*, **9**, 2215–2235.

40. Dervan,P.B. and Edelson,B.S. (2003) Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr. Opin. Struct. Biol.*, **13**, 284–299.

41. White,S., Szewczyk,J.W., Turner,J.M., Baird,E.E. and Dervan,P.B. (1998) Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature*, **391**, 468–471.

42. Chaires,J.B. (1997) Energetics of drug-DNA interactions. *Biopolymers*, **44**, 201–215.

43. Neidle,S. (2001) DNA minor-groove recognition by small molecules. *Nat. Prod. Rep.*, **18**, 291–309.

44. Haq,I., Ladbury,J.E., Chowdhry,B.Z., Jenkins,T.C. and Chaires,J.B. (1997) Specific Binding of Hoechst 33258 to the d(CGCAAATTTGCG)2 Duplex: Calorimetric and Spectroscopic Studies. *J. Mol. Biol.*, **271**, 244–257.

45. Han,F., Taulier,N. and Chalikian,T. V (2005) Association of the Minor Groove Binding Drug Hoechst 33258 with d(CGCGAATTCGCG) 2 : Volumetric, Calorimetric, and Spectroscopic Characterizations. *Biochemistry*, **44**, 9785–9794.

46. Downs,T.R. and Wilfinger,W.W. (1983) Fluorometric quantification of DNA in cells and tissue. *Anal. Biochem.*, **131**, 538–547.

47. Goracci,L., Germani,R., Savelli,G. and Bassani,D.M. (2005) Hoechst 33258 as a pH-sensitive probe to study the interaction of amine oxide surfactants with DNA. *Chembiochem*, **6**, 197–203.

48. Barooah,N., Mohanty,J., Pal,H., Sarkar,S.K., Mukherjee,T. and Bhasikuttan,A.C. (2011) pH and temperature dependent relaxation dynamics of Hoechst-33258: a time resolved fluorescence study. *Photochem. Photobiol. Sci.*, **10**, 35–41.

49. Kubbies,M. (1990) Flow cytometric recognition of clastogen induced chromatin damage in G0/G1 lymphocytes by non-stoichiometric Hoechst fluorochrome binding. *Cytometry*, **11**, 386–394.

50. Loontiens,F.G., Regenfuss,P., Zechel,A., Dumortier,L. and Clegg,R.M. (1990) Binding characteristics of Hoechst 33258 with calf thymus DNA, poly[d(A-T)] and d(CCGGAATTCCGG): multiple stoichiometries and determination of tight binding with a wide spectrum of site affinities. *Biochemistry*, **29**, 9029–9039.

51. Harshman,K.D. and Dervan,P.B. (1985) Molecular recognition of B-DNA by Hoechst 33258. *Nucleic Acids Res.*, **13**, 4825–4835.

52. Tsunoda,M., Sakaue,T., Naito,S., Sunami,T., Abe,N., Ueno,Y., Matsuda,A. and Takénaka,A. (2010) Insights into the structures of DNA damaged by hydroxyl radical: crystal structures of DNA duplexes containing 5-formyluracil. *J. Nucleic Acids*, **2010**, 1–10.

53. Spink,N., Brown,D.G., Skelly,J. V and Neidle,S. (1994) Sequence-dependent effects in drug-DNA interaction : the crystal structure of Hoechst 33258 bound to the. *Nucleic Acids Res.*, **22**, 1607–1612.

54. Bazhulina,N.P., Nikitin, a M., Rodin,S. a, Surovaya, a N., Kravatsky,Y. V, Pismensky,V.F., Archipova,V.S., Martin,R. and Gursky,G. V (2009) Binding of Hoechst 33258 and its derivatives to DNA. *J. Biomol. Struct. Dyn.*, **26**, 701–718.

55. Higgins,L.D. and Searle,M.S. (1999) Site-specificity of bis-benzimidazole Hoechst 33258 in A-tract recognition of the DNA dodecamer duplex d(GCAAAATTTTGC)2. *Chem. Commun.*, **1**, 1861–1862.

