2012

Noncovalent interactions in the E. coli replisome: novel aspects of single-stranded DNA-binding protein (SSB) and the β2 sliding clamp

Claire E. Mason
University of Wollongong

Recommended Citation

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au
UNIVERSITY OF WOLLONGONG

COPYRIGHT WARNING

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site. You are reminded of the following:

Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.
Noncovalent Interactions in the *E. coli* Replisome: Novel Aspects of Single-Stranded DNA-Binding Protein (SSB) and the $\beta_2$ Sliding Clamp

A thesis submitted in fulfilment of the requirements for the award of the degree Doctor of Philosophy from University of Wollongong by Claire E. Mason Bachelor of Medicinal Chemistry (Advanced, Honours) School of Chemistry 2012
DECLARATION

I, Claire E. Mason, declare that this thesis, submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Chemistry, University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. This document has not been submitted for qualifications at any other academic institution.

Claire E. Mason

14th June, 2012
ACKNOWLEDGEMENTS

A PhD is a long and often difficult process, and there are many people who have contributed significantly to this work.

Firstly, I feel very fortunate to have had the assistance of supervisors who were extremely generous with their time and advice: Professor Jenny Beck and Professor Nick Dixon. Jenny, thank you for being a wonderful supervisor throughout the whole process, and especially for your advice and words of encouragement when things were difficult. Nick, thank you for introducing me to the world of DNA replication, and for your ideas, exceptional knowledge and guidance.

The work in the latter part of this thesis could not have been completed without the assistance, supervision and ideas of Dr Slobodan Jergic. Boban – I can’t thank you enough for sharing your work with me, teaching me so much, and for sharing your interesting perspectives on life, bacterial feelings, and insect sacrifice! Many thanks also to Nick Horan, who contributed significantly to work described in this thesis.

To all the members of the Dixon and Mass Spec. groups, past and present, thank you for your help and friendship, and for brightening the PhD journey. One of the best parts of doing a PhD was meeting so many great friends, who I hope to stay in touch with for many years to come. Thitima and Ben, thank you for keeping the mass spec lab running smoothly and for help when instruments were being temperamental. Thanks also to Allen Lo and Yao Wang for assistance with protein purifications and cloning.

Finally, thank you to my long-suffering family and friends for your love, support and patience. I am especially indebted to my parents, and last but not least, to Tim – thank you for your love and endless understanding, and most of all for putting up with me through my many PhD-induced meltdowns and stopping me from going (too) crazy. I am very grateful.
PUBLICATIONS


Contribution to paper: Designed experiments with assistance from J. Beck and N. Dixon; carried out all experiments; wrote manuscript with input from J. Beck and N.Dixon.


Contribution to paper: Created, purified and characterised mixed β₂ dimers for use in DNA replication assays; carried out selected replication assays (with S. Jergic); carried out surface plasmon resonance assays (with S. Jergic).
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Adenine</td>
</tr>
<tr>
<td>Å</td>
<td>Ångström(s)</td>
</tr>
<tr>
<td>aa</td>
<td>Amino acid</td>
</tr>
<tr>
<td>A&lt;sub&gt;λ&lt;/sub&gt;</td>
<td>Absorbance at a wavelength of λ nm</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>ATPγS</td>
<td>Adenosine 5′-(3-thiotriphosphate)</td>
</tr>
<tr>
<td>Bio</td>
<td>Biotin</td>
</tr>
<tr>
<td>bp</td>
<td>Base pair(s)</td>
</tr>
<tr>
<td>C</td>
<td>Cytosine</td>
</tr>
<tr>
<td>CBM</td>
<td>Clamp-binding motif</td>
</tr>
<tr>
<td>CID</td>
<td>Collision-induced dissociation</td>
</tr>
<tr>
<td>Da</td>
<td>Dalton(s)</td>
</tr>
<tr>
<td>DEAE</td>
<td>Diethylaminoethyl</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxynucleotide triphosphate</td>
</tr>
<tr>
<td>dsDNA</td>
<td>Double-stranded DNA</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>ε&lt;sub&gt;λ&lt;/sub&gt;</td>
<td>Molar absorption coefficient at a wavelength of λ nm</td>
</tr>
<tr>
<td>E. coli</td>
<td>Escherichia coli</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediamine tetraacetic acid</td>
</tr>
<tr>
<td>ESI</td>
<td>Electrospray ionisation</td>
</tr>
</tbody>
</table>
FPLC: Fast protein liquid chromatography
FRET: Fluorescence resonance energy transfer
G: Guanine
g: Gravity
GuHCl: Guanidine hydrochloride
HEPES: 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid
HPLC: High performance liquid chromatography
HSP: Heat shock protein
ITC: Isothermal titration calorimetry
K_D: Equilibrium dissociation constant
kDa: Kilodalton(s)
LB: Lysogeny broth/Luria-Bertani medium
MES: 2-(4-morpholino)ethanesulfonic acid
MS: Mass spectrometry
MWCO: Molecular weight cut-off
m/z: Mass-to-charge ratio
nanoESI: Nanoelectrospray ionisation
NMR: Nuclear magnetic resonance
Nt: Nucleotide(s)
OB: Oligonucleotide/oligosaccharide-binding (fold)
PAGE: Polyacrylamide gel electrophoresis
PCNA: Proliferating cell nuclear antigen
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHP</td>
<td>Polymerase and histidinol phosphatase</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PMSF</td>
<td>Phenylmethylsulfonyl fluoride</td>
</tr>
<tr>
<td>Pol III HE</td>
<td>DNA polymerase III holoenzyme</td>
</tr>
<tr>
<td>Q-ToF</td>
<td>Quadrupole time-of-flight</td>
</tr>
<tr>
<td>RBS</td>
<td>Ribosome binding site</td>
</tr>
<tr>
<td>RFC</td>
<td>Replication factor C</td>
</tr>
<tr>
<td>$R_{\text{max}}$</td>
<td>Response at saturation (SPR)</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RPA</td>
<td>Replication protein A</td>
</tr>
<tr>
<td>RU</td>
<td>Resonance units</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SPR</td>
<td>Surface plasmon resonance</td>
</tr>
<tr>
<td>SSB</td>
<td>Single-stranded DNA-binding protein</td>
</tr>
<tr>
<td>ssDNA</td>
<td>Single-stranded DNA</td>
</tr>
<tr>
<td>T</td>
<td>Thymine</td>
</tr>
<tr>
<td>Taq</td>
<td><em>Thermus aquaticus</em></td>
</tr>
<tr>
<td>TLS</td>
<td>Translesion synthesis</td>
</tr>
<tr>
<td>ToF</td>
<td>Time-of-flight</td>
</tr>
<tr>
<td>Tris</td>
<td>Tris(hydroxymethyl)aminomethane</td>
</tr>
<tr>
<td>Tris-HCl</td>
<td>Tris(hydroxymethyl)aminomethane hydrochloride</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
</tbody>
</table>
ABSTRACT

