CHALMERS

Aspects of DNA Strand Exchange: Recombination Proteins and Model System Studies

KAROLIN FRYKHOLM

Department of Chemical and Biological Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden 2010

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Aspects of DNA Strand Exchange: Recombination Proteins and Model System Studies

KAROLIN FRYKHOLM

Department of Chemical and Biological Engineering

CHALMERS UNIVERSITY OF TECHNOLOGY

Göteborg, Sweden 2010

Aspects of DNA Strand Exchange:
Recombination Proteins and Model System Studies

KAROLIN FRYKHOLM

ISBN 978-91-7385-448-1

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 3129
ISSN 0346-718X

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

Cover:
Left: a structure of the human Rad51 protein filament constructed from molecular modeling with angular constraints from SSLD data, as described on page 29ff of this Thesis. *Right:* a schematic illustration of strand exchange of a 20-mer oligonucleotide. Artificial DNA strand exchange in model systems is presented on page 35ff of this Thesis.

Printed by Chalmers Reproservice
Göteborg, Sweden 2010

Aspects of DNA Strand Exchange:
Recombination Proteins and Model System Studies

KAROLIN FRYKHOLM

Department of Chemical and Biological Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY

ABSTRACT

DNA recombination is of fundamental importance to all living cells; it is part of the DNA repair machinery and a means to generate genetic diversity. DNA strand exchange, the exchange of strands between homologous DNA molecules, is the central reaction of the recombination process. The work presented in this Thesis has the aim of gaining insight into the mechanism of this reaction by investigating different aspects of DNA strand exchange. Structural studies of recombinase nucleoprotein filaments, which constitute the scaffold for the reaction *in vivo*, are reported together with investigations of artificial strand exchange in two different model systems.

The structures of active RecA and Rad51 nucleoprotein filaments have been studied by Site-Specific Linear Dichroism (SSLD), a spectroscopic approach based on linear dichroism in combination with molecular replacement of individual amino acids. In this Thesis it is shown how LD data of systematically engineered proteins can provide angular orientations for specific residues and how these coordinates can be used to build a structural model of the protein. From SSLD data of RecA it is concluded that the protein adopts similar structures in the initial and final states of strand exchange, indicating a static role for RecA during the reaction. The study of the human Rad51 protein illustrates how experimental data from SSLD can be successfully combined with theoretical molecular modelling. The outcome is a model structure of the protein in its active complex with DNA, the first detailed structure reported so far for the complete human Rad51 protein.

This Thesis reports on artificial strand exchange catalysis aided by cationic lipid vesicles and it is shown that DNA opens up in a zipper-like fashion, which facilitates strand exchange. It is further concluded that the exchange mechanism on the liposome surface is fundamentally different from that in bulk solution. Non-ionic catalysis has been investigated by the use of polyethylene glycol (PEG) to induce molecular crowding and provide possibilities for hydrophobic interactions. PEG accelerates strand exchange dramatically and the results emphasize the importance of hydrophobic interactions between DNA and its environment for the dynamic behaviour of DNA strands.

Keywords: DNA strand exchange, RecA, Rad51, Site-Specific Linear Dichroism, cationic liposome, molecular crowding

LIST OF PUBLICATIONS

This Thesis is based on the work presented in the following papers:

Paper I

Conserved conformation of RecA protein after executing the DNA strand-exchange reaction. A site-specific linear dichroism structure study

Karolin Frykholm, Katsumi Morimatsu and Bengt Nordén

Biochemistry, **2006***, 45 (37), 11172-11178*

Paper II

Structure of human Rad51 protein filament from molecular modeling and site-specific linear dichroism spectroscopy

Anna Reymer, Karolin Frykholm, Katsumi Morimatsu, Masayuki Takahashi and Bengt Nordén

Proceedings of the National Academy of Sciences of the United States of America, **2009***, 106 (32), 13248-13253*

Paper III

Enhanced DNA strand exchange on positively charged liposomes

Karolin Frykholm, Francesca Baldelli Bombelli, Bengt Nordén and Fredrik Westerlund

Soft Matter, **2008***, 4 (12), 2500-2506*

Paper IV

Mechanism of DNA strand exchange at liposome surfaces investigated using mismatched DNA

Karolin Frykholm, Bengt Nordén and Fredrik Westerlund

Langmuir, **2009***, 25 (3), 1606-1611*

Paper V

DNA strand exchange catalyzed by molecular crowding in PEG solutions

Bobo Feng, Karolin Frykholm, Bengt Nordén and Fredrik Westerlund

Chemical Communications, **2010**, *DOI: 10.1039/C0CC03117H*

CONTRIBUTION REPORT

Paper I

The author of this Thesis, Karolin Frykholm (KF), performed the LD experiments, did most of the data analysis and wrote the paper.

Paper II

KF performed most of the experimental work, analysed the LD data and participated in interpreting the results and writing the paper. KF did not perform the molecular modelling.

Paper III

KF performed all the experimental work, analysed the data and wrote the paper.

Paper IV

KF performed all the experimental work, analysed the data and participated in writing the paper.

Paper V

KF supervised the experimental work and participated in interpreting the results and writing the paper.

CONTENTS

1. INTRODUCTION

IN ALL LIVING cells DNA is the storage molecule for the hereditary information and by replication of DNA this information is passed on to coming generations. According to the central dogma of molecular biology DNA is transcribed into RNA that eventually is translated into proteins. Proteins act as building blocks of the cells and are responsible for molecular recognition and catalysis of biochemical reactions, thus of importance for both cellular structure and function. Damage to DNA will affect both the replication of DNA and the downstream processes, possibly resulting in misfolded and malfunctioning proteins or even failed protein synthesis. A repair system for DNA is therefore crucial for the survival of individual cells. In a longer perspective an organism is dependent on adaption to its surroundings and a changing environment by rearrangement of its genetic material, creating new combinations of genes and modifying the regulations of their expression. Accordingly, alongside the DNA repair system there is a need for a system to generate genetic diversity. These two polarized events, preventing or causing changes in the genetic material, are both facilitated by DNA recombination.

Considering the importance of DNA recombination for cell survival, it is not surprising that all organisms have machinery for that purpose. In fact, a common system is found in species from all domains of life, where the main recombination protein has been evolutionary conserved from bacteria and archaea to eukaryotes like yeast and human [1]. The RecA family of recombinases, including bacterial RecA, archaeal RadA and eukaryotic Dmc1 and Rad51, show structural and functional similarities, such as formation of a helical nucleoprotein filament on DNA, where DNA is stretched and unwound, DNA-dependent ATPase activity and catalysis of the central reaction of recombination: the exchange of strands between homologous DNA molecules [2-6].

In vivo DNA recombination is a complex process: in *E. coli* bacteria at least 25 different proteins are involved [7]. One approach to understanding a matter of this kind, a multistep reaction in a complicated biological context, is to study structure and function of individual proteins and their responses to different stimuli *in vitro*. The bacterial RecA, specifically the one from *E. coli*, has for long served as a model protein for studies of recombinases and their activities. Structural information provides an important basis for mechanistic insights into DNA recombination and strand exchange and the structural characteristics of the RecA nucleoprotein complex have during the last decades been explored by various methods including electron microscopy [8-9], small-angle neutron scattering [10-11] and linear dichroism spectroscopy [12-14]. High resolution crystal structures of the RecA monomer and polymer were determined almost

twenty years ago [15] and recently crystal structures of RecA complexes with single- and double-stranded DNA were presented [16]. However, despite the rich knowledge gained over years of extensive research there are still many details of the reaction mechanism that remain to be elucidated. Moreover, investigations regarding structure as well as function of the non-bacterial RecA protein homologues in general and the human Rad51 protein in particular, are far from complete.

In this Thesis structural information on recombination proteins in their nucleoprotein complexes with DNA has been derived by site-specific linear dichroism (SSLD), a spectroscopic approach based on linear dichroism in combination with molecular replacement. In the SSLD methodology single amino acid substitutions in a protein, where an aromatic residue is selectively replaced by an optically "invisible" residue, allow for determination of orientation angles of the replaced residue. The approach has previously been applied to RecA [12, 17] and RecA was also the target protein of the study in Paper I, where structural similarities and differences between complexes with single- and double-stranded DNA are discussed. The SSLD technique can, in combination with molecular modelling, serve as a powerful tool to determine the three-dimensional structure of fibrous nucleoprotein complexes, systems not amenable to X-ray crystallography or NMR, as shown for the human Rad51 protein in Paper II.

A further simplified approach to the study of DNA recombination is to focus on the main reaction itself, the DNA strand exchange. By designing simple model systems individual parameters of importance to the reaction can be systematically investigated. *In vitro* DNA strand exchange has been found to be accelerate by cationic polypeptides [18] or a synthetic cationic polymer, with varied functionalities, as catalysing agent [19-20]. In the work presented in the second part of this Thesis two different model systems for artificial DNA strand exchange have been investigated. In Papers III and IV the concept of cationic catalysis, aided by cationic liposomes, was applied. In Paper V catalysis was instead obtained by molecular crowding and hydrophobic interactions using polyethylene glycol (PEG), mimicking the environment inside the recombinase nucleoprotein filament.

In summary, the work presented in this Thesis deals with two aspects of DNA strand exchange; recombination proteins and artificial strand exchange in model systems. Recombinases like RecA and Rad51 constitute the scaffold for the DNA strand exchange reaction *in vivo*. When using a model system an artificial scaffold is designed with the intention to identify and investigate individual parameters affecting the reaction. Both of these perspectives on DNA strand exchange can contribute with small pieces to the challenging puzzle of understanding in detail the fundamental and over millions of years of evolution conserved process of DNA recombination.

2. BACKGROUND

This chapter is intended to describe the basics of DNA and DNA strand exchange, and introduce the reader to the RecA family of recombinases, their activities and structural characteristics.

2.1. DNA

THAT DNA, OR deoxyribonucleic acid, is the holder of the genetic information was demonstrated in 1944 by Avery and co-workers [21], although it was not evident how this relatively simple polymeric molecule could carry all the information needed for a cell and how that information could be transmitted and passed on from one generation to the next. In the classical paper from 1953, Watson and Crick proposed a helical structure of two complementary DNA strands with a specific base pairing pattern, suggesting a possible copying mechanism that would facilitate heredity of the genetic material [22]. Following this discovery the central dogma of molecular biology, the transfer of genetic information from DNA to protein via RNA, was formulated [23] and it was described how three DNA bases make up a codon for one of the twenty amino acids used by nature for protein synthesis [24].

The DNA polymer is built up of four different monomeric units, the nucleotides adenosine-, guanosine-, cytidine- and thymidine monophosphate. Each nucleotide consists of a sugar, a deoxyribose, with a phosphate group attached to its 5' carbon and one of the aromatic purine (adenine, A, and guanine, G) or pyrimidine (cytosine, C, and thymine, T) bases at its 1' carbon. The phosphate of one nucleotide connects to the 3' carbon of the sugar of another nucleotide to form a polynucleotide chain with a backbone of alternating sugar and phosphate groups, as illustrated in Figure 2.1 (left). Two such chains form the characteristic double helix of DNA, with the negatively charged phosphates on the outside and the hydrophobic bases on the inside of the helix. The strands are held together by the specific base pairing between purine and pyrimidine bases: A is always paired, with two hydrogen bonds, with T and G is always paired, with three hydrogen bonds, with C (Figure 2.1, right). Hydrophobic interactions between the bases stacking onto each other stabilize the helix [25-27].

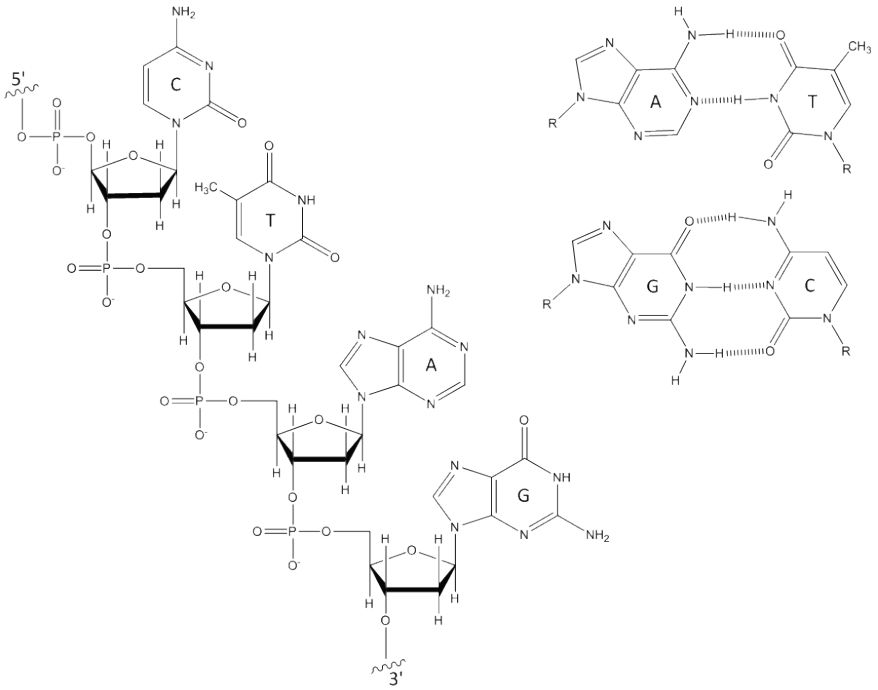

Figure 2.1 A fragment of a DNA polynucleotide chain, left, showing the backbone of alternating sugar and phosphate groups, and the four bases cytosine (C), thymine (T), adenine (A) and guanine (G). The specific base pairing pattern is shown to the right (R = backbone).

In its most common conformation, known as B-form and shown in Figure 2.2, DNA forms a double-stranded, right-handed helix with the two strands wound about each other in anti-parallel direction and the bases oriented with their planes almost perpendicular to the helix axis. It comprises ten base pairs per helical turn with a helical rise of 3.4 Å per base pair. Two deep grooves run along the helix side, one wide and one narrow, referred to as the major and minor groove, respectively. There are other, biologically relevant, conformations of DNA, *e.g.* the wider and flatter A-form and the left-handed Z-form [28].

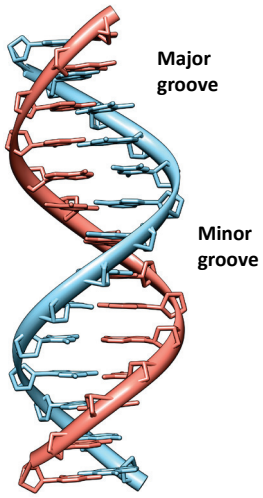

Figure 2.2 A representation of a DNA double helix in the B-form conformation.

2.1.1. DNA Recombination and Strand Exchange

The correct replication of DNA, facilitated by the complementarity of the two strands in the DNA duplex, is fundamental for the heredity of genetic information. The matching in sequence between the two strands also allows for homologous recombination, the exchange of strands between DNA molecules. Homologous recombination is the process resulting in new genetic combinations in the offspring upon the union of maternal and paternal cells, thereby generating genetic diversity. It is also an important pathway for DNA repair; in bacteria recombinational DNA repair is the major function of homologous recombination [29].

Several models describing DNA recombination have been proposed. The double-strand break repair model, illustrated in Figure 2.3, applies to homologous recombination in all organisms [30-31]. At a double-stranded (ds) break, repair is initiated by processing of the DNA to give single-stranded (ss) overhangs. A recombinase and supporting proteins are recruited to the ssDNA, forming a nucleoprotein filament to promote homologous pairing and strand exchange. Repair synthesis and capture of the second end of the double strand break results in the formation of a branched intermediate, known as a Holliday junction. Following branch migration and completion of the repair process by DNA synthesis and ligation, the recombined DNA molecules are separated. Many proteins are involved in directing the recombinational repair pathways *in vivo*, common to most of the proposed paths and highly important to the central step of strand exchange is a recombinase of the RecA family [3, 7, 32]. This protein family, conserved over all domains of life, will be further introduced below.

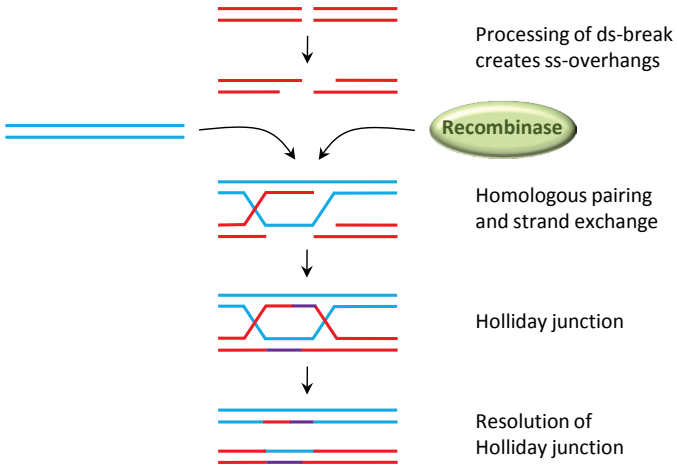

Figure 2.3 The general steps of homologous recombination according to the double-strand break repair model. Repair of a double-strand break is initiated by processing of the DNA to create single-stranded overhangs. Homologous pairing and strand exchange takes place inside the recombinase filament formed on ssDNA. Repair synthesis and ligation results in a characteristic branched intermediate, known as a Holliday junction. Upon resolution of the Holliday junction the recombined DNA molecules are separated.

2.2. RECOMBINATION PROTEINS

DNA STRAND EXCHANGE, the pairing of two DNA molecules of identical or nearly identical sequence and the transfer of a single strand between them, is one of the main activities of the recombinases of the RecA family. The importance of strand exchange in recombinational repair of DNA damages causing interrupted replication, a fundamental process for cell survival, is most likely the reason for the evolutionary conservation of these proteins. The mediation of information exchange between the genomes of parental cells is another important role of DNA recombination and the *recA* gene of *E. coli* was first identified following isolation of mutants deficient in conjugal recombination [33]. RecA is present in virtually all bacteria and homologous proteins have been identified in archaea as well as eukaryotes [2, 4-6, 34]. The sequence is well conserved among bacterial RecA proteins with identity ranging from 49% to 100% [29]. The eukaryotic recombinases also show high mutual sequence similarity, with 89% homology and 68% identity in sequence between yeast and human Rad51 proteins (given by a ClustalW [35] sequence alignment). Higher eukaryotic proteins are almost identical in sequence; mouse and human Rad51 proteins differ by four amino acids only [34]. The sequence conservation from bacterial to archaeal and eukaryotic recombinases is limited to the core region of RecA, as shown in the comparison in Figure 2.4. The RadA, Rad51 and Dmc1 proteins lack the C-terminal domain present in RecA, while having an extended N-terminal region with a domain conserved within the group. For the human Rad51 protein the sequence homology with the core region of RecA is about 50%.