56. Roos,W.P. and Kaina,B. (2006) DNA damage-induced cell death by apoptosis. *Trends Mol. Med.*, **12**, 440–450.

57. Loeb,L.A. and Monnat,R.J. (2008) DNA polymerases and human disease. *Nat. Rev. Genet.*, **9**, 594–604.

58. Jackson,S.P. and Bartek,J. (2009) The DNA-damage response in human biology and disease. *Nature*, **461**, 1071–1078.

59. Shrivastav,M., De Haro,L.P. and Nickoloff,J.A. (2008) Regulation of DNA double-strand break repair pathway choice. *Cell Res.*, **18**, 134–147.

60. Heyer,W.-D., Ehmsen,K.T. and Liu,J. (2010) Regulation of homologous recombination in eukaryotes. *Annu. Rev. Genet.*, **44**, 131–139.

61. San Filippo,J., Sung,P. and Klein,H. (2008) Mechanism of Eukaryotic Homologous Recombination. *Annu. Rev. Biochem.*, **77**, 229–257.

62. Candelli,A., Modesti,M., Peterman,E.J.G. and Wuite,G.J.L. (2013) Single-molecule views on homologous recombination. *Q. Rev. Biophys.*, **46**, 323–348.

63. Gustafsson,C.M. (2015) Scientific Background on the Nobel Prize in Chemistry 2015. Mechanistic sttudies of DNA repair The Royal Swedish Academy of Sciences.

64. Lord,C.J. and Ashworth,A. (2012) The DNA damage response and cancer therapy. *Nature*, **481**, 287–294.

65. Pâques,F. and Haber,J.E. (1999) Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae. *Microbiol. Mol. Biol. Rev.*, **63**, 349–404.

66. Chapman,J.R., Taylor,M.R.G. and Boulton,S.J. (2012) Playing the End Game: DNA Double-Strand Break Repair Pathway Choice. *Mol. Cell*, **47**, 497–510.

67. Davis,L. and Smith,G.R. (2001) Meiotic recombination and chromosome segregation in Schizosaccharomyces pombe. *Proc. Natl. Acad. Sci. U. S. A.*, **98**, 8395–8402.

68. Slupianek,A., Schmutte,C., Tombline,G., Nieborowska-Skorska,M., Hoser,G., Nowicki,M.O., Pierce,A.J., Fishel,R. and Skorski,T. (2001) BCR/ABL regulates mammalian RecA homologs, resulting in drug resistance. *Mol. Cell*, **8**, 795–806.

69. Sonoda,E., Sasaki,M.S., Buerstedde,J.M., Bezzubova,O., Shinohara,A., Ogawa,H., Takata,M., Yamaguchi-Iwai,Y. and Takeda,S. (1998) Rad51-deficient vertebrate cells accumulate chromosomal breaks prior to cell death. *EMBO J.*, **17**, 598–608.

70. Collis,S.J., Tighe, a, Scott,S.D., Roberts,S. a, Hendry,J.H. and Margison,G.P. (2001) Ribozyme minigene-mediated RAD51 down-regulation increases radiosensitivity of human prostate cancer cells. *Nucleic Acids Res.*, **29**, 1534–1538.

71. Maacke,H., Opitz,S., Jost,K., Hamdorf,W., Henning,W., Krüger,S., Feller, a C., Lopens,A., Diedrich,K., Schwinger,E., et al. (2000) Over-expression of wild-type Rad51 correlates with histological grading of invasive ductal breast cancer. *Int. J. cancer*, **88**, 907–913.

72. Richardson,C., Stark,J.M., Ommundsen,M. and Jasin,M. (2004) Rad51 overexpression promotes alternative double-strand break repair pathways and genome instability. *Oncogene*, **23**, 546–553.

73. Lim,D.S. and Hasty,P. (1996) A mutation in mouse rad51 results in an early embryonic lethal that is suppressed by a mutation in p53. *Mol. Cell. Biol.*, **16**, 7133–7143.