This thesis presents analyses of key protein-protein and protein-nucleic acid interactions in the *E. coli* replication machinery. The *E. coli* single-stranded DNA-binding protein (SSB) is a central mediator of DNA metabolism, binding to regions of exposed single-stranded DNA (ssDNA) and playing crucial roles in DNA replication, recombination and repair. SSB is a stable homotetramer, held together by interactions between the N-terminal domains of its subunits. Structurally, the N-terminal domains form classic oligonucleotide-binding folds (OB-folds), and are also responsible for ssDNA binding. The SSB-ssDNA complexes may adopt a number of configurations, differing in the extent of SSB-ssDNA contact, and are heavily influenced by salt concentration and protein-DNA ratio. The C-terminus of SSB appears to be unstructured, and its highly conserved tip (~eight residues) is responsible for interacting with a host of other proteins involved in DNA metabolism and regulating their functions. Recent evidence suggests that the SSB C-termini, in the absence of bound DNA, may be able to occupy the OB-folds of the tetramer via an electrostatic interaction.

The first aim of this thesis was to use electrospray ionisation mass spectrometry (ESI-MS) to characterise SSB and its complexes with ssDNA, and to assess the validity of ESI-MS for the analysis of SSB-ssDNA interactions. ESI-MS was used to determine the rate of subunit exchange between unlabelled and uniformly $^{15}$N-labelled SSB, revealing a slow subunit exchange rate. Subunit exchange in a truncated version of SSB missing the last eight residues, SSBΔC8, proceeded markedly more rapidly in low salt, suggesting for the first time that the C-terminus plays a role in stabilising the SSB tetramer. This is proposed to occur via transient intersubunit interactions of the C-termini with the OB-folds from adjacent subunits, and may have implications for how other proteins access the SSB C-terminus.
ESI-MS analysis of SSB-ssDNA complexes formed under a variety of salt conditions allowed complexes corresponding to different SSB binding modes to be observed. Subunit exchange was dramatically inhibited when SSB was bound to ssDNA oligonucleotides, suggesting that DNA-binding locks the tetramer in place, even when only half the SSB subunits are DNA-bound. In addition, analysis of SSB transfer between discrete ssDNA molecules by ESI-MS revealed a dramatic preference for transfer to occur from an initial complex containing unoccupied SSB subunits, confirming the predictions from solution phase experiments which determined the mechanism of direct transfer via a ternary complex. This work showed, therefore, that SSB-ssDNA complexes observed by mass spectrometry offer an accurate reflection of the specific contacts formed in the different binding modes.

The protein clamp responsible for the processivity of *E. coli* DNA replication, $\beta_2$, also contains a well-conserved site through which it interacts with a large number of binding partners. The $\beta_2$ dimer operates by encircling and sliding along ssDNA, thus acting as a DNA tether for its binding partners. It contains an identical, primarily hydrophobic binding pocket on each of its subunits, enabling it to interact with a multiple polymerases and accessory subunits along with its primary contact with the replicative polymerase subunit, $\alpha$. Recently in our laboratory, a previously undescribed interaction between $\beta_2$ and the replicative exonuclease subunit, $\varepsilon$, was discovered.

The second aim of this thesis was to characterise the strength and several functional aspects of the newly identified $\varepsilon$-$\beta_2$ interaction: particularly whether the interaction is mediated thorough the common protein binding cleft on the clamp. This required the creation of $\beta_2$ dimers in which the putative $\varepsilon$-binding cleft was disrupted in either one or both subunits. Attempts to create such dimers by fusing both subunits into a single gene product were unsuccessful. Instead, the subunit exchange properties of the $\beta_2$ dimer in various salt concentrations were determined by ESI-MS, and this information was used to create and purify hybrid dimers containing a single, intact
binding site. Functional replication assays using these dimers showed that ε-dependent rolling circle replication reactions were dependent on the presence of both intact binding clefts in the β₂ dimer, indicating that the ε-β₂ interaction occupies the predicted binding pocket on β₂, and occurs concurrently with the β₂-polymerase interaction. These assays also suggested that the role played by the ε-β₂ interaction during DNA replication is to stabilise the replication fork. Surface plasmon resonance measurements of the affinity of the interaction between β₂ and a peptide corresponding to the proposed clamp binding motif from ε showed that the interaction was weak (Kᵦ ~200 μM). Thus, the ε-β₂ interaction, whilst transient alone, likely gains its significance as part of a complex network of protein-protein interactions at the replication fork.
TABLE OF CONTENTS

Declaration i
Acknowledgements ii
Publications iii
Abbreviations iv
Abstract vii
Table of Contents x
List of Figures xiv
List of Tables xviii

Chapter 1 : General introduction 1
  1.1 DNA Replication in E. coli 1
    1.1.1 Initiation of replication 4
    1.1.2 The primosome 6
    1.1.3 The DNA polymerase III holoenzyme 11
      1.1.3.1 The Pol III core 11
      1.1.3.2 The β2 sliding clamp 20
      1.1.3.3 The clamp loader complex 21
    1.1.4 Dynamics of DNA replication 28
    1.1.5 Single-stranded DNA-binding protein (SSB) 30
      1.1.5.1 SSB structure 31
      1.1.5.2 SSB-ssDNA interactions 33
      1.1.5.3 The SSB C-terminus and SSB-protein interactions 35
  1.2 Mass Spectrometry for the Study of Noncovalent Protein Interactions 42
    1.2.1 Electrospray ionisation mass spectrometry for analysis of biomolecules 42
    1.2.2 Applications of ESI-MS to intact protein complexes and stoichiometry determination 48
    1.2.3 Applications of ESI-MS to protein-nucleic acid interactions 52
    1.2.4 ESI-MS for the study of macromolecular dynamics 54
  1.3 Aims and scope of this thesis 56

Chapter 2 : General materials and methods 58
  2.1 Materials 58
    2.2 Plasmids, Strains and Growth Media 58
3.4 Conclusions

Chapter 4: Examination of SSB-ssDNA complexes by mass spectrometry and surface plasmon resonance

4.1 Introduction

4.2 Materials and Methods:
   4.2.1 Proteins and oligonucleotides
   4.2.2 Mass spectrometry competition experiments.
   4.2.3 Surface plasmon resonance competition experiments.