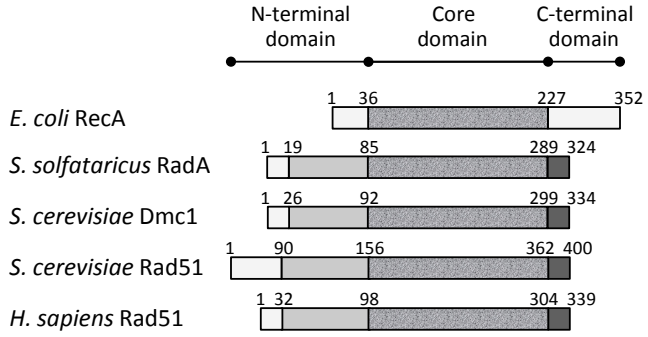

Figure 2.4 A schematic sequence comparison of recombinases from all domains of life. The core domain (spotted regions) is conserved among all species in the comparison. The archaeal and eukaryotic proteins share a conserved N-terminal domain (light grey regions) without counterpart in the bacterial RecA, and are homologous also in their C-terminus (dark grey regions). The C-terminal domain present in RecA is absent in the recombinases from archaea, yeast and human. White regions indicate non-homologous sequences. Amino acid numberings are taken from the multiple sequence alignments by Brendel *et al.* [1].

Along with the sequence conservation the RecA family of recombinases share structural features and have common functionalities although there are also differences between them in these respects, particularly between proteins of prokaryotic and eukaryotic origin. These recombinases all promote DNA strand exchange and the functional structure for this activity is a right-handed helical nucleoprotein filament. In presence of ATP (or an ATP analogue) bacterial RecA as well as yeast and human Rad51 assemble onto single- or double-stranded DNA, forming an extended helical filament that can span thousands of bases or base pairs. The overall structural characteristics of the nucleoprotein filaments formed by RecA and Rad51 proteins are similar; one helical turn is completed by about six protein monomers and harbours approximately 18 bases or base pairs. The filament is narrow and has a pitch of about 95±5 Å [8-9, 36]. The DNA bound in the complex adopts an extended conformation; it is stretched by approximately 50% compared to normal B-DNA, to an average axial rise of 5.2 Å per base or base pair, and dsDNA is unwound by 15° per base pair step [37-38]. The recent crystal structures of RecA-DNA complexes show that the DNA elongation is non-uniform, with a conformation similar to B-DNA for triplets of nucleotides compensated by a larger rise and negative twist at every third base or base pair step [16]. These findings verify earlier suggestions based on theoretical arguments [39]. The extended nucleoprotein filament is often referred to as an "active" filament, in contrast to the compressed and wider filaments observed in absence of DNA or with ADP as cofactor (Figure 2.5).

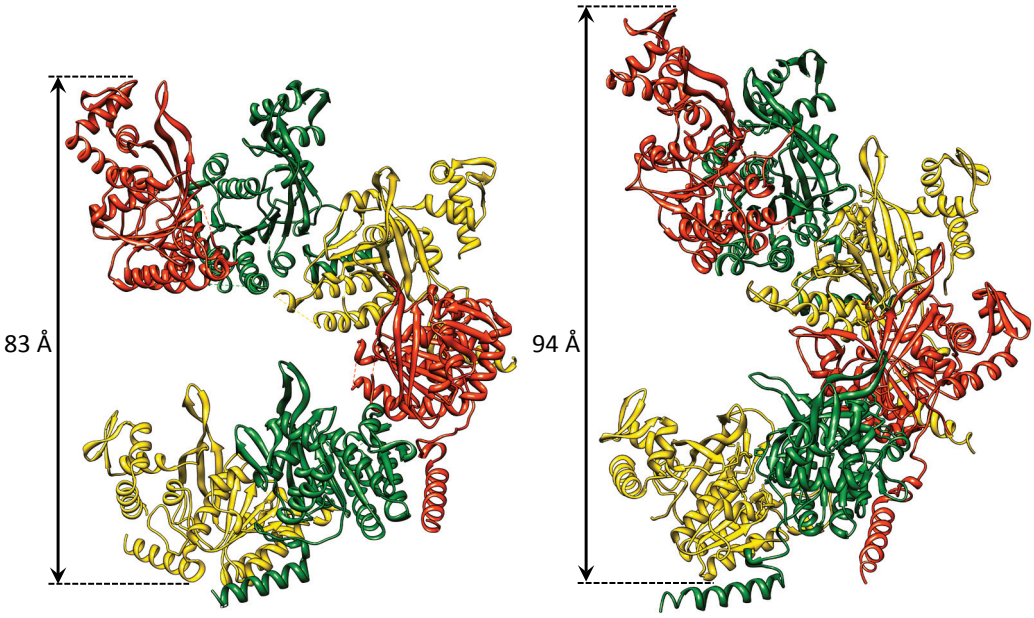

Figure 2.5 A representation of an inactive (left) and an active (right) RecA filament. The inactive filament formed in absence of DNA with ADP as cofactor is compressed and wider than the active filament, which appears in the presence of DNA with ATP as cofactor. (PDB entries: 2REB and 3CMU, respectively.)

Both RecA and Rad51 proteins exhibit DNA dependent ATPase activity, although ATP is hydrolysed at significantly higher rates by RecA than by the eukaryotic recombinases [40]. The role of ATP hydrolysis is not fully clear, and might differ between the prokaryotic and eukaryotic proteins. Several models that couple the hydrolysis of ATP by RecA to DNA strand exchange have been proposed. RecA has been suggested to have an ATP-driven motor function, where it rotates the axes of two DNA molecules, one bound on the inside and the other on the outside of the filament, around each other [41-42]. In other models ATP hydrolysis is coupled to dissociation of RecA at the disassembling end of the filament [43], in some models in connection to a redistribution of the RecA monomers [44-46]. None of the eukaryotic recombinases have been ascribed a motor protein function, and findings coupling ATP hydrolysis to dissociation of protein monomers from the DNA support the dissociation model for these proteins [47-50].

For the non-bacterial RecA homologues ATP hydrolysis is the only documented activity, beside DNA-pairing and strand exchange. The RecA protein however, is a multifunctional enzyme with a range of activities coupled to the maintenance of genomic integrity. It has a regulatory role in induction of the SOS response to DNA damage, acting as a co-protease in the autocatalytic cleavage of LexA, UmuD and other proteins. RecA is also involved in SOS mutagenesis, by activation of the translesion DNA polymerase V, and it has been proposed to be active, possibly indirectly, in chromosome partitioning at cell division [3, 29].

Structural investigations of the *E. coli* RecA protein and the human Rad51 protein are part of the work of this Thesis; some characteristics of these proteins will therefore be highlighted below.

2.2.1. RecA

The RecA protein of *E. coli* is a 352 amino acid polypeptide of 38 kDa, but it polymerizes in solution into very high molecular weight filaments. The isoelectric point of the protein has been determined to 5.6, resulting in a net negative charge at physiological pH [29]. The first crystal structures of the *E. coli* RecA monomer and polymer were solved in 1992 at 2.3 Å resolution [15, 51] and since then several other high resolution structures of RecA from *E. coli* and other bacterial origin have been determined [16, 52-59]. As discussed above, the sequence similarity among bacterial RecA proteins is high and they are also structurally conserved. General structural features, and their functional implications, have been reviewed many years ago by Roca and Cox [29] and more recently by Cox [60].

A representation of an *E. coli* RecA monomer, as a superposition of the structure from the nucleoprotein filament with ssDNA solved by Chen *et al.* [16] on the crystal structure presented by Story *et al*. [15], is shown in Figure 2.6. The core domain, which includes the ATP binding site, is the most highly conserved part among bacterial RecA proteins, with structural homology also to eukaryotic recombinases. To promote DNA strand exchange the RecA nucleoprotein filament can bind three DNA strands: a primary ssDNA molecule and an incoming DNA duplex. The primary DNA binding has been ascribed mainly to the among bacteria well conserved regions around loops L1 and L2 (residues 151-176 and 190-227, respectively), based on mutagenesis [61-62] and DNA cross-linking studies [63-65] and confirmed by the recently presented crystal structures by Chen *et al.* [16]. Other residues that have been shown to be involved in DNA binding include Tyr65 [66], Tyr103 and Tyr264 [67] and the region of residues 233-243 [68]. Whereas the primary DNA binding site is located in the interior of the protein filament, a second binding site is proposed to be located on the exterior of the filament [69]. This site, involving the C-terminal domain, possibly acts as a gateway for the incoming dsDNA [70-71].

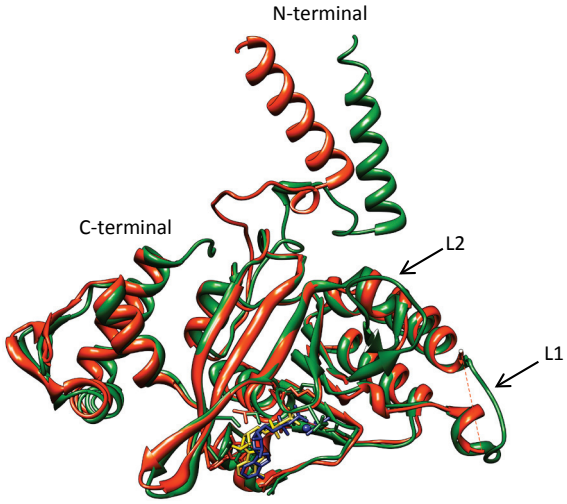

Figure 2.6 A superposition of two RecA monomer structures, one taken from an inactive filament formed with ADP in absence of DNA (red, with ADP shown in yellow; PDB entry 2REB) and one from an active filament with ssDNA and ADP-AlF_4-Mg (green, with ADP-AlF_4-Mg shown in blue; PDB entry 3CMU). Differences between the two structures are observed for the N-terminal domain and for the DNA binding loops L1 and L2, which are not visible in the inactive structure.

The appearance and activity of the helical filament formed by RecA is dependent on the presence of DNA and nucleotide cofactor [8, 60]. In absence of DNA, or at high levels of ADP, the protein is in an inactive state, characterized by a compressed, wide filament with a low pitch. The elongated active filament is formed with ssDNA and ATP. The ability of this state to interact with an incoming duplex DNA, for DNA strand exchange, is dependent on the magnesium ion

concentration. A high level of Mg^{2+}, accompanied by a conformational change of mainly the C-terminal domain, is required for elevated strand exchange capacity [72]. Finally, the postsynaptic state of the filament, with dsDNA bound, is structurally similar to the active, presynaptic state [16] but shows higher rates of disassembly.

2.2.2. RAD51

THE RAD51 PROTEIN is a eukaryotic homologue of the RecA protein. It is similar to RecA in size; the human Rad51 protein, HsRad51, has 338 amino acids and a molecular weight of 37 kDa. The core domain of RecA is shared by Rad51 but the eukaryotic protein has a truncated C-terminal and an extended N-terminal region. Rad51 proteins from yeast and human have, by electron microscopy, been observed to form nucleoprotein filaments similar to those formed by RecA [2, 9, 36]. High resolution structures, however, are so far limited to the core [73] and N-terminal domains [74] of HsRad51 and two crystal structures of ScRad51, the Rad51 protein from yeast *Saccharomyces cerevisiae*. [75-76] These ScRad51 structures are both high-pitch filaments, with a pitch of 130 Å compared to the usually observed pitch of approximately 95 Å for the extended, active filaments of RecA or Rad51 [8-9, 16, 36]. High pitches have, however, been reported from electron microscopy studies of both RecA [77] and Rad51 [78].

Similar to bacterial RecA, the eukaryotic Rad51 proteins have two flexible loops, L1 and L2, facing the interior of their helical filaments and putatively being involved in DNA binding. These loops are not visible in any of the filamentous crystal structures of ScRad51 [75-76] or in the structure of the core domain from HsRad51 [73], but mutagenesis and tryptophan fluorescence analysis of the loop regions (residues 230-236 for L1 and residues 268-292 for L2 in HsRad51) have confirmed the L1 loop being essential for DNA binding [79-80]. The L2 loop was in these studies concluded to be close to the DNA binding site but of less importance for DNA binding. However, in a mutational study of ScRad51 the L2 loop was shown to be involved in binding of dsDNA [81]. The N-terminal domain of HsRad51 has also been shown to bind DNA [74]. Mutations in this region do not prevent DNA binding but affect strand exchange activity [80] and the N-terminal has therefore been suggested as the second DNA binding site of HsRad51, possibly with a role similar to that of the C-terminal domain of RecA. The ATP binding site of Rad51 is, similar to that of RecA, positioned at the subunit-subunit interface in the filament [16, 75, 82]. Binding of ATP has been found to promote functionally important conformational changes, involving interactions between residues of adjacent protein monomers at this interface [83-84].

The strand exchange promoted by RecA is, as mentioned above, dependent on Mg^{2+}. In presence of Mg^{2+} the human Rad51 protein has a weaker strand exchange activity than RecA, but it has been shown to be activated by Ca^{2+} [48]. Rapid ATP hydrolysis in presence of Mg^{2+} turns the HsRad51 filament into the inactive ADP-bound state, while Ca^{2+} inhibits the ATPase activity and thereby preserves the active nucleoprotein filament and stimulates strand exchange activity. This effect could not be seen for ScRad51 protein but a stimulatory effect, although not by the same mechanism, by Ca^{2+} on the human recombinase Dmc1 has been observed [47].

2.3. Artificial DNA Strand Exchange

As described above, homologous recombination and DNA strand exchange *in vivo* are complex processes, and recombinases like RecA and Rad51 catalyse the transfer of strands between DNA molecules of very long sequence. The study of strand exchange in simple model systems *in vitro*, using short DNA oligonucleotides and artificial catalysing agents can increase the fundamental understanding of the reaction. Maruyama and co-workers have shown that a comb-type copolymer with a cationic poly(L-lysine) backbone and a dextran graft chain accelerates strand exchange of oligonucleotides *in vitro*, possibly by stabilization of the intermediate triplex [19, 85]. Also simple arginine-rich peptides have been demonstrated to exhibit an accelerating effect on the strand exchange reaction [18]. Notably, polyarginine showed a considerably higher activity than an equally charged polylysine peptide, an effect speculated to be due to hydrogen bonding interactions between arginine and DNA, in contrast to the primarily electrostatic interaction of lysine. Likewise, modification of the comb-type copolymer with guanidino groups, *i.e.* exchanging lysine for arginine in the polymer backbone, markedly increased its enhancing effect on strand exchange [20].

In the work of this Thesis artificial catalysis of DNA strand exchange has been studied in two different model systems. Catalysis based on electrostatic interactions using cationic liposomes was the approach taken in Papers III and IV. In Paper V the catalytic effect was instead due to molecular crowding and hydrophobic interactions induced by PEG.

3. FUNDAMENTAL CONCEPTS AND METHODOLOGY

In this chapter some fundamental concepts of importance to the work of this Thesis are presented and the main experimental techniques used are described.

3.1. SPECTROSCOPY

SPECTROSCOPY IS THE study of the interaction between electromagnetic radiation and matter. Several methods relying on the absorption or emission of light by molecules have been extensively used throughout the work of this Thesis. In the following section the theoretical basis of the applied spectroscopic techniques will be briefly described and some methodological aspects will be discussed. A more comprehensive presentation of spectroscopy in general, fluorescence spectroscopy and polarized spectroscopy can be found in the textbooks by Hollas, Lakowicz and Nordén *et al.*, respectively [86-88].

3.1.1. ABSORPTION SPECTROSCOPY

ELECTROMAGNETIC RADIATION CAN be described as a wave with one electric and one magnetic field component oscillating in phase, mutually perpendicular to each other and perpendicular to the direction of propagation. The energy of the radiation is quantized into discrete packages, energy quanta, known as photons. A molecule can absorb light only if the energy of the photon, given by the oscillation frequency ν, corresponds exactly to the energy gap, ΔE, between two energy states of the molecule, as stated by the Bohr frequency condition:

$$\Delta E = h\nu = \frac{hc}{\lambda} \qquad \text{(Equation 1)}$$

where h is Planck's constant, c is the speed of light and λ is the wavelength of the radiation. A second prerequisite for absorption to occur, *i.e.* for the transition of an electron in a molecule from its ground state to an excited state, is constructive interference between the electric field of the light and the transient oscillation of the electron as it changes state. If the initial and final states are represented by the wavefunctions Ψ_i and Ψ_f, respectively, this transient oscillation can be described by the transition dipole moment $\vec{\mu}_{fi}$ as

$$\vec{\mu}_{fi} = \int \Psi_f \mu \Psi_i \, d\tau \qquad \text{(Equation 2)}$$

where μ is the electric dipole operator. The probability, P, for a transition to occur upon radiation, and thus for the absorbance, A, is proportional to the

magnitude of the transition dipole moment as well as to its orientation relative to the incident light:

$$A \propto P \propto \left|\vec{\mu}_{fi}\right|^2 cos^2\,\theta \qquad \text{(Equation 3)}$$

where θ is the angle between $\vec{\mu}_{fi}$ and the electric field vector of the electromagnetic radiation.

The absorbance is related to the concentration, *c*, of the absorbing species, according to the Beer-Lambert law:

$$A(\lambda) = log\frac{I_0}{I} = \varepsilon(\lambda) c\, l \qquad \text{(Equation 4)}$$

where I_0 and I are the intensities of the incident and transmitted light, respectively, l is the optical path length and $\varepsilon(\lambda)$ is the molar absorption coefficient at wavelength λ.

3.1.2. Polarized Spectroscopy

Standard absorption spectroscopy makes use of isotropic, or non-polarized, light. However, the use of polarized light can provide information about molecular structure, orientation and symmetry. Circularly polarized light, where the end of the electric field vector describes a circle as the wave of radiation progresses in time, is used in circular dichroism spectroscopy. Correspondingly, in linear dichroism spectroscopy linearly polarized light, with the electric field of the radiation oscillating in one plane only, is used. Linear dichroism, with its special application site-specific linear dichroism (SSLD), has played an important role for the work of this Thesis in studying the structure of recombination proteins.

