DEVELOPING NOVEL SINGLE MOLECULE ANALYSES OF THE SINGLE-STRANDED DNA BINDING PROTEIN FROM *SULFOLOBUS SOLFATARICUS*

Michael J. Morten

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews

2015

Developing novel single molecule analyses of the single-stranded DNA binding protein from *Sulfolobus solfataricus*

Michael J. Morten

This thesis is submitted in partial fulfilment for the degree of PhD

at the University of St Andrews

University of
St Andrews

May 2015

Table of Contents

1. Introduction 1

1.1. Double-stranded DNA 1

1.2. DNA damage and mutations 2

1.3. DNA repair 3

1.4. Single-stranded nucleic acids 5

1.5. Single-stranded DNA binding proteins 7

1.5.1. Oligonucleotide/oligosaccharide binding folds. 9

1.5.2. Loops 10

1.5.3. Human replication protein A 12

1.5.4. Bacterial SSB 19

1.6. Archaea 23

1.6.1. Archaeal SSB 26

1.7. Single molecule assays 31

1.8. Fluorescence 33

1.9. FRET 37

1.9.1. A brief history 38

1.9.2. FRET theory 42

1.9.3. FRET as a spectroscopic ruler 47

1.10. TIRF microscopy 49

1.11. Surface modifications 52

1.12. Protein & DNA labelling.....58

1.13. Fluorophores.....59

1.14. Single molecule fluorescence investigations into SSBs.....62

2. Materials and methods 64

2.1. SSB expression & purification 64

2.2. Protein Labelling.....65

2.3. Labelling and annealing DNA 67

2.4. Fluorescence experiments.....68

2.4.1. Ratio $_A$.....68

2.5. Binding to a one dimensional lattice.....69

2.5.1. 1:1 binding.....69

2.5.2. Cooperative binding 70

2.6. Wormlike chain model.....75

2.7. Single molecule experiments 76

2.7.1. Cleaning of slides 76

2.7.2. Aminosilation of glass 76

2.7.3. PEGylation of amino-covered slides 77

2.7.4. Microchannel preparation.....77

2.7.5. TIRF experiments.....77

2.7.6. Rate analysis.....78

2.8. Nucleic acid sequences 79

2.9. Exosome degradation 79

3. Ensemble studies of SsoSSB 81

3.1. Introduction 81

3.2. Expression and purification 82

3.3. Tryptophan quenching 83

3.4. Dimensions of SSB/ssDNA complexes 84

3.5. ssDNA Labelling 87

3.6. Protein induced fluorescence enhancement 88

3.7. Protein labelling 90

3.8. FRET 91

3.9. Quenching 92

3.10. PELDOR 93

3.11. Persistence length of naked and decorated ssDNA 98

3.12. Dissociation constants and occupancy 101

3.13. Cooperativity parameters 108

3.14. Quenching suggests monomeric structure 110

3.15. Discussion 111

3.16. Conclusions 116

4. Single molecule analysis of SsoSSB 118

4.1. Non-specific binding to the slide 120

4.2. PIFE 121

4.3. FRET 127

4.4. Full in-trace interpretation 129

4.5. Binding kinetics 132

4.6. Discussion 138

4.7. Conclusions 144

5. SsoSSB binds to RNA 146

5.1. Ensemble fluorescence results 152

5.2. Single molecule fluorescence results 154

5.3. Possible roles for RNA binding 156

5.4. Discussion 158

5.5. Conclusions 160

6. Further work 162

6.1. Applications in high temperature and salt concentrations 162

6.1.1. Ensemble fluorescence 163

6.1.2. Discussion 165

6.2. The role of the acidic C terminal tail 167

6.2.1. *In silico* model of SsoSSB tail binding to the OB fold 170

6.2.2. Ensemble measurement of k_{off} 171

6.2.3. Discussion 172

6.3. Conclusions for further work 176

7. Conclusions 179

8. References 181

List of figures

Figure 1.1: Crystal structure of B DNA. 2

Figure 1.2: A simple diagram showing the pathways and the main intermediates of BER, NER and MMR. 5

Figure 1.3: The structures of RNA 7

Figure 1.4: Cartoon of an OB fold.....10

Figure 1.5: The structure of an OB fold from RPA. 11

Figure 1.6: The structure and domain organisation of RPA. 13

Figure 1.7: A cartoon of RPA binding to ssDNA and exchanging its position with another RPA protein. 16

Figure 1.8: A schematic of the main steps involved NER.....18

Figure 1.9: A cartoon showing the tetrameric structure of EcoSSB bound to ssDNA. 20

Figure 1.10: A cartoon showing two of EcoSSBs binding modes. 22

Figure 1.11: Phylogenetic tree showing the distinction between the 16SRNA sequences of the archaea.....25

Figure 1.12: The crystal structure of a SsoSSB monomer and the relative electrostatic charges on its surface.....27

Figure 1.13: Part of the OB fold from SsoSSB. 28

Figure 1.14: A plot of the molecular potential energy against the internuclear separation during an electronic transfer to an excited state. 35

Figure 1.15: A simple Jablonksi plot. 36

Figure 1.16: A Jablonski plot showing FRET between a donor and an acceptor, which is capable of fluorescing once excited. 38

Figure 1.17: Vectors that describe the transition dipoles involved in FRET between donor and acceptor molecules. 43

Figure 1.18: A schematic showing how the relative intensities of donor and acceptor fluorescent dyes, with R_0 = 6 nm, change with respect to distance. 49

Figure 1.19: A schematic to show how light refracts as it travels into a medium with a lower refractive index (n_1>n_2.) 50

Figure 1.20: Schematic of a prism based FRET microscope. 52

Figure 1.21: A schematic of a surface being passivated with a SAM. 54

Figure 1.22: ssDNA can be specifically immobilised on a quartz slide through a biotin with an eight carbon linker to the 5' end. 55

Figure 1.23: DNA curtains imaged using RPA tagged with m-Cherry. 56

Figure 1.24: Schematics of RecA binding to ssDNA inside a porous lipid nanocontainer. 57

Figure 1.25: The structure of Cy3 and Cy5 dyes. 61

Figure 2.1: Curly arrow mechanism for the maleimide thiol coupling used for attaching fluorescent dyes to exposed thiol groups on cysteine residues on the surface of proteins. 66

Figure 2.2: Four fluorescence spectra to illustrate how the Ratio A value is calculated. 69

Figure 2.3: Schematic of ssNA with a length of 7 nt (M=7) where a protein that has a binding site of 5 nt (n=5), gives a total of 3 possible overlapping binding sites. 71

Figure 2.4: Schematic of two proteins with binding sites of 2 nt on a lattice of length 7 nt. 71

Figure 2.5: Schematic showing a free ligand binding to a) a singly and b) a doubly contiguous binding site (n=2, M=6 and for (a) k=2,j=1 and (b) k=3, j=2.).....72

Figure 3.1: The purification of wild type SsoSSB. 83

Figure 3.2: A cartoon to show the volumes that fluorescent dyes can occupy bound to ssDNA. 85

Figure 3.3: Minimum energy model of two SsoSSB monomers, each bound to 10nt ssDNA Cy3 with a gap of a single nucleotide between the two proteins.....87

Figure 3.4: The viscosity of the solution is a barrier to the rotation around the double bond in the alkyl chain between the *E* and *Z* conformations cyanine dyes can adopt. 89

Figure 3.5: Labelling of SsoSSB with Alexa 647 in the presence of 8 M urea increases the efficiency of the labelling reaction.....91

Figure 3.6: FRET can be used to track SsoSSB binding to ssDNA by using ensemble spectroscopy. 92

Figure 3.7: MALDI-TOF analysis of A114C SsoSSB before (top) and after (bottom) labelling with spin label MTSSL.....94

Figure 3.8: PELDOR can be used to verify the structure of proteins such as Influenza A Virus NS1. 95

Figure 3.9: PELDOR data from SsoSSB MTSL experiments.....97

Figure 3.10: Ensemble FRET spectra of 50 nM 39 nt ssDNA Cy3 Cy5 in the absence and presence of SsoSSB showing a decrease in $ratio_A$, that is consistent with the ssDNA being straightened and increasing the end-to-end distance. 99

Figure 3.11: From the data from PELDOR and end to distance measurements this suggests that SsoSSB binds as a monomer and produces a relatively unordered nucleofilament. 101

Figure 3.12: Tryptophan quenching due to SsoSSB binding to ssDNA..... 103

Figure 3.13: Enhancement of the fluorescence from 10 nM ssDNA Cy3 was measured at different concentrations of unlabelled SsoSSB, and a two fold increase in the intensity was observed. 106

Figure 3.14: Ensemble FRET titrations were also completed adding a SsoSSB Alexa 647 to 10 nM ssDNA Cy3, as shown in the cartoon..... 107

Figure 3.15: A reverse titration adding ssDNA to 50 nM SsoSSB Alexa 647 was completed in triplicate. 108

Figure 3.16: PIFE titrations could also be fitted to Epstein's model that describes the cooperative effects of SsoSSB binding to adjacent sites along ssDNA, taking into account overlapping binding sites and the end effects of a short oligo. 109

Figure 3.17: Quenching of Alexa 647 fluorescence. 110

Figure 4.1: The fluorescence intensities from single molecules of 12C ssDNA Cy3 molecules. 119

Figure 4.2: Images taken from TIRF microscope of a clean PEG slide before and after loading 100 nM SsoSSB labelled with Alexa 647 on to the surface..... 121

Figure 4.3: The intensities of the fluorescence emitted from single molecules of 12C ssDNA Cy3 immobilised on quartz slide have been normalised and plotted against time. 123

Figure 4.4: Histograms, each from approximately 1000 single molecule traces, which show the fluorescence intensity of immobilised 12C ssDNA Cy3 on a quartz slide being exposed to increasing concentrations of unlabelled SsoSSB. 124

Figure 4.5: Histograms and a single molecule trace to show that at high concentrations of SsoSSB the traces showed no dynamics but were approximately twice the intensity of 12C ssDNA Cy3 in the absence of protein. 125

Figure 4.6: Gaussian peaks fitted to histograms showing the frequency distribution of the fluorescence intensity of 12C ssDNA Cy3 at 0 nM and 1 nM unlabelled SsoSSB for approximately 200 molecules. 126

Figure 4.7: Single molecule traces showing SsoSSB Alexa 647 monomers binding to 12C ssDNA Cy3 as bursts of increased acceptor fluorescence accompanied by quenching of the donor intensity, typical of highly efficient FRET events. 128

Figure 4.8: An example of a single molecule trace showing the changes in fluorescence when one SsoSSB Alexa 647 monomer and two monomers bind to 12C ssDNA Cy3.....130

Figure 4.9: An example of how the FRET, quenching and PIFE events can be interpreted in terms of the number of SsoSSB Alexa 647 monomers bound to 12C ssDNA Cy3.....131

Figure 4.10: Two SsoSSB monomers can bind to or dissociate from ssDNA either sequentially or simultaneously. 132

Figure 4.11: The rates of k_{off} and k_{on} for the a single monomer binding to 12C ssDNA Cy3 were calculated by measuring the time (a) immediately before a single monomer binding to or (b) dissociating from ssDNA.....133

Figure 4.12: The dwell times of approximately 200 events showing a single SsoSSB Alexa 647 monomer binding to 12C ssDNA Cy3 were measured at each concentration (0.05, 2.5 and 10 nM) and plotted as histograms. 134

Figure 4.13: The rates of k_{off} and k_{on} for a second monomer binding to 12C ssDNA Cy3 were calculated by measuring the time (a) immediately before the second monomer bound to or (b) dissociated from ssDNA. 136

Figure 4.14: The dwell times of approximately 200 events showing a second SsoSSB Alexa 647 monomer binding to 12C ssDNA Cy3 were measured at each concentration (0.05, 2.5 and 10 nM) and plotted as histograms. 137

Figure 5.1: The results from ITC, performed by Dr Lisa Cubbedu, showing the interaction of SsoSSB with ssDNA and RNA: with 21dA (a) and 21dU (b). 149

Figure 5.2: NMR analysis of SsoSSB binding to RNA and ssDNA by Dr Roland Gamsjaeger. 150

Figure 5.3: Molecules of ssDNA and RNA bound to the OB fold of SsoSSB. Figure made by Dr Roland Gamsjeager. 151

Figure 5.4: Ensemble fluorescence titrations describing the similarities between SsoSSB binding to RNA and ssDNA. 153

Figure 5.5: Single molecule traces from (a) 12C ssDNA Cy3 and (b) 12C RNA Cy3 immobilised on to a microscope slide in the presence of SsoSSB monomers labelled with Alexa 647. 155

Figure 5.6: SSoSSB binding to RNA may suggest a role *in vivo* such as acting as a chaperone of RNA during transport or metabolism. 157

Figure 6.1: The enhanced fluorescence from 50 nM 20C ssDNA Cy3 was observed upon the addition of unlabelled SsoSSB in 1.5 M LiCl. 164

Figure 6.2: PIFE was used again to study SsoSSB binding at temperatures closer to the native environment of *S. solfataricus*. 165

Figure 6.3: Cartoons of the four C terminal tails of EcoSSB. 167

Figure 6.4: The end sequence of C terminal tails from the SSBs from various organisms, showing that the acidic residues (D and E) are commonly found at the tip of the disordered tails. 169

Figure 6.5: Left shows how SsoSSB has a 33 amino acid C terminal tail that is long enough to allow the acidic tip to interact with the OB fold. Right shows a model produced by Vina Autodock and MGL Tools. 171

Figure 6.6: The decrease in PIFE was used to track the dissociation of 50 nM unlabelled SsoSSB from 25 nM 12C ssDNA Cy3 as it binds to a 1000 times excess of unlabelled 12C ssDNA. 172

List of tables

Table 2.1: Names and sequences of nucleic acids used. 79

Table 3.1: Dissociation constants from tryptophan quenching experiments with increasing lengths of ssDNA that can respectively allow up to 1, 2 and 4 monomers to bind to them at any one time. Dissociation constants were calculated from a modified one to one binding model. 104

Table 4.1: Table of observed association constants for the first and second SsoSSB monomers and dimers to bind to ssDNA, represented as $k_{on,1}$, $k_{on,2}$ and $k_{on,1.2}$ respectively. 136

Table 4.2: Table of observed dissociation constants for the first and second SsoSSB monomers and dimers to bind to ssDNA, represented as $k_{off,1}$, $k_{off,2}$ and $k_{off,1.2}$ respectively. 136

Table 5.1: The observed rate constants for SsoSSB monomers binding to RNA. .. 155

List of abbreviations

$(SSB)_{35}$	SSB binding mode occluding 35 nt
APS	Ammonium persulfate
ATP	Adenosine triphophate
BER	Base excision repair
bp	Base pair
DBD	DNA binding domain
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DNA	Deoxyribonucleic acid
dsDNA	Double stranded DNA
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EcoSSB	*E. coli* SSB
EDTA	Ethylenediaminetetraacetic acid
EMCCD	Electron multiplying charge coupled device
EPR	Electron paramagnetic resonance
FRET	Förster resonance energy transfer
FWHM	Full width half maximum
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HSQC NMR	Heteronuclear single quantum coherence NMR
HtH	Helix turn helix
ITC	Isothermal caloimetery
L_{12}	Loop connecting beta sheets 1 and 2
LB	Luria-Bertani
MALDI-TOF	Matrix associated laser desorption ionisation time of flight
MMR	DNA mismatch repair

MTSSL	(1-Oxyl-2,2,5,5-tetramethylpyrroline-3-methyl) methanethiosulfonate
NA	Nucleic acid
NER	Nucleotide excision repair
NHS	N-HydroxySuccinimide
NMR	Nuclear magnetic resonance
nt	Nucleotide
OB	Oligonucleotide/oligosaccharide binding
PDB	Protein database
PEG	Polyethylene glycol
PELDOR	Pulsed electron–electron double resonance
PIFE	Protein induced fluroescence enhancement
RecA	DNA recombination and repair protein A
RNA	Ribonucleic acid
RPA	Replication protein A
Rrp	Ribosomal RNA processing protein
S. solfataricus	*Sulfolobus solfataricus*
SAM	Self assembled monolayer
SCOP	Structural classification of proteins
smFRET	single molecule FRET
smTIRF	single molecule TIRF
ss	Single stranded
SSB	ssDNA binding protein
ssDNA	single stranded DNA
ssNA	single stranded NA
SsoSSB	*S. solfataricus* SSB
STM	Site targeted mutagenesis
TBE	TRIS borate EDTA buffer
TBP	TATA binding protein

TEMED	Tetramethylethylenediamine
TFB	Transcription factor B
TIRF	Total internal reflection fluoresence
TRIS	Tris(hydroxymethyl)aminomethane

Declarations:

1. Candidates's declaration:

I, Michael Morten, hereby certify that this thesis, which is approximately 50,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in September 2011 and as a candidate for the degree of PhD in September 2011; the higher study for which this is a record was carried out in the University of St Andrews between 2011 and 2015.

Date:

Michael Morten

2. Supervisor's declaration:

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date:

Prof. Malcolm White

Date:

Dr Carlos Penedo

3. Permission for publication: (to be signed by both candidate and supervisor)

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

PRINTED COPY

No embargo on printed copy.

ELECTRONIC COPY

Embargo on all of electronic copy for a period of 1 year on the following ground(s):

Publication would preclude future publication

Date:

Michael Morten

Date:

Prof. Malcolm White

Date:

Dr Carlos Penedo

Acknowledgements

First and foremost I would like to thank my supervisors Prof. Malcolm White and Dr Carlos Penedo. They have both been incredibly supportive and have given me the opportunity and encouragement to study an interesting and engaging project.

I am extremely grateful to the members of both the MFW and JCP groups for helping me and keeping me going throughout the project. Special mentions go to Dr Jose Peregrina and Biljana Petrovic-Stojanovska as they were the first ones to try and teach me the techniques used in my PhD, however everyone in the labs was always very kind with their time when I needed their help - so thank you!

With respect to the work presented in this thesis, I am very grateful for Dr Bela Bodewho carried out the PELDOR work, Dr Elena Evguenieva-Hackenberg who kindly donated the exosome, Dr Jose Peregrina who made the models of SsoSSB, Dr Catherine Botting and Dr Sally Shirran who carried out the MALDI-TOF experiments and I have also been fortunate enough to benefit from a collaboration on the RNA work with Dr Liza Cubeddu and Dr Roland Gamsjaeger.

I would also like to thank my family and the pals in and out of the lab who made the St Andrews experience so enjoyable, as well as Hannah for being so supportive during the write up, and finally my dad for proof reading the final draft.

Abstract

Single-stranded DNA binding proteins (SSB) bind to single-stranded DNA (ssDNA) that is generated by molecular machines such as helicases and polymerases. SSBs play crucial roles in DNA translation, replication and repair and their importance is demonstrated by their inclusion across all domains of life. The homotetrameric *E. coli* SSB and the heterotrimeric human RPA demonstrate how SSBs can vary structurally, but all fulfil their roles by employing oligonucleotide/oligosaccharide binding (OB) folds. Nucleofilaments of SSB proteins bound to ssDNA sequester the ssDNA strands, and in doing so protect exposed bases, keep the ssDNA in conformations favoured by other proteins that metabolise DNA and also recruit other proteins to bind to ssDNA.

This thesis focuses on the SSB from the archaeon *S. solfataricus* (SsoSSB), and has found SsoSSB to be a monomer that binds cooperatively to ssDNA with a binding site size of 4-5 nucleotides. Tagging ssDNA and SsoSSB with fluorescent labels allowed the real time observation of single molecule interactions during the initial nucleation event and subsequent binding of an adjacent SsoSSB monomer. This was achieved by interpreting fluorescent traces that have recorded combinations of FRET, protein induced fluorescent enhancement (PIFE) and quenching events. This novel analysis gave precise measurements of the dynamics of the first and second monomers binding to ssDNA, which allowed affinity and cooperativity constants to be quantified for this important molecular process.

SsoSSB was also found to have a similar affinity for RNA, demonstrating a promiscuity not found in other SSBs and suggesting further roles for SsoSSB in the cell - possibly exploiting its capacity to protect nucleic acids from degradation. The extreme temperatures that *S. solfataricus* experiences and the strength of the interaction with ssDNA and RNA make exploring the application of SsoSSB for industrial uses an interesting prospect; and its rare monomeric structure provides an opportunity to investigate the action of OB folds in a more isolated environment than in higher order structures.

1. Introduction

1.1. Double-stranded DNA

All cellular life depends on DNA – a macromolecule that is used as a code for an organism's genomic information. Its structure was famously solved by Franklin, Watson and Crick in 1953 which sparked a rapid gain in knowledge about biological systems, and the central dogma of biology of 'DNA makes RNA makes proteins' was soon coined.[1, 2]

The DNA molecule itself is a polymer and usually exists in a conformation that comprises two strands woven around each other creating a double helix. The most common form is B-DNA: a right handed double helix has a pitch of 34 Å and a radius of 10 Å. Each DNA polymer consists of monomer units called nucleotides, which are conjugated together through a sugar phosphate backbone that run around the outside of the double stranded molecule. The covalent bonds that run through one nucleotide to the next pass through the 3' and 5' carbon of the sugar ring, giving a single strand of DNA an asymmetric characteristic and thus directionality. The strands in the double helix run in opposite directions to each other, which allows more efficient packing of the molecule.

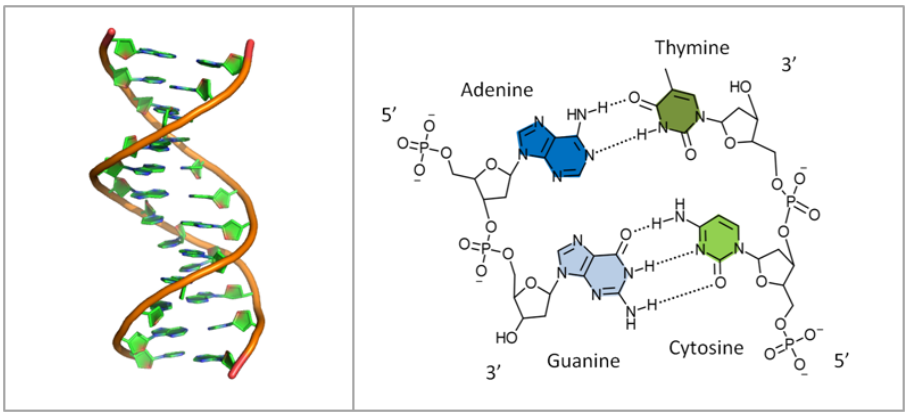


Figure 1.1: Crystal structure of B DNA.
The crystal structure of B DNA is shown on the left (pdb: 1bna) with the cartoon on the right showing the chemical structure of adenine, thymine, cytosine and guanine with hydrogen bonds between complementary base pairs drawn as dashed lines.

In the centre of the two strands lie aromatic bases that project inwards and lie normal to the direction propagated by the backbones. The bases make up a four letter alphabet that code for an organism's genome, and in DNA they are cytosine (C), guanine (G), thymine (T) and adenine (A). These bases may pair through specific hydrogen bonds so that G-C and A-T pairs are formed. The complementary base pairing contributes to holding the two strands together, and facilitates the replication of genetic material – a fact that had not 'escaped the notice' of Watson and Crick in 1953.[1] It since been shown that complementary base pairing of nucleic acids is also integral in transcription, translation and repair processes.

1.2. DNA damage and mutations

The role of DNA in the cell is to faithfully store genomic information, however this information also needs to be accessible, which requires the macromolecule to present the sequence of bases during replication and transcription. This allows mistakes to be

made during replication and gives opportunities for damaging reagents to interact with DNA.

Naturally occurring reagents such as reactive oxygen, nitrogen, or carbonyl species as well as lipid peroxidation products and alkylating reagents can produce abnormal chemical structures of DNA which can disrupt and stall DNA replication and transcription.[3] Hydrolytic processes in the cell cause breakages in DNA strands which can also be fatal for the cell.[4] Mispaired bases, insertions and deletions can arise from faulty DNA replication or spontaneous degradation and can also result in irregular conformations of DNA, where bases can be partially flipped out of the double helix.[5] Exogenous causes of DNA damage include mutagenic chemicals, toxins, thermal disruption and exposure to UV, X-ray and gamma radiation.[3]

A mutation is a heritable alteration to the original nucleotide sequence of a genome - either through an insertion, deletion or substitution of one or more nucleotides. This can occur through spontaneous molecular degradation of DNA; through error prone translesion synthesis where replication proteins that can bypass DNA damage also tend to offer low fidelity as they process along undamaged templates; through faulty DNA repair; and through mutagenic chemicals and UV light.

1.3. DNA repair

Each day there could be up to 1 million events that damage DNA in a single human cell.[6] This still only affects a small fraction of the genome; however the damage to DNA could potentially be fatal if left unchecked. Cells have evolved the ability to

repair DNA in multiple pathways to cope with the variation in types of damage.[7] These pathways can be classed as either direct reversal of the DNA damage or excision repair. Direct reversal of the chemical steps that occurred during DNA damage can be more efficient than excision repair, however it is only possible to repair a few particular types of damage using a direct reversal pathway. Examples of direct reversal pathways include the removal of alkyl adducts from guanine bases and the reversal of pyrimidine dimer formation, by methyl guanine methyl transferase proteins and photolyases respectively.[8, 9] Humans do not possess a photolysase enzyme and rely on excision repair pathways to correct for these and other damaged DNA structures.[8]

During excision repair processes the damage to the DNA is recognised, the damaged DNA is removed and replaced with newly generated DNA. The three types of excision repair are base excision repair (BER), nucleotide excision repair (NER) and mismatch repair (MMR). As their names suggest BER involves the removal of a damaged base, NER includes the removal of a damaged nucleotide and MMR corrects undamaged but mismatched base pairs. The roles of each pathway overlap with each other and as a result they can cooperate to keep the genomic information faithful.

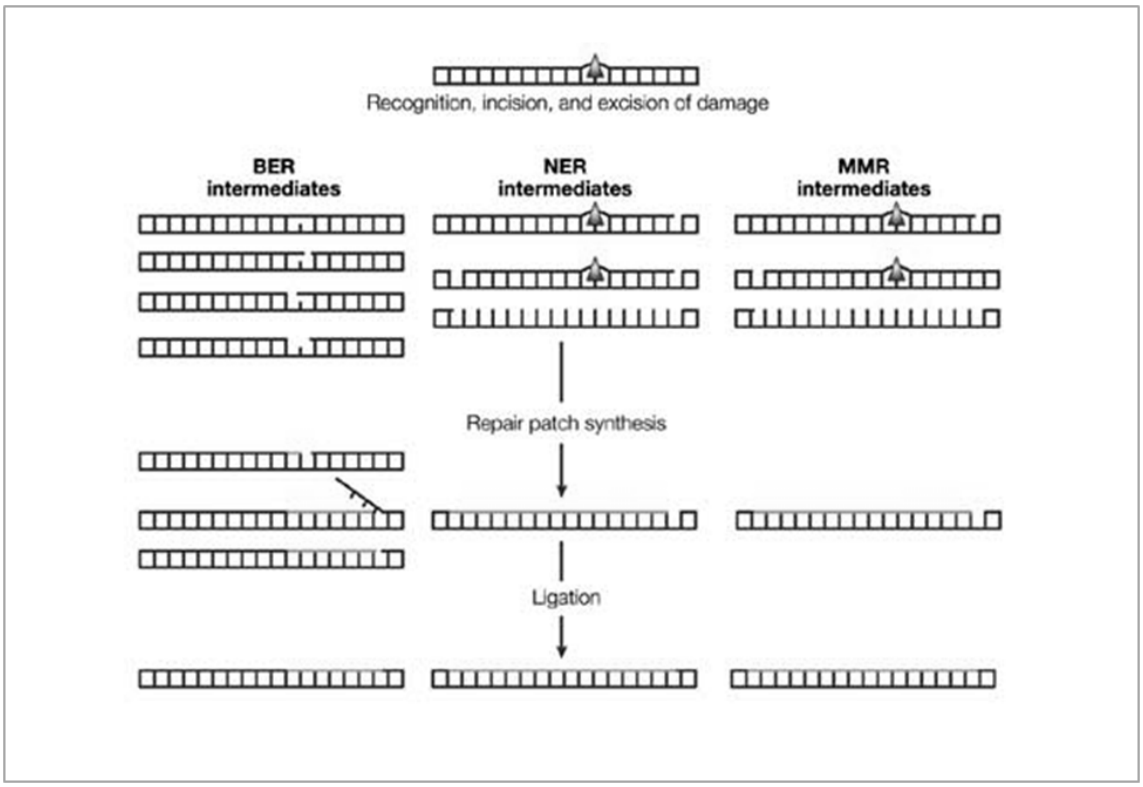


Figure 1.2: A simple diagram showing the pathways and the main intermediates of BER, NER and MMR.
The damage or mismatch (arrow) is first recognised and a repair patch containing the damage is removed from DNA. The non-damaged strand is used as a template to re-synthesise the missing patch of DNA and the repair of DNA is completed by the ligation of the nascent patch to the DNA backbone of the now repaired strand. Figure modified from Cline and Hanawalt.[10]

1.4. Single-stranded nucleic acids

The planar arrangement of the aromatic bases in a dsDNA double helix lends itself to π-π stacking where the overlap of the p orbitals contribute to the stability of the double helix.[11] The arrangement of the sugar-phosphate backbone on the outside of the double helix provides a barrier between the bases and potential damaging agents, which gives a certain amount of protection and ensures the fidelity of the genetic information. For example, hydrolytic deamination of cytosine occurs three orders of magnitude slower in dsDNA than ssDNA as the double helix structure prevents solvent access to the bases.[12] Unfortunately, the arrangement of dsDNA also presents a barrier for proteins that are involved in essential processes that metabolise DNA. In

order to complete these vital processes, the two strands may have to be separated to expose the bases and allow access to the genetic code.

Single-stranded nucleic acid (ssNA) are more flexible and, unlike double stranded DNA, a helix is no longer the lowest energy conformation.[13] A single strand of NA collapses into less defined forms, but still adopts conformations that minimise like charge interactions along the phosphate backbone.[14, 15] As a result, a single strand of NA will generally be more compact in high salt concentrations, allowing the phosphates to exist in closer proximity to each other since ions in solution can shield the electrostatic repulsion along the sugar phosphate backbone. Depending on the sequence of the bases, complementary pairing between bases on the same strand may be possible and this coupled with base stacking and the ions in solution could result in the molecule being able to fold up to give a more defined structure. This allows certain ssNA's to perform cellular functions, as is the case for riboswitches, tRNA and other hairpins and pseudoknots.[16, 17]

ssNA are present in the cell as single-stranded DNA (ssDNA) or RNA. RNA has a similar structure to DNA and it is thought that RNA pre-dated DNA as a genomic material in an earlier RNA/protein world.[18] Each RNA nucleotide also consists of a sugar-base nucleoside conjugated to a monophosphate, however, the sugar has an extra hydroxyl group at the 2' carbon of the ribose sugar. The base thymine is not present in RNA which uses uracil (U) instead. The differences between the two NA strands result in RNA being a shorter, stiffer polymer than its ssDNA analogue due to increased base stacking and sugar pucker effects, as shown in Figure 1.3.[14]

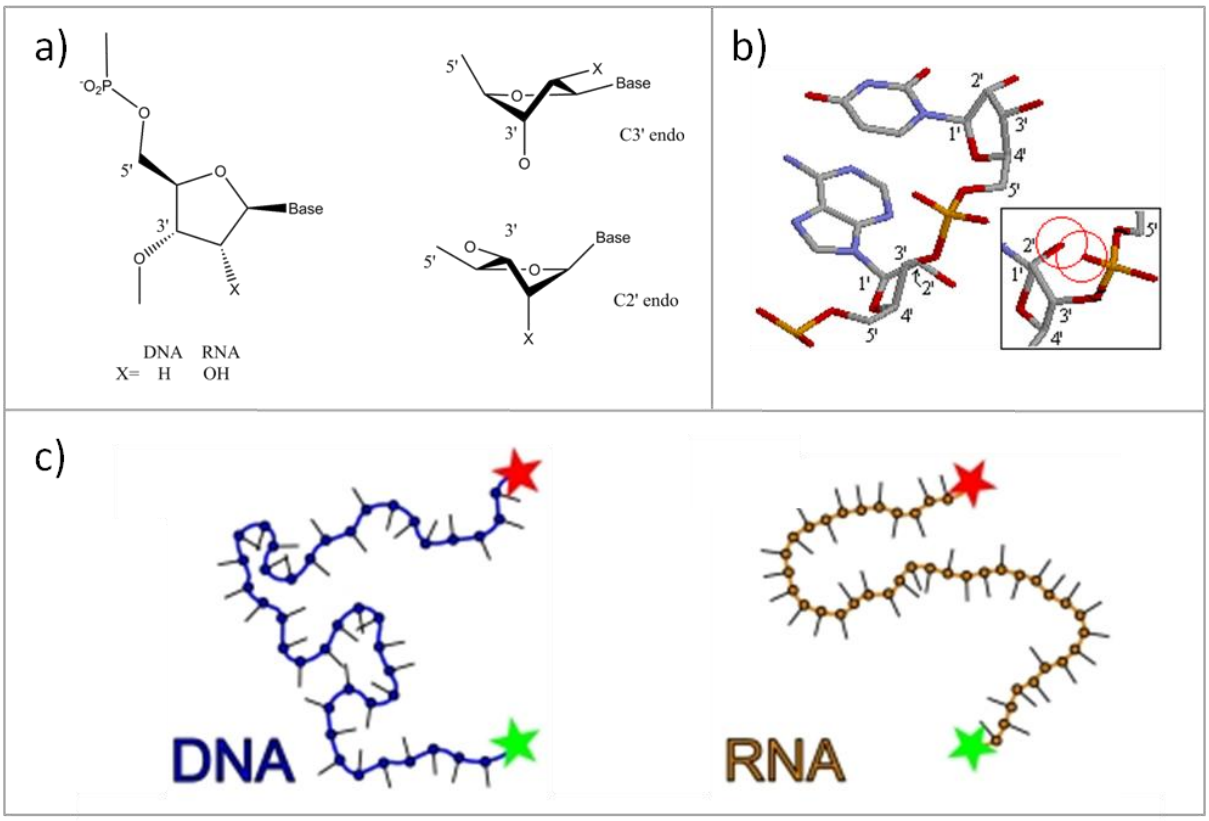


Figure 1.3: The structures of RNA
(a) The difference between ribose and deoxyribose sugars is the 2' hydroxyl group. ssDNA can interconvert relatively easily between the C3' endo and C2' endo conformations. (b) RNA shows more of a preference towards the C3' endo as shown by the 3D structure. The inset shows the clash between the 2' hydroxyl and the phosphate group whilst the RNA adopts the C2' endo conformation. (c) The increased base stacking due to the presence of uracil rather than thymine, and the barrier for the ribose sugar adopting the C2' endo ring pucker results in polymers of RNA being shorter and stiffer than ssDNA analogues. This was observed by Chen et al. who measured the distance between fluorescence dyes (green and red stars) conjugated to the 3' and 5' ends of ssDNA and RNA.[14]

1.5. Single-stranded DNA binding proteins

The propensity of ssDNA to fold and adopt compact conformations could potentially cause fatal stalls of molecule machinery involved in DNA repair, replication and transcription.[19] Breakages in ssDNA occur more frequently than in dsDNA and are extremely disruptive to cellular life.[20] Specialist proteins named single stranded DNA binding proteins (SSB) are expressed to protect the ssDNA,[21] to sequester the strands and to remove any secondary structure of the ssDNA in order to help maintain a

conformation of the polymer that is conducive to other proteins that metabolise or manipulate DNA.[22]

By definition SSBs have a high affinity for ssDNA and they form filaments that decorate the single strands through sequence independent binding,[23, 24] allowing them to provide an essential role in a large number of processes. The high affinity of SSBs for ssDNA rather than dsDNA or RNA, automatically gives the cell a marker for ssDNA if it can somehow detect and exploit SSB's presence on ssDNA.[25, 26] This potentially could enhance the efficacy of essential DNA processes as proteins with lower affinity to ssDNA than SSB as an initial interaction with the SSB, rather than the DNA, could increase the rate of binding and stabilise the protein DNA interaction. SSBs must also be removed readily from ssDNA at some point in DNA metabolism, either to allow other proteins access to the genetic code or simply for the regeneration of the double helix at the end of replication, transcription or repair. There are a great range of SSBs varying in structure and function, from monomers to pentamers.[27-32]

Here a brief overview is presented of some of the common features shared amongst SSBs as well as short overviews on two of the more well known SSBs: the human replication protein A (RPA) and the SSB from *E. coli* (EcoSSB.) These two proteins can both alternate between different ssDNA binding modes, to allow the re-distribution of proteins along a nucleofilament, and also engage with other proteins metabolising ssDNA – however, the different structures of these two SSBs define the

different strategies which each protein employs in order to carry out the same functions.

1.5.1. Oligonucleotide/oligosaccharide binding folds.

SSBs are ubiquitous in all cellular life and there are also examples of viral SSBs,[33] which reflects the vital roles these proteins fulfil.[34] SSB sequence and overall structure vary greatly between organisms but all bind to ssDNA through similar oligonucleotide/oligosaccharide binding (OB) folds.[35, 36] These folds all consist of loops that can interact with the ssDNA as it enters the OB fold, forming hydrogen bonds with the sugar phosphate backbone, anchoring the ssDNA in place. Aromatic residues in the heart of the OB fold act like teeth in a zip, stacking between the bases on the DNA strand, thus cementing the position of the OB fold along the strand.[35]

OB folds are not restricted to SSBs. Evolution has produced OB folds that have been reported to be involved in inorganic pyrophosphatases in yeast, archaea and bacteria to bind to a host of ligands - including RNA in anti-codon binding domains, oligosaccharides in AB_5 bacterial toxins, and proteins in superantigens.[37, 38]. A schematic of an OB fold is shown in Figure 1.4, where the core of the OB fold is made up of five β sheets typically forming a barrel with a Greek key motif, sometimes capped by an α helix. This topology is conserved throughout nucleic acid binding proteins and a total of sixteen Structural Classification of Proteins (SCOP) superfamilies.[39] This has led to the suggestion that OB folds are an example of an ancient topology that is highly capable of withstanding a range of mutations, enabling proteins to evolve to accommodate a different ligand.[37] Mutations in OB

folds can be tolerated in residues that are not integral to the structure of the β sheets, away from the hydrophobic core. There is a wide variety observed in the length and sequence of the hydrophilic loops that link the β sheets together that could account for the exploitation of similar OB fold cores to fulfil a range of different functions. The OB folds from the nucleic acid binding superfamily have been subjected to a level of evolutionary pressure to remain unchanged since before the existence of the last universal common ancestor, before the divergence of bacteria, eukaryotes and archaea which could account for the similarity of these OB folds across all domains of life.[37]

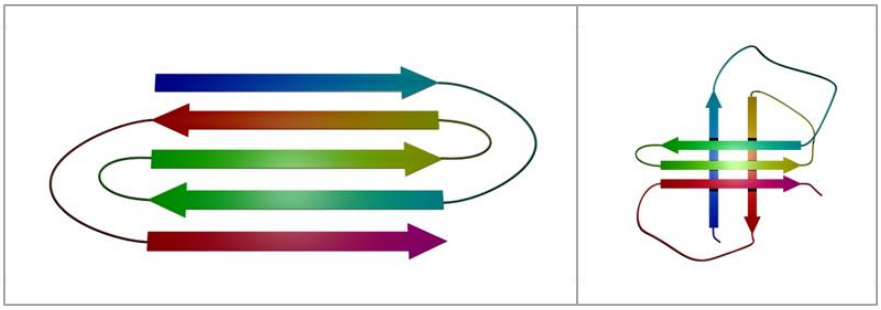


Figure 1.4: Cartoon of an OB fold
Five β sheets in Greek key topology (left) and folded to produce a cartoon example of an OB fold (right).

1.5.2. Loops

The rigid β strands of the OB fold barrel are linked via flexible hydrophilic loops, which do not interfere with the hydrophobic core of the OB fold. Mutations in these residues would not require the rearrangement of a significant number of atomic contacts and are not as crucial during the folding of the protein structure.[40] As a result, a large variation is observed due to adaptations to the loops on the periphery

of the β sheets, where the extension and sequence of the loops can determine which ligand the OB fold may bind and therefore help define the protein's function and superfamily.[36, 41] One example is seen in human RPA, where the L_{12} and L_{45} loops are flexible enough to move closer to the ssDNA in order to form strong interactions, shown in Figure 1.5.[35]

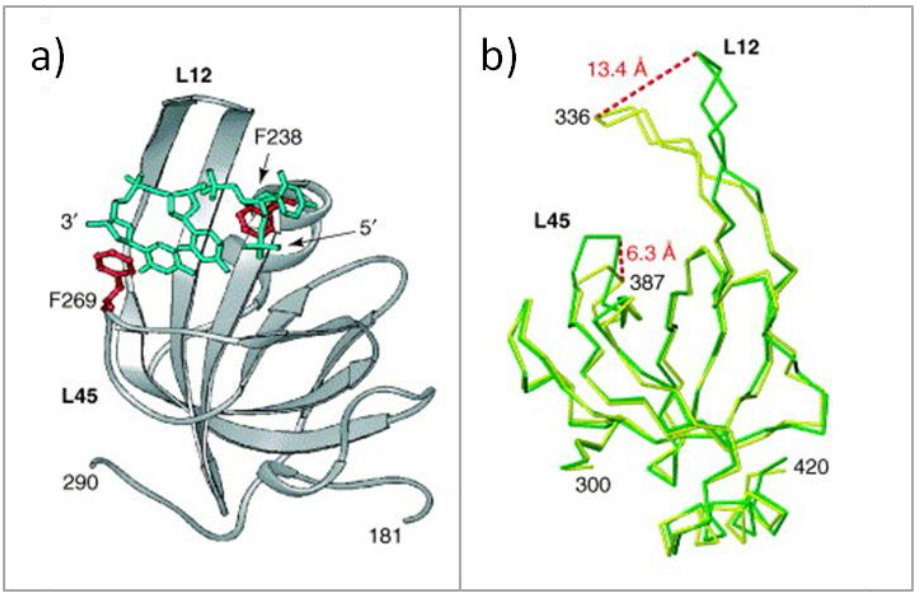


Figure 1.5: The structure of an OB fold from RPA.
(a) Crystal structure of an OB fold from RPA showing the bound ssDNA as a green stick model. Aromatic residues are shown in red that take part in base stacking with the ssDNA. (b) Two Cα traces from crystal structure of RPA OB fold in DNA binding domain A (DBD-A) which is bound (yellow) and unbound (green) to ssDNA. The flexible loops L_{12} and L_{45} change position in order to interact with the ssDNA, and the distance moved by each loop is labelled in red. Figure modified from Bochkarev *et al*.[35]

A variety of loops can be seen in members of the same group of proteins in the same superfamily, and is exemplified by SSBs where loops are explicitly involved in protein-protein interfaces and protein-ssDNA contacts outside the OB fold. As a result, loops between β sheets are heavily implicated in determining the ssDNA binding modes, the multimeric state of the SSB and the processes in which the SSB is able to participate in.[27, 32]

1.5.3. Human replication protein A

A well studied example of a eukaryotic SSB is the *Homo sapiens* replication protein A (RPA). No entire crystal structure of human RPA has been solved, however structure containing the four OB folds`of an analogue RPA from *Ustilago maydis*, bound to 32 nt ssDNA as shown in Figure 1.6(a).[30] Crystal structures of truncated human RPA have been solved, showing the first two OB folds from the bound to ssDNA.[42] The heterotrimer of RPA consists of six OB folds in separate ssDNA binding domains (DBD) spread across the three different subunits, as shown in Figure 1.6(c). Each subunit is connected covalently to the next through flexible linkers, and each is named RPA70, RPA32 and RPA14 to reflect their molecular weight. Four OB folds are distributed around RPA70, one in RPA32 and another in RPA14. There also is a winged helix-turn-helix (HtH) domain in RPA32 that assists in binding to ssDNA but is primarily involved in protein-protein interactions.[43, 44]

RPA subunits bind sequentially to ssDNA in a 5'-3' direction beginning with the two OB folds found in ssDNA binding domain A (DBD-A) and DBD-B respectively, to form the complex shown in Figure 1.6(b).[42, 45, 46] DBD-A binds initially and is tethered to DBD-B through a flexible loop, which results in the binding of DBD-B quickly following DBD-A. NMR data suggest that these two domains tumble independently of each other in solution and a thermodynamic analysis shows that they bind to ssDNA in a non-cooperative fashion.[47, 48] The tethering effectively increases the concentration of ssDNA in DBD-B's local environment and results in an increased rate of binding once DBD-A has bound.[47] RPA then binds ssDNA cooperatively using two other OB-folds, DBD-C and DBD-D, found in the subunits RPA70 and RPA32 respectively, again following a 5'-3' direction. Binding of DBD-

C occurs with a conformational change in the RPA structure where the zinc ribbon in DBD-C stabilises the trimeric core and the interaction with ssDNA. The conformational change also allows DBD-D to bind with ssDNA, with the whole protein occluding 27 nt in the fully bound state. RPA14 and the OB fold found in the N terminus of RPA70 lack two phenylalanines, the aromatic residues found that are conserved in the other OB folds that are used to stack in between the bases of ssDNA, and it is known the RPA14 does not contribute to ssDNA binding, and participates in heterotrimer formation,[49] but there is some controversy surrounding the N terminus of RPA70 and its role in ssDNA binding.[50]

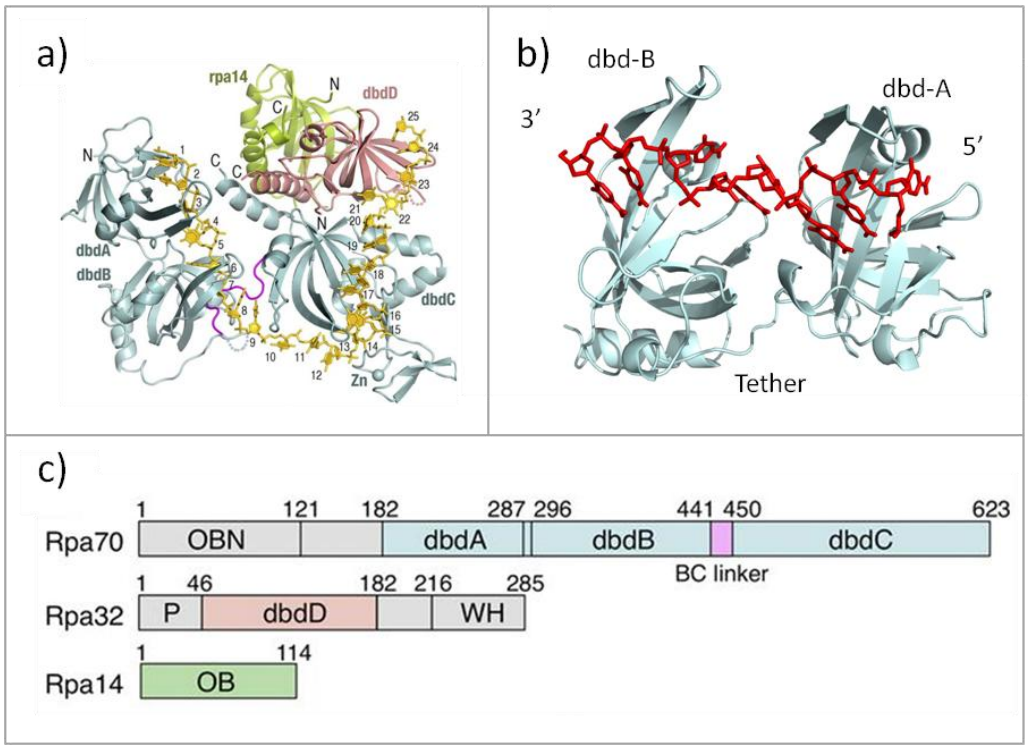


Figure 1.6: The structure and domain organisation of RPA.
(a) The RPA from *Ustilago maydis*, bound to 32 nt ssDNA, showing the U-shaped conformation that the ssDNA is held in as it is bound to the RPA. (b) The crystal structure of the truncated human RPA, showing the two OB folds that are found in DBD-A and DBD-B bound to 8 nt ssDNA that initiates RPA binding. (c) The distribution of ssDNA binding domains and OB folds throughout the three subunits of human RPA. Figures are modified from Fan et al. and pdb: 1jmc.[30, 42]

The different binding domains of RPA exhibit different affinities for ssDNA, with the overall complex binding to ssDNA with a dissociation constant (K_d) that ranges from picomolar to low nanomolar depending on the experimental conditions and sequence and length of ssDNA,[51, 52] which is two to three orders of magnitude lower than DBD-A.[47] The result of the conformational change when binding in the fully bound state can be observed as the protein holds the ssDNA along a U-shaped path through the four OB folds, as shown in Figure 1.6(a).[30] Multiple trimers of RPA are known to only bind to ssDNA with low cooperativity,[52] which could be a result of the high abundance of the protein inside the cell that negates any requirement for the heterotrimers to prefer to bind adjacently to each other.[52] The overall conformation of an RPA-ssDNA nucleofilament is therefore relatively unordered, with kinks of bare ssDNA in between the heterotrimers.[53]

A nucleofilament needs to both protect ssDNA and also allow other parts of the cellular machinery access to the ssDNA. The proteins in the nucleofilament therefore need to be easily redistributed or removed once they interact with other proteins involved in transcription, replication or repair processes. The fully bound RPA is in equilibrium with a partially bound state where only 8-10 nt are in contact with DBD-A and DBD-B, which are the two DBDs with the highest affinity for ssDNA.[42] These two binding modes are shown in Figure 1.7, as well as a model of how they could be exploited during the exchange of RPA heterotrimers. The different affinities of RPA's binding domains allow the protein to bind to ssDNA with different modes and fulfil different roles in the cell, facilitating diffusion along ssDNA, melting dsDNA, and also exchange with other proteins recruited to ssDNA by RPA.[54, 55] This can be

somewhat controlled through phosphorylation of RPA32, which can favour a conformation of RPA that promotes the 8-10 nt binding mode over the fully bound mode. RPA32 can be phosphorylated at a number of sites. The extent of phosphorylation varies depending on the cell's stage in mitosis and also in response to different levels of DNA damage, leading to different affinities for ds and ssDNA.[43]

The winged helix and DBD-F primarily form contacts with other proteins in order to fulfil RPA's role to recruit proteins to ssDNA. The N terminus of RPA70 has been shown to directly interact with proteins involved in DNA damage repair and processes at replication checkpoints including p53, ATRIP, Mre11 and Rad9.[56] RPA's name comes from its role in replication that was identified when it was first purified in 1988.[57] In the initial replication phases, RPA's high affinity for ssDNA and its ability to melt dsDNA is used to initiate and recruit helicase and polymerase alpha activity.[58] During elongation, RPA recruits the polymerases delta and epsilon to DNA through contacts with PCNA, as well as interacting with Dna2 during Okazaki fragment processing.[59, 60] Removal of RPA also allows FEN1 access to cleave the remaining flap left by Dna2.[60]

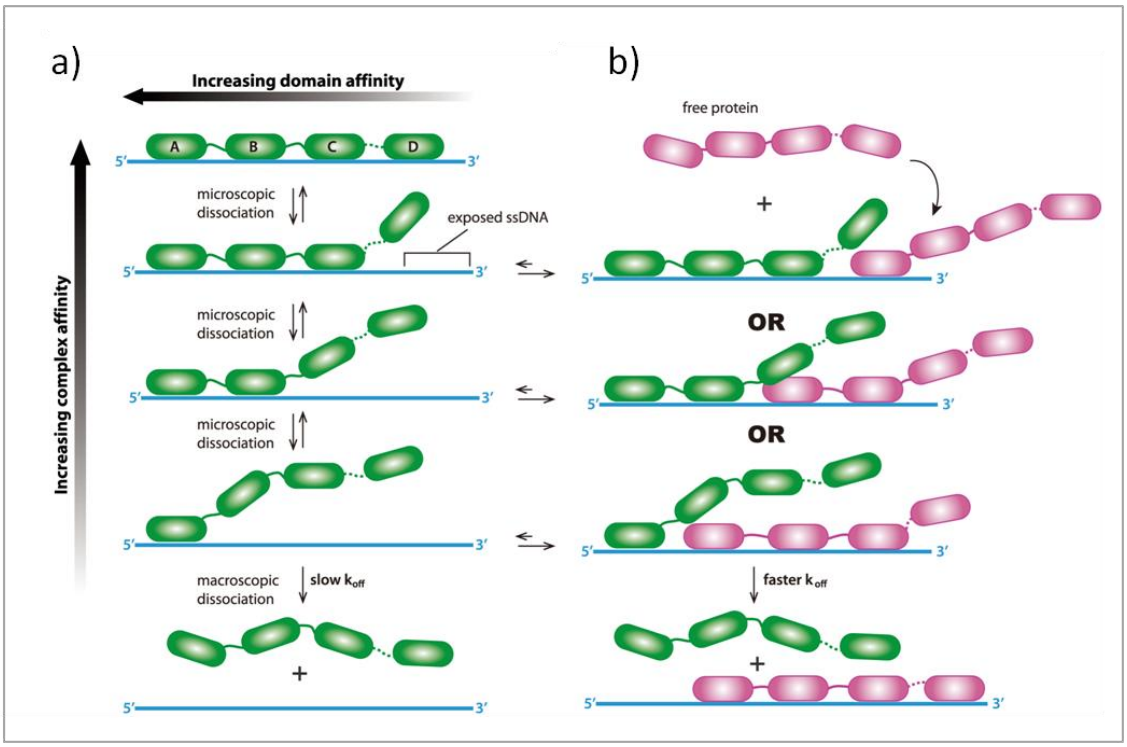


Figure 1.7: A cartoon of RPA binding to ssDNA and exchanging its position with another RPA protein.
The different affinities for ssDNA of the different domains of RPA are exploited during exchange with other RPA heterotrimers and other proteins. The strength of RPA affinity increases with the number of DBDs bound to ssDNA. The affinity of each domain also differs, with DBD-A and DBD-B exhibiting the strongest affinity for ssDNA. RPA is modelled as dissociating from ssDNA as a sequence of microscopic dissociations of individual DBDs, shown on the left by the green protein. Similarly, the mechanism of RPA binding to ssDNA can also be broken down into the microscopic binding of each DBD. The exchange of two RPA proteins (one green, the other purple) can also be viewed in stages of DBDs binding and unbinding. As each green DBD dissociates, this exposes an increasing number of nucleotides available for other proteins to bind to, increasing the strength of interaction between the purple RPA and ssDNA as the affinity for ssDNA of the green RPA is weakened. Figure modified from Gibb *et al*.[61]

RPA also has roles in DNA repair, contributing towards NER, BER, MMR and double strand break repair. For example, RPA displays a weaker affinity towards damaged DNA, therefore during NER RPA preferentially binds to the undamaged strand, protecting it from endonucleases (Figure 1.8).[62, 63] The binding polarity, coupled with the weaker affinity for damaged sites, helps to position the RPA to the 5’ side of the damaged strand which allows RPA to specifically recruit the nucleases XPG to the 3’ end and also XPA and XPF-ERCC1 to the 5’ end site of damage.[46, 63,

[64] Following the removal of the damaged section of DNA, RPA is poised to recruit polymerases involved in DNA replication similar to the processes described above.[65] The phosphorylation of RPA32 in response to DNA damage decreases RPA's affinity for ssDNA, which prevents RPA melting unnecessary lengths of dsDNA, regulating and containing the nucleofilament structure around the DNA lesion. The weakened binding also facilitates the removal of RPA by DNA polymerases at the end of the repair process.[66]

RPA's role in BER is less understood; however, its involvement is certainly critical since substitution of mutants for wild-type RPA increases sensitivity to methyl methane sulfonate, a lesion repaired through BER.[67] Interactions between RPA and DNA glycosylases, its capacity to enhance primer extension, unwind the downstream strand at the 5' end, and stimulate DNA ligase I all reinforce the importance of RPA's contributions to BER.[68, 69]

During MMR, RPA protects the template strand and has a role in stimulating and regulating the nuclease activity that removes the mismatched base.[70] The phosphorylation of RPA32 leads to a decrease in RPA's affinity for ssDNA, inducing a conformation change where RPA32 interacts with DBD-F, moving DBD-C and D away from the ssDNA.[43, 71] Phosphorylation occurs before the re-annealing of DNA and facilitates the removal of RPA by polymerase delta similar to NER.[66]

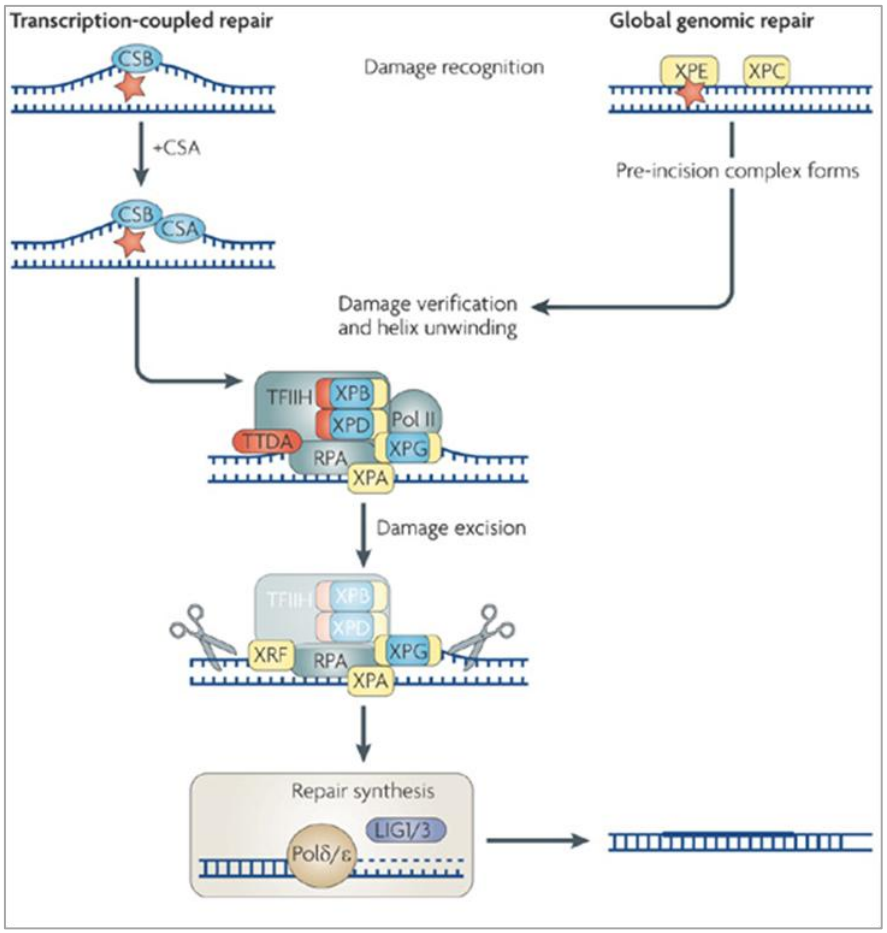


Figure 1.8: A schematic of the main steps involved NER.
A repair response is triggered by the stalling of a transcription complex when it encounters damage or through global genomic repair (GGR) when a lesion is present on non-transcribed DNA. In transcription-coupled repair (TCR), the damage is encountered by the RNA polymerase II complex and the initiation of NER is carried out by CSB (also known as ERCC6). In GGR the lesion is detected by xeroderma pigmentosum complementation group E (XPE, also known as DDB1) or by XPC. RPA or XPA are involved in pre-incision events that precede the unwinding of a repair bubble by transcription factor II human (TFIIH). XPF and XPG are recruited to either side of the damage and incisions are made sequentially, initially by XPG at the 3' site of the damage followed by XPF at the 5' site. The section of nucleotides containing the lesion is then removed, with RPA able to assist in recruiting DNA polymerase (Polδ) and proliferating cell nuclear antigen, PCNA, with Polε to fill the gap with undamaged nucleotides. DNA ligase I (LIG1) is then able to complete the repair process by conjugating the new nucleotides to the original DNA strand. Figure modified from McKinnon.[72]

After 3' overhangs are generated in response to a double strand break, RPA binds to the ssDNA to protect and to prevent the formation of any unwanted secondary structures.[73] RPA recruits RAD52 through an interaction between the C terminus of RPA32, stabilising the nucleoprotein complex, possibly through RAD52 interacting with DBD-A and B, and ensuring stoichiometric interactions.[69, 74] The interaction

with RAD52, coupled with the phosphorylation of RPA32 and subsequent change in RPA conformation and resulting shift to the 8-10 bt binding mode, aids the exchange between RAD51 and RPA on ssDNA.[55, 75]

1.5.4. Bacterial SSB

E. coli SSB (EcoSSB) is arguably the most intensively studied SSB with many extensive reviews detailing its structure, functions and roles.[76] It displays a well characterised homotetramer structure with each 19 kDa monomer consisting of a single OB fold, surrounded by extended loops that support contacts between the monomers and other tetramers in addition to guiding the ssDNA around the tetramer.[32, 77, 78] Each monomer also has an acidic C terminal tail that has been shown to be involved in protein recruitment, and also has a possible role in binding to ssDNA although there are conflicting reports to the extent of its contribution.[78-81] The homotetramer structure is shown in Figure 1.9 which also displays the four intrinsically disordered C terminal tails and the sequence of the acidic tips.