4.3 Results and Discussion:
   4.3.1 Conversion between SSB ssDNA-binding modes observed by mass spectrometry
   4.3.2 Transfer of SSB between ssDNA oligonucleotides detected by mass spectrometry
   4.3.3 Transfer of SSB between ssDNA oligonucleotides detected by surface plasmon resonance

4.4 Conclusions:

Chapter 5: Production of modified $\beta_2$ clamps for examination of the $\varepsilon$-$\beta_2$ interaction

5.1 Introduction

5.2 Materials and Methods
   5.2.1 Oligonucleotide primers
   5.2.2 Construction of plasmids
   5.2.3 Buffers for protein purification
   5.2.4 Overproduction and purification of His$_6$-single-chain $\beta_2$
   5.2.5 Overproduction and purification of (His$_6$-$\beta_2$)$_2$
   5.2.6 Overproduction and purification of (His$_6$-$\beta^C_2$)$_2$
   5.2.7 Subunit exchange between ($\beta_{WT}$)$_2$ and (His$_6$-$\beta_2$)$_2$
   5.2.8 Preparation of heterodimeric His$_6$-$\beta_{WT}/\beta_{WT}$ and His$_6$-$\beta^C_2/\beta_{WT}$

5.3 Results and Discussion
   5.3.1 Problems with single-chain $\beta_2$ constructs
   5.3.2 Subunit exchange in the wild-type $\beta_2$ dimer
      5.3.2.1 $\beta_2$ subunit exchange is minimal under native mass spectrometry conditions
      5.3.2.2 Development of solution conditions for formation and purification of mixed $\beta_2$ dimers
      5.3.2.3 $\beta_2$ subunit exchange under replication assay conditions
   5.3.3 The truncated mutant $\beta^C_2$ forms a stable dimer
   5.3.4 Production and purification of mixed $\beta_2$ dimers using insights gained from subunit exchange.

5.4 Conclusions
Chapter 6: Functional investigations of the *E. coli* ε-β₂ interaction 169

6.1 Introduction 169
6.1.1 DNA replication assays as a means for dissecting the functional properties of polymerase subunits 169
6.1.2 A strand displacement reaction with an absolute dependence on ε 173
6.1.3 Discovery of an ε-β₂ interaction 178

6.2 Materials and Methods 182
6.2.1 Proteins 182
6.2.2 Oligonucleotides and DNA templates 182
6.2.3 Buffers 184
6.2.4 DNA replication assays 184

6.3 Results and Discussion 186
6.3.1 Both protein binding clefts in the β₂ dimer are required for ε-dependent strand displacement DNA synthesis 186
6.3.2 Dependence of strand displacement synthesis on β₂ concentration 190
6.3.3 The ε-β₂ interaction is also required for ‘normal’ DNA replication 193
6.3.4 A model for the ε-β₂ interaction during replication 202

6.4 Conclusions 208

Chapter 7: The β₂ clamp interacts directly with the predicted clamp binding motif in ε 210

7.1 Introduction 210

7.2 Materials and Methods 220
7.2.1 Peptides and proteins 220
7.2.2 Surface plasmon resonance 221
7.2.3 Determination of equilibrium binding affinities 224

7.3 Results and Discussion 226
7.3.1 Direct interactions between β₂ and ε peptides can be detected by SPR 226
7.3.1.1 Interaction between the ε₁ peptide and β₂ 226
7.3.1.2 There is a direct interaction between the ε₁⁰⁷ peptide and β₂ 229
7.3.1.3 The interaction between the ε₀ peptide and β₂ is extremely weak 232
7.3.2 A proposed role for the ε-β₂ interaction in replication 235

7.4 Conclusions 238

Chapter 8: Concluding remarks 239

References 243
LIST OF FIGURES

Figure 1.1. Bidirectional replication of the *E. coli* chromosome. 2
Figure 1.2. Schematic representation of the components at the *E. coli* replication fork. 3
Figure 1.3. Initiation of replication from the *E. coli* oriC. 5
Figure 1.4. Semidiscontinuous DNA replication of the leading and lagging strands. 7
Figure 1.5. The *E. coli* primosome. 9
Figure 1.6. Domain organisation and structure of the α subunit. 13
Figure 1.7. Structures of the ε proofreading subunit. 17
Figure 1.8. The clamp loader complex. 22
Figure 1.9. The *dnaX* gene produces two protein products. 23
Figure 1.10. Dynamics at the *E. coli* replication fork. 29
Figure 1.11. Domain organisation and structure of SSB. 32
Figure 1.12. Proposed structures of the (SSB)$_{35}$ and (SSB)$_{65}$ ssDNA binding modes. 34
Figure 1.13. The sequence of the SSB C-terminus is highly conserved amongst bacteria. 36
Figure 1.14. The SSB C-terminus possesses similar binding sites on different proteins. 40
Figure 1.15. The electrospray ionisation process. 43
Figure 1.16. Effect of pH on protein folding detected by ESI-MS. 45
Figure 1.17. Arrangement of a quadrupole-time-of-flight (Q-ToF) mass spectrometer used for the transmission of proteins and protein complexes. 47
Figure 1.18. Mass spectrometry for measuring protein dynamics. 55
Figure 3.1. Model for the proposed interaction between the SSB C-terminus and the structured OB-fold domain. 67
Figure 3.2. Monitoring subunit exchange by mass spectrometry. 72
Figure 3.3. Native nanoESI-mass spectra of SSB and $^{15}$N-SSB tetramers. 85
Figure 3.4. Subunit exchange between SSB and $^{15}$N-SSB monitored by nanoESI-MS. 87
Figure 3.5. SSB subunit exchange rates in 1 M ammonium acetate. 89
Figure 3.6. SSB subunit exchange rates in 10 mM ammonium acetate. 90
Figure 3.7. The SSB subunit exchange rate is independent of protein concentration. 92
Figure 3.8. Conformations for occupation of SSB OB-folds by the C-terminal tails. 93
Figure 3.9. Native nanoESI mass spectra of SSBΔC8 and 15N-SSBΔC8 tetramers. 95
Figure 3.10. Subunit exchange between SSBΔC8 and 15N-SSBΔC8 by nanoESI-MS. 96
Figure 3.11. SSB-oligonucleotide complexes observed by nanoESI-MS. 98
Figure 3.12. Subunit exchange between SSB and 15N-SSB bound to oligonucleotides. 100
Figure 4.1. SSB can diffuse along ssDNA by reptation. 105
Figure 4.2. Direct transfer of SSB tetramers between oligonucleotides. 106
Figure 4.3. Conversion between DNA-binding modes of SSB observed by nanoESI-MS. 111
Figure 4.4. Transfer of SSB from a 1:1 SSB:(dT)₇₀ complex onto added (dT)₃₅ does not occur. 113
Figure 4.5. Transfer of SSB from a 1:1 SSB:(dT)₇₀ complex onto added Bio-(dT)₇₀ can occur at elevated salt concentration. 114
Figure 4.6. SSB bound to (dT)₇₀ in the (SSB)₃₅ mode can be transferred to added Bio-(dT)₇₀. 115
Figure 4.7. Transfer of SSB from complexes with (dT)₃₅ onto added (dT)₇₀ or Bio-(dT)₇₀ readily occurs. 116
Figure 4.8. Transfer of SSB tetramers from immobilised (dT)₇₀ detected by surface plasmon resonance. 119
Figure 4.9. Transfer of SSB tetramers from immobilised (dT)₃₅ detected by surface plasmon resonance. 121
Figure 5.1. The crystal structure of the β₂ sliding clamp. 125
Figure 5.2. A hydrophobic cleft in β mediates its protein-protein interactions. 129
Figure 5.3. Plasmid pCM1513. 136
Figure 5.4. Plasmid pCM1503. 137
Figure 5.5. Plasmid pCM1510. 138
Figure 5.6. Plasmid pCM1531. 139
Figure 5.7. Positive ion nanoESI-mass spectrum of denatured His₆-single chain β₂. 150
Figure 5.8. Positive ion nanoESI mass spectra of β₂ and (His₆-β₂). 155
Figure 5.9. Subunit exchange between β₂ and (His₆-β),. 156
Figure 5.10. Subunit exchange between β₁ and (His₆-β), at high salt concentrations. 158
Figure 5.11. Subunit exchange between β₂ and (His₆-β), at low salt concentrations. 159
Figure 5.12. Subunit exchange between β₁ and (His₆-β), under replication assay conditions. 161
Figure 5.13. The truncated mutant βᶜ clamp folds normally, but has a disrupted protein-binding cleft. 163
Figure 5.14. (His₆-βᶜ), is a stable dimer. 164
Figure 5.15. Production of hybrid β₂ dimers. 165
Figure 5.16. NanoESI-MS of β heterodimers following purification. 166
Figure 6.1. A simple DNA replication assay. 170
Figure 6.2. Strand displacement rolling-circle DNA replication assays. 171
Figure 6.3. Proposed model for stabilisation of helicase-independent strand displacement replication. 172
Figure 6.4. A representative coupled circle-filling/strand displacement assay. 174
Figure 6.5. Helicase-independent strand displacement is dependent on the presence of the ε subunit. 175
Figure 6.6. The contribution of ε to strand displacement synthesis is independent of its proofreading exonuclease activity. 177
Figure 6.7. A potential β₁ binding site in ε based on sequence analysis. 179
Figure 6.8. Variations in the putative β₁ binding site of ε affect strand displacement synthesis according to their predicted β₁ binding strength. 180
Figure 6.9. The primed M₁₃ template DNA. 183
Figure 6.10. Both protein binding sites in the β₂ dimer are required to support strand displacement synthesis. 188
Figure 6.11. Strand displacement requires an unexpectedly high concentration of β₂. 191
Figure 6.12. Strand displacement is observed at a very high concentration of His-βᶜ/βʷT. 192
Figure 6.13. Variations in the predicted β₁-binding motif in ε also affect primer extension DNA replication. 195
Figure 6.14. Processive replication is mediated by the internal clamp binding motif in $\alpha$. 197