3.1.2.1. Linear Dichroism

Linear dichroism, LD, is defined as the differential absorption,

$$LD = A_{\parallel} - A_{\perp} \qquad \text{(Equation 5)}$$

between orthogonal forms of plane polarized light, where the polarization vector of the incident light beam is oriented parallel (∥) and perpendicular (⊥) to a macroscopic orientation axis of the sample. To give rise to an LD signal the sample needs to be anisotropic, *i.e.* the molecules in the sample have to be macroscopically oriented. Orientation can be achieved in different ways, *e.g.* by stretching the molecules in a polymer film, by placing the molecules in an electric field or by introducing a flow gradient in the sample. The principles of flow LD, which is the method used in the work of this Thesis, are schematically shown in

Figure 3.1. The sample cell is a Couette flow cell consisting of two coaxial transparent (quartz) cylinders separated by a narrow gap into which the sample solution is loaded. Rotation of one of the cylinders creates a laminar shear flow that aligns the sample molecules in the flow direction.

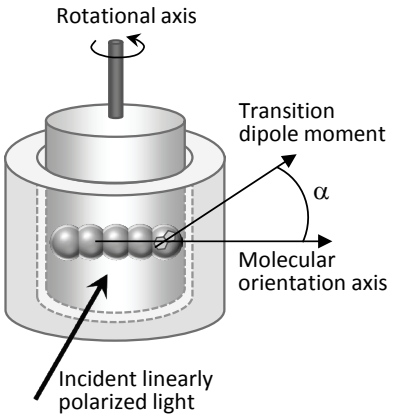

Figure 3.1 An illustration of the principles of flow LD using a Couette flow cell. One of the coaxial cylinders is rotated, creating a shear flow in which the sample molecules are aligned. A chromophore of an aligned molecule is indicated and the angle α (Equation 6) is defined.

For a uniaxially oriented sample, the so-called reduced LD (LD^r) is related to the orientation of the transition dipole moment of the light-absorbing chromophore as

$$LD^r = \frac{LD}{A_{iso}} = \frac{3}{2} S(3\,cos^2\,\alpha - 1) \qquad \text{(Equation 6)}$$

where A_{iso} is the absorption of the corresponding isotropic sample and α is the angle between the transition dipole moment and the molecular orientation axis. S is an orientation parameter indicating the degree of orientation of the sample; $S = 1$ for a perfectly oriented sample and $S = 0$ for a randomly oriented sample. Thus, given that the transition dipole moment direction in the chromophore is known, LD can be used to deduce the orientation of the chromophore relative to the macroscopic orientation axis, *i.e.* the orientation of a DNA-binding ligand relative to the DNA helix axis or the orientation of an aromatic amino acid relative to the fibre axis of a filamentous protein.

The helical filaments formed by recombination proteins like RecA or Rad51 are easily aligned in a shear flow and thus well suited for LD studies. The main intrinsic protein chromophores are the aromatic side chains of the amino acids tyrosine, tryptophan and phenylalanine, with transitions in the near UV region between 200-300 nm. In the far UV region, around and below 200 nm, transitions in the peptide bond give rise to absorption and there are minor contributions from other amino acids [89]. However, a protein LD spectrum over the UV wavelengths will be an average with contributions from all aromatic

residues and all peptide bonds in the protein, impeding the possibility to extract specific structural information. This problem is circumvented by the site-specific linear dichroism methodology.

3.1.2.2. Site-Specific Linear Dichroism

SITE-SPECIFIC LINEAR DICHROISM (SSLD) is a special application of LD where the spectroscopic technique is used in combination with molecular replacement of amino acid chromophores inside a retained protein structure, in order to deduce structural information on these specific residues. The SSLD is determined as the differential LD between wild-type protein and a modified protein in which one residue has been replaced by another, optically "invisible", amino acid (Figure 3.2). If the LD spectra of wild-type and modified proteins are normalized with respect to orientation, the SSLD can be evaluated to determine the angular coordinates of the replaced residue.

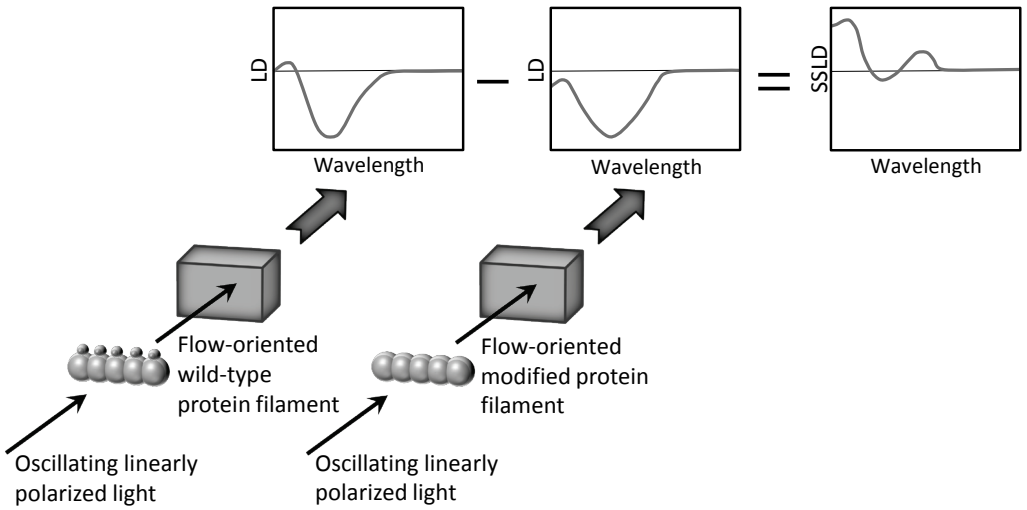

Figure 3.2 Schematic description of the SSLD principle. SSLD is determined as the differential LD of wild-type and modified proteins. In the modified protein a single amino acid is replaced by an optically "invisible" residue and the SSLD corresponds in principle to the LD of the replaced amino acid.

The total LD of the protein will be a function of wavelength, λ, with contributions from all chromophores, *i*, and transitions, *u*:

$$LD(\lambda) = \sum_i \sum_u LD_{iu}(\lambda) \qquad \text{(Equation 7)}$$

Recalling the expression for LD^r (Equation 6), the orientation angle of a specific transition dipole moment in a chromophore, α_{iu}, relative to the orientation axis of the sample, can be related to the total LD as

$$LD(\lambda) = \sum_i \sum_u {A^{iso}}_{iu}(\lambda)\, {LD^r}_{iu}(\lambda) = \frac{3}{2} S \sum_i \sum_u \left[{A^{iso}}_{iu}(\lambda)(3\cos^2 \alpha_{iu} - 1)\right]$$

(Equation 8)

Considering an SSLD experiment including the wild-type and a modified protein with a single amino acid substitution, and noting that (Equation 7) and (Equation 8) are valid only for non-overlapping absorption bands, (Equation 8) can, for the wild-type protein, be expressed as

$$LD_{WT}(\lambda) = \frac{3}{2} S_{WT} \sum_u \left[{A^{iso}}_{1u}(\lambda)(3\cos^2 \alpha_{1u} - 1) + \sum_{i=2}^{N} {A^{iso}}_{iu}(\lambda)(3\cos^2 \alpha_{iu} - 1) \right]$$

(Equation 9)

where chromophore 1 is the SSLD target and *WT* denotes the wild-type protein. If, in the mutant protein, chromophore 1 is replaced by a chromophore with zero absorption at the wavelength (or wavelengths) of interest the corresponding expression for the mutant protein will be

$$LD_{MUT}(\lambda) = \frac{3}{2} S_{MUT} \sum_u \left[0 + \sum_{i=2}^{N} {A^{iso}}_{iu}(\lambda)(3\cos^2 \alpha_{iu} - 1) \right] \qquad \text{(Equation 10)}$$

If the orientation parameters for the wild-type and modified proteins can be assumed to be the same, or if information on these parameters is known so that the LD spectra can be normalized with respect to orientation, the differential between (Equation 9) and (Equation 10), *i.e.* the SSLD, will correspond to the LD spectrum of the replaced chromophore:

$$SSLD(\lambda) = \frac{3}{2} S \sum_u \left[{A^{iso}}_{1u}(\lambda)(3\cos^2 \alpha_{1u} - 1) \right] \qquad \text{(Equation 11)}$$

Thus, the orientation angles of the transitions in the substituted residue, α_{1u}, can be identified, provided that the orientation parameter can be independently determined. Note that two angles are needed to specify the orientation of an object in space (except if it is a linear object, in which case one angle is enough). Thus, two different absorption bands, having non-parallel transition moment directions, need to be studied in order to determine the orientation of, say, a planar chromophore like phenol in tyrosine or indole in tryptophan.

A prerequisite for the SSLD principle is that the molecular replacement of an amino acid residue does not significantly change the internal structure of the protein or its arrangement in the nucleoprotein filament. This is not easily verified but a carefully considered choice of replacing residue together with a verification of retained function of the modified protein can justify such an assumption. Furthermore, SSLD is dependent on the normalization of LD spectra with respect to the degree of macroscopic orientation, represented by the orientation parameter *S*. The orientation may vary between different sample preparations and even small variations in *S* will, if not corrected for, give rise to large variations in the SSLD spectrum. An inserted probe chromophore with separable absorption and known transition dipole moment orientation that can be used as an internal standard is one way to meet with the requirement for

normalization. This method was applied in a previous SSLD study of RecA in complex with single-stranded poly(dεA), where the ethenoadenine chromophore served as an internal standard for the fibre orientation [12]. In the studies presented in this Thesis another approach, applicable if the substitute and substituted residue have identical absorption coefficients at some "magic wavelength", was used.

In the work on RecA in Paper I as well as in the work on HsRad51 in Paper II, the tyrosine residues in the proteins were replaced, one at a time, by phenylalanine. These amino acids are structurally very similar, differing by one hydroxyl group only, and thus this substitution is expected not to significantly affect the protein structure. Furthermore, the replacement of tyrosine by phenylalanine will in principal correspond to the replacement of the tyrosine transition dipole moments by the phenylalanine transition dipole moments. The tyrosine and phenylalanine transitions show coinciding absorption intensities at certain wavelengths (Figure 3.3). At these "magic wavelengths" the LD intensity, after normalization with respect to orientation, of wild-type and modified proteins is expected to be the same. Thus, normalizing the LD spectra to the same LD intensity at such a "magic wavelength" is equivalent to a normalization of orientation parameters.

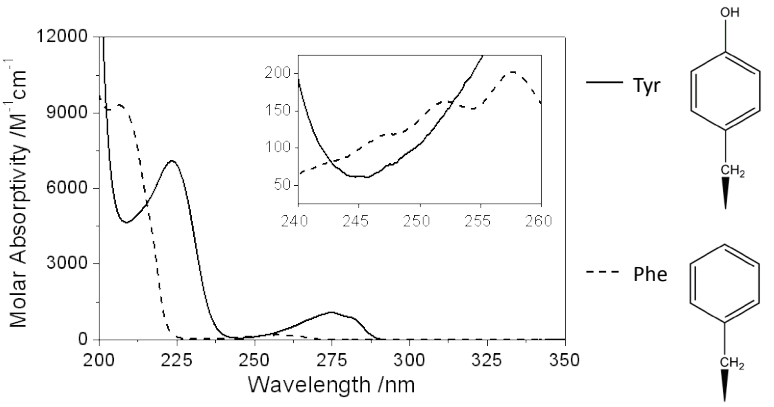

Figure 3.3 Absorption spectra of tyrosine and phenylalanine, showing the "magic wavelengths" where the chromophores have identical absorption coefficients. The molecular structures of the tyrosine and phenylalanine side chains are shown to the right.

3.1.3. FLUORESCENCE SPECTROSCOPY

FOLLOWING AN ABSORPTION event, the excited molecule will eventually return to its ground state by dissipation of the absorbed energy. There are several possible paths for the deactivation process, as illustrated in the Jablonski diagram in Figure 3.4. Upon excitation from the ground state, S_0, to a higher electronic state, S_n, the molecule rapidly relaxes to the lowest vibrational state of S_1 by non-radiative vibrational relaxation and internal conversion. From this state one of

three events can return the molecule to the ground state: non-radiative internal conversion and vibrational relaxation, radiative relaxation by fluorescence emission or spin conversion (intersystem crossing) to the first excited triplet state, T_1. Depopulation of the triplet state can then occur non-radiatively or by phosphorescence emission. Phosphorescence is a much slower process than fluorescence since the intersystem crossing spin-change is a forbidden transition. Due to the initial loss of energy through vibrational relaxation, the emitted light is always lower in energy than the absorbed light. This red-shift of the emission spectrum compared to the absorption spectrum is called the Stoke's shift.

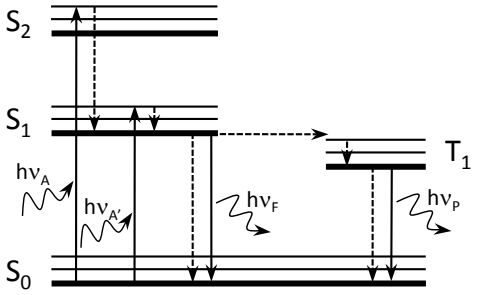

Figure 3.4 A Jablonski diagram illustrating different deactivation paths following electronic excitation by absorption of light. Solid arrows indicate radiative processes and dashed arrows indicate non-radiative processes, see text for details.

3.1.3.1. Förster Resonance Energy Transfer

FÖRSTER (OR FLUORESCENCE) resonance energy transfer (FRET) is a useful application of fluorescence spectroscopy; in the work of this Thesis FRET was used to monitor the strand exchange reaction in the model system studies of Papers III-V. FRET is a non-radiative transfer of the excitation energy from a donor fluorophore to an acceptor chromophore that is possible if part of the donor emission spectrum overlaps with the absorption spectrum of the acceptor. The energy transfer occurs through dipole-dipole interactions between the donor and acceptor and its efficiency, E_{ET}, is strongly dependent on the donor-acceptor distance, *r*, according to

$$E_{ET} = \frac{R_0^6}{R_0^6 + r^6} \qquad \text{(Equation 12)}$$

where R_0 is the characteristic Förster distance for a given donor-acceptor pair, defined as the distance where E_{ET} equals 0.5. Typical Förster distances range from 20 to 90 Å and for FAM and TAMRA, the chromophore pair in the FRET set-up used in the work of this Thesis, R_0 is about 50 Å [90]. Due to its dependency on the distance between the donor and acceptor FRET can be used to monitor distances within or between molecules. Furthermore, molecular interactions can be probed using FRET, as donor emission can serve as an on-off switch triggered by release or binding of a molecule carrying the FRET acceptor.

3.2. PROTEIN TECHNIQUES

IN PRINCIPLE ANY gene can be manipulated in the laboratory by recombinant DNA technology, and its encoded protein can be over-expressed and purified for *in vitro* studies. Useful protocols for such work can be found in the handbook by Sambrook and Russell [91]. For the work presented in Paper II the human Rad51 gene was systematically modified to produce proteins with a specific alteration in amino acid sequence, compared to the wild-type protein. Wild-type and modified proteins used for the work of this Thesis were expressed in bacteria and purified as described in this section.

3.2.1. SITE-DIRECTED MUTAGENESIS

FOR THE SSLD study of HsRad51 the ten tyrosine residues of this protein were replaced, one at a time, by phenylalanine. The amino acid substitution was made by site-directed mutagenesis on KM40, an M13-vector carrying the human Rad51 gene, using the oligonucleotide-directed method described by Kunkel [92]. The basic approach of this method, illustrated in Figure 3.5, is to use a single-stranded DNA template, prepared in a modified *E. coli* strain to contain several uracil residues instead of thymine.

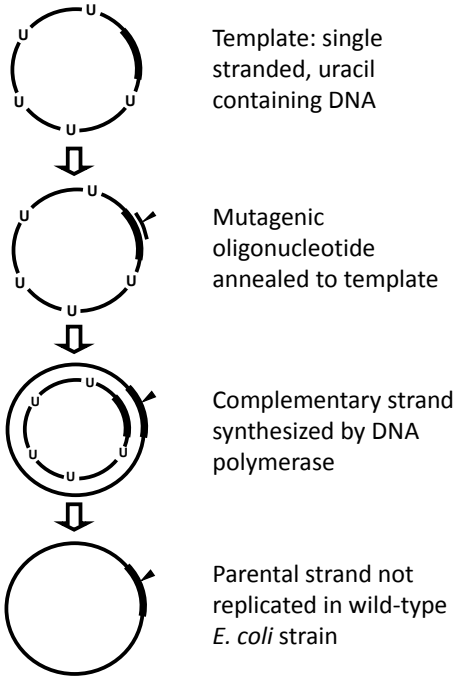

Figure 3.5 The basic principles of the method used for site-directed mutagenesis. Uracil containing DNA is prepared from a modified *E. coli* strain. Following annealing of a mutagenic oligonucleotide and elongation of the complementary strand, the double-stranded plasmid is transformed into a wild-type *E. coli* strain, in which the uracil containing strand is inactivated and only the mutated strand will be replicated.

The mutation is introduced by annealing an oligonucleotide, carrying the specific mutation with flanking complementary sequence, to the single-stranded template DNA, followed by elongation and ligation of the complementary strand. Amplification of the covalently closed circular dsDNA in a wild-type *E. coli* strain produces non-uracil containing DNA with the mutated sequence. By the introduction of a restriction enzyme cleavage site close to the mutation successful mutagenesis was conveniently screened for. The mutations were then verified by DNA sequencing.

3.2.2. Expression of Rad51 Proteins

Wild-type HsRad51 was expressed in *E. coli* BL21(DE3) cells carrying pLysE [93] freshly transformed with the plasmid pT7-HsRad51 (a gift from Dr. Shinohara, Osaka University, Japan), where the gene encoding HsRad51 is inserted under the T7 promoter between the NcoI and BamHI restriction sites in the pET3d vector. Cells were grown in LB broth at 25°C and at $OD_{600} \approx 0.5$ protein expression was induced by addition of IPTG to a final concentration of 50 mg l^{-1}. After about 18 h of incubation cells were harvested and stored at -20°C until purification. Expression of HsRad51 was checked by *SDS-PAGE*.