Similar to RPA, the multiple sites EcoSSB employs to bind to ssDNA allow the diffusion and re-distribution of tetramers in a nucleofilament. Coupled with many possible protein-protein interactions, this allows EcoSSB to assist in many different roles in the cell. EcoSSB can wrap ssDNA around all four of its monomers, using multiple points along the surface of EcoSSB both in and around the OB folds.[32] These contacts provide both electrostatic and base stacking interactions, contributing to a strong overall affinity towards the ssDNA. EcoSSB has multiple binding modes that are dependent on monovalent salt concentration, pH, divalent and multivalent

cation concentrations, temperature and protein concentrations, with the $(SSB)_{35}$ and $(SSB)_{65}$ modes being the two most prevalent, which are shown in Figure 1.10.[82, 83] The number of nucleotides occluded is dependent on whether two or four monomers are involved in binding, with each respective mode binding to 35 and 65 nt, thus giving the different modes their name. The $(SSB)_{35}$ mode is favoured under low salt conditions and high EcoSSB concentrations, showing highly cooperative binding between tetramers along a strand of ssDNA.[84] The $(SSB)_{65}$ mode has a limited cooperativity producing beads consisting of two tetramers bound on ssDNA, resulting from low protein concentrations and high salt concentrations.[32]

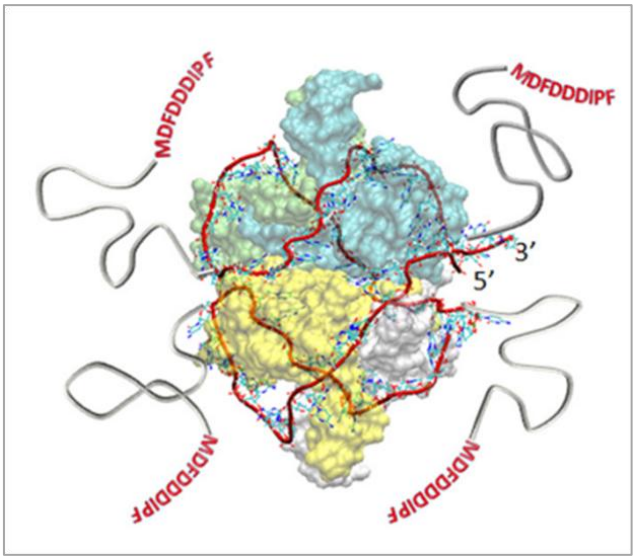


Figure 1.9: A cartoon showing the tetrameric structure of EcoSSB bound to ssDNA.
The four monomers are coloured blue, green, yellow and white with the four unstructured C terminal tail modelled as grey ribbons and the acidic C terminal residues are shown in red. The ssDNA in the $(SSB)_{65}$ binding mode is represented as a red ribbon wrapped around the tetramer. Figure adapted from Kozlov *et al*.[85]

The interaction of SSB with other proteins as well as ssDNA is clearly advantageous, yet modifications to an OB fold to achieve this must be done without destabilising the hydrophobic core that is occupied with providing a strong interaction with the

ligand. Loops that act as linkers between β sheets are more likely to be involved in ssDNA contacts and with neighbouring SSB monomers in the SSB filament.

Many SSBs have developed an intrinsically disordered C terminal tail that is characteristically acidic towards its extreme end, including EcoSSB which has one unstructured tail per monomer shown in Figure 1.9.[31, 81, 86, 87] This hydrophilic tail is thought to be able to recruit other proteins to ssDNA by protruding out from the SSB as it is bound to ssDNA, which provides a surface for an electrostatic interaction with positively charged areas of other proteins that are involved in DNA processes, such as transcription, replication and DNA repair.[26, 29, 86, 88] For example, EcoSSB interacts directly with the clamp loader within DNA polymerase III holoenzyme (Pol III HE), which assists clamp loading, aids processivity and allows the efficient removal of tetramers from ssDNA that could potentially upset the polymerases efficiency during DNA replication.[89] EcoSSB's association with *E. coli* primase strengthens the primase's interaction with the nascent RNA primer.[90] Dissociation of EcoSSB from the primase also destabilises the primase's hold on ssDNA, and allows the clamp assembly to occur. During DNA recombination, EcoSSB stimulates RecQ helicase activity via interactions with the EcoSSB C terminal tail.[91] EcoSSB also has a role in stabilising the binding and promoting the activities of the exonuclease RecJ,[92] and the RecG helicase,[93] as well as mediating the formation of RecA filaments through interactions with RecO.[94]

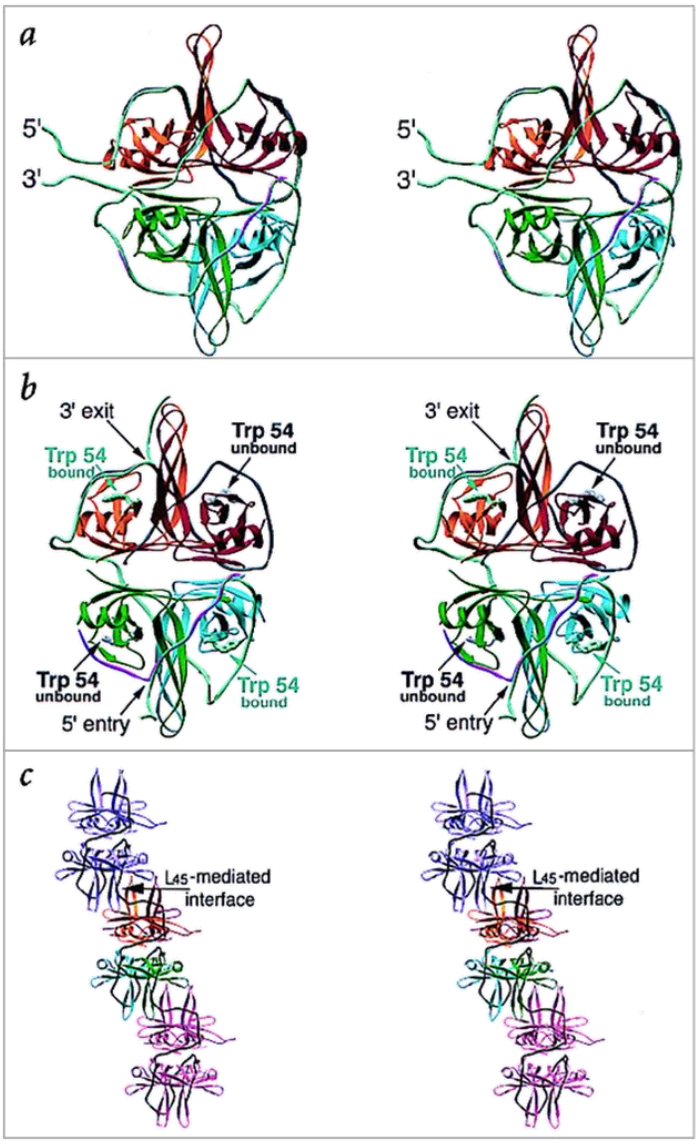


Figure 1.10: A cartoon showing two of EcoSSBs binding modes.
EcoSSB is shown binding to ssDNA where (a) wraps 65 nt ssDNA around all four monomers and (b) where only 35 nt ssDNA is wrapped around two EcoSSB monomers. (c) The fully cooperative binding of three tetramers (the middle tetramer has its four monomers colour coded to match the tetramers in (a) and (b)) to ssDNA produced by $(SSB)_{35}$, with strong interactions between proteins assisted by the L_{45} loop. Figure modified from Raghunrathan *et al.*[32]

In DNA repair processes, EcoSSB recruits the exonuclease *E. coli* ExoI during MMR,[95] and interacts with Uracil DNA glycosylase during BER.[96] DNA polymerase II (pol II) is involved in a variety of responses to damaged DNA and requires EcoSSB to process efficiently along DNA and to stimulate pol II-associated nuclease activity.[97] Pol II was the first protein to be identified that interacted with EcoSSB,

being co-purified in 1972.[26] Their interaction is also shown by the formation of a pol II/SSB complex in the absence of ssDNA.[22, 97]

1.6. Archaea

The domains of life were historically classed as simply either plants or animals; however this grouping could not take into account fungi, protists or bacteria. A new classification attempt in the mid-20th century split five kingdoms into two domains distinguishing between prokaryotic and eukaryotic organisms. This was again revised following an analysis of ribosomal RNA which led to the discovery of a third domain in 1977 by Woese and Fox.[98] At the time they named this new domain the archaeabacteria, which was eventually shortened to archaea to further distinguish the domain from the bacteria.

Archaea are unicellular organisms that lack a nucleus but typically possess a singular circular chromosome, similar to bacteria. Common characteristics of eukaryotes can also be found when comparing their molecular machinery, where protein complexes involved in processes such as transcription, DNA replication and translation share the same fundamental features.[99] Yet a clear genetic distinction of archaea from both other domains can be seen through their 16S (18S) rRNA, which is present throughout all types of self-replication systems and changes slowly over time so that any link between distant species can be made.

The archaea was split further into Crenarchaeota and Euryarchaeaota, and additional phyla were distinguished as the Korarchaeota, Thaumarchaeota and Aigarchaeota, as

shown in Figure 1.11.[100] It has also been suggested that the Nanoarchaea could be classed as a separate phylum, however this has been countered with the proposal that they are a rapidly evolving branch of the Euryarchaeaota.[100] The Euryarchaeatoa consists of the methanogens and their relatives, namely extreme halophiles, sulfate-reducing species and thermophiles and also methangenic species, whereas the Crenarchaeaota include (hyper)thermophiles, thermoacidophiles, eocytes and sulfur-dependent archaea.[101] The capacity of many archaea to survive in extreme temperatures suggests their last common ancestor also was a thermophile; and it is their tolerance for harsh conditions which makes them a desirable and viable alternative to eukaryotes for use in research.

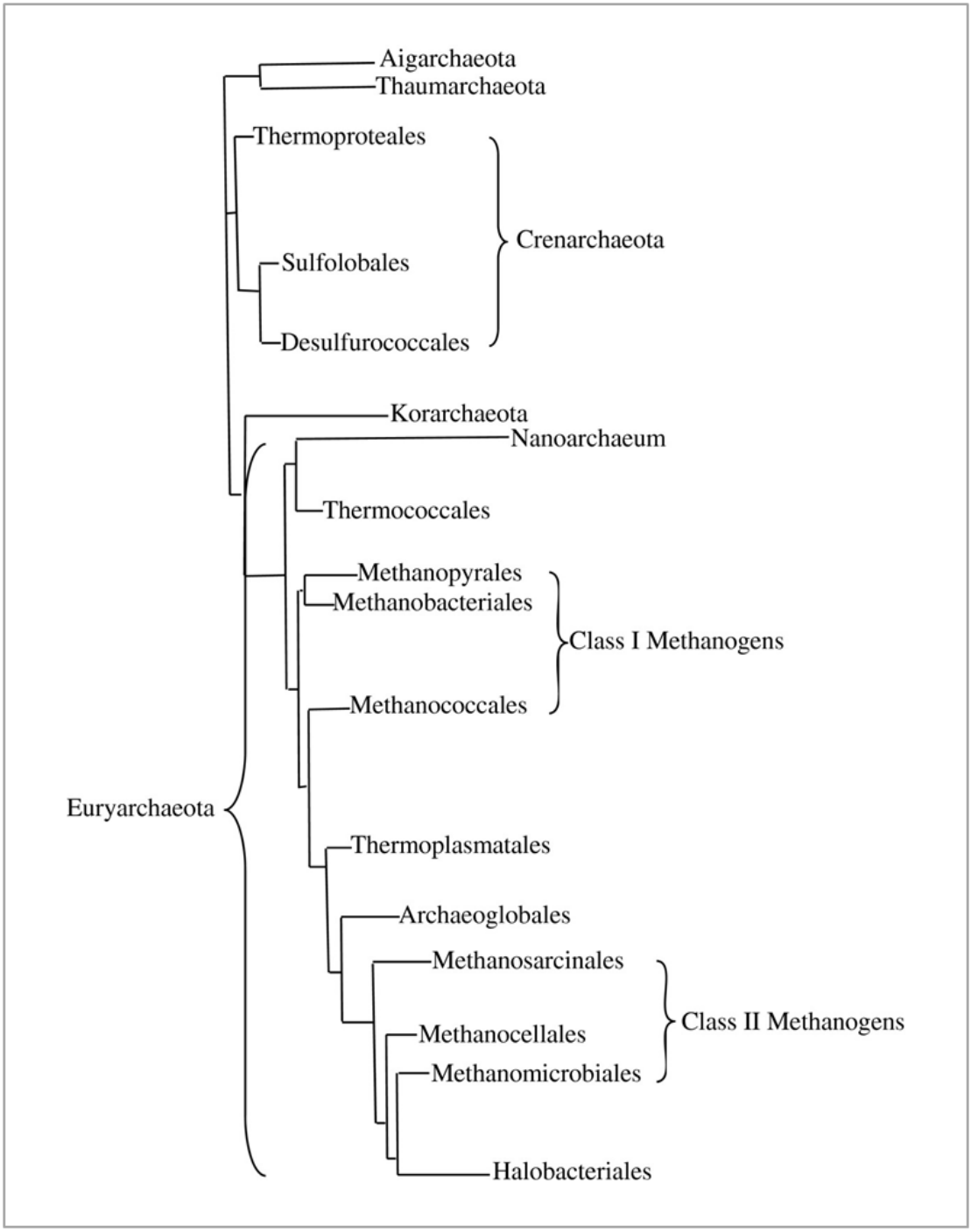


Figure 1.11: Phylogenetic tree showing the distinction between the 16SRNA sequences of the archaea.
This- shows the established phyla (−ota), major orders (−ales), and two classes of methanogens. Figure modified from Godde.[100]

1.6.1. Archaeal SSB

SSBs from archaea show a large variation in structure and sequence, displaying common features with both eukaryote and prokaryote SSBs, typically employing OB folds. The euryarchaeon *Methanosarcina acetivorans* has three SSBs that all show similarities to human RPA, and each has the capacity to form homomultimeric complexes.[102] These three proteins alone exhibit different binding modes and differ in their molecular weight, multimeric state and the number of OB folds and zinc fingers. Other methanogenic species that exhibit SSBs again show a range of different ratios of OB folds and zinc fingers, including *Methanopyrus kandleri*,[103] *Methanocaldococcus jannaschii*,[104] and *Methanothermobacter thermautotrophicus*.[103] Proteins from non-methanogenic species of euryarchaea also show similar zinc fingers and OB folds, including examples from *Haloferax volcanii*,[105] *Ferroplasma acidarmanus*,[106] *Thermoplasma volcanium*,[107] and *Pyrococcus furiosus*.[108] Examples of SSBs from crenarchaea are not present or unknown, the most well studied being from *Sulfolobus solfataricus* and a biophysical analysis of this protein is presented in chapter 3 of this thesis.

1.6.1.1. *Sulfolobus solfataricus* SSB

The SSB from *Sulfolobus solfataricus* (SsoSSB) was identified by Wadsworth and White, after purification from a cell extract.[109] The mass of a monomer was measured as 16.184 kDa, and in gel filtration of the recombinant protein was eluted with an estimated size of 20 ± 4 kDa. The K_d for ssDNA-monomer complex was shown to be in the low nanomolar region, and fluorescent titrations saturated ssDNA

strands with SSB monomers at a ratio of a single SSB monomer to approximately 4-5 nt of ssDNA, which was independent of the length of the ssDNA.[109]

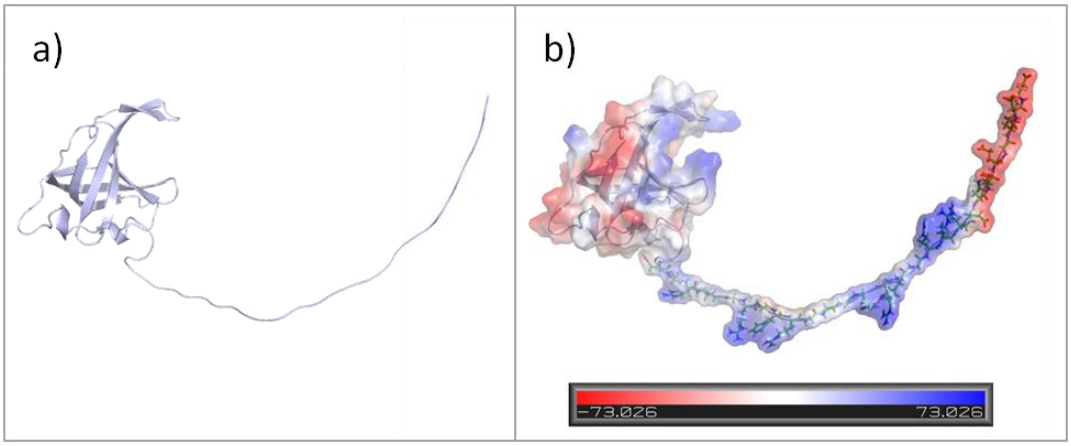


Figure 1.12: The crystal structure of a SsoSSB monomer and the relative electrostatic charges on its surface.
(a) Cartoon of SsoSSB monomer with unstructured C terminal tail modelled on to the crystal structure. (b) Electrostatic surface generated on top of the structure shown in (a), clearly showing the acidic nature of the residues at the C terminus. (pdb: 1o7i)

The crystal structure of a truncated SsoSSB was solved as a monomer and is shown in Figure 1.12(a). Crystals were produced from a construct where the highly flexible C-terminal tail was removed by proteolytic digestion, at residue P115.[27] This is consistent with previous attempts to crystallise EcoSSB where the acidic terminal tail had to be removed before the first structures were solved.[110] A subsequent structure of the full length *E.coli* protein could not provide any structure for the tail, confirming its disordered nature.[81] Previous work in the White lab has shown that SsoSSB is expected to exist primarily as a monomer and multi-angle laser light scattering experiments completed by Gamsjaeger *et al*. agree with SsoSSB adopting a monomeric arrangement in solution.[109] However Haseltine *et al.* also characterised this protein and reported a mixture of dimeric, tetrameric and monomeric species present during gel filtration.[111] Further electrophoresis analysis by Haiyan Shi *et al.*

also suggested that this SSB is a tetramer, so clearly there is some debate in the literature as to what multimeric form this protein takes.[112]

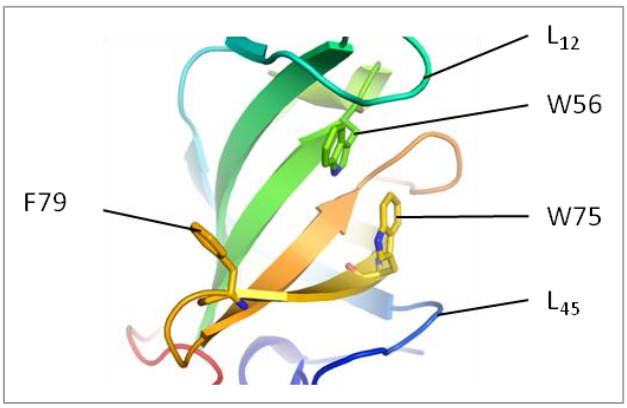


Figure 1.13: Part of the OB fold from SsoSSB.
An image of part of the crystal structure of SsoSSB showing the three aromatic residues (W56, W75 and F79) that can stack in between the bases of ssDNA. The loops that interact with the sugar phosphate backbone of ssDNA are also labelled as L_{12} and L_{45}.

The crystal structure of SsoSSB shows the presence of a single OB fold, producing a hydrophobic core for the binding of a single strand of DNA. Inside the core lie three aromatic residues (W56, W75 and F79) which are involved in stacking between bases, providing a major contribution to ssDNA binding, and are labelled in Figure 1.13. W75 appears to be both involved with base stacking and interacting with the phosphate backbone.[27, 113] SsoSSB has also been shown by Gamsjaeger *et al.* to bind to ssDNA with a defined polarity identical to RPA.[114] In the case of RPA, this plays a crucial role in directing molecular traffic along ssDNA and SsoSSB could potentially exploit this directionality in a similar vein.

The loop linking the β strands 1 and 2 (L_{12}) and between 4 and 5 (L_{45}) are expected to close around ssDNA as it enters the OB fold, which is typical of other OB folds.[27] L_{23} and L_{45} are not as extended in SsoSSB compared to EcoSSB, which suggests that

protein-protein interactions are weaker and supports the idea SsoSSB is not as capable of forming multimeric states which exploit the loops to strengthen the interaction between monomers and guide the ssDNA around sets of monomers.[27]

The C terminal tail of SsoSSB has been shown to interact with RNA polymerase, showing that SSB has the capacity to recruit RNA polymerase and/or initiate transcription.[86] SSB has been also shown to be able to substitute for the TATA binding protein (TBP) and assist transcription factor II B (TFB) dependent transcription.[86] Proteolytic removal of the tail does not prevent SSB binding to ssDNA but stops any interaction with other proteins and the deletion of the tail is expected to cause severe repair-deficient phenotypes, similar to those seen in EcoSSB.[109]

The level of SsoSSB expression *in vivo* only increases by a modest amount after exposure to UV radiation, but has been shown to be able to selectively unwind damaged dsDNA *in vitro* making SsoSSB a candidate for signalling DNA damage.[88, 115]

It seems that SsoSSB is arguably one of the simplest SSBs structurally and yet it contains all the necessary ingredients that an SSB requires to fulfil its role. A detailed analysis of this stripped down model could provide a benchmark for answers to be extrapolated into how other SSBs containing multiple OB folds bind to ssDNA and facilitate essential cellular processes. It is clear that a nucleofilament must be a highly dynamic environment, and that sensitive experiments are required in order to

capture the subtleties of how SsoSSB persists on ssDNA. Fluorescence microscopy is capable of resolving single molecules and could provide the necessary level of detail to be able to reach reliable conclusions about how SsoSSB and ssDNA interact, and it is with these objectives in mind that single molecule techniques are discussed for the remainder of the introduction.

1.7. Single molecule assays

Achieving single molecule resolution allows the observer to simply look at the structure and movement of individual molecules, which allows the observation of temporal and spatial details that are otherwise lost in ensemble averaging. Difficult ensemble experiments studying dynamic processes are made much easier as synchronisation is not required when studying each molecule individually. Many single molecule experiments can be completed quickly with ease producing large amounts of data that can be used to construct frequency histograms, which are effectively probability distribution functions. The shape and position of these histograms not only give the 'average' result but can also reveal identify transient states and heterogeneous behaviours that are difficult to observe with conventional ensemble techniques.[116]

'It is very easy to answer many of these fundamental biological questions; you just look at the thing!'

Richard Feynman, 'There's plenty of room at the bottom,'1959[117]

Feynman was famously referring to electron microscopy in the quote above, but there are a number of other single molecule techniques used to study biological systems. A non-exhaustive list includes manipulation techniques (eg. atomic force microscopy,[118] scanning tunnelling microscopy,[119] optical tweezers[120] and magnetic tweezers[121]), nanopore technology, and spectroscopic techniques such as surface enhanced Raman spectroscopy[122] and fluorescence microscopy, which is the primary technique used in this thesis.

The resolution of an optical microscope is usually limited to the diffraction limit of the wavelength of light used to illuminate the sample, typically around 200 nm. This is defined by the Abbe diffraction limit, which is given in equation 1.1 and describes the diameter of a spot observed by microscope (d) in terms of the wavelength of light (λ) that is travelling through a medium with a refractive index (n) and converges at a point at an angle (θ). The numerical aperture of a microscope is further defined as $NA = 2nsin\theta$.

$$d = \frac{\lambda}{2nsin\theta} = \frac{\lambda}{2(NA)} \tag{1.1}$$

Observing two points that are separated by shorter distances than the diffraction limit requires some form of super-resolution microscopy. It is essential for these types of microscopy that individual fluorophores are excited sequentially so that the point spread function of each emitter does not overlap with another. Once one fluorophore has bleached or deactivated, another may be excited and the process repeated. This type of technique has been exploited by a number of different types of super-resolution microscopies, including PALM,[123] STORM,[124] STED[125] and PAINT.[126] Super-resolution techniques have the potential to resolve features that are separated by approximately less than 20 nm,[127] however they sometimes require hours of exposure to achieve this scale of resolution, and is therefore problematic when trying to study interactions on a millisecond time scale.[128]

Other techniques that use the intensities of fluorophores to directly report on their microenvironment, are used to deduce structural changes below the diffraction limit. Fluorescence is a very dynamic process, with some fluorophores emitting tens of

thousands of photons per second.[129] Emission of radiation is one of many possible processes that an excited fluorophore can go through. The probability that a molecule does emit a photon is very sensitive to subtle changes in the local environment. For example, changes in temperature, the surrounding solution's viscosity and transfers of energy to nearby molecules all vary the rate of fluorescence from a fluorophore.[130] This can be exploited to provide other methods to resolve structural details below the diffraction limit, and track any changes in real time on a millisecond time scale.

1.8. Fluorescence

According to atomic and molecular orbital theory, electrons can exist in distinct orbitals that surround the nuclei of the atom(s).[131] The energy of an electron depends on the shape of the orbital and its proximity to a nucleus. The lowest energy electronic arrangement of a molecule or atom is named the ground state, but excited states are available where electrons occupy higher energy orbitals.

To enter an excited state, an electronic transfer can be caused through the absorption of a photon that has the same energy as the difference in energy between the two energy levels involved in the transition. The probabilities of these transitions are subjected to the Franck-Codon principle, which is described briefly as follows.[131] The ground and excited states of a molecule can be written as initial and final wavefunctions respectively (ψ_i and ψ_f), and the transition dipole moment (μ_{fi}) can be modelled where the initial wavefunction is perturbed by the transition dipole moment operator ($\hat{\mu}$). The Franck-Condon principle follows the Born-Oppenheimer approximation, which states that electrons move much faster than nuclei.[131] The electronic transition can be viewed as a vertical transition without movement of the

nuclei during the transition in Figure 1.14, which shows an energy diagram of different vibrational and electronic states with respect to the coordinates of the nuclei. The overall wavefunctions can therefore be represented as products of the vibrational, electronic and spin wavefunctions and the electronic shown in equation 1.3 and transition dipole moment operator can also be approximated by splitting it into electronic and vibrational parts, shown in equations 1.4.

$$\mu_{fi} = \langle \psi_f | \hat{\mu} | \psi_i \rangle \tag{1.2}$$

where:

$$\psi_i = \psi_{i,e} \psi_{i,v} \psi_{i,s} \tag{1.3}$$

$$\hat{\mu} = \hat{\mu}_e + \hat{\mu}_N = -e \sum_i \hat{r}_i + e \sum_I Z_I \hat{R}_I \tag{1.4}$$

therefore:

$$\mu_{fi} = -e \sum_i \langle \psi_{f,e} | \hat{r}_i | \psi_{i,e} \rangle \langle \psi_{f,v} | \psi_{i,v} \rangle \langle \psi_{f,s} | \psi_{i,s} \rangle + e \sum_i Z_I \langle \psi_{f,e} | \psi_{i,e} \rangle \langle \psi_{f,v} | \widehat{R_I} | \psi_{i,v} \rangle \langle \psi_{f,s} | \psi_{i,s} \rangle \tag{1.5}$$

As a result of the Born-Oppenheimer approximation, the transition probability can be written as shown in equation 1.5, where the zero term arises at the start of the second term of three integrals because different electronic states are orthogonal to each other ($\langle \psi_{f,e} | \psi_{i,e} \rangle = 0$). This leaves the first term, whose magnitude is governed by the vibrational overlap integral called the Franck-Condon factor and the orbital and spin selection rules. As a result, an electronic transition is more likely to occur without a

change in the spin state ($\Delta S = 0$) between states that have a large vibrational overlap that are not symmetrically equivalent ($\Delta l = \pm 1$).[131]

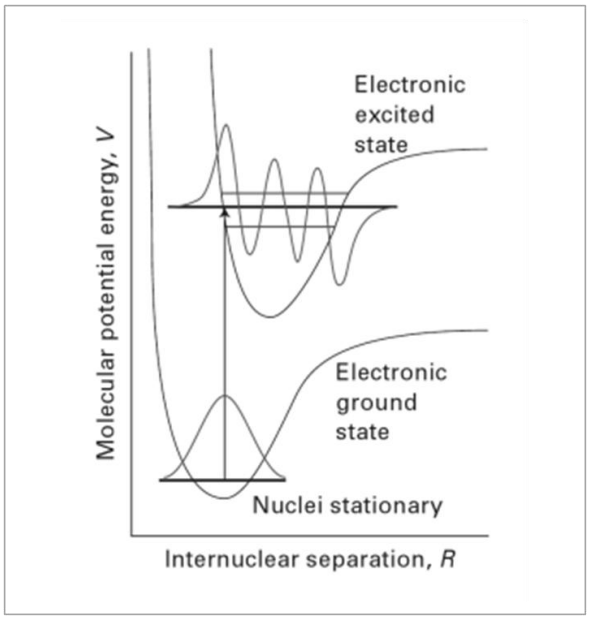


Figure 1.14: A plot of the molecular potential energy against the internuclear separation during an electronic transfer to an excited state.
The Born-Oppenhiemer approximation assumes an electronic transition occurs at speeds fast enough to neglect the movement of the nuclei involved, therefore is represented above by a vertical transition from the ground state to an excited state. The Franck Condon factor describes how transitions with the most overlap between the integral of the respective vibrational wavefinctions are more likely to occur, shown above. Figure modified from Atkins' Physical Chemistry.[131]

Once a molecule enters an excited state there are number of processes that could happen. It is possible to absorb another photon and be excited to a further higher electronic state, or go through processes that do not change the energy of a system such as intersystem crossing and internal conversion.[132] Alternatively the molecule could relax back to the ground state, by finding a way to transfer its excess energy. This relaxation can be done in a number of ways which include collisions with other molecules, vibrational relaxations, chemical reactions, phosphorescence and fluorescence. All these different vertical and horizontal pathways compete with one

another to be the most efficient and they can be described by a Jablonski plot, such as the one shown in Figure 1.15.

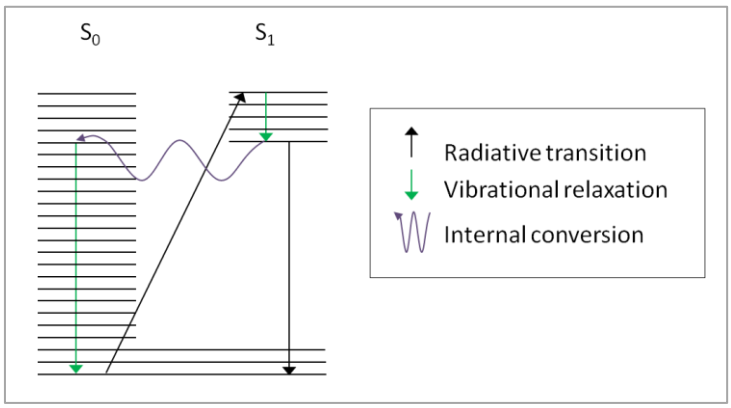


Figure 1.15: A simple Jablonksi plot.
The absorption of a photon is followed by vibrational relaxation to the lowest vibrational energy level of the excited singlet state. This in turn is then either followed by the emission of radiation in the form of fluorescence or by internal conversion back to the ground singlet state which relaxes to the ground vibrational state non-radiatively. More complex Jablonksi plots could also include transitions to and from states of higher multiplicities including intersystem crossing and phosphorescence, as well as the absorption and emission of multiple photons.

Fluorescence is the emission of radiation during the transition of a system to a lower energy electronic state of the same multiplicity as the higher energy state.[132] The fluorescence lifetime, τ_f, of a molecule is the average time the molecule spends in its excited state before emitting a photon, and in the absence of any other quenching processes is defined as:

$$\tau_f = \frac{1}{k_f} \tag{1.6}$$

The lifetime can be used to define the quantum yield of fluorescence which can be viewed as a measure of how efficient fluorescence is compared to all the other

competing processes available to an excited molecule. The quantum yield of fluorescence, Φ_f, is therefore the proportion of photons a molecule absorbs that result in a photon being emitted due to fluorescence.[132]

$$\Phi_f = \frac{k_f}{\sum k_i} \tag{1.7}$$

1.9. FRET

Förster radiative energy transfer (FRET) is a non-radiative transfer of energy from a donor molecule to an acceptor, through a long range dipole-dipole interaction. The efficiency of the transfer process greatly depends on the distance between the two species and has therefore been used in experiments that are designed to report on the relative positions of two labelled macromolecules. It is named after Theodor Förster, who provided a fully quantum mechanical explanation of energy transfer in series of papers, starting from 1946.[133-135] Again it can be described using a Jablonski plot as seen in Figure 1.16, where the acceptor molecule is capable of fluorescing once excited by the transfer of energy.

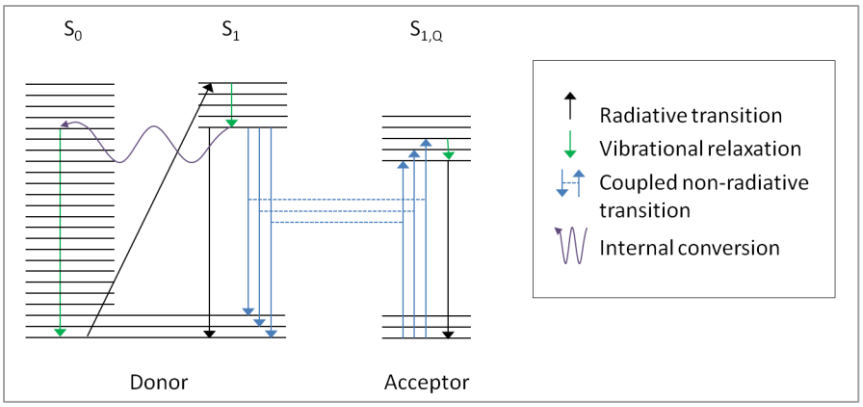


Figure 1.16: A Jablonski plot showing FRET between a donor and an acceptor, which is capable of fluorescing once excited.
The coupled transitions show resonance between the energy levels involved. Figure modified from Förster's 1948 paper.[133]

1.9.1. A brief history

Förster's seminal papers that successfully outlined the theory behind FRET were published in the 1940's and 50's, however the idea that molecules could donate and accept energy at distances beyond their collision cross section was well established and was first observed experimentally shortly after 1900. Below is a short account that briefly mentions some of the major events and people involved in understanding how FRET came to be understood; a more in depth analysis is beyond the scope of this thesis but is given by Clegg.[135]

The foundation of the theory describing FRET relies on an electromagnetic interaction between two species and requires the notion of an electromagnetic field. The movement of objects as a result of magnetism is an ancient knowledge, and has been used practically in navigation for centuries. A link between electricity and magnetism had been suspected since before the 1800's, however the first mention in the literature that describes communication between two places via electromagnetic interactions was published in 1822, when Ampère observed wires would attract or

repel each other once currents were passed through them. Later Faraday's experiments on magnetic lines of force enabled Maxwell to piece together his famous mathematical description of an electromagnetic field. It is his equations that are familiar to all physics students and were used by Hertz to derive a mathematical description of an electromagnetic field emanating from a vibrating electric dipole. These experiments were published in 1889 and the Hertzian oscillating dipole serves as a beginning point for the classical description of fluorescence and energy transfer.[135]

The emergence of quantum theory was evident at the turn of the 20th century. Plank had solved the blackbody radiation problem, by suggesting energy changes in matter could only take place through well defined leaps, before Einstein then suggested light itself could be quantified, by using wave-particle duality to explain the photoelectric effect.[135] In 1913 Bohr managed to piece together many of these ideas as well as others, including notable contributions from Conway, to present his atomic model. Electromagnetic interactions of an electron orbiting an atomic nucleus provided a cornerstone for spectroscopy, and its extension to multi-electron systems was made possible through Heisenberg's and Schrödinger's work on quantum mechanics during the 1920's. This was in turn was applied by Kallman and London, F. Perrin and ultimately Förster to describe transfers of energy in vapours and solution.[136, 137]

The first observation of energy transfer was recorded as the emission of radiation from thallium atoms that were indirectly stimulated through the excitation of mercury atoms. A comparison between the atomic radii and the calculated

'spectroscopic' cross section showed that the donor and acceptor species were too far apart to transfer energy through collisions. Cario and Franck published this sensitised luminescence in the 1920's,[138] and the dependence on the resonance between the energy levels of the sensitiser and acceptor was demonstrated by Beutler and Josephi shortly afterwards through spectroscopic studies of many pairs of acceptors and donors.[139]

Many classical theories of energy transfer involving dipole interactions were being developed in 1920's and some were arriving at the r^{-6} distance dependence typical of Förster long range dipole coupling. Kallman and London published the first quantum explanation of energy transfer in vapours in 1928,[136] which was the basis for F. Perrin's quantum theories on energy transfer in condensed systems in 1932 and 1933.[137] Both explanations were impractical to use, and F. Perrin overestimated the distances involved, but laid the foundations for Förster to extend and improve a theory of energy transfer after World War II.

Like many, Förster was partly motivated to study energy transfer by the supposedly overly efficient mechanism for plant photosynthesis. The surface area of 'reaction centres' on a leaf where electron transfer reactions took place was considered small when compared to the number of photons being absorbed. It was reasoned that the larger surrounding area was capable of absorbing photons and transferring energy to the reaction centre.[135]

Oppenheimer, also interested in photosynthesis, arrived at a correct solution and actually published before Förster. However, due to the short length of the abstract he did not arrive at an expression for general energy transfer and because it was placed in a relatively obscure journal during the war, the publication was largely ignored by the spectroscopic community.[140]

What sets Förster's work apart is that he managed to integrate experimentally observable spectra into his theory. Kallman and London included unrealistically sharp spectra which did not take in to account spectral broadening due to the fluorphores' interaction with the environment and therefore made any practical use of their theory difficult.[136] Both J. and F. Perrin, using classical and quantum mechanical explanations respectively, again only looked at sharp spectra from over simplified and identical acceptor and donor species which gave exact resonance and overestimated the distance required for energy transfer between the pairs as a result. Broadening of the spectra through collisions in solution was considered by F. Perrin and did bring his distance closer to Förster's but it was still too large.[135, 137]

Förster realised that broadening of the spectra of the donor and acceptor species reduced the chances of resonance between the two species, and that he could use statistics to calculate the number of pairs that were well matched in energy at any one time. The probability that a pair is in perfect resonance with each other can be calculated from the spectral overlap between emission of the donor and the absorption of the acceptor, and is known as the overlap integral (J). This probability is typically much less than one, and so dramatically shortens the required distance for

energy transfer to occur. Förster also included a dependence on the orientation of the two dipoles into his expression for the rate of energy transfer, which in solution can be average due to molecular tumbling of the donor and acceptor. Overall, Förster pieced together a theory and used a fully quantum mechanical approach to produce expressions that used experimentally obtainable values in its derivation which made it both accessible and extremely useful.[133]

1.9.2. FRET theory

The original quantum mechanical derivation of Förster's expression for energy transfer between donor and acceptor molecules was described by Förster in German in 1948 and an English translation was published by Knox in 1993.[133, 134] The analysis presented below follows van der Meer's interpretation that was published in 2013.[141] There can be found more detailed accounts of the theory behind FRET including discussions on the classical descriptions, the overlap integral, the orientation factor and the units used in the different forms of the equation.

Atoms consists of electrons that are essentially elastically bound to a nucleus. As electrons and the nucleus both carry charge, the movement of the electrons relative to the 'stationary' nucleus creates an oscillating electric dipole. The electrical dipoles of two interacting molecules (a donor and acceptor) can be approximated as two ideal dipoles if the distance between the two molecules is large compared to the radius of the spheres containing all the individual donor and acceptor dipoles respectively. This can be shown as vectors similar to the ones in Figure 1.17, and these dipoles

should be viewed classically as oscillating dipoles, and as transition dipoles in quantum mechanics.

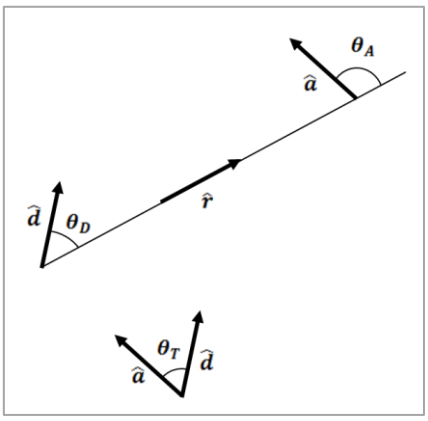


Figure 1.17: Vectors that describe the transition dipoles involved in FRET between donor and acceptor molecules.

According to quantum mechanics, there are only discrete levels of energy that are available to an electron in a confined volume. Similarly, donor and acceptor molecules can only exist in energy states that do not form a continuous spectrum. A transfer of energy from an excited donor to an acceptor in its ground state, $|D^*A\rangle$, may take place that returns the donor to its ground state and leaves the acceptor in an excited state, $|DA^*\rangle$, through a dipole dipole interaction, U. This can be expressed using the refractive index of the medium, n, the distance between the donor and acceptor, r_{DA}, the vector pointing from the donor to the acceptor, $\vec{r}$, and the dipole moment vectors of the donor and acceptor charge distribution, $\overrightarrow{p_D}$ and $\overrightarrow{p_A}$ respectively.

$$U = \frac{1}{4\pi\varepsilon_0 n^2 r_{DA}^3}[\overrightarrow{p_D}.\overrightarrow{p_A} - 3(\overrightarrow{p_D}.\vec{r})(\vec{r}.\overrightarrow{p_A})] \tag{1.8}$$

According to time-resolved perturbation theory, the rate transfer in terms of the energies of the donor and acceptor molecules, $k_T(W_D W_A)$, in the 'very weak coupling' limit defined by Förster is given in equation 1.9. The very weak coupling limit is defined where the interaction energy between the dipoles is much less than the vibronic bandwidth of the donor, which would make the requirement of resonance extremely rigorous. In practice however, solvent effects broaden the vibrational levels and FRET in the very weak limit tends to be very common.[142]

$$k_T(W_D W_A) = \frac{1}{h}\int \langle D^* A|U|DA^*\rangle^2 dW \tag{1.9}$$

As a result of the Born-Oppenheimer approximation, the expectation value of equation 1.9 can be expressed in terms of electronic transition dipole moments only, $\overrightarrow{m_D}$ and $\overrightarrow{m_A}$ for the donor and acceptor respectively, which can be written in terms of their magnitudes, $|m_i|$, and their directions, $\hat{\imath}$, to give $\overrightarrow{m_D} = |m_D|\hat{d}$ and $\overrightarrow{m_A} = |m_A|\hat{a}$. The integrand can therefore be written including an orientation factor, κ:

$$\langle D^* A|U|DA^*\rangle^2 = \frac{\kappa^2 m_D^2 m_A^2}{(4\pi\varepsilon_0)^2 n^4 r_{DA}^6} S_D^2 S_A^2 \tag{1.10}$$

where
$$\kappa = \hat{d}.\hat{a} - 3(\hat{d}.\hat{r})(\hat{r}.\hat{a}) \tag{1.11}$$

or
$$\kappa = cos\theta_T - 3cos\theta_D cos\theta_A \tag{1.12}$$

The S_D factor represents the overlap integral between the initial vibrational donor state with energy W_D^* and the final state $W_D^* - W$, and the S_A factor similarly

represents the overlap integral between the initial vibrational acceptor state with energy W_A and the final state $W_A + W$. The energy level diagram in Figure 1.16 shows the resonance between coupled transitions as energy is transferred from the donor to the acceptor.

By multiplying equation 1.10 by suitable Boltzmann factors, $g^*(W_D^*)dW_D^*$ and $g(W_A)dW_A$, integrating over all W_D^* and W_A, and changing the integration variable from energy, W, to frequency, υ, the following expression for energy transfer can be obtained:

$$k_T(W_D W_A) = \frac{\kappa^2}{(4\pi\varepsilon_0)^2 n^4 r_{DA}^6} \int_0^\infty M_D(\upsilon) L_A(\upsilon) d\upsilon \tag{1.13}$$

where:

$$M_D(\upsilon) = m_D^2 \int g^*(W_D^*) S_D^2(W_D^*, W_D^* - h\upsilon) dW_D^* \tag{1.14}$$

$$L_A(\upsilon) = m_A^2 \int g(W_A) S_A^2(W_A, W_A + h\upsilon) dW_A \tag{1.15}$$

The probability of an emission of a photon with energy between W and $W + dW$, from an excited state with energy W_D^*, is given by Einstein's spontaneous emission theory and can linked to the fluorescence properties of the donor. The coefficient of spontaneous emission (A) is given in equation (1.16), where B is Einstein's coefficient of stimulated emission. As a result equation 1.14 can be re-written in terms of the donor lifetime, τ_D, the quantum yield of the donor, Φ_D, and its fluorescence spectrum, $f_D(\upsilon)$.

$$A = \left(\frac{8\pi h\upsilon^3}{c^3}\right)B \tag{1.16}$$

$$B = \frac{\left|\mu_{fi}\right|^2}{6\varepsilon_0\hbar^2} \tag{1.17}$$

$$M_D(\upsilon) = \frac{(4\pi\varepsilon_0)3h\Phi_{\mathrm{D}}\mathrm{c}^3\mathrm{f}_{\mathrm{D}}(\upsilon)}{32\pi^3 n\tau_D\upsilon^3} \tag{1.18}$$

Similarly the second integral, equation1.15, is related to the absorption spectrum of the acceptor, $\varepsilon_A(\upsilon)$, and can be re-written as below.

$$L_A(\upsilon) = \frac{(4\pi\varepsilon_0)3(\ln 10)nhc\varepsilon_A(\upsilon)}{4\pi^2 N'\upsilon} \tag{1.19}$$

Substitution of these expressions for $M_D(\upsilon)$ and $L_A(\upsilon)$ into equation 1.13 yields Förster's expression for k_T.

$$k_T = \frac{9(\ln 10)\kappa^2\Phi_D c^4}{128\pi^5 n^4 N'\tau_D r_{DA}^6}\int_0^\infty \frac{\mathrm{f}_{\mathrm{D}}(\upsilon)\varepsilon_{\mathrm{A}}(\upsilon)}{\upsilon^4}d\upsilon \tag{1.20}$$

R_0 is defined as the distance between a donor and an acceptor at which the probability of the excited donor fluorescing is equal to the probability of a transfer of energy to the acceptor, and can be expressed as:

$$R_0^6 = \frac{9(\ln 10)\kappa^2\Phi_D J}{128\pi^5 n^4 N'} \tag{1.21}$$

where:

$$J = c^4 \int_0^\infty \frac{f_D(\upsilon)\varepsilon_A(\upsilon)}{\upsilon^4} d\upsilon \tag{1.22}$$

This simplifies Förster's expression for k_T:

$$k_T = \frac{1}{\tau_D}\left(\frac{R_0}{r_{DA}}\right)^6 \tag{1.23}$$

Similar to the definition of the quantum yield, the efficiency of this transfer can be given as a function as the rate of energy transfer over the total of all of the rates of the processes that can occur once a donor species has absorbed an electron.

$$E = \frac{k_T}{\sum k_i} \tag{1.24}$$

Substituting in $\tau_{DA} = 1/k_i$

$$E = k_T \tau_{DA} \tag{1.25}$$

Since the lifetime of a donor in the absence of an acceptor, τ_D, is defined as $\tau_{DA} + \frac{1}{k_T}$ the efficiency of energy transfer can also be written as:

$$E = \frac{k_T}{k_T + 1/\tau_D} = \frac{R_0^6}{R_0^6 + r_{DA}^6} = \frac{1}{1 + \left(\frac{r_{DA}}{R_0}\right)^6} \tag{1.26}$$

1.9.3. FRET as a spectroscopic ruler

The r^{-6} dependence of FRET and its sensitivity to changes in distance typically ranging in between 10-80 Å, means that it is ideal to be used as a reporter to changes

in protein structure and the kinetics of binding to other biological molecules. This is shown in Figure 1.18 where the changes in intensities of donor and acceptor molecules with $R_0 = 60$ Å were modelled using equation 1.19 in Figure 1.18, where R_0 is plotted with respect to distance, in terms of nm and numbers of nucleotides in a straight ssDNA. Conventional FRET experiments involve the labelling of proteins on specific locations of interest and using ensemble or single molecule spectroscopy to measure the FRET efficiency and calculate the inter-dye distances. One of the first examples of using FRET as a ruler was Stryer and Haugland in 1967, who determined the length of poly-L-proline oligomers using organic chromophores.[143] In their publication they also acknowledge the work done by Latt, Cheung, and Blout who used FRET to study an inflexible steroid.[144]

Since then many more FRET pairs of fluorescence dyes have been developed to study a wide range of biological systems, with the first FRET observed from a single pair of molecules by Ha *et al.* in 1996.[145] By observing individual pairs of dyes, it easily allowed the genuine donor-acceptor pairs to be distinguished from the donor-donor pairs and the donor molecules paired with bleached acceptor dyes that would distort any FRET value calculated from an ensemble. Single molecule FRET (smFRET) has been successfully used to study many protein and nucleic acid structures and dynamics from measuring the movement of helicases to elucidating the structures of riboswitches.[17, 146]

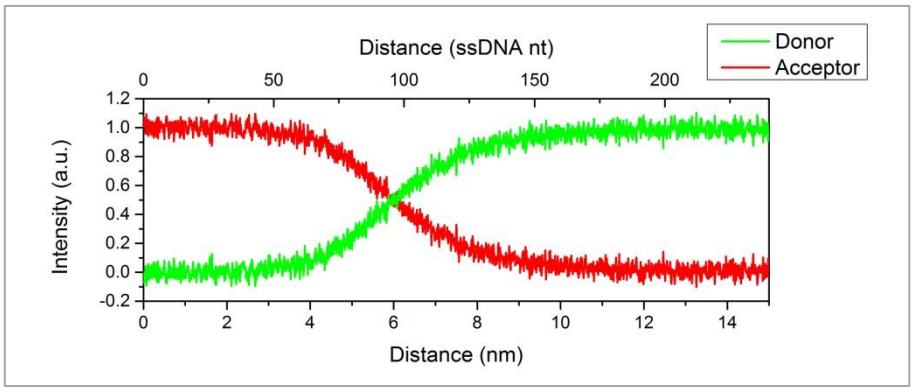


Figure 1.18: A schematic showing how the relative intensities of donor and acceptor fluorescent dyes, with R_0 = 6 nm, change with respect to distance.
The distance has been presented in terms of both nanometres and the number of nucleotides of a ssDNA where the length of a single nucleotide was taken as 0.63 nm, which was calculated from five different crystal structures.[15]

1.10. TIRF microscopy

The fluorescence intensity from a single molecule is obviously much weaker than the cumulative intensity of an ensemble of fluorophores, and therefore requires more sensitive methods of detection that provide a greater signal to noise ratio. There are several techniques available, including confocal microscopy, where lenses are used to limit the volume that is illuminated thus reducing the intensity of fluorescence originating from the bulk solution enabling the detection of single molecules.[147] This can be applied to viewing molecules on a surface but also allows the observation of freely diffusing molecules and restricts the time of observation to the duration of time that the molecule of interest spends diffusing through the excitation volume.

An alternative is total internal reflection fluorescence (TIRF) microscopy, where a laser is directed on to a surface at a critical angle that sets up an evanescent wave propagating into the sample solution.[132, 148] An evanescent wave is defined as a near field wave that has an intensity that has an exponential decay that is dependent on the

distance the wave has travelled from the boundary it has been created from.[132] Normally, as light travels into a medium with a lower refractive index, ($n_1 > n_2$, shown in Figure 1.19) refraction occurs and a certain proportion of light is reflected back off the interface between the two media. At a certain critical angle, θ_c, total internal reflection occurs and only the evanescent wave propagates into the medium of the lower refractive index. The critical angle can be calculated by rearranging Snell's law, where $\theta_1 = \theta_c$ and $\theta_2 = 90°$.

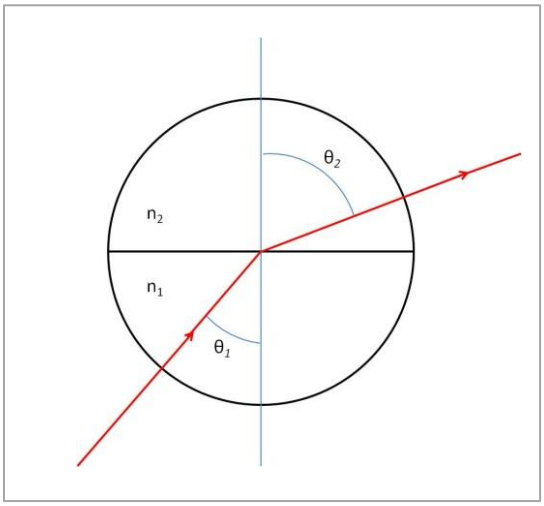


Figure 1.19: A schematic to show how light refracts as it travels into a medium with a lower refractive index ($n_1 > n_2$.)
θ_i is the angle to the normal of the surface.

Snell's law:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{1.27}$$

$$\theta_c = \sin^{-1}\left(\frac{n_2}{n_1}\right) \tag{1.28}$$

This is exploited in TIRF microscopy to only illuminate the volume next to the surface of the sample chamber, typically with a depth of 200 nm.[132] The intensity of an evanescent wave (I_z) decays exponentially as it propagates into a medium, as

described by equation 1.29, and this only excites a small fraction of fluorophores, reducing the levels of noise recorded from the bulk solution. The penetration depth is represented by d_p and the incident intensity by I_0. Using an excitation wavelength of 532 nm, which is used to excite Cy3 fluorescence dyes, d_p= 77 nm.

$$I_z = I_0 e^{-\frac{z}{d_p}} \tag{1.29}$$

where:

$$d_p = \frac{\lambda}{4\pi\sqrt{n_1^2 sin^2\theta_1 - n_2^2}}$$

A cartoon of prism type TIRF microscopy is shown in Figure 1.20, where light passes through a quartz prism before the evanescent wave is established at the boundary between the quartz slide and the solution in the sample chamber. The molecules of interest can be specifically immobilised to the surface to stop them from diffusing away from the excitation volume, maximising their exposure time. Dichroic mirrors are used to distinguish the fluorescence emitted from the donor and acceptor fluorophores and the intensities of each dye are recorded and the FRET between the two dyes can be calculated. By reducing the concentration of the target molecule to the picomolar regime, only one fluorophore is present in each pixel recorded and single molecule resolution is achieved.