Figure 6.15. Potential roles for the $\varepsilon$-$\beta_2$ contact during replication. 199

Figure 6.16. The absence of a second protein binding cleft in $\beta_2$ results in weakened synthesis. 200

Figure 6.17. The $\varepsilon$-$\beta_2$ interaction plays a stabilising role in replication. 201

Figure 6.18. Models for the $\alpha$-$\beta_2$ complex in the absence and presence of DNA. 203

Figure 6.19. Proposed model for the $\alpha\varepsilon\theta\beta_2$-DNA complex. 205

Figure 7.1. NanoESI-mass spectra of $\varepsilon$-$\beta_2$ mixtures. 212

Figure 7.2. NanoESI-mass spectra of $\alpha\varepsilon\theta$-$\beta_2$ mixtures. 213

Figure 7.3. Principles of surface plasmon resonance applied to the detection of macromolecular interactions. 215

Figure 7.4. A typical SPR sensorgram. 216

Figure 7.5. The ProteOn XPR36 system allows analysis of multiple interactions simultaneously. 218

Figure 7.6. Experimental setup for surface plasmon resonance experiments. 223

Figure 7.7. ProteOn sensorgrams showing binding of $\beta_2$ to immobilised $\varepsilon_L$ peptide. 227

Figure 7.8. Binding isotherm for the binding of $\beta_2$ to immobilised $\varepsilon_L$ peptide. 228

Figure 7.9. ProteOn sensorgrams showing binding of $\beta_2$ to immobilised $\varepsilon_{WT}$ peptide. 230

Figure 7.10. Binding isotherm for the binding of $\beta_2$ to immobilised $\varepsilon_{WT}$ peptide. 231

Figure 7.11. ProteOn sensorgrams showing binding of $\beta_2$ to immobilised $\varepsilon_Q$ peptide. 234

Figure 7.12. Binding curves for the $\varepsilon_{L\text{pep}}\beta_2$, $\varepsilon_{WT\text{pep}}\beta_2$ and $\varepsilon_{Q\text{pep}}\beta_2$ interactions measured by SPR. 235
LIST OF TABLES

Table 2.1. Properties of oligonucleotides used in this thesis. 63

Table 2.2. Typical instrumental parameters used for nanoESI-MS. 65

Table 3.1. Rate constants for decay of homotetrameric species during SSB subunit exchange. 90
Chapter 1 : **GENERAL INTRODUCTION**

### 1.1 DNA Replication in *E. coli*

DNA replication, the process by which organisms duplicate their genetic information, is essential for cell division, and thus for the propagation of life. The faithful duplication of a cell’s genetic information by replication of its chromosomal DNA is a huge task, and is able to be accomplished only by the precisely coordinated actions of a large number of enzymes and accessory proteins, which together form a molecular machine known as the replisome[1-5]. The bacterium *Escherichia coli* (*E. coli*) has served as a model organism for teasing out the molecular details of bacterial DNA replication, many of which find parallels in the more complex DNA replication of eukaryotic organisms. An understanding of the numerous protein-protein interactions involved in bacterial DNA replication is important not only from a basic scientific standpoint; the essential and conserved nature of replication also presents it as a potential drug target for antibacterial agents[6]. Although a relatively comprehensive picture of DNA replication in *E. coli* has been constructed, the complexity of the process means that details continue to be added.