The mutant human Rad51 proteins were expressed from KM40, an M13 phage vector carrying the modified human Rad51 gene under control of the T7 promoter. *E. coli* XL1 Blue cells [94] carrying plasmid pTI1219 expressing the T7 RNA polymerase were grown in LB broth at 37°C to $OD_{600} \approx 0.5$. Expression of T7 RNA polymerase was induced by addition of IPTG to a final concentration of 50 mg l^{-1} and cells were infected with the modified KM40 phages at a multiplicity of infection of about 5-20 for expression of the Rad51 protein. After 30 min incubation at 25°C without shaking, shaking was turned on and incubation continued at 25°C for about 18 h. Cells were then harvested and stored at -20°C until purification. Expression of HsRad51 was checked by *SDS-PAGE*.

3.2.3. Purification of Rad51 Proteins

Wild-type and modified human Rad51 proteins were purified according to the following protocol. Throughout the purification process all solutions were kept cold on ice. Frozen cells were thawed and suspended in three volumes (*i.e.* 21 ml for 7 g cell mass) of saccharose buffer (25% sucrose, 50 mM Tris-HCl, pH 8.0, 1 mM EDTA, 5 mM β-mercaptoethanol), containing sugar to stabilize the protein. Cells were lysed by addition of lysozyme to a final concentration of 0.6 mg ml^{-1} and after 15 min of incubation on ice the suspension was subjected to sonication using a tip sonicator. Cell debris was centrifuged down (21000×g, 40 min, 4°C) and the supernatant was diluted with saccharose buffer to adjust the absorption at 260 nm to 60. 10% polyethyleneimine, which is positively

charged and thereby precipitates negatively charged DNA and proteins, was added to a final concentration of 0.15%. After incubation with stirring at 4°C for one hour the precipitate was collected by centrifugation (10000×g, 10 min, 4°C) and suspended in ten volumes, compared to the original cell mass, of 1 M NaCl buffer (50 mM Tris-HCl, pH 7.5, 1 M NaCl, 1 mM EDTA, 10% glycerol, 5 mM β-mercaptoethanol). 1 M NaCl inhibits the interaction between protein and polyethyleneimine, leaving the protein in solution rather than in the precipitate. The suspension was incubated with stirring at 4°C for one hour and after centrifugation (10000×g, 10 min, 4°C) the supernatant was collected and the protein was precipitated again by addition of 0.22 mg ml^{-1} solid $(NH_4)_2SO_4$. After incubation with stirring at 4°C for one hour the precipitate was collected by centrifugation (10000×g, 10 min, 4°C) and dissolved in 2.5-3.5 ml of HA-buffer A (2.5 mM K_2HPO_4, 2.5 mM KH_2PO_4, 200 mM NaCl, 10% glycerol, 5 mM β-mercaptoethanol).

The solution was applied to a hydroxyapatite column (CHT® Ceramic Hydroxyapatite, Bio-Rad, volume 15 ml), which has both positively and negatively charged functional groups and separates proteins according to the strength of their electrostatic interactions with the resin. Proteins were eluted with a 5-600 mM phosphate gradient, with the Rad51 protein eluted at 180-400 mM phosphate. The peak fractions of Rad51 protein, verified by *SDS-PAGE*, were collected and precipitated by dialysis against a 70% $(NH_4)_2SO_4$ buffer (70% $(NH_4)_2SO_4$, 50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 5 mM β-mercaptoethanol) at 4°C overnight.

The precipitated protein was collected by centrifugation (10000×g, 10 min, 4°C), dissolved in 3 ml S200-buffer (20 mM Tris-HCl, pH 7.5, 1 M NaCl, 1 mM EDTA, 10% glycerol, 5 mM β-mercaptoethanol) and applied to a Superdex™ 200 column (Amersham, volume 120 ml), which is a gel filtration column separating proteins by size (separation range 10-600 kDa). The Rad51 protein peak fractions, verified by *SDS-PAGE*, were pooled and dialyzed against magnesium buffer (20 mM Tris-HCl pH 7.5, 20 mM $MgCl_2$, 10% glycerol, 5 mM β-mercaptoethanol) at 4°C overnight.

After dialysis the resulting precipitate was collected by centrifugation (10000×g, 10 min, 4°C) and dissolved in 1.75 ml DEAE-buffer A (20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 1 mM EDTA, 10% glycerol, 5 mM β-mercaptoethanol). The solution was applied to an ion-exchange DEAE-5PW column (Tosoh, volume 3.8 ml). The resin is positively charged with a pore size of 0.1 µm and separation is accomplished by an ionic strength gradient. The Rad51 protein can enter the column pores and be eluted with a 200-1000 mM NaCl gradient. The Rad51 peak fractions, eluted at 300-400 mM NaCl, were verified by *SDS-PAGE*, collected and dialyzed against DEAE-buffer A at 4°C over night. Finally the protein was dialyzed against stock buffer (50 mM Tris-HCl, pH 7.5, 200 mM NaCl, 1 mM EDTA,

50% glycerol, 1 mM dithiothreitol) at 4°C overnight and thereafter stored at -80°C.

The concentration of the Rad51 proteins was determined by *DC* Protein Assay (Bio-Rad), which is a colorimetric assay based on the reaction of protein with an alkaline copper tartrate solution and Folin reagent. Bovine serum albumin was used as a standard. As a control, *SDS-PAGE* was run to verify the relative concentrations of all proteins.

3.3. LIPID MEMBRANES

LIPIDS ARE AMPHIPHILIC molecules consisting of a hydrophilic head group and one or two hydrophobic aliphatic tails. They self-assemble in aqueous solution forming aggregates of different shape and structure such as the micelle, lipid bilayer and liposome depicted in Figure 3.6. The morphology of the aggregates is determined, apart from external factors like concentration and temperature, by the molecular geometry of the lipids, defined by the critical packing parameter, *CPP*, or surfactant number, N_s, that relates the head group area, a_0, to the length, l, and volume, v, of the hydrophobic part of the amphiphile:

$$N_s = \frac{v}{l\, a_0} \qquad \text{(Equation 13)}$$

From their surfactant number, lipids can be classified as being cone-shaped ($N_s > 1$), cylindrical ($N_s \approx 1$) or inverted cone-shaped ($N_s < 1$) with cylindrical lipids being most prone to form lamellar phases such as bilayers [95]. Liposomes are spherical assemblies of bilayers formed in solution enclosing an aqueous volume. The nomenclature of such lipid vesicles relates to their size and lamellarity (the number of bilayers they are composed of); the large unilamellar vesicles (LUVs) used in the work of this Thesis contain one bilayer and have a diameter of approximately 100 nm.

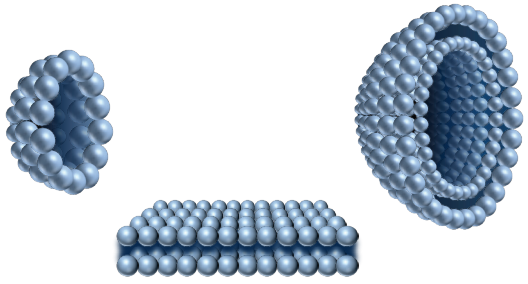

Figure 3.6 Schematic illustration of lipid aggregates. In a micelle, section shown to the left, lipids assemble in a spherical structure with the hydrophobic tails directed towards the centre of the sphere to avoid the aqueous surroundings. A bilayer, centre, consists of two sheets of lipids, packed together to shield the hydrophobic tails from water. To the right a section of a unilamellar liposome, a spherical assembly of a bilayer enclosing an aqueous volume, is shown. The lipid tails are omitted to increase clarity.

Biological membranes, such as the bilayer membranes surrounding the cytoplasm and the cellular compartments, have phospholipids as their main constituent. Taking into account these and other natural lipids, as well as synthetically modified lipids, there is an enormous diversity in structure characteristics and chemical properties of lipids available. Thus, the surface displayed by liposomes can be varied by altering lipid headgroup properties, such as charge and size. Cationic lipids have been widely used in vesicles for gene delivery, since they form stable complexes with negatively charged DNA [96-98]. To avoid aggregation, lipids with PEG-modified headgroups that provide steric hindrance between liposomes can be introduced [99-101].

In the model system studies of Papers III and IV cationic liposomes, which attract DNA to their surface, were shown to have an enhancing effect on the strand exchange reaction. The lipids used for that work were the zwitterionic DOPC, the cationic DOTAP and the PEG-modified lipid DSPE-MPEG, depicted in Figure 3.7. Among other lipids of interest for possible liposome surface functionalisation, although not used here, are nucleolipids, with potential to interact specifically with DNA [102-104].

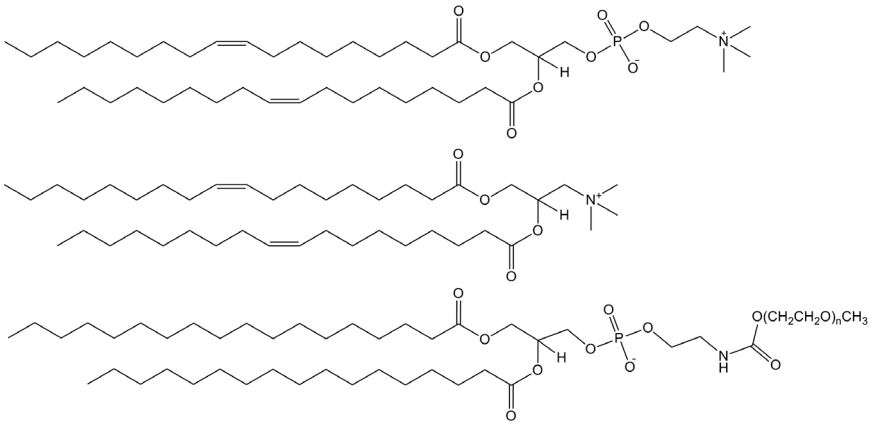

Figure 3.7 Molecular structures of the lipids used in the work of this Thesis. Top: 1,2-Dioleoyl-sn-glycero-3-phosphatidylcholine (DOPC), centre: N-[1-(2,3-Dioleoyloxy)propyl]-N,N,N-trimethylammonium chloride (DOTAP), bottom: 1,2-Distearoyl-sn-glycero-3-phosphoethanolamine-N-[poly(ethyleneglycol)2000] (DSPE-MPEG).

3.3.1. Preparation of Lipid Vesicles

Large unilamellar vesicles (LUVs) were prepared by the thin-film hydration method followed by freeze-thawing and extrusion [105]. Lipids dissolved in chloroform were mixed at the desired molar ratios in a round bottom flask and solvent was then evaporated under reduced pressure using a rotary evaporator to produce a thin lipid film on the flask walls. The film was kept under vacuum for at least two hours or overnight to remove any remaining traces of solvent before hydration in aqueous buffer. Upon hydration and under vigorous shaking by vortexing, the lipids “swell” and sheets of bilayers detach and self-close to

form onion-like structures of large multilamellar vesicles (LMVs). The vesicles were disrupted by freeze-thawing in five cycles (freezing in liquid nitrogen and thawing at 45°C in a heat block, with the final thawing at room temperature) prior to down-sizing by extrusion. The lipid solution was extruded 21 times through polycarbonate filters with a pore size of 100 nm, using a hand-held syringe extruder. This method produced LUVs with a relatively narrow size distribution and a mean diameter of approximately 110 nm, verified by dynamic light scattering.

3.4. Molecular Crowding

A general and indeed significant difference between reactions *in vitro* and the corresponding process *in vivo* is the fact that the intracellular environment is crowded with soluble and insoluble macromolecules such as proteins, nucleic acids, sugars and lipids. Up to 40% of the cellular volume can be occupied by macromolecules, with a total concentration that may be orders of magnitude higher than the dilute solutions normally used in biochemical studies *in vitro* [106]. This "molecular crowding" affects structure, function and stability of biomolecules like proteins and nucleic acids, as well as reaction rates and equilibria of biochemical processes, and there is an emerging interest in these effects [107-110].

Molecular crowding is more accurately termed an excluded volume effect and the extent of the volume exclusion is size dependent (Figure 3.8). Considering the high intracellular concentration of macromolecules, there is an un-negligible restriction in the volume accessible to other molecules, especially large biopolymers. As a result of this there will be an increased effective concentration of the solutes in the cell.

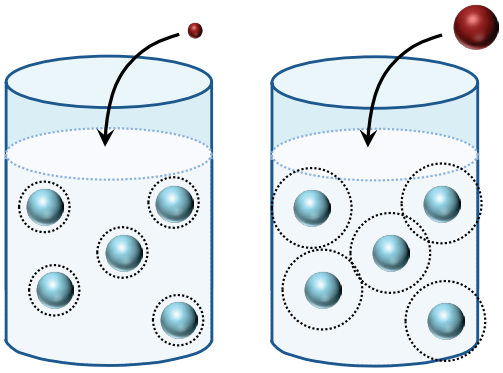

Figure 3.8 Illustration of the excluded volume effect and how it depends on molecular size. Spherical macromolecules in solution occupy a certain volume that other molecules are excluded from. The volume accessible to an added molecule depends on its size; the restriction in accessible volume is much larger for a large molecule than a small. Dotted circles indicate the sum of radii of macromolecule and added molecule, thus the closest distance between molecular centres.

Addition of high concentrations of macromolecules to the sample solution is a way to mimic the crowded intracellular conditions *in vitro*. The crowding agent should ideally be water soluble, inert towards the system under study and not prone to self aggregate. PEG, dextran, polyvinyl alcohol and Ficoll (a copolymer of sucrose and epichlorohydrin) are examples of polymers commonly used as crowding agents. Proteins such as albumin, haemoglobin and lysozyme are also used [108-110].

PEG as a crowding agent has been demonstrated to stabilize or destabilize dsDNA, depending on duplex length as well as PEG size and concentration [111-113]. Suggested factors governed by PEG that affect DNA stability are, beside purely sterical effects, decreased water activity, an altered behaviour of DNA counterions, a disruption of base stacking, or a combination of multiple effects. As a cautionary note on using PEG as a crowding agent it has been suggested that PEG not only causes an excluded volume effect but also interacts attractively with non-polar or hydrophobic side chains of proteins [110]. Similar interactions could possibly occur also between PEG and the hydrophobic bases of DNA, especially at high PEG concentration. In aqueous two-phase systems of PEG and dextran, or hydroxypropyl starch, nucleic acids, especially of high molecular weight, avoid the PEG-rich phase, although the partition is strongly influenced by the ionic strength and electrolyte composition of the system [114-115]. At high concentrations of NaCl the nucleic acids favour the PEG-rich phase.

As a parallel to the molecular crowding present in the intracellular matrix, DNA molecules that are subject to strand exchange inside the recombinase filament will also experience a crowded environment. To mimic this effect in a model system study PEG was used as a catalysing agent in the artificial strand exchange set-up presented in Paper V.

4. RESULTS

In this chapter the work of Papers I-V is summarized and the main results presented therein are discussed. Some preliminary, unpublished results of relevance will also be presented.

4.1. RECOMBINATION PROTEINS

STRUCTURAL INVESTIGATIONS OF proteins and their complexes with different ligands are an important source to provide mechanistic information. To obtain high resolution structures, needed for detailed mechanistic conclusions, techniques like X-ray crystallography and NMR spectroscopy are traditionally applied. However, not all proteins, or protein complexes, are amenable to these methods. Specifically, the filamentous complexes of the RecA family recombinases and DNA are outside the size limitations of NMR and have, with one recent exception [16], proven difficult to crystallize with DNA intact. In this Thesis an alternative approach to the investigation of these structures is presented.

The SSLD technique was first introduced almost twenty years ago, when a single amino acid substituted RecA protein was analysed by LD and orientation parameters for the substituted residue was obtained [17]. A decade later a more complete SSLD study of RecA in its nucleoprotein complex with ssDNA was presented [12]. The work of this Thesis further illustrates the applicability of the method, with a continuation of the previous work on RecA presented in Paper I and a study of the human Rad51 protein in Paper II that shows how experimental data from SSLD can be successfully combined with theoretical molecular modelling.

4.1.1. RECA

FOR THE DNA strand exchange reaction to occur the RecA protein initially polymerizes onto a single-stranded DNA, forming the so called pre-synaptic filament. This nucleoprotein complex was the target of the previously presented SSLD study [12]. The subsequent binding of a double-stranded DNA to the pre-synaptic filament results in the formation of the synaptic filament where search for homology takes place and, upon homologous pairing, strand exchange occurs. The resulting post-synaptic RecA filament hosts the recombined dsDNA. Following the successful use of SSLD to gain structural information on the pre-synaptic filament, there was an interest in studying the nucleoprotein filament of RecA and dsDNA using the same technique. This is the work presented in Paper I.

Wild-type and modified *E. coli* RecA proteins, where each of the seven tyrosine residues was replaced by phenylalanine, in complex with double-stranded calf-thymus DNA and the non-hydrolysable ATP analogue ATPγS were subjected to flow LD measurements. The LD spectra were normalized at a "magic wavelength" of 250-255 nm, as described in Section 3.1.2.2., and the differential LD of wild-type and modified protein complexes, *i.e.* the SSLD of the substituted amino acids, was determined.

The SSLD spectrum of each tyrosine residue, computed by subtracting the normalized LD spectrum of each modified nucleoprotein complex from that of the wild-type nucleoprotein complex, is shown in Figure 4.1. The peaks are centred at 230 nm and 280 nm, as expected for the L_a and L_b transitions[a] of tyrosine, respectively [116]. According to (Equation 11) the orientation angles of these transitions, relative to the filament helix axis, can be determined from the SSLD, provided that the orientation parameter, S, is known. An approximate absolute value of S was estimated from the normalized LD intensity at the "magic wavelength", under the assumption that the effective orientation angle of the DNA bases in the RecA filament is 80 ± 10°, as has been determined by small-angle neutron scattering measurements [13].

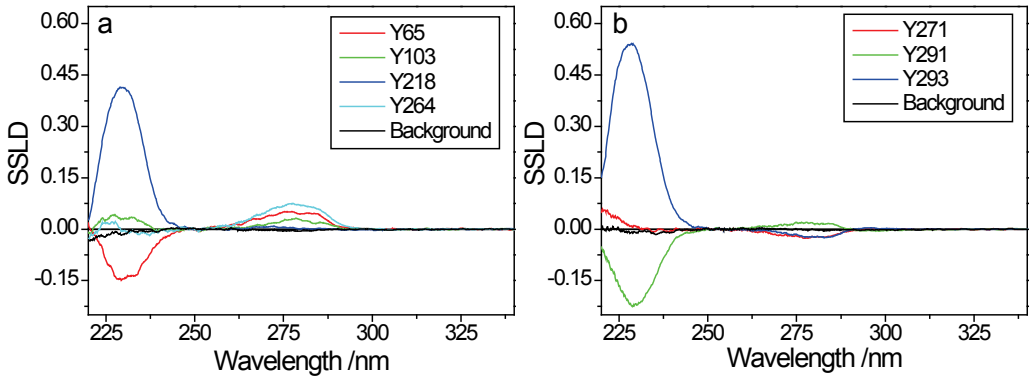

Figure 4.1 SSLD spectra of the tyrosine residues of RecA computed by subtraction: (LD of wild-type RecA complex) - (LD of modified RecA complex). Background indicates the differential spectrum of the two wild-type spectra that were recorded first and last in a series of measurements.