Electron multiplied charge coupled device (EMCCD) cameras are widely used since they offer a high signal to noise ratio to detect single fluorophores.[149] Photons are detected by a metal oxide semiconductor capacitor that acts as a photodiode and storage device. The amount of stored charge is linearly dependent on the number of

photons detected, and after an allotted exposure time the charge is released to produce an image calculated by computer software. The dark noise is minimised by cooling the EMCCD to -80 ºC and higher quantum efficiency is achieved through using a back illuminated CCD chip.[149]

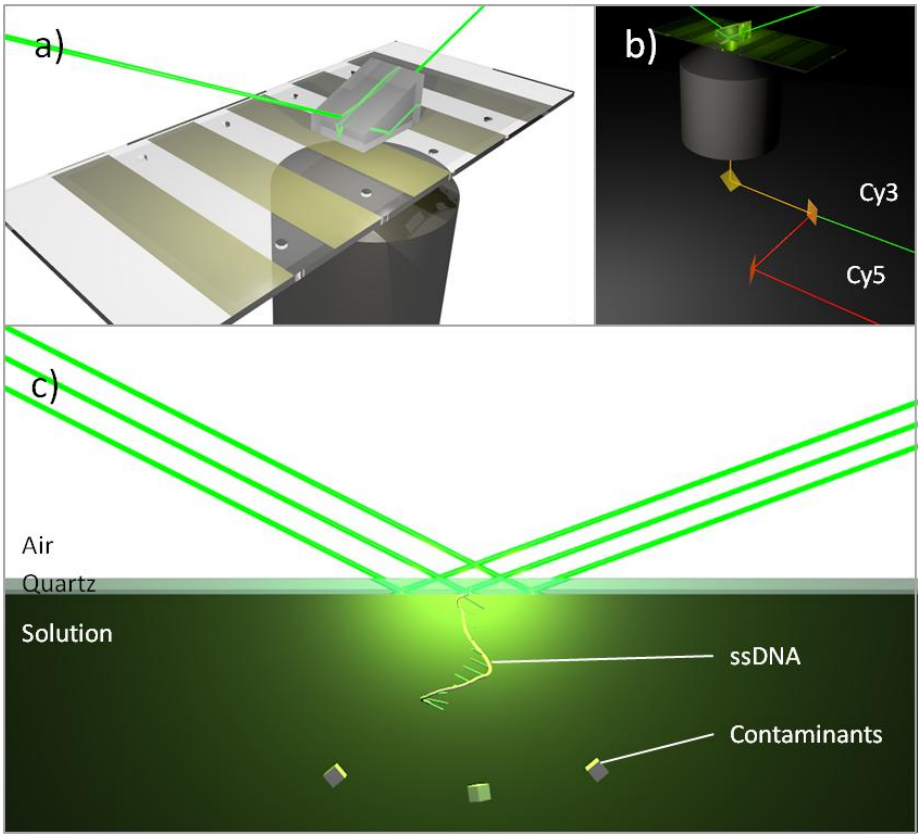


Figure 1.20: Schematic of a prism based FRET microscope.
(a) Cartoon of prism sat on top of a quartz slide positioned above an objective. The quartz slide has been modified to produce four flow chambers. (b) A cartoon showing how the fluorescence from each dye travels down the objective and is then split up by a dichroic mirror and aligned into two separate beams of light for the emission of each dye to be imaged separately. (c) A schematic of an evanescent wave only illuminating a volume within 200 nm of the surface of the quartz due to the exponential decay of the laser throughout the aqueous medium. Contaminants in the bulk solution are not illuminated, increasing the signal to noise ratio.

1.11. Surface modifications

The boundary where the evanescence wave is created in TIRF microscopy is usually the interface between the quartz surface of a microfluidic device and the solution

inside the sample chamber.[150] When proteins and other biological molecules are loaded into the sample chamber, there is a tendency for them and contaminants in solution to non-specifically stick to the quartz surface itself. In order to decrease the likelihood that unwanted fluorophores bind to the surface, the quartz is modified with a self-assembled monolayer (SAM) to produce a dense molecular brush that passivates the surface and prevents access to the quartz and therefore any non-specific binding.

A suitable molecule to use in a SAM is made up of an anchoring group, a backbone and a terminal group. The anchoring group should form a strong covalent bond between the tether and surface. The backbone is usually a hydrocarbon chain and will ideally produce a dense closely packed surface as a result of either cross linking or van der Waal's interactions between chains. The terminal group is chosen to either specifically bind a target molecule to surfaces or to provide an electrostatic repulsion to any charged protein surfaces which also helps minimise non-specific interactions. Usually the composition of the SAM is designed to produce a dispersion of anchors for the specific immobilisation of the target molecule amongst a crowd of repulsive head groups, as depicted in Figure 1.21. Examples of SAM's include alkylthiols on gold and silver and alkylsilanes on glass and quartz. Chemical modifications of SAM's can produce monolayers such as alkyls, alcohols, carboxylic acids, amines, azides, or alkynes that can securely attach a plethora of substrates to a surface.[151, 152]

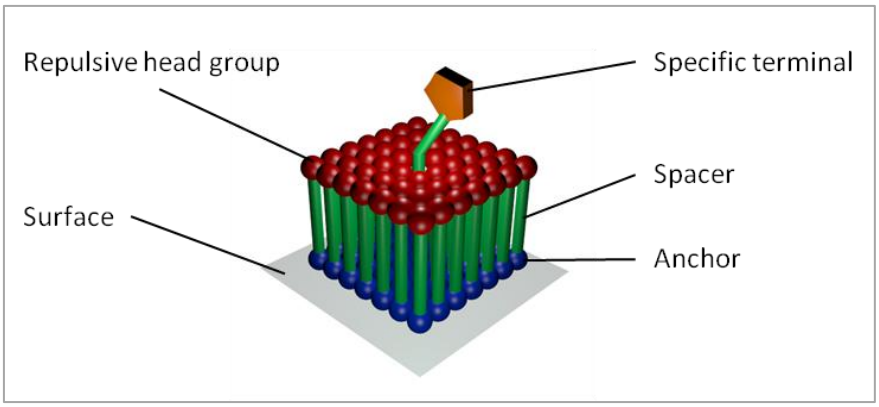


Figure 1.21: A schematic of a surface being passivated with a SAM.
SAMs can provide a specific binding point for molecules of interest as well as presenting a dense brush of repulsive head groups to minimise non-specific binding of molecules to the surface.

There are many examples of single molecule fluorescence experiments that employ a quartz surface that has been functionalised with silanes, which in turn has been modified to produce a uniform surface of poly(ethylene glycol), PEG.[146, 153-155] A common approach has been to use the interaction between biotin and avidin to exclusively bind biomolecules of interest to the PEG surface.[146, 156] Although this interaction is non-covalent it still has a dissociation constant approximately 1 fM and therefore provides a secure and specific anchor for adsorbates.[157] Figure 1.22 shows how ssDNA can be modified with a biotin at the 5' terminus and how this can be used to bind to a streptavidin molecule. By cross-linking branched PEG molecules, a denser layer of hydrocarbons is achieved and fewer non-specific binding events occur. This allows decreased levels of contamination on a surface used in single molecule experiments.[151, 158]

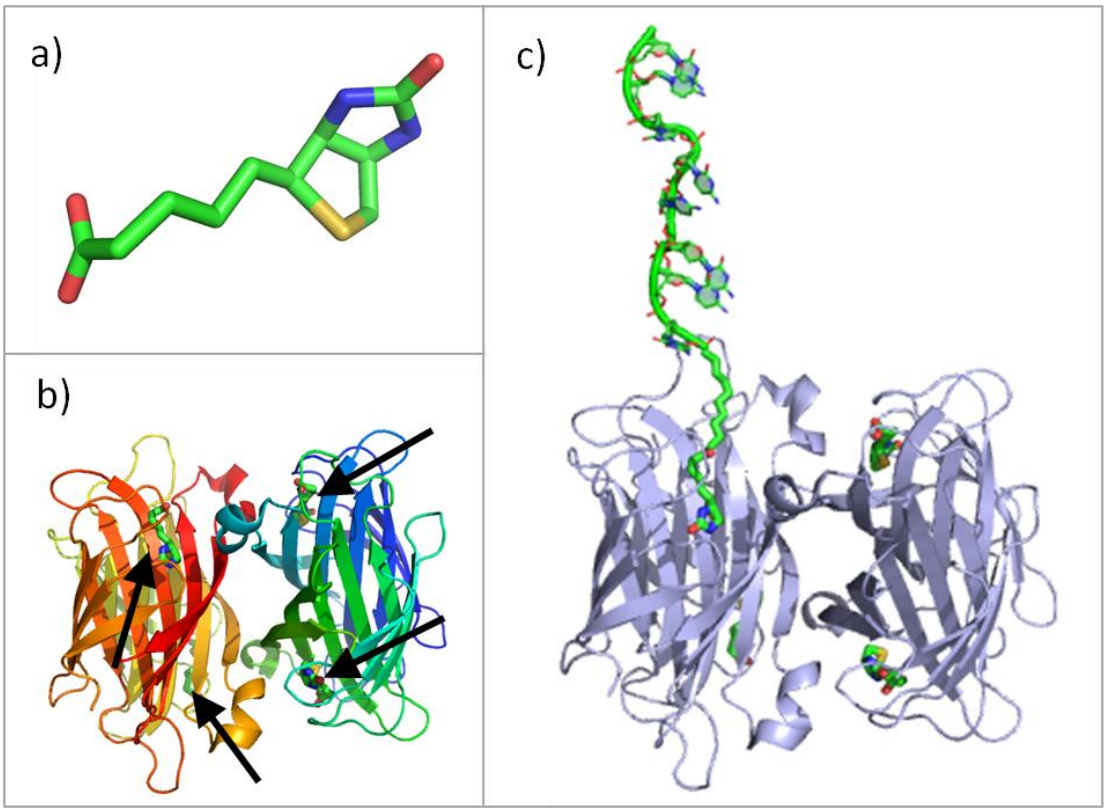


Figure 1.22: ssDNA can be specifically immobilised on a quartz slide through a biotin with an eight carbon linker to the 5' end.
(a) The crystal structure of biotin which interacts with hydrophobic and hydrophilic residues to produce an extremely strong interaction. (b) The crystal structure of streptavidin bound to four biotin molecules, one per monomer. (c) A cartoon of C8 ssDNA bound to streptavidin via a 5' biotin modification with an eight carbon linker. (Pdb 1mep and 1jmc)

Other common materials used to modify a surface are lipid molecules.[159, 160] Lipid molecules spontaneously form a layer and adhere to a surface as a result of arrangement of the molecules to optimise favourable interactions between hydrophobic chains, hydrophilic terminal groups and solution molecules. The lipid layer is a 2D fluid bound to a surface that can be made less viscous by using different lipids or by the addition of cholesterol. This provides a static anchor for the immobilisation of an adsorbate. One different and elegant solution to reducing the movement of lipid molecules on a surface has been pioneered by Greene.[161] A score is made into the quartz slide and a flow is supplied over the surface. As the lipids are

pushed along the surface, the score acts as a barrier and the lipid molecules are trapped in position against the score. Attached to the lipid molecules are strands of DNA which are pushed flat against the surface, producing a curtain of DNA that is within range of an evanescent wave in a TIRF experiment. It is therefore possible to directly view the action of DNA motors on a micrometre scale or the binding of SSB like proteins to view nucleofilaments such as RPA, as shown in Figure 1.23.[162]

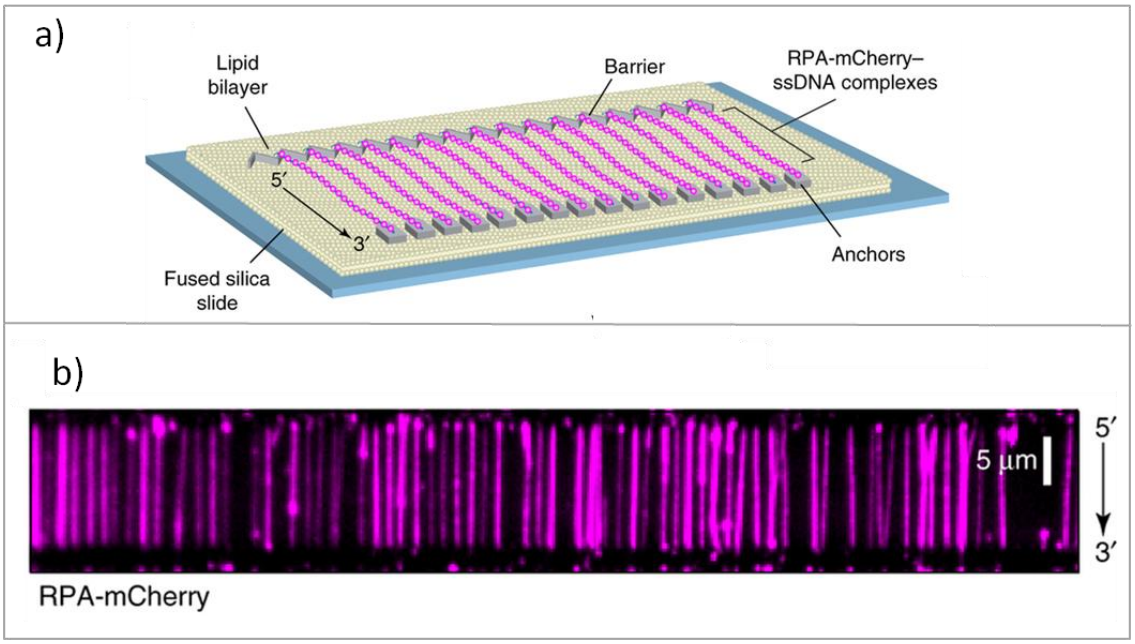


Figure 1.23: DNA curtains imaged using RPA tagged with m-Cherry.
(a) A schematic showing the experimental setup to produce (b) images of individual ssDNA molecules lying flat against the surface of a slide. Figure modified from Gibb *et al.*[163]

Single molecule experiments that investigate proteins that have low affinity to their substrate cannot usually be immobilised by a standard tether, since the concentrations of labelled substrate or protein would saturate the camera at the concentrations required to observe binding events. While an ensemble experiment would just increase the concentration of the proteins or substrate to compensate for large dissociation constants, this is unsuitable for single molecule experiments as it

would lead to large numbers of fluorophores that are not participating in binding events, essentially increasing the background fluorescence. Trapping the protein and its substrate inside a vesicle with a volume of a few attolitres effectively increases the concentrations required for binding to occur, but also localises the areas of high concentrations so that individual binding events can still be resolved.[154, 159, 160, 164] It provides a more natural environment in which to study single molecules since the protein and substrate are still in solution and do not interact with the surface. This bypasses any issues that may occur with the orientation or position that the adsorbate adopts as it binds to the surface. The vesicles are unilamellar and are typically 50 nm in diameter. The walls of these nanocontainers can be made porous either through hydrophobic/hydrophilic interactions of the membrane with the solution,[164] or through the introduction of membrane proteins that act as a molecular transport channel as shown in Figure 1.24.[159]

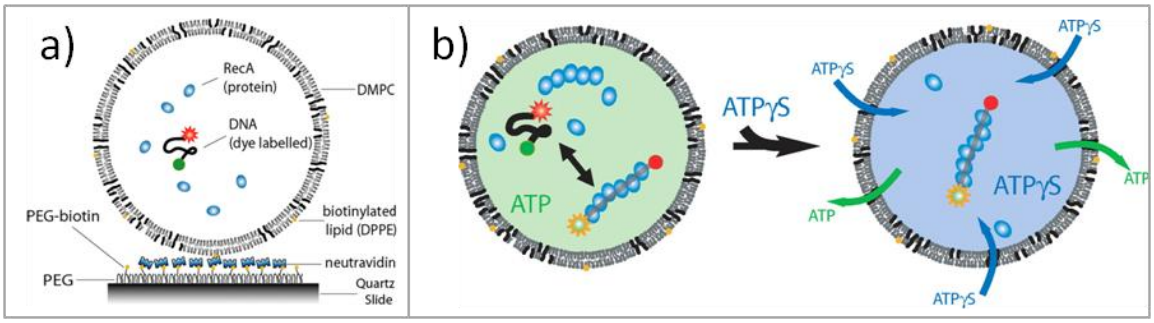


Figure 1.24: Schematics of RecA binding to ssDNA inside a porous lipid nanocontainer.
(a) A DMPC lipid nanopore is immobilised on a PEG surface through biotin/neutravidin interactions. Inside the nanopore are RecA proteins and a dual labelled ssDNA. (b) α-hemolysin in the DMPC membrane allow the passage of ATP and ATPγS across the membrane which influences the binding of RecA to ssDNA. Figure adapted from Cisse *et al*.[159]

1.12. Protein & DNA labelling

Conjugating fluorophores to a protein or nucleic acid requires a labelling reaction that is specific and efficient under relatively mild conditions. This requires a careful selection of the chemistry involved to avoid over or under labelling and also the position of the label must be considered as the presence of a dye must not interfere with the action of the target molecule.

A common technique is to express a fluorescent protein such as green fluorescent protein (GFP) on to the terminus of a protein.[132] This has been smartly exploited by Chen *et al.* who have managed to track the movement of telomeres *in vivo* by tagging a Cas9 mutant with GFP and infecting the cells with the appropriate guide siRNA that is complementary to the telomere base sequence and results in the protein forming a stable complex with the telomeres.[165] Another example is the relatively new method that includes unnatural amino acids to the protein's primary structure. These can be fluorescent themselves,[166] or provides new sites for the dyes to conjugate to.[167] This increases the chemistry available so that the specificity and efficiency of the labelling processes can be increased. By introducing an alkyne to the protein surface, copper free click chemistry can be used to bind a dye with an azide group.[168] This can be done with high specificity and selectivity since other unnatural amino acids can be incorporated into the protein structure that can bind to other fluorescent dyes via different covalent linkages.

As with unnatural amino acids, site targeted mutagenesis (STM) can also be used to include naturally occurring amino acids that can act as a linker for a fluorophore to bind to. Cysteine residues provide a thiol that can be conjugated to a maleimide

groups on a fluorophore.[146] Cysteines occur at a relatively low frequency in a protein structure and the reaction can be controlled through the pH of the reaction buffer that ensures the dye reacts specifically and selectively with the cysteines and no lysine residues are labelled.[169]

Similarly nucleic acids can be labelled utilising amine modifications on the nucleic acid and activated ester groups on the fluorophore, such as N-HydroxySuccinimide (NHS) and sulfo-NHS esters.[170] Other dyes are capable of labelling dsDNA such as YOYO-1,[24] and there are also examples of unnatural fluorescent bases that can be incorporated in to nucleic acid sequence.[171]

1.13. Fluorophores

To be useful in single molecule experiments a fluorophore must have a high quantum yield and be stable with respect to blinking and photobleaching. Blinking is the temporary switching of a molecule between fluorescent and non-fluorescent states under constant excitation, whilst photobleaching of a dye is the photo-destruction of a fluorescent dye that permanently prevents any further emission from that molecule.

Examples of intrinsic fluorophores that are naturally occurring in protein wild type structures are fluorescent amino acids such as tryptophan,[109] tyrosine and phenylalanine; and fluorescent co-factors, including flavin and NADH.[132] The appearance of these fluorophores are commonly exploited in ensemble experiments but are not usually suitable for single molecule experiments which require artificial dyes that are more photostable under the illumination of a laser.

The acronym laser stands for 'light amplification by stimulated emission of radiation' and a laser produces a highly concentrated beam of light that therefore achieves high irradiance and increases the rate of photons that are emitted by any fluorophores it excites compared to excitation by other light sources. This increases the fluorescence intensity of the fluorophores but also shortens the time before a dye is bleached, since bleaching is a stochastic process that can occur each time a fluorophore is excited.

Common extrinsic fluorophores used are quantum dots, fluorescent proteins and aromatic dyes. Quantum dots are nanocrystals that confine their excitons in all three dimensions.[172] They exhibit a bright fluorescent intensity and relatively long lifetimes before they are bleached. Their composition and size are changeable which allows them to be tuned to emit a range of emission profiles, as well as giving them a potential to be modified with a variety of surfaces and tethered to many different substrates. This gives them advantages over other fluorophores in situations where their size and toxicity are not an issue, however their increased rate in blinking limits their application in smTIRF but makes them more suited to uses in super resolution microscopy.[173] Fluorescent proteins have the advantage of their versatility since they have virtually zero cytotoxocity and mutations to GFP have produced fluorescence proteins that span the entire visible range. They are certainly widely used *in vivo* single molecule studies but again their size means that they are not suitable in crowded environments and their photostability can also limit their use in single molecule experiments.

Smaller organic molecules whose fluorescence originates from their aromaticity have been developed since the 19th century to improve their photostablilty and brightness.[132] They are easily conjugated to a wide range of biomolecules and there is a wide range of different fluorophores to choose from to match an experiment's requirements. A range of cyanine dyes derived from two semi saturated indole rings connected by an alkene, whose length determines the optical properties, are widely used due to their greater water solubility, photostability, pH tolerance and high quantum yields compared to other synthetic fluorescent dyes such as rhodamine. The wide range of cyanine dyes allows two or more dyes to be paired together for use in FRET experiments. In general they show sensitivity towards external conditions such as temperature, pH, salt concentration and also display different behaviours such as auto-quenching and dimer formation at high concentrations.[130] A common pairing is Cy3 and Cy5, whose structures are shown in below. The R_0 of this pair has been found to be approximately 60 Å and the efficiency of the energy transfer is therefore sensitive to changes in distances between 20-80 Å.

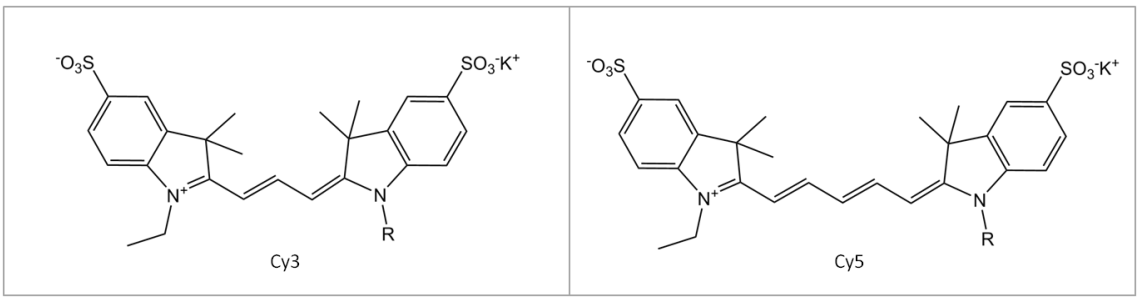


Figure 1.25: The structure of Cy3 and Cy5 dyes.
Their names are derived from the number of carbons that link the two indole rings. The R group can be a functional group used to conjugate these dyes to molecules of interest, usually with an aliphatic chain acting as a spacer between the dye and the molecules surface.

1.14. Single molecule fluorescence investigations into SSBs

ssDNA is easily immobilised on to a microscope slide and SSBs' high affinity for ssDNA allows fluorescence experiments to still observe binding events at relatively low protein concentrations, which is ideal for single molecule research. There are various single molecule experiments that are presented in the literature that describe the behaviour of SSBs from a range of organisms. FRET has been employed to report on the end-to-end distance of ssDNA overhangs during the filament growth of EcoSSB and other homotetramer SSBs from *Plasmodium falciparum* and *Thermus thermophilus*, whilst the 1D diffusion of the homotetramer SSBs and also RPA have also been characterised using similar techniques.[61, 174-177] The growth of EcoSSB filaments has also been observed on lambda DNA by applying a flow to the sample chamber and exposing the DNA to fluorescently labelled SSB. These techniques have also been applied to other proteins that produce similar nucleofilaments, notably RecA, RAD51 and Brca2.[24]

It is the purpose of this thesis to use similar single molecule techniques to investigate the binding of SsoSSB to ssDNA and how SsoSSB forms nucleofilaments. SsoSSB's monomeric structure gives a unique opportunity to focus on the interaction between individual OB folds and ssDNA, which could be extrapolated to uncover details about the behaviour of SSBs that contain multiple OB folds.

Following this introduction, a guide to how these studies were completed is given in the material and methods chapter, before the ensemble results are initially presented to investigate the binding of SsoSSB to ssDNA that were both labelled with fluorescent dyes. SsoSSB's preference to adopt a monomeric structure was

confirmed using fluorescence and pulsed electron-electron double resonance (PELDOR) experiments while dissociation and cooperativity constants were quantified from a number of different fluorescent titrations. These showed how the presence of proteins and other fluorescent dyes were able to affect the photophysical properties of a dye molecule as its own microenvironment changes. FRET, protein induced fluorescent enhancements and fluorescent quenching were all observed in different ensemble titrations to characterise SsoSSB's affinity for ssDNA. These concepts were then built upon in the next chapter which presents how single molecule experiments allowed a new in-trace analysis which combined all three fluorescence phenomena. The analysis was developed to accurately determine the dwell times of the first individual SsoSSB monomers binding to ssDNA, which confirmed and complemented the conclusions made from the ensemble data. Studies into SsoSSB's promiscuity are given in chapter 5, which describes how analogous ensemble and single molecule fluorescence experiments were used to show the SsoSSB has a comparable affinity for RNA as it does for ssDNA, suggesting extended roles for SsoSSB in the cell. Future experiments are then suggested to investigate how SsoSSB tolerance for high temperatures and extreme salt concentrations could be applied commercially before the conclusions are summarised in the final chapter.

2. Materials and methods

2.1. SSB expression & purification

Wild-type SSB and variant A114C from *S. sulfolobus* were transformed and expressed as described previously.[8] They were transformed by addition of 50 ng of of DNA to 50 μl of C43 competent cells. These were plated on to plates containing kanamycin and incubated upside down overnight at 37 °C. A single colony was extracted and grown overnight in 100 ml Luria-Bertani (LB) medium with kanamycin (final concentration 35 µg/ml) at 37 °C with shaking. 15 ml of the culture was used to inoculate 1 L of LB and 35µg/ml kanamycin for approximately 3 hr at 37 °C with shaking. Once the optical density at 600 nm was between 0.6 and 0.8, IPTG was introduced to a final concentration of 0.4 mM to induce expression of SSB for a further 4 hr at 37 °C. The cells were pelleted by centrifugation (Beckmann, JLA 8.1000 rotor) at 4°C, 5000 rpm for 20 min and frozen at -20 °C until required.

Buffer A was made from 20 mM Tris-HCl (pH 7.40), 1 mM EDTA and 1 mM DTT; buffer B was made from buffer A and 1 M NaCl; gel filtration buffer was made from buffer A and 500 mM NaCl; and lysis buffer from buffer A, 500 mM NaCl, 1 tablet of protease inhibitors (EDTA free, Roche, cOmplete), and 0.1 % Triton x-100 detergent.

Lysis buffer was used to re-suspend the thawed pellet up to a volume of 50 ml and immediately sonicated (Soniprep 150, MSE (UK) Ltd) at 4 °C for 6 x 30 sec with 30 s cooling between each sonication. The pellet was heat treated at 70 °C for 20 min

and the denatured proteins were precipitated by centrifugation (Beckman, JA 25.50 rotor) at 4 °C, 20000 rpm for 20 min. The supernatant was diluted with an equivalent volume of buffer A, to reduce the NaCl concentration below 250 mM, and filtered (Millex, 0.22 µm).

The solution was passed through two 5 ml heparin-Sepharose column (GE Healthcare) equilibrated with buffer A. The protein was eluted off the column along a concentration gradient of buffer B and the protein was collected in fractions. The fractions containing protein were concentrated down to below 10 ml and passed through a HiLoad 26/ 60 Superdex 200 size exclusion column (GE Healthcare) equilibrated with gel filtration buffer. To remove any remaining impurities, the protein was passed through a 5 ml sepharose column (GE Healthcare) equilibrated with buffer A, and eluted with buffer B. The protein concentration in each fraction was monitored using UV-vis absorption at 260 nm and proteins peaks were investigated using an SDS-PAGE gel (Invitrogen, precast NuPAGE 4-12% Bis-Tris, 200 V for 35 min).

The concentration of pure SsoSSB in solution was determined using a Cary Varian UV-vis spectrophotometer and an extinction coefficient of 12660 $cm^{-1}M^{-1}$. All binding experiments were carried out in 50 mM Tris-HCl (pH7.5), 50 mM KCl unless stated otherwise.

2.2. Protein Labelling

SSB A114C was labelled with maleimide derivatives of Alexa 647 (equivalent to the Cy5 dye). The thiol group on the A114C variant was exploited to conjugate the dyes

to the protein, the mechanism of which is shown below in Figure 2.1. The labelling was carried out by unfolding the protein in 8 M urea and reducing the cysteine groups in a 9 molar excess of TCEP that had been purged of air by flowing nitrogen through the solution for 30 min. The protein was added to the eppendorf tube, given a blanket of nitrogen and left to incubate at room temperature for 20 min. A 10 molar excess of maleimide functionalised dye was added to the solution and the eppendorf tube was again flushed with nitrogen. This was left at room temperature overnight. The solution was diluted with buffer A (see protein purification of SSB) to re-fold the protein and loaded on to a heparin column before being eluted off by a concentration gradient of buffer B.

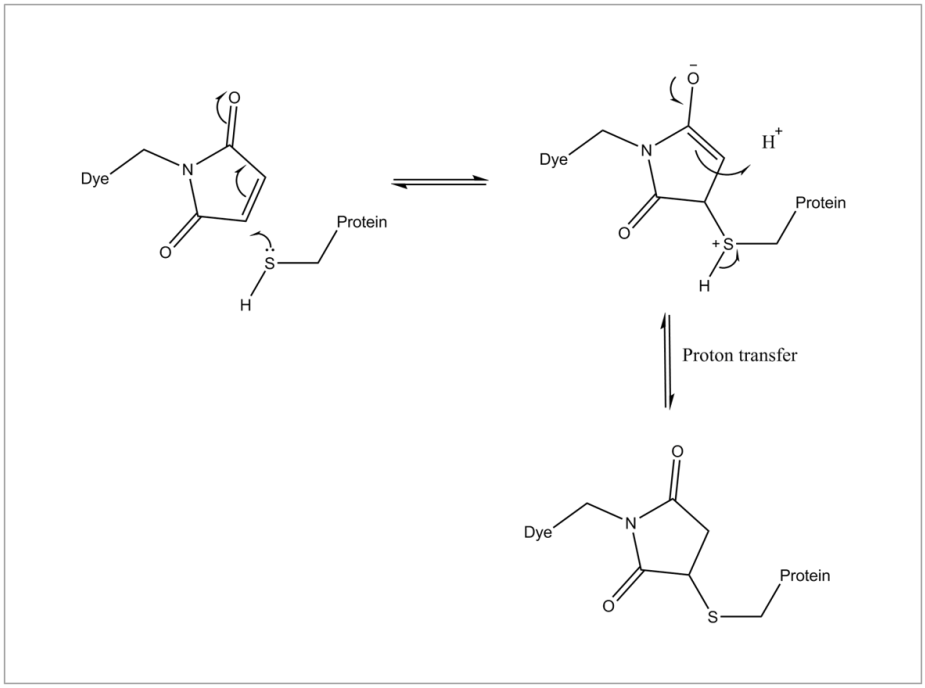


Figure 2.1: Curly arrow mechanism for the maleimide thiol coupling used for attaching fluorescent dyes to exposed thiol groups on cysteine residues on the surface of proteins.

2.3. Labelling and annealing DNA

Single stranded DNA (ssDNA) was labelled with Cy3 and Cy5 dyes (GE Healthcare) via NHS-ester linkages. Initially the dry pellet of DNA was dissolved in deionized water to a concentration of 25 µg/µl. The dye (either Cy3 or Cy5), 1mg, was dissolved in 60 µl of DMSO, split into 4 tubes and dried in the SpeedVac system (several hours) before re-suspending each pellet in 14 µl of DMSO. In the vial with DMSO and the dye 4 µl a solution of 25 µg/µl oligonucleotide, 7 µl of deionized water and 75 µl of labeling buffer (0.1 M sodium tetraborate, pH 8.5) was added. These were incubated at room temperature overnight wrapped in aluminum foil. The DNA was precipitated by adding 10 µl of 3M of sodium acetate in 100 µl of labeling reaction, which was mixed by pipetting. Absolute ethanol was added and mixed gently by inversion. The tube was wrapped in aluminium foil and incubated overnight at -20 ℃. The mixture was then spun at 13000 rpm for 1 hour. The supernatant was removed and the pellet dissolved in 50 mM Tris-HCl pH 7.5. The DNA was separated from any unbound dye using a denaturing polyacrylamide gel 12% with 7 M urea, 300 µl of APS (ammonium persulfate), and 30 µl of TEMED with a total gel volume of 50 ml. The gel was run for 1 hour in TBE before the DNA was loaded in 50% of formamide at 18 W with a set temperature threshold of 25 °C. Following electrophoresis the band was cut from the gel, transferred to a vial, chopped and re-suspended in 400 ml 50 mM Tris pH 7.5. The DNA was incubated overnight wrapped in aluminium foil in a shaker. The DNA was then spun at 13000 rpm for 10 min, and the supernatant retrieved using a micro spin column. The concentration was measured on Cary Varian UV-vis spectrophotometer. If it was required, ssDNA was annealed by mixing 200 pmoles of each strand in 50 mM Tris-

HCl, 1mM EDTA and 50mM KCl buffer. The DNA was put into a water bath pre-heated to 70 °C and then switched off overnight. This was then run through a native 12% polyacrylamide gel (as before without urea) and the DNA was retrieved using the crush and soak method. This was then precipitated a final time.

2.4. Fluorescence experiments

All bulk fluorescent experiments were carried out on a Cary Varian spectrophotometer.

2.4.1. Ratio $_A$

Förster resonance energy transfer (FRET) is a dipole-dipole interaction between a donor and acceptor that has r^{-6} dependence, where r is distance between the two dyes. It also has a dependence on the spectral overlap of the emission of the donor and the absorbance of the acceptor. Cy3 and Cy5 have appropriate optical properties to act as a donor and acceptor respectively, as well as their capacity to act as stable, long lived fluorophores. The relative intensities of these dyes can therefore be tracked so FRET can be used as an optical ruler to measure the distance between the dyes as SSB binds to ssDNA. Figure 2.2 shows the relationship between the efficiency of the energy transfer from Cy3 to Cy5. R_0 is the distance where the energy transfer is 50%, and for Cy3 and Cy5 this occurs when the separation between the dyes is 5.6 nm.

The efficiency of the energy transfer between the two dyes may be tracked by measuring the changes in the intensity of the acceptor peaks. $Ratio_A$ is the ratio between the integrals of the Cy5 peaks that arise due to energy transfer from Cy3 and excited directly. There is an overlap with the Cy3 peak which has to be subtracted to obtain a true intensity of Cy5.

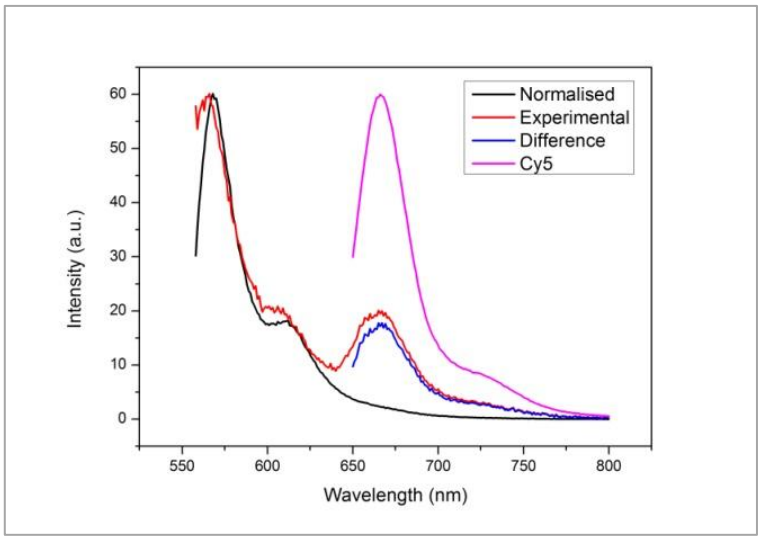


Figure 2.2: Four fluorescence spectra to illustrate how the Ratio A value is calculated.
During the experiment the emission spectrum is recorded using excitation wavelengths 550 nm and 640 nm which excite the Cy3 and Cy5 dyes respectively. Cy3 emits at 570 nm and Cy5 emits at 660 nm. The Cy5 peak of the 550 nm spectrum (red) also contains a small intensity peak from the Cy3 peak (black), which has to be subtracted to give a true intensity of the Cy5 peak. The area underneath the curve with the Cy3 subtrated (blue) is then divided by the area of the Cy5 peak that is excited at 640 nm (green) to give the Ratio A value.

2.5. Binding to a one dimensional lattice

2.5.1. 1:1 binding

The strength of how tightly a protein NA complex is held together can be described by the dissociation constant (K_d). This can be modelled for a general reaction, shown below, since the rate of any reaction at equilibrium is zero.

$$A_xB_y \rightleftharpoons xA + yB$$

at equilibrium

$$\frac{d[A_xB_y]}{dt} = k_{assn}[A]^x[B]^y - k_{diss}[A_xB_y] = 0 \tag{2.1}$$

$$K_d = \frac{k_{diss}}{k_{assn}} \tag{2.2}$$

$$K_d = \frac{[A]^x[B]^y}{[A_xB_y]} \tag{2.3}$$

For a 1:1 binding where $x = y = 1$, an expression that can be used practically in the form of equation 2.4, since $[AB]$ can be obtained using the relationships that $[A]_{tot} = [A] + [AB]$ and $[B]_{tot} = [B] + [AB]$. For lattices that were able to accommodate more than one SsoSSB monomer, the total concentration of ssDNA was multiplied by the number monomers able to bind simultaneously to an individual strand of ssDNA.

$$[AB] = \frac{-(K_d + [A]_{tot} + [B]_{tot}) \pm \sqrt{(K_d + [A]_{tot} + [B]_{tot})^2 - 4[A]_{tot}[B]_{tot}}}{2} \tag{2.4}$$

2.5.2. Cooperative binding

All binding events are considered to be equal and independent in the above equation; however a model of ligands (proteins) binding to a finite one dimensional lattice such as ssDNA or RNA will require an account of overlapping binding sites, cooperative binding and end effects. Below is a summary of the model that outlines ligands binding to a finite lattice as described by Epstein.[178]

The majority of proteins that bind to ssNA will occupy a binding site of more than a single nucleotide. Therefore the size of the binding site (n), which includes the area rendered inaccessible to other proteins through steric clashes in addition to those directly bound to the protein, as well as the length of the ssNA (M) can all be measured in nucleotides. If binding along a ssNA strand is non-specific according to

the sequence of bases a protein can bind to, a protein has $(M - n + 1)$ number of binding sites that overlap with each other, as depicted in Figure 2.3.

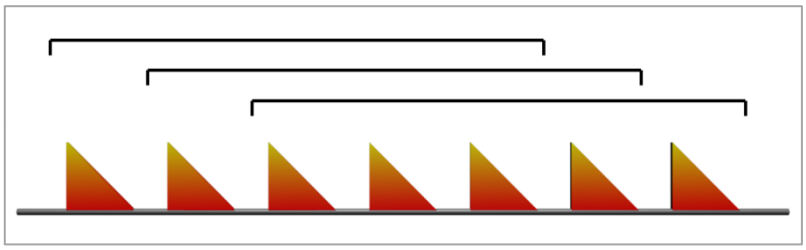


Figure 2.3: Schematic of ssNA with a length of 7 nt (M=7) where a protein that has a binding site of 5 nt (n=5), gives a total of 3 possible overlapping binding sites.

Binding sites on the same lattice are not identical, for example ligands may be more or less likely to bind to a binding site adjacent to an already occupied site. This could be due to protein protein interactions, or a distortion of ssNA resulting in a more or less favourable conformation of ssNA for binding to occur. The number of adjacencies (j) on a lattice defines the number of times a protein binds directly next to another on a single lattice. The number of proteins on the lattice is denoted as k and is shown in Figure 2.4.

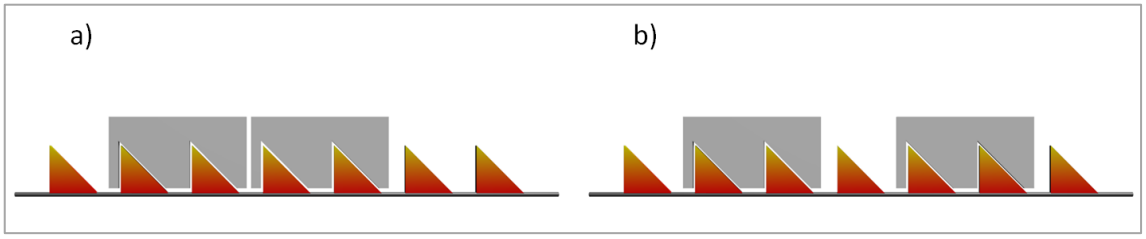


Figure 2.4: Schematic of two proteins with binding sites of 2 nt on a lattice of length 7 nt.
The arrangement on the left shows the two proteins occupying adjacent binding sites whereas on the right the proteins are bound to the lattice in an arrangement that gives no adjacencies (n=2, M=7, k=2 and a) j=1 and b) j=0.)

The cooperativity factor, ω, is an equilibrium parameter that helps define the equilibrium constant where a free ligand binds to a site adjacent to an already occupied site. The equilibrium constant becomes $K\omega$ for single contiguous regions shown in Figure 2.5(a) and $K\omega^2$ for doubly contiguous regions, such as (b).

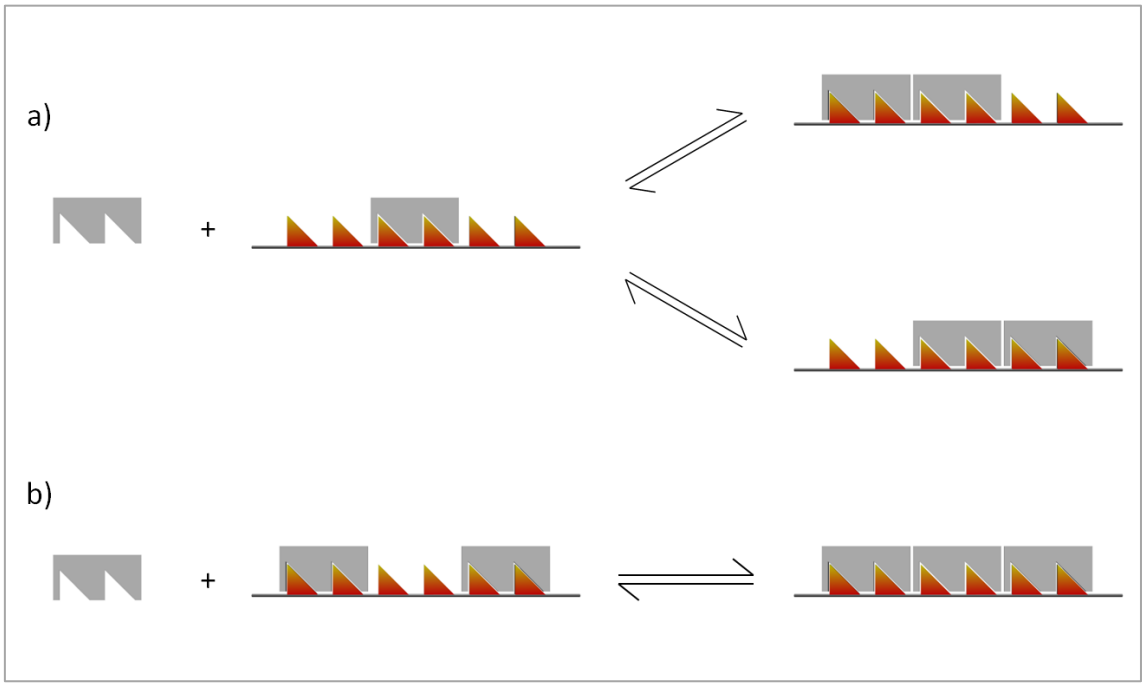


Figure 2.5: Schematic showing a free ligand binding to a) a singly and b) a doubly contiguous binding site (n=2, M=6 and for (a) k=2,j=1 and (b) k=3, j=2.)

The greatest integer, less than or equal to, M/n is the maximum number of proteins (g) that bind to a lattice at any one time. When the lattice is as fully decorated with proteins as possible, the number of nucleotides remaining (r) can be defined as $r = M - ng$. Futhermore, the concentration of free ligand (protein) is written as L, and can be calculated from the total ligand concentration, L_T, and the fraction of change in signal measured, as shown below.

$$L = L_T . \frac{(Change\ in\ signal)}{(Change\ in\ signal)_T} \tag{2.5}$$

End effects must also be taken into account as realistic lattices are finite which can limit the binding of proteins depending on the position of bound proteins. For any arrangement of ligands on a lattice $j \leq k - 1$ and $k \leq g$, which produces an increasing number of possible arrangement of k ligands as the length of the lattice increases. The relative probability of a single arrangement is $(KL)^k \omega^j$, where the relative unit probability of a set of (k, j) is set to a naked lattice.

The total number of arrangements in each set of k and j values can be denoted as $P_M(k, j)$ and can be calculated using a combinatorial analysis. The number of runs of ligands on a lattice can be given by $k - j$, and in order to keep those runs separate $k - j - 1$ sites that are 'attached' to the right hand side of each run must be left vacant. The remaining unattached vacant sites (N_u) are therefore:

$$N_u = M - nk - (k - j - 1) \tag{2.6}$$

The number of possible configurations of runs and unattached vacant sites is denoted N_c and given by equation 2.7.

$$N_c = \frac{(N_u + k - j)!}{N_u!\,(k - j)!} \tag{2.7}$$

Run lengths will not be identical in non-ideal situations and different lengths in runs (N_c) must be multiplied by the number of ways the run lengths can be arranged (N_p) in order to calculate $P_M(k,j)$ (equation 2.10), which represents the number of ways $k-j$ may be split up into positive integers. This is given by:

$$N_p = \frac{(k-1)!}{j!\,(k-j-1)!} \tag{2.8}$$

$$P_M(k,j) = N_c N_p \tag{2.9}$$

therefore

$$P_M(k,j) = \frac{(M-nk+1)!\,(k-1)!}{(M-nk-k+j+1)!\,(k-j)!\,(k-j-1)!} \tag{2.10}$$

$P_M(k,j)$ can be used to calculate $\bar{k}$, the average number of ligands on a lattice at a given concentration of free ligands.

$$\bar{k} = \frac{\sum_{k=0}^{g}\sum_{j=0}^{k-1} kP_M(k,j)\,(KL)^k\omega^j}{\sum_{k=0}^{g}\sum_{j=0}^{k-1} P_M(k,j)\,(KL)^k\omega^j} \tag{2.11}$$

This in turn can be used to calculate the fraction of occupied sites (θ) on the lattice which is directly proportional to experimental changes in fluorescence:

$$\theta = \frac{n\bar{k}}{M} \tag{2.12}$$

2.6. Wormlike chain model

Unless any significant complementary base pairing is possible, ssDNA and RNA can be considered to behave as disordered polymers. In an ideal polymer each monomer is free to position itself in any degree of rotation relative to the other monomers in the polymer. The average length between the two ends of ideal polymers is zero as a result. A single nucleotide has constraints placed on it due to the sterics of being bound to the rest of the chain which gives the strand at least a vague directionality. Electrostatic interactions and π-π stacking result in certain conformations being more favourable energetically and deviations away from these conformations costs energy. ssNA therefore cannot be described as an ideal polymer. An appropriate model for ssNA is the wormlike chain model which takes into account the energy costs that are required to change conformations.[15] An impression of the flexibility of a wormlike chain can be described using the persistence length, (L_P) which is a measure of the average length of each 'link' in the chain and can be calculated from probability function of the end-to-end distance ($p(r,t)$), given below in terms of the end-to-end distance (r) and contour length (L) expressed in multiples of persistence length ($t = L/L_P$).[15]

RatioA to FRET:

$$E_F = \frac{\varepsilon_A(\lambda_A)}{\varepsilon_D(\lambda_D)} * \left(R_A - \frac{\varepsilon_A(\lambda_D)}{\varepsilon_A(\lambda_A)}\right) \tag{2.13}$$

FRET to distance:

$$r = R_0 \left(\frac{1}{E_F - 1}\right)^{1/6} \tag{2.14}$$

Persistence length:

$$p(r,t) = \frac{4\pi A r^2}{(1-r^2)^{9/2}} exp\left(-\frac{3t}{4(1-r^2)}\right) \quad (2.15)$$

$$A = \frac{4(3t/4)^{3/2} exp(3t/4)}{\pi^{3/2}\left(4 + \frac{12}{(3t/4)} + \frac{15}{(3t/4)^2}\right)}$$

Where A is a normalisation constant, r is the end-to-end distance and $\varepsilon_i(\lambda_j)$ represents the absorption coefficient of the acceptor or donor at the wavelength corresponding to the direct excitation of the acceptor or donor.

2.7. Single molecule experiments

2.7.1. Cleaning of slides

Holes were drilled into the microscope slide to create what would become 4 channels per slide. Slides were cleaned manually using kimwipes and acetone then methanol. They were then immersed in etchant solution (1:1 of H_2O_2 and HCl) for 2 x 30 min before being washed in a sonicator bath for periods of 15 min immersed in detergent (2% Alconox in Millipore water), acetone, 1 M KOH, methanol, and finally 1 M KOH. In between each step the slides were rinsed with Millipore water and then washed in fresh Millipore water in the sonicator bath for 5 min.

2.7.2. Aminosilation of glass

The slides are cleaned as described above and coverslips are cleaned using the sonication bath only. A slide was dried under a stream of nitrogen then immediately burnt, each side of the slide being held under a blue flame for 2 min. The slide was

cooled under nitrogen and stored in a falcon tube. A coverslip was also dried and burnt, but only held in the flame for less than a second. The slides and coverslips were placed in clean jars and 2.5 ml of acetic acid and 200 μl of aminosilane was added. The jar was placed in a sonicator bath for no more than10 min before the slides and coverslips were rinsed with methanol, water and methanol again before the slides and coverslips were dried under nitrogen.

2.7.3. PEGylation of amino-covered slides

1.5 mg of biotinylated PEG was added to 320 μl of PEG buffer (100 mM sodium bicarbonate pH 8.5). This solution was placed on the slide where it was to be used as a channel, and a coverslip was placed on top. The ensemble was placed in a humidity box and left for between 1.5 and 2 hours. The coverslip and slide were separated and again rinsed with methanol, water and methanol before being dried and put back together creating a microchannel as described below.

2.7.4. Microchannel preparation

Double sided tape was applied either side of the region of the slide to be used as the channel. A coverslip was placed on top of the tape and pressure was applied with a pipette tip to remove any air bubbles. Araldite epoxy (50:50 mix) was smeared to the corners of the coverslip to close the channel. This was left to dry for 5 min before the channel was loaded.

2.7.5. TIRF experiments

Single molecule experiments were carried out using a custom built prism type total internal reflection fluorescence (TIRF) microscope. In all solutions an imaging buffer was used in order to exclude oxygen and other known fluorescence quenchers. This

consisted of 50 mM Tris-HCl pH 7.5, 6% glucose, glucose oxidase, catalase and Trolox. Neutravidin was used to bind biotinylated oligos to biotinylated PEG molecules at a concentration so that a single oligo could be resolved with a 60x magnification lense. Fluorescence intensities of Cy3 and Cy5 could be simultaneously measured by an electron multiplied charge coupled device (EMCCD) camera and recorded and analysed on a computer. The fluorescent intensities were analysed using IDL and MATLAB custom written scripts.

2.7.6. Rate analysis

Dwell times were measured from the single molecule traces. For example, a histogram was created from the dwell times that a single SsoSSB monomer spent bound to ssDNA. The histogram was then fitted to single exponential equation to calculate the rate constant, k.

:

$$y = Ae^{-kx} + y_0 \tag{2.16}$$

where:

$$k_{on,i} = \frac{1}{\tau_{on,i}}[SSB] \tag{2.17}$$

$$k_{off,i} = \frac{1}{\tau_{off,i}} \tag{2.18}$$

where i is the first or second monomer or dimer.

2.8. Nucleic acid sequences

All nucleic acids were purchased from Integrated DNA technologies (IDT) and their sequences are given below, where 5Biosg, 5AmMC6T, 3AmMC6T and 3Cy3sp are abbreviations used by IDT to represent 5' biotin, 5' amino linker C6 dT, 3' amino linker C6 dT and 3' Cy3 modifications.

Table 2.1: Names and sequences of nucleic acids used.

Name	5' 3'
ssDNA C_{20}	/5Biosg/CC CCC CCC CCC CCC CCC CCC /3AmMC6T/
ssDNA C_{12}	/5Biosg/ CCC CCC CCC CCC /3AmMC6T/
SSBbind1	/5Biosg/ AGC GAC GGT ATT CGT ATC GA /3AmMC6T/
SSBbind2	/5AmMC6T/ TT TTT TTT TTT TCG ATA CGA ATA CCG TCG CT
SSBbind3	/5AmMC6T/ TT TTT TTT TTT TTT TTT CGA TAC GAA TAC CGT CGC T
RNA 6-FAM	/56-FAM/CUU UCA AUU CUA UAG UAG AUU AGC
RNA C_{20}	/5Biosg/CC CCC CCC CCC CCC CCC CCC /3AmMC6T/
HeAc1	/5Biosg/TT TTT TTT TTT TTT TTT TTT TCG ATA CGA ATA CCG TCG CT/5AmMC6T/
Trac1	/5AmMC6T/T TTT TTT TTT TTT TTT TTT CG ATA CGA ATA CCG TCG CT /3Cy3sp/

2.9. Exosome degradation

200 nM RNA labelled with fluorescein (RNA-FAM) at the 5' end was incubated with wild type SSB (concentrations from 0 – 480 µM) for 5 min at room temperature in 20 mM HEPES (pH 7.9), 0.1 mM EDTA, 60 mM KCl, 8 mM $MgCl_2$, 2 mM DTT, 10 mM K_2HPO_4. To each aliquot, 0.5 µl *Sulfolobus solfataricus* Rrp41-Rrp42 hexameric ring and 0.4 µl Rrp4 protein was added. Both the Rrp41-Rrp42 and

hexameric ring were kindly donated by Dr Elena Evguenieva-Hackenberg. The total volume of each aliquot was 10 µl. The reaction was left to incubate at 60 °C for 1 hour. 10 µl of each sample was added to acid phenol (Ambion) and mixed thoroughly, then spun at 13000 rpm for 1 min. 5 µl from the resulting bubble formed was added to 5 µl formamide (Promega) and loaded on to a denaturing gel (25 % polyacrylamide, 7 M urea, 300 µl of ammonium persulfate (APS), and 30 µl of TEMED, 5 ml TBE, total volume 50 ml) run at 85 W with a temperature threshold of 50 °C for 2.5 hr. The image was scanned using a FLA Typhoon 5000 and analysed using ImageJ software.

3. Ensemble studies of SsoSSB

3.1. Introduction

SSBs were originally perceived as relatively simple proteins that only took on a role as a caretaker of ssDNA. However more thorough reviews of SSBs' interaction with ssDNA highlight some of the more subtle aspects involved in controlling the growth of a nucleofilament. A high affinity for ssDNA is critical for the formation of SSB/ssDNA complexes; however an SSB too tightly bound to ssDNA prevents access to the base sequence of the ssDNA and is essentially a road block to replication, repair and transcription. SSBs' binding to areas of DNA damage also assist in melting the double helix to provide a large enough bubble for DNA repair machinery to act upon.[88] However the size of this nucleofilament must be controlled so that the resulting bubble is not unnecessarily large, and the binding of SSB monomers must be weak enough to allow the substitution of SSBs with the appropriate nucleases and polymerases.[86] This is clearly demonstrated by the multiple binding modes of RPA, where the weaker 8-10 nt binding mode helps regulate the size of the repair bubble.[66] It is clear that SSBs' affinity for ssDNA must compromise between being strong enough for any nucleofilament to offer effective protection against insults, as well as being weak enough for the protein to be removed at the appropriate time. This suggests that the nucleofilament is a more sophisticated and dynamic environment than originally thought and that the strength of binding must be finely balanced in order to carry out the many roles that SSBs undertake.

SsoSSB stands out as a model protein in the literature because it has been reported to be a monomer in solution and could therefore present a unique opportunity to study individual OB folds binding to ssDNA.[114] Its similarities to the binding domains in both RPA and EcoSSB allow comparisons to be made to more established models, and possibly some details of SsoSSB binding can be extrapolated to explain the behaviour of these more complex arrangements of SSBs. The following chapter will discuss how ensemble fluorescence and other techniques can be used to explore the validity of the claims in the literature that SsoSSB is a functional monomer and how it is possible to follow the initial stages of building a nucleofilament through a novel single molecule combination of FRET, quenching and enhancement events.