The entire ~4.6 million base pair circular chromosome of *E. coli* is replicated in about 40 minutes, requiring the replisome to progress at the remarkable speed of ~1000 base pairs per second[7]. The *E. coli* chromosome contains a single origin of replication, oriC, at which replication is initiated[8,9]. Unwinding at this origin results in the formation of two replication forks which serve as the points of assembly for two separate replisomes which proceed in opposite directions around the chromosome[8] (Figure 1.1). Replication is terminated at specific Ter sites within the chromosome to which the terminator protein Tus binds tightly, forming a trap which arrests replication[10-11].
The protein components required at each replication fork during DNA replication are: a replicative helicase (DnaB in *E. coli*), responsible for unwinding the duplex DNA ahead of the replication fork to allow it to be copied[14]; a primase (DnaG), which synthesises short RNA primers from which DNA synthesis can proceed[15], and either two or three copies of a polymerase, which carries out DNA synthesis. The replicative polymerase in *E. coli* is the DNA polymerase III holoenzyme (Pol III HE), an assembly of 10 different proteins, each of which carry out specialised roles[2,5,16,17]. The single-stranded DNA-binding protein (SSB) is also an important player in DNA replication, binding to the single-stranded DNA which is an intermediate in the process, and participating in multiple interactions which contribute to the overall organisation of replication. A model of the *E. coli* replication fork is presented in Figure 1.2.

This introduction will give an overview of the processes and components involved in *E. coli* DNA replication, with the intention of providing a context for the two proteins which form
Figure 1.2. Schematic representation of the components at the *E. coli* replication fork. DnaB helicase encircles the lagging strand DNA and unwinds the parental duplex in a 5'→3' direction. DnaG primase synthesises short RNA primers every 1-2 kb on the lagging strand. The leading and lagging strand core polymerases (αεθ) are tethered to their respective DNA template strands by the β2 sliding clamp. The lagging strand is synthesised as a series of discontinuous Okazaki fragments, each beginning from a new RNA primer. Single-stranded DNA-binding protein (SSB) coats the exposed single-stranded DNA generated on the lagging strand. The τ subunits of the clamp loader complex tether the leading and lagging strand polymerases together. The τδδ'χψ version of the clamp loader complex is shown here, with the third τ subunit tethering a third ‘reserve’ core polymerase.

The basis of the work described in this thesis: the β2 sliding clamp and single-stranded DNA binding protein (SSB). The β2 clamp will be briefly described in this introduction (section 1.1.3.2), but a more complete treatment can be found in the introduction to Chapter 5.
An overview of the general features of SSB will be given here (section 1.1.5), and more specific aspects of its function will be described in sections 3.1 and 4.1.

### 1.1.1 Initiation of replication

The origin of replication in *E. coli*, oriC, is a ~245 base pair (bp) sequence\(^\text{[18]}\) containing five 9 bp repeats (DnaA boxes), which serve as recognition sequences for the initiator protein, DnaA\(^{[19,20]}\). DnaA is a member of the AAA+ ATPase family\(^{[21]}\), and ATP binding and hydrolysis are integral to its function as an initiator\(^{[22]}\). Following origin recognition, ATP-dependent binding of DnaA to additional, weaker recognition sites within the origin\(^{[23,24]}\) results in the cooperative assembly of multiple DnaA molecules into a helical nucleoprotein filament\(^{[21]}\). This filament, along with the binding of other DNA-bending architectural proteins (HU and IHF)\(^{[25]}\) introduces topological strain into the DNA, and facilitates localised melting of an AT-rich region located adjacent to oriC, forming a small ‘bubble’ of unwound DNA\(^{[9,26]}\) (Figure 1.3).

Once the strands are separated, DnaA fulfils the second part of its role as an initiator – recruitment of the replicative helicase, DnaB, to the newly formed replication forks\(^{[27]}\) (Figure 1.3).

DnaB is a 52 kDa protein\(^{[28]}\) consisting of two domains: a ~12 kDa N-terminal domain and a larger, ~33 kDa C-terminal domain, separated by a flexible linker region\(^{[29]}\). The C-terminal domain (CTD) contains residues required for single-stranded DNA binding, ATPase activity, and oligomerisation\(^{[29,11]}\). The N-terminal domain (NTD) appears to play a more regulatory role, containing regions responsible for the interaction of DnaB with DnaG\(^{[12]}\) and DnaA\(^{[13]}\), as well as contributing to hexamer formation\(^{[14]}\). Six monomers of DnaB assemble into a ring-shaped homohexamer\(^{[15-17]}\) which encircles single-stranded DNA (ssDNA)\(^{[18,19]}\) and unwinds duplex DNA by translocating in a 5´→3´ direction along the strand to which it is bound (thus
Figure 1.3: Initiation of replication from the *E. coli* oriC. OriC contains a number of DnaA recognition sequences, or DnaA boxes, adjacent to an AT-rich region which serves as the initial unwinding site. After recognising and binding to the DnaA boxes, DnaA assembles into a helical nucleoprotein filament which induces unwinding of the adjacent double-stranded DNA. Two hexamers of DnaB helicase are loaded onto the unwound region with the assistance of the helicase-loader DnaC, ATP and DnaA. After dissociation of DnaC, each helicase is able to begin unwinding in opposite directions.
excluding the second strand from its central channel). This process is powered by ATP hydrolysis\textsuperscript{[14]}. DnaA-mediated loading of DnaB onto ssDNA \textit{in vivo} requires it to be bound to its loading partner, DnaC, in a complex which contains up to six monomers of DnaC associated with the DnaB hexamer\textsuperscript{[40-43]}. A direct interaction between DnaA and DnaB enables DnaA to recruit DnaB to the unwound region near the origin\textsuperscript{[27,44]} (Figure 1.3). Two DnaB hexamers must be recruited – one for each nascent replication fork – and need to be oriented in opposite directions such that they proceed away from each other once replication begins\textsuperscript{[45,46]}. Recent work has shown that DnaC is also able to interact directly with DnaA, indicating that both DnaB-DnaA and DnaC-DnaA interactions may be involved in ensuring that each of the DnaB hexamers are loaded in the proper orientation\textsuperscript{[47]}. The actual mechanism by which ssDNA enters the centre of the DnaB ring is unclear, but DnaC is likely to provide assistance, potentially by acting as a scaffold which encourages opening of the ring\textsuperscript{[47]}.

Following assembly of DnaB onto ssDNA, the primase, DnaG, binds to DnaB and induces the release of DnaC from the initiation complex\textsuperscript{[48]}, relieving the inhibitory effect of DnaC on helicase activity\textsuperscript{[49]} and allowing DnaB to begin unwinding. This point represents the shift from the initiation to the elongation phase of replication, with DnaB positioned to direct the assembly of the remaining replisome components at each of the two freshly minted replication forks.