The orientations of the tyrosine transition moments relative to the helix axis in RecA in complex with single- [12] and double-stranded DNA calculated from SSLD data are shown in Table 4.1. The angles determined for the different RecA-DNA complexes are in general similar. However, significant differences are found for Tyr65 and Tyr264 and some minor changes in orientation are observed for Tyr103 and Tyr271. The general similarity indicates that the overall orientation

[a] In Paper I the transitions are denoted B and L_b, respectively. However, according to the Platt nomenclature L_a and L_b are the correct terms.

and internal structure of RecA in the nucleoprotein filament is not significantly altered from the initial to the final state of the strand exchange reaction. Local conformational changes around Tyr65 and Tyr264, as the bound DNA changes from single-stranded to double-stranded are not unlikely, since these residues have been suggested to be involved in DNA binding, as has Tyr103 [66-67]. Tyr271 has so far not been assigned any particular function in RecA, the small change in orientation observed for this residue might be related to the conformational change of Tyr264.

Table 4.1 Orientation angles of tyrosine transition moments relative to the helix axis of the RecA filaments. Angles in filament with ssDNA are taken from [12], angles in crystal structure are taken from [15].

transition moment	wavelength, nm	angle α, ° dsDNA	angle α, ° ssDNA	angle α, ° in crystal
ssDNA	320-350	--	70 (80-60)[a]	--
dsDNA	250-255	80 (90-70)[a]	--	--
Tyr65-L_a	227	67 (68-63)	81 (90[b]-62)	73
Tyr65-L_b	278	32 (30-39)	47 (44-52)	40
Tyr103-L_a	227	51 (51-52)	55 (55-55)	81
Tyr103-L_b	278	41 (40-45)	46 (42-51)	32
Tyr218-L_a	227	19 (13-31)	32 (21-46)	20
Tyr218-L_b	278	53 (53-53)	58 (60-56)	80
Tyr264-L_a	227	53 (53-53)	63 (67-58)	82
Tyr264-L_b	278	18 (12-31)	37 (29-48)	12
Tyr271-L_a	227	53 (53-54)	59 (61-56)	71
Tyr271-L_b	278	68 (70-64)	77 (90[b]-62)	70
Tyr291-L_a	227	79 (86-70)	77 (90[b]-62)	78
Tyr291-L_b	278	46 (45-48)	50 (48-53)	19
Tyr293-L_a	227	0[b] (0[b]-19)	10 (0[b]-40)	26
Tyr293-L_b	278	66 (67-62)	66 (72-59)	65

[a] Angles of transition moments determined assuming ssDNA and dsDNA orientations in the filaments are 70° ± 10° and 80° ± 10°, respectively. [b] LD^r/S value out of range.

Differences in orientation angles of tyrosine transition moments between the active nucleoprotein filaments studied by SSLD and the crystal structure of the inactive RecA-ADP complex (Table 4.1) can, for residues located in the core of the protein or known not to be involved in DNA binding, be considered as a result of a change in orientation of the whole RecA unit in the filament. Thus a concerted degree of rotation of all residues of a protein monomer, relative to their orientation in the crystal structure can be determined and a structural model of the RecA-DNA filament constructed. Using the SSLD data for Tyr218, Tyr291 and Tyr293 in the RecA-DNA complex with dsDNA a rotation of (36°, 349°), *i.e.* 36° rotation around an axis orthogonal to the long axis of the filament followed by 349° rotation around the long axis, was calculated. This is within experimental

uncertainty identical to the (36°, 351°) rotation previously determined for the complex with ssDNA [12]. In that study, SSLD data for the same tyrosine residues and two tryptophans were used for the calculation. However, omitting the transition moments of the tryptophans and recalculating the rotation degree results in a similar value.

The results presented in Paper I, suggesting that the overall orientation and internal structure of RecA is the same for the initial and final state of the strand exchange reaction, could indicate a static role for RecA in the recombination process. It may be that the protein merely acts as a scaffold inducing the proper DNA geometry and providing a proficient environment for strand exchange to take place. Similar conclusions are drawn by Chen *et al.*, who report structural similarity between RecA complexes with single- and double-stranded DNA from crystallographic data [16].

The structures of the RecA-ssDNA and RecA-dsDNA complexes reported by Chen *et al.* differ mainly by a slight deviation of protein monomer body orientation relative to the filament axis. The orientation angles of the tyrosine transition moments averaged over all monomers in the crystallographic structures appear very similar, with deviations less than five degrees between complexes with single- and double-stranded DNA. On the other hand, the orientation angles as determined from SSLD data indicate noticeable deviations between structures of complexes formed with ssDNA and dsDNA (Table 4.1). It is too early to draw any conclusions from this discrepancy between the crystal and solution structures but it is an exciting thought that dynamics in the nucleoprotein fiber, appearing in aqueous solution, might be lost in the more frozen crystal structure.

4.1.2. RAD51

THE RAD51 PROTEIN, the human homologue of RecA, was in focus of the SSLD study presented in Paper II. This was, in many ways, a challenging project. Firstly, the in-house experience of working with the Rad51 protein was not as extensive as that concerning RecA. While the SSLD study of RecA presented in Paper I could take advantage of the preceding work done on this protein, the study of Rad51 required preparatory work like site-directed mutagenesis for single amino acid substitutions, protein expression and purification to be set up from scratch. Secondly, the construction of a structural model of the active RecA nucleoprotein filament was facilitated by a straightforward rotation of the known crystal structure of the protein in its inactive state according to the SSLD data. For HsRad51, however, no high resolution structure of the complete protein is available and thus a more sophisticated modelling procedure was needed.

The human Rad51 protein contains ten tyrosine residues that were chosen as targets for the SSLD study. Site-directed mutagenesis was performed to replace each of the tyrosine residues by phenylalanine. The expression of modified proteins resulted in varying levels of soluble protein, in general lower than for the wild-type protein. For two of the mutants, Y191F and Y301F, the total yield of purified protein was very low and they were therefore excluded from the study. Flow LD spectra of wild-type and modified human Rad51 protein in complex with double-stranded calf-thymus DNA in presence of ATP, representing the post-synaptic filament, were measured. The pre-synaptic complex with ssDNA would also have been of interest to study, but the LD signal obtained from that complex was too weak to allow for a quantitative analysis. The spectra were normalized according to the "magic wavelength" approach described in Section 3.1.2.2. and the SSLD spectra of the tyrosine residues were computed as the spectral difference between wild-type and modified protein complexes. As seen in Figure 4.2 the SSLD spectra show peaks characteristic for the tyrosine transition moments, centred at 230 nm and 280 nm, respectively.

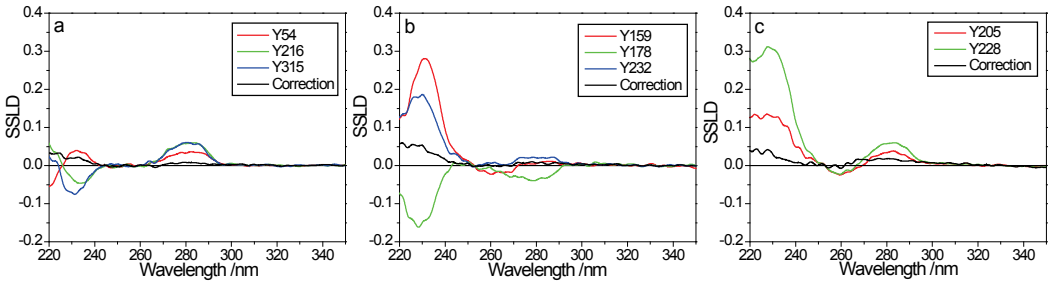

Figure 4.2 SSLD spectra of eight tyrosine residues of HsRad51 computed by subtraction: (LD of wild-type HsRad51 complex)-(LD of modified HsRad51 complex). Correction indicates the differential spectrum of the two wild-type spectra that were recorded first and last in a series of measurements.

The determination of angular orientations from SSLD data requires an estimation of the orientation parameter *S*. Assuming that the orientation of DNA in complex with Rad51 is similar to that in the RecA-DNA complex, which has been assessed from small-angle neutron scattering measurements to 80°±10° for dsDNA [13], an approximate absolute value of *S* could be determined. Doing so, the α angles for the transition moments of the substituted tyrosine residues, relative to the helix axis of the nucleoprotein filament, were calculated from the SSLD data. As shown in the SSLD studies of RecA, earlier by Morimatsu *et al.* [12] and in Paper I of this Thesis, the angular data can be used as constraints when constructing a structural model of the active nucleoprotein filament from a known structure of the protein in its inactive state.

The available high resolution structures of HsRad51 include a crystal structure of the central domain as a fusion complex with a Rad51-binding motif

of the BRCA2 protein [73] and an NMR structure of the N-terminal domain [74]. However, these two structures do not make up a complete protein monomer; important pieces missing are the two putative DNA binding loops and the polymerization motif. By homology modelling, using ScRad51 as a template structure [75], the domain structures were merged together and the structural gaps not covered by the domains were filled. This monomer model could then, after a molecular dynamics structure optimization, be used to construct the protein filament.

To find an axis around which the filament could be assembled a trial axis was rotated to minimize a functional defined as the square root of the squared differences between tyrosine orientation angles calculated from the model and measured by SSLD. The trial axis that best fitted the experimental data was set as the filament axis. The monomer model structures were assembled to a filament with 6.39 monomers per helical turn and a pitch of 99 Å, as suggested from electron microscopy data [9, 78]. Internal monomer orientations were adjusted to be consistent with results from tryptophan-scanning mutagenesis of strategically positioned amino acid residues [79-80, 84, 117]. Moreover, the two loops suggested to be involved in DNA binding [79, 84], which are not visible in the crystal structures of either the central domain of HsRad51 or the ScRad51 filament, were subject to careful modelling. Finally, a molecular dynamics simulation of a filament fragment of three monomers was performed to include effects from subunit-subunit interactions. The refined structure of the central monomer was taken as the final monomeric structure.

The resulting model structure of the filament is shown in Figure 4.3, which also illustrates the angular orientation of the tyrosine transition moments for one of the tyrosine residues. An alignment, with good agreement, to a density map of a three-dimensional reconstruction of an electron micrograph of a HsRad51-DNA filament [78] verified the quality of the modelled filament. As the orientation angles of the tyrosine transition moments determined by SSLD were used as initial input when constructing the filament and minor adjustments of the monomer orientations in the final filament were made to refine the angular fit relative to the experimental data, the angles calculated from the model and from SSLD data are, not surprisingly, in good agreement. However, significant discrepancies are observed for two residues, Tyr178 and Tyr216. These residues are both located on the surface of the protein filament, hence exposed to solvent water and could be regarded as flexible, adopting conformations that might vary between different monomers in the filament, which could explain the angular variations between the model structure and experimental data. The orientation of a particular residue measured by the SSLD technique will be an average of all orientations of that residue in all monomers along the filament, whereas the computer modelling yields one distinct conformation for each residue.

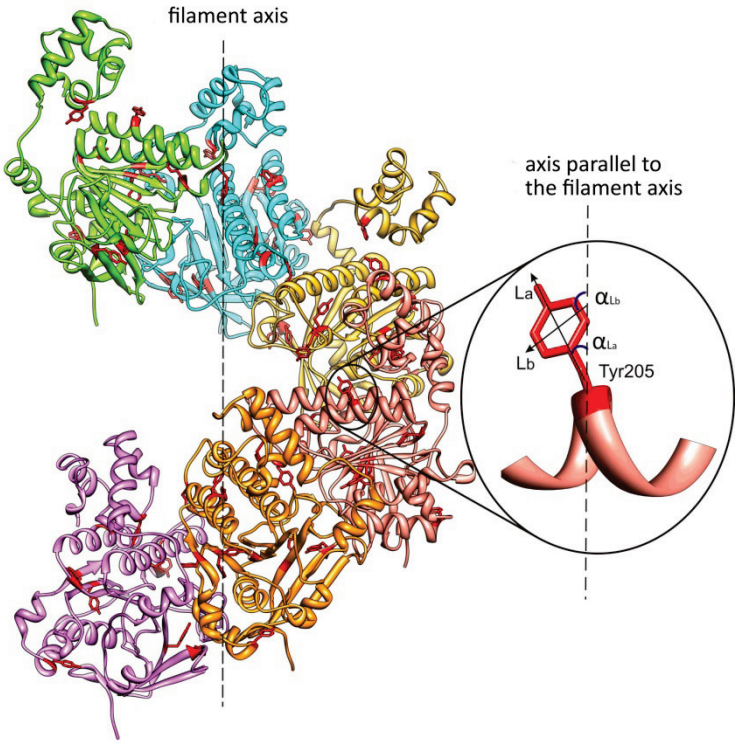

Figure 4.3 Model structure of the human Rad51 helical filament constructed from a monomer structure based on known domain structures of HsRad51 and homology modelling with ScRad51. The filament was assembled to have 6.39 monomers per turn and a helical pitch of 99 Å. The angular orientations of the tyrosine transition moments L_a and L_b relative to the filament axis are illustrated for Tyr205.

The filament model highlights some interesting structural features and opens for mechanistic speculations. At the monomer-monomer interface Phe126 and Phe129 of one subunit and His294 of a neighbouring monomer, residues suggested to be close to the ATP binding site [84], are inserted into a pocket. This pocket is well suited to snugly fit the nucleotide cofactor, although the cofactor was not explicitly included in the current modelling. A similar positioning of the ATP analogue, in a pocket between adjacent monomers, was suggested in the recently presented structure of the RecA-DNA complex filaments [16]. Locating the ATP binding site in this position, in proximity to the L2 loop that is one of the proposed DNA binding sites, raises the speculation that hydrolysis of ATP to ADP after completion of the strand exchange reaction could promote filament disassembly by destabilizations of both DNA-monomer and monomer-monomer interfaces.

The proposed DNA binding loops, L1 and L2, are in the model structure facing the central cavity of the helical protein filament, which has the right dimensions to fit an extended dsDNA molecule. The L2 loop has a compact conformation with some ordered α-helical structure and is in a position that enables direct contact with DNA. Two residues of the L1 loop, Tyr232 and Arg235, that have been shown to be of importance for DNA binding and strand

exchange activity [79-80], are in the model strategically positioned for interaction with DNA (Figure 4.4). Tyr232 is intercalating between two DNA bases of one strand and Arg235 is located close to the sugar-phosphate backbone of the other strand. It is speculated in Paper II that the L1 loop, by the aid of these residues, may serve as a positioning pin for the monomer during filament assembly.

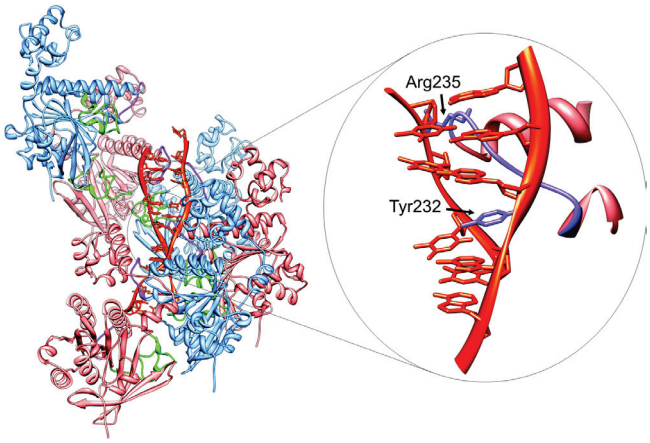

Figure 4.4 Loop orientation in the structural model of HsRad51. Both L1 (magenta) and L2 (green) loops are facing the interior of the filament, which allows the housing of a dsDNA. The close-up illustrates how Tyr232 and Arg235 of loop L1 interact with DNA.

4.1.2.1. Structural Effects Induced by Nucleotide Cofactor

THE RAD51 NUCLEOPROTEIN complexes studied in Paper II were formed with dsDNA in presence of ATP and Mg^{2+}. Based on comparison of LD spectra of the wild-type protein complex measured first and last in each series of measurements, which showed only minor structural differences, ATP hydrolysis was considered to be sufficiently slow to be negligible on the time-scale of the LD experiments. However, it has been shown that HsRad51 has significant ATPase activity in presence of Mg^{2+}. ATP in an HsRad51-ssDNA complex was to a large extent converted into ADP in less than one hour, whereas Ca^{2+} was found to inhibit ATP hydrolysis [48]. To investigate whether the structure of HsRad51 in Paper II is that of a nucleoprotein filament in the active ATP-bound state or if the complex by rapid ATP hydrolysis has been converted to the inactive ADP-bound form, LD spectra of HsRad51-DNA complexes in presence of ATP or ADP have been examined. In a preliminary, unpublished study nucleoprotein complexes with both double- and single-stranded DNA were investigated under three different sample conditions; with ATP and Mg^{2+}, with ATP and Ca^{2+} and with ADP and Ca^{2+}. Complexes formed in presence of ATP and Mg^{2+} correspond, regarding sample conditions, to the filament studied in Paper II. Exchanging Mg^{2+} for Ca^{2+} should minimize ATP hydrolysis and thus the transformation of the complexes to the inactive ADP-bound state. Finally, complexes formed in presence of ADP and Ca^{2+} represent the inactive form of the nucleoprotein filaments.

Figure 4.5 shows flow LD spectra of HsRad51 in complex with double-stranded calf-thymus DNA or single-stranded poly(dT) in presence of ATP or ADP and Mg^{2+} or Ca^{2+}, as described above. From the raw data (Figure 4.5 a and c) it can be concluded that the filaments formed with dsDNA attain a higher degree of orientation compared to those formed with ssDNA. A higher degree of orientation could originate from the nucleoprotein filaments with dsDNA being stiffer and more rigid, thereby aligning better in the flow, than those with ssDNA. The low LD signal from the complexes with ssDNA could also be due to a less efficient filament formation or weaker binding to ssDNA than to dsDNA in this *in vitro* set-up. The LD signal from the filament formed on dsDNA is significantly decreased when the cofactor is changed from ATP to ADP, indicating a change in filament stiffness and/or DNA binding.