3.2. Expression and purification

The purification of wild type and variant SsoSSB were followed by UV-vis and SDS-PAGE analysis. The gel filtration step only reported a single peak, shown in Figure 3.1(a), that coincided with where a protein with a molecular mass of approximately 16 kDa was expected to elute from the column, in agreement with previous experiments by Wadsworth *et al*.[109] A single band from SDS-PAGE analysis shown in Figure 3.1(b), which indicated that the purification produced a sample of suitably pure protein for ensemble and single molecule use, after the heparin, sepharose and size exclusion column. MALDI-TOF mass spectrometry carried out by Dr Catherine Botting's group, University of St Andrews, confirmed that the variant A114C carried a molecular mass of 16.170 kDa, compared to 16.1699 kDa calculated from the total molecular weight of the sum of amino acids.

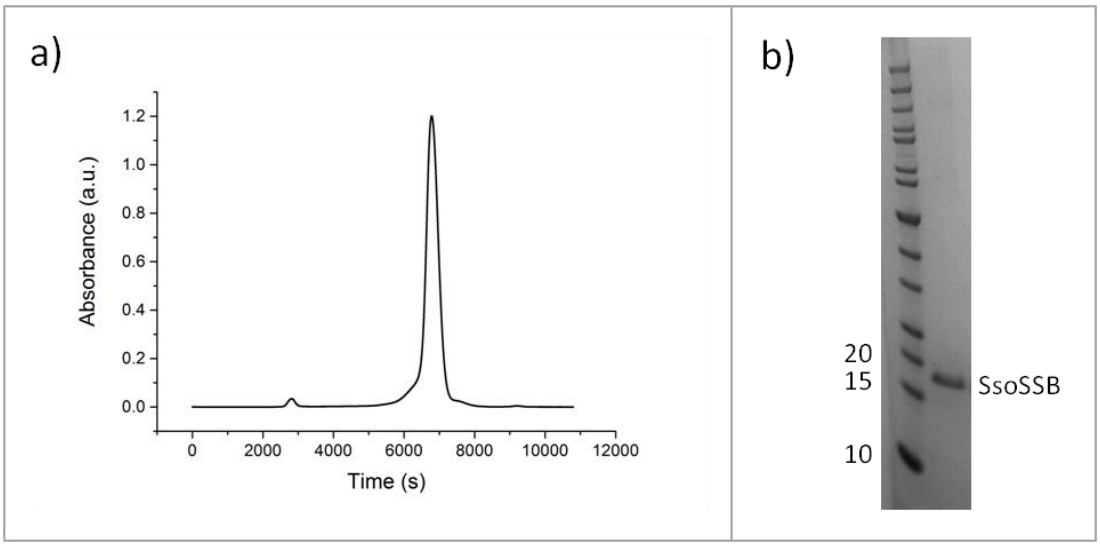


Figure 3.1: The purification of wild type SsoSSB.
Only a single peak from the absorption at 280 nm was observed during size exclusion chromatography after the purification of SsoSSB. Following a further purification step with a Sepharose affinity column, SDS-PAGE indicated that the purification produced a pure protein sample, with a monomer mass between 15-20 kDa.

3.3. Tryptophan quenching

Tryptophan is a naturally occurring residue that has the capacity to fluoresce after excitation to either the ^1L_a or ^1L_b states. Upon excitation to ^1L_a the electronic density shifts from the pyrrole to the benzene ring, which produces a relatively polar arrangement of electronic charge that can be stabilised through hydrogen bonding to polar solvent molecules such as water, which positions the ^1L_a energy level below the ^1L_b. Experimentally, tryptophan is excited at a wavelength of 295 nm, which predominately excites tryptophan residues over other fluorescent residues such as phenylalanine and tyrosine. This will generally produce a spectrum that has combination of characteristics from the fluorescence of both ^1L_a and ^1L_b states, depending on the extent of solvation of the tryptophan residues.

SsoSSB uses two tryptophan residues and a single phenylalanine residue to stack in between the bases of ssDNA and the presence of a strand of DNA in the OB fold excludes water molecules from surrounding these aromatic residues.[114] This in turn raises the energy level of the tryptophan ^1L_a state relative to the ground state, and an excitation of 295 nm no longer produces an excited state of tryptophan with the same optical properties. The ^1L_b state is now lower in energy than the ^1L_a, which is less sensitive to changes in polarity, and a decrease in fluorescence emitted at 350 nm is observed.

Naturally occurring fluorescent residues that can report on ssDNA binding can be convenient tools in order to analyse SSBs in ensemble experiments; however they are not ideal fluorophores in terms of quantum yield and photostability for more sensitive techniques such as single molecule microscopy. Synthetic dyes are designed to emit photons at a higher rate and to be less susceptible to blinking or photobleaching. They can be tethered to ssDNA and SsoSSB to gain detailed reports on their relative positions, and therefore elucidate the mechanisms of binding of these proteins to their substrates.

3.4. Dimensions of SSB/ssDNA complexes

The affinity of SSBs for ssDNA have resulted in dwell times for bound states that persist for several seconds, however the dynamics of proteins in nucleofilament have shown structural changes separated by microseconds.[155, 179] To study SSB complexes it is therefore necessary to have a technique capable of a high temporal resolution but which also has the capacity to follow longer lived complexes. The lifetime of excited

states capable of fluorescing are normally in the range of nanoseconds and therefore reliably report any changes in the environment of the protein in real time. The use of fluorescence microscopy is therefore suitable to track the initial stages of the formation of a nucleofilament. The intensities of fluorescent dyes are also particularly sensitive to changes in their local environment, which again makes them well suited to watching SSB monomers binding to ssDNA.

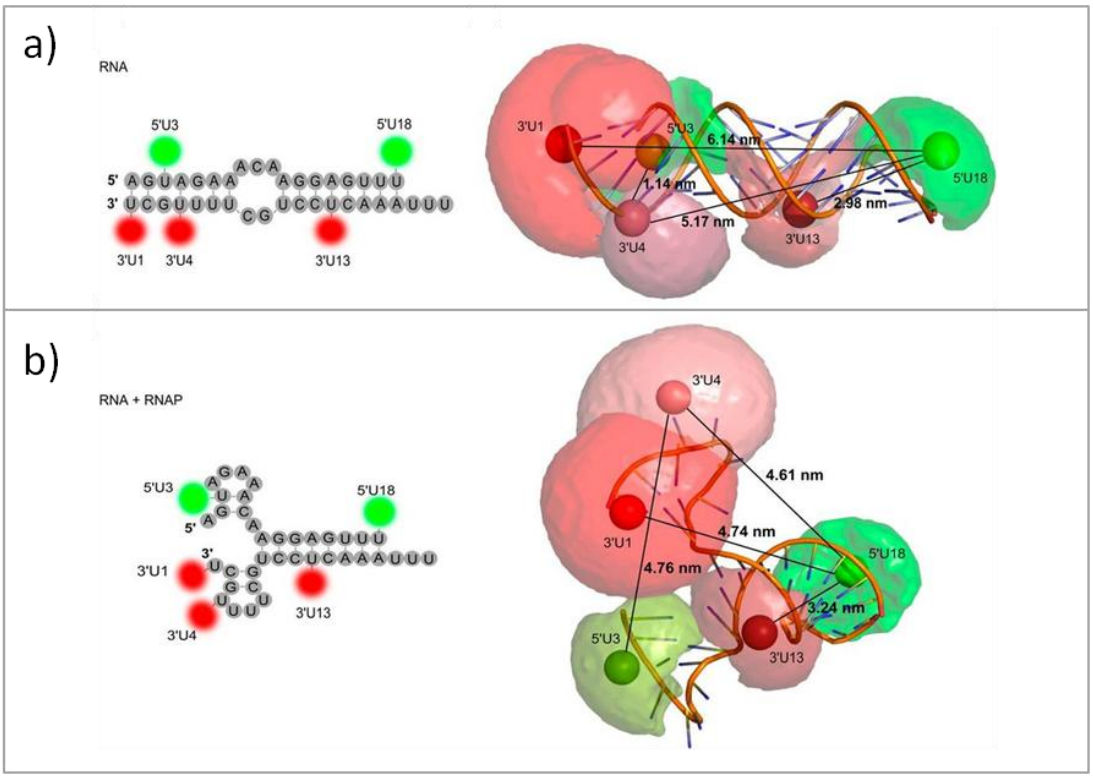


Figure 3.2: A cartoon to show the volumes that fluorescent dyes can occupy bound to ssDNA. The fluorescence dyes are tethered to the biomolecules of interest and are able to rotate. The volumes which these dyes can occupy are shown in the 3D representations on the right, which are clearly larger than the points in space that the 2D representations on the left show. These volumes have to be taken into consideration when measuring distances between dyes, and have been used in this figure by Tomescu *et al.* to determine the structure of a vRNA promoter (a) before and (b) after binding to RNA polymerase.[180]

In order to avoid steric clashing between the bulky dye structure and the protein surface which would potentially prevent efficient labelling of proteins and possibly

affect the activity of the proteins, the functional group used to conjugate the dye to the protein is usually attached to the dye via a flexible carbon chain tether. The particular fluorescent dyes used in this thesis employ a maleimide or NHS anchoring group at the end of a 6 carbon linker, in order to bind to a cysteine residue at the surface of the protein's structure or a primary amine at a modified nucleotide respectively. The volume that the fluorescent dye could possibly occupy is relatively large as a result, shown in Figure 3.2. This allows the dyes to interact with other molecules that also occupy the same volume, such as residues at the surface of other proteins, base stacking with nucleotides and other fluorescent dyes. A scale model of two SsoSSB monomers labelled with Cy5, bound to adjacent sites on ssDNA labelled with Cy3, is presented in Figure 3.3. This model was created by Dr Jose Peregrina and clearly shows that the distance between the Cy3 and Cy5 dyes would be approximately 35-65 Å and would therefore result in an efficient energy transfer from the donor (Cy3) to an acceptor dye (Cy5) wherever the SsoSSB binds to the ssDNA. The Cy3 carbon linker is also long enough for the dye to interact with the protein surface, potentially influencing the molecular conformation of the dye. The Cy5 labels also have linkers that give a large volume that the dye can potentially occupy, allowing the possibility of the dyes experiencing short range interactions with each other and possibly also the protein surfaces. As more proteins bind to ssDNA and lengthen the nucleofilament, the more crowded the molecular environment will become and this will be reflected in the relative fluorescence intensities of the donor and acceptor dyes. This indicates that the fluorescent labelling of SsoSSB is likely to provide a detailed description of how nucleofilaments are formed on ssDNA.

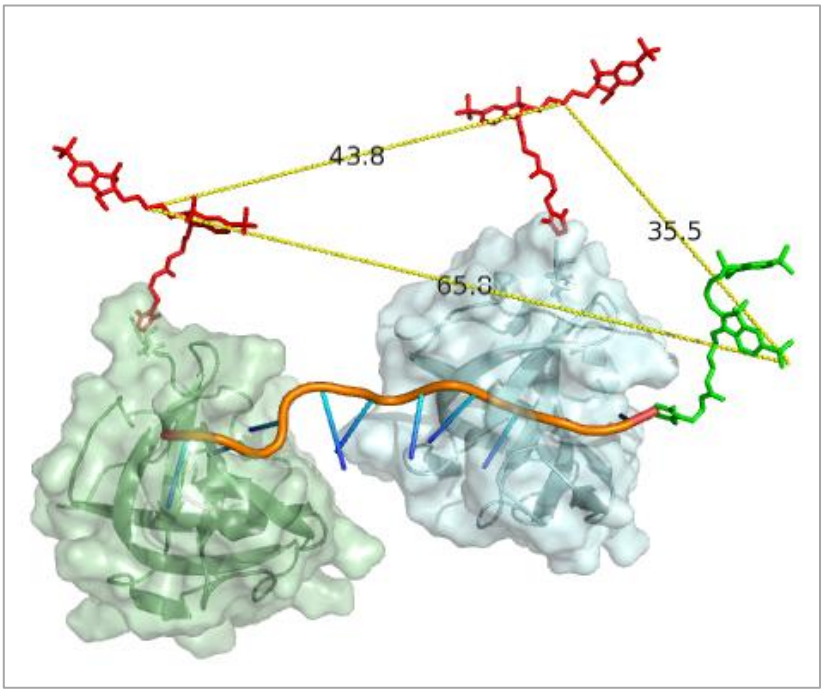


Figure 3.3: Minimum energy model of two SsoSSB monomers, each bound to 10nt ssDNA Cy3 with a gap of a single nucleotide between the two proteins.
The dyes have been introduced to the model using Pymol and the distances between the dyes are labelled as 43.8, 35.5 and 65.8 Å respectively, which give an approximate indication that the distances between them are relatively small. The ssDNA was modelled into the structures of SsoSSB monomers by aligning the protein structures to DBD-A and DBD-B domains from RPA bound to ssDNA.[42] Additional nucleotides were added and a energy minimisation was completed using VMD.[181] Figure adapted from Dr Jose Peregrina.

3.5. ssDNA Labelling

The results from experiments involving multiple interactions between the biomolecules and fluorescence dyes are potentially complicated by heterogeneities and impure samples. In order to minimise these influences, several steps were taken to ensure that homogeneous samples of labelled proteins and ssDNA were obtained. A pure sample of labelled ssDNA was achieved by passing the labelling mixture containing labelled and unlabelled ssDNA through acrylamide denaturing gels. The labelled ssDNA was separated from the unlabelled ssDNA upon identification by UV shadowing and extraction of the appropriate gel band. The extent of labelling was

analysed through UV-vis spectroscopy and was typically found to be in excess of 90%.

3.6. Protein induced fluorescence enhancement

Changes in the local environment of fluorescent dyes are reflected in their quantum yield and therefore the intensity of their fluorescence. The fluorescence intensities of cyanine dyes are dependent on the conformation of the double bonds that link the two aromatic indole rings - and the interconversion between *E* and *Z* conformations can modulate the dye between two fluorescent and non-fluorescent states respectively, shown in Figure 3.4(a).

The viscosity of the medium surrounding the dye influences the energy barrier to rotation of its double bonds and an increase in viscosity can effectively trap the dye in the fluorescent *E* state, and an increase in fluorescence intensity would be observed as a result. The increase in viscosity can be mimicked by the proximity of a protein, and the same increase in fluorescence is observed.[182] This short range interaction can therefore be exploited to report on the presence and position of proteins along a strand of DNA labelled with a cyanine dye such as Cy3. This protein induced fluorescence enhancement (PIFE) has been previously used in a variety of systems and has been developed as an alternative and complementary method to FRET to report on proteins that bind to DNA, as exemplified by BamHI and RecA.[182] The binding of SsoSSB to labelled ssDNA also enhances the fluorescence intensity of Cy3 and Cy5 by a factor of two, as shown in Figure 3.4(b-c), which was comparable to the enhancement observed by Hwang.[182] It has been shown that PIFE

is able to report with higher sensitivity the binding of proteins at distances 1 to 10 nucleotides from the position of the dye. Thus, PIFE complements FRET measurements as a molecular ruler with short range sensitivity. However, in order to observe FRET from the donor Cy3 on ssDNA, an acceptor dye must be present on SsoSSB therefore the protein itself must also be labelled.

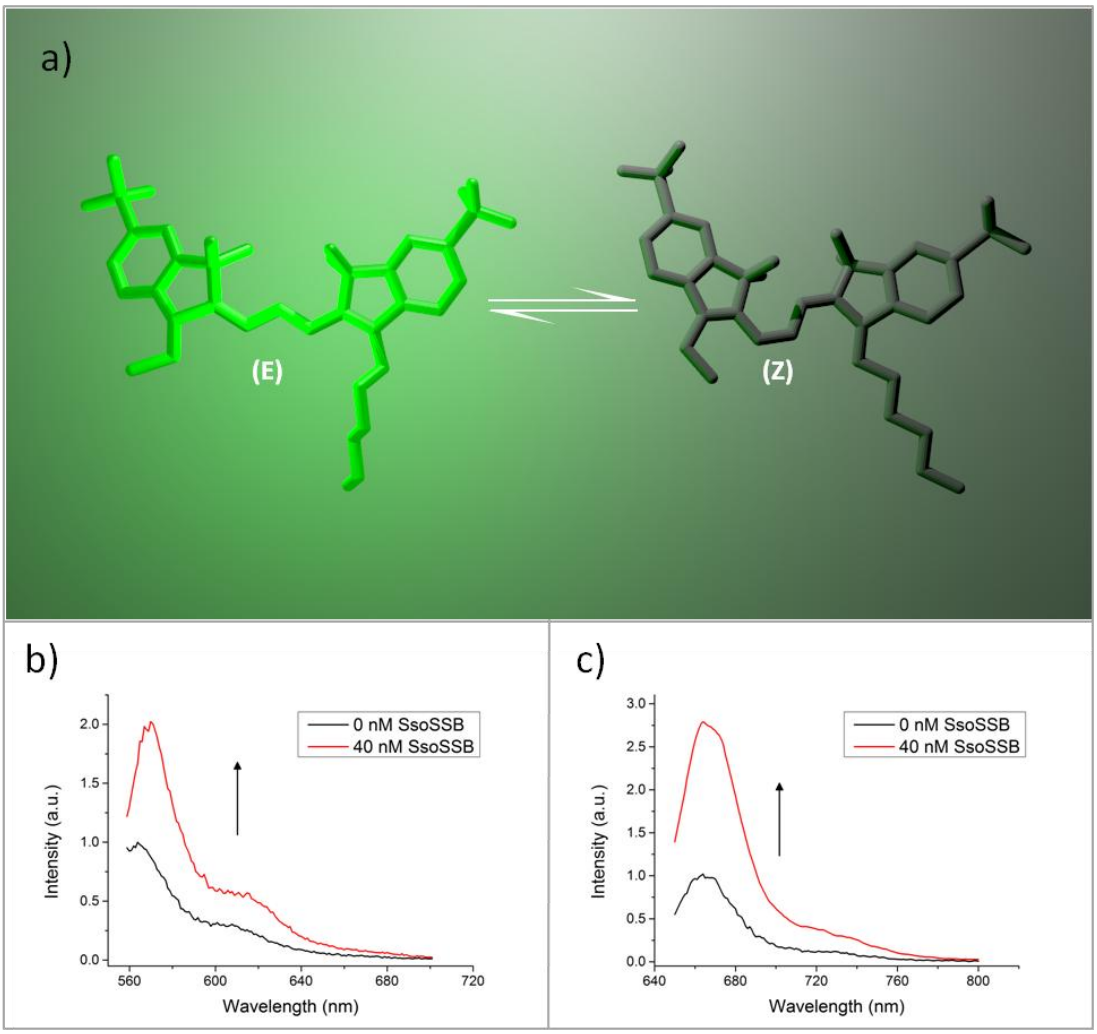


Figure 3.4: The viscosity of the solution is a barrier to the rotation around the double bond in the alkyl chain between the *E* and *Z* conformations cyanine dyes can adopt.
(a) shows the fluorescent and non-fluorescent conformations of Cy3. The proximity of a protein surface can result in a similar barrier to rotation, trapping out the fluorescent *E* conformation. Since there is no more interconversion between *E* and *Z* a protein induced fluorescent enhancement (PIFE) is observed. This is shown in (a) and (b) where 40 nM SsoSSB induces a two to threefold increase in fluorescence to 10 nM ssDNA labelled at the 3' end with Cy3 and Cy5 dyes respectively.

3.7. Protein labelling

In order to ensure high levels of purity were reached with protein samples; the labelling mixture was passed through a heparin affinity column where the labelled and unlabelled proteins were observed as two separate peaks, as seen in Figure 3.5 The labelled protein exhibited a slightly weaker binding to the column and eluted off at a lower NaCl concentration than the unlabelled protein. The SDS-PAGE gel showed that the two peaks produced two identical bands and MALDI-TOF and UV-vis confirmed that all the protein in the labelled fractions was bound to a single fluorescent dye. The efficiency of labelling was increased by adding urea to a final concentration of 8 M to the labelling mixture. Urea is a known reagent used for denaturing and unfolding protein structures, which was used in this case to further expose the cysteine residues during the labelling reaction. The increase in the yield of the labelling reaction can be viewed qualitatively in the ratio of labelled and unlabelled UV-vis peaks of Figure 3.5. The respective fractions collected from the affinity column were again analysed by SDS-PAGE electrophoresis, UV-vis spectroscopy and MALDI-TOF mass spectrometry and a sample of approximately 100% labelled protein was obtained.

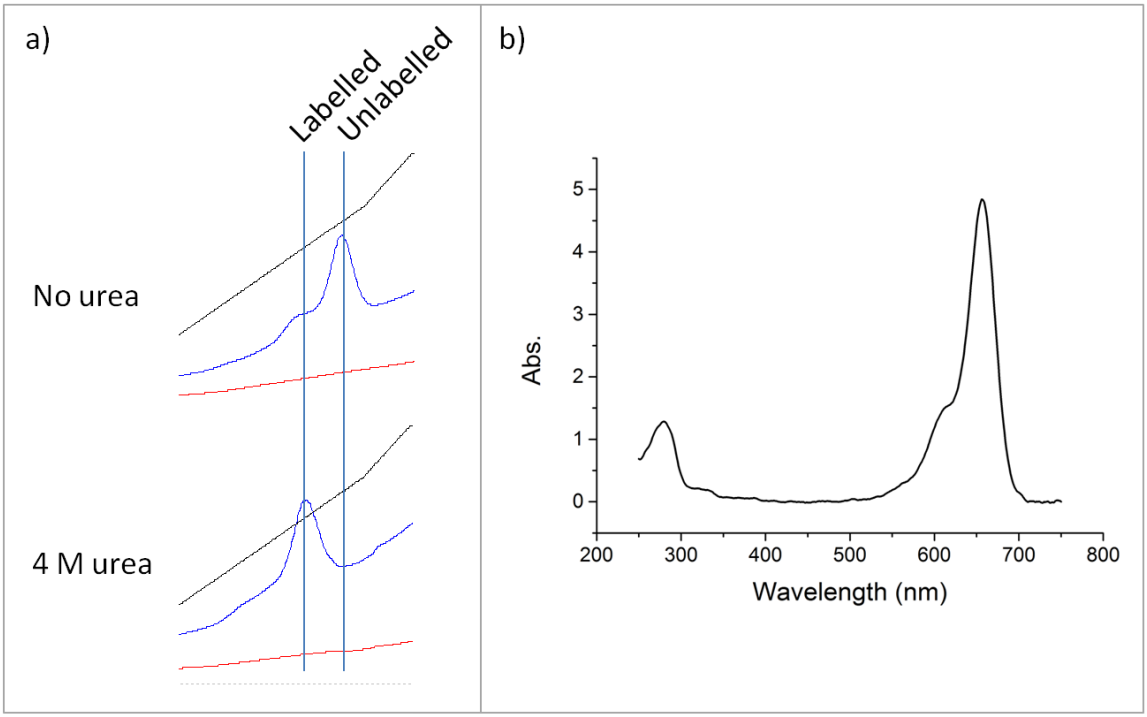


Figure 3.5: Labelling of SsoSSB with Alexa 647 in the presence of 8 M urea increases the efficiency of the labelling reaction.
The UV-vis absorption of the fractions at 280 nm is shown in (a) as the the blue line. The labelled and unlabelled SsoSSB monomers elute from the heparin column at different salt concentrations (black line) and can be resolved by tracking the UV-vis absorption at 280 nm. The UV-vis absorption spectrum of the labelled SsoSSB is shown in (b), where the relative peak heights show that the dye and SsoSSB monomers are present at a 1:1 ratio. Note that the urea concentration was diluted to 4 M before passed through the column, as stated in materials and methods.

3.8. FRET

The experimental details and theory concerning FRET experiments have been discussed in the introduction, and FRET has been employed in many published experiments to investigate the binding of SSBs and similar proteins to ssDNA. The results presented in this chapter show ensemble experiments that confirm that the Alexa 647 dye bound to SsoSSB can be indirectly excited via an energy transfer from an excited Cy3 dye conjugated to ssDNA. This confirmed that FRET is an appropriate method of detecting SsoSSB bound to ssDNA, and was also observed during single molecule experiments.

In the absence of an acceptor, the emission spectrum of Cy3 appears as a single peak at 565 nm with a broad shoulder at 615 nm as shown by the black line in Figure 3.6(a). The spectrum of Cy3 in the proximity of Alexa 647 showed an additional peak in the red line in Figure 3.6(b) that was centred at 665 nm, which corresponded to the sensitised emission of the acceptor dye. The FRET efficiency was calculated by the $ratio_A$, the ratio of the integral of this additional peak and the integral of the Alexa 647 directly excited at 640 nm, as described in the materials and methods.

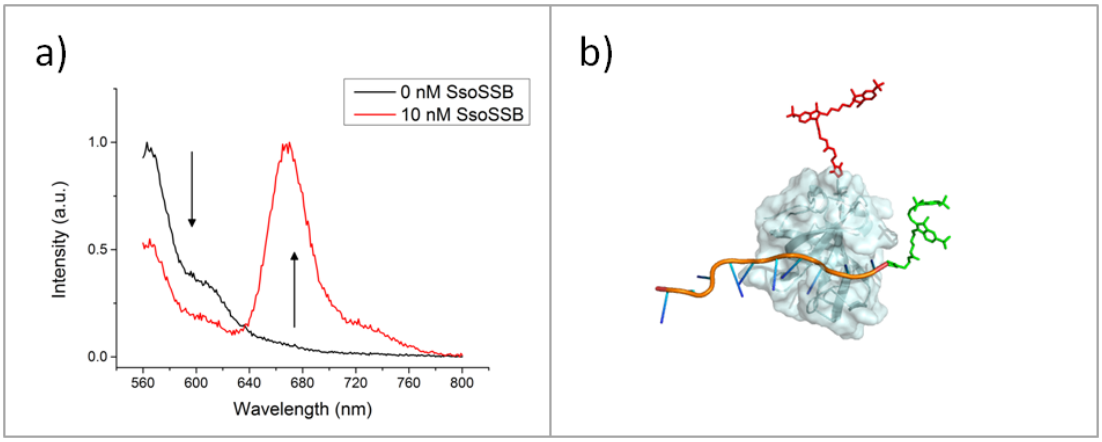


Figure 3.6: FRET can be used to track SsoSSB binding to ssDNA by using ensemble spectroscopy. The respective decrease and increase of the peaks at 565 nm and 665 nm before and after ssDNA Cy3 was introduced to SsoSSB Alexa647, is shown in (a). The SsoSSB binding to the ssDNA and the energy transfer from the Cy3 dye (green) to the Alexa 647 dye (red) can be visualised in the cartoon in (b).

3.9. Quenching

In order to fully characterise the fluorescent properties of the labelled proteins, labelled SsoSSB was introduced to ssDNA without the donor dye. From section 3.4 it is clear that two labelled SsoSSB monomers bound adjacent to one another also place their fluorescent dyes in close proximity to one another. This was observed as a quenching of the fluorescence from the Alexa 647 and is discussed in more detail in paragraph 3.14. The quenching provides another method of tracking SsoSSB binding

to ssDNA in ensembles; however it is difficult to utilise this in single molecule experiments since this quenching would resemble reversible blinking events and would also confuse bleaching experiments that are used to count monomers bound to ssDNA.

3.10. PELDOR

Although the crystal structure of SsoSSB heavily implies that it exists as a monomer in solution, there is still a dichotomy in the published literature as to whether the protein predominately exists as a monomer or a tetramer. In order to address this, the variant A114C (the A114 residue is situated on the surface of the protein at edge of OB fold) was labelled with the spin label MTSSL and purified by size exclusion. The masses of unlabelled and MTSSL labelled SsoSSB were investigated by Dr Catherine Botting and her group, using MALDI-TOF and the extent of labelling was verified to be approximately 100%, shown in Figure 3.7. The interaction between spin labels was then analysed by PELDOR spectroscopy, carried out by Dr Bela Bode.

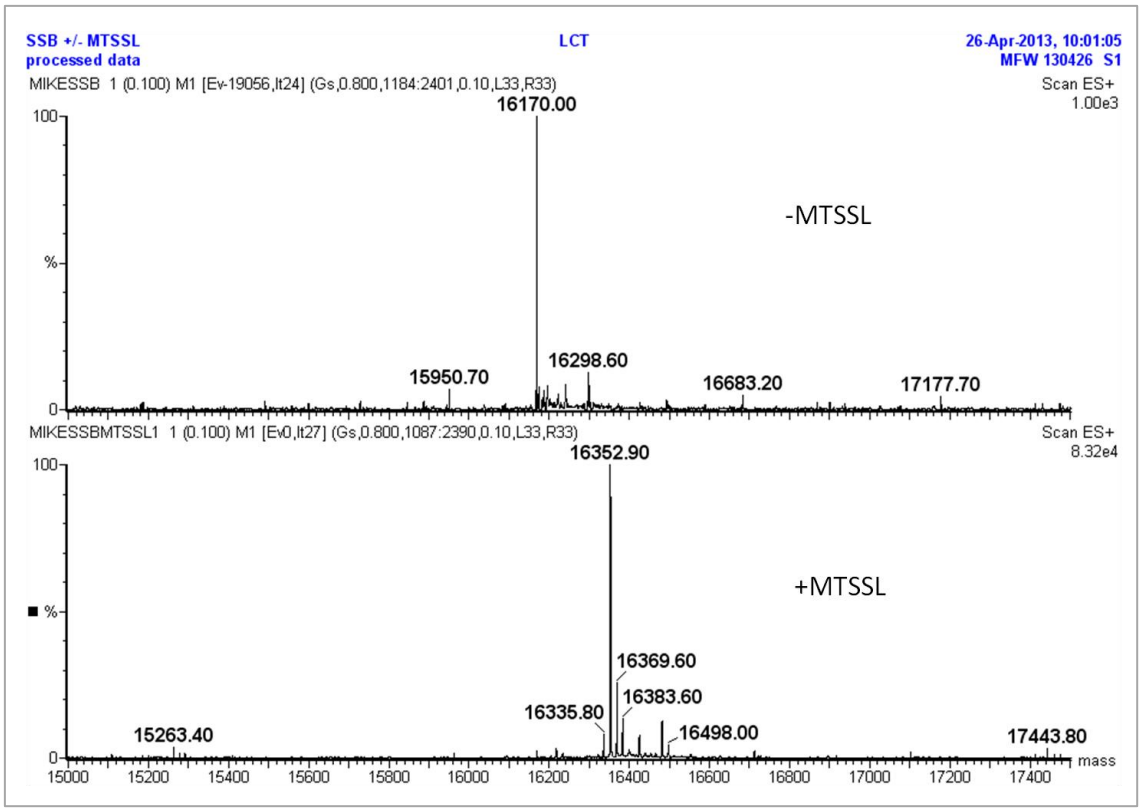


Figure 3.7: MALDI-TOF analysis of A114C SsoSSB before (top) and after (bottom) labelling with spin label MTSSL.
The peak on the bottom has shifted by 182 Da, corresponding to the molecular weight of a single spin label.

Pulsed electron-electron double resonance (PELDOR) is a pulsed electron paramagnetic resonance (EPR) spectroscopic technique that is generally used to measure the interaction between unpaired electrons more than 2 nm apart.[183, 184] Unpaired electrons are rarely found in native proteins and spin labels are generally conjugated to specific amino acids, much like fluorescent labelling. Upon applying pulses of microwaves to the sample the spin state of the unpaired electrons can be switched, and the relaxation back to its original state is measured as a spin echo in a PELDOR experiment. Dipolar coupling to a second unpaired electron is represented by a modulation and an increase in the depth of the exponential decay of the echo intensity compared to a lone unpaired electron. These effects are dependent on the

distance between the two paramagnetic centres and their relative orientation, analogous to the FRET observed between two fluorescent dyes. Previous distances have been used to investigate the structure of Influenza A Virus NS1, the structure of which is shown Figure 3.8(a) with possible conformations of the spin label shown as red and blue balls.[183] The interaction between the two labels is shown in Figure 3.8(b) as an exponential decay that exhibits a sigmoidal modulation. This can be used to calculate the distance distribution between the two labels, shown in the lower panel of Figure 3.8(b), which can be compared to calculations made *in silico*.[183]

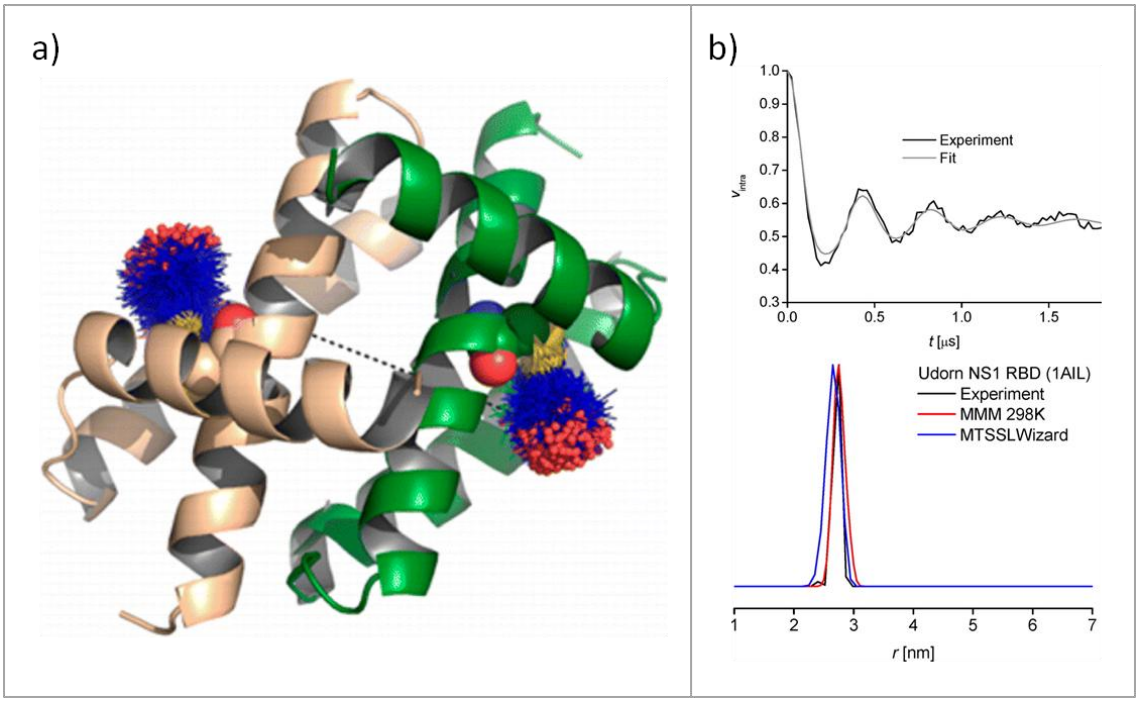


Figure 3.8: PELDOR can be used to verify the structure of proteins such as Influenza A Virus NS1. (a) MTSL Wizard is able to map the conformations of the spin label (red and blue cones) that are bound to specific residues on the protein surface. (b) shows the modulation of the spin echo produced through the dipolar coupling of the two paramagnetic species (top) and the comparison of the distances measured experimentally and those calculated from the *in silico*model. Figure modified from Kerry *et al*.[183]

The spin labelled SsoSSB in solution did not produce any modulation in the absence of ssDNA, as shown in Figure 3.9(a), which indicates that SsoSSB predominately

exists as independent monomers in solution without forming any other higher order structures at a concentration of 100 μM. The depth of the decay increased upon addition of 50 μM ssDNA was consistent with a broad distribution of distances between spin labels centred around 3-4 nm. This increase in depth observed upon the introduction of ssDNA also supports the possibility that a single SsoSSB monomer only comes into proximity with other monomers once they are bound to the same strand of ssDNA. The distance calculated by PELDOR roughly agrees with the distances predicted by the model of two SsoSSB monomers at adjacent binding sites on ssDNA (4 nm) but the broad distribution does not imply that there was an overwhelming preference for SsoSSB monomers to bind snugly next to each other; however further investigations are necessary to confirm this conclusion as there are many factors that could produce this broadening other than low cooperative binding to ssDNA.

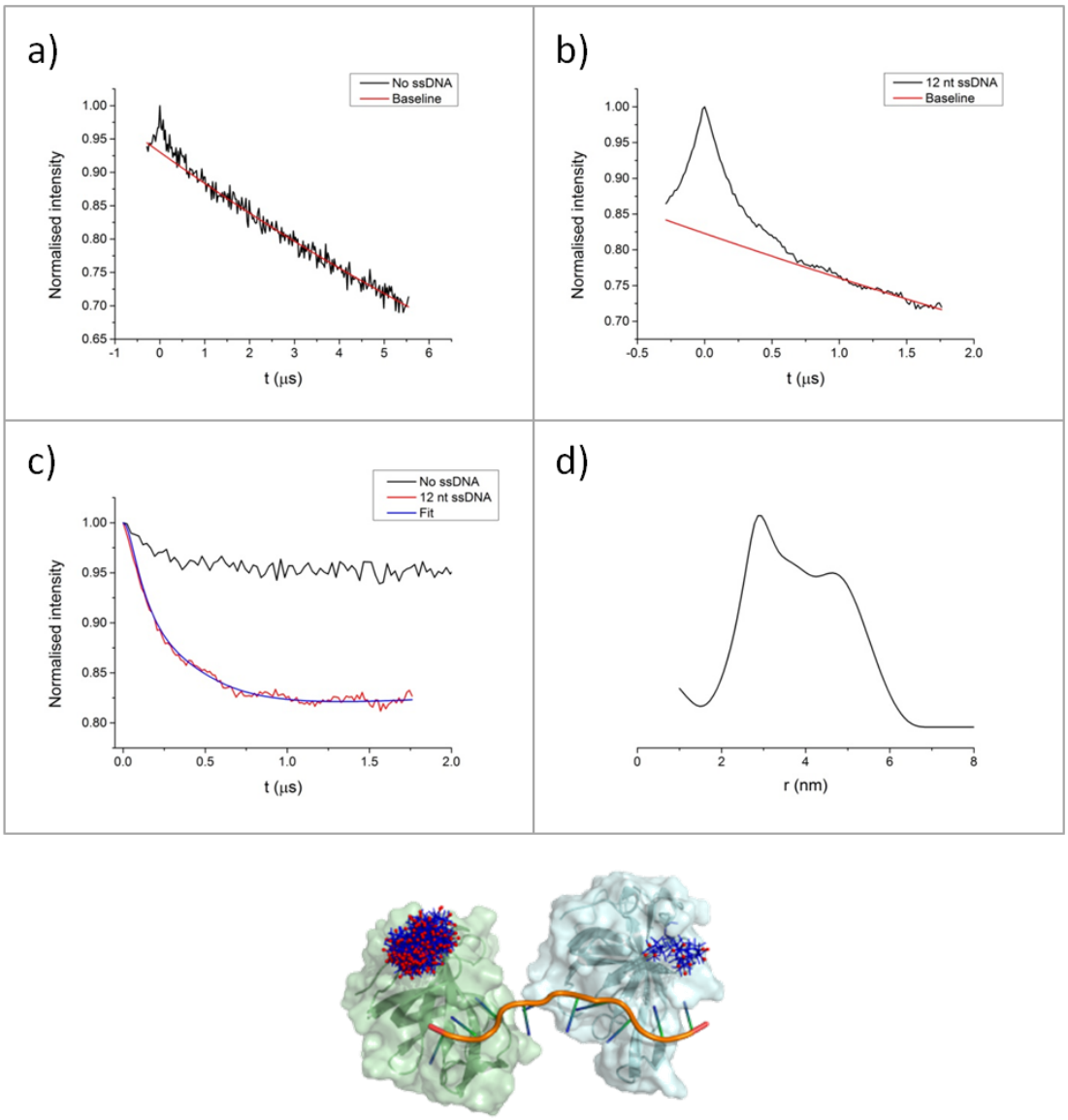


Figure 3.9: PELDOR data from SsoSSB MTSL experiments.
(a) and (b) show the uncorrected data before and after SsoSSB MTSL was introduced to 12C ssDNA. (c) demonstrates the increase in the depth of modulation observed when ssDNA was added to SsoSSB MTSL. The depth of modulation from SsoSSB alone suggested that 90% of the species under observation were monomers, while the data from the addition of ssDNA was consistent with the formation of dimers that made up approximately 40% of the SsoSSB in solution. This was translated as a broad distribution of distances centred around 3-4 nm, as shown in (d). Below shows the MTSSL wizard applied to the model used in section 3.4, which gives a mean distance between the possible configurations of MTSSL dyes of 4.1 nm.

3.11. Persistence length of naked and decorated ssDNA

The shape of a nucleofilament is largely dependent on the quaternary structure of a SSB, and the binding mode that is being employed by the proteins. The end-to-end distance of ssDNA bound to proteins is often an indication to what binding mode a protein employs. For example, the change in distance between the 3' and 5' ends of ssDNA has been used to follow the $(SSB)_{65}$ and $(SSB)_{35}$ binding modes of EcoSSB in single molecule FRET experiments.[155] The conformation of ssDNA has been studied in solution and comparisons to Monte Carlo simulations *in silico* by Chen *et al.* have shown that the wormlike chain model is a suitable model to describe single stranded nucleic acids.[14] This means that the end-to-end distance can also be used to quantitatively measure the degree of flexibility of ssDNA.

The end-to-end distance of a 39 nt ssDNA was calculated from the FRET observed from 3' and 5' terminal ends labelled with Cy3 and Cy5 dyes respectively, shown in Figure 3.10. The sensitised emission decreased upon addition of SsoSSB, which was shown by the ensemble fluorescence spectra in (a). The $ratio_A$ was calculated from the spectra and the average values from three titrations are shown in Figure 3.10(b), which exhibited a decrease in $ratio_A$ as SsoSSB was added to the ssDNA in solution. This corresponded to an increase in the distance between the two dyes and the $Ratio_A$, the FRET value, and end-to-end distance were used to calculate the persistence length of the polymer, which in turn described the stiffness of the ssDNA, effectively measuring the length of the rigid links in the chain.

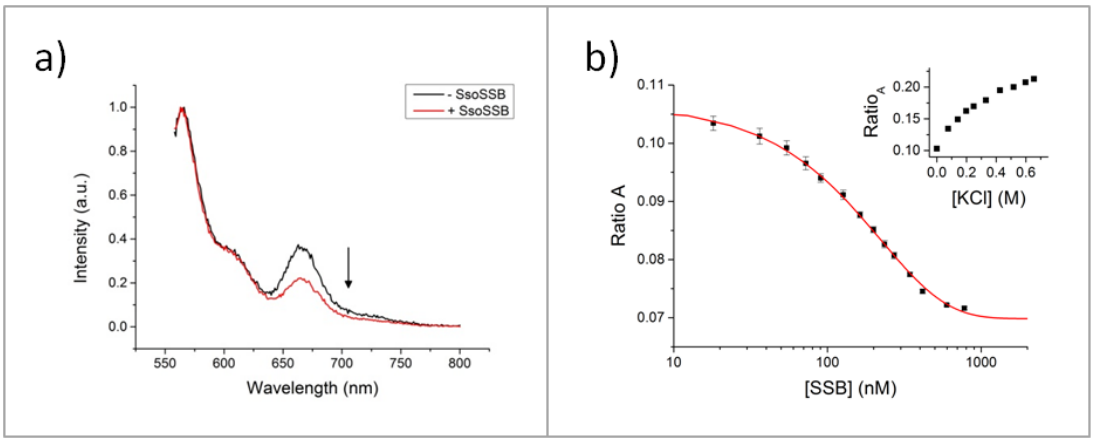


Figure 3.10: Ensemble FRET spectra of 50 nM 39 nt ssDNA Cy3 Cy5 in the absence and presence of SsoSSB showing a decrease in ratio$_A$, that is consistent with the ssDNA being straightened and increasing the end-to-end distance.
(a) Shows reduced intensity of the peak representing the sensitised emission of the acceptor dye. The ratio$_A$ was measured as increasing concentrations of SsoSSB was introduced to the ssDNA, which was repeated in triplicate, and the mean values are plotted in (b) with the standard error as error bars. The red line holds no significance and is meant to help guide the eye. The inset shows the ratio$_A$ changes as KCl is titrated into the same oligo.

The average contour length of a single nucleotide from crystal structures has been shown to be 0.63 nm, so if a 39 nt ssDNA was perfectly rigid and straight, its end-to-end distance would be approximately 24.5 nm, which would also match its persistence length. On the other hand, if each nucleotide had complete rotational freedom from its neighbours the persistence length would equal half the length of a single nucleotide. In reality, ssDNA is capable of adopting compact conformations but must also minimise like charge interactions along the phosphate sugar backbone, which prevents the polymer from completely compacting.

For the 39 nt ssDNA, the persistence length was calculated as 1.76 nm (2.79 nt) and 1.01 nm (1.63 nt) in 0 mM and 650 mM KCl concentrations respectively as control experiments for the SsoSSB titrations. The ratio$_A$ values are shown in the inset of Figure 3.10. This correlates to an end-to-end distance of 9.16 nm and 7.16 nm respectively. This agreed with the persistence lengths of other oligonucleotides in the

literature and confirmed that this was a reliable technique to use to investigate the persistence length of nucleofilaments.[15] SsoSSB was titrated against ssDNA in a 100 mM KCl solution, and an increase in end-to-end distance and persistence length was observed to give a final end-to-end distance of 11.38 nm and persistence length 2.67 nm (4.24 nt).

The distances calculated from the SsoSSB nucleofilaments seem to suggest a linear arrangement of SsoSSB monomers on ssDNA. In solution ssDNA naturally curls up on itself but still maintains at least some directionality, as shown by the middle cartoon in Figure 3.11(a) rather than the rigid linear structure shown on top. The ssDNA appeared to be straightened out by the binding of SsoSSB but retained a degree of flexibility seen in the bottom cartoon. This is in contrast to the globular assembly observed from similar experiments with EcoSSB, which wraps the ssDNA around itself, shown in Figure 3.11(b). This produced an increase in FRET as EcoSSB bound to ssDNA brings the two ends nearer to each other in both the $(SSB)_{35}$ and $(SSB)_{65}$ binding modes. The SsoSSB nucleofilament does not appear to be a rigid rod but keeps a degree of flexibility between monomers, rather like a beaded chain in which length of each link in the chain is one SSB monomer in length, possibly similar to RPA where each protein holds the ssDNA in a U-shaped conformation, shown in Figure 3.11(c) but produces an overall beaded nucleofilament due to the low cooperativity of binding between each protein.

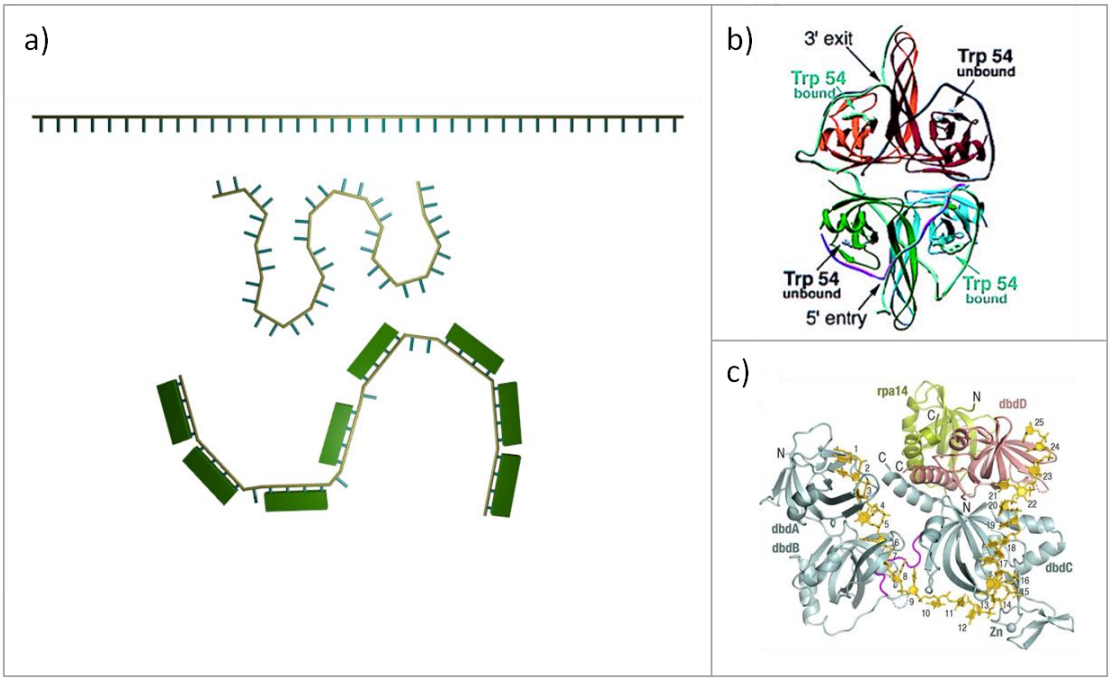


Figure 3.11: From the data from PELDOR and end to distance measurements this suggests that SsoSSB binds as a monomer and produces a relatively unordered nucleofilament.
This is shown in (a), where cartoons show the difference between perfectly rigid ssDNA (top), a collapsed naked ssDNA with persistence length approximately 2 nt (middle) and a strand of ssDNA decorated with SsoSSB with a persistence length of approximately 4 nt (bottom). The binding of SsoSSB to at ssDNA sequence of approximately 40 nt can be compared to EcoSSB and RPA binding to similar lengths of ssDNA in (b) and (c) respectively. Figures adapted from Fan et al., Raghunathan *et al.* (pdb: 1jmc.)[30, 32, 42]

3.12. Dissociation constants and occupancy

The affinity of SsoSSB for ssDNA has previously been shown to be in the low nanomolar range by isothermal calorimetery (ITC), fluorescence anisotropy, BioLayer Interferometry and quenching of the tryptophan fluorescence.[27, 109, 114, 185] The dissociation constants presented in the literature were generally calculated through a 1:1 binding model shown in materials and methods. The stoichiometry of SsoSSB binding to ssDNA has also been investigated and one monomer has been found to occupy approximately 4-5 nt,[109] drawing similarities to the RPA crystal structure which showed two OB folds in DBD-A and DBD-B bound to 10 nt. The

NMR structure of SsoSSB bound to ssDNA indicates that the aromatic residues stack with 3 nt, that are spread over a region of 4 nt.[114] Since the aromatic residues are located in the middle of the OB fold, it is unlikely that monomers in a nucleofilament would only occupy 4 nt each.[113] The model of two SsoSSB monomers aligned with the RPA does suggest that there could be steric clashing between neighbouring proteins that only bind to 4 nt, therefore it seems intuitive that a slightly larger stoichiometry would be more realistic.

The quenching of tryptophan has been used in the literature to report on the stoichiometry of SsoSSB binding to the oligonucleotides and was used to confirm binding by SsoSSB alongside the PIFE, FRET and Alexa 647 quenching assays. The concentration of SsoSSB used was a least an order of magnitude above the dissociation constant in order to achieve efficient binding of SsoSSB to the oligos. The fluorescence emission from the tryptophan residues taken from a titration where ssDNA was added to SsoSSB is shown in Figure 3.12(a), and the average degree of quenching from three experiments was plotted against total ssDNA concentration in (b), showing that 80% of the initial tryptophan emission was quenched by SsoSSB binding to ssDNA. These titrations can be plotted against ssDNA concentration in terms of nt per SsoSSB monomer in Figure 3.12(c), which shows that the stoichiometry is approximately 5 nt and independent of the length of ssDNA, consistent with other values in the literature.

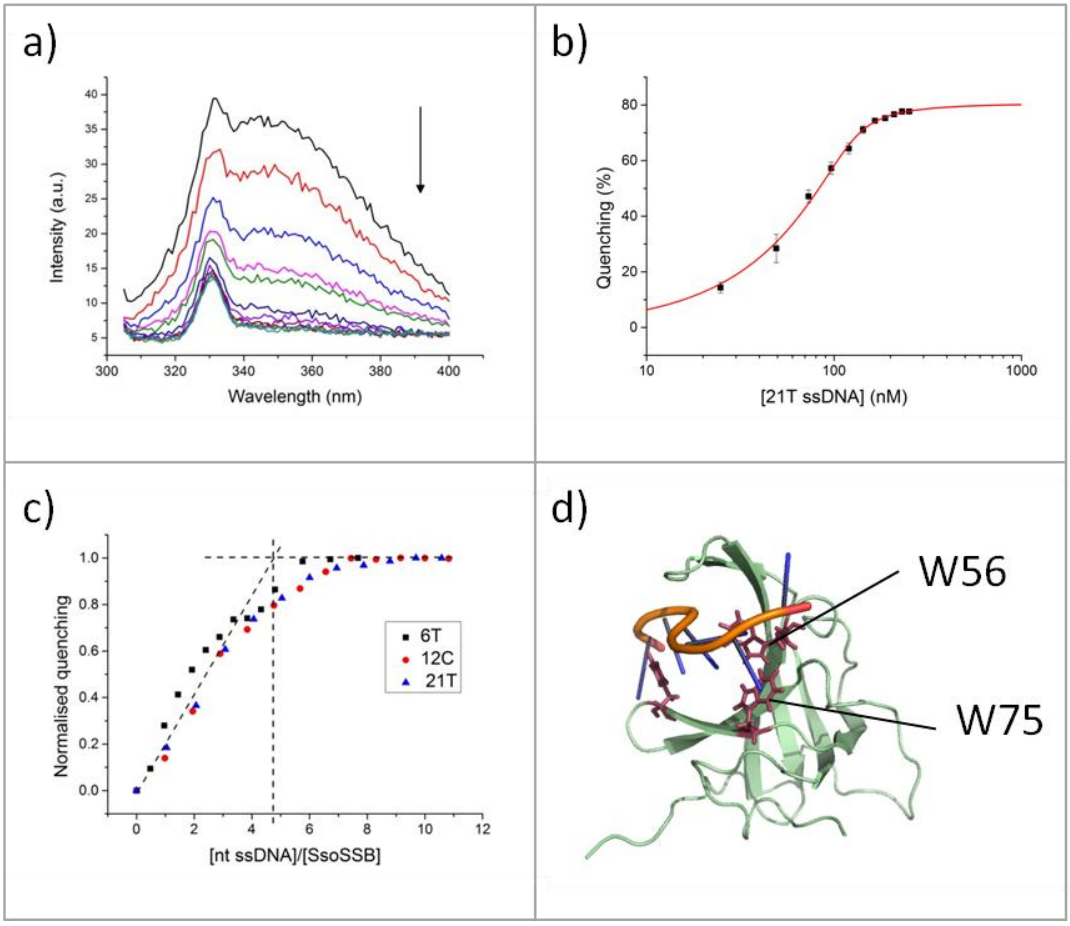


Figure 3.12: Tryptophan quenching due to SsoSSB binding to ssDNA.
An example of a titration where 21T ssDNA is added to 500 nM SsoSSB unlabelled wildtype where the quenching of fluorescence from the ^1L_a can be visualised by the decrease in fluorescence at 350 nm (λ_{ex} = 295 nm) shown in (a). The fluorescence from the ^1L_b state can be seen at 330 nm which remains relatively constant as ssDNA occupies the OB fold and excludes water moelcules. The quenching can be plotted against the total concentration of ssDNA, as shown for 21T ssDNA in (b). This was repeated in triplicate and the mean values were plotted with the standard errors displayed as error bars. The quenching can also be plotted as nucleotides per SsoSSB monomer, and was repeated for 12C and 6T ssDNA. The mean values are plotted in (c) with the error bars omitted for clarity. The stoichiometry remains approximately 5 nt ssDNA per SsoSSB monomer for all three lengths of ssDNA. (d) shows the NMR structure from Gamsjaeger *et al.* showing the three aromatic residues as stick models, base stacking with three nucleotides of ssDNA.[114] The tryptophans that fluoresce in the absence of ssDNA are labelled as W56 and W75.

Approximate dissociation constants for SsoSSB can be calculated for the various lengths of ssDNA from 1:1 binding models fitted to the average quenching values from the titrations, and are shown in Table 3.1, which show a modest decrease as the length of the ssDNA grows to accommodate up to four monomers on the same strand of ssDNA. The difference between dissociation constants from nucleofilaments of

one monomer and four monomers is over an order of magnitude, which is comparable to the cooperativity of RPA but these experiments alone do not necessarily confirm a cooperative binding mechanism between SsoSSB monomers. Further investigation is required to quantify this cooperativity, but it is highly likely given the reports in the literature and this evidence that SsoSSB only demonstrates a limited cooperativity compared to proteins such as EcoSSB or bacteriophage T4 gene 32 protein, whose cooperativity factors can be multiple orders of magnitude.[186]

Table 3.1: Dissociation constants from tryptophan quenching experiments with increasing lengths of ssDNA that can respectively allow up to 1, 2 and 4 monomers to bind to them at any one time. Dissociation constants were calculated from a modified one to one binding model.

ssDNA	Kd (nM)
6T	91 ± 4.0
12C	15 ± 3.0
21T	5.2 ± 0.7

As a fluorophore, tryptophan does not exhibit a quantum yield high enough for it to be used at low nanomolar concentrations, making it impractical to use at concentrations around SsoSSB's dissociation constant and also at a single molecule level. As a result, alternative methods were used that exploit artificial dyes whose fluorescent intensities are more suitable for applications at lower concentrations.

Titrations of unlabelled SsoSSB against ssDNA Cy3 were conducted in order to confirm that PIFE was a suitable method for following the binding of SsoSSB. An approximate two fold enhancement of 10 nM 12C ssDNA Cy3 was observed up to a

concentration of 20 nM SsoSSB, indicating that two monomers were able to persist on 12 nt ssDNA, agreeing with the tryptophan quenching experiments shown in Figure 3.12(c). Enhancement of cyanine dyes are reported to be a short range phenomenon, as the protein surface has to be close enough to exert a steric hindrance to the rotation of the indole groups.[182] In the literature this has been seen up to 12 base pairs away, and the spectra in Figure 3.13 seem to suggest that a protein binding to C12 ssDNA was sufficiently close to the dye to influence its local environment and produce a two fold enhancement. The intensity of Cy3 fluorescence allowed the concentration of ssDNA to be reduced down to approximately 10 nM, which was a more appropriate concentration to be investigating dissociation constants. From Figure 3.13(b) a dissociation constant of 8.1 ± 0.2 nM was calculated for two SsoSSB monomers binding to ssDNA which agrees with the previously calculated K_d values.