### 1.1.2 The primosome

Unwinding at the origin of replication produces two single-strands of DNA, which serve as the templates for synthesis of two new ‘daughter’ DNA strands by DNA polymerase III, such that the two chromosomes produced by a round of replication each contain one strand from the parent duplex and one newly synthesised strand\textsuperscript{[50]}. Synthesis of both strands must be
Figure 1.4. Semidiscontinuous DNA replication of the leading and lagging strands. (a) The leading strand is synthesised continuously in the same direction as the movement of the overall replication fork. The lagging strand, which is synthesised in the opposite direction to the movement of the fork, is synthesised as a discontinuous series of Okazaki fragments, each 1-2 kb long. The gaps between the fragments are subsequently sealed by the actions of DNA Pol I and DNA ligase. (b) The lagging strand is thought to form a ‘trombone loop’ which allows the directional alignment of leading and lagging strand synthesis. In both diagrams, template DNA strands are shown in black, RNA primers are in red, and newly synthesised daughter DNA is in blue.

coordinated, but the antiparallel arrangement of the template strands presents a challenge to the progressing replication fork. The polymerase is only able to synthesise new DNA in a 5´ to 3´ direction\textsuperscript{[17]}, meaning that only one of the strands (termed the leading strand) is able to be replicated continuously. The orientation of the second strand (the lagging strand) makes it necessary for the polymerase to replicate it as numerous short (1000-2000 nucleotide) pieces, known as Okazaki fragments\textsuperscript{[51]}, which are subsequently stitched into a continuous DNA strand by the actions of DNA polymerase I\textsuperscript{[52]}, and DNA ligase\textsuperscript{[53]} (Figure 1.4a). The lagging stand is thought to form a so-called ‘trombone loop’ to accommodate the opposing directionality of leading and lagging strand synthesis\textsuperscript{[54,55]} (this concept is illustrated in Figure 1.4b; see also Figure 1.2).
The polymerase at a newly formed replication fork encounters a further problem in that it is not able to begin synthesis of a DNA strand \textit{de novo}\[17,54\]. Rather, it requires a primer – a piece of RNA about 10-12 nucleotides in length\[57\] – to use as the starting point for its own synthesis\[17,56,58\]. This is relatively straightforward on the leading strand, which requires only a single primer to initiate synthesis, but the requirement is significantly more complex on the lagging strand, where each of the thousands of Okazaki fragments synthesised during a round of replication must be started with a new primer, requiring repeated primer synthesis every 1-2 seconds\[59\]. The enzyme responsible for synthesis of these primers in \textit{E. coli} is DnaG, an RNA polymerase which serves as the replicative primase\[15,58\]. DnaG is a 65 kDa protein which is monomeric in solution\[58\] and consists of three domains: an N-terminal zinc-binding domain, thought to possess regulatory functions\[58,60\]; an RNA polymerase domain, containing the catalytic active site\[61-63\]; and a C-terminal helicase-binding domain, responsible for interacting with the helicase, DnaB\[64-67\] (Figure 1.5a).

The association of DnaG with DnaB is critical to its priming action at the replication fork, and the two proteins are collectively referred to as the primosome (additional proteins are associated with the primosome formed during replication restart, but will not be discussed here). In the intact replisomal machinery, DnaB encircles the lagging strand and remains associated with the replication fork for the duration of replication, processively unwinding the duplex ahead of the fork\[59,68\]. Primase, in contrast, acts in a distributive fashion on the lagging strand as it is extruded from DnaB, priming each Okazaki fragment by repeated cycles of association with and dissociation from the template\[59\]. This cycling of the primase from one site to the next on the lagging strand is mediated by the direct but transient interaction of its C-terminus with surfaces of DnaB oriented away from the fork\[32,58,64-66,69\]. In the absence of this interaction, DnaG is an extremely slow polymerase, capable of synthesising approximately one primer every 1000 seconds\[70\]. However, its interaction with DnaB results in a very
Figure 1.5. The *E. coli* primosome. (a) Domain organisation of DnaB helicase and DnaG primase. DnaB is hexameric in solution whilst DnaG is a monomer. DnaB and DnaG interact with one another via their N-terminal and C-terminal domains, respectively. (b) Proposed model for primosome action. In this model, three monomers of DnaG associate with the DnaB hexamer (*right*). Interaction of the zinc-binding domain from one DnaG monomer with the catalytic RNA polymerase domain of an adjacent monomer may modulate primer synthesis (*left*). Figure (b) adapted from reference 77.

marked stimulation of its activity (more than 1000-fold), largely by increasing its affinity for template DNA[71]. This interaction, therefore, enables the rapid primer synthesis required *in vivo*. In turn, DnaG also stimulates the helicase activity of DnaB[14].

The transient association of primase with the replication fork provides a mechanism for regulating Okazaki fragment size – the frequency with which DnaG interacts with the replication fork, mediated by its association with DnaB, determines the frequency of primer synthesis on the lagging strand, which in turn governs Okazaki fragment length[72-74]. Several lines of evidence suggest that three DnaG molecules bind to the DnaB hexamer[66,75,76], which may add an additional level of regulation: interactions between adjacent primase molecules have been proposed to modulate both the initiation of primer synthesis and to restrict the
length of the primers synthesised to their physiological length of 10-15 nucleotides[^60,77]. Thus, in this model, DnaB serves as a localisation platform for the multiple primases needed for optimal primer synthesis (Figure 1.5b). The close coordination between DnaB and DnaG required for productive priming is also a likely mechanism for ensuring that the template is not primed indiscriminately, and that primase action is limited to the replication fork[^77].

The interaction of multiple primase molecules with DnaB may also serve additional regulatory roles. Single-molecule experiments have shown that the cooperative interaction of two or three DnaG molecules with DnaB results in halting of leading strand synthesis[^78]. Given the significantly higher complexity of lagging strand synthesis relative to that on the leading strand, this could provide a mechanism for coordinating the rates of leading and lagging strand synthesis (preventing the leading strand from ‘outrunning’ the lagging strand[^78]).

A more complete picture of the operation of the *E. coli* primosome awaits atomic resolution structures of the full-length DnaB and DnaG proteins from *E. coli*, and of the two in complex. Thus far, structures for the C-terminal helicase-binding domain[^66,67] and catalytic RNA polymerase domain[^61-63] of DnaG, and of the N-terminal domain of DnaB[^79,80] have been solved—intriguingly, the C-terminal domain from DnaG and the N-terminal domain from DnaB adopt extremely similar folds[^87]. The DnaB-DnaG complex from the thermophilic bacterium *Geobacillus stearothermophilus* is significantly more stable at room temperature than the transient *E. coli* complex (presumably due to the need for it to form at higher temperatures), and the crystal structure of the helicase-binding domain (HBD) of *G. stearothermophilus* primase bound to the full-length hexameric helicase has been solved[^74], providing further clues to the arrangement of the complex. In the structure, three primase HBDs pack against the outside of the ring formed by the N-terminal domains of DnaB, locking the conformation of DnaB, which may explain the contribution of DnaG binding to helicase activity[^76]. Modelling of the complex
with full-length DnaG based on the positioning of the HBDs in the structure is also consistent with the suggestion that adjacent DnaG molecules are able to interact with one another in trans [76].