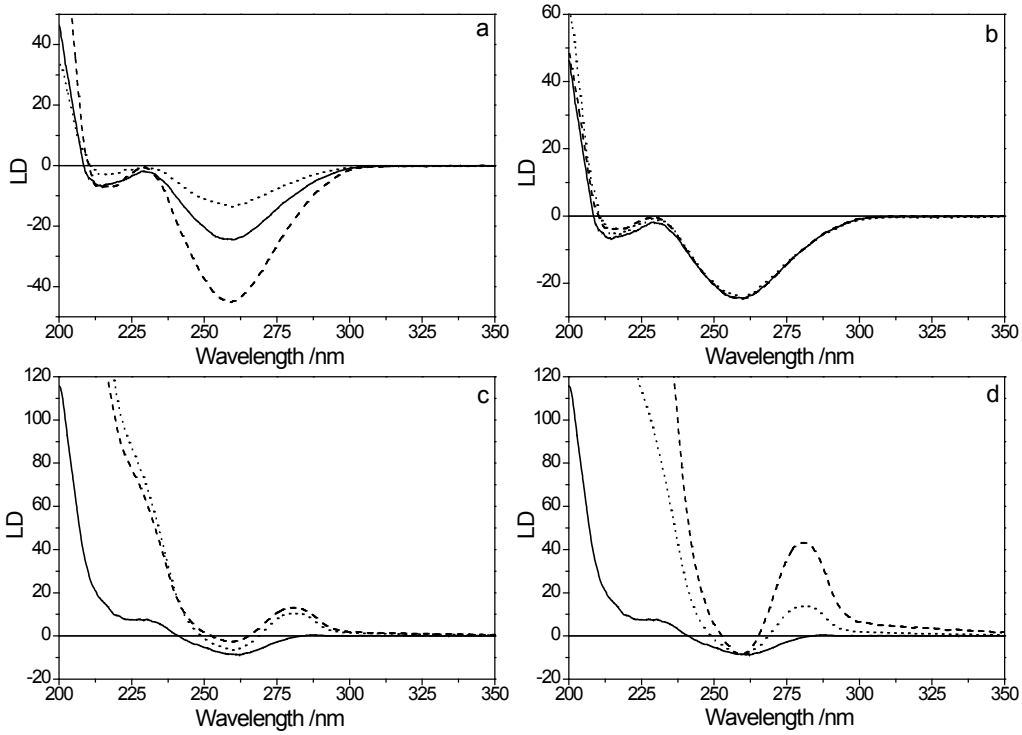

Figure 4.5 Flow LD spectra of HsRad51 nucleoprotein filaments with double-stranded (**a** and **b**) or single-stranded (c and d) DNA under different conditions, varying the nucleotide cofactor and divalent cation: ATP+Ca^{2+} (solid line), ATP+Mg^{2+} (dashed line) or ADP+Ca^{2+} (dotted line). The spectra in **b** and **d** are scaled to the same intensity at 260 nm, taking the ATP+Ca^{2+} conditions as reference. All samples (4 µM HsRad51, 12 µM poly(dT) or calf-thymus DNA, 300 µM ATP or ADP in buffer, pH 6.9, containing 20mM Tris-HCl, 40 mM NaCl, 30% glycerol, 1.2 mM $MgCl_2$ or $CaCl_2$, 0.2 mM EDTA, 0.2 mM DTT) were incubated at room temperature for 2 h before measurement.

A comparison of the LD spectra of the nucleoprotein complexes formed with double- and single-stranded DNA does not only reveal differences in the degree of orientation, there are also considerable dissimilarities in the spectral shapes observed in response to the different cofactor/cation conditions. Whereas a change in nucleotide cofactor or replacing Mg^{2+} by Ca^{2+} seems to affect only the degree of orientation for the complex with dsDNA, the nucleoproteins with ssDNA seem to adopt two significantly different filament structures depending on the sample conditions. Scaling the LD spectra to the same intensity at 260 nm (taking the ATP-Ca^{2+} condition as reference) clearly illustrates these observations (Figure 4.4 b and d). While the spectra of the HsRad51-dsDNA complexes are nearly identical there are notable differences in the spectra of the complexes formed with ssDNA under different conditions. The HsRad51-ssDNA filament formed in presence of ATP and Ca^{2+} exhibits an LD spectrum of similar shape as the spectra of complexes with dsDNA, although with a stronger positive signal at shorter wavelengths. The spectra of nucleoprotein complexes with ssDNA formed in presence of ADP or ATP and Mg^{2+} differ from that formed in presence of ATP and Ca^{2+} but are mutually similar in shape with a lower signal at 260 nm and more pronounced positive signals at shorter wavelengths and around 280 nm.

Although preliminary, these data support the claim that the model structure presented in Paper II is that of an active Rad51 nucleoprotein filament. The structure of the filament formed by HsRad51 on dsDNA seems, from these data, to be very similar regardless if the nucleotide cofactor is ATP or ADP. With ssDNA, however, the filaments seem to be more structurally sensitive, with the LD spectrum of the active HsRad51-ssDNA filament formed in presence of ATP and Ca^{2+} distinct from that of the inactive complex formed in presence of ADP and Ca^{2+}. The similar structural shapes observed for the HsRad51-ssDNA complexes formed in presence of ADP and ATP and Mg^{2+} are consistent with the report of rapid conversion of the active filament into the inactive state in presence of Mg^{2+} due to ATP hydrolysis [48]. The LD spectra of the HsRad51-dsDNA complexes are comparable to that of the active nucleoprotein filament with ssDNA rather than to that of the inactive filament, in agreement with results presented for RecA in Paper I and by Chen *et al.* [16] that suggest overall similar structures of the active filaments formed on single- and double-stranded DNA.

4.2. Model System Studies

Structural studies of recombinases, the scaffold for DNA strand exchange *in vivo*, serve as a tool for gaining insight into the reaction mechanism, as discussed above. A complementary approach is to increase the fundamental knowledge about the behaviour of DNA in the reaction and identify factors of

importance for dissociation and hybridization of DNA strands. A model system where strand exchange of oligonucleotides can be investigated under controlled conditions with the possibility of systematic variation of individual parameters could provide such knowledge. In this Thesis two different model systems for studies of DNA strand exchange are presented. In both systems a catalysing agent is present with the aim of enhancing the strand exchange reaction, although with distinctively separate modes of action.

Two steps of importance for an accelerated strand exchange can be identified: accumulation of DNA and destabilization of the double helix in order to allow for strand invasion. In the model system described in Papers III and IV cationic liposomes serve as catalysing agent. By electrostatic interactions DNA is efficiently accumulated onto the surface of lipid vesicles. Electrostatic interactions are also expected to affect the stability of the duplex. In Paper V the catalytic approach to enhance strand exchange is to use PEG to induce a crowded environment, thus increasing the effective DNA concentration, and provide the possibility of hydrophobic interactions to increase duplex breathing.

4.2.1. CATIONIC CATALYSIS ON LIPOSOMES

IN PAPER III positively charged liposomes were introduced as a platform for model studies of DNA strand exchange. The inspiration to this approach was the finding that a cationic comb-type copolymer accelerates DNA strand exchange *in vitro* [19-20] and the fact that cationic lipid vesicles form stable complexes with DNA, so called lipoplexes that have been employed as non-viral carriers for gene delivery [96-98].

100 nm-sized unilamellar lipid vesicles containing the positively charged lipid DOTAP were used in the study and DNA strand exchange of 20-mer oligonucleotides was monitored using FRET. In a standard experiment dsDNA labelled with FAM on the 5'-end of one strand and TAMRA on the 3'-end of the other was mixed with liposomes. Addition of a five-fold excess of unlabelled ssDNA, complementary to the FAM-labelled strand, resulted in increased FAM emission, indicating separation of the FRET pair due to strand exchange. In presence of cationic liposomes with a surface charge of 25-50% a significant increase in strand exchange rate, compared to the same reaction in bulk without liposomes, was observed. The highest and most reproducible rates were obtained for liposomes with 35% positive charge. Typical kinetic traces at different lipid:DNA ratios are shown in Figure 4.6. The data could be fitted to a mono-exponential expression, indicative of a single step process. It should be noted, though, that there was an instant increase of FAM emission upon ssDNA addition before the measurement was resumed.

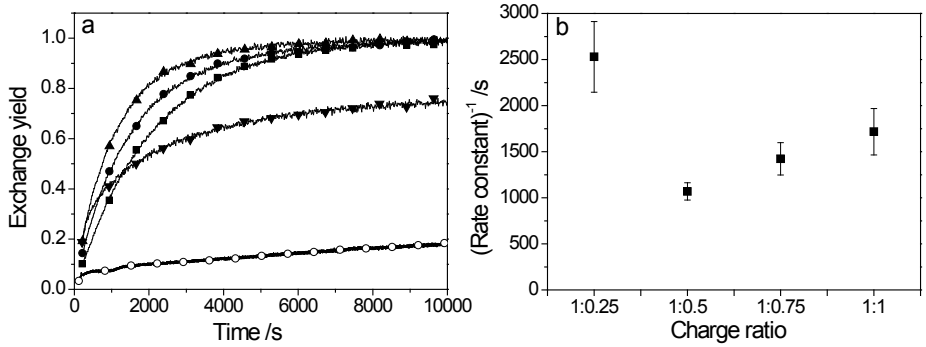

Figure 4.6 (a) Typical kinetic traces for strand exchange in the presence of liposomes containing 35% positively charged lipid, at different lipid:DNA charge ratios: 1:1 (■), 1:0.75 (●), 1:0.5 (▲), 1:0.25 (▼), when adding a five-fold excess of unlabelled single strand to a doubly labelled duplex. For comparison the exchange reaction in absence of liposomes with the same amount of DNA as for a lipid:DNA charge ratio of 1:0.5 is shown (○). **(b)** Inverse rate constants at different lipid:DNA charge ratios. Error bars are one standard deviation based on at least four measurements. Measurements were performed at 25°C.

The fastest rate of strand exchange was observed for a lipid:DNA charge ratio of 1:0.5. Considering that approximately half of the lipid content is in the inner leaflet of the liposome this corresponds to a lipid:DNA charge ratio of 1:1 on the surface. Thus, at this charge ratio all of the DNA strands can bind to the liposome surface, which facilitates a fast exchange. At the lower lipid:DNA charge ratio there is excess positive charge on the liposome and the bound DNA molecules will on average be more separated. The limited access to single strands close to a destabilized duplex on the surface will decrease the rate of strand exchange. At lipid:DNA charge ratios higher than 1:0.5 the single- and double-stranded DNA molecules will compete for the binding sites on the liposome. SsDNA, being more flexible and hydrophobic than dsDNA, has been suggested to interact more strongly with cationic liposomes [118-119]. A weaker binding of dsDNA will decrease the average time spent on the liposomes and this, together with the fact that ssDNA molecules are in excess, slows down the strand exchange. This hypothesis was verified by an experiment where the unlabelled ssDNA was added to the liposomes prior to the addition of doubly labelled duplex, resulting in a decreased strand exchange rate.

The addition of DNA to the positively charged liposomes caused an immediate aggregation of the system. To investigate the importance of this behaviour a PEG-modified lipid was added to the liposomes and sample aggregation as well as strand exchange was monitored. Dynamic light scattering measurements verified that aggregation was significantly reduced for liposomes containing the PEG-modified lipid. The rates of DNA strand exchange obtained with PEG-modified liposomes were lower than with unmodified liposomes, but exchange was still efficient compared to in bulk without liposomes. This shows that strand exchange actually takes place on the surface of single liposomes, not only at the interface between aggregated vesicles.

It is concluded in Paper III that the availability of ssDNA in close proximity to the dsDNA on the liposome surface is important for efficient strand exchange, although the destabilization of the initial duplex is considered to be the rate limiting step. The mechanism for strand exchange on the liposome surface, compared to in bulk, was further investigated in Paper IV, where the effect of single base mismatches, particularly the position of the mismatch, on strand exchange rate was studied.

The correction of a mismatched duplex, *i.e.* the strand exchange reaction of a doubly labelled duplex **[t+mX]**, where **t** is the FAM-labelled strand and **mX** is the TAMRA-labelled strand with a mismatch at position **X** (counting from the 5'-end), and a five-fold excess of the unlabelled, fully complementary strand **c**, is highly sensitive to the position of the mismatch. As shown by the inverse rate constants in Table 4.2 the exchange is very fast if the mismatch is positioned close to the end of the duplex (**[t+m4] [t+m5]**), whereas a mismatch in the middle of the sequence (**[t+m10]**) has a minor effect on the strand exchange rate compared to the reaction of the fully complementary duplex **[t+c]** with single strand **c**. This is not an effect of differences in intrinsic duplex stability, since these three mismatched duplexes have similar melting temperatures in solution. Furthermore, for a mismatch positioned close to the end of the duplex (**[t+m3]**, **[t+m4]** or **[t+m5]**) there is a high kinetic discrimination favouring exchange with the fully complementary strand **c** over exchange with an ssDNA with the same mismatch (**m3**, **m4** or **m5**, respectively)**.** This was not observed for the duplex with a centrally positioned mismatch (**[t+m10]**), where the strand exchange rates obtained were similar regardless if the added single strand was the fully complementary strand **c** or the mismatched strand **m10** (Table 4.2).

Table 4.2 Melting temperatures, T_m,[a] for the fully complementary duplex **[t + c]** and the single base mismatched duplexes **[t + mX]** with the mismatch at position **X**, and inverse rate constants, τ, for strand exchange of a doubly labelled duplex **[t + c]** or **[t + mX]** when adding a five-fold excess of the unlabelled fully complementary strand (**c**) or a strand with an identical mismatch (**mX**)[b]

			ssDNA added	
			τ /s	
dsDNA	T_m /°C, dl	T_m /°C, sl	**c**	**mX**
[t + c]	53.3	51.3	1030 ± 80	-
[t + m3]	51.8	49.8	430 ± 90	1220 ± 90
[t + m4]	46.0	43.8	<200[c]	940 ± 70
[t + m5]	46.0	43.8	<200[c]	770 ± 150
[t + m10]	47.0	44.0	900 ± 190	660 ± 110

[a] Measured on doubly labelled duplexes (dl) and duplexes only labelled with FAM on the **t** strand (sl). [b] Measurements were performed at 25°C. Standard deviations are based on at least three measurements. [c] Rate constant too high to be determined properly.

The results presented in Paper IV indicate that dsDNA opens up in a zipper-like fashion and it is concluded that the fraying ends are largely stabilized on the liposome surface. Polyelectrolyte theory and the strong interactions with liposomes observed for ssDNA support this hypothesis [118-119]. Contrary observations, that mismatches in the middle of the sequence largely affect duplex dissociation and strand exchange rates in different solution set-ups, have been reported [120-122]. These variations suggest that exchange mechanisms are fundamentally different in bulk solution and on the liposome surface.

Based on the proposed mechanism the rate of strand exchange for a duplex with a mismatch at position 3 (**[t+m3]**) with the fully complementary strand should be at least as high as for **[t+m4]** and **[t+m5]**. The observed rate was, however, significantly slower for this mismatched duplex. This can be explained by the **[t+m3]** duplex having a significantly higher melting temperature than the other mismatched duplexes and therefore being more stable (Table 4.2). Considering the higher intrinsic stability of this duplex it is notable that the **m3** strand was exchanged significantly faster than the **m10** strand, supporting the suggested mechanism.

The addition of an excess of mismatched strand (**m3** or **m4**) to a fully complementary duplex resulted in a small, but still significant, increase in FAM emission (Figure 4.7), indicating an exchange from the fully complementary duplex **[t+c]** to a mismatched duplex **[t+m3]** or **[t+m4]**. The exchange occurs even though the doubly labelled **[t+c]** duplex is more stable than the singly labelled **[t+m3]** and **[t+m4]** duplexes (Table 4.2). The driving force for the strand exchange reaction can be considered a balance between the higher concentration of single strands and a greater product duplex stability. Apparently, the former to some extent overcomes even the 9.5°C difference in melting temperature between **[t+c]** and **[t+m4]**. Although the exchange yield was close to the theoretical yield, *i.e.* the final composition was close to the equilibrium composition, when adding **c** as well as **m3** or **m4** to **[t+c]**, the reactions with mismatched single strands showed a lower final emission (Figure 4.7). This suggests a non-quantitative exchange due to the lower stability of the mismatched duplexes. In fact, the degree of exchange is directly related to the stability of the product duplex, illustrated by the higher final emission of the reaction with **m3**, resulting in a more stable duplex (**[t+m3]**) than the exchange with **m4**. Interestingly, the observed strand exchange rate was similar regardless if the **c**, **m3** or **m4** strand was added to the **[t+c]** duplex. The results presented in Paper IV thus support the conclusions from Paper III that the rate of strand exchange on the liposome surface is directed by the stability of the initial duplex rather than the formation of the product duplex. Contrary to these findings, mismatched DNA has been reported to show slower hybridization kinetics in solution than fully complementary DNA [123], again indicating a different behaviour of DNA on the liposome surface compared to in bulk solution.

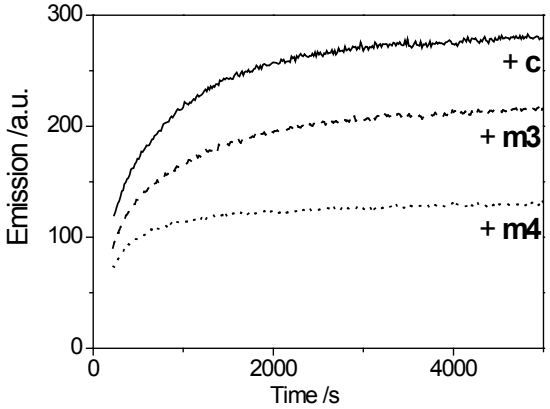

Figure 4.7 Representative kinetic traces for strand exchange when adding a five-fold excess of fully complementary strand (**c**, solid line) or single base mismatched strands (**m3**, dashed line, and **m4**, dotted line) to fully complementary duplex **[t+c]**. Measurements were performed at 25°C.

4.2.2. CATALYSIS BY MOLECULAR CROWDING

EXAMINING THE STRUCTURES of the nucleoprotein filaments of RecA [16] and Rad51 (Paper II) reveals that DNA in the interior of the filaments is exposed to hydrophobic surfaces in a predominantly crowded and non-polar environment. In the model system presented in Paper V the strand exchange reaction was studied in a set-up aiming to mimic these conditions. PEG, which is commonly used to simulate intracellular macromolecular crowding, was found to markedly enhance the rate of strand exchange.