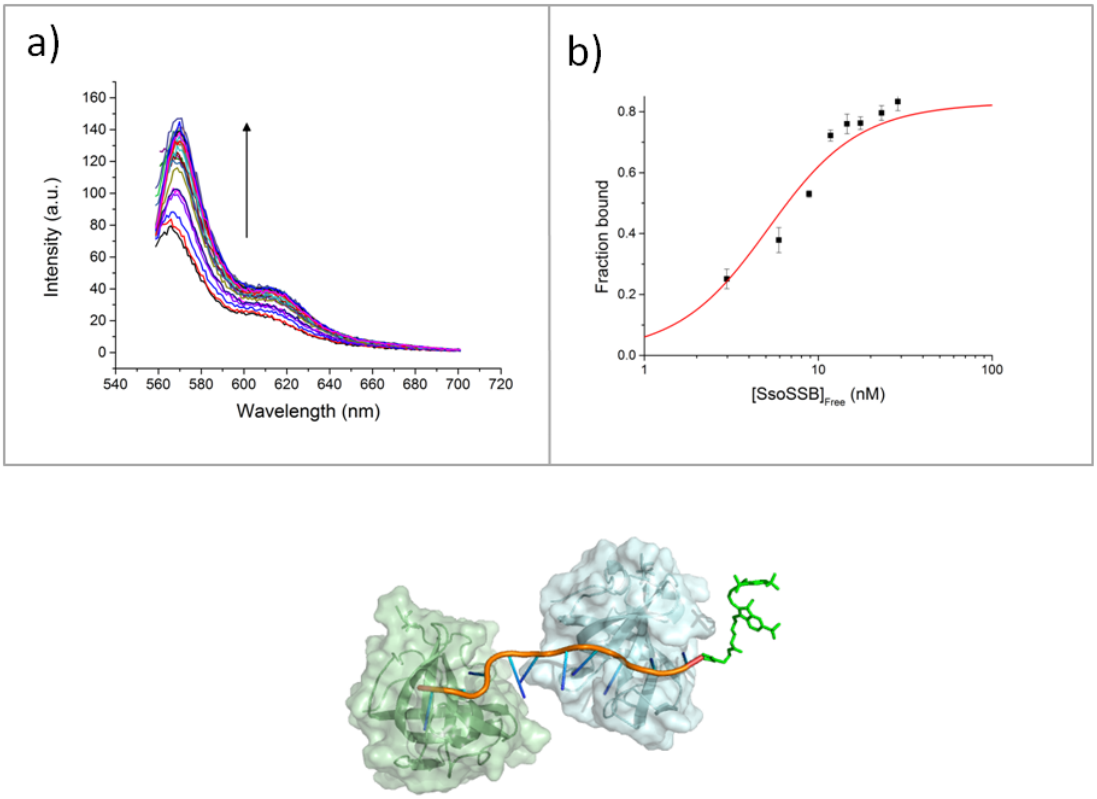


Figure 3.13: Enhancement of the fluorescence from 10 nM ssDNA Cy3 was measured at different concentrations of unlabelled SsoSSB, and a two fold increase in the intensity was observed.
The fluorescence emission spectra are presented in (a) and the intensities were used to calculate the fraction of nucleotides bound to SsoSSB monomers, with the arrow representing increasing SSB concentration.The results were completed in triplicate and the mean values were plotted in (b) with the standard error as error bars. This was fitted to a one-site binding curve and a K_d = 8.1 ± 0.2 nM was calculated. A cartoon is shown below of two SsoSSB monomers bound to ssDNA Cy3.

By adding a fluorophore to the SsoSSB that was capable of accepting energy from the Cy3, the possibility that FRET could be used to track binding was explored. FRET was observed from a 10 nM C12 ssDNA Cy3, as shown in Figure 3.14(a), but the curve plotting the $ratio_A$ in (b) clearly did not agree with the observed occupancy from the literature or the tryptophan and PIFE experiments. The $ratio_A$ plateaued significantly before the expected 2:1 ratio of SsoSSB monomers to strands of ssDNA, therefore the fluorescence of the acceptor was investigated in more detail in order to try and explain this observation.

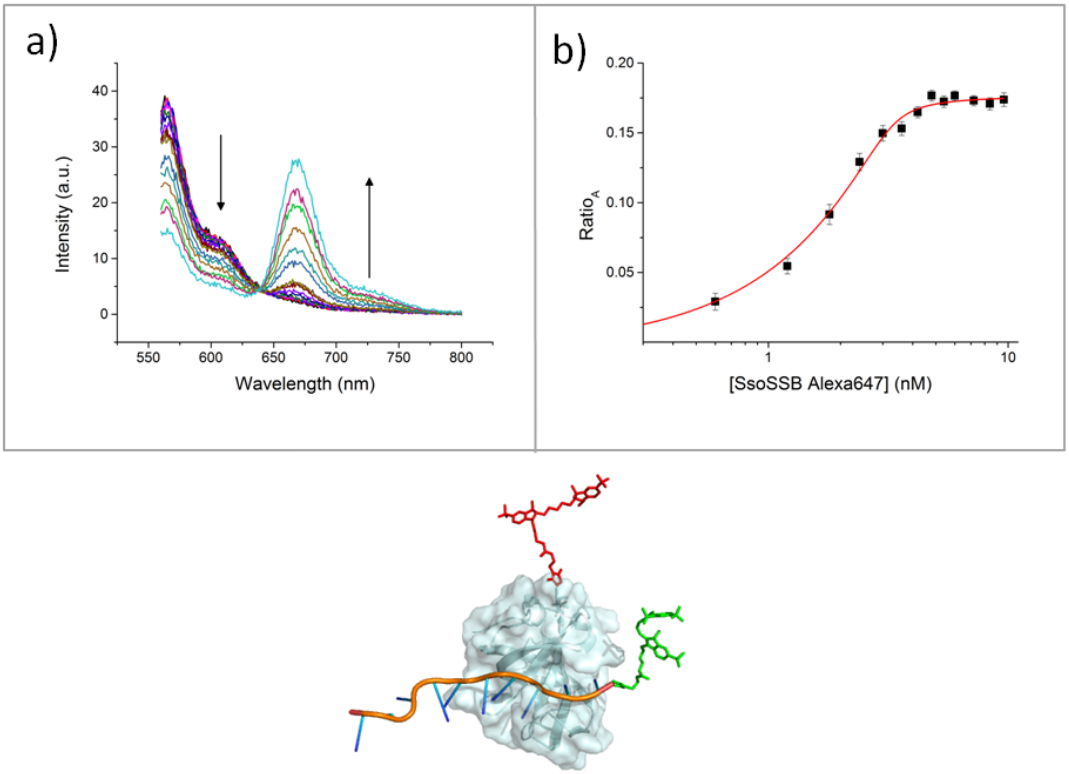


Figure 3.14: Ensemble FRET titrations were also completed adding a SsoSSB Alexa 647 to 10 nM ssDNA Cy3, as shown in the cartoon.
The spectra from a typical titration are plotted in (a), and the mean ratio$_A$ values from triplicate experiments are plotted in (b) with the standard error plotted as error bars. The data was plotted to one to one binding site model and an apparent K_d of 1.5 ± 0.3 nM was calculated. The data suggested that only approximately a tenth of the ssDNA could accommodate SsoSSB, which did not agree with the tryptophan and PIFE titrations, where higher concentrations of SsoSSB resulted in the SsoSSB completely decorating the ssDNA. A cartoon is shown below of an SsoSSB Alexa 647 monomer bound to ssDNA Cy3.

In section 3.9 the quenching of Alexa 647 was described when SsoSSB was bound to a length of ssDNA that was long enough to accommodate up to two monomers. This interaction was also exploited in order to define an occupancy and dissociation constant. The results from the titration of ssDNA to SsoSSB Alexa 647 are presented in Figure 3.15, and again they indicated that two monomers could persist on 12C ssDNA with a dissociation constant of 10.5 ± 0.3 nM. This also demonstrates that the presence of the dyes in the nucleofilament did not significantly affect the binding of

SsoSSB, since tryptophan quenching, PIFE and Alexa 647 quenching all produced similar dissociation constants.

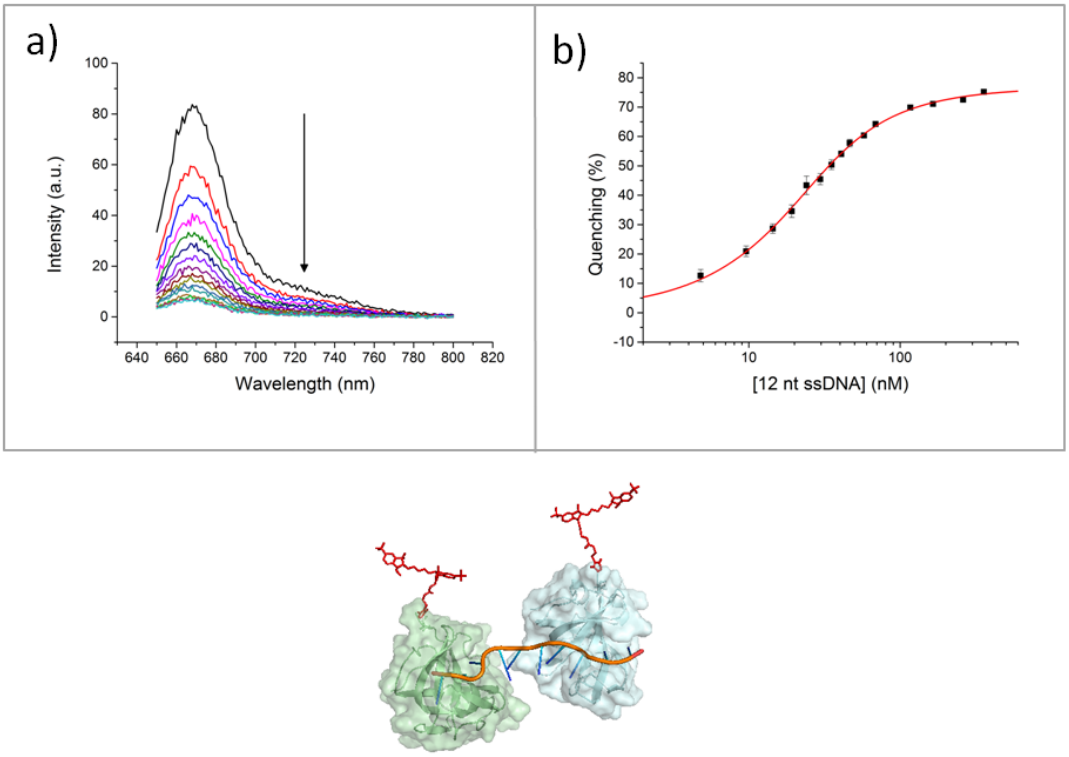


Figure 3.15: A reverse titration adding ssDNA to 50 nM SsoSSB Alexa 647 was completed in triplicate.
The mean values are plotted in (b) with the standard error. The spectra from a typical experiment is plotted in (a). This was fitted to one to one binding model and a K_d = 11 ± 0.3 nM was calculated, and a total of two SsoSSB monomers were able to bind per strand ssDNA, as shown in the cartoon.

3.13. Cooperativity parameters

Since PIFE reports on binding to ssDNA at low nanomolar concentrations, it was used in ensemble experiments to explore to what extent SSB binding to ssDNA is a cooperative process, as shown in Figure 3.16. The experiments used ssDNA strands that were a maximum of 20 nt in length, with a binding site set at 5 nt. A model that accounts for binding to a finite lattice rather than an infinite lattice was used. The free protein concentration *L* was calculated from the fraction of the total

enhancement of fluorescence, and the reciprocal of the dissociation constant from the SsoSSB tryptophan quenching in the presence of 6T ssDNA was used as K. The short length of the ssDNA negated any cooperative effects of binding and this finite lattice model gave a dissociation constant of 90.4 ± 10.9 nM, agreeing with the 1:1 binding model used in section 3.12. A dissociation constant of 90 nM was therefore used following the method described in materials and methods, in order to obtain a value of 100 ± 16 for the cooperativity factor, ω, which describes a modest cooperative binding.

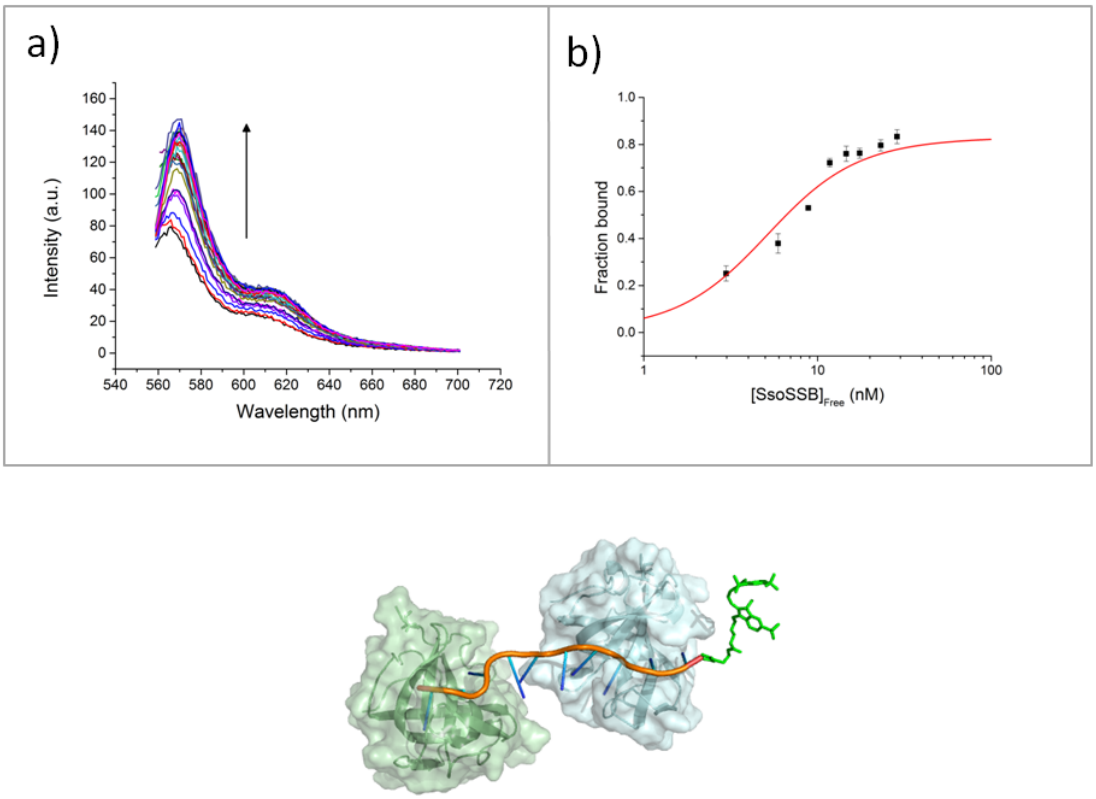


Figure 3.16: PIFE titrations could also be fitted to Epstein's model that describes the cooperative effects of SsoSSB binding to adjacent sites along ssDNA, taking into account overlapping binding sites and the end effects of a short oligo.
The spectra from the fluroescence scans are shown in (a) and the mean values are plotted in (b) with the standard errors shown as error bars. The cooperative parameter to the equilibrium constant was found to be 100 ± 16, which indicated a modest positive cooperativity.

3.14. Quenching suggests monomeric structure

The ssDNA was shortened to 6T which could only support a single SsoSSB Alexa 647 monomer, and the quenching of Alexa 647 was not observed, as shown in Figure 3.17. The binding of the protein was still tracked by the quenching of tryptophan, which suggested that the cause of the quenching was due to the proximity of two monomers.

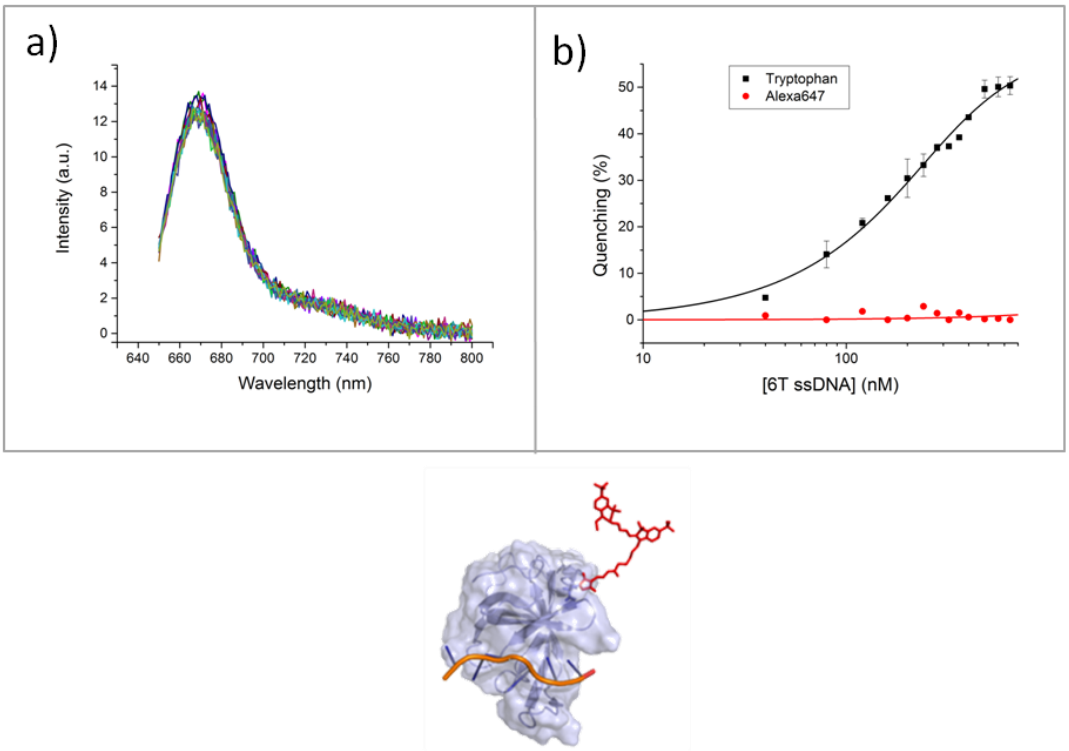


Figure 3.17: Quenching of Alexa 647 fluorescence.
Quenching of Alexa 647 was not observed in (a) when the oligonucleotide was shortened to 6T, which could not support two monomers at any one time. The binding to ssDNA was tracked using tryptophan quenching as shown in (b), which was approximately 80% compared to the negligible quenching observed for Alexa 647. Below is a cartoon of a single monomer of SsoSSB Alexa 647 bound to 6T ssDNA.

The fluorescence intensity of SsoSSB Alexa 647 was quenched by approximately 80% when introduced to 12C ssDNA, shown in Figure 3.15(b). If SsoSSB did form any higher order structures, such as a tetramer, it is unlikely that the same extent of

quenching of Alexa 647 would have been observed as the monomers would already be in contact with each other without ssDNA.

From this experiment, it was deduced that the origin of the Alexa 647 quenching was due the presence of a second labelled monomer. The similarity in the UV-vis spectra of SsoSSB Alexa 647 in the presence and absence of ssDNA suggested that the Alexa 647 molecules do not form a non-fluorescent dimer and that the quenching was achieved through some other mechanism. This could be explained by the flexibility of the nucleofilament observed in the previous FRET experiment, which would allow the fluorescent dyes to avoid contacting each other. Cyanine dyes have been shown to predominately self-quench through an energy transfer mechanism, even when dimer formation is possible. Furthermore, the nature of the protein surface interactions with the dyes in crowded environments could also have played a role in the decreased fluorescence intensity of the Alexa 647.

3.15. Discussion

In these ensemble studies, *Sulfolobus solfataricus* SSB binding to ssDNA was primarily investigated through changes in fluorescence intensity of artificial dyes that were conjugated to the protein and ssDNA. The oligomeric state of SsoSSB is a topic debated in the literature and the evidence presented here strongly suggests that this SSB exists primarily as a monomer in solution, which was also addressed by Gamsjaeger.[114] The labelling process was optimised so that approximately 100% of the protein monomers were conjugated to either a fluorescent dye or spin label, which was confirmed by mass spectrometry and optical absorbance where

appropriate. Any unlabelled proteins were removed by ion exchange and size exclusion columns, so that no unlabelled proteins could cause any artefact signals. Gel filtration, PELDOR and ensemble FRET and quenching experiments all failed to observe any significant interactions between monomers in solution which agree with the published crystal structure. The addition of ssDNA to these experiments showed clear evidence that the monomers were typically only in proximity to one another when they were bound to the same strand of DNA.

The distribution of distances obtained from PELDOR was too broad to confirm the model depicted in Figure 3.3. The low signal to noise ratio meant that the data collected could not determine if one or multiple distances were measured. Results from Gamsjaeger *et al.* suggested that it was likely that SsoSSB monomers did have an orientation relative to the ssDNA, so a more defined distance would have been resolved if all SsoSSB monomers were bound in a strict, rigid nucleofilament.[114] Protein dynamics would be minimal at the low temperatures that these experiments were completed; therefore any poorly defined distances could be a result from either a wide range of distances present or artefacts from other unbound species in solution. A weak protein protein interface could produce a large variation in the degree of rotational freedom that the proteins experience whilst bound to the ssDNA. Also, weak protein protein interaction could also result in anti-, non-, or only mildly cooperative behaviour that would generate a wide range of distances between the dyes on the SsoSSBs bound to ssDNA as well as increasing the noise from unbound proteins in solution.

Alone, the PELDOR results did not point towards a strong protein protein interface. Even with the stabilising presence of ssDNA the results seemed to discount the formation of rigid filaments. This was somewhat intuitive given the steric clashing expected if SsoSSB monomers were forced together. However, the quality of the PELDOR data was not credible enough on its own. The end-to-end distance calculated also implied that there was some degree of flexibility between monomers on a 39 nt ssDNA, since the persistence length matched reasonably well with the binding site of a SsoSSB monomer. These results were not under the stringent restrictions that programs such as MTSSL Wizard apply, however these conclusions have taken into account the length and flexibility of the dyes' tethers. A qualitative analysis of the $ratio_A$ values measured clearly showed that SsoSSB increases the end-to-end distance of ssDNA, and that the final distance was much too long for a ssDNA stabilised *E.coli* like tetrameric arrangement of monomers and also certainly too short to be a fully extended strand. Coupled with the distribution of distances of SsoSSB monomers bound to ssDNA measured by PELDOR, these studies imply that the SSB/ssDNA complex was a reasonably flexible linear filament. Together the large distribution of distances between monomers and the short persistence length suggested that there was a degree of freedom as to where SsoSSB monomers bind relative to each other and that they bind without a strongly defined protein protein interface.

Previous experiments investigated the effect SsoSSB had on staining the ssDNA with ethidium bromide. These experiments concluded that an SsoSSB nucleofilament effectively covered the ssDNA and prevented the intercalation of ethidium bromide

between the ssDNA bases.[109] The strands of ssDNA used were hundreds of nucleotides long which could not have given much indication of the action of a single monomer with respect to its neighbours; therefore the complete decoration could have been an example of cooperative binding or a result of the vast excess of protein, since the concentration of SsoSSB was well above the reported dissociation constants of 10 to 100 nM. To investigate both cooperativity factors and dissociation constant in more detail, the length of ssDNA was reduced to observe a maximum of four monomers binding at any one time as well as reducing the concentration of SsoSSB to examine binding at low nanomolar concentrations.

The binding of SsoSSB monomers was initially investigated by the quenching of the fluorescence from the tryptophan residues. The apparent decrease in dissociation constant of SsoSSB observed as the length of ssDNA was increased to allow up to four monomers to bind on the same strand was indicative of a binding mechanism that displayed a modest cooperativity.

SSBs typically bind to ssDNA in a nonspecific manner with regard to nucleotide sequence and results in multiple overlapping binding sites for SSBs to bind on ssDNA, as described in Materials and Methods. The McGhee-von Hippel model for cooperative binding to infinite lattices has been widely used to describe cooperative proteins binding to a one dimensional lattice, including EcoSSB and bacteriophage T4 gene 32 protein binding to DNA.[186] It has also been shown to be unsuitable for cooperative binding to small lattices and so Epstein's model was used as an alternative.[178] This fully describes overlapping binding sites, end effects of a finite

lattice and cooperative binding; however this was more appropriate for a forward titration where protein is added to ssDNA – which is impractical when measuring the quenching of protein fluorescence.

Binding of two SsoSSB monomers to 12 nt ssDNA was also followed using PIFE, which allowed the concentration of ssDNA and SsoSSB to be reduced closer to the K_d, and therefore to investigate binding isotherms in more detail. The addition of protein to ssDNA allowed the use of Epstein's finite lattice model that was not possible for reverse titrations such as those observed for tryptophan quenching experiments. The forward PIFE titration again showed that SsoSSB can exhibit modest cooperative binding to ssDNA, which ensures that the nucleofilaments are effective at completely decorating and protecting the ssDNA from insults. The limit on the cooperativity value observed could be a result of the high concentration of SsoSSB *in vivo* which negates the necessity for highly cooperative factors that would efficiently coat the ssDNA with a limited numbers of proteins.

FRET and Alexa 647 quenching both showed measurable changes upon the addition of ssDNA. The change in signal of FRET was over a smaller concentration than expected which prompted a closer look at the acceptor dye during SsoSSB binding. The quenching was only observed upon the addition of ssDNA long enough to accommodate two SsoSSB monomers, so was attributed to the presence of a second SsoSSB monomer in the nucleofilament. Alexa 647 is structurally similar to Cy5 and both exhibit non-fluorescence states brought about by aggregation, resonance energy transfer, photoisomersiation, intersystem crossing and photo-induced charge

transfers. The lack of any change to the profile of the UV-vis absorption discounted the formation of non-fluorescent dimers, possibly as a result of the flexibility between monomers in the nucleofilament. However, in a nucleofilament the Alexa 647 molecules were still in a crowded environment, and despite the ambiguity of the explanation of the quenching, the proximity of charged protein surfaces and other fluorescent molecules clearly had a significant effect on the optical properties of the dyes.

3.16. Conclusions

This chapter has focused on how the photophysics of the cyanine dyes in crowded environments can be attributed to relatively small proteins binding to ssDNA. Significant quenching and enhancement of the fluorescence intensities of Cy3 and Alexa 647 due to the relative positions and orientation of the labelled monomers of SsoSSB have been exploited to confirm the tight binding of SsoSSB to ssDNA in addition to providing evidence that suggested a monomeric structure for the SsoSSB protomer and a moderately cooperative binding mechanism that resulted in a loosely linear nucleofilament.

There have been no results observed that strongly support a well defined interface between SsoSSB monomers, either free in solution or bound to ssDNA. In contrast, the ensemble data strongly implied that SsoSSB is a monomeric species in solution, and that once bound to ssDNA the monomers still experienced a degree of rotational freedom from each other and did not produce an inflexible linear nucleofilament. The lack of a defined and favourable protein-protein interaction suggested that the cooperativity observed could also be due to the nature of ssDNA itself rather than an

overriding energetic preference to make contact between monomers. The nucleotides of ssDNA neighbouring a monomer of SsoSSB already bound are likely to behave differently to a completely free ssDNA, rendering them more accessible for another SsoSSB monomer. The opposite trend was noticed for EcoSSB, whose OB folds are in a well defined position and orientation relative to each other and the binding of the second EcoSSB OB fold was identified as anti-cooperative due to energetic penalties of constraining ssDNA in that manner.

The FRET, quenching and enhancements detailed in this chapter rather serendipitously provided the basis for a novel single molecule fluorescent assay. Using a combination of the three different fluorescence processes was advantageous in that it was clearly able to distinguish between nucleation and filament growth, which provided a precise method of studying the initial stages of OB folds forming a nucleofilament.

4. Single molecule analysis of SsoSSB

TIRF microscopy allowed the observation of single molecules of fluorescently labelled ssDNA immobilised on a quartz slide. This technique measures the fluorescence intensity of fluorophores and any binding of proteins to the ssDNA has to be inferred from the changes in the rate of photons emitted from the fluorescent dyes. A series of filters, mirrors and dichroic mirrors allows the fluorescence from two different dyes to be separated and measured simultaneously, so that any transfer of energy between the two dyes can be calculated. The fluorescence of the two different dyes are reflected onto different halves of a EM-CCD camera, giving two images of the same slide, viewing the donor and acceptor dyes respectively. These were 2D representations of the slide, so ideally a fluorophore resembled a singular bright spot on a dark background, so the ssDNA was immobilised on to the slide at approximately low picomolar concentrations in order for each fluorophore to be resolved separately.

Single molecule results are commonly presented as traces of the fluorescent intensities of the two fluorescent dyes plotted against time. In the absence of an acceptor dye and other external quenchers, the donor dye should fluoresce at a constant intensity for the duration of the experiment, until photobleaching of the donor molecule occurs, which results in the fluorescence intensities effectively becoming zero. An example of this is a single ssDNA molecule immobilised to a quartz slide labelled with Cy3 dye at the 3’ end, shown in Figure 4.1(a). This is compared to that of multiple fluorophores where discrete steps clearly show distinct

photobleaching events of more than one fluorescent dye. Traces showing multiple donor dyes were excluded from further analysis to ensure that conclusions made were from the fluorescence emitting from a single ssDNA. Throughout this thesis Cy3 was used as donor dye and its fluorescence intensity is represented by the green trace. Similarly, Alexa 647 was employed as the acceptor dye and is represented by red traces.

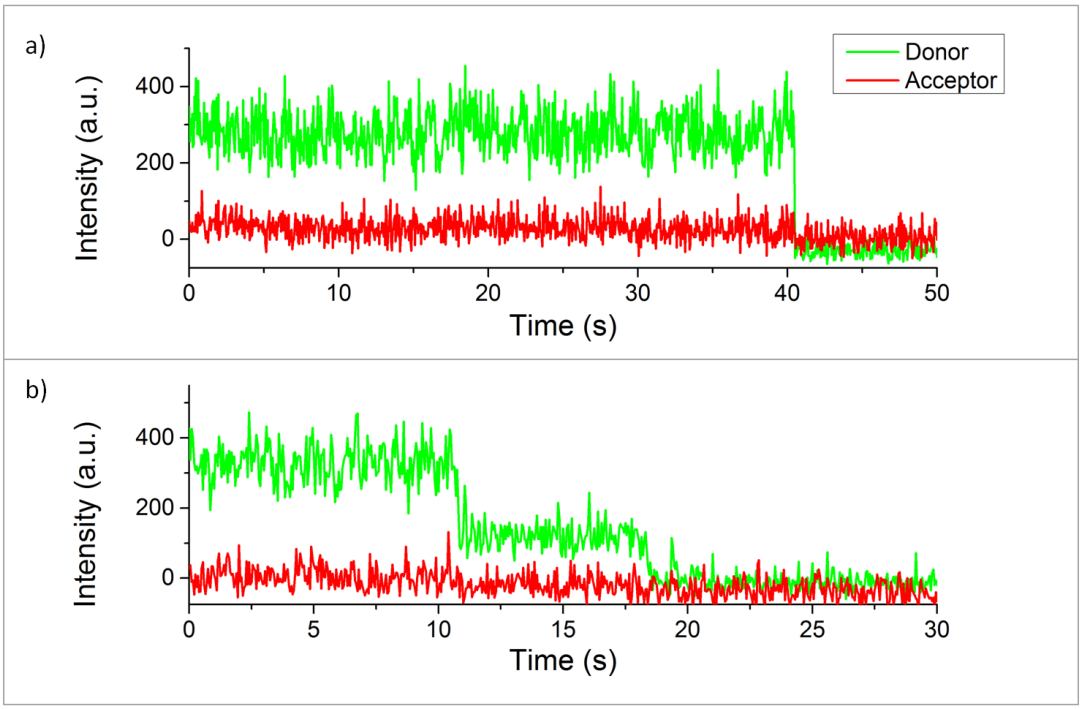


Figure 4.1: The fluorescence intensities from single molecules of 12C ssDNA Cy3 molecules. Donor and acceptor emission are represented by the green and red traces respectively, and this will be consistent throughout all figures. Top trace (a) shows a single photobleaching step at 40 s, whilst (b) shows two at 10 s and 17 s. This double photobleaching suggested that there were two fluorescent molecules present in a single spot and the data from traces showing similar behaviour are disregarded to avoid confusion. The near zero intensity of the fluorescence from the 640 nm channel (acceptor channel) suggested that there were no acceptor dyes present for FRET to occur.

Similar to the ensemble experiments, single molecule binding events are viewed as increases in FRET, where an anti-correlation of the donor and acceptor dyes show a transition to an increase in acceptor intensity coupled with a decrease in donor

intensity. Many of these events are examined from numerous molecules of ssDNA in various experimental conditions to build a comprehensive picture of the mechanism SsoSSB employs to bind to ssDNA.

4.1. Non-specific binding to the slide

Initially SsoSSB labelled with Alexa 647 was washed on and off the PEG surface of the microscope slide to ensure that any signals received in later experiments were not due to unspecific surface interactions. The number of spots observed on the surface that persisted after washing was comparable to a clean surface, as shown in Figure 4.2. The change in fluorescent intensity before and after the addition of SsoSSB Alexa 647 was also compared and the affinity of SsoSSB for the surface of the slide seemed insignificant, demonstrated by the facile removal of SsoSSB from the PEG surface, also shown in Figure 4.2. This was compared with the spots observed after the addition of ssDNA modified with biotin and Cy3 at the 5’ and 3’ ends respectively, where the specific immobilisation of the ssDNA meant that the presence of ssDNA persisted through multiple washes. Subsequent single molecule studies showed the same changes in the fluorescent intensity of the dyes that had been seen in the ensemble experiments, and the FRET, PIFE and quenching events were therefore attributed to specific interactions with ssDNA.

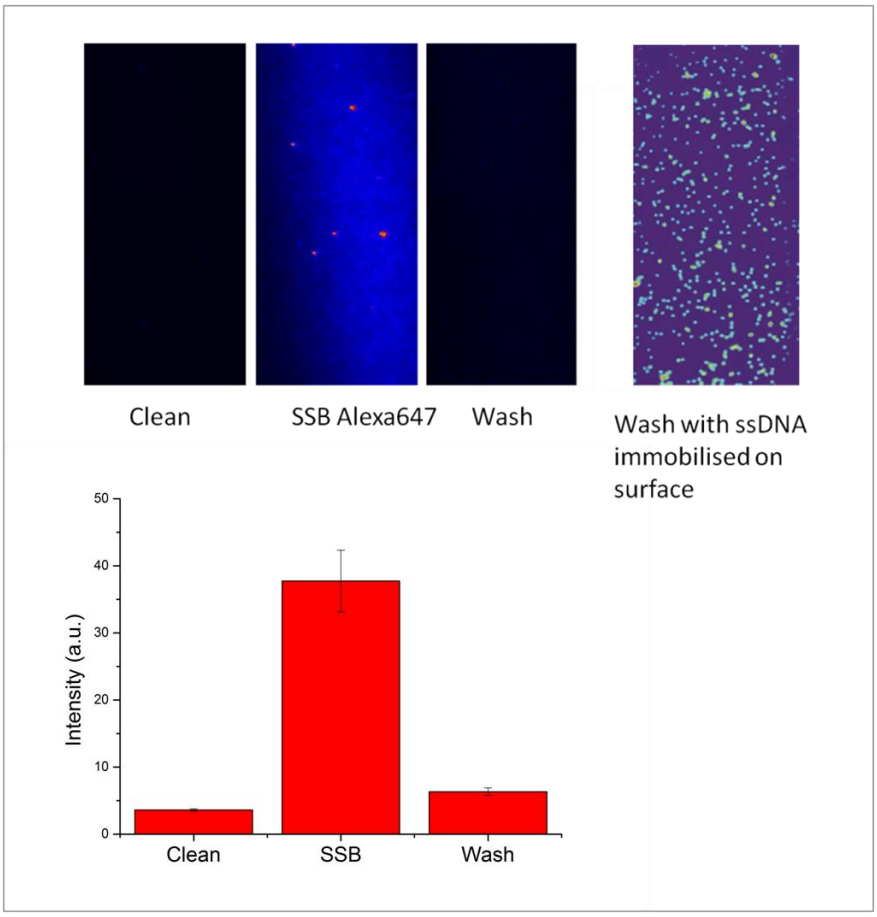


Figure 4.2: Images taken from TIRF microscope of a clean PEG slide before and after loading 100 nM SsoSSB labelled with Alexa 647 on to the surface.
The SsoSSB was subsequently washed off the surface of the slide. This is compared with trying to wash off ssDNA that has been immobilised by biotin/neutravidin interactions.

4.2. PIFE

The fluorescence from ssDNA alone was relatively constant until photobleaching occured, when the fluorescence intensity was totally and irreversibly quenched. No enhancements were observed under these conditions and a typical trace in the absence of SsoSSB is shown in the bottom trace of Figure 4.3. The single molecule traces were normalised so that fluorescence intensity from the Cy3 in the absence of SsoSSB equalled one. This state was assumed to be the average fluorescence intensity of the lowest level of fluorescence before photobleaching, and histograms

were produced to quantify the fluorescence increase due to presence of SsoSSB. Upon the addition of unlabelled SsoSSB, PIFE was observed as a 2-3 fold increase in the intensity of fluorescence of Cy3 conjugated to molecules of ssDNA, clearly shown in the other traces in Figure 4.3. There the normalised traces are also shown from a range of SsoSSB concentrations. The frequency of transitions between low and high fluorescence intensities exhibited a positive correlation with the concentrations of SsoSSB, which suggested that these were the PIFE binding events observed by ensemble experiments. SsoSSB concentrations approaching 50 nM resulted in constant enhancement of the recorded Cy3 fluorescence, indicating either a continuous presence of the same monomers on ssDNA or a rapid exchange of proteins that is too fast to resolve (on and off binding below 33 ms).

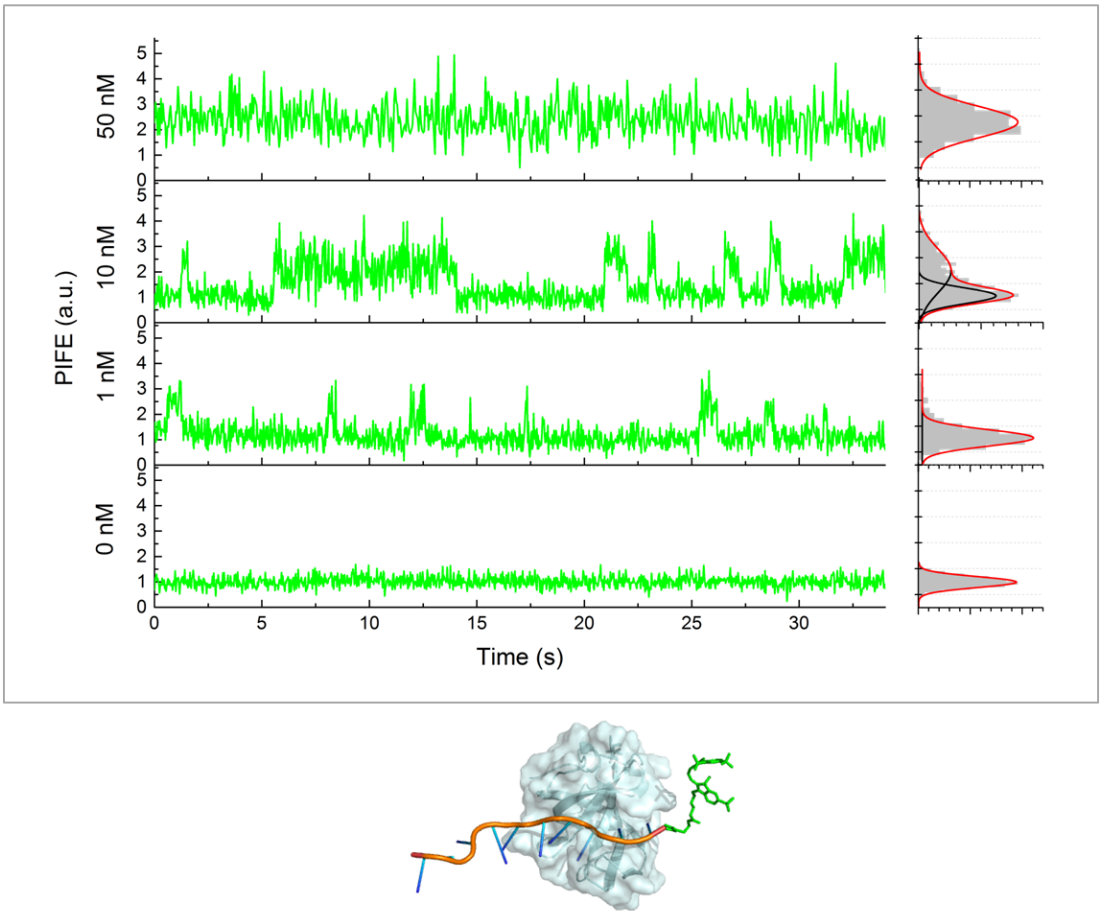


Figure 4.3: The intensities of the fluorescence emitted from single molecules of 12C ssDNA Cy3 immobilised on quartz slide have been normalised and plotted against time.
These traces show binding events as a 2-3 fold increase in fluorescence. The proportion of time the unlabelled SsoSSB was bound to ssDNA increased with protein concentration so that at 50 nM the ssDNA was always decorated with proteins. The histograms to the right of the traces show the frequency distribution of the fluorescence intensities, again showing an increase in enhancement increasing with protein concentration. Below is a cartoon of unlabelled SsoSSB bound to ssDNA, inducing a fluorescence enhancement of the Cy3 dye (green.)

The histograms from approximately 1000 traces are shown Figure 4.4, where in the absence of SsoSSB only a single peak is observed centred at the normalised value of 1. A second peak was observed upon addition of SsoSSB which indicated a two-fold increase in fluorescence and which became increasingly pronounced with increasing concentrations of SsoSSB, representative of the increasing fraction of time that ssDNA was bound by SsoSSB monomers.

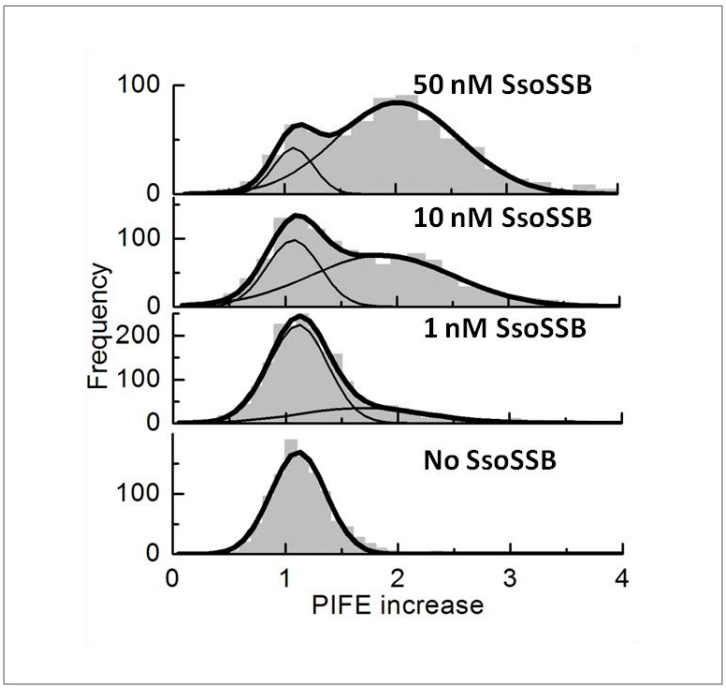


Figure 4.4: Histograms, each from approximately 1000 single molecule traces, which show the fluorescence intensity of immobilised 12C ssDNA Cy3 on a quartz slide being exposed to increasing concentrations of unlabelled SsoSSB.
Upon the addition of SsosSSB a two-fold increase in the fluorescence intensity of Cy3 was observed. This two-fold increase became increasingly dominant with the increasing concentrations of SsoSSB, coupled with the relative decrease in size of the original peak at 0 nM SsoSSB.

The lack of binding dynamics at an SsoSSB concentration of 50 nM could have been potentially confusing when trying to identify when the fluorescence was enhanced; therefore after traces at 50 nM were recorded, imaging buffer was washed through the sample chamber to reduce the SsoSSB concentration, as depicted in the cartoon in Figure 4.5(a). The histograms in Figure 4.5(b) from the washed ssDNA showed an additional peak, approximately half the intensity of the ssDNA Cy3 in the presence of 50 nM SsoSSB and the traces from the washed ssDNA exhibited clear dissociation and association events as a result of reducing the protein concentration, shown in Figure 4.5(c). This indicated that fluorescence intensity observed at high protein

concentrations was constantly enhanced due to the presence of SsoSSB monomers bound to ssDNA.

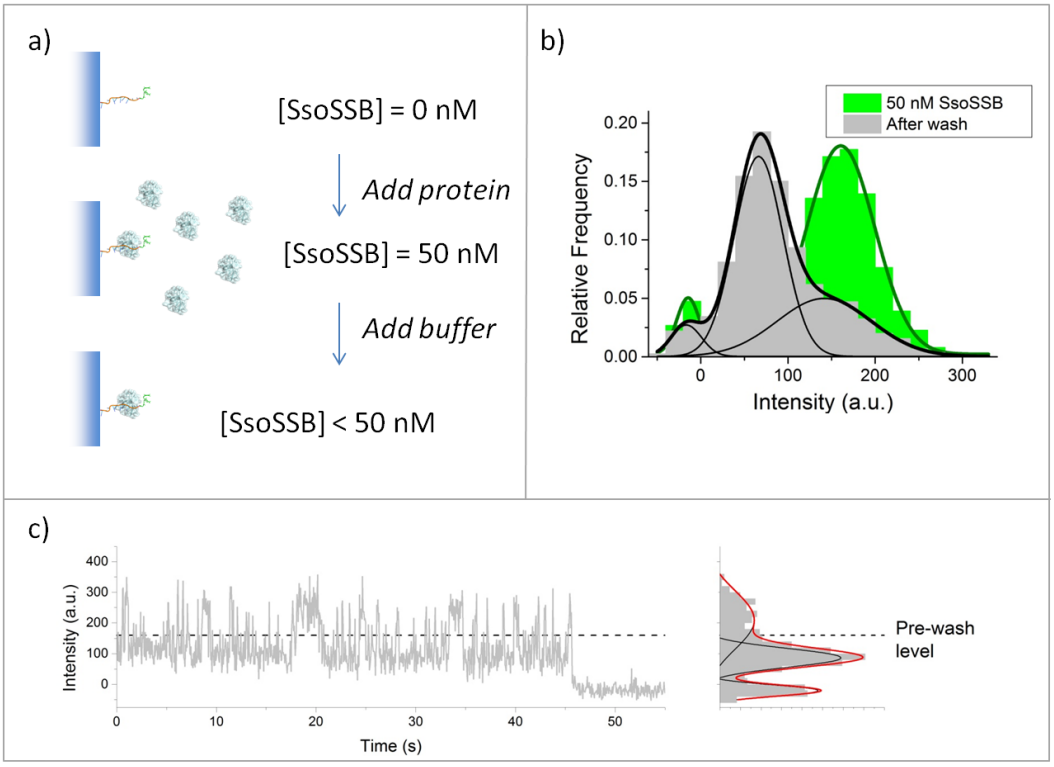
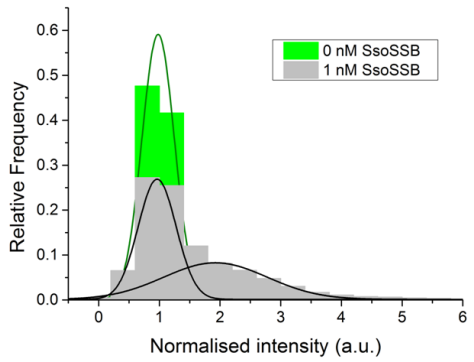


Figure 4.5: Histograms and a single molecule trace to show that at high concentrations of SsoSSB the traces showed no dynamics but were approximately twice the intensity of 12C ssDNA Cy3 in the absence of protein.
To ensure that the fluorescence observed at 50 nM was increased by protein binding to ssDNA, the sample chamber was washed with imaging buffer to remove the unbound protein from the bulk solution, lowering the protein concentration. This enabled the bound protein to dissociate from the ssDNA and allowed the fluorescence intensity to return to its previous intensity. Figure (a) shows a schematic of the process, (b) is a frequency histogram of fluorescence intensities from ssDNA Cy3 in the prescence of 50 nM SsoSSB and from after washing with imaging buffer, and (c) is an example of a trace showing the fluorescent intensity of ssDNA Cy3 after washing with buffer. The three intensities clearly show that the ssDNA now spends most of its time undecorated with protein, but also shows enhancements consistent with SsoSSB binding events. The near zero intensity was a result of photobleaching of the dye, the one step profile of these photobleaching events suggested that these traces were recorded from single molecules.

The broadness of the peaks which represent the enhanced fluorescent intensity can be attributed to the flexible nature of the ssDNA leading to transient interactions with the protein surface. This meant that it was difficult to differentiate overlapping binding sites that the SsoSSB could bind to along ssDNA. The enhancement due to

protein binding has been shown to be distance dependent, and that it is possible to discriminate between binding positions of RecA, which occupies only 3 nt per monomer.[182] SsoSSB proteins bind to a total of 5 nt along the same length of ssDNA. This resulted in only subtle changes in fluorescence intensity and the increased noise observed during PIFE did not allow the unambiguous discrimination between enhancements from different numbers of monomers bound to ssDNA. The difference between the levels of noise recorded from decorated and undecorated ssDNA is shown more clearly by the histograms in Figure 4.6. The full width at half maximum heights (FWHM) are also presented and the presence of SsoSSB on the ssDNA caused approximately a three-fold increase in the FWHM. This was most likely to due, as previously stated, to the transient nature of the interaction between the protein surface and the dye, leading to fluctuations in the strength of the interaction between the dye and protein surface.



Peak centre	FWHM
0.98 ± 0.002 (0 nM)	0.63 ± 0.009
0.96 ± 0.006 (1 nM)	0.73 ± 0.03
1.92 ± 0.1 (1 nM)	2.01 ± 0.1

Figure 4.6: Gaussian peaks fitted to histograms showing the frequency distribution of the fluorescence intensity of 12C ssDNA Cy3 at 0 nM and 1 nM unlabelled SsoSSB for approximately 200 molecules.
The table shows the centre and full width at half maximum height (FWHM) values for the Gaussian peaks representing the fluorescence intensities for undecorated ssDNA and for when SsoSSB was bound to the ssDNA. The intensities were normalised so that undecorated ssDNA gave a fluorescence intensity centred at 1. At 1 nM the ssDNA still clearly spent a significant proportion of its time undecorated but the fluorescence intensity also was enhanced approximately two fold due to the presence of SsoSSB. The FWHM values increased upon SsoSSB binding, demonstrating the dynamic interaction between the dye and protein surface.

4.3. FRET

When SsoSSB Alexa 647 was introduced to 12C ssDNA Cy3 immobilised on a quartz slide, a very efficient energy transfer was observed, with E_F typically adopting values close to 1. These were attributed to SsoSSB binding to ssDNA and typical traces showing only changes in FRET are shown in Figure 4.7(a). The frequency of the bursts of high FRET intensity were again positively correlated with increasing SsoSSB concentration similar with the PIFE binding events, which confirmed that the observed FRET was due to labelled SsoSSB Alexa 647 interacting with ssDNA. Approximately 350 binding events were analysed and gave a frequency distribution of E_F that yielded two states centred at E_F = 0.11 and 0.97, shown in Figure 4.7(b), which were assigned to undecorated ssDNA and a single monomer bound to ssDNA respectively. A non-zero E_F value for the unbound ssDNA was most likely due to cross-talk between the donor and acceptor channels rather than the presence of an acceptor dye. FRET is generally insensitive to changes below 4 nm, and the overlapping binding sites and the flexibility of the ssDNA again did not give significantly different E_F values in order to indentify where on the ssDNA SsoSSB was bound. These binding events persisted with an approximate lifetime of 200 ms, which was a short period of time compared to the reported dwell times of EcoSSB and RPA on ssDNA, which were found to be able to maintain 400 and 5 seconds of contact with ssDNA respectively.[187]

The quenching of the acceptor fluorescence when a second monomer binds to the ssDNA, as seen in the ensemble titrations, was potentially problematic and the

overall kinetics of the initial growth of a filament, including nucleation is discussed in the next section of this chapter.

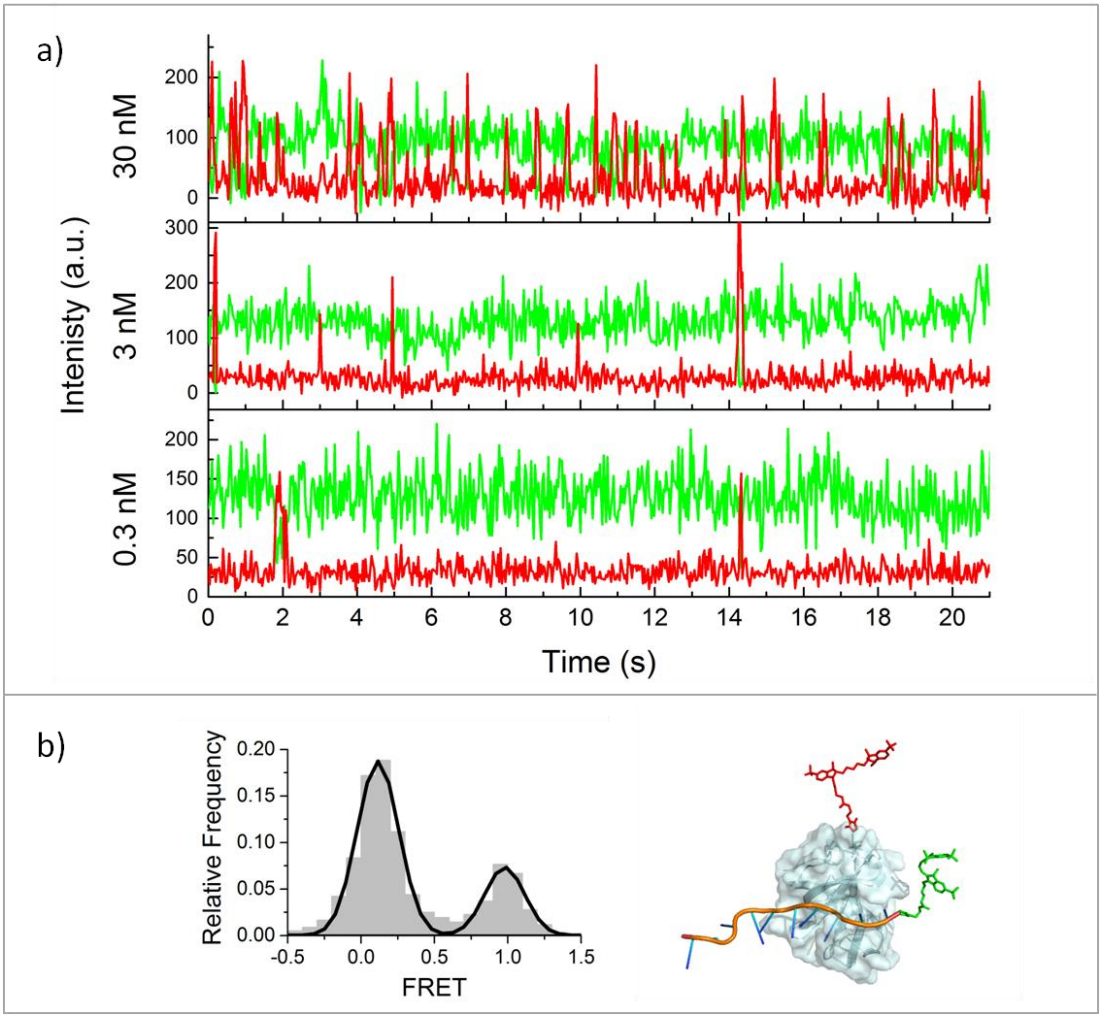


Figure 4.7: Single molecule traces showing SsoSSB Alexa 647 monomers binding to 12C ssDNA Cy3 as bursts of increased acceptor fluorescence accompanied by quenching of the donor intensity, typical of highly efficient FRET events.
Figure (a) shows an increasing rate of binding as SsoSBB Alexa 647 concentration was increased, and (b) shows a frequency histogram of the FRET values from approximately 200 binding events, giving a low and high FRET values that correspond to ssDNA undecorated and decorated with SsoSSB Alexa 647 respectively. (b) also shows a cartoon of SsoSSB Alexa 647 bound to ssDNA Cy3 demonstrating the short inter-dye distances.

4.4. Full in-trace interpretation

This thesis has used FRET, PIFE and self-quenching to individually describe SsoSSB and its mechanism of binding to ssDNA. They were, however, not mutually exclusive techniques in this instance. PIFE and self-quenching also influenced the fluorescence intensities of donor and acceptor dyes in the single molecule FRET experiments. Many traces exhibited all three behaviours, as shown in Figure 4.8. In this exemplary trace, an instance of high FRET is labelled, as is a section which showed quenching of the acceptor with a simultaneous enhancement of the donor dye emission. These were interpreted as a single monomer binding to ssDNA and a second monomer binding to ssDNA respectively.