74. Wold,M.S. (1997) Replication protein A: A heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu. Rev. Biochem.*, **66**, 61–92.

75. Richardson,P.R., Chapman,M. a., Wilson,D.C., Bates,S.P. and Jones,A.C. (2002) The nature of conformational preference in a number of p-alkyl phenols and p-alkyl benzenes. *Phys. Chem. Chem. Phys.*, **4**, 4910–4915.

76. Renkawitz,J., Lademann,C. a., Kalocsay,M. and Jentsch,S. (2013) Monitoring Homology Search during DNA Double-Strand Break Repair In Vivo. *Mol. Cell*, **50**, 261–272.

77. Forget,A.L. and Kowalczykowski,S.C. (2012) Single-molecule imaging of DNA pairing by RecA reveals a three-dimensional homology search. *Nature*, **482**, 423–427.

78. Renkawitz,J., Lademann,C.A. and Jentsch,S. (2014) Mechanisms and principles of homology search during recombination. *Nat. Rev. Mol. Cell Biol.*, **15**, 369–383.

79. Ragunathan,K., Liu,C. and Ha,T. (2012) RecA filament sliding on DNA facilitates homology search. *Elife*, **1**, e00067.

80. Qi,Z., Redding,S., Lee,J.Y., Gibb,B., Kwon,Y., Niu,H., Gaines,W.A., Sung,P. and Greene,E.C. (2015) DNA Sequence Alignment by Microhomology Sampling during Homologous Recombination. *Cell*, **160**, 856–869.

81. Nandakumar,D. and Patel,S.S. (2015) Finding the Right Match Fast. *Cell*, **160**, 809–811.

82. Holthausen,J.T., Wyman,C. and Kanaar,R. (2010) Regulation of DNA strand exchange in homologous recombination. *DNA Repair (Amst).*, **9**, 1264–1272.

83. van der Heijden,T., Modesti,M., Hage,S., Kanaar,R., Wyman,C. and Dekker,C. (2008) Homologous Recombination in Real Time: DNA Strand Exchange by RecA. *Mol. Cell*, **30**, 530–538.

84. Game,J.C. and Mortimer,R.K. (1974) A genetic study of x-ray sensitive mutants in yeast. *Mutat. Res.*, **24**, 281–292.

85. Haynes,R.H. and Kunz,B.A. (1981) DNA Repair and Mutagenesis in Yeast. In *The Molecular Biology of the Yeast Saccharomyces: Life Cycle and Inheritance*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, Vol. 28, pp. 371–414.

86. Game,J.C. (1983) Radiation-Sensitive Mutants and Repair in Yeast. In *Yeast Genetics*. Springer Series in Molecular Biology, New York, pp. 109–137.

87. Aboussekhra, a, Chanet,R., Adjiri, a and Fabre,F. (1992) Semidominant suppressors of Srs2 helicase mutations of Saccharomyces cerevisiae map in the RAD51 gene, whose sequence predicts a protein with similarities to procaryotic RecA proteins. *Mol. Cell. Biol.*, **12**, 3224–3234.

88. Basile,G., Aker,M. and Mortimer,R.K. (1992) Nucleotide sequence and transcriptional regulation of the yeast recombinational repair gene RAD51. *Mol. Cell. Biol.*, **12**, 3235–3246.

89. Shinohara,A., Ogawa,H. and Ogawa,T. (1992) Rad51 protein involved in repair and recombination in S. cerevisiae is a RecA-like protein. *Cell*, **69**, 457–470.

90. Tombline,G. and Fishel,R. (2002) Biochemical characterization of the human RAD51 protein. I. ATP hydrolysis. *J. Biol. Chem.*, **277**, 14417–14425.

91. Ogawa,T., Yu,X., Shinohara,A. and Egelman,E.H. (1993) Similarity of the yeast RAD51 filament to the bacterial RecA filament. *Science (80-.).*, **259**, 1896–1899.