DNA strand exchange was monitored using the same FRET assay as was used in Papers III and IV and typical kinetic traces for the reaction in aqueous solutions of PEG at different concentrations are shown in Figure 4.8. The accelerating effect by PEG is significant, with a rapid increase in strand exchange rate above a critical concentration of 40% (w/w) PEG.

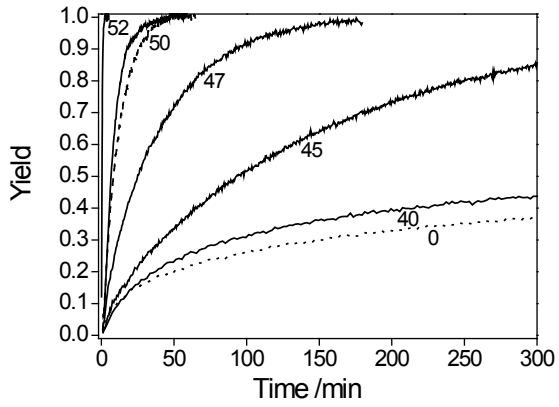

Figure 4.8 Normalized kinetic traces for DNA strand exchange in phosphate buffer (50 mM Na^+, dotted line), in phosphate buffer solutions containing 40%, 45%, 47%, 50%, and 52% (w/w) PEG (solid lines), and 50% (w/w) PEG in phosphate buffer with additional salt (200 mM Na^+ in total, dashed line). Measurements were performed at 37°C.

Table 4.3 shows the inverse rate constants from a mono-exponential fit of the kinetic data, along with melting temperatures at the different PEG concentrations. Although the addition of PEG results in decreased duplex stability, this alone does not explain the accelerating effect since it is preserved in a sample with additional salt that increases duplex stability.

Table 4.3 Inverted rate constants of strand exchange (τ) and DNA melting temperatures (T_m) at different PEG concentrations.

% PEG (w/w)	τ /s	T_m [b] /°C
0	$1.3 \cdot 10^4 \pm 470$	52 ± 0.4
40	$9.6 \cdot 10^3 \pm 290$	53 ± 0.6
45	$8.6 \cdot 10^3 \pm 580$	50 ± 0.7
47	$2.2 \cdot 10^3 \pm 17$	48 ± 0.5
50	440 ± 40	46 ± 0.6
52	34 ± 2.6	44 ± 0.3
50[a]	640 ± 20	55 ± 0.6

[a] 200 mM Na^+. [b] Standard deviations based on two independent experiments.

The dramatic effect of PEG on the rate of strand exchange, while only moderately affecting the melting temperature, could possibly be explained by the large stabilizing effect of base stacking in a DNA double helix [25, 27]. The addition of salt will shield the electrostatic backbone repulsion, but will not affect base stacking. Non-polar cosolutes have, however, been found to disrupt base stacking forces [124-125] and PEG has due to its hydrocarbon groups a non-polar character. In Paper V it is speculated that PEG not only induces molecular crowding that increases the effective DNA concentration, but in addition provides a possibility for hydrophobic interactions with the DNA bases. This would allow more extensive duplex breathing, facilitating strand exchange. The strong dependency of strand exchange rate on PEG concentration is ascribed a threshold effect in duplex breathing, such that water activity in the bulk, affected by the presence of PEG, must be sufficiently decreased to match the water activity in the DNA helix interior and reduce the stacking forces. It could also be that the PEG concentration needs to be sufficiently high to overcome sterical repulsion between DNA and PEG molecules.

The results presented in Paper V show that DNA strand exchange may be driven, and enhanced, solely by non-ionic forces. This emphasizes that the dynamic behaviour of DNA strands could be substantially dependent on hydrophobic interactions between DNA and its environment.

5. CONCLUDING REMARKS AND FUTURE OUTLOOK

THE WORK PRESENTED in this Thesis deals with different aspects of DNA strand exchange. Structural studies of recombination proteins and investigations of artificial strand exchange in model systems have been performed, with the aim to elucidate the mechanism of the reaction. The results obtained highlight some interesting details on the matter but also raise important questions that remain to be answered.

In the model structure of the HsRad51 filament presented in this Thesis the orientations of the DNA binding loops are indicated, suggesting possible points of protein-DNA interaction. The L1 loop is positioned to have direct contact with DNA and the L2 loop is located in proximity of a putative ATP pocket identified at the monomer-monomer interface. As DNA itself might have a prominent role in the reaction complex, the specific interactions between protein and DNA deserve further investigation. The non-uniformly stretched conformation of DNA in the nucleoprotein filament, suggested by Prévost and Takahashi [39] as well as Chen *et al.* [16], seems to be a key feature in the strand exchange catalysis. The role of the recombinase would thus be to act as a scaffold for the DNA to adopt this structure and to locate the DNA molecules in position for exchange. The structural model of HsRad51 ascribes the L1 loop a positioning function, interacting with DNA in its stretched conformation. The SSLD results for RecA presented in this Thesis also support the hypothesis of a static role for the protein during strand exchange.

HsRad51 was in the work of this Thesis bound to dsDNA. It would be of great interest to investigate the complex with ssDNA using SSLD, particularly to examine any structural variations compared to the present model structure. Furthermore, the preliminary data presented in this Thesis indicate that ATP hydrolysis induces structural changes in HsRad51 and that these effects are dependent on the bound DNA being single-stranded or double-stranded. This study needs to be continued in order to establish the coupling between protein interactions with DNA and nucleotide cofactor and the role of ATP hydrolysis in the strand exchange reaction.

In this Thesis it is demonstrated that SSLD in combination with molecular modelling can serve as a powerful tool to determine detailed three-dimensional structures of proteins in systems not amenable to traditional methods like X-ray crystallography or NMR spectroscopy. The filamentous recombinase-DNA complexes have proven well suited for the technique, but any other protein or protein complex could be an SSLD target provided that it can be macroscopically oriented. Membrane bound peptides and proteins have been studied by LD using liposomes, which can be oriented in a shear flow, as a model membrane [126-

127] and could therefore be obvious candidates for future studies using the SSLD technique.

Two model systems for artificial DNA strand exchange catalysis are described in this Thesis. The liposome based system constitutes a promising platform for studies of DNA behaviour and strand exchange. It is concluded that the exchange mechanism on the liposome surface is fundamentally different from that in bulk solution. The ends of the DNA duplex are shown to breathe significantly more when DNA is bound to the liposome than in solution, leading to a zipper-like opening of dsDNA that facilitates a rapid strand exchange. Duplex destabilization was identified as the rate limiting step and it would thus be of interest to elaborate on this parameter. This could be done by introduction of functional moieties on the surface that could provide *e.g.* hydrophobic or hydrogen bonding interactions.

The rate of DNA strand exchange can also be enhanced by several orders of magnitude in a non-ionic artificial system using PEG as a crowding and catalysing agent, as shown in the work of this Thesis. The proposed mechanism includes increased effective concentration of DNA and more frequent duplex breathing due to a disruption of base stacking forces by hydrophobic interactions between PEG and the DNA bases. The study of strand exchange in presence of PEG indicates that the reaction may be very sensitive to small environmental variations. The crowded conditions with possibilities of hydrophobic interactions induced by PEG are suggested to be mimicking the environment experienced by DNA in the recombinase nucleoprotein filaments. Thus from the results presented it may be speculated that the efficiency of DNA strand exchange *in vivo* could be determined, or at least affected, by a few hydrophobic amino acid residues at the active sites of the recombinases.

That the artificial strand exchange is favoured by molecular crowding can be noted also in the model system based on cationic catalysis. The lipoplexes rapidly form aggregates, although DNA strand exchange is still efficiently enhanced. In fact, the enhancing effect is partly lost when aggregation is prevented. DNA is accumulated and the reaction does take place on the surface of individual lipid vesicles, as shown in the work of this Thesis, but at the crowded interface between aggregated liposomes strand exchange is further accelerated possibly due to an even higher local concentration of DNA and shortened diffusion paths for single strands approaching the duplex.

The model system studies presented in this Thesis may not only contribute to shedding light on the mechanism of DNA strand exchange. Experiences regarding DNA stability and dynamics made in this work may also be useful in other contexts including DNA-lipid, or DNA-surface, interactions and assembly of DNA nanostructures.

6. AUTHOR'S ACKNOWLEDGEMENTS

This Thesis has a single author but many people have, in one way or another, contributed to the work presented herein. I would like to use this space to mention some of them.

Bengt, I am happy that you took me along on this journey. Thank you for giving me the opportunity to work on this fascinating project and for believing in it all the way.

Fredrik, you have meant a lot to me. I really enjoyed working with you and I appreciate your co-supervision. Thank you for caring so much and for your support whenever I have needed it.

Katsumi, I am most grateful for your hands-on supervision in the lab and for teaching me a lot about protein engineering and purification, in particular about recombination proteins.

Anna, you have been invaluable to the Rad51 project. Your expertise brought it a great step forward and I am glad you joined the group. Thank you for all your efforts with the paper when I was busy with "baby stuff", no one else could have done it better!

Bobo, it was a pleasure supervising your master's thesis project, you did a great job and I learnt a lot meanwhile.

Masa and Francesca are acknowledged for co-authorship. Masa, you have been a valuable source of knowledge on recombination proteins. Thank you for advice and experimental contributions.

Mikael, Helene and Niklas, you have been excellent company in the office during the years. Special thanks to Helene for therapeutic conversations!

Louise, thanks for your efforts with the Rad51 protein and for carrying on the project. I am sure it will be great!

All the people, present and former, that make Physical Chemistry such a nice place to work are acknowledged. You provide entertaining lunch room chats as well as valuable advice on research matters and emergency assistance with bad behaving instruments. Thank you!

Finally, I would like to acknowledge the ones that did not contribute to the Thesis as such but who are most important to me: my family. Thank you for love and support, at all times. Special thanks go to P-A, Linnéa and Albin, for reminding me of what is really important in life.

7. BIBLIOGRAPHY

1. Brendel, V.*, et al.*, *Evolutionary comparisons of RecA-like proteins across all major kingdoms of living organisms.* J Mol Evol, 1997. **44**(5): 528-41.
2. Benson, F.E.*, et al.*, *Purification and characterization of the human Rad51 protein, an analogue of E. coli RecA.* Embo J, 1994. **13**(23): 5764-71.
3. Lusetti, S.L. and M.M. Cox, *The bacterial RecA protein and the recombinational DNA repair of stalled replication forks.* Annu Rev Biochem, 2002. **71**: 71-100.
4. Sehorn, M.G.*, et al.*, *Human meiotic recombinase Dmc1 promotes ATP-dependent homologous DNA strand exchange.* Nature, 2004. **429**(6990): 433-37.
5. Seitz, E.M.*, et al.*, *RadA protein is an archaeal RecA protein homolog that catalyzes DNA strand exchange.* Genes Dev, 1998. **12**(9): 1248-53.
6. Shinohara, A.*, et al.*, *Rad51 protein involved in repair and recombination in S. cerevisiae is a RecA-like protein.* Cell, 1992. **69**(3): 457-70.
7. Kowalczykowski, S.C.*, et al.*, *Biochemistry of homologous recombination in Escherichia coli.* Microbiol Rev, 1994. **58**(3): 401-65.
8. Egelman, E.H. and A. Stasiak, *Electron-Microscopy of Reca-DNA Complexes - 2 Different States, Their Functional-Significance and Relation to the Solved Crystal-Structure.* Micron, 1993. **24**(3): 309-24.
9. Yu, X.*, et al.*, *Domain structure and dynamics in the helical filaments formed by RecA and Rad51 on DNA.* Proc Natl Acad Sci U S A, 2001. **98**(15): 8419-24.
10. DiCapua, E.*, et al.*, *Complexes of RecA protein in solution. A study by small angle neutron scattering.* J Mol Biol, 1990. **214**(2): 557-70.
11. Ellouze, C.*, et al.*, *Evidence for elongation of the helical pitch of the RecA filament upon ATP and ADP binding using small-angle neutron scattering.* Eur J Biochem, 1995. **233**(2): 579-83.
12. Morimatsu, K.*, et al.*, *Arrangement of RecA protein in its active filament determined by polarized-light spectroscopy.* Proc Natl Acad Sci U S A, 2002. **99**(18): 11688-93.
13. Nordén, B.*, et al.*, *Structure of RecA-DNA complexes studied by combination of linear dichroism and small-angle neutron scattering measurements on flow-oriented samples.* J Mol Biol, 1992. **226**(4): 1175-91.
14. Takahashi, M.*, et al.*, *Binding stoichiometry and structure of RecA-DNA complexes studied by flow linear dichroism and fluorescence spectroscopy. Evidence for multiple heterogeneous DNA co-ordination.* J Mol Biol, 1989. **205**(1): 137-47.
15. Story, R.M.*, et al.*, *The structure of the E. coli recA protein monomer and polymer.* Nature, 1992. **355**(6358): 318-25.
16. Chen, Z.*, et al.*, *Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures.* Nature, 2008. **453**(7194): 489-94.
17. Hagmar, P.*, et al.*, *Structure of DNA-RecA complexes studied by residue differential linear dichroism and fluorescence spectroscopy for a genetically engineered RecA protein.* J Mol Biol, 1992. **226**(4): 1193-205.
18. Tajima, K.*, et al.*, *Simple basic peptides activate DNA strand exchange.* Chemistry Letters, 2003. **32**(5): 470-71.
19. Kim, W.J.*, et al.*, *Comb-type cationic copolymer expedites DNA strand*

exchange while stabilizing DNA duplex. Chemistry, 2001. **7**(1): 176-80.
20. Choi, S.W.*, et al.*, *Activation of DNA strand exchange by cationic comb-type copolymers: effect of cationic moieties of the copolymers.* Nucleic Acids Res, 2008. **36**(1): 342-51.
21. Avery, O.T.*, et al.*, *Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types : Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii.* J Exp Med, 1944. **79**(2): 137-58.
22. Watson, J.D. and F.H. Crick, *Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid.* Nature, 1953. **171**(4356): 737-38.
23. Crick, F.H., *On protein synthesis.* Symp Soc Exp Biol, 1958. **12**: 138-63.
24. Crick, F.H.*, et al.*, *General nature of the genetic code for proteins.* Nature, 1961. **192**: 1227-32.
25. Kool, E.T., *Hydrogen bonding, base stacking, and steric effects in dna replication.* Annu Rev Biophys Biomol Struct, 2001. **30**: 1-22.
26. Petersheim, M. and D.H. Turner, *Base-stacking and base-pairing contributions to helix stability: thermodynamics of double-helix formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp.* Biochemistry, 1983. **22**(2): 256-63.
27. Yakovchuk, P.*, et al.*, *Base-stacking and base-pairing contributions into thermal stability of the DNA double helix.* Nucleic Acids Res, 2006. **34**(2): 564-74.
28. Bloomfield, V.A.*, et al.*, *Nucleic acids: structures, properties, and functions*. 2000, Sausalito, Calif.: University Science Books.
29. Roca, A.I. and M.M. Cox, *RecA protein: structure, function, and role in recombinational DNA repair.* Prog Nucleic Acid Res Mol Biol, 1997. **56**: 129-223.
30. Kowalczykowski, S.C., *Initiation of genetic recombination and recombination-dependent replication.* Trends Biochem Sci, 2000. **25**(4): 156-65.
31. Sung, P. and H. Klein, *Mechanism of homologous recombination: mediators and helicases take on regulatory functions.* Nat Rev Mol Cell Biol, 2006. **7**(10): 739-50.
32. San Filippo, J.*, et al.*, *Mechanism of eukaryotic homologous recombination.* Annu Rev Biochem, 2008. **77**: 229-57.
33. Clark, A.J. and A.D. Margulies, *Isolation And Characterization Of Recombination-Deficient Mutants Of Escherichia Coli K12.* Proc Natl Acad Sci U S A, 1965. **53**: 451-59.
34. Shinohara, A.*, et al.*, *Cloning of human, mouse and fission yeast recombination genes homologous to RAD51 and recA.* Nat Genet, 1993. **4**(3): 239-43.
35. Larkin, M.A.*, et al.*, *Clustal W and Clustal X version 2.0.* Bioinformatics, 2007. **23**(21): 2947-48.
36. Ogawa, T.*, et al.*, *Similarity of the yeast RAD51 filament to the bacterial RecA filament.* Science, 1993. **259**(5103): 1896-99.
37. Stasiak, A. and E. Di Capua, *The helicity of DNA in complexes with recA protein.* Nature, 1982. **299**(5879): 185-86.
38. Stasiak, A.*, et al.*, *Elongation of duplex DNA by recA protein.* J Mol Biol, 1981. **151**(3): 557-64.
39. Prévost, C. and M. Takahashi, *Geometry of the DNA strands within the RecA nucleofilament: role in homologous recombination.* Q Rev Biophys, 2003. **36**(4): 429-53.
40. Bianco, P.R.*, et al.*, *DNA strand exchange proteins: a biochemical and physical comparison.* Front Biosci, 1998. **3**.
41. Cox, M.M., *The bacterial RecA protein as a motor protein.* Annu Rev Microbiol, 2003. **57**: 551-77.