1.1.3 The DNA polymerase III holoenzyme

The multisubunit Pol III HE assembly functions in a highly coordinated manner, but can be broken down into three functional subassemblies: the catalytic core, consisting of the α, ε and θ subunits [81], which is responsible for DNA synthesis and proofreading; the β₂ sliding clamp, responsible for conferring processivity to the core and allowing it to replicate long stretches of DNA without dissociation [82]; and the seven-subunit clamp loader complex (τγδδ’χψ or τδδ’χψ), responsible for loading β₂ onto DNA, as well as for a host of organisational functions [3]. Two (or potentially three; see later) cores are present per holoenzyme, allowing simultaneous replication of the leading and lagging strands, and are tethered together by the τ subunits of the clamp loader complex [81,84] (Figure 1.2).

1.1.3.1 The Pol III core

The three subunits of the Pol III core, α, ε and θ, form a complex in which ε binds to both α and θ, but α and θ do not contact each other [85]. The α subunit (the product of the dnaE gene) contains the polymerase activity of the enzyme [86,87], while ε (encoded by dnaQ) provides the polymerase with a proofreading function by acting as a 3’→5’ exonuclease [88]. The function of the third core subunit, θ, the product of holE [85], is unclear – it is not essential to cell viability [89], though it has been shown to slightly stimulate the activity of ε [90].
The α subunit

Based on a combination of sequence, structural and biochemical analyses, DNA polymerases have been divided into six families, designated A, B, C, D, X and Y [91,92]. The α subunit of Pol III, a 1160 amino acid, 130 kDa protein [93], is a member of the C family, which also includes the replicative polymerases from other bacteria [91,92]. Alone, α incorporates nucleotides, in a 5’→3’ direction, at a rate of only ~8 nt/s [87,94]. However, its role within the assembled replicase relies heavily upon its interactions with a number of other proteins; as well as its contact with ε, it interacts with the β2 sliding clamp [82,95] and the τ subunit of the clamp loader complex [96].

Until relatively recently, structural information for the members of the C family of polymerases was lacking. The crystal structure of a truncated form of E. coli α (residues 1-917) [97] and the structures of full-length α from Thermus aquaticus (which shares 39% sequence identity with E. coli α and possesses a highly similar structure) alone [98] and bound to primer-template DNA and an incoming nucleotide [99] have provided many clues to the mechanism of α as a polymerase, and helped to pin down the domain organisation of the protein in detail (Figure 1.6a). The crystal structures of polymerases from other families had shown that, despite significant differences in structure, their overall shape could be compared to the shape of a cupped right hand, with domains corresponding to the ‘palm’, ‘fingers’ and ‘thumb’ of the hand [92]. The hand analogy can be extended to the mechanism of action of polymerases, which are able to clasp their substrate, DNA, and position it effectively for polymerisation to occur. The palm domain contains the active site, providing residues which coordinate the two metal ions (generally Mg2+) required for catalysis, whilst the fingers and thumb domains play roles in interacting with the incoming nucleotide, and positioning the newly produced duplex DNA, respectively [92].
Figure 1.6. Domain organisation and structure of the α subunit. (a) Domain organisation of *E. coli* α. Numbers above the figure indicate domain boundaries. Figure adapted from reference 3. (b) Crystal structures of the full-length α subunit from *Thermus aquaticus*, in the absence (left) and presence (right) of DNA. Domains are coloured as per the schematic in (a). The active site aspartate residues within the palm domain are coloured cyan. The internal β₂-binding sequence (corresponding to residues 920-924 in *E. coli*, and residues 975-979 in *T. aquaticus*) is highlighted in pale yellow. The clear differences between the structures in the absence and presence of DNA highlight the conformational changes in α which occur upon DNA-binding, particularly in the β₂-binding and OB-fold domains. Both figures were produced using PyMol[98] and the PDB coordinates 2HPI (apo α)[99] and 3E0D (α-DNA)[99]. (c) More detailed view of the conformational changes induced by DNA binding.
Only the domains which make significant movements are shown. Domains are coloured as in (a) and (b); domains positions from the unliganded structure are shown in grey. The β2-binding domain rotates by about 20°, bringing it into contact with the DNA. The fingers and thumb domains also make smaller movements in towards the DNA. Note the orientation of this structure relative to those in (b) (rotated 180° both horizontally and vertically). Part (c) adapted from reference 99.

The structures of α reveal that, whilst adopting a unique fold, it also contains palm, thumb and fingers domains\footnote{\cite{97-99}} (Figure 1.6b), with the catalytic palm domain exhibiting a similar arrangement to the palm domain of the X family polymerases\footnote{\cite{97,98}}. The highly conserved aspartate residues which form the active site (Asp401, Asp403 and Asp555 in \textit{E. coli} \footnote{\cite{101}}) are located close together within the palm domain, in a deep cleft formed by the fingers and thumb domains\footnote{\cite{97-99}} (Figure 1.6b).

The α subunit also includes a number of accessory domains important to its function. The barrel-shaped PHP domain, located at the N-terminus (residues 1-272), packs against the thumb and palm domains\footnote{\cite{97}}. Whilst this domain was originally named for its resemblance to histidinol phosphatase (PHP – polymerase and histidinol phosphatase) and predicted to possess a role in pyrophosphate hydrolysis\footnote{\cite{102}}, such a role has not been found\footnote{\cite{97}}. Deletion studies show that the PHP domain is required for polymerase activity\footnote{\cite{103}}. In addition, it is through this domain that α interacts with the proofreading subunit of the core, ε\footnote{\cite{104,105}}. The PHP domain of α from \textit{Thermus thermophilus} (whose replication machinery closely resembles that from \textit{E. coli} \footnote{\cite{106}}) has been found to contain an intrinsic Zn$^{2+}$-dependent 3′→5′ exonuclease activity\footnote{\cite{107}}. Whilst this activity does not substitute for the proofreading exonuclease activity of the ε subunit, it is possible that, given their proximity, the PHP domain of α and the ε subunit may work together in a complementary fashion\footnote{\cite{107}}.
Binding studies using truncated mutants of α provided the first evidence that binding to both β₂ and τ is accomplished though the C-terminal region of α[^95,108]. The crystal structures have allowed this region to be further divided into a β₂-binding domain (residues 778-964), an oligonucleotide binding (OB)-fold domain (residues 964-1078) and, at the extreme C-terminus, a τ-binding domain (residues 1078-1160)[^98] (Figure 1.6a). The interaction between α and the β₂ sliding clamp is an extremely important one: alone, the core polymerase is only able to synthesise 10-15 nucleotides continuously before dissociating from the template[^109], but its association with the β₂ clamp converts it into an remarkably processive enzyme, capable of replicating for tens of thousands of nucleotides without dissociation[^109].