The aim of these experiments was to characterise the binding of SsoSSB to ssDNA and the length of the ssDNA was initially kept at 12 nt in order to only view a small number of monomers binding, ideally simplifying the interpretation of the traces. The flexibility of ssDNA and its short length meant that when the first SsoSSB Alexa 647 monomer bound to ssDNA, the protein surface disturbed the local environment of the donor dye regardless of which binding site the protein was bound to. The resultant enhancement of the donor fluorescence was not observed due to the efficient energy transfer to the acceptor dye, which stimulated the acceptor fluorescence but quenched the donor fluorescence and demonstrated the anti-correlation that is typical of pairs of fluorescence dyes capable of FRET. These periods of high FRET efficiency were attributed to the presence of a single labelled monomer, and an example of these events is shown by the region labelled high FRET in Figure 4.8.

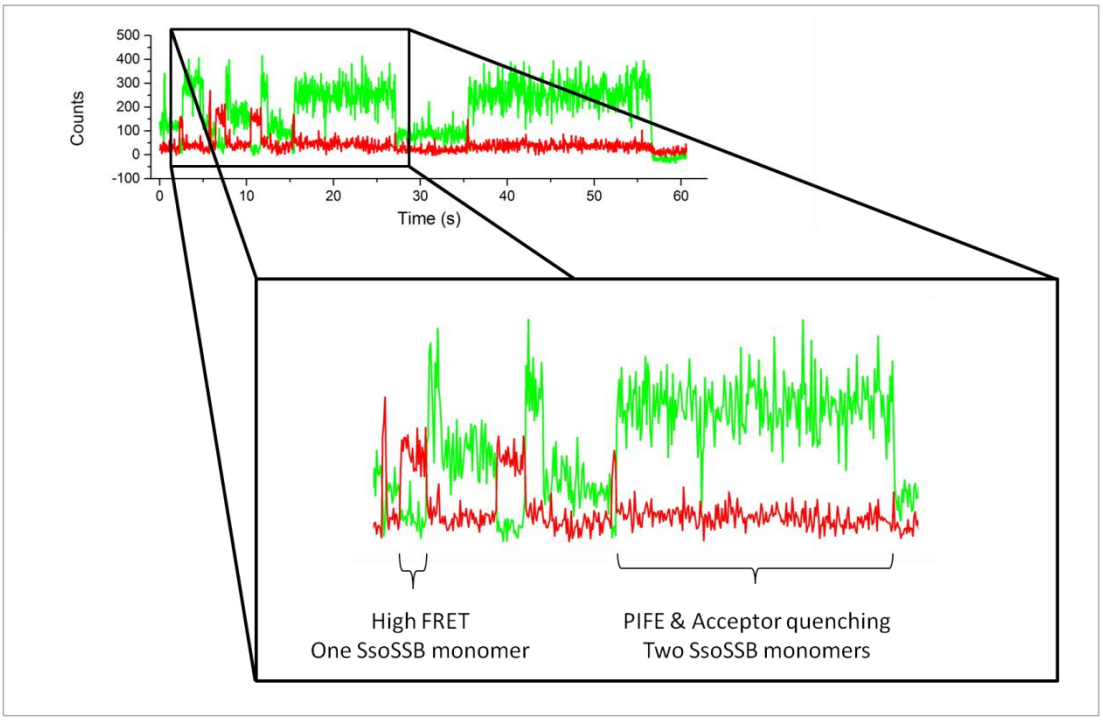


Figure 4.8: An example of a single molecule trace showing the changes in fluorescence when one SsoSSB Alexa 647 monomer and two monomers bind to 12C ssDNA Cy3.
A single monomer bound to ssDNA was shown by an increase in acceptor fluorescence (Alexa647, red) and a simultaneous decrease in donor emission (Cy3, green), typical of a highly efficient FRET. There are four instances in the inset where a single monomer has bound to the ssDNA, the second of which has been labelled. If a second monomer binds, the acceptor fluorescence is quenched however the interaction of the donor dye with the protein surface results in an enhancement of fluorescence (PIFE.) Similarly there are three instances where two SsoSSB monomers are bound shown in the inset, the final one is labelled.

These periods of high FRET efficiencies were relatively brief, and were either immediately followed by the donor and acceptor dyes returning to intensity levels that were characteristic of undecorated ssDNA, or by a period of a 2-3 fold enhancement of the donor fluorescence intensity which was concomitant with the quenching of the acceptor fluorescence. The return to typical levels of fluorescence was distinctive of the single monomers dissociating from ssDNA, as seen in the previous paragraph 4.3. The regular occurrences of these periods of enhancement, the low percentage of unlabelled protein in the sample and the observations of self-quenching during ensemble experiments suggest that it was most likely that these

high PIFE, acceptor quenching events were a result of a second SsoSSB Alexa 647 monomer binding to ssDNA, and are represented by the area labelled PIFE in Figure 4.8. This novel in-trace analysis has allowed the binding of the first SsoSSB monomer to be unambiguously separated from the sequential binding of further monomers during growth of the filament. An example of how these FRET, PIFE and quenching events can be used to track the number of monomers on 12C ssDNA Cy3 is given in Figure 4.9. This technique was used to characterise the differences in affinities for ssDNA of monomers that are binding to ssDNA as an individual monomer or adjacently to other proteins already bound.

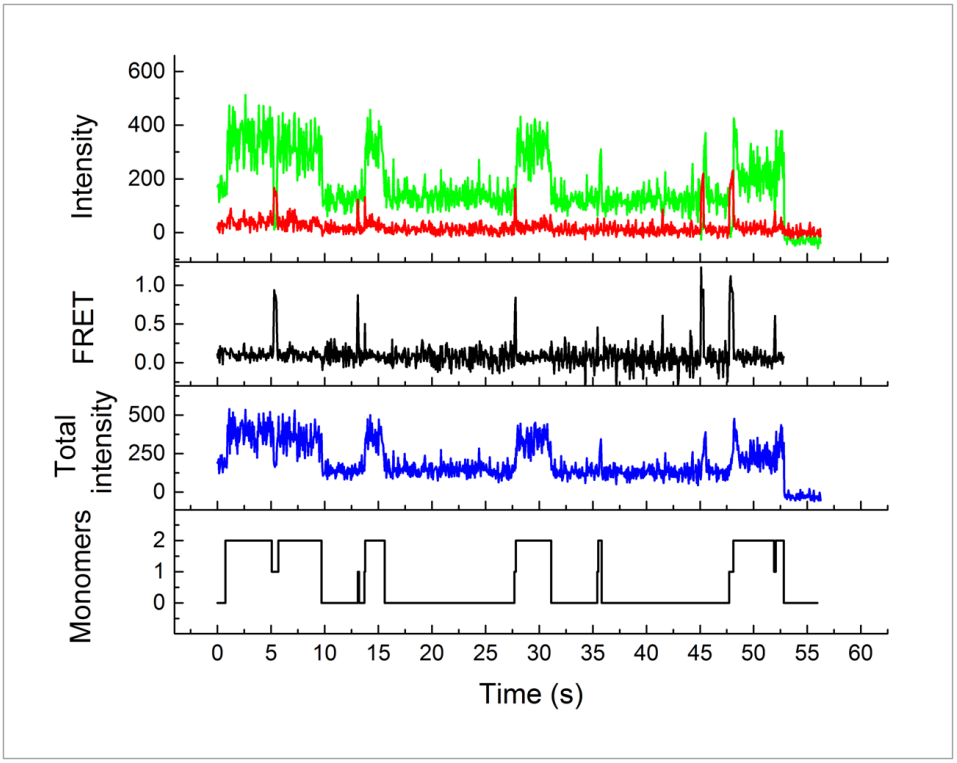


Figure 4.9: An example of how the FRET, quenching and PIFE events can be interpreted in terms of the number of SsoSSB Alexa 647 monomers bound to 12C ssDNA Cy3.
One and two monomers bound are represented by high FRET and high PIFE regions respectively. The top trace shows the raw data taken from a typical single molecule trace where the donor and acceptor fluorescence intensities are given by green and red traces respectively. The second plot (top black line) is a trace the FRET calculated from the intensities of the donor and acceptor, while the third plot (blue line) is the sum of the donor and acceptor intensities. The bottom plot shows the number of SsoSSB monomers bound to the ssDNA, and shows how the high FRET events correspond to a single monomer bound to ssDNA. The PIFE can be clearly shown in the changes to the total intensity of the two dyes, and was concomitant with quenching of the acceptor dye.

4.5. Binding kinetics

This unique way of clearly observing the nucleation of SsoSSB on a single strand of 12C ssDNA Cy3, allowed the dwell times of the first and second monomer to be measured. Binding events can be viewed using the scheme in Figure 4.10, where SSB_1 and SSB_2 refer to the first SsoSSB monomers to bind to the ssDNA. The rates that each monomer binds to ssDNA can be calculated from measuring the times that monomers of SsoSSBs dwell on ssDNA. For example, the observed rate of the first SsoSSB monomer to bind to ssDNA, $k_{on,1}$, is proportional to the inverse of the time it takes for that SsoSSB monomer to bind $\tau_{on,1}$, ie. the time the ssDNA spends undecorated. Similarly, the observed rate of the dissociation of the first monomer, $k_{off,1}$, is related to the inverse of the time the first monomer is bound the ssDNA before it is to released, $\tau_{off,1}$. These events are described pictorially in Figure 4.11.

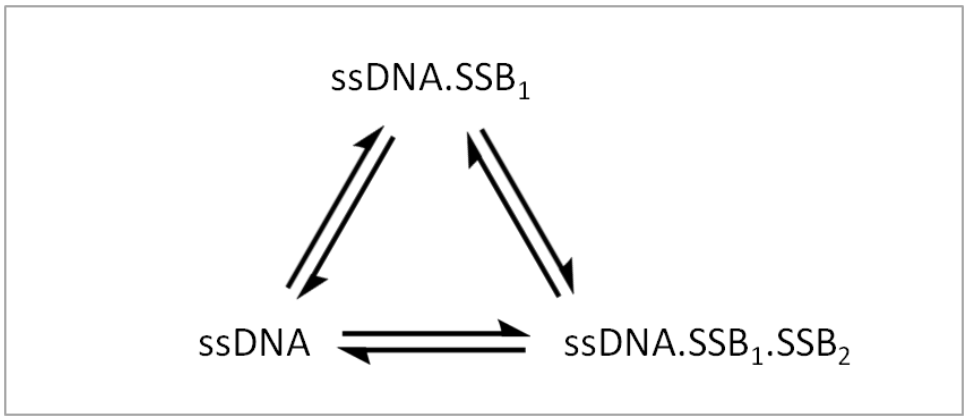


Figure 4.10: Two SsoSSB monomers can bind to or dissociate from ssDNA either sequentially or simultaneously.
The SSB_1 and SSB_2 represent the first and second monomers to bind to ssDNA respectively.

The duration of approximately 300-400 binding events were measured and the rate of binding on and off from ssDNA was calculated by fitting single exponential decays to frequency distribution histograms of the dwell times. The histograms for the first

monomer binding on and off across a concentration range 0.05 - 10 nM are given in Figure 4.12 and the rate constants are given in Table 4.1. There clearly was a concentration dependence on the rate of the first monomer binding to ssDNA, which suggests that these FRET events are caused by the SsoSSB Alexa 647 interacting with ssDNA Cy3. The dissociation of the first monomer was largely independent of protein concentration, which was in agreement with a rate equation that is first order with respect to protein concentration. Approximately 85% of binding occurred as an individual monomer, which suggested that SsoSSB predominately behaved as a monomer in solution in agreement with PELDOR and work published by Gamsjaeger.[114]

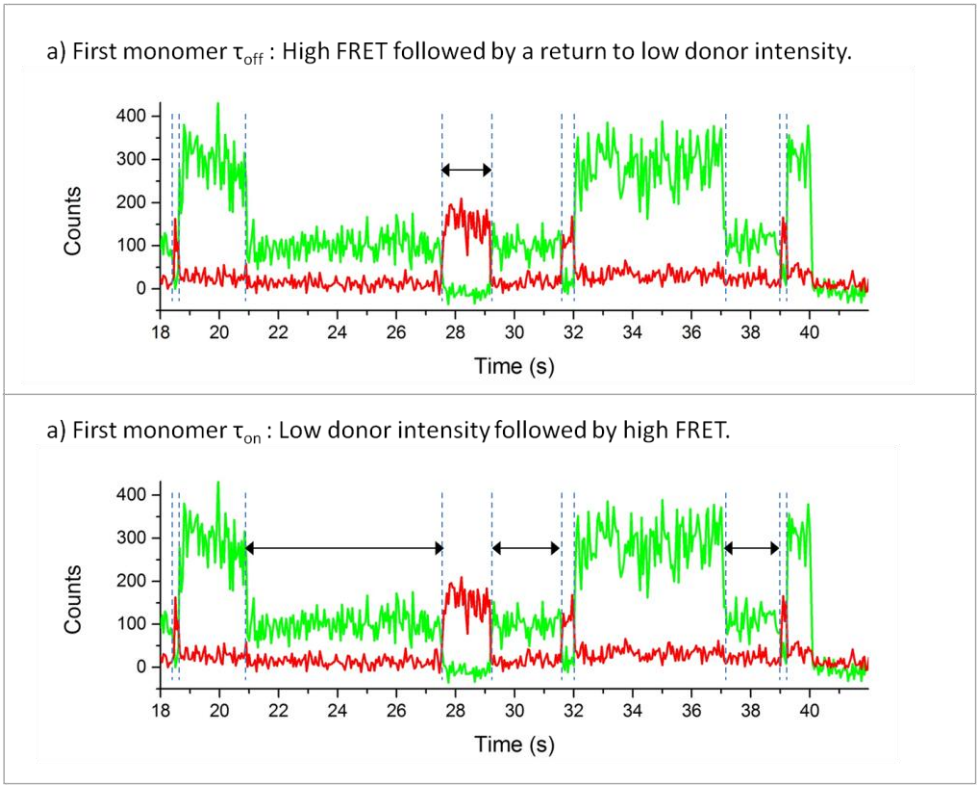


Figure 4.11: The rates of k_{off} and k_{on} for the a single monomer binding to 12C ssDNA Cy3 were calculated by measuring the time (a) immediately before a single monomer binding to or (b) dissociating from ssDNA.

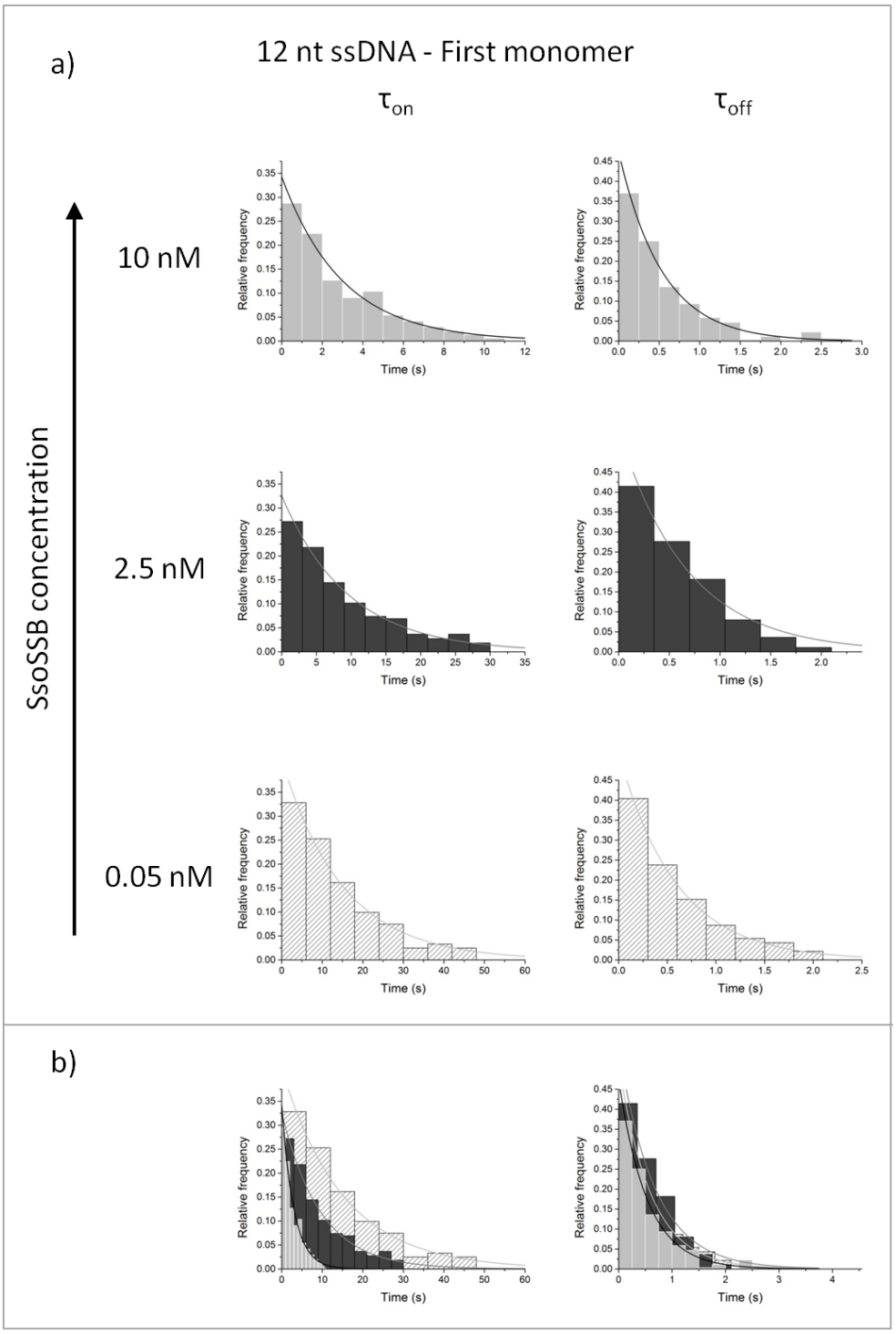


Figure 4.12: The dwell times of approximately 200 events showing a single SsoSSB Alexa 647 monomer binding to 12C ssDNA Cy3 were measured at each concentration (0.05, 2.5 and 10 nM) and plotted as histograms.

Single exponential curves were fitted and rate constants were calculated. The fitting was not improved by including a second exponential term. The histograms from (a) were overlaid on top of each other in (b) to clearly show that increasing the concentration of SsoSSB increases the rate at which monomers bind to ssDNA but the time before dissociation remains relatively constant.

The binding events of the second monomer were analysed in a similar manner, where the rate of binding of the second monomer, $k_{on,2}$, corresponded to the time it took from the binding of the first monomer until a second SsoSSB bound adjacently to the first, causing a quenching of the acceptor dyes' fluorescence and inducing a high donor intensity due to PIFE. The length of this enhancement represented the duration the ssDNA accommodated two SsoSSB monomers, and was generally followed by both SsoSSB monomers dissociating from the ssDNA with the fluorescence intensities of the donor and acceptor returning to low FRET, no PIFE states. These histograms are shown in Figure 4.14. It was not uncommon to observe two monomers bound to ssDNA to dissociate separately however it was more frequent to see the synchronised dissociation of both monomers. The rate constant for the second monomer binding to ssDNA showed a slight increase with the protein concentration, but was not as pronounced as the first monomer, whereas the rate of dissociation of the second monomer was largely independent of protein concentration. The rate of a monomer dissociating from a nucleofilament of two monomers was approximately two to three times slower than a single monomer dissociating from ssDNA. This pointed towards a cooperativity constant of approximately 10-100, roughly in agreement with the ensemble titrations.

Table 4.1: Table of observed association constants for the first and second SsoSSB monomers and dimers to bind to ssDNA, represented as $k_{on,1}$, $k_{on,2}$ and $k_{on,1.2}$ respectively.

[SsoSSB] (nM)	$k_{on,1}$ (s^{-1})	$k_{on,2}$ (s^{-1})	$k_{on,1.2}$ (s^{-1})
0.05	0.06 ± 0.005	1.75 ± 0.1	0.09 ± 0.001
2.5	0.1 ± 0.01	2.55 ± 0.2	0.12 ± 0.008
10	0.32 ± 0.04	3.37 ± 0.3	0.15 ± 0.003

Table 4.2: Table of observed dissociation constants for the first and second SsoSSB monomers and dimers to bind to ssDNA, represented as $k_{off,1}$, $k_{off,2}$ and $k_{off,1.2}$ respectively.

[SsoSSB] (nM)	$k_{off,1}$ (s^{-1})	$k_{off,2}$ (s^{-1})	$k_{off,1.2}$ (s^{-1})
0.05	1.6 ± 0.12	0.22 ± 0.01	0.33 ± 0.03
2.5	1.41 ± 0.10	0.31 ± 0.02	0.31 ± 0.003
10	1.56 ± 0.15	0.30 ± 0.008	0.36 ± 0.003

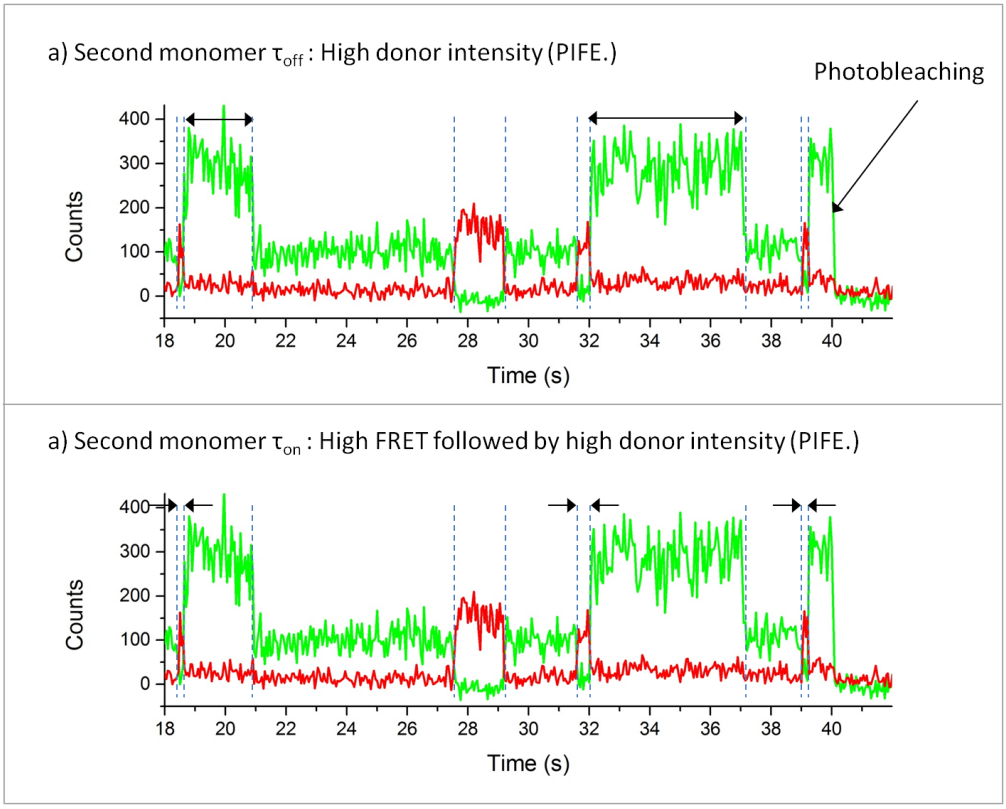


Figure 4.13: The rates of k_{off} and k_{on} for a second monomer binding to 12C ssDNA Cy3 were calculated by measuring the time (a) immediately before the second monomer bound to or (b) dissociated from ssDNA.

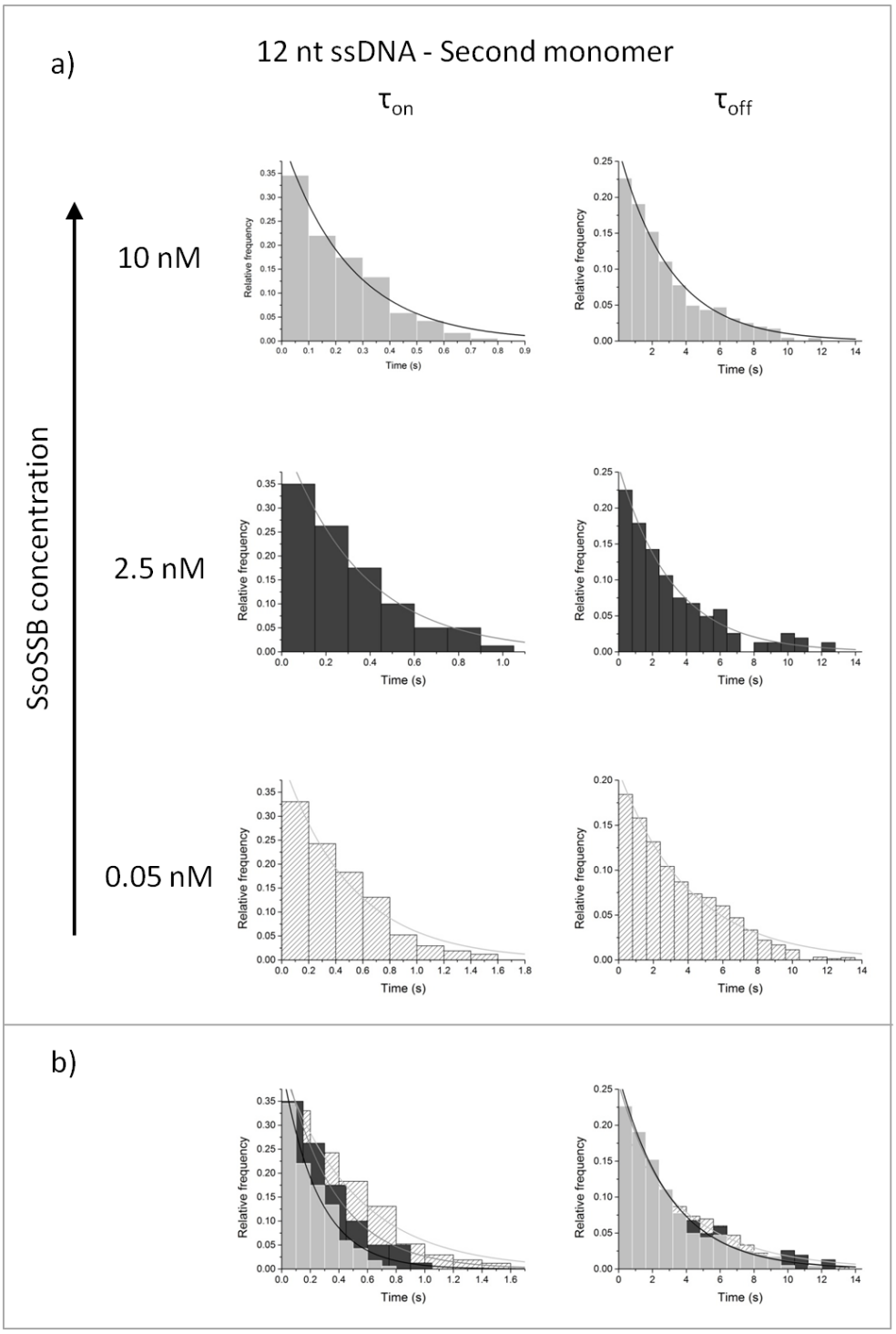


Figure 4.14: The dwell times of approximately 200 events showing a second SsoSSB Alexa 647 monomer binding to 12C ssDNA Cy3 were measured at each concentration (0.05, 2.5 and 10 nM) and plotted as histograms.

Single exponential curves were fitted and rate constants were calculated. The fitting was not improved by including a second exponential term. Again the histograms from (a) were overlaid on top of each other in (b) to show that increasing the concentration of SsoSSB slightly increased the rate at which the second monomer bound to ssDNA but the time before dissociation remained relatively constant.

4.6. Discussion

There are numerous examples of how single molecule fluorescence microscopy has been employed to observe proteins such as EcoSSB and RecA form nucleofilaments in real time.[24, 179] Strands of ssDNA were captured and moved in and out of channels in a microfluidic device, and the rates of binding and dissociation were calculated after essentially fishing for proteins. Other assays are more similar to those described in this thesis, measuring the FRET or PIFE resulting from the binding of the labelled or unlabelled protein, respectively.[179, 182]

PIFE compared to FRET demonstrates a greater sensitivity to changes in a closer range to a fluorescent dye's environment. This had been shown by Myong *et al.* who showed that FRET was unresponsive to changes below 4 nm whereas the PIFE showed a clear distance dependence in that regime.[182] This was more evident on dsDNA than ssDNA, since the more rigid double helix holds the dye in a more fixed position compared to ssDNA - whose innate flexibility leads to larger movements of the dye and therefore greater noise being recorded in the fluorescence signal.[182] The different enhancements following the action of a total of four RecA monomers could be distinguished upon binding to 13 nt ssDNA.[182] RecA preferentially builds a nucleofilament in a direction towards the 3' end of DNA, allowing a clear progression of enhancement to be viewed as the nucleofilament steadily advances along DNA closer to the dye. RecA is also a relatively large protein that binds to a small number of nucleotides compared to SsoSSB, presenting a larger protein surface for more significant interactions with the dye and therefore producing larger enhancements in the fluorescent intensity. RPA is also a much larger protein than

SsoSSB and PIFE has been attributed to changes in the protein's position due to RPA sliding along the ssDNA, again showing the capacity to use PIFE as a tool to track events occurring close to the fluorescent dye.[54] These much larger proteins have produced much more noticeable changes in the enhancement of fluorescent dyes than were observed with SsoSSB. Monomers of SsoSSB binding to ssDNA did effect the conformation of the dye sufficiently to produce a visible enhancement, however this technique could not make a distinction between different number of monomers bound to each strand or the position of a monomer along the ssDNA since, in this case, PIFE did not display any dependence on the distance between the dye and the protein surface.

RPA and EcoSSB have been shown to be able to diffuse along ssDNA, and similar PIFE experiments have been used to characterise RPA's movement along ssDNA.[54, 177] RPA and EcoSSB both contact ssDNA at multiple sites on their protein surfaces, which possibly allows the protein to slide along nucleotides without fully dissociating.[176] SsoSSB only employs a single OB fold and is less likely to be able to partially disengage from ssDNA to allow 1D diffusion and the larger fluctuations observed in Cy3 intensity were therefore attributed to the flexibility of ssDNA and the hydrocarbon chain linking Cy3 to ssDNA, producing a dynamic interaction with SsoSSB monomers. SsoSSB is adept at recruiting other proteins through its acidic terminal tail and it was possible that the constant PIFE observed at 50 nM was due to rapid exchange of unlabelled monomers that was unresolved in the experiment.

The procedure used to label the protein was optimised to promote the efficiency of the labelling reaction and the separation of labelled and unlabelled protein, which produced samples of approximately 100% labelled SsoSSB monomers. Ensemble experiments suggested that the quenching events were due to an interaction between labelled SsoSSB monomers brought into proximity to each other by binding to adjacent sites on the same strand of ssDNA. Self-quenching of cyanine dyes is well reported in the literature and examples of non-fluorescent dimers and resonance energy transfer between like dyes have both been shown to reduce the fluorescence of Cy5and Alexa 647, who share very similar structures.[130, 188] The small size of SsoSSB monomers bound to neighbouring sites placed these acceptor dyes within 3 nm of each other, which clearly had a profound effect on their optical properties. The fluorescence emission from SsoSSB Alexa 647 can be directly excited by using a continuous wave laser with a wavelength at 642 nm. This allows the fluorophores to be imaged on the slide and can be used as a method to count proteins localised on a surface; however this was an impractical method to monitor SsoSSB binding kinetics to ssDNA. Firstly, all observed fluorophores would have to be assumed to be SsoSSB Alexa 647 monomers specifically interacting ssDNA rather than adhering non-specifically to the slide. Secondly, the quenching observed upon the binding of a second adjacent monomer would have been indistinguishable from loss of fluorescence following the dissociation of the first monomer before a second could bind. These difficulties would therefore make it impossible to reliably interpret these traces and draw any unambiguous conclusions and alternative methods were used to describe SsoSSB Alexa 647 binding to ssDNA.

Dye-dye interactions can possibly complicate single molecule fluorescent experiments, and great care is required to choose appropriate dyes that do not give potentially confusing results.[130, 189, 190] In this instance, serendipitous dye-dye interactions gave different relative fluorescence intensities of donor and acceptor dyes that could be used to deduce whether a single or multiple monomers were bound to ssDNA. As a result of two labelled SsoSSB monomers bound to the same ssDNA molecule, the fluorescence from the acceptor dyes was quenched and the donor dye fluorescence was reinstated to a level of intensity consistent with the presence of a protein bound to ssDNA, observed in the PIFE experiments. This unique method of observing SsoSSB binding to ssDNA was used to characterise the initial stages of assembling nucleofilaments and was used to find that nucleofilaments which consisted of a single monomer did not persist as long as nucleofilaments that were two or more monomers in length. This suggested a degree of cooperativity is involved in SsoSSB binding to ssDNA, which was supported by the ensemble results.

From these single molecule experiments, approximately 85% of all binding events observed showed the separate binding of SsoSSB monomers to ssDNA. This indicated that SsoSSB predominately exists as a monomer in solution, and was supported by PELDOR data and other articles in the literature.[114] The 15% of binding events that showed two SsoSSB monomers binding simultaneously could have easily resulted from poor temporal resolution and it was difficult to confidently assign them to any pre-organisation of a multimeric protein. The crystal structure of SsoSSB, despite its similarities to EcoSSB, does not imply that a strong protein

protein interaction exists between monomers and any such cooperative binding was therefore more likely a result of a reduced entropic penalty for a monomer binding to a position on ssDNA adjoining an occupied site, possibly due to the conformations of nucleotides neighbouring a proteinbeing restricted compared to undecorated ssDNA.[27, 191] A cooperative binding process is probably more heavily relied on at increased temperatures,[84] and would result in a more tightly packed nucleofilament providing a more secure protection in an environment that increases the rate insults to ssDNA.

The acidic tail of EcoSSB and other bacterial SSBs has been shown to be involved in determining the extent of cooperativity through interactions with OB folds, and there is a possibility that the SsoSSB flexible tail also plays a role in recruiting other monomers and maintaining a stable nucleofilament.[192, 193] There is a debate in the literature to the extent of the interaction between OB folds and flexible acidic tails which is likely to vary between SSBs.[85, 109] It is improbable that these acidic tails interact strongly enough with OB folds to compete with ssDNA binding; however, the strength of their interaction could be enough to influence the binding of other SsoSSB monomers on to ssDNA.[80] Clearly a nucleofilament is a highly dynamic and carefully balanced environment, where SSBs are required to bind to ssDNA but are also removed at the appropriate time. The strength of the interactions between SSB monomers and the oligonucleotides involved in a nucleofilament are dictated by contributions from base stacking between OB fold residues and bases on ssDNA, electrostatic attractions between ssDNA phosphates and loops on the OB fold, potential interactions between protein surfaces, any entropic and enthalpic penalties

experienced by ssDNA as it is contorted to fit the shape of the nucleofilament, and the proximity of flexible acidic tails potentially recruiting proteins to ssDNA who also show a slight affinity for OB folds.[192, 194] Changing the pH, salt concentration, protein concentration, temperature, or modifying the structure of DNA could alter how one or many of these factors determine the binding modes of SSBs.[82, 195] It is this subtle balance that allows SSBs to bind strongly enough to protect ssDNA but also be removed easily upon the completion of DNA repair or replication. This chapter primarily supports the conclusions made in the previous one, and lends credence to the observation that SsoSSB monomers bind to ssDNA in a cooperative manner. It is also probably accurate to say that there was evidence to show that the dissociation of SsoSSB monomers was also a cooperative process, since a high proportion of filament dissociation occurred by the simultaneous dissociation of both monomers bound to ssDNA. The removal of a neighbouring monomer could possibly weaken the interaction of other monomers with ssDNA through cumulative effects from steric clashing, the disappearance of a potentially stabilising protein protein interface and increased movement of ssDNA.

The decreased affinity of a single monomer could also be related to the weaker binding mode of RPA, where this attenuated affinity restricts the binding of RPA proteins to the region of damaged DNA and prevents the unnecessary melting of large sections of dsDNA and allows the substitution of RPA for nucleases, polymerases and other proteins that metabolise DNA. The RPA 8-10 nt binding mode is regulated by phosphorylation of RPA32 and only employs two OB folds.[43] These OB folds are tethered together to increase the strength of the protein binding to

ssDNA and are structurally remarkably similar to the SsoSSB's OB fold.[27, 42] The monermeric nature of SsoSSB could be a result of an adaption to the high temperatures that are typical in to *S. solfataricus*' environment, since DNA would less likely be tolerant of being restricted to the path that a protein with many DNA binding domains would try and impose on its confirmation. SsoSSB also takes advantage of a third aromatic residue to increase the strength of the base stacking interaction, and this combined with the limited cooperativity observed in SsoSSB could be an alternative to the tethering of OB folds and the phosphorylation of multimeric structures that is more effective in regulating SSB association with ssDNA and is therefore more efficient at fulfilling the roles of an archaeal SSB at elevated temperatures.

4.7. Conclusions

Single molecule TIRF microscopy has been used to supplement ensemble experiments characterising the binding of SsoSSB to ssDNA. Observing single proteins binding on and off DNA may sound trivial given the availability of single molecule techniques - however there were several complicating factors that had to be taken into account to come to an understanding of how this protein fulfils its function; notably, the small size of the protein demanded a closer investigation of the photophysics of the dyes in crowded environments.

Single molecule traces showed FRET, PIFE and quenching events which were similar when compared to ensemble experiments and were interpreted to describe the SsoSSB monomers binding to ssDNA. An efficient FRET from the donor dye on the

ssDNA to the acceptor dye on the first SsoSSB monomer to bind was observed and it was subsequently quenched upon the binding of a second labelled monomer. The second SsoSSB to bind to the ssDNA also allowed the PIFE to be clearly observed and therefore the combination of FRET, quenching and PIFE distinctly showed the first monomer binding to the ssDNA, followed by a second. To the best of the author's knowledge, this is the first time that PIFE, FRET and quenching have been used in this manner. The dwell times of zero, one and two monomers were measured and the rate of an initial monomer binding to ssDNA increased as the concentration of SSB increased. The lifetime of a single monomer was independent of the protein concentration and a similar concentration dependence was observed with the rate of second monomer binding. Two monomers on a single strand of ssDNA were significantly more stable on ssDNA than a single monomer, signifying that SsoSSB exhibited a cooperative behaviour when binding to ssDNA. This cooperativity between independent OB folds possibly indicated a method of producing effective nucleofilaments at high temperatures and highlighted how the structural differences between SSBs influence how they employ similar OB folds to fulfil similar roles maintaining the integrity and regulating the metabolism of ssDNA in different organisms.

5. **SsoSSB binds to RNA**

It was unsurprising, given its name, to find that SsoSSB binds to ssDNA. SsoSSB's high affinity towards ssDNA and the cooperative binding mechanism it employs permit a stable and protective filament to form along a strand of ssDNA. Its unusual monomeric structure weakens the interaction with ssDNA by limiting the number of contact points each protein has with other proteins and the ssDNA; therefore allowing each OB fold to be removed, giving other proteins access to the genetic code. It is this rare monomeric structure coupled with SsoSSB's high affinity for nucleobases and its capacity to withstand extreme temperatures make it a candidate for further roles in nature and in the laboratory outside cellular DNA replication and repair.

ssDNA is not the only single stranded nucleic acid present in *S. solfataricus* cells, and the high temperatures this organism commonly experiences due its lifestyle also puts similar stress on RNA as well as ssDNA.[196] The high concentration of SsoSSB in the cell would make it an ideal chaperone for RNA if SsoSSB's affinity for RNA is comparable for its affinity for ssDNA. There are many examples of OB folds that bind specifically to RNA, such as the ones found in the RNA polymerase subunit RBP8, anitcodon binding domains and RNA capping enzymes,[37] however it is uncommon for SSBs and OB folds in general to be able to tightly bind to both RNA and ssDNA.

The structure of RNA results in a slightly stiffer polymer than that of ssDNA, due to the introduction of the 2' hydroxyl group accompanied by the substitution of the base thymine for the pyrimidine uracil, which essentially results in the loss of methyl groups along the polymer. These changes affect the ring puckering of RNA and increase the base stacking of RNA compared to ssDNA analogues as discussed previously.[14]

Since the interaction between monomers of SsoSSB appears to be minimal in solution, it is anticipated that the cooperative binding is largely due to the restricted freedom of ssDNA at sites adjacent to an already bound SsoSSB.[27, 191] The nearest nucleotides to a bound SsoSSB monomer experience a larger number of steric clashes during changes in conformation, reducing the number of low energy conformations and possibly also raising the energetic penalties in transferring between them. This decrease in movement of the ssDNA would lower the entropic cost of a second monomer binding to these adjacent nucleotides, and facilitate a cooperative binding mechanism.[191]

This could only be an advantage for monomeric species since pre-organised dimers, trimers and tetramers would require the polymer to be guided around the protein to interact with the static OB folds. This would be extremely energetically expensive since the enthalpy gained due to the OB fold binding to the nucleotides must compete with the loss of entropy and any enthalpic penalties accrued in positioning the polymer in high energy conformations. This is observed in EcoSSB where the binding of the second OB fold to ssDNA is an anti-cooperative process due to the

increased repulsion experienced by the phosphate group along the backbone of ssDNA.[197] These penalties could be even more pronounced as the stiffness of the polymer is increased, and could contribute to the reason why multimeric SSBs frequently demonstrate a diminished affinity for RNA compared to ssDNA.

SsoSSB's atypical monomeric structure identified it as a candidate for an SSB that also shows a strong affinity for RNA; however its OB fold also needed to be inspected to see if it was suitable to accommodate ribonucleotides. Other OB folds have been adapted to primarily bind only to ssDNA and the presence of the extra hydroxyl groups on RNA causes steric clashes and again weakens the overall interaction with OB fold.[198] Similarly, in some specialised OB folds, thymine produces a stronger hydrophobic interface than that of uracil and binding to RNA leaves vacancies due to the absence of methyl groups, decreasing the strength of the OB fold's association with RNA.[198] Significant binding of SSBs to RNA would therefore require an OB fold promiscuous enough to not discriminate between ssDNA and RNA. Isothermal calorimetry was used by Dr Lisa Cubbedu to confirm that SsoSSB can bind to RNA and the heat affect from introducing ssDNA and RNA to a solution of SsoSSB can be seen in Figure 5.1.

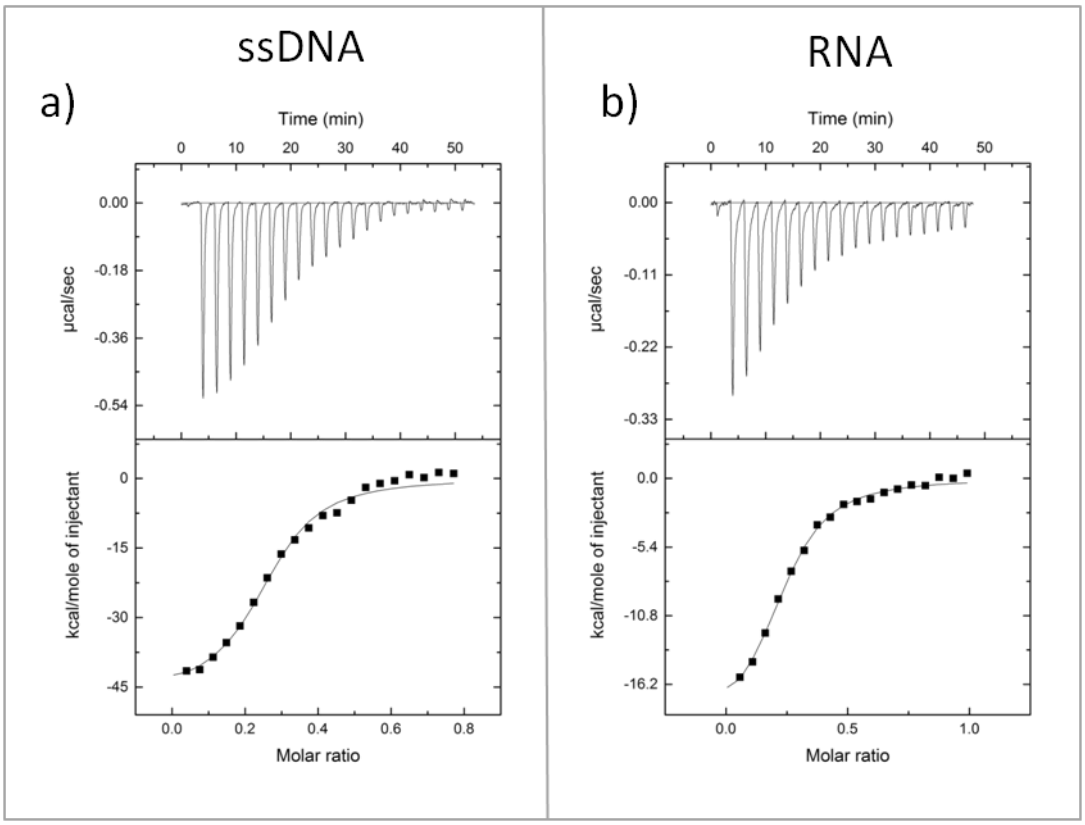


Figure 5.1: The results from ITC, performed by Dr Lisa Cubbedu, showing the interaction of SsoSSB with ssDNA and RNA: with 21dA (a) and 21dU (b).
The figure shows heat effects from injection of ssDNA and RNA, respectively, into buffer. A lower reference power used in these experiments yielded small exothermic peaks at the end of the titration (representing heats of dilution) and no reversion to endothermic peaks.

The residues involved in ssDNA and RNA binding were compared by Dr Roland Gamsjaeger using 15N HSQC NMR spectroscopy and the changes in chemical shifts of residues observed upon binding to RNA were similar to the shifts observed due to binding to ssDNA and are shown in Figure 5.2. The residues involved in RNA and ssDNA binding were mapped onto the crystal structure of SsoSSB and it is clear that ssDNA and RNA interact with similar areas on the surface of the protein.

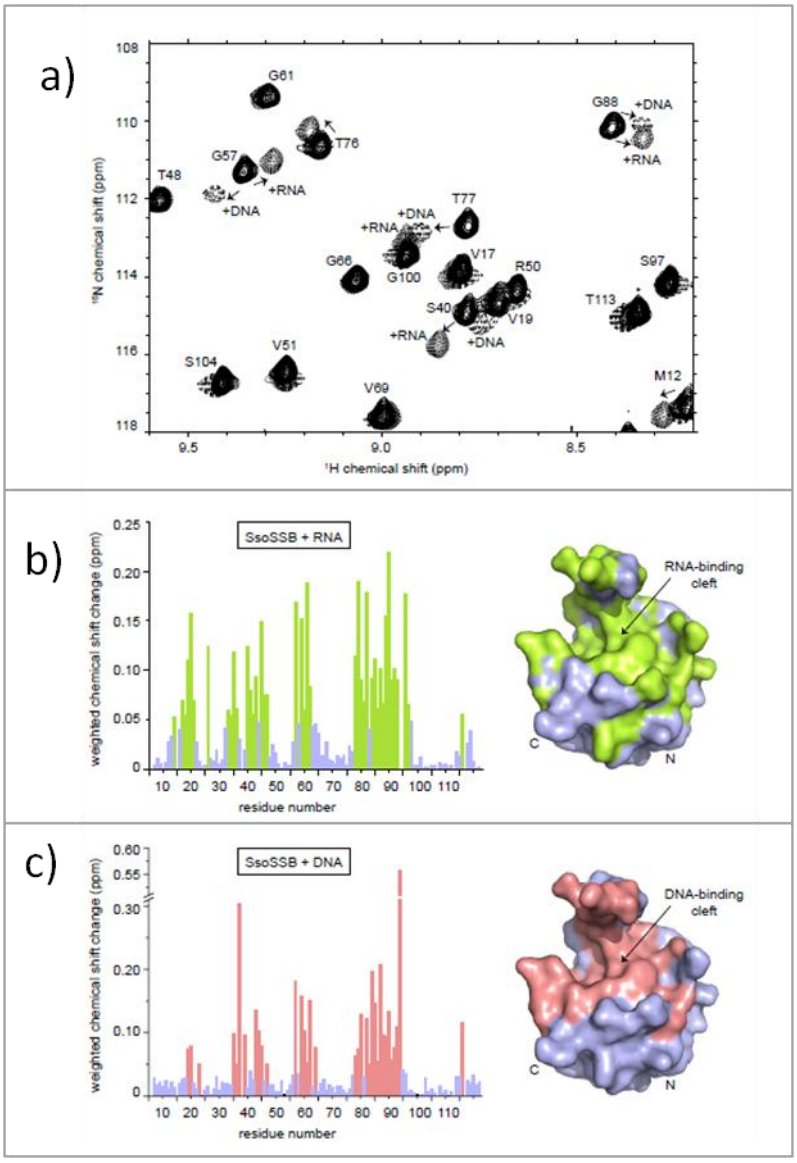


Figure 5.2: NMR analysis of SsoSSB binding to RNA and ssDNA by Dr Roland Gamsjaeger.
Figure (a) shows a section of a 15N HSQC spectrum of ~0.8-1 mM SsoSSB alone (solid lines) and a 1:1 mixture of SsoSSB with 6U ssRNA (dotted lines) as well 6T ssDNA (dashed lines). Assignments and directions of movement are indicated. B/D. Weighted backbone chemical shift changes of HN and N, [199] for SsoSSB upon binding to ssRNA (b) and ssDNA (c), respectively. Residues exhibiting changes larger than the average (binding residues) are coloured in green for RNA and salmon for DNA. Space-filling representation of the crystal structure of SsoSSB (PDB 1O7I) with binding residues coloured in green for RNA (b) and salmon for DNA (c). Note the high similarity of the binding site for RNA compared to DNA.

Gamsjaeger also produced a low energy model of SsoSSB bound to RNA from the NMR structure of SsoSSB bound to ssDNA and there was no suggestion that the 2' hydroxyl group causes any significant clashes within the OB fold. Models of ssDNA

and RNA interacting with the OB fold can be seen in Figure 5.3 and SsoSSB clearly can accommodate the 2' hydroxyl groups without disrupting the position of the nucleic acid inside the OB fold. The dynamics of this system were investigated in this chapter to confirm that SsoSSB employs an equivalent binding mechanism when interacting with RNA to the mechanism observed with ssDNA.

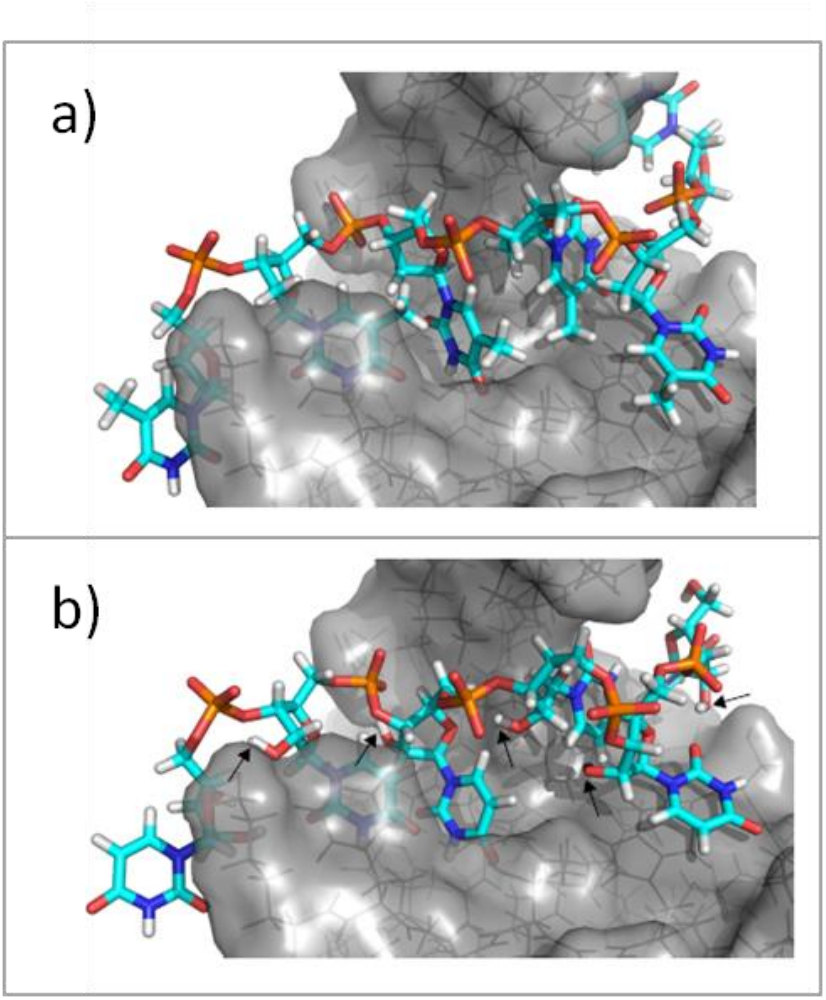


Figure 5.3: Molecules of ssDNA and RNA bound to the OB fold of SsoSSB. Figure made by Dr Roland Gamsjeager.
(a) Energy-lowest NMR structure (PDB ID 2MNA) (Gamsjaeger, Kariawasam et al. 2014) of SsoSSB-DNA complex structure. (b). Model of SsoSSB-RNA structure based on DNA-bound structure. The location of the 2' hydroxyl groups are indicated by arrows.

5.1. Ensemble fluorescence results

Tryptophan quenching and PIFE were used to track SsoSSB binding to 12C RNA Cy3 in investigations that were designed to be analogous to the ensemble experiments described in chapter 3. Again the fluorescence intensity recorded from the tryptophan residues decreased with increasing SsoSSB concentration until the total concentration of RNA was half the total concentration of SsoSSB. The results from this titration can be seen in Figure 5.4(a) as well as titrations of SsoSSB against 20C RNA which showed a similar level of quenching that demonstrated a plateau at around a concentration a quarter of the total concentration of SsoSSB. In Figure 5.4(a), the x-axis has been scaled with respect to the number of nucleotides per SsoSSB monomer present, with each curve indicating that each monomer occupies approximately 5 nt when bound to ssDNA or RNA, irrespective of the length. A modified single binding model was fitted to the titrations and dissociation constants of 3.2 ± 0.3 and 0.9 ± 0.1 nM were calculated for 12C and 20C RNA Cy3 respectively, which were again similar to the dissociation constant previously found for ssDNA and were in agreement with the ITC data.

Similarly, PIFE was used to follow SsoSSB binding to 12C RNA Cy3 and the intensity of Cy3 fluorescence increased with SsoSSB concentration in a study comparable to the previous ssDNA titration, as shown in Figure 5.4(b). The increase in Cy3 fluorescence halted at twice the concentration of RNA and the final Cy3 intensity was again approximately twice the initial value in the absence of SsoSSB. Again this suggested that this length of RNA could only accommodate two SsoSSB monomers and that PIFE was a viable method of tracking SsoSSB binding to RNA.

The dissociation and cooperativity constants for SsoSSB binding to RNA were calculated from fitting the PIFE titration results to the Epstein's finite lattice model, and gave a dissociation constant of 94 ± 6.4 nM and a cooperativity constant of 121 ± 19 which were similar to the dissociation and cooperation constants for ssDNA. This also supported the ITC and NMR data and suggested that SsoSSB uses a similar mechanism to bind to RNA.