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

93. Bugreev,D. V and Mazin,A. V (2004) Ca2+ activates human homologous recombination protein Rad51 by modulating its ATPase activity. *Proc. Natl. Acad. Sci. U. S. A.*, **101**, 9988–9993.

94. Sung,P. and Robberson,D.L. (1995) DNA strand exchange mediated by a RAD51-ssDNA nucleoprotein filament with polarity opposite to that of RecA. *Cell*, **82**, 453–461.

95. Yu,X., Jacobs,S. a, West,S.C., Ogawa,T. and Egelman,E.H. (2001) Domain structure and dynamics in the helical filaments formed by RecA and Rad51 on DNA. *Proc. Natl. Acad. Sci. U. S. A.*, **98**, 8419–8424.

96. Danilowicz,C., Peacock-Villada,A., Vlassakis,J., Facon,A., Feinstein,E., Kleckner,N. and Prentiss,M. (2014) The differential extension in dsDNA bound to Rad51 filaments may play important roles in homology recognition and strand exchange. *Nucleic Acids Res.*, **42**, 526–533.

97. Chen,Z., Yang,H. and Pavletich,N.P. (2008) Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature*, **453**, 489–494.

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

99. Chabbert,M., Cazenave,C. and Hélène,C. (1987) Kinetic studies of recA protein binding to a fluorescent single-stranded polynucleotide. *Biochemistry*, **26**, 2218–2225.

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

101. Krejci,L., Altmannova,V., Spirek,M. and Zhao,X. (2012) Homologous recombination and its regulation. *Nucleic Acids Res.*, **40**, 5795–5818.

102. Lisby,M., Barlow,J.H., Burgess,R.C. and Rothstein,R. (2004) Choreography of the DNA damage response: Spatiotemporal relationships among checkpoint and repair proteins. *Cell*, **118**, 699–713.

103. Sung,P., Krejci,L., Van Komen,S. and Sehorn,M.G. (2003) Rad51 recombinase and recombination mediators. *J. Biol. Chem.*, **278**, 42729–42732.

104. Sung,P. and Klein,H. (2006) Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat. Rev. Mol. Cell Biol.*, **7**, 739–750.

105. New,J.H., Sugiyama,T., Zaitseva,E. and Kowalczykowski,S.C. (1998) Rad52 protein stimulates DNA strand exchange by Rad51 and replication protein A. *Nature*, **391**, 407–410.

106. Shinohara, a and Ogawa,T. (1998) Stimulation by Rad52 of yeast Rad51-mediated recombination. *Nature*, **391**, 404–407.

107. Sugiyama,T. and Kowalczykowski,S.C. (2002) Rad52 protein associates with replication protein A (RPA)-single-stranded DNA to accelerate Rad51-mediated displacement of RPA and presynaptic complex formation. *J. Biol. Chem.*, **277**, 31663–31672.

108. Carreira,A., Hilario,J., Amitani,I., Baskin,R.J., Shivji,M.K.K., Venkitaraman,A.R. and Kowalczykowski,S.C. (2009) The BRC repeats of BRCA2 modulate the DNA-binding selectivity of RAD51. *Cell*, **136**, 1032–1043.

109. Akamatsu,Y. and Jasin,M. (2010) Role for the mammalian Swi5-Sfr1 complex in DNA strand break repair through homologous recombination. *PLoS Genet.*, **6**, e1001160.

110. Kuwabara,N., Murayama,Y., Hashimoto,H., Kokabu,Y., Ikeguchi,M., Sato,M., Mayanagi,K., Tsutsui,Y., Iwasaki,H. and Shimizu,T. (2012) Mechanistic insights into the activation of Rad51-mediated strand exchange from the structure of a recombination activator, the Swi5-Sfr1 complex. *Structure*, **20**, 440–449.

111. Kokabu,Y., Murayama,Y., Kuwabara,N., Oroguchi,T., Hashimoto,H., Tsutsui,Y., Nozaki,N., Akashi,S., Unzai,S., Shimizu,T., et al. (2011) Fission yeast Swi5-Sfr1 protein complex, an activator of Rad51 recombinase, forms an extremely elongated dogleg-shaped structure. *J. Biol. Chem.*, **286**, 43569–43576.