42. Cox, M.M., *Motoring along with the bacterial RecA protein.* Nat Rev Mol Cell Biol, 2007. **8**(2): 127-38.
43. Howard-Flanders, P.*, et al.*, *Role of RecA protein spiral filaments in genetic recombination.* Nature, 1984. **309**(5965): 215-19.
44. Kowalczykowski, S.C. and R.A. Krupp, *DNA-strand exchange promoted by RecA protein in the absence of ATP: implications for the mechanism of energy transduction in protein-promoted nucleic acid transactions.* Proc Natl Acad Sci U S A, 1995. **92**(8): 3478-82.
45. Menetski, J.P. and S.C. Kowalczykowski, *Enhancement of Escherichia coli RecA protein enzymatic function by dATP.* Biochemistry, 1989. **28**(14): 5871-81.
46. Rehrauer, W.M. and S.C. Kowalczykowski, *Alteration of the nucleoside triphosphate (NTP) catalytic domain within Escherichia coli recA protein attenuates NTP hydrolysis but not joint molecule formation.* J Biol Chem, 1993. **268**(2): 1292-97.
47. Bugreev, D.V.*, et al.*, *Activation of human meiosis-specific recombinase Dmc1 by Ca2+.* J Biol Chem, 2005. **280**(29): 26886-95.
48. Bugreev, D.V. and A.V. Mazin, *Ca2+ activates human homologous recombination protein Rad51 by modulating its ATPase activity.* Proc Natl Acad Sci U S A, 2004. **101**(27): 9988-93.
49. Chi, P.*, et al.*, *Roles of ATP binding and ATP hydrolysis in human Rad51 recombinase function.* DNA Repair (Amst), 2006. **5**(3): 381-91.
50. Ristic, D.*, et al.*, *Human Rad51 filaments on double- and single-stranded DNA: correlating regular and irregular forms with recombination function.* Nucleic Acids Res, 2005. **33**(10): 3292-302.
51. Story, R.M. and T.A. Steitz, *Structure of the recA protein-ADP complex.* Nature, 1992. **355**(6358): 374-76.
52. Datta, S.*, et al.*, *Structural studies on MtRecA-nucleotide complexes: insights into DNA and nucleotide binding and the structural signature of NTP recognition.* Proteins, 2003. **50**(3): 474-85.
53. Datta, S.*, et al.*, *Crystal structures of Mycobacterium smegmatis RecA and its nucleotide complexes.* J Bacteriol, 2003. **185**(14): 4280-84.
54. Datta, S.*, et al.*, *Crystal structures of Mycobacterium tuberculosis RecA and its complex with ADP-AlF(4): implications for decreased ATPase activity and molecular aggregation.* Nucleic Acids Res, 2000. **28**(24): 4964-73.
55. Krishna, R.*, et al.*, *Crystallographic identification of an ordered C-terminal domain and a second nucleotide-binding site in RecA: new insights into allostery.* Nucleic Acids Res, 2006. **34**(8): 2186-95.
56. Krishna, R.*, et al.*, *Snapshots of RecA protein involving movement of the C-domain and different conformations of the DNA-binding loops: crystallographic and comparative analysis of 11 structures of Mycobacterium smegmatis RecA.* J Mol Biol, 2007. **367**(4): 1130-44.
57. Prabu, J.R.*, et al.*, *Functionally important movements in RecA molecules and filaments: studies involving mutation and environmental changes.* Acta Crystallogr D Biol Crystallogr, 2008. **64**(Pt 11): 1146-57.
58. Xing, X. and C.E. Bell, *Crystal structures of Escherichia coli RecA in a compressed helical filament.* J Mol Biol, 2004. **342**(5): 1471-85.
59. Xing, X. and C.E. Bell, *Crystal structures of Escherichia coli RecA in complex with MgADP and MnAMP-PNP.* Biochemistry, 2004. **43**(51): 16142-52.
60. Cox, M.M., *The bacterial RecA protein: Structure, function and*

regulation. Topics in Current Genetics, 2007. **17**: 53-94.
61. Hortnagel, K.*, et al.*, *Saturation mutagenesis of the E. coli RecA loop L2 homologous DNA pairing region reveals residues essential for recombination and recombinational repair.* J Mol Biol, 1999. **286**(4): 1097-106.
62. Nastri, H.G. and K.L. Knight, *Identification of residues in the L1 region of the RecA protein which are important to recombination or coprotease activities.* J Biol Chem, 1994. **269**(42): 26311-22.
63. Malkov, V.V. and R.D. Cameriniotero, *Photocross-Links between Single-Stranded-DNA and Escherichia-Coli Reca Protein Map to Loops L1 (Amino-Acid-Residues 157-164) and L2 (Amino-Acid-Residues 195-209).* Journal of Biological Chemistry, 1995. **270**(50): 30230-33.
64. Morimatsu, K. and T. Horii, *DNA-binding surface of RecA protein photochemical cross-linking of the first DNA binding site on RecA filament.* Eur J Biochem, 1995. **234**(3): 695-705.
65. Wang, Y. and K. Adzuma, *Differential proximity probing of two DNA binding sites in the Escherichia coli recA protein using photo-cross-linking methods.* Biochemistry, 1996. **35**(11): 3563-71.
66. Morimatsu, K.*, et al.*, *Interaction of tyrosine 65 of RecA protein with the first and second DNA strands.* J Mol Biol, 2001. **306**(2): 189-99.
67. Morimatsu, K.*, et al.*, *Interaction of Tyr103 and Tyr264 of the RecA protein with DNA and nucleotide cofactors. Fluorescence study of engineered proteins.* Eur J Biochem, 1995. **228**(3): 779-85.
68. Rehrauer, W.M. and S.C. Kowalczykowski, *The DNA binding site(s) of the Escherichia coli RecA protein.* J Biol Chem, 1996. **271**(20): 11996-2002.
69. Kubista, M.*, et al.*, *Stoichiometry, base orientation, and nuclease accessibility of RecA.DNA complexes seen by polarized light in flow-oriented solution. Implications for the mechanism of genetic recombination.* J Biol Chem, 1990. **265**(31): 18891-97.
70. Aihara, H.*, et al.*, *An interaction between a specified surface of the C-terminal domain of RecA protein and double-stranded DNA for homologous pairing.* J Mol Biol, 1997. **274**(2): 213-21.
71. Kurumizaka, H.*, et al.*, *A possible role of the C-terminal domain of the RecA protein. A gateway model for double-stranded DNA binding.* J Biol Chem, 1996. **271**(52): 33515-24.
72. Lusetti, S.L.*, et al.*, *Magnesium ion-dependent activation of the RecA protein involves the C terminus.* J Biol Chem, 2003. **278**(18): 16381-88.
73. Pellegrini, L.*, et al.*, *Insights into DNA recombination from the structure of a RAD51-BRCA2 complex.* Nature, 2002. **420**(6913): 287-93.
74. Aihara, H.*, et al.*, *The N-terminal domain of the human Rad51 protein binds DNA: structure and a DNA binding surface as revealed by NMR.* J Mol Biol, 1999. **290**(2): 495-504.
75. Conway, A.B.*, et al.*, *Crystal structure of a Rad51 filament.* Nat Struct Mol Biol, 2004. **11**(8): 791-96.
76. Chen, J.*, et al.*, *Insights into the mechanism of Rad51 recombinase from the structure and properties of a filament interface mutant.* Nucleic Acids Res, 2010. **38**(14): 4889-906.
77. VanLoock, M.S.*, et al.*, *Complexes of RecA with LexA and RecX differentiate between active and inactive RecA nucleoprotein filaments.* J Mol Biol, 2003. **333**(2): 345-54.

78. Galkin, V.E., *et al.*, *BRCA2 BRC motifs bind RAD51-DNA filaments.* Proc Natl Acad Sci U S A, 2005. **102**(24): 8537-42.
79. Matsuo, Y., *et al.*, *Roles of the human Rad51 L1 and L2 loops in DNA binding.* Febs J, 2006. **273**(14): 3148-59.
80. Prasad, T.K., *et al.*, *Visualizing the assembly of human Rad51 filaments on double-stranded DNA.* J Mol Biol, 2006. **363**(3): 713-28.
81. Zhang, X.P., *et al.*, *Loop 2 in Saccharomyces cerevisiae Rad51 protein regulates filament formation and ATPase activity.* Nucleic Acids Res, 2009. **37**(1): 158-71.
82. VanLoock, M.S., *et al.*, *ATP-mediated conformational changes in the RecA filament.* Structure, 2003. **11**(2): 187-96.
83. Grigorescu, A.A., Vissers, J.H.A., Ristic, D., Pigli, Y.Z., Lynch, T.W., Wyman, C., Rice, P.A., *Inter-subunit interactions that coordinate Rad51's activities.* Nucleic Acids Research, 2009. **37**(2): 557-67.
84. Renodon-Corniere, A., *et al.*, *Structural analysis of the human Rad51 protein-DNA complex filament by tryptophan fluorescence scanning analysis: transmission of allosteric effects between ATP binding and DNA binding.* J Mol Biol, 2008. **383**(3): 575-87.
85. Kim, W.J., *et al.*, *DNA strand exchange stimulated by spontaneous complex formation with cationic comb-type copolymer.* J Am Chem Soc, 2002. **124**(43): 12676-77.
86. Hollas, J.M., *Modern spectroscopy*. 3rd ed. 1996, Chichester: J. Wiley.
87. Lakowicz, J.R., *Principles of fluorescence spectroscopy*. 3rd ed. 2006, New York: Springer.
88. Nordén, B., *et al.*, *Linear Dichroism and Circular Dichroism. A Textbook on Polarized-Light Spectroscopy*. 2010, Cambridge: RSC Publishing.
89. Wetlaufer, D.B., *Ultraviolet spectra Of Proteins and Amino Acids.* Advances in Protein Chemistry, 1963. **17**: 303-90.
90. Wu, P. and L. Brand, *Resonance energy transfer: methods and applications.* Anal Biochem, 1994. **218**(1): 1-13.
91. Sambrook, J. and D.W. Russell, *Molecular cloning: a laboratory manual*. 3rd ed. 2001, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.
92. Kunkel, T.A., *Rapid and efficient site-specific mutagenesis without phenotypic selection.* Proc Natl Acad Sci U S A, 1985. **82**(2): 488-92.
93. Studier, F.W., *et al.*, *Use of T7 RNA polymerase to direct expression of cloned genes.* Methods Enzymol, 1990. **185**: 60-89.
94. Bullock, W.O., *et al.*, *Xl1-Blue - a High-Efficiency Plasmid Transforming Reca Escherichia-Coli Strain with Beta-Galactosidase Selection.* Biotechniques, 1987. **5**(4): 376-79.
95. Holmberg, K., *Surfactants and polymers in aqueous solution*. 2nd ed. 2003, Chichester, West Sussex, England ; Hoboken, NJ: John Wiley & Sons.
96. Felgner, P.L., *et al.*, *Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure.* Proc Natl Acad Sci U S A, 1987. **84**(21): 7413-17.
97. Chesnoy, S. and L. Huang, *Structure and function of lipid-DNA complexes for gene delivery.* Annu Rev Biophys Biomol Struct, 2000. **29**: 27-47.
98. Safinya, C.R., *Structures of lipid-DNA complexes: supramolecular assembly and gene delivery.* Curr Opin Struct Biol, 2001. **11**(4): 440-48.
99. Allen, T.M., *et al.*, *Liposomes containing synthetic lipid derivatives of poly(ethylene glycol) show prolonged circulation half-*

lives in vivo. Biochim Biophys Acta, 1991. **1066**(1): 29-36.
100. Klibanov, A.L.*, et al.*, *Amphipathic polyethyleneglycols effectively prolong the circulation time of liposomes.* FEBS Lett, 1990. **268**(1): 235-37.
101. Torchilin, V.P. and V.S. Trubetskoy, *Which Polymers Can Make Nanoparticulate Drug Carriers Long-Circulating.* Adv Drug Deliv Rev, 1995. **16**(2-3): 141-55.
102. Rosemeyer, H., *Nucleolipids: natural occurrence, synthesis, molecular recognition, and supramolecular assemblies as potential precursors of life and bioorganic materials.* Chem Biodivers, 2005. **2**(8): 977-1063.
103. Banchelli, M.*, et al.*, *Molecular recognition drives oligonucleotide binding to nucleolipid self-assemblies.* Angew Chem Int Ed Engl, 2007. **46**(17): 3070-73.
104. Milani, S.*, et al.*, *Intercalation of single-strand oligonucleotides between nucleolipid anionic membranes: a neutron diffraction study.* Langmuir, 2009. **25**(7): 4084-92.
105. Mayer, L.D.*, et al.*, *Vesicles of variable sizes produced by a rapid extrusion procedure.* Biochim Biophys Acta, 1986. **858**(1): 161-68.
106. Ellis, R.J. and A.P. Minton, *Cell biology: join the crowd.* Nature, 2003. **425**(6953): 27-28.
107. Zimmerman, S.B. and A.P. Minton, *Macromolecular crowding: biochemical, biophysical, and physiological consequences.* Annu Rev Biophys Biomol Struct, 1993. **22**: 27-65.
108. Ellis, R.J., *Macromolecular crowding: obvious but underappreciated.* Trends Biochem Sci, 2001. **26**(10): 597-604.
109. Miyoshi, D. and N. Sugimoto, *Molecular crowding effects on structure and stability of DNA.* Biochimie, 2008. **90**(7): 1040-51.
110. Zhou, H.X.*, et al.*, *Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences.* Annu Rev Biophys, 2008. **37**: 375-97.
111. Karimata, H.*, et al.*, *Stabilization of a DNA duplex under molecular crowding conditions of PEG.* Nucleic Acids Symp Ser (Oxf), 2004(48): 107-08.
112. Nakano, S.*, et al.*, *The effect of molecular crowding with nucleotide length and cosolute structure on DNA duplex stability.* J Am Chem Soc, 2004. **126**(44): 14330-31.
113. Spink, C.H. and J.B. Chaires, *Selective stabilization of triplex DNA by poly(ethylene glycols).* Journal of the American Chemical Society, 1995. **117**(51): 12887-88.
114. Albertsson, P.A., *Partition Studies on Nucleic Acids. I. Influence of Electrolytes, Polymer Concentration and Nucleic Acid Conformation of the Partition in the Dextran-Polyethylene Glycol System.* Biochim Biophys Acta, 1965. **103**: 1-12.
115. Sturesson, S.*, et al.*, *Partition of macromolecules and cell particles in aqueous two-phase systems based on hydroxypropyl starch and poly(ethylene glycol).* Appl Biochem Biotechnol, 1990. **26**(3): 281-95.
116. Platt, J.R., *Classification of Spectra of Cata-Condensed Hydrocarbons.* Journal of Chemical Physics, 1949. **17**(5): 484-95.
117. Selmane, T.*, et al.*, *Identification of the subunit-subunit interface of Xenopus Rad51.1 protein: similarity to RecA.* J Mol Biol, 2004. **335**(4): 895-904.
118. Jonsson, M. and P. Linse, *Polyelectrolyte-macroion complexation. II. Effect of chain flexibility.* J Chem Phys, 2001. **115**(23): 10975-85.
119. Costa, D.*, et al.*, *Responsive polymer gels: double-stranded versus single-*

stranded DNA. J Phys Chem B, 2007. **111**(37): 10886-96.
120. Tibanyenda, N.*, et al.*, *The effect of single base-pair mismatches on the duplex stability of d(T-A-T-T-A-A-T-A-T-C-A-A-G-T-T-G) . d(C-A-A-C-T-T-G-A-T-A-T-T-A-A-T-A).* Eur J Biochem, 1984. **139**(1): 19-27.
121. Kim, W.J.*, et al.*, *Cationic comb-type copolymers for DNA analysis.* Nat Mater, 2003. **2**(12): 815-20.
122. Yamana, K.*, et al.*, *DNA mismatch detection using a pyrene-excimer-forming probe.* Chem Commun (Camb), 2005(19): 2509-11.
123. Jensen, K.K.*, et al.*, *Kinetics for hybridization of peptide nucleic acids (PNA) with DNA and RNA studied with the BIAcore technique.* Biochemistry, 1997. **36**(16): 5072-77.
124. Sen, A. and P.E. Nielsen, *On the stability of peptide nucleic acid duplexes in the presence of organic solvents.* Nucleic Acids Res, 2007. **35**(10): 3367-74.
125. Sen, A. and P.E. Nielsen, *Hydrogen bonding versus stacking stabilization by modified nucleobases incorporated in PNA.DNA duplexes.* Biophys Chem, 2009. **141**(1): 29-33.
126. Ardhammar, M.*, et al.*, *Invisible liposomes: refractive index matching with sucrose enables flow dichroism assessment of peptide orientation in lipid vesicle membrane.* Proc Natl Acad Sci U S A, 2002. **99**(24): 15313-17.
127. Caesar, C.E.*, et al.*, *Assigning membrane binding geometry of cytochrome C by polarized light spectroscopy.* Biophys J, 2009. **96**(8): 3399-411.

CHALMERS UNIVERSITY OF TECHNOLOGY
SE-412 96 Göteborg, Sweden
Telephone: +46-(0)31 772 10 00
www.chalmers.se

DNA-strängutbyte, den centrala reaktionen i en process där bitar av DNA byter plats och kan kombineras ihop på nya sätt, är fundamentalt viktigt för alla levande celler. Processen utgör en del av det maskineri som används för att reparera DNA men är också ett sätt att åstadkomma genetisk variation. Arbetet som presenteras i den här avhandlingen har som mål att kartlägga reaktionsmekanismen genom att undersöka DNA-strängutbyte ur olika aspekter. Strukturen av de proteiner som katalyserar (påskyndar) reaktionen i cellen har studerats, liksom strängutbyte i två olika modellsystem.

Det bakteriella rekombinationsproteinet RecA och den mänskliga motsvarigheten Rad51 bildar båda, i sin aktiva form, fiberliknande komplex med DNA. Dessa har studerats med en spektroskopisk teknik baserad på planpolariserat ljus. I den här avhandlingen visas att orienteringen av specifika aminosyror i proteinet kan beräknas utifrån spektroskopiska mätningar gjorda på systematiskt modifierade proteiner och att den informationen kan användas för att bygga en modell av det aktiva proteinets struktur. För RecA har de här studierna lett till slutsatsen att proteinets struktur är lika både före och efter det att strängutbytet har skett, vilket indikerar att RecA endast har en statisk roll under reaktionens gång. Studien av Rad51 illustrerar hur experimentella data kan kombineras med teoretisk molekylmodellering. Resultatet, som presenteras i den här avhandlingen, är en modell av proteinets struktur i sitt aktiva komplex med DNA. Det är den första detaljerade strukturen som rapporterats för det humana Rad51 proteinet i sin helhet.

I en av de modellstudier av strängutbyte som presenteras i den här avhandlingen har det framkommit att reaktionen katalyseras på en positivt laddad partikelyta och att den negativt laddade dubbelsträngade DNA-molekylen öppnas som ett blixtlås från ena änden för att möjliggöra strängutbyte. I ett annat modellsystem har katalys åstadkommits genom att låta reaktionen ske i närvaro av en oladdad polymer, som reducerar utrymmet för DNA-molekylerna och ger möjlighet till hydrofoba interaktioner, det vill säga interaktioner mellan vattenskyende grupper utan positiva och negativa laddningar. Polymeren har visat sig påskynda strängutbytet avsevärt och resultaten belyser att hydrofoba interaktioner mellan DNA och dess omgivning kan ha stor betydelse för DNA-strängarnas dynamik.