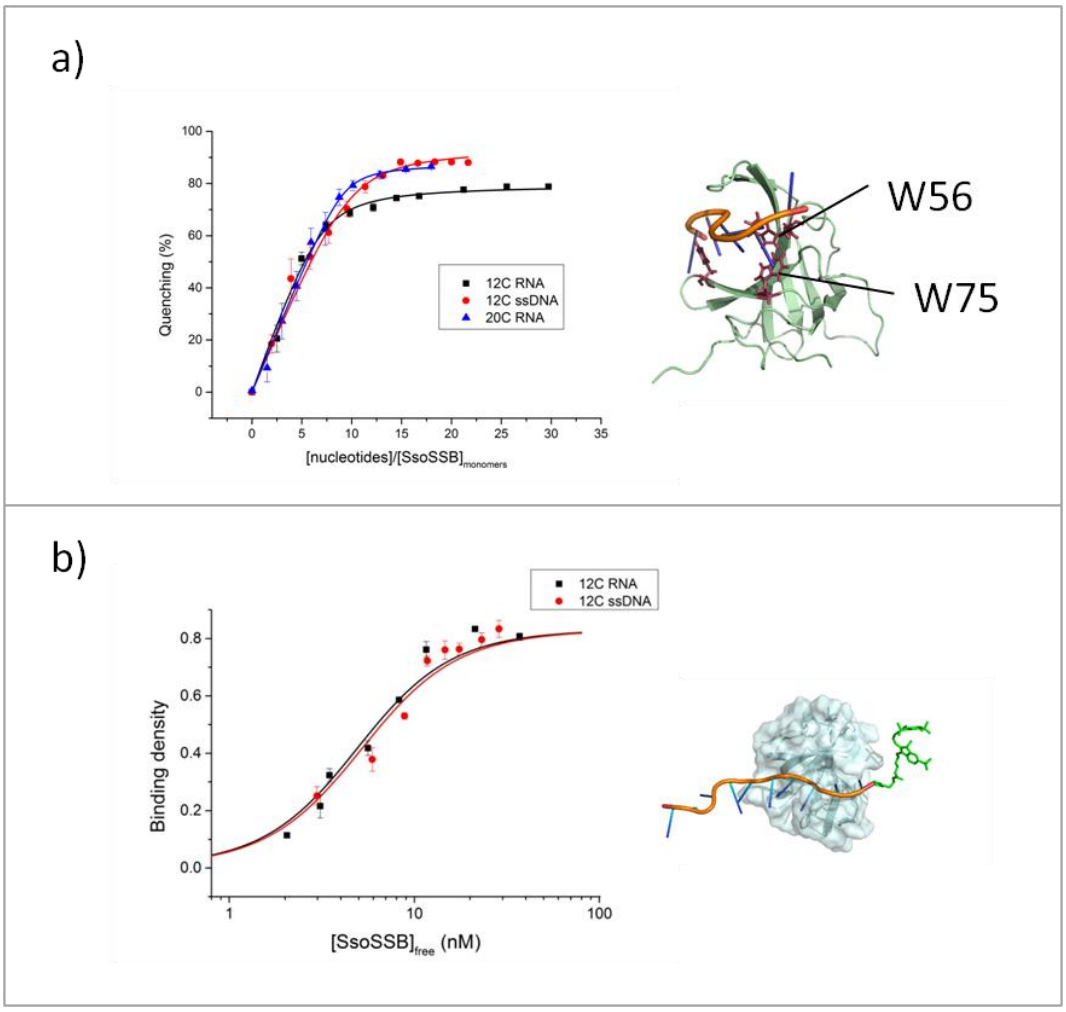


Figure 5.4: Ensemble fluorescence titrations describing the similarities between SsoSSB binding to RNA and ssDNA.
All titrations were completed in triplicate and the mean values are plotted with the standard errors shown as error bars. Figure (a) shows the quenching of the tryptophan residues as 20C RNA, 12C RNA and 12C ssDNA are introduced to SsoSSB in a reverse titration. The x-axis is scaled to show how the quenching of the tryptophan fluorescence was related to the number of nucleotides present per SsoSSB monomer and the data are fitted to a modified single site binding model. Figure (b) shows the PIFE observed as SsoSSB was added to 50 nM 12C RNA Cy3 compared to the titration with 12C ssDNA Cy3. The data is plotted with respect to the concentration of free SsoSSB present in solution and is fitted to Epstein's finite lattice cooperative binding model.

5.2. Single molecule fluorescence results

Single molecule TIRF experiments were carried out to follow the binding of SsoSSB monomers labelled with Alexa 647 to 12C RNA Cy3 immobilised on to a microscope slide passivated with PEG, again similar to the investigation of SsoSSB binding to ssDNA. The observation of brief bursts of high acceptor intensity was coupled with an almost complete quenching of the donor fluorescence - typical of a high FRET state representing the temporary presence of a labelled monomer of SsoSSB bound to the RNA. The average durations of these high FRET states were similar to those measured with ssDNA and were independent of protein concentration. The frequency of these binding events plainly increased with increasing protein concentration, which also suggested that SsoSSB binds to RNA in a similar manner as it does to ssDNA. The frequency of the FRET events can be seen increasing with protein concentration in Figure 5.5(a), where traces from single molecules of 12C RNA Cy3 are presented from SsoSSB Alexa 647 concentrations of 1, 20 and 40 nM respectively. Approximately 200 binding events were measured and average dwell times and observed rate constants were calculated from fitting single exponential curves to the resulting histograms. The average dwell times of the unbound state can be clearly seen to decrease in the green bar charts below the traces, whilst the dwell time of a SsoSSB Alexa 647 monomer on RNA remains constant and is represented by the red bars. Similar traces and bar charts are presented for 12C ssDNA for comparison, and are interpreted to suggest that SsoSSB binds to RNA in a similar manner to ssDNA. The corresponding observed rate constants for RNA binding are given in Table 5.1.

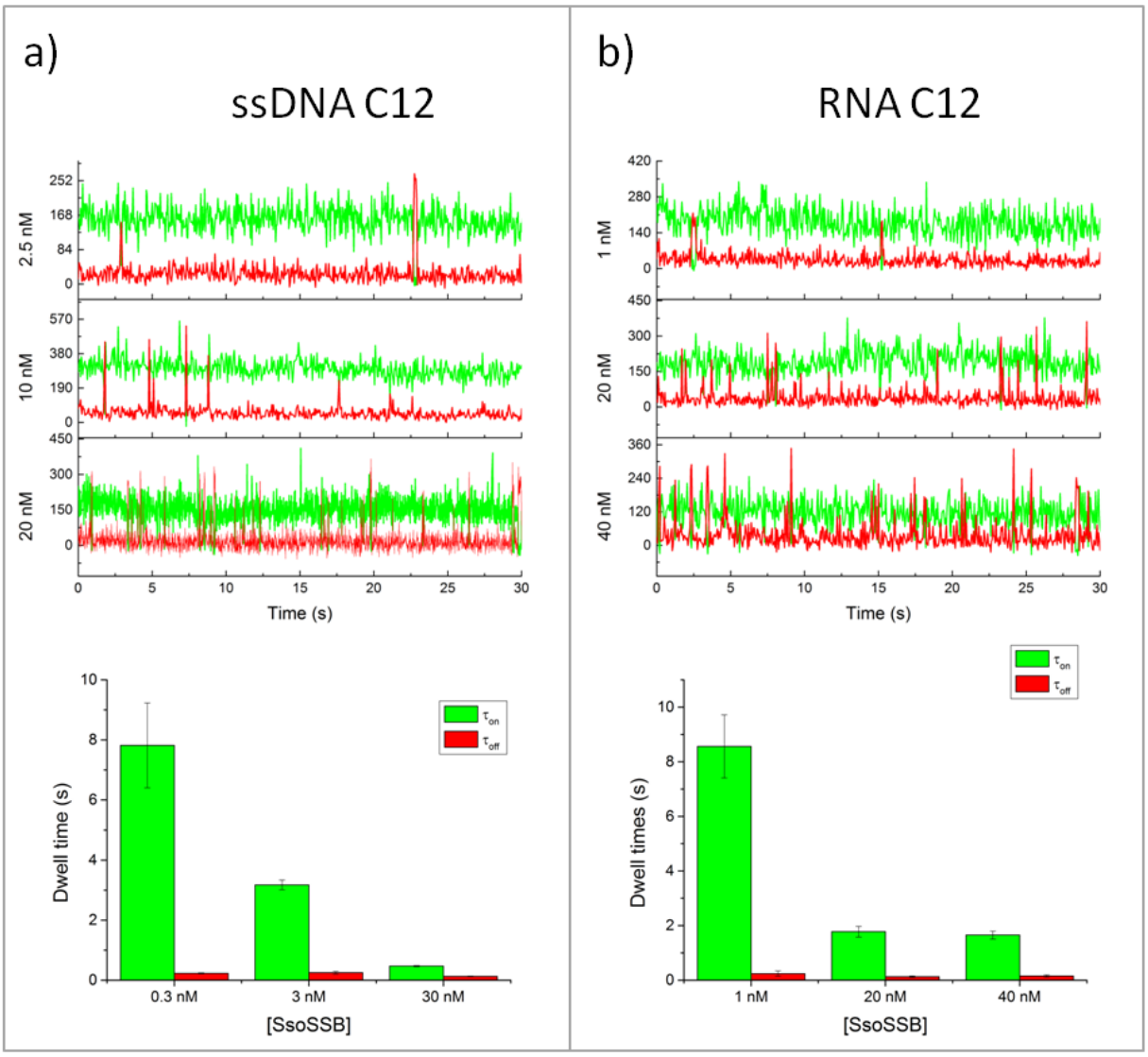


Figure 5.5: Single molecule traces from (a) 12C ssDNA Cy3 and (b) 12C RNA Cy3 immobilised on to a microscope slide in the presence of SsoSSB monomers labelled with Alexa 647.
Single molecule traces displayed an increasing frequency of FRET events typical of SsoSSB Alexa 647 binding specifically to ssDNA or RNA. Below the traces are bar charts showing the average dwell times that represent the times before SsoSSB binds to the oligos (green) and the average time an SsoSSB monomer spends bound to ssDNA or RNA (red).

Table 5.1: The observed rate constants for SsoSSB monomers binding to RNA.

	RNA	
[SsoSSB] (nM)	**k_{on} (s^{-1})**	**k_{off} (s^{-1})**
1	0.11 ± 0.02	4.13 ± 1.6
20	0.56 ± 0.06	7.63 ± 1.8
40	0.61 ± 0.05	6.37 ± 1.5

5.3. Possible roles for RNA binding

SsoSSB has been shown to form filaments on ssDNA to protect its structure whilst it is exposed during DNA metabolism. The previous chapters have demonstrated SsoSSB's capacity to melt the secondary structure of ssDNA and form a linear like filament. As a result, the end-to-end distance of the oligo increases, which was followed using ensemble FRET experiments detailed in chapter 3. SsoSSB was introduced to a 26 nt RNA oligo designed to form a hairpin structure with Cy3 and Cy5 dyes conjugated at each end.The structure of a possible hairpin which it could adopt is shown in Figure 5.6(c). The spectrum of the Cy3 fluorescence at the beginning and end of the titration and the single site binding curve fitted to the data are shown in Figure 5.6(a), again showing a decrease in FRET correlating with increasing SsoSSB concentration. The curve indicates that the fraction of RNA melted reached a maximum at an SsoSSB concentration that is roughly five times the concentration of RNA.

The image in Figure 5.6(b) is of the fluorescence recorded from an electrophoresis gel where RNA labelled with fluorescein was exposed to a *S. solfataricus* exosome. The activity of the Rrp41-Rrp42 hexameric ring and Rrp4 protein on the RNA can be observed by comparing the RNA in the absence of the exosome in the control lane and in the adjacent lane, which shows the difference in the length of the fully digested RNA after it has been introduced to the exosome in the absence of SsoSSB. Increasing the concentration of SsoSSB decreased the concentration of fully digested RNA and produced increasingly concentrated fluorescent signals that were representative of only partially digested RNA strands which can be observed in the

other lanes. This was an example of how SsoSSB could possibly protect RNA molecules against damage and improve the stability of RNA molecules.

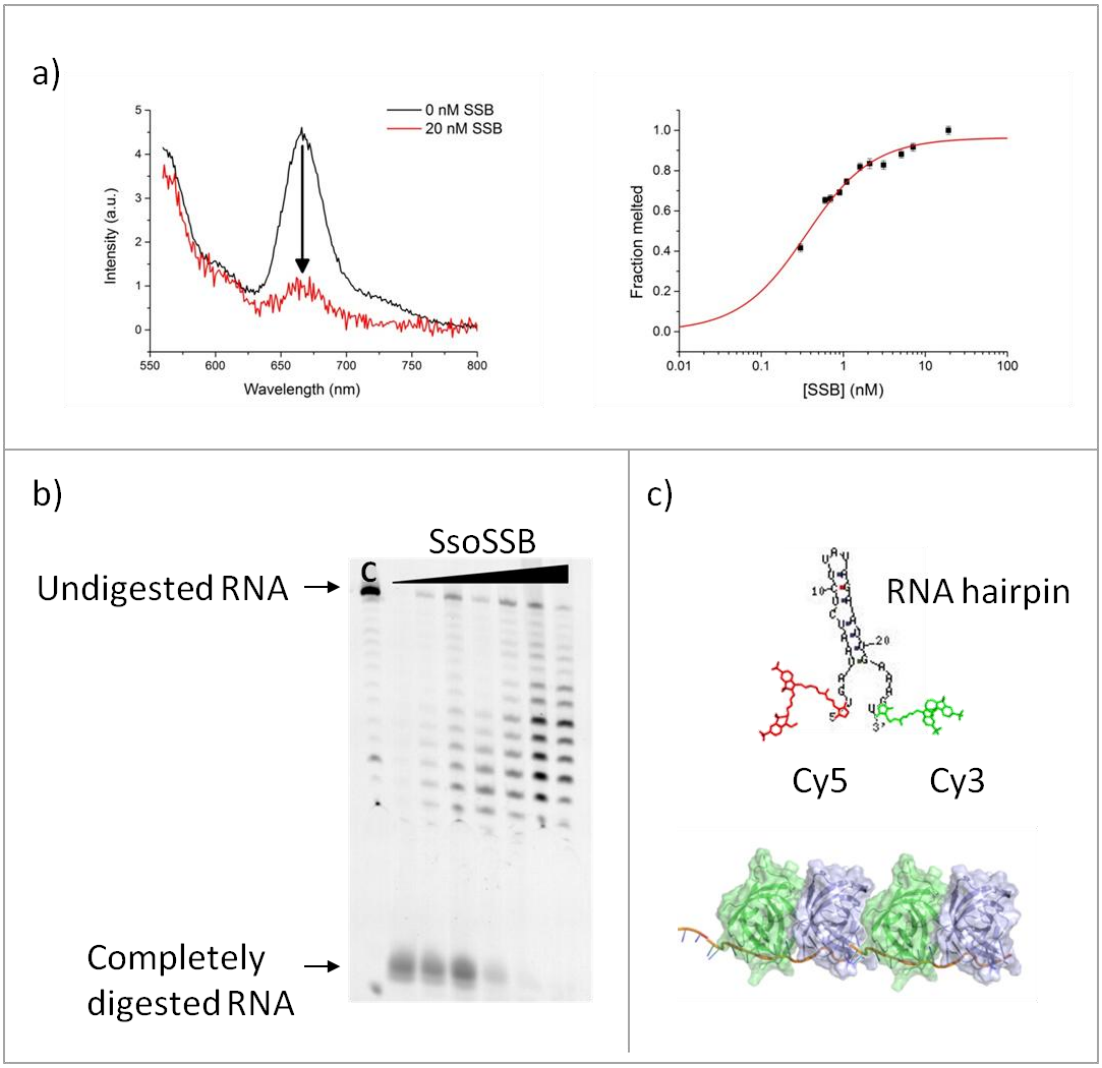


Figure 5.6: SSoSSB binding to RNA may suggest a role *in vivo* such as acting as a chaperone of RNA during transport or metabolism.

This was explored by investigating SsoSSB's capability to melt RNA hairpins. Figure (a) shows the results from ensemble fluorescence experiments, following the FRET from 10 nM doubly labelled hairpin as SsoSSB bound to and removed the secondary structure to form a roughly linear filament. The experiment was completed in triplicate, the standard errors are plotted and a 1:1 binding curve was fitted to the data. Figure (b) shows an image from polyacrylamide gel which was used to determine if SsoSSB could protect RNA from degradation by an archaeal exosome. The 200 nM 25 nt RNA labelled with fluorescein shown in the absence of the exosome in the control C, was fully digested by the archaeal exosome, but is increasingly inhibited as the SsoSSB concentration is increased from 0, 10, 120, 240, 360, 420 to 480 μM. A cartoon of a possible hairpin conformation and the sequence of the RNA hairpin is shown in (c) as well as a schematic of a linear SsoSSB/RNA nucleofilament.

5.4. Discussion

Ensemble and single molecule fluorescence SsoSSB has been used to show that SsoSSB displays a similar affinity to RNA and ssDNA, which has been supported by ITC and NMR data from collaborators. Comparable dissociation constants and cooperative constants suggested that SsoSSB forms analogous nucleofilaments along RNA as it does along ssDNA. Fluorescence intensities from single molecules showed that single monomers persisted on RNA for similar dwell times as ssDNA which again supports SsoSSB interacting with RNA in an equivalent manner to ssDNA.

The double helix is the predominant form that DNA adopts due to its increased stability, and therefore cellular ssDNA only appears transiently, for example, as an intermediate between damaged and repaired DNA. In contrast, RNA is more omnipresent throughout the cell and many mesophilic organisms employ SSBs that discriminate against binding to the more abundant RNA. This maintains a high level of SSBs in the cell that are available to protect the ssDNA when it is exposed and it avoids risking a deficit of SSB due to any unnecessary occupation of SSBs along strands of RNA. The unwanted removal of vital RNA secondary structures at moderate temperatures through SSB binding would also be minimised by the specific SSB binding to ssDNA. Hyperthermophiles, such as *S. solfataricus*, have been known to express specialised RNA binding proteins that are able to assist the RNA in maintaining secondary structures at temperatures approaching and above the melting point of the RNA.[200] This would hinder SSBs access to RNA and prevent them from

interfering with critical RNA molecules that rely on their structures, which would in part negate the necessity to employ an SSB that only binds to ssDNA.

The high temperatures that *S. solfataricus* endure also induce higher rates of damage to unstructured RNA throughout the cell and it appears that SsoSSB is equipped to form protective filaments along ribopolynucleotides in addition to deoxypolynucleotides. The half-life of mRNA has been shown to be greater in *S. solfataricus* than in bacteria, possibly as a result of interactions with proteins such as SsoSSB. The levels of SsoSSB constitute approximately 0.1% of total soluble protein,[201] potentially high enough to cater for both roles protecting ssDNA and RNA but it seems unlikely that SsoSSB is the only RNA binding protein that acts as a chaperone for unstructured RNA molecules. *S. sulfolobus* also expresses a protein named alba, which has been to shown to bind to RNA and is thought to be also employed by the cell to protect nucleic acids *in vivo*.[200] SsoSSB has also demonstrated its ability to melt RNA hairpins and the presence of more than one protein fulfilling similar roles raises the possibility that SsoSSB belongs to a collection of proteins that is responsible for maintaining the integrity of unstructured RNA. Molecules of RNA are employed throughout the cell in various processes and the strain on the workload of SsoSSB would be lightened by different proteins contributing towards protecting RNA at specific points in the lifetime of the riboxyoligonucleotides.

This affinity towards RNA could also be an advantageous feature of OB folds that could have been initially developed in a pre-DNA world rather than a result of

specific evolution to survive in extreme temperatures. There is evidence that RNA pre-dates DNA, [202] and there are many examples of OB folds that selectively bind to RNA such as tRNA binding proteins described as chaperones to protect bacterial tRNAs [203, 204], N-terminal anti-codon binding domains of various class II tRNA synthetases [203, 205], translational initiation factors and bacterial and archaeal ribosomal proteins.[206]

SsoSSB has already been shown to interact with RNA polymerase via SsoSSB's acidic C terminal tail and stimulate transcription *in vitro*,[86] and is a likely candidate to protect mRNA but it remains to be seen if SsoSSB is involved in shielding other unstructured RNA strands from damage. It is certainly possible that SsoSSB interacts with other proteins that metabolise RNA, such as RNA helicases, but whether or not SsoSSB is involved in processes such as ribosome biogenesis requires further investigation.

5.5. Conclusions

SsoSSB has been shown by a number of techniques to bind to RNA with an affinity and mechanism that is comparable to SsoSSB's binding to ssDNA. SsoSSB's OB fold was shown to be able to accommodate the differences between the molecular structures of ssDNA and RNA through *in silico* models and NMR experiments *in vitro*. Ensemble fluorescence results were supported by ITC to show that SsoSSB demonstrated similar dissociation constants and cooperative constants when binding to ssDNA and RNA. SsoSSB monomers were observed binding to RNA, displaying similar dwell times to those measured on ssDNA. This pointed towards filaments

being formed on RNA, which were observed by the extension of end-to-end distances of RNA hairpins being melted by SsoSSB seen by ensemble fluorescence experiments and were analogous to those that had been observed on ssDNA. SsoSSB's role as a chaperone was further highlighted by SsoSSB's ability to retard the activity of an exosome that digests unstructured RNA, again demonstrating the formation of protective SsoSSB nucleofilaments along the length of RNA.

6. Further work

6.1. Applications in high temperature and salt concentrations

It has already been established that SsoSSB has a high affinity for ssDNA,[27, 86, 109] and that it can withstand temperatures as high as 95 ºC without loss of structure or function.[112] SSBs from other organism have previously been used to bind to ssDNA during nucleotide synthesis, as biomarkers for ssDNA on Au surfaces and it would be likely that SsoSSBs structure would allow it to be used in similar roles.[207] Its ability to withstand high temperatures could possibly lend SsoSSB certain advantages over other SSBs, possibly in improving the efficiency of PCR in a similar manner to other SSBs.[208]

Many other SSBs only use two aromatic residues in their OB folds to bind to ssDNA. By employing a phenylanaline and two tryptophans residues, SsoSSB's OB fold is probably less reliant on electrostatic interactions to anchor the ssDNA in place and it would also experience a greater contribution from π-π stacking during binding events.[114] As a result, SsoSSB's OB fold could offer some resistance to any increases in salt concentrations that would adversely affect binding to ssDNA.[209]

One avenue of research that could potentially be explored could be sequencing of single molecules of DNA. Nanopore technology is highly competitive and can produce potentially very lucrative projects and SsoSSB's robustness could make it a reasonable candidate for use in these experiments.[210] The movement of nucleic acids passing through nanopores can allow observations to be made about the DNA's

sequence, structure and the proteins bound on that strand.[211] Nanopores can be produced in a variety of membranes, and one particular route of fabricating them is through controlling the dielectric breakdown of a membrane to produce nanopores of a defined size.[212] The salt concentration used throughout this thesis was 5 mM; however, for SsoSSB to be useful in the development of these laboratory practices it needs to be able to operate in LiCl concentrations of 1.5 M and above.

6.1.1. Ensemble fluorescence

PIFE was used to assess if SsoSSB can still bind to ssDNA in high LiCl concentrations and Figure 6.1 shows an approximately two fold enhancement indicating that there was still significant binding of SsoSSB in a buffer containing 1.5 M LiCl. This was only achieved binding to 20C ssDNA and not to 12C. The binding of SsoSSB should be independent of the length of oligonucleotide with a non-cooperative binding mechanism, and previous experiments have shown that SsoSSB monomers do not express any anti-cooperative behaviour on 12C ssDNA; therefore the stronger affinity for a longer ssDNA again supports a cooperative binding mechanism given SsoSSB's monomeric structure.

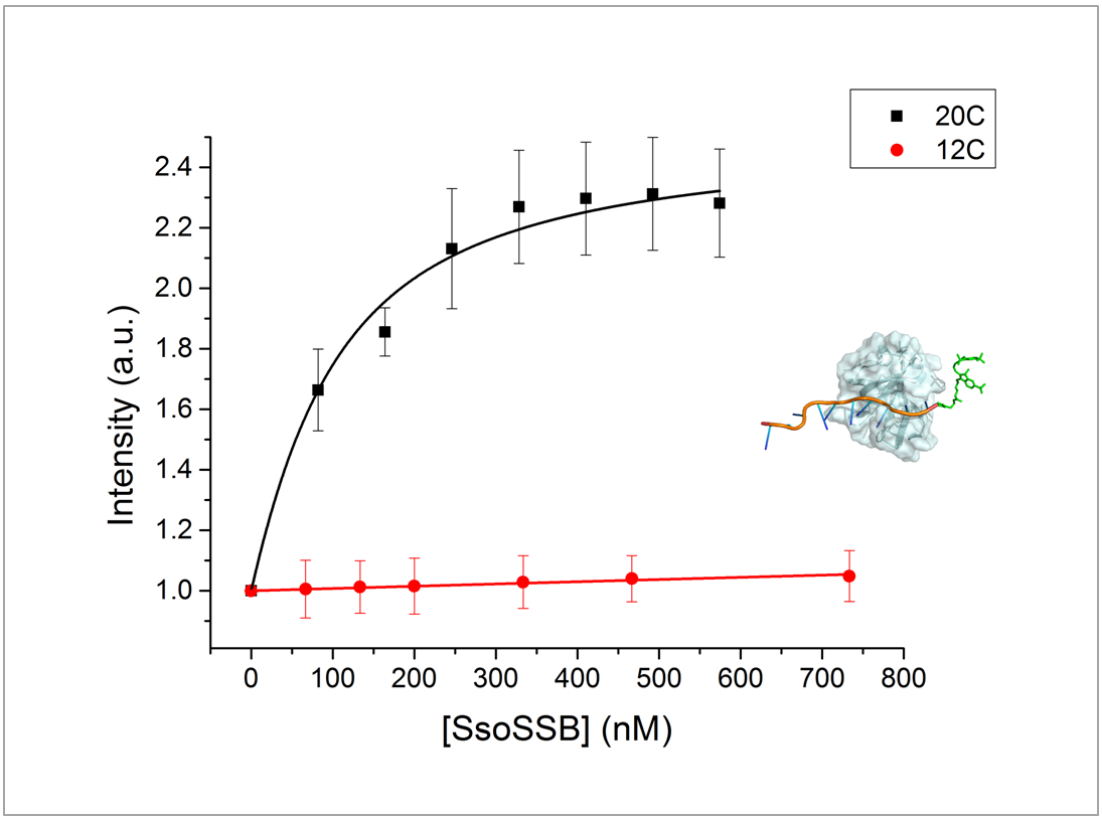


Figure 6.1: The enhanced fluorescence from 50 nM 20C ssDNA Cy3 was observed upon the addition of unlabelled SsoSSB in 1.5 M LiCl.
The 2.3 fold increase in fluorescence plateaus at approximately 200 nM. In contrast, similar additions of SsoSSB to 12C ssDNA indicated that SsoSSB could not bind to shorter lengths of ssDNA at high salt concentrations. The titrations were completed in triplicate and the mean values are plotted with the standard error shown as error bars.

Ensemble fluorescence experiments were also used to demonstrate that an increase in temperature also weakened SsoSSB's affinity for ssDNA. Figure 6.2 compares the fluorescent enhancement observed upon SsoSSB binding to 12C ssDNA Cy3 at 25 °C and 65 °C. At both temperatures SsoSSB was still able to induce a two-fold increase in fluorescence by binding to the ssDNA, however the titration clearly suggested that SsoSSB's affinity for ssDNA was weakened at the elevated temperature.

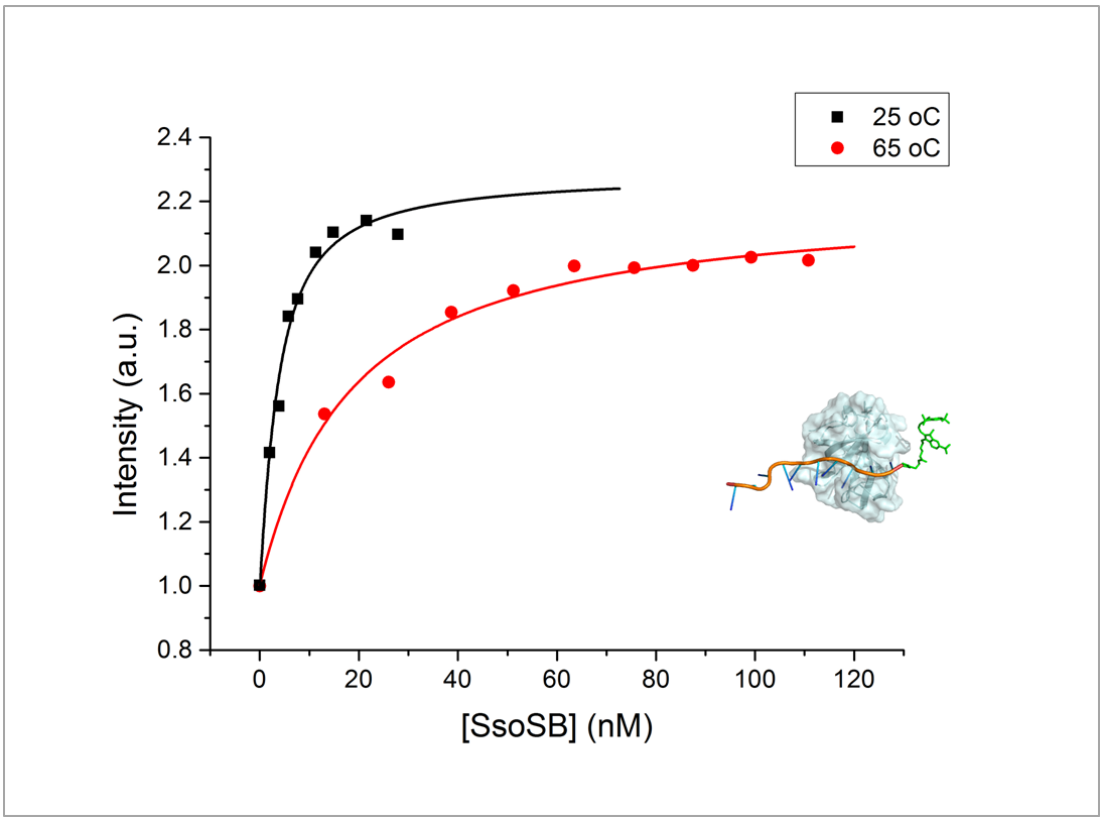


Figure 6.2: PIFE was used again to study SsoSSB binding at temperatures closer to the native environment of *S. solfataricus*.
65 °C was chosen as a compromise so that the fluorescence dyes were not significantly affected by the increase in temperature. Titrations at both temperatures showed an approximately 2 fold enhancement when unlabelled SsoSSB was added to 10 nM 12C ssDNA, however the SsoSSB exhibited a weaker binding at the elevated temperatures.

6.1.2. Discussion

The enhancement of the fluorescence observed upon SsoSSB binding to the longer ssDNA suggested that SsoSSB monomers bind tighter to longer strands of ssDNA. SsoSSB has been found to bind to ssDNA as a monomer with a footprint of approximately 5 nt, to form linear-like filaments. Increasing the length of ssDNA allowed more than two SsoSSB monomers to bind to the oligo at any one time. This permitted SsoSSB monomers to contact more than one other monomer whilst bound to ssDNA, since there was sufficient room on either side of monomers binding to the middle of the strand for another monomer to bind at an adjacent binding site. The

difference between binding to 12C and 20C oilgos was very similar in low salt environments, however SsoSSB's high affinity for ssDNA in solutions with low ionic strengths and the relative sensitivity of ensemble fluorescence techniques could have masked any difference between binding events to the different oligos. Increasing the salt concentration shielded electrostatic charges, weakening the attraction between the sugar-phosphate backbone of ssDNA and the OB fold, and also increased the number of cations that are condensed during binding, again diminishing SsoSSB's overall affinity for ssDNA.[209] This could have possibly exaggerated the differences between SsoSSB binding to 12C and 20C, exhibiting a tighter binding to the longer oligo. This hinted at an applied role *in vitro* for SsoSSB in high salt environments, and together with SsoSSB's tolerance of high temperatures, these results could make SsoSSB a candidate as a substitute for other SSBs that are already applied in more moderate conditions. Regardless of future uses of SsoSSB, further studies at different salt concentrations, with monovalent and multivalent cations, and temperatures are required to guarantee that the binding behaviour observed in this thesis is an accurate portrayal of SsoSSB's native activity.

6.2. The role of the acidic C terminal tail

SsoSSB shares structural features with both bacterial and eukaryotic SSBs. The OB fold is a common motif found throughout all SSBs, whereas the acidic flexible C terminal tail is more frequently seen in examples from bacteria. There is no doubt that these unstructured domains play a crucial role in recruiting other proteins to ssDNA - SsoSSB alone has been shown use its own tail to interact with RNA polymerase.[86] There is also a debate in the literature that suggests further roles for the C terminal tail, implying that the acidic tip of the flexible tails could also engage with positively charged OB folds on SSBs in addition to interacting with other proteins.[79, 85, 192] These mechanisms are shown in Figure 6.3, and propose that C terminal tails directly compete with ssDNA to occupy the OB fold, which facilitates dissociation from ssDNA and also regulates the extent of cooperative binding through interactions with OB folds on adjacent proteins.

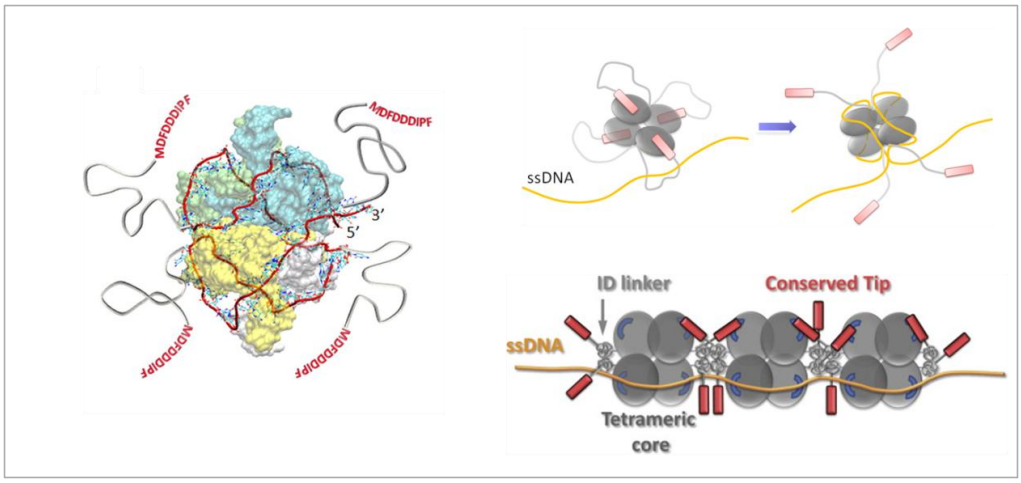


Figure 6.3: Cartoons of the four C terminal tails of EcoSSB.
Left shows the four monomers coloured in blue, green, yellow and white and the four intrinsically disordered (ID) linkers ending in an acidic tip. The acidic residues are shown in red font at the end of each tail (MDFDDDIPF) and the path of ssDNA being wrapped around the four OB folds is shown as a red line. Top right shows how the C terminal tails (grey, red tip) are displaced by ssDNA (yellow line) as it binds to the four OB folds. In this arrangement the C terminal tails now protrude out in to solution and are available to interact with and recruit other proteins to ssDNA. Bottom right shows how the C terminal tail, made up of the ID linker (grey) and the conserved acidic tip (red) influences cooperative binding through interactions with the OB folds (blure) on the adjacent tetramers bound to ssDNA (yellow). Figures modified from Kozlov *et al*. (2010) and Kozlov *et al*. (2015)[85, 192]

EcoSSB is the classic model used to study interactions between C terminal tails and OB folds, but the extent of the interaction varies between SSBs. Other examples include SSBs from *Plasmodium falciparum*, *Klebsiella pneumonia*, *Salmonella enteric*, *Pseudomonas aeruginosa* and *Thermus thermophilus*.[175, 192, 193] The last eight to ten amino acids of the C terminal tail are generally acidic, whereas the intrinsically disordered linker varies in length and composition. The sequences of the final residues at the C terminal from several SSBs are shown in Figure 6.4 and demonstrate the conserved acidic nature of the tip of the disordered tail. The overall lengths of the tails from different SSBs have been recorded between 25-135 amino acids, and their sequence defines whether the polymer is more likely to adopt globular or more extended formations.[192] This also influences the strength of the interaction between C terminal tails and OB folds, which has been shown to govern which binding mode the SSB employs and influences the level of cooperativity those binding modes exhibit.[192] Attempting to predict the function or behaviour of disordered peptide chains from their sequence is extremely challenging and even with the support of empirical evidence there are opposing views to the degree of interaction these C terminal tails have with OB folds.

Organism	C terminal tip
T. aquaticus	FPPEEE-LPF
T. thermophilus	FPPEED-LPF
D. radiodurans	FPPEEDDLPF
M. tuberculosis	FGFFDDEPPF
M. smegmatis	FSGAADEPPF
K. pneumonia	MD-FDDDIPF
S. enteric	MD-FDDDIPF
P. aeruginosa	YDSFDDDIPF
E. coli	MD-FDDDIPF
P. falciparum	DKMNVQEFEE
S. sulfolobus	RQENEEGEEE

Figure 6.4: The end sequence of C terminal tails from the SSBs from various organisms, showing that the acidic residues (D and E) are commonly found at the tip of the disordered tails.

The contact of the tails with the OB fold must be transient by definition, since a too strong affinity for C terminal tails would impede OB folds from effectively binding to ssDNA and would inhibit SSB's efficiency at fulfilling its roles. Crystal structures of the full length EcoSSB do not indicate that its C terminal tail adopts a rigid position, and its inherent flexibility suggests that any interaction with OB fold is likely to be dynamic and/or disordered.[81] Sensitive experimental techniques must therefore be used in attempting to define the extent of the role that the C terminal tail plays with the OB fold in binding to ssDNA. A truncated SsoSSB, whose tail has been removed at A114 through trypsin digestion, showed comparable behaviour to the full length protein at the ensemble level.[109] Wadsworth observed similar binding kinetics by following the quenching of tryptophan produced by the addition of ssDNA to the truncated and full length SsoSSBs. The fluorescence emission from tryptophan residues is possibly too weak to observe any subtle changes in binding behaviour and further enquiries are needed to review what level, if any, the C terminal tail has an effect on SsoSSB's binding to ssDNA.

6.2.1. *In silico* model of SsoSSB tail binding to the OB fold

As an initial investigation, an *in silico* model was produced to explore whether or not the binding of the tail was viable. Autodock Vina and MGLTools were used to visualise possible low energy conformations that the tip of the C terminal tail could occupy within the OB fold.[213] The length of the full terminal tail (33 amino acids) is comfortably long enough to allow the interaction between the tip and the OB fold without placing the tail under steric stress, which is shown pictorially in Figure 6.5. It was therefore reasonable to treat the tip of the tail as an independent ligand and the final eight residues of the tail were placed in box of water (28 x 38 x 32 Å) that was positioned around the surface of the OB fold. An energy minimisation gave a number of possible conformations and Figure 6.5 shows the lowest energy conformations available to the final eight residues that the acidic tail can adopt inside the OB fold as determined by Autodock Vina.[213] This was used to find nine different low energy conformations of the C terminal tail inside the OB fold, with binding affinities ranging from $\Delta G =$ -6.4 to -5.9 kcal/mol. These can be translated into dissociation constants, Kd, simply using the relationship $\Delta G = RTln(Kd)$, and on average were found to be approximately 40 μM at $T = 298.15$ K and where $R = 1.99$ cal/(mol.K.)

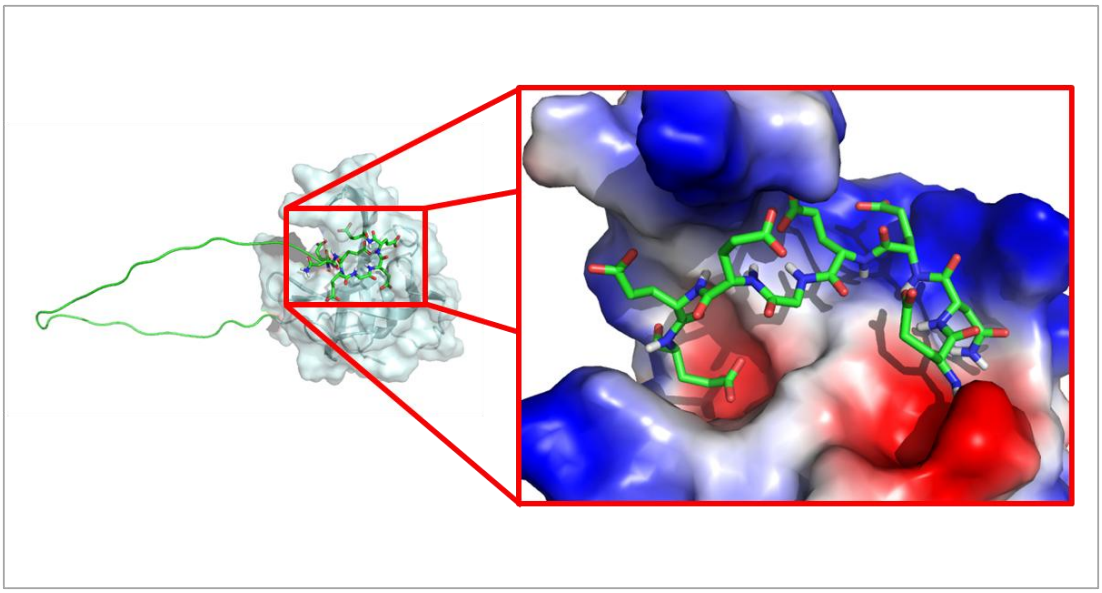


Figure 6.5: Left shows how SsoSSB has a 33 amino acid C terminal tail that is long enough to allow the acidic tip to interact with the OB fold. Right shows a model produced by Vina Autodock and MGL Tools.
The figure displays the lowest energy conformation available to the 8 amino acids from the acidic tip of the C terminal tail inside a 28 x 38 x 32 Å box surrounding the OB fold. The colours blue to red show the negative and positive nature of the electrostatic surface of the OB fold respectively.

6.2.2. Ensemble measurement of k_{off}

The truncated and full length SsoSSBs were introduced to 12C ssDNA Cy3 and an increase in fluorescence was observed in both cases that were typical of SsoSSB binding. An excess of unlabelled 12C ssDNA was then added to the solution and a decrease in enhancement was observed as more SsoSSB monomers dissociated from the labelled ssDNA and bound to the more abundant unlabelled oligo. This decrease of fluorescence was followed over time and fitted to a single exponential, shown in Figure 6.6. The truncated SsoSSB exhibited a slightly slower rate of dissociation from the labelled ssDNA, possibly hinting towards a mechanism where the tail does play a role in facilitating the dissociation from ssDNA.

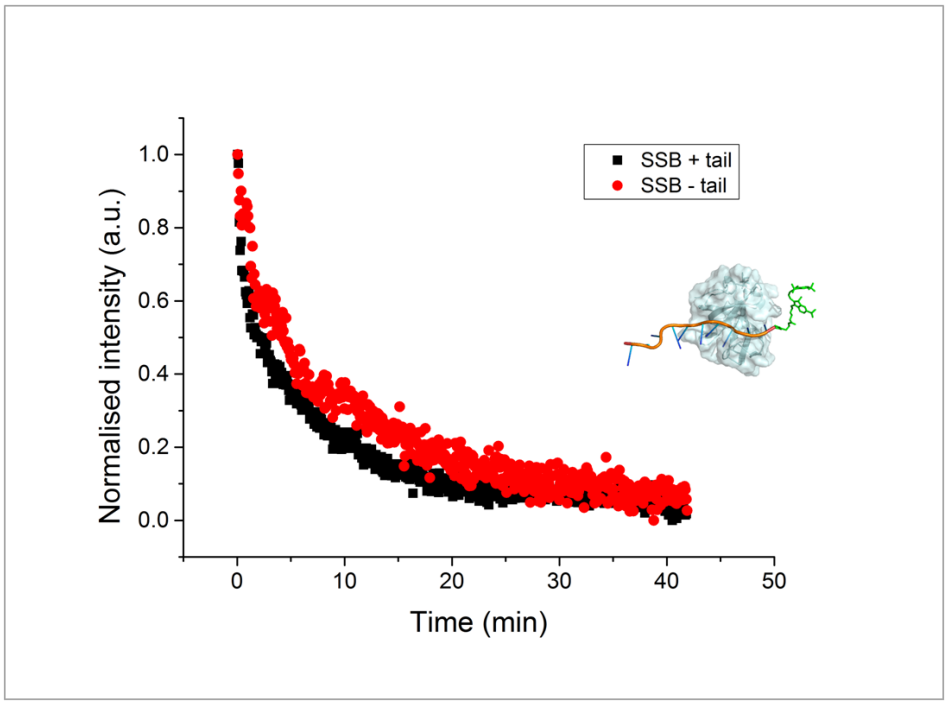


Figure 6.6: The decrease in PIFE was used to track the dissociation of 50 nM unlabelled SsoSSB from 25 nM 12C ssDNA Cy3 as it binds to a 1000 times excess of unlabelled 12C ssDNA.
The truncated SsoSSB (red) had its C terminal tail removed by proteolytic digestion and shows a similar behaviour to the full length protein (black.) The protein was allowed to bind to the 12C Cy3 before the addition of the excess unlabelled oligos at t=0.

6.2.3. Discussion

The evidence for a significant electrostatic interaction between SsoSSB's OB fold and the tip of its C terminal tail appears to be weak. Any affinity that the OB fold had for the C terminal tail only had a minor effect on the binding of ssDNA. The OB fold's affinity for the tip of the C terminal seemed to be three orders of magnitude weaker than the affinity demonstrated for ssDNA. This comparison was made between an *in silico* model and *in vitro* studies which are not equivalent, since the *in silico* model has to make many assumptions in order to minimise the atomic positions, yet show a distinct difference in the strength of interactions. Further *in silico* experiments to calculate the binding affinities of ssDNA with OB folds from

SsoSSB and other SSBs would give more trustworthy conclusions and possibly quantify the relative affinities of the OB fold for the competing molecules.

Ensemble fluorescence studies displayed a slight increase in the average times it took for truncated SsoSSB to dissociate from ssDNA compared to the full length SsoSSB, hinting at a possible role in removing SsoSSB monomers. The OB fold's affinity for the acidic tip on the same monomer was possibly only strong enough to have a slight influence when SsoSSB dissociated from the ssDNA - thus playing a limited role, if any, in the binding mechanism to ssDNA.

Single molecule techniques provide powerful tools used to investigate proteins. The results from the single molecule studies from chapter 5 appeared to show that SsoSSB binding was a positively cooperative process. There single molecule fluorescence traces were used to distinguish whether a monomer had bound adjacent to an existing monomer on the ssDNA, by investigating FRET, PIFE and quenching events. It has been suggested that EcoSSB uses interactions between C terminal tails and the OB folds on different tetramers to ensure that binding to ssDNA is highly cooperative.[192] Further investigations at the single molecule level with the truncated SsoSSB could be completed using the analysis in chapter 5 to measure the dwell times of two monomers on ssDNA and delineate the C terminal's role in determining the extent of cooperative binding between monomers.

This could also be applied to the C terminal's role in dissociation since many traces showed the dissociation of two monomers within 33 ms of each other. Due to the

limited temporal resolution of the microscope, these events could not be defined as two monomers leaving simultaneously or as two separate events that occurred in very quick succession. Investigating the relative frequencies of simultaneous and sequential dissociation for truncated and full length SsoSSBs could explore whether these events are truly cooperative dissociation and how much of a role the C terminal tail has.

This work could be extended to investigate the effect salts have on ssDNA binding since the electrostatic interaction between the OB fold and the C terminal tail is highly dependent on the ionic strength of its environment.[209] If C terminal tails were having an effect on adjacent monomers, this could be partially responsible for SsoSSB's cooperative binding and dissociation to and from ssDNA. There are many other factors that govern the binding behaviour of SsoSSB, but with analogous C terminal tails potentially having large effects on other SSBs there possibly are unanswered questions surrounding the level of responsibility that the C terminal tail has in determining how SsoSSB binds to ssDNA.

If the C terminal is found to be effectively a tether that recruits other monomers to ssDNA, it would be potentially interesting to explore whether mimicking RPA and covalently linking SsoSSB monomers together would affect the binding of ssDNA. There are no examples in the literature of studies involving this type of mutant of SsoSSB, and a more detailed investigation on correlating the oligo length with SsoSSB's affinity for ssDNA could determine how controlling the size of a

nucleofilament could be used to potentially tune the strength of binding to an oligo and any possible applications that this could be used in.

6.3. Conclusions for further work

S. solfataricus is an extremophile, surviving in temperatures of 95 °C and a pH of 2.0.[214] It has adapted to living in harsh environments by employing specialised protein machinery, that can perform their functions at elevated temperatures in order to survive in such a harsh environment.[215] EcoSSB experiences weaker affinities for ssDNA at high temperatures and its binding mechanism can shift to rely on a more cooperative binding mechanism to effectively form complete nucleofilaments.[84] This thesis has focused on demonstrating that SsoSSB can exhibit cooperative binding to ssDNA, but further work is required to fully understand if similar effects are truly implemented at temperatures more similar to those experienced by *S. solfataricus* in its natural environment.

The kinetics of SsoSSB binding to ssDNA has been explored in chapters 3 and 4, which has been used along with other reports in the literature to paint a more complete picture of how many factors influence the association between SsoSSB and ssDNA. There has been a focus on how the electrostatic and base stacking interactions between ssDNA and the OB fold provide an enthalpic driving force for the SsoSSB to bind to ssDNA,[194] and how the presence of the OB fold on ssDNA distorts neighbouring nucleotides and possibly permits cooperative binding between monomers.[27, 193] The C terminal tail has been designated a largely non-binding role;[109] however, this could possibly need to be re-addressed in the light of other reports that suggested a possible role in ssDNA binding in other organisms. SSBs from different organisms demonstrate varying extents of how much the electrostatic interactions, base stacking and C terminal tail are involved in binding to ssDNA,[192]

and further work is required to fully characterise SsoSSB's binding mechanism, and whether the degree of influence the OB fold and C terminal tail have on that binding mechanism changes when *in vitro* studies are performed in conditions more comparable to those experienced *in vivo*.

Once these binding mechanisms are better understood, the application of SsoSSB in extreme temperatures and possibly high salt concentrations could be explored. There are already examples of other SSBs that are employed in PCR and nucleotide synthesis reactions,[208] and it would be interesting to see what future roles that SsoSSB could take on.

The novel in-trace analysis of FRET, PIFE and quenching events could possibly help in quantifying the more subtle details of SsoSSB binding to ssDNA highlighted above. Single molecule fluorescence is a powerful technique capable of providing comprehensive information on the kinetics of molecular systems. It is also allows a closer look at the molecular behaviour of the fluorescence dyes themselves, and there are many examples of how the photophysical properties of fluorescent dyes are perturbed through events such as dye-dye interactions, base stacking and photoisomerisation.[182, 190] The work in this thesis has been able to exploit the behaviour of dyes in crowded environments and it would be interesting to further define the physical origin of the quenching as well as exploring other potential molecular systems where this combination of techniques can be used. There are other examples of monomeric SSBs to which this technique could be applied as well as

systems that exhibit self assembly, aggregation or structural changes that would place more than two fluorescence dyes in proximity with each other.

7. **Conclusions**

This thesis has largely described the process of how SsoSSB binds to ssDNA and has discussed how this relates to its function of maintaining the structure of intermediate ssDNA during vital processes such as DNA repair, replication and transcription. SsoSSB was shown as a predominately monomeric protein capable of binding to ssDNA to form linear-like nucleofilaments. The duration of time that SsoSSB monomers were bound to ssDNA was measured in real time at single molecule resolution and also at an ensemble level. It was found that SsoSSB demonstrated a moderately cooperative binding mechanism when associating with ssDNA. SsoSSB was also shown to be capable of binding to RNA in an analogous manner to ssDNA which indicated a parallel role for SsoSSB acting as a chaperone for RNA in addition to its responsibilities preserving the integrity of ssDNA.

SsoSSB's monomeric structure presented a rare opportunity to examine single OB folds binding to ssDNA. The higher order tertiary structures of other SSBs place their OB folds in more well defined positions relative to one another and the relatively weak protein protein interactions observed for SsoSSB showed the relatively diminished affinity of a single monomer compared to the cumulative strength of multiple OB folds bound to ssDNA. This enhanced stability of more than one SsoSSB monomer appears to originate predominately from the perturbation of oligonucleotides upon the presence of OB folds however the affinity for ssDNA is governed by a balance between other factors including protein protein interactions, the flexibility of the ssDNA, base stacking and electrostatic interactions between the

ssDNA and OB folds and also the extent of the role of C terminal tails in ssDNA binding. The relative importance of these factors varies across the wide range of SSBs and all are in turn influenced by the structure of the SSB, salt concentration, temperature, pH and protein concentrations.

This all points towards SSBs bound to ssDNA in a filament being a highly dynamic and subtly balanced environment rather than a simplistic static nucleofilament. Monomers of SSBs may in fact only have relatively transient interactions with ssDNA which allows SSBs with higher order structures capable of multiple contacts along ssDNA to slide along oligos. The presence of other proteins could disturb a number of the factors that influence SSB affinity for ssDNA mentioned above and could facilitate the substitution of SSB for other proteins that are involved in metabolising DNA.

An investigation into the photophysics of fluorescent dyes was also undertaken and a novel analysis of fluorescent intensities from single molecules was developed that combined FRET, PIFE and self-quenching. The fluorescence properties of dyes report on their microenvironment and remain a powerful tool used to investigate the movement of molecules from nanomachines to protein complexes. In this thesis, this was exploited to unambiguously interpret single molecule data to show sequential SsoSSB monomers binding to ssDNA. This type of analysis has the potential to be used by the single molecule community to take advantage of behaviour of fluorescent dyes in crowded environments and deduce the dynamics of similar biologically relevant systems.

8. References

1. Watson, J. D.; Crick, F. H., Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **1953,** *171* (4356), 737-8.
2. Crick, F., Central dogma of molecular biology. *Nature* **1970,** *227* (5258), 561-3.
3. Hoeijmakers, J. H., DNA damage, aging, and cancer. *N Engl J Med* **2009,** *361* (15), 1475-85.
4. Abbondandolo, A.; Dogliotti, E.; Lohman, P. H.; Berends, F., Molecular dosimetry of DNA damage caused by alkylation. I. Single-strand breaks induced by ethylating agents in cultured mammalian cells in relation to survival. *Mutat Res* **1982,** *92* (1-2), 361-77; Dogliotti, E.; Lakhanisky, T.; van der Schans, G. P.; Lohman, P. H., Molecular dosimetry of DNA damage caused by alkylation. II. The induction and repair of different classes of single-strand breaks in cultured mammalian cells treated with ethylating agents. *Mutat Res* **1984,** *132* (1-2), 41-9.
5. Imhof, P.; Zahran, M., The effect of a G:T mispair on the dynamics of DNA. *PLoS One* **2013,** *8* (1), e53305.
6. Lodish, H., *Molecular cell biology*. Macmillan: 2008.
7. Friedberg, E. C.; Walker, G. C.; Siede, W., *DNA repair and mutagenesis*. American Society for Microbiology (ASM): 1995.
8. Lucas-Lledó, J. I.; Lynch, M., Evolution of mutation rates: phylogenomic analysis of the photolyase/cryptochrome family. *Mol Biol Evol* **2009,** *26* (5), 1143-53.
9. Kaina, B.; Christmann, M.; Naumann, S.; Roos, W. P., MGMT: key node in the battle against genotoxicity, carcinogenicity and apoptosis induced by alkylating agents. *DNA Repair (Amst)* **2007,** *6* (8), 1079-99.
10. Cline, S. D.; Hanawalt, P. C., Who's on first in the cellular response to DNA damage? *Nat Rev Mol Cell Biol* **2003,** *4* (5), 361-373.
11. Matta, C. F.; Castillo, N.; Boyd, R. J., Extended Weak Bonding Interactions in DNA: π-Stacking (Base−Base), Base−Backbone, and Backbone−Backbone Interactions†. *The Journal of Physical Chemistry B* **2005,** *110* (1), 563-578.
12. Frederico, L. A.; Kunkel, T. A.; Shaw, B. R., A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy. *Biochemistry* **1990,** *29* (10), 2532-7; Lindahl, T.; Nyberg, B., Heat-induced deamination of cytosine residues in deoxyribonucleic acid. *Biochemistry* **1974,** *13* (16), 3405-10.
13. Rivetti, C.; Walker, C.; Bustamante, C., Polymer chain statistics and conformational analysis of DNA molecules with bends or sections of different flexibility. *J Mol Biol* **1998,** *280* (1), 41-59.
14. Chen, H.; Meisburger, S. P.; Pabit, S. A.; Sutton, J. L.; Webb, W. W.; Pollack, L., Ionic strength-dependent persistence lengths of single-stranded RNA and DNA. *Proc Natl Acad Sci U S A* **2012,** *109* (3), 799-804.
15. Murphy, M. C.; Rasnik, I.; Cheng, W.; Lohman, T. M.; Ha, T., Probing single-stranded DNA conformational flexibility using fluorescence spectroscopy. *Biophys J* **2004,** *86* (4), 2530-7.
16. Mitas, M.; Yu, A.; Dill, J.; Kamp, T. J.; Chambers, E. J.; Haworth, I. S., Hairpin properties of single-stranded DNA containing a GC-rich triplet repeat: (CTG)15. *Nucleic Acids Res* **1995,** *23* (6), 1050-9; Tan, E.; Wilson, T. J.; Nahas, M. K.; Clegg, R. M.; Lilley, D. M. J.; Ha, T., A four-way junction accelerates hairpin ribozyme folding via a discrete intermediate. *Proc Natl Acad Sci U S A* **2003,** *100* (16), 9308-9313.