112. Kurokawa,Y., Murayama,Y., Haruta-Takahashi,N., Urabe,I. and Iwasaki,H. (2008) Reconstitution of DNA strand exchange mediated by Rhp51 recombinase and two mediators. *PLoS Biol.*, **6**, e88.

113. Tsai,S.-P., Su,G.-C., Lin,S.-W., Chung,C.-I., Xue,X., Dunlop,M.H., Akamatsu,Y., Jasin,M., Sung,P. and Chi,P. (2012) Rad51 presynaptic filament stabilization function of the mouse Swi5-Sfr1 heterodimeric complex. *Nucleic Acids Res.*, **40**, 6558–6569.

114. Kowalczykowski,S.C. and Eggleston, a K. (1994) Homologous pairing and DNA strand-exchange proteins. *Annu. Rev. Biochem.*, **63**, 991–1043.

115. Chi,P., Van Komen,S., Sehorn,M.G., Sigurdsson,S. and Sung,P. (2006) Roles of ATP binding and ATP hydrolysis in human Rad51 recombinase function. *DNA Repair (Amst).*, **5**, 381–391.

116. Modesti,M., Ristic,D., van der Heijden,T., Dekker,C., van Mameren,J., Peterman,E.J.G., Wuite,G.J.L., Kanaar,R. and Wyman,C. (2007) Fluorescent human RAD51 reveals multiple nucleation sites and filament segments tightly associated along a single DNA molecule. *Structure*, **15**, 599–609.

117. Symington,L.S. and Heyer,W.-D. (2006) Some disassembly required: role of DNA translocases in the disruption of recombination intermediates and dead-end complexes. *Genes Dev.*, **20**, 2479–2486.

118. Cox,M.M. (2003) The bacterial RecA protein as a motor protein. *Annu. Rev. Microbiol.*, **57**, 551–577.

119. Hilario,J., Amitani,I., Baskin,R.J. and Kowalczykowski,S.C. (2009) Direct imaging of human Rad51 nucleoprotein dynamics on individual DNA molecules. *Proc. Natl. Acad. Sci. U. S. A.*, **106**, 361–368.

120. van Mameren,J., Modesti,M., Kanaar,R., Wyman,C., Peterman,E.J.G. and Wuite,G.J.L. (2009) Counting RAD51 proteins disassembling from nucleoprotein filaments under tension. *Nature*, **457**, 745–748.

121. Su,G.-C., Chung,C.-I., Liao,C.-Y., Lin,S.-W., Tsai,C.-T., Huang,T., Li,H.-W. and Chi,P. (2014) Enhancement of ADP release from the RAD51 presynaptic filament by the SWI5-SFR1 complex. *Nucleic Acids Res.*, **42**, 349–358.

122. Nordén,B., Rodger,A. and Dafforn,T. (2010) Linear Dichroism and Circular Dichroism: A Textbook on Polarized-Light Spectroscopy. Royal Society of Chemistry, Cambridge, UK.

123. Abbe,E. (1873) Beitrage zur Theorie des Mikroskops und der mikroskopischen Wahrmehmung. *Arch. für Mikroskopische Anat. (in Ger.*, **9**, 413–420.

124. Ehrenberg,M. (2014) Scientific Background on the Nobel Prize in Chemistry 2014. Super-Resolved Fluorescence Microscopy The Royal Swedish Academy of Sciences.

125. Persson,F., Utko,P., Reisner,W., Larsen,N.B. and Kristensen,A. (2009) Confinement spectroscopy: probing single DNA molecules with tapered nanochannels. *Nano Lett.*, **9**, 1382–1385.

126. Levy,S.L. and Craighead,H.G. (2010) DNA manipulation, sorting, and mapping in nanofluidic systems. *Chem. Soc. Rev.*, **39**, 1133–1152.