The site in α responsible for its association with β₂ has been the subject of some contention: an ‘internal’ β₂-binding site (residues 920-924) was originally identified by sequence analysis[^110]. However, a second site, located at the extreme C-terminus (residues 1154-1160) was also identified as being able to bind to β₂[^111,112]. It was subsequently shown that only the internal site was required for processive replication[^113], and that the extreme C-terminus of α played a much more significant role in binding to the τ subunit of the clamp loader complex[^113]. The β₂-binding domain also contains a helix-hairpin-helix (HhH) motif (residues 836-854)[^98], a common non-sequence-specific DNA binding motif[^114].

The structure of a ternary complex of Taq α with primer-template DNA and an incoming nucleotide (dATP) revealed that several domains of α undergo significant conformational changes upon substrate binding[^99] (Figure 1.6 b, c). The thumb and β₂-binding domains both make significant movements ‘in’ towards the DNA, enabling them to establish a pincer-like grip on the backbone of the duplex: two α-helices from the thumb domain engage with the minor groove of the DNA, whilst the entire β₂-binding domain rotates by about 20°, placing its HhH motif in close proximity to the DNA[^99]. The fingers domain also moves, forming one
side of a pocket (confined on its other sides by the palm domain and the 3’ end of the primer) which acts to position the incoming nucleotide correctly for incorporation\[99\]. Since OB-folds are classic single-stranded DNA recognition motifs\[115\], its identification at the C-terminus suggested a role for this region in binding to the single-stranded template DNA downstream of the active site. The structure of the ternary complex revealed how this is achieved: upon binding to the primer-template, the OB-fold moves by about 10°, bringing it into a position close to the template strand, which appears to be sharply bent as it emerges from the active site of the polymerase\[99\] (Figure 1.6b).

**The ε subunit**

The α subunit, like other replicative polymerases, is able to copy DNA with high fidelity, only incorporating an incorrect nucleotide, on average, every 10^4-10^6 nucleotides\[116,117\]. However, given that the *E. coli* genome contains ~4.6 million nucleotides\[117\], this level of fidelity is insufficient to prevent mistakes from being made during replication, and an editing mechanism is required. The exonuclease action of the proofreading subunit, ε, allows more than 90% of misincorporated nucleotides to be excised immediately, with the result that the intact Pol III holoenzyme has an average error rate of just 10^{-7}-10^{-8} per base pair\[116-118\] (this rate is decreased even further, to ~10^{-10} per base pair, by post-replication mismatch repair).

The ε subunit was first identified as the replicative proofreader following characterisation of *E. coli* strains with mutations in the gene encoding ε, *dnaQ*\[119,120\]. These strains exhibited greatly increased mutation rates\[119\], which were attributed to defective proofreading during replication\[120\]. Subsequently, ε was found to contain the 3’→5’ exonuclease activity of Pol III\[88\], a finding which was initially surprising given that the proofreading activities of previously
characterised polymerases (e.g. \(E.\ coli\) Pol I and the replicative polymerase from phage T4) were found in the same polypeptide chain as their polymerase activities\textsuperscript{17}.

The \(\epsilon\) subunit is a 27 kDa, 243 amino acid protein with two functional domains\textsuperscript{122} (see Figure 6.7a). The N-terminal domain, consisting of the first 180 amino acids, contains the catalytic exonuclease activity of the protein, and is responsible for the interaction of \(\epsilon\) with the \(\Theta\) subunit\textsuperscript{122}. The C-terminal domain (residues 202-243) mediates the interaction of \(\epsilon\) with the PHP domain of \(\alpha\)\textsuperscript{122,123}. The domains are connected by a ‘Q-linker’ (residues 180-201)\textsuperscript{124} – a member of a class of flexible, glutamine-rich sequences which are frequently used as hinge regions between the functional domains of proteins\textsuperscript{125}.

The crystal structure of the N-terminal catalytic domain (\(\epsilon186\)) in complex with two Mn\(^{2+}\) ions and a molecule of thymidine-5’-monophosphate (TMP, representing the product of the

**Figure 1.7. Structures of the \(\sigma\) proofreading subunit.** (a) Crystal structure of the N-terminal domain of \(\epsilon\) (\(\epsilon186\)) bound to TMP (pink). Two Mn\(^{2+}\) ions (grey spheres) are situated in the active site of the enzyme. The protein is coloured from blue at the N-terminus to red at the C-terminus. Figure produced using PyMol\textsuperscript{100} and PDB coordinates 1J53\textsuperscript{121}. (b) Crystal structure of a fusion protein containing the \(\epsilon\) C-terminus (residues 209-243, in green) fused to the PHP domain of \(\alpha\) (residues 2-270, in blue). The flexible linker connecting the two domains is shown in yellow. Part (b) reproduced from reference 105.
exonuclease reaction) has been solved\[^{121}\] and reveals that it shares a similar topology with the exonuclease domains from other polymerases: a five-stranded $\beta$-sheet forms the core of the domain, and is surrounded by several helices\[^{121}\] (Figure 1.7a). The active site contains four carboxylate residues (Asp12, Glu14, Asp103 and Asp167) along with a catalytically important histidine (His162), and two divalent metal ions ($\text{Mg}^{2+}$ or $\text{Mn}^{2+}$) which are coordinated by the carboxylates\[^{121}\]. The position of TMP in the structure suggests that the phosphate group of the nucleotide to be excised from the 3' end of the growing nucleotide chain is bound in the active site by interactions with both metal ions, whilst other residues help to position the deoxyribose and thymine sections of the TMP molecule\[^{121}\]. Hydrolysis involves nucleophilic attack of the phosphate group by a water molecule coordinated by one of the metal ions\[^{121,126}\]. This two-metal mechanism of catalysis is common to other exonucleases, including the exonuclease domains of DNA Pol I\[^{127}\] and T4 DNA polymerase\[^{128}\].

The flexible nature of the C-terminal domain of $\epsilon$\[^{124}\] has hampered efforts to probe its structure. Recently, the structure of a fusion protein containing the PHP domain of $\alpha$ (residues 2-270) linked to the C-terminus of $\epsilon$ (residues 209-243) via a flexible 9 amino acid linker was solved in our laboratory\[^{105}\], providing some clues to the nature of the $\alpha$-$\epsilon$ interface. In the structure, residues 219-236 of $\epsilon$ form an $\alpha$-helix which lies along one side of the PHP domain from $\alpha$, whilst residues