17. Lemay, J.-F.; Penedo, J. C.; Tremblay, R.; Lilley, D. M. J.; Lafontaine, D. A., Folding of the adenine riboswitch. *Chem Biol* **2006,** *13* (8), 857-868.
18. Forterre, P., The origin of DNA genomes and DNA replication proteins. *Current Opinion in Microbiology* **2002,** *5* (5), 525-532.
19. Anand, R. P.; Shah, K. A.; Niu, H.; Sung, P.; Mirkin, S. M.; Freudenreich, C. H., Overcoming natural replication barriers: differential helicase requirements. *Nucleic Acids Research* **2012,** *40* (3), 1091-1105; Mirkin, S. M., DNA replication: Driving past four-stranded snags. *Nature* **2013,** *497* (7450), 449-450; Woodford, K. J.; Howell, R. M.; Usdin, K., A novel K(+)-dependent DNA synthesis arrest site in a commonly occurring sequence motif in eukaryotes. *Journal of Biological Chemistry* **1994,** *269* (43), 27029-27035.
20. Chan, K.; Sterling, J. F.; Roberts, S. A.; Bhagwat, A. S.; Resnick, M. A.; Gordenin, D. A., Base damage within single-strand DNA underlies in vivo hypermutability induced by a ubiquitous environmental agent. *PLoS Genet* **2012,** *8* (12), e1003149; Kuzminov, A., Single-strand interruptions in replicating chromosomes cause double-strand breaks. *Proc Natl Acad Sci U S A* **2001,** *98* (15), 8241-6.
21. Molineux, I. J.; Gefter, M. L., Properties of the Escherichia coli DNA-binding (unwinding) protein interaction with nucleolytic enzymes and DNA. *J Mol Biol* **1975,** *98* (4), 811-25.
22. Molineux, I. J.; Gefter, M. L., Properties of the Escherichia coli in DNA binding (unwinding) protein: interaction with DNA polymerase and DNA. *Proc Natl Acad Sci U S A* **1974,** *71* (10), 3858-62.
23. Hamon, L.; Pastré, D.; Dupaigne, P.; Breton, C. L.; Cam, E. L.; Piétrement, O., High-resolution AFM imaging of single-stranded DNA-binding (SSB) protein—DNA complexes. *Nucleic Acids Research* **2007,** *35* (8), e58.
24. Bell, J. C.; Plank, J. L.; Dombrowski, C. C.; Kowalczykowski, S. C., Direct imaging of RecA nucleation and growth on single molecules of SSB-coated ssDNA. *Nature* **2012,** *491* (7423), 274-8.
25. Jensen, D. E.; Kelly, R. C.; von Hippel, P. H., DNA "melting" proteins. II. Effects of bacteriophage T4 gene 32-protein binding on the conformation and stability of nucleic acid structures. *J Biol Chem* **1976,** *251* (22), 7215-28.
26. Sigal, N.; Delius, H.; Kornberg, T.; Gefter, M. L.; Alberts, B., A DNA-unwinding protein isolated from Escherichia coli: its interaction with DNA and with DNA polymerases. *Proc Natl Acad Sci U S A* **1972,** *69* (12), 3537-41.
27. Kerr, I. D.; Wadsworth, R. I. M.; Cubeddu, L.; Blankenfeldt, W.; Naismith, J. H.; White, M. F., Insights into ssDNA recognition by the OB fold from a structural and thermodynamic study of Sulfolobus SSB protein. *EMBO J* **2003,** *22* (11), 2561-2570.
28. Norais, C. A.; Chitteni-Pattu, S.; Wood, E. A.; Inman, R. B.; Cox, M. M., DdrB protein, an alternative Deinococcus radiodurans SSB induced by ionizing radiation. *J Biol Chem* **2009,** *284* (32), 21402-11; Shamoo, Y.; Friedman, A. M.; Parsons, M. R.; Konigsberg, W. H.; Steitz, T. A., Crystal structure of a replication fork single-stranded DNA binding protein (T4 gp32) complexed to DNA. *Nature* **1995,** *376* (6538), 362-366.
29. Bernstein, D. A.; Eggington, J. M.; Killoran, M. P.; Misic, A. M.; Cox, M. M.; Keck, J. L., Crystal structure of the Deinococcus radiodurans single-stranded DNA-binding protein suggests a mechanism for coping with DNA damage. *Proc Natl Acad Sci U S A* **2004,** *101* (23), 8575-8580.
30. Fan, J.; Pavletich, N. P., Structure and conformational change of a replication protein A heterotrimer bound to ssDNA. *Genes Dev* **2012,** *26* (20), 2337-47.
31. Antony, E.; Weiland, E. A.; Korolev, S.; Lohman, T. M., Plasmodium falciparum SSB tetramer wraps single-stranded DNA with similar topology but opposite polarity to E. coli SSB. *J Mol Biol* **2012,** *420* (4-5), 269-83.

32. Raghunathan, S.; Kozlov, A. G.; Lohman, T. M.; Waksman, G., Structure of the DNA binding domain of E. coli SSB bound to ssDNA. *Nat Struct Biol* **2000,** *7* (8), 648-52.
33. Boehmer, P. E.; Lehman, I. R., Herpes simplex virus type 1 ICP8: helix-destabilizing properties. *Journal of Virology* **1993,** *67* (2), 711-715.
34. Glassberg, J.; Meyer, R. R.; Kornberg, A., Mutant single-strand binding protein of Escherichia coli: genetic and physiological characterization. *Journal of Bacteriology* **1979,** *140* (1), 14-19.
35. Bochkarev, A.; Bochkareva, E., From RPA to BRCA2: lessons from single-stranded DNA binding by the OB-fold. *Current opinion in structural biology* **2004,** *14* (1), 36-42.
36. Suck, D., Common fold, common function, common origin? *Nat Struct Biol* **1997,** *4* (3), 161-5.
37. Arcus, V., OB-fold domains: a snapshot of the evolution of sequence, structure and function. *Current Opinion in Structural Biology* **2002,** *12* (6), 794-801.
38. Murzin, A. G., OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J* **1993,** *12* (3), 861-7.
39. Lo Conte, L.; Ailey, B.; Hubbard, T. J. P.; Brenner, S. E.; Murzin, A. G.; Chothia, C., SCOP: a Structural Classification of Proteins database. *Nucleic Acids Research* **2000,** *28* (1), 257-259.
40. Theobald, D. L.; Mitton-Fry, R. M.; Wuttke, D. S., Nucleic acid recognition by OB-fold proteins. *Annu Rev Biophys Biomol Struct* **2003,** *32*, 115-33.
41. Draper, D. E., Themes in RNA-protein recognition. *Journal of Molecular Biology* **1999,** *293* (2), 255-270.
42. Bochkarev, A.; Pfuetzner, R. A.; Edwards, A. M.; Frappier, L., Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA. *Nature* **1997,** *385* (6612), 176-81.
43. Fanning, E.; Klimovich, V.; Nager, A. R., A dynamic model for replication protein A (RPA) function in DNA processing pathways. *Nucleic Acids Res* **2006,** *34* (15), 4126-37.
44. Zou, Y.; Liu, Y.; Wu, X.; Shell, S. M., Functions of human replication protein A (RPA): from DNA replication to DNA damage and stress responses. *J Cell Physiol* **2006,** *208* (2), 267-73.
45. Iftode, C.; Borowiec, J. A., 5' --> 3' molecular polarity of human replication protein A (hRPA) binding to pseudo-origin DNA substrates. *Biochemistry* **2000,** *39* (39), 11970-81.
46. de Laat, W. L.; Appeldoorn, E.; Sugasawa, K.; Weterings, E.; Jaspers, N. G.; Hoeijmakers, J. H., DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. *Genes Dev* **1998,** *12* (16), 2598-609.
47. Arunkumar, A. I.; Stauffer, M. E.; Bochkareva, E.; Bochkarev, A.; Chazin, W. J., Independent and coordinated functions of replication protein A tandem high affinity single-stranded DNA binding domains. *J Biol Chem* **2003,** *278* (42), 41077-82.
48. Zhou, H. X., The affinity-enhancing roles of flexible linkers in two-domain DNA-binding proteins. *Biochemistry* **2001,** *40* (50), 15069-73.
49. Bochkarev, A.; Bochkareva, E.; Frappier, L.; Edwards, A. M., The crystal structure of the complex of replication protein A subunits RPA32 and RPA14 reveals a mechanism for single-stranded DNA binding. *EMBO J.* **1999,** *18* (16), 4498-504.
50. Gomes, X. V.; Wold, M. S., Functional domains of the 70-kilodalton subunit of human replication protein A. *Biochemistry* **1996,** *35* (32), 10558-68; Daughdrill, G. W.; Buchko, G. W.; Botuyan, M. V.; Arrowsmith, C.; Wold, M. S.; Kennedy, M. A.; Lowry, D. F., Chemical shift changes provide evidence for overlapping single-stranded DNA- and

XPA-binding sites on the 70 kDa subunit of human replication protein A. *Nucleic Acids Res* **2003,** *31* (14), 4176-83.
51. Kim, C.; Paulus, B. F.; Wold, M. S., Interactions of human replication protein A with oligonucleotides. *Biochemistry* **1994,** *33* (47), 14197-14206; Liu, Y.; Yang, Z.; Utzat, C. D.; Geacintov, N. E.; Basu, A. K.; Zou, Y., Interactions of human replication protein A with single-stranded DNA adducts. *Biochem J* **2005,** *385* (Pt 2), 519-26.
52. Kim, C.; Wold, M. S., Recombinant Human Replication Protein A Binds to Polynucleotides with Low Cooperativity. *Biochemistry* **1995,** *34* (6), 2058-2064.
53. Treuner, K.; Ramsperger, U.; Knippers, R., Replication protein A induces the unwinding of long double-stranded DNA regions. *J Mol Biol* **1996,** *259* (1), 104-12.
54. .
55. New, J. H.; Sugiyama, T.; Zaitseva, E.; Kowalczykowski, S. C., Rad52 protein stimulates DNA strand exchange by Rad51 and replication protein A. *Nature* **1998,** *391* (6665), 407-10.
56. Xu, X.; Vaithiyalingam, S.; Glick, G. G.; Mordes, D. A.; Chazin, W. J.; Cortez, D., The basic cleft of RPA70N binds multiple checkpoint proteins, including RAD9, to regulate ATR signaling. *Mol Cell Biol* **2008,** *28* (24), 7345-53; Olson, E.; Nievera, C. J.; Liu, E.; Lee, A. Y.; Chen, L.; Wu, X., The Mre11 complex mediates the S-phase checkpoint through an interaction with replication protein A. *Mol Cell Biol* **2007,** *27* (17), 6053-67; Ball, H. L.; Ehrhardt, M. R.; Mordes, D. A.; Glick, G. G.; Chazin, W. J.; Cortez, D., Function of a conserved checkpoint recruitment domain in ATRIP proteins. *Mol Cell Biol* **2007,** *27* (9), 3367-77; Namiki, Y.; Zou, L., ATRIP associates with replication protein A-coated ssDNA through multiple interactions. *Proc Natl Acad Sci U S A* **2006,** *103* (3), 580-5; Bochkareva, E.; Kaustov, L.; Ayed, A.; Yi, G. S.; Lu, Y.; Pineda-Lucena, A.; Liao, J. C.; Okorokov, A. L.; Milner, J.; Arrowsmith, C. H.; Bochkarev, A., Single-stranded DNA mimicry in the p53 transactivation domain interaction with replication protein A. *Proc Natl Acad Sci U S A* **2005,** *102* (43), 15412-7; Lin, Y. L.; Chen, C.; Keshav, K. F.; Winchester, E.; Dutta, A., Dissection of functional domains of the human DNA replication protein complex replication protein A. *J Biol Chem* **1996,** *271* (29), 17190-8.
57. Wold, M. S.; Kelly, T., Purification and characterization of replication protein A, a cellular protein required for in vitro replication of simian virus 40 DNA. *Proc Natl Acad Sci U S A* **1988,** *85* (8), 2523-7.
58. Braun, K. A.; Lao, Y.; He, Z.; Ingles, C. J.; Wold, M. S., Role of protein-protein interactions in the function of replication protein A (RPA): RPA modulates the activity of DNA polymerase alpha by multiple mechanisms. *Biochemistry* **1997,** *36* (28), 8443-54.
59. Waga, S.; Stillman, B., Anatomy of a DNA replication fork revealed by reconstitution of SV40 DNA replication in vitro. *Nature* **1994,** *369* (6477), 207-12; Waga, S.; Bauer, G.; Stillman, B., Reconstitution of complete SV40 DNA replication with purified replication factors. *J Biol Chem* **1994,** *269* (14), 10923-34; Bae, S. H.; Bae, K. H.; Kim, J. A.; Seo, Y. S., RPA governs endonuclease switching during processing of Okazaki fragments in eukaryotes. *Nature* **2001,** *412* (6845), 456-61.
60. Kao, H. I.; Campbell, J. L.; Bambara, R. A., Dna2p helicase/nuclease is a tracking protein, like FEN1, for flap cleavage during Okazaki fragment maturation. *J Biol Chem* **2004,** *279* (49), 50840-9.
61. Gibb, B.; Ye, L. F.; Gergoudis, S. C.; Kwon, Y.; Niu, H.; Sung, P.; Greene, E. C., Concentration-dependent exchange of replication protein A on single-stranded DNA revealed by single-molecule imaging. *PLoS One* **2014,** *9* (2), e87922.
62. Evans, E.; Moggs, J. G.; Hwang, J. R.; Egly, J. M.; Wood, R. D., Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. *EMBO J* **1997,** *16* (21), 6559-73.

63. Patrick, S. M.; Turchi, J. J., Replication protein A (RPA) binding to duplex cisplatin-damaged DNA is mediated through the generation of single-stranded DNA. *J Biol Chem* **1999,** *274* (21), 14972-8.
64. Matsunaga, T.; Park, C. H.; Bessho, T.; Mu, D.; Sancar, A., Replication protein A confers structure-specific endonuclease activities to the XPF-ERCC1 and XPG subunits of human DNA repair excision nuclease. *J Biol Chem* **1996,** *271* (19), 11047-50; He, Z.; Henricksen, L. A.; Wold, M. S.; Ingles, C. J., RPA involvement in the damage-recognition and incision steps of nucleotide excision repair. *Nature* **1995,** *374* (6522), 566-9.
65. Shivji, M. K.; Podust, V. N.; Hübscher, U.; Wood, R. D., Nucleotide excision repair DNA synthesis by DNA polymerase epsilon in the presence of PCNA, RFC, and RPA. *Biochemistry* **1995,** *34* (15), 5011-7.
66. Oakley, G. G.; Patrick, S. M., Replication protein A: directing traffic at the intersection of replication and repair. *Front Biosci (Landmark Ed)* **2010,** *15*, 883-900.
67. Umezu, K.; Sugawara, N.; Chen, C.; Haber, J. E.; Kolodner, R. D., Genetic analysis of yeast RPA1 reveals its multiple functions in DNA metabolism. *Genetics* **1998,** *148* (3), 989-1005.
68. DeMott, M. S.; Zigman, S.; Bambara, R. A., Replication protein A stimulates long patch DNA base excision repair. *J Biol Chem* **1998,** *273* (42), 27492-8; Ranalli, T. A.; DeMott, M. S.; Bambara, R. A., Mechanism underlying replication protein a stimulation of DNA ligase I. *J Biol Chem* **2002,** *277* (3), 1719-27.
69. Mer, G.; Bochkarev, A.; Gupta, R.; Bochkareva, E.; Frappier, L.; Ingles, C. J.; Edwards, A. M.; Chazin, W. J., Structural basis for the recognition of DNA repair proteins UNG2, XPA, and RAD52 by replication factor RPA. *Cell* **2000,** *103* (3), 449-56.
70. Ramilo, C.; Gu, L.; Guo, S.; Zhang, X.; Patrick, S. M.; Turchi, J. J.; Li, G. M., Partial reconstitution of human DNA mismatch repair in vitro: characterization of the role of human replication protein A. *Mol Cell Biol* **2002,** *22* (7), 2037-46; Genschel, J.; Modrich, P., Functions of MutLalpha, replication protein A (RPA), and HMGB1 in 5'-directed mismatch repair. *J Biol Chem* **2009,** *284* (32), 21536-44; Genschel, J.; Modrich, P., Mechanism of 5'-directed excision in human mismatch repair. *Mol Cell* **2003,** *12* (5), 1077-86.
71. Binz, S. K.; Lao, Y.; Lowry, D. F.; Wold, M. S., The phosphorylation domain of the 32-kDa subunit of replication protein A (RPA) modulates RPA-DNA interactions. Evidence for an intersubunit interaction. *J Biol Chem* **2003,** *278* (37), 35584-91.
72. McKinnon, P. J., DNA repair deficiency and neurological disease. *Nat Rev Neurosci* **2009,** *10* (2), 100-12.
73. Park, M. S.; Ludwig, D. L.; Stigger, E.; Lee, S. H., Physical interaction between human RAD52 and RPA is required for homologous recombination in mammalian cells. *J Biol Chem* **1996,** *271* (31), 18996-9000.
74. Jackson, D.; Dhar, K.; Wahl, J. K.; Wold, M. S.; Borgstahl, G. E., Analysis of the human replication protein A:Rad52 complex: evidence for crosstalk between RPA32, RPA70, Rad52 and DNA. *J Mol Biol* **2002,** *321* (1), 133-48.
75. Deng, X.; Prakash, A.; Dhar, K.; Baia, G. S.; Kolar, C.; Oakley, G. G.; Borgstahl, G. E., Human replication protein A-Rad52-single-stranded DNA complex: stoichiometry and evidence for strand transfer regulation by phosphorylation. *Biochemistry* **2009,** *48* (28), 6633-43.
76. ; Shereda, R. D.; Kozlov, A. G.; Lohman, T. M.; Cox, M. M.; Keck, J. L., SSB as an organizer/mobilizer of genome maintenance complexes. *Crit Rev Biochem Mol Biol* **2008,** *43* (5), 289-318; Meyer, R. R.; Laine, P. S., The single-stranded DNA-binding protein of Escherichia coli. *Microbiol Rev* **1990,** *54* (4), 342-80.
77. Chase, J. W.; L'Italien, J. J.; Murphy, J. B.; Spicer, E. K.; Williams, K. R., Characterization of the Escherichia coli SSB-113 mutant single-stranded DNA-binding protein. Cloning of the gene, DNA and protein sequence analysis, high pressure liquid

chromatography peptide mapping, and DNA-binding studies. *J Biol Chem* **1984,** *259* (2), 805-14.
78. Cheng, Z.; Caillet, A.; Ren, B.; Ding, H., Stimulation of Escherichia coli DNA damage inducible DNA helicase DinG by the single-stranded DNA binding protein SSB. *FEBS Lett* **2012,** *586* (21), 3825-30.
79. Antony, E.; Weiland, E.; Yuan, Q.; Manhart, C. M.; Nguyen, B.; Kozlov, A. G.; McHenry, C. S.; Lohman, T. M., Multiple C-Terminal Tails within a Single E. coli SSB Homotetramer Coordinate DNA Replication and Repair. *J Mol Biol* **2013,** *425* (23), 4802-19.
80. Shishmarev, D.; Wang, Y.; Mason, C. E.; Su, X. C.; Oakley, A. J.; Graham, B.; Huber, T.; Dixon, N. E.; Otting, G., Intramolecular binding mode of the C-terminus of Escherichia coli single-stranded DNA binding protein determined by nuclear magnetic resonance spectroscopy. *Nucleic Acids Res* **2014,** *42* (4), 2750-7.
81. Savvides, S. N.; Raghunathan, S.; Fütterer, K.; Kozlov, A. G.; Lohman, T. M.; Waksman, G., The C-terminal domain of full-length E. coli SSB is disordered even when bound to DNA. *Protein Sci* **2004,** *13* (7), 1942-7.
82. Lohman, T. M.; Overman, L. B., Two binding modes in Escherichia coli single strand binding protein-single stranded DNA complexes. Modulation by NaCl concentration. *J Biol Chem* **1985,** *260* (6), 3594-603.
83. Griffith, J. D.; Harris, L. D.; Register, J., Visualization of SSB-ssDNA complexes active in the assembly of stable RecA-DNA filaments. *Cold Spring Harb Symp Quant Biol* **1984,** *49*, 553-9; Lohman, T. M.; Overman, L. B.; Datta, S., Salt-dependent changes in the DNA binding co-operativity of Escherichia coli single strand binding protein. *J Mol Biol* **1986,** *187* (4), 603-15; Wei, T. F.; Bujalowski, W.; Lohman, T. M., Cooperative binding of polyamines induces the Escherichia coli single-strand binding protein-DNA binding mode transitions. *Biochemistry* **1992,** *31* (26), 6166-74.
84. Ferrari, M. E.; Bujalowski, W.; Lohman, T. M., Co-operative binding of Escherichia coli SSB tetramers to single-stranded DNA in the (SSB)35 binding mode. *J Mol Biol* **1994,** *236* (1), 106-23.
85. Kozlov, A. G.; Cox, M. M.; Lohman, T. M., Regulation of single-stranded DNA binding by the C termini of Escherichia coli single-stranded DNA-binding (SSB) protein. *J Biol Chem* **2010,** *285* (22), 17246-52.
86. Richard, D. J.; Bell, S. D.; White, M. F., Physical and functional interaction of the archaeal single-stranded DNA-binding protein SSB with RNA polymerase. *Nucleic Acids Res* **2004,** *32* (3), 1065-74.
87. Bouthier de la Tour, C.; Boisnard, S.; Norais, C.; Toueille, M.; Bentchikou, E.; Vannier, F.; Cox, M. M.; Sommer, S.; Servant, P., The deinococcal DdrB protein is involved in an early step of DNA double strand break repair and in plasmid transformation through its single-strand annealing activity. *DNA Repair (Amst)* **2011,** *10* (12), 1223-31; Eggington, J. M.; Haruta, N.; Wood, E. A.; Cox, M. M., The single-stranded DNA-binding protein of Deinococcus radiodurans. *BMC Microbiol* **2004,** *4*, 2-2.
88. Cubeddu, L.; White, M. F., DNA damage detection by an archaeal single-stranded DNA-binding protein. *J Mol Biol* **2005,** *353* (3), 507-16.
89. Kelman, Z.; Yuzhakov, A.; Andjelkovic, J.; O'Donnell, M., Devoted to the lagging strand-the subunit of DNA polymerase III holoenzyme contacts SSB to promote processive elongation and sliding clamp assembly. *EMBO J* **1998,** *17* (8), 2436-49.
90. Yuzhakov, A.; Kelman, Z.; O'Donnell, M., Trading places on DNA--a three-point switch underlies primer handoff from primase to the replicative DNA polymerase. *Cell* **1999,** *96* (1), 153-63.
91. Shereda, R. D.; Bernstein, D. A.; Keck, J. L., A central role for SSB in Escherichia coli RecQ DNA helicase function. *J Biol Chem* **2007,** *282* (26), 19247-58.

92. Han, E. S.; Cooper, D. L.; Persky, N. S.; Sutera, V. A.; Whitaker, R. D.; Montello, M. L.; Lovett, S. T., RecJ exonuclease: substrates, products and interaction with SSB. *Nucleic Acids Res* **2006,** *34* (4), 1084-91.
93. Slocum, S. L.; Buss, J. A.; Kimura, Y.; Bianco, P. R., Characterization of the ATPase activity of the Escherichia coli RecG protein reveals that the preferred cofactor is negatively supercoiled DNA. *J Mol Biol* **2007,** *367* (3), 647-64.
94. Hobbs, M. D.; Sakai, A.; Cox, M. M., SSB protein limits RecOR binding onto single-stranded DNA. *J Biol Chem* **2007,** *282* (15), 11058-67; Umezu, K.; Kolodner, R. D., Protein interactions in genetic recombination in Escherichia coli. Interactions involving RecO and RecR overcome the inhibition of RecA by single-stranded DNA-binding protein. *J Biol Chem* **1994,** *269* (47), 30005-13.
95. Lu, D.; Keck, J. L., Structural basis of Escherichia coli single-stranded DNA-binding protein stimulation of exonuclease I. *Proc Natl Acad Sci U S A* **2008,** *105* (27), 9169-74.
96. Kumar, N. V.; Varshney, U., Contrasting effects of single stranded DNA binding protein on the activity of uracil DNA glycosylase from Escherichia coli towards different DNA substrates. *Nucleic Acids Res* **1997,** *25* (12), 2336-43.
97. Weiner, J. H.; Bertsch, L. L.; Kornberg, A., The deoxyribonucleic acid unwinding protein of Escherichia coli. Properties and functions in replication. *J Biol Chem* **1975,** *250* (6), 1972-80.
98. Woese, C. R.; Fox, G. E., Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci U S A* **1977,** *74* (11), 5088-90.
99. White, M. F., Archaeal DNA repair: paradigms and puzzles. *Biochem Soc Trans* **2003,** *31* (Pt 3), 690-693; Barry, E. R.; Bell, S. D., DNA Replication in the Archaea. *Microbiology and Molecular Biology Reviews* **2006,** *70* (4), 876-887; Bell, S. D.; Jackson, S. P., Transcription and translation in Archaea: a mosaic of eukaryal and bacterial features. *Trends Microbiol* **1998,** *6* (6), 222-8.
100. Godde, J. S., Breaking through a phylogenetic impasse: a pair of associated archaea might have played host in the endosymbiotic origin of eukaryotes. *Cell Biosci* **2012,** *2* (1), 29.
101. Woese, C. R.; Kandler, O.; Wheelis, M. L., Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci U S A* **1990,** *87* (12), 4576-9.
102. Robbins, J. B.; Murphy, M. C.; White, B. A.; Mackie, R. I.; Ha, T.; Cann, I. K., Functional analysis of multiple single-stranded DNA-binding proteins from Methanosarcina acetivorans and their effects on DNA synthesis by DNA polymerase BI. *J Biol Chem* **2004,** *279* (8), 6315-26.
103. Robbins, J. B.; McKinney, M. C.; Guzman, C. E.; Sriratana, B.; Fitz-Gibbon, S.; Ha, T.; Cann, I. K., The euryarchaeota, nature's medium for engineering of single-stranded DNA-binding proteins. *J Biol Chem* **2005,** *280* (15), 15325-39.
104. Kelly, T. J.; Simancek, P.; Brush, G. S., Identification and characterization of a single-stranded DNA-binding protein from the archaeon Methanococcus jannaschii. *Proc Natl Acad Sci U S A* **1998,** *95* (25), 14634-14639.
105. Skowyra, A.; MacNeill, S. A., Identification of essential and non-essential single-stranded DNA-binding proteins in a model archaeal organism. *Nucleic Acids Res* **2012,** *40* (3), 1077-90.
106. Pugh, R. A.; Lin, Y.; Eller, C.; Leesley, H.; Cann, I. K.; Spies, M., Ferroplasma acidarmanus RPA2 facilitates efficient unwinding of forked DNA substrates by monomers of FacXPD helicase. *J Mol Biol* **2008,** *383* (5), 982-98.
107. Lin, Y.; Robbins, J. B.; Nyannor, E. K. D.; Chen, Y.-H.; Cann, I. K. O., A CCCH Zinc Finger Conserved in a Replication Protein A Homolog Found in Diverse Euryarchaeotes. *Journal of Bacteriology* **2005,** *187* (23), 7881-7889.

108. Komori, K.; Ishino, Y., Replication protein A in Pyrococcus furiosus is involved in homologous DNA recombination. *J Biol Chem* **2001,** *276* (28), 25654-60.
109. Wadsworth, R. I. M.; White, M. F., Identification and properties of the crenarchaeal single-stranded DNA binding protein from Sulfolobus solfataricus. *Nucleic Acids Research* **2001,** *29* (4), 914-920.
110. Raghunathan, S.; Ricard, C. S.; Lohman, T. M.; Waksman, G., Crystal structure of the homo-tetrameric DNA binding domain of Escherichia coli single-stranded DNA-binding protein determined by multiwavelength x-ray diffraction on the selenomethionyl protein at 2.9-A resolution. *Proc Natl Acad Sci U S A* **1997,** *94* (13), 6652-7.
111. Haseltine, C. A.; Kowalczykowski, S. C., A distinctive single-strand DNA-binding protein from the Archaeon Sulfolobus solfataricus. *Mol Microbiol* **2002,** *43* (6), 1505-15.
112. Shi, H.; Zhang, Y.; Zhang, G.; Guo, J.; Zhang, X.; Song, H.; Lv, J.; Gao, J.; Wang, Y.; Chen, L., Systematic Functional Comparative Analysis of Four Single-Stranded DNA-Binding Proteins and Their Affection on Viral RNA Metabolism. *PLoS One* **2013,** *8* (1), e55076.
113. Gamsjaeger, R.; Kariawasam, R.; Touma, C.; Kwan, A. H.; White, M. F.; Cubeddu, L., Backbone and side-chain (1)H, (13)C and (15)N resonance assignments of the OB domain of the single stranded DNA binding protein from Sulfolobus solfataricus and chemical shift mapping of the DNA-binding interface. *Biomol NMR Assign* **2013**.
114. Gamsjaeger, R.; Kariawasam, R.; Gimenez, A. X.; Touma, C.; McIlwain, E.; Bernardo, R. E.; Shepherd, N. E.; Ataide, S. F.; Dong, Q.; Richard, D. J.; White, M. F.; Cubeddu, L., The structural basis of DNA binding by the single stranded DNA binding protein from Sulfolobus solfataricus. *Biochem J* **2014**.
115. Götz, D.; Paytubi, S.; Munro, S.; Lundgren, M.; Bernander, R.; White, M. F., Responses of hyperthermophilic crenarchaea to UV irradiation. *Genome Biol* **2007,** *8* (10), R220.
116. Deniz, A. A.; Mukhopadhyay, S.; Lemke, E. A., Single-molecule biophysics: at the interface of biology, physics and chemistry. *J R Soc Interface* **2008,** *5* (18), 15-45.
117. Feynman, R. P., There's plenty of room at the bottom. *Engineering and Science* **1960,** *23* (5), 22-36.
118. Giessibl, F. J., Advances in atomic force microscopy. **2003,** *75* (July).
119. Tersoff, J.; Hamann, D. R., Theory and application for the scanning tunneling microscope. *Physical review letters* **1983,** *50* (25), 1998-2001.
120. Ashkin, A., Acceleration and trapping of particles by radiation pressure. *Physical Review Letters* **1970,** *24* (4), 156-159.
121. Gosse, C.; Croquette, V., Magnetic tweezers: micromanipulation and force measurement at the molecular level. *Biophysical journal* **2002,** *82* (6), 3314-29.
122. Blackie, E. J.; Le Ru, E. C.; Etchegoin, P. G., Single-molecule surface-enhanced Raman spectroscopy of nonresonant molecules. *Journal of the American Chemical Society* **2009,** *131* (40), 14466-72.
123. Hess, S. T.; Girirajan, T. P.; Mason, M. D., Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophys J* **2006,** *91* (11), 4258-72.
124. Rust, M. J.; Bates, M.; Zhuang, X., Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat Methods* **2006,** *3* (10), 793-5.
125. Hell, S. W.; Wichmann, J., Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt Lett* **1994,** *19* (11), 780-2.

126. Jungmann, R.; Steinhauer, C.; Scheible, M.; Kuzyk, A.; Tinnefeld, P.; Simmel, F. C., Single-molecule kinetics and super-resolution microscopy by fluorescence imaging of transient binding on DNA origami. *Nano Lett* **2010,** *10* (11), 4756-61.
127. Rittweger, E.; Han, K. Y.; Irvine, S. E.; Eggeling, C.; Hell, S. W., STED microscopy reveals crystal colour centres with nanometric resolution. *Nat Photon* **2009,** *3* (3), 144-147.
128. Hosny, N. A.; Song, M.; Connelly, J. T.; Ameer-Beg, S.; Knight, M. M.; Wheeler, A. P., Super-resolution imaging strategies for cell biologists using a spinning disk microscope. *PLoS One* **2013,** *8* (10), e74604.
129. Thompson, M. A.; Biteen, J. S.; Lord, S. J.; Conley, N. R.; Moerner, W. E., Molecules and methods for super-resolution imaging. *Methods Enzymol* **2010,** *475*, 27-59.
130. Gruber, H. J.; Hahn, C. D.; Kada, G.; Riener, C. K.; Harms, G. S.; Ahrer, W.; Dax, T. G.; Knaus, H. G., Anomalous fluorescence enhancement of Cy3 and cy3.5 versus anomalous fluorescence loss of Cy5 and Cy7 upon covalent linking to IgG and noncovalent binding to avidin. *Bioconjug Chem* **2000,** *11* (5), 696-704.
131. Atkins, P.; De Paula, J., *Physical Chemistry*. Oxford University Press: 2006.
132. Lakowicz, J. R., *Principles of Fluorescence Spectroscopy*. 3 ed.; Springer: New York, 2006.
133. Förster, T., Zwischenmolekulare energiewanderung und fluoreszenz. *Annalen der physik* **1948,** *437* (1-2), 55-75.
134. Mielczarek, E. V.; Greenbaum, E. S.; Knox, R. S., *Biological physics*. Springer: 1993.
135. Clegg, R. M., The history of FRET. In *Reviews in Fluorescence 2006*, Springer: 2006; pp 1-45.
136. Kallman, H.; London, F., Über quantenmechanische Energieübertragung zwischen atomaren Systemen. *Z. Phys. Chem* **1929,** *2* (207-243), 275.
137. Perrin, F., Théorie quantique des transferts d'activation entre molécules de même espèce. Cas des solutions fluorescentes. *Ann. Phys.(Paris)* **1932,** *12*, 283-314; Perrin, F. In *Interaction entre atomes normal et active. Transferts d'activation. Formation d'une molecule activee*, Annales de l'institut Henri Poincaré, Presses universitaires de France: 1933; pp 279-318.
138. Cario, G.; Franck, J., Über zerlegung von wasserstoffmolekülen durch angeregte quecksilberatome. *Zeitschrift für Physik A Hadrons and Nuclei* **1922,** *11* (1), 161-166.
139. Beutler, H.; Josephy, B., Resonanz bei Stößen in der Fluoreszenz und Chemilumineszenz. *Zeitschrift für Physik* **1929,** *53* (11-12), 747-765.
140. Arnold, W.; Oppenheimer, J. R., Internal conversion in the photosynthetic mechanism of blue-green algae. *J Gen Physiol* **1950,** *33* (4), 423-35; Oppenheimer, J. R., Internal conversion in photosynthesis. *Physical Review* **1941,** *60* (2), 158-165.
141. van der Meer, B. W., Förster Theory. **2013**, 23--62.
142. Valeur, B.; Berberan-Santos, M. N., *Molecular fluorescence: principles and applications*. John Wiley & Sons: 2012.
143. Stryer, L.; Haugland, R. P., Energy transfer: a spectroscopic ruler. *Proceedings of the National Academy of Sciences* **1967,** *58* (2), 719-726.
144. Latt, S. A.; Cheung, H. T.; Blout, E. R., Energy transfer. A system with relatively fixed donor-acceptor separation. *J Am Chem Soc* **1965,** *87*, 995-1003.
145. Ha, T.;Enderle, T.;Ogletree, D. F.;Chemla, D. S.;Selvin, P. R.;Weiss, S ., Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proceedings of the National Academy of Sciences* **1996,** *93* (13), 6264-6268.
146. Park, J.; Myong, S.; Niedziela-Majka, A.; Lee, K. S.; Yu, J.; Lohman, T. M.; Ha, T., PcrA Helicase Dismantles RecA Filaments by Reeling in DNA in Uniform Steps. *Cell* **2010,** *142* (4), 544-555.

147. Talaga, D. S.; Lau, W. L.; Roder, H.; Tang, J.; Jia, Y.; DeGrado, W. F.; Hochstrasser, R. M., Dynamics and folding of single two-stranded coiled-coil peptides studied by fluorescent energy transfer confocal microscopy. *Proceedings of the National Academy of Sciences of the United States of America* **2000,** *97* (24), 13021-6.
148. Paige, M. F.; Bjerneld, E. J.; Moerner, W. E., A comparison of through-the-objective total internal reflection microscopy and epifluorescence microscopy for single-molecule fluorescence Imaging. *Single Molecules* **2001,** *2* (3), 191-201; Rüttinger, S.; Lamarre, B.; Knight, A. E., Single molecule genotyping by TIRF microscopy. *Journal of fluorescence* **2008,** *18* (5), 1021-6; Beaumont, V., Visualizing membrane trafficking using total internal reflection fluorescence microscopy. *Biochemical Society Transactions* **2003,** *31*, 819-823; Funatsu, T.; Harada, Y.; Tokunaga, M.; Saito, K., Imaging of single fluorescent molecules and individual ATP turnovers by single myosin molecules in aqueous solution. *Nature* **1995,** *374*, 555-559.
149. Walter, N. G., *Single Molecule Tools, Part A: Fluorescence Based Approaches: Fluorescence Based Approaches*. Academic Press: 2010; Vol. 472.
150. Joo, C.; Ha, T., Preparing sample chambers for single-molecule FRET. *Cold Spring Harb Protoc* **2012,** *2012* (10), 1104-8; Mattheyses, A. L.; Simon, S. M.; Rappoport, J. Z., Imaging with total internal reflection fluorescence microscopy for the cell biologist. *J Cell Sci* **2010,** *123* (Pt 21), 3621-8.
151. Heyes, C. D.; Groll, J.; Moller, M.; Nienhaus, G. U., Synthesis, patterning and applications of star-shaped poly(ethylene glycol) biofunctionalized surfaces. *Molecular BioSystems* **2007,** *3* (6), 419-430.
152. Nuzzo, R. G.; Allara, D. L., Adsorption of bifunctional organic disulfides on gold surfaces. *Journal of the American Chemical Society* **1983,** *105* (13), 4481-4483; Liu, G.; Chockalingham, M.; Khor, S. M.; Gui, A. L.; Gooding, J. J., A Comparative Study of the Modification of Gold and Glassy Carbon Surfaces with Mixed Layers of In Situ Generated Aryl Diazonium Compounds. *Electroanalysis* **2010,** *22* (9), 918-926; Blümmel, J.; Perschmann, N.; Aydin, D.; Drinjakovic, J.; Surrey, T.; Lopez-Garcia, M.; Kessler, H.; Spatz, J. P., Protein repellent properties of covalently attached PEG coatings on nanostructured SiO(2)-based interfaces. *Biomaterials* **2007,** *28* (32), 4739-47.
153. Roy, R.; Hohng, S.; Ha, T., A practical guide to single-molecule FRET. *Nat Meth* **2008,** *5* (6), 507-516.
154. Koopmans, W. J. A.; Schmidt, T.; van Noort, J., Nucleosome Immobilization Strategies for Single-Pair FRET Microscopy. *ChemPhysChem* **2008,** *9* (14), 2002-2009.
155. Roy, R.; Kozlov, A. G.; Lohman, T. M.; Ha, T., Dynamic structural rearrangements between DNA binding modes of E. coli SSB protein. *J Mol Biol* **2007,** *369* (5), 1244-1257.
156. Iqbal, A.; Arslan, S.; Okumus, B.; Wilson, T. J.; Giraud, G.; Norman, D. G.; Ha, T.; Lilley, D. M. J., Orientation dependence in fluorescent energy transfer between Cy3 and Cy5 terminally attached to double-stranded nucleic acids. *Proceedings of the National Academy of Sciences of the United States of America* **2008,** *105* (32), 11176-81; Craggs, T. D.; Hutton, R. D.; Brenlla, A.; White, M. F.; Penedo, J. C., Single-molecule characterization of Fen1 and Fen1/PCNA complexes acting on flap substrates. *Nucleic Acids Res* **2014,** *42* (3), 1857-72; Altman, R. B.; Terry, D. S.; Zhou, Z.; Zheng, Q.; Geggier, P.; Kolster, R. A.; Zhao, Y.; Javitch, J. A.; Warren, J. D.; Blanchard, S. C., Cyanine fluorophore derivatives with enhanced photostability. *Nature Methods* **2011,** *9* (1), 68-71.
157. Livnah, O.; Bayer, E. A.; Wilchek, M.; Sussman, J. L., Three-dimensional structures of avidin and the avidin-biotin complex. *Proc Natl Acad Sci U S A* **1993,** *90* (11), 5076-80.
158. Ameringer, T.; Hinz, M.; Mourran, C.; Seliger, H.; Groll, J.; Moeller, M., Ultrathin Functional Star PEG Coatings for DNA Microarrays. *Biomacromolecules* **2005,** *6* (4), 1819-1823; Groll, J.; Ameringer, T.; Spatz, J. P.; Moeller, M., Ultrathin Coatings from

Isocyanate-Terminated Star PEG Prepolymers: Layer Formation and Characterization. *Langmuir* **2005,** *21* (5), 1991-1999; Götz, H.; Beginn, U.; Bartelink, C. F.; Grünbauer, H. J. M.; Möller, M., Preparation of isophorone diisocyanate terminated star polyethers. *Macromolecular Materials and Engineering* **2002,** *287* (4), 223-230.
159. Cisse, I.; Okumus, B.; Joo, C.; Ha, T., Fueling protein DNA interactions inside porous nanocontainers. *Proceedings of the National Academy of Sciences of the United States of America* **2007,** *104* (31), 12646-12650.
160. Okumus, B.; Arslan, S.; Fengler, S. M.; Myong, S.; Ha, T., Single Molecule Nanocontainers Made Porous Using a Bacterial Toxin. *Journal of the American Chemical Society* **2009,** *131* (41), 14844-14849.
161. Gibb, B.; Silverstein, T. D.; Finkelstein, I. J.; Greene, E. C., Single-stranded DNA curtains for real-time single-molecule visualization of protein-nucleic acid interactions. *Anal Chem* **2012,** *84* (18), 7607-12; Greene, E. C.; Wind, S.; Fazio, T.; Gorman, J.; Visnapuu, M. L., DNA curtains for high-throughput single-molecule optical imaging. *Methods Enzymol* **2010,** *472*, 293-315; Fazio, T.; Visnapuu, M. L.; Wind, S.; Greene, E. C., DNA curtains and nanoscale curtain rods: high-throughput tools for single molecule imaging. *Langmuir* **2008,** *24* (18), 10524-31.
162. Finkelstein, I. J.; Visnapuu, M. L.; Greene, E. C., Single-molecule imaging reveals mechanisms of protein disruption by a DNA translocase. *Nature* **2010,** *468* (7326), 983-7.
163. Gibb, B.; Ye, L. F.; Kwon, Y.; Niu, H.; Sung, P.; Greene, E. C., Protein dynamics during presynaptic-complex assembly on individual single-stranded DNA molecules. *Nat Struct Mol Biol* **2014**.
164. Ishitsuka, Y.; Okumus, B.; Arslan, S.; Chen, K. H.; Ha, T., Temperature-independent porous nanocontainers for single-molecule fluorescence studies. *Anal Chem* **2010,** *82* (23), 9694-701.
165. Chen, B.; Gilbert, L. A.; Cimini, B. A.; Schnitzbauer, J.; Zhang, W.; Li, G. W.; Park, J.; Blackburn, E. H.; Weissman, J. S.; Qi, L. S.; Huang, B., Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **2013,** *155* (7), 1479-91.
166. Pantoja, R.; Rodriguez, E. A.; Dibas, M. I.; Dougherty, D. A.; Lester, H. A., Single-molecule imaging of a fluorescent unnatural amino acid incorporated into nicotinic receptors. *Biophys J* **2009,** *96* (1), 226-37.
167. Chalker, J. M.; Bernardes, G. J.; Davis, B. G., A "tag-and-modify" approach to site-selective protein modification. *Acc Chem Res* **2011,** *44* (9), 730-41.
168. Milles, S.; Tyagi, S.; Banterle, N.; Koehler, C.; VanDelinder, V.; Plass, T.; Neal, A. P.; Lemke, E. A., Click strategies for single-molecule protein fluorescence. *J Am Chem Soc* **2012,** *134* (11), 5187-95; Tyagi, S.; Lemke, E. A., Genetically encoded click chemistry for single-molecule FRET of proteins. *Methods Cell Biol* **2013,** *113*, 169-87.
169. Joo, C.; Ha, T., Labeling proteins for single-molecule FRET. *Cold Spring Harb Protoc* **2012,** *2012* (9), 1009-12.
170. Joo, C.; Ha, T., Labeling DNA (or RNA) for single-molecule FRET. *Cold Spring Harb Protoc* **2012,** *2012* (9), 1005-8; McCluskey, K.; Shaw, E.; Lafontaine, D. A.; Penedo, J. C., Single-molecule fluorescence of nucleic acids. *Methods Mol Biol* **2014,** *1076*, 759-91.
171. Wilson, J. N.; Kool, E. T., Fluorescent DNA base replacements: Reporters and sensors for biological systems. *Org Biomol Chem* **2006,** *4* (23), 4265-74.
172. Chan, W. C.; Maxwell, D. J.; Gao, X.; Bailey, R. E.; Han, M.; Nie, S., Luminescent quantum dots for multiplexed biological detection and imaging. *Curr Opin Biotechnol* **2002,** *13* (1), 40-6; Jamieson, T.; Bakhshi, R.; Petrova, D.; Pocock, R.; Imani, M.; Seifalian, A. M., Biological applications of quantum dots. *Biomaterials* **2007,** *28* (31), 4717-32.
173. Dubertret, B.; Skourides, P.; Norris, D. J.; Noireaux, V.; Brivanlou, A. H.; Libchaber, A., In vivo imaging of quantum dots encapsulated in phospholipid micelles.

Science **2002,** *298* (5599), 1759-62; Medintz, I. L.; Uyeda, H. T.; Goldman, E. R.; Mattoussi, H., Quantum dot bioconjugates for imaging, labelling and sensing. *Nat Mater* **2005,** *4* (6), 435-46.
174. Antony, E.; Kozlov, A. G.; Nguyen, B.; Lohman, T. M., Plasmodium falciparum SSB tetramer binds single-stranded DNA only in a fully wrapped mode. *J Mol Biol* **2012,** *420* (4-5), 284-95.
175. Zhang, J.; Zhou, R.; Inoue, J.; Mikawa, T.; Ha, T., Single molecule analysis of Thermus thermophilus SSB protein dynamics on single-stranded DNA. *Nucleic Acids Res* **2014,** *42* (6), 3821-32.
176. Zhou, R.; Kozlov, A. G.; Roy, R.; Zhang, J.; Korolev, S.; Lohman, T. M.; Ha, T., SSB functions as a sliding platform that migrates on DNA via reptation. *Cell* **2011,** *146* (2), 222-32.
177. Roy, R.; Kozlov, A. G.; Lohman, T. M.; Ha, T., SSB protein diffusion on single-stranded DNA stimulates RecA filament formation. *Nature* **2009,** *461* (7267), 1092-7.
178. Epstein, I. R., Cooperative and non-cooperative binding of large ligands to a finite one-dimensional lattice. A model for ligand-oligonucleotide interactions. *Biophys Chem* **1978,** *8* (4), 327-39.
179. Joo, C.; McKinney, S. A.; Nakamura, M.; Rasnik, I.; Myong, S.; Ha, T., Real-time observation of RecA filament dynamics with single monomer resolution. *Cell* **2006,** *126* (3), 515-27.
180. Tomescu, A. I.; Robb, N. C.; Hengrung, N.; Fodor, E.; Kapanidis, A. N., Single-molecule FRET reveals a corkscrew RNA structure for the polymerase-bound influenza virus promoter. *Proceedings of the National Academy of Sciences* **2014,** *111* (32), E3335-E3342.
181. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K., Scalable molecular dynamics with NAMD. *J Comput Chem* **2005,** *26* (16), 1781-802.
182. Hwang, H.; Kim, H.; Myong, S., Protein induced fluorescence enhancement as a single molecule assay with short distance sensitivity. *Proceedings of the National Academy of Sciences* **2011,** *108* (18), 7414-7418.
183. Kerry, P. S.; Turkington, H. L.; Ackermann, K.; Jameison, S. A.; Bode, B. E., Analysis of Influenza A Virus NS1 Dimer Interfaces in Solution by Pulse EPR Distance Measurements. *J Phys Chem B* **2014,** *118* (37), 10882-8.
184. Pannier, M.; Veit, S.; Godt, A.; Jeschke, G.; Spiess, H. W., Dead-Time Free Measurement of Dipole-Dipole Interactions between Spins. *J. Magn. Reson.* **2000,** *142* (2), 331-340; Bode, B. E.; Margraf, D.; Plackmeyer, J.; Dürner, G.; Prisner, T. F.; Schiemann, O., Counting the Monomers in Nanometer-Sized Oligomers by Pulsed Electron-Electron Double Resonance. *J. Am. Chem. Soc* **2007,** *129* (21), 6736-6745.
185. Kernchen, U.; Lipps, G., Thermodynamic analysis of the single-stranded DNA binding activity of the archaeal replication protein A (RPA) from Sulfolobus solfataricus. *Biochemistry* **2006,** *45* (2), 594-603.
186. Kowalczykowski, S. C.; Lonberg, N.; Newport, J. W.; von Hippel, P. H., Interactions of bacteriophage T4-coded gene 32 protein with nucleic acids. I. Characterization of the binding interactions. *J Mol Biol* **1981,** *145* (1), 75-104.
187. Patrick, S. M.; Turchi, J. J., Stopped-flow kinetic analysis of replication protein A-binding DNA: damage recognition and affinity for single-stranded DNA reveal differential contributions of k(on) and k(off) rate constants. *J Biol Chem* **2001,** *276* (25), 22630-7; Kozlov, A. G.; Lohman, T. M., Kinetic mechanism of direct transfer of Escherichia coli SSB tetramers between single-stranded DNA molecules. *Biochemistry* **2002,** *41* (39), 11611-27.
188. Berlier, J. E.; Rothe, A.; Buller, G.; Bradford, J.; Gray, D. R.; Filanoski, B. J.; Telford, W. G.; Yue, S.; Liu, J.; Cheung, C.-Y.; Chang, W.; Hirsch, J. D.; Beechem Rosaria P.

Haugland, J. M.; Haugland, R. P., Quantitative Comparison of Long-wavelength Alexa Fluor Dyes to Cy Dyes: Fluorescence of the Dyes and Their Bioconjugates. *Journal of Histochemistry & Cytochemistry* **2003,** *51* (12), 1699-1712.
189. Di Fiori, N.; Meller, A., The Effect of dye-dye interactions on the spatial resolution of single-molecule FRET measurements in nucleic acids. *Biophys J* **2010,** *98* (10), 2265-72.
190. Quinn, S. D.; Dalgarno, P. A.; Cameron, R. T.; Hedley, G. J.; Hacker, C.; Lucocq, J. M.; Baillie, G. S.; Samuel, I. D.; Penedo, J. C., Real-time probing of β-amyloid self-assembly and inhibition using fluorescence self-quenching between neighbouring dyes. *Mol Biosyst* **2014,** *10* (1), 34-44.
191. Huang, Y. H.; Huang, C. Y., Characterization of a single-stranded DNA-binding protein from Klebsiella pneumoniae: mutation at either Arg73 or Ser76 causes a less cooperative complex on DNA. *Genes Cells* **2012,** *17* (2), 146-57.
192. Kozlov, A. G.; Weiland, E.; Mittal, A.; Waldman, V.; Antony, E.; Fazio, N.; Pappu, R. V.; Lohman, T. M., Intrinsically Disordered C-Terminal Tails of E. coli Single-Stranded DNA Binding Protein Regulate Cooperative Binding to Single-Stranded DNA. *J Mol Biol* **2015,** *427* (4), 763-74.
193. Huang, Y. H.; Huang, C. Y., C-terminal domain swapping of SSB changes the size of the ssDNA binding site. *Biomed Res Int* **2014,** *2014*, 573936.
194. Mishra, G.; Levy, Y., Molecular determinants of the interactions between proteins and ssDNA. *Proc Natl Acad Sci U S A* **2015**.
195. Bujalowski, W.; Overman, L. B.; Lohman, T. M., Binding mode transitions of Escherichia coli single strand binding protein-single-stranded DNA complexes. Cation, anion, pH, and binding density effects. *J Biol Chem* **1988,** *263* (10), 4629-40.
196. Hethke, C.; Bergerat, A.; Hausner, W.; Forterre, P.; Thomm, M., Cell-free transcription at 95 degrees: thermostability of transcriptional components and DNA topology requirements of Pyrococcus transcription. *Genetics* **1999,** *152* (4), 1325-1333.
197. Lohman, T. M.; Bujalowski, W., Negative cooperativity within individual tetramers of Escherichia coli single strand binding protein is responsible for the transition between the (SSB)35 and (SSB)56 DNA binding modes. *Biochemistry* **1988,** *27* (7), 2260-5.
198. Nandakumar, J.; Podell, E. R.; Cech, T. R., How telomeric protein POT1 avoids RNA to achieve specificity for single-stranded DNA. *Proc Natl Acad Sci U S A* **2010,** *107* (2), 651-6.
199. Ayed, A.; Mulder, F. A.; Yi, G. S.; Lu, Y.; Kay, L. E.; Arrowsmith, C. H., Latent and active p53 are identical in conformation. *Nat Struct Biol* **2001,** *8* (9), 756-60.
200. Jelinska, C.; Petrovic-Stojanovska, B.; Ingledew, W. J.; White, M. F., Dimer-dimer stacking interactions are important for nucleic acid binding by the archaeal chromatin protein Alba. *Biochem J* **2010,** *427* (1), 49-55.
201. Paytubi, S.; McMahon, S. A.; Graham, S.; Liu, H.; Botting, C. H.; Makarova, K. S.; Koonin, E. V.; Naismith, J. H.; White, M. F., Displacement of the canonical single-stranded DNA-binding protein in the Thermoproteales. *Proc Natl Acad Sci U S A* **2012,** *109* (7), E398-405.
202. Lamond, A. I.; Gibson, T. J., Catalytic RNA and the origin of genetic systems. *Trends Genet* **1990,** *6* (5), 145-149; Orgel, L. E., The origin of life—a review of facts and speculations. *Trends in biochemical sciences* **1998,** *23* (12), 491-495; Pace, N. R.; Marsh, T. L., RNA catalysis and the origin of life. *Orig Life Evol Biosph* **1985,** *16* (2), 97-9116.
203. Swairjo, M. A.; Morales, A. J.; Wang, C. C.; Ortiz, A. R.; Schimmel, P., Crystal structure of trbp111: a structure-specific tRNA-binding protein. *EMBO J* **2000,** *19* (23), 6287-6298.
204. Simos, G.; Sauer, A.; Fasiolo, F.; Hurt, E. C., A conserved domain within Arc1p delivers tRNA to aminoacyl-tRNA synthetases. *Mol Cell* **1998,** *1* (2), 235-242.

205. Cusack, S.; Yaremchuk, A.; Krikliviy, I.; Tukalo, M., tRNA(Pro) anticodon recognition by Thermus thermophilus prolyl-tRNA synthetase. *Structure* **1998,** *6* (1), 101-108.
206. Li, W.; Hoffman, D. W., Structure and dynamics of translation initiation factor aIF-1A from the archaeon Methanococcus jannaschii determined by NMR spectroscopy. *Protein Sci* **2001,** *10* (12), 2426-2438; Sette, M.; van Tilborg, P.; Spurio, R.; Kaptein, R.; Paci, M.; Gualerzi, C. O.; Boelens, R., The structure of the translational initiation factor IF1 from E.coli contains an oligomer-binding motif. *EMBO J* **1997,** *16* (6), 1436-1443; Wu, B.; Yee, A.; Pineda-Lucena, A.; Semesi, A.; Ramelot, T. A.; Cort, J. R.; Jung, J.-W.; Edwards, A.; Lee, W.; Kennedy, M.; Arrowsmith, C. H., Solution structure of ribosomal protein S28E from Methanobacterium thermoautotrophicum. *Protein Sci* **2003,** *12* (12), 2831-2837.
207. Perales, C.; Cava, F.; Meijer, W. J.; Berenguer, J., Enhancement of DNA, cDNA synthesis and fidelity at high temperatures by a dimeric single-stranded DNA-binding protein. *Nucleic Acids Res* **2003,** *31* (22), 6473-80;
208. Olszewski, M.; Rebała, K.; Szczerkowska, Z.; Kur, J., Application of SSB-like protein from Thermus aquaticus in multiplex PCR of human Y-STR markers identification. *Mol Cell Probes* **2005,** *19* (3), 203-5; Dąbrowski, S.; Kur, J., Cloning, Overexpression, and Purification of the Recombinant His-Tagged SSB Protein ofEscherichia coliand Use in Polymerase Chain Reaction Amplification. *Protein Expression and Purification* **1999,** *16* (1), 96-102.
209. Lohman, T. M., Kinetics of protein-nucleic acid interactions: use of salt effects to probe mechanisms of interaction. *CRC Crit Rev Biochem* **1986,** *19* (3), 191-245.
210. Wang, Y.; Yang, Q.; Wang, Z., The evolution of nanopore sequencing. *Front Genet* **2014,** *5*, 449; Japrung, D.; Bahrami, A.; Nadzeyka, A.; Peto, L.; Bauerdick, S.; Edel, J. B.; Albrecht, T., SSB binding to single-stranded DNA probed using solid-state nanopore sensors. *J Phys Chem B* **2014,** *118* (40), 11605-12.
211. Clamer, M.; Höfler, L.; Mikhailova, E.; Viero, G.; Bayley, H., Detection of 3'-end RNA uridylation with a protein nanopore. *ACS Nano* **2014,** *8* (2), 1364-74; Rodriguez-Larrea, D.; Bayley, H., Multistep protein unfolding during nanopore translocation. *Nat Nanotechnol* **2013,** *8* (4), 288-95; Venkatesan, B. M.; Bashir, R., Nanopore sensors for nucleic acid analysis. *Nat Nanotechnol* **2011,** *6* (10), 615-24; Clarke, J.; Wu, H. C.; Jayasinghe, L.; Patel, A.; Reid, S.; Bayley, H., Continuous base identification for single-molecule nanopore DNA sequencing. *Nat Nanotechnol* **2009,** *4* (4), 265-70.
212. Briggs, K.; Charron, M.; Kwok, H.; Le, T.; Chahal, S.; Bustamante, J.; Waugh, M.; Tabard-Cossa, V., Kinetics of nanopore fabrication during controlled breakdown of dielectric membranes in solution. *Nanotechnology* **2015,** *26* (8), 084004.
213. Trott, O.; Olson, A. J., AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **2010,** *31* (2), 455-61.
214. Zaparty, M.; Esser, D.; Gertig, S.; Haferkamp, P.; Kouril, T.; Manica, A.; Pham, T. K.; Reimann, J.; Schreiber, K.; Sierocinski, P.; Teichmann, D.; van Wolferen, M.; von Jan, M.; Wieloch, P.; Albers, S. V.; Driessen, A. J.; Klenk, H. P.; Schleper, C.; Schomburg, D.; van der Oost, J.; Wright, P. C.; Siebers, B., "Hot standards" for the thermoacidophilic archaeon Sulfolobus solfataricus. *Extremophiles* **2010,** *14* (1), 119-42.
215. Grogan, D. W., Stability and repair of DNA in hyperthermophilic Archaea. *Curr Issues Mol Biol* **2004,** *6* (2), 137-44.