127. Nyberg,L.K., Persson,F., Berg,J., Bergström,J., Fransson,E., Olsson,L., Persson,M., Stålnacke,A., Wigenius,J., Tegenfeldt,J.O., et al. (2012) A single-step competitive binding assay for mapping of single DNA molecules. *Biochem. Biophys. Res. Commun.*, **417**, 404–408.

128. Nilsson,A.N., Emilsson,G., Nyberg,L.K., Noble,C., Svensson Stadler,L., Fritzsche,J., Moore,E.R.B., Tegenfeldt,J.O., Ambjörnsson,T. and Westerlund,F. (2014) Competitive binding-based optical DNA mapping for fast identification of bacteria - multi-ligand transfer matrix theory and experimental applications on Escherichia coli. *Nucleic Acids Res.*, **42**, 1–12.

129. Persson,F., Fritzsche,J., Mir,K.U., Modesti,M., Westerlund,F. and Tegenfeldt,J.O. (2012) Lipid-based passivation in nanofluidics. *Nano Lett.*, **12**, 2260–2265.

130. Bianco,P.R., Tracy,R.B. and Kowalczykowski,S.C. (1998) DNA strand exchange proteins: a biochemical and physical comparison. *Front. Biosci.*, **3**, D570–D603.

131. Fasman,G.D. (1975) Nucleic Acids. In *Handbook of Biochemistry and Molecular Biology*. CRC Press, Cleveland, Ohio, pp. 66–215.

132. Fasman,G.D. (1976) Proteins. In *Handbook of Biochemistry and Molecular Biology*. CRC Press, Cleveland, Ohio, pp. 183–203.

133. Nomme,J., Takizawa,Y., Martinez,S.F., Renodon-Cornière,A., Fleury,F., Weigel,P., Yamamoto,K., Kurumizaka,H. and Takahashi,M. (2008) Inhibition of filament formation of human Rad51 protein by a small peptide derived from the BRC-motif of the BRCA2 protein. *Genes to cells*, **13**, 471–481.

134. Fujise,H., Cruz,P., Reo,N. V and Lauf,P.K. (1991) Relationship between total magnesium concentration and free intracellular magnesium in sheep red blood cells. *Biochim. Biophys. Acta*, **1094**, 51–54.

135. Keith,C.H., Maxfield,F.R. and Shelanski,M.L. (1985) Intracellular free calcium levels are reduced in mitotic Pt K2 epithelial cells. *Proc. Natl. Acad. Sci. U. S. A.*, **82**, 800–804.

136. Maravall,M., Mainen,Z.F., Sabatini,B.L. and Svoboda,K. (2000) Estimating intracellular calcium concentrations and buffering without wavelength ratioing. *Biophys. J.*, **78**, 2655–2667.

137. Rogers,D.M. and Hirst,J.D. (2003) Ab Initio Study of Aromatic Side Chains of Amino Acids in Gas Phase and Solution. *J. Phys. Chem. A*, **107**, 11191–11200.

138. Reymer,A., Frykholm,K., Morimatsu,K., Takahashi,M. and Nordén,B. (2009) Structure of human Rad51 protein filament from molecular modeling and site-specific linear dichroism spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.*, **106**, 13248–13253.

139. Freyer,M.W., Buscaglia,R., Cashman,D., Hyslop,S., Wilson,W.D., Chaires,J.B. and Lewis,E.A. (2007) Binding of netropsin to several DNA constructs: evidence for at least two different 1:1 complexes formed from an -AATT-containing ds-DNA construct and a single minor groove binding ligand. *Biophys. Chem.*, **126**, 186–196.

140. Freyer,M.W., Buscaglia,R., Hollingsworth,A., Ramos,J., Blynn,M., Pratt,R., Wilson,W.D. and Lewis,E.A. (2007) Break in the heat capacity change at 303 K for complex binding of netropsin to AATT containing hairpin DNA constructs. *Biophys. J.*, **92**, 2516–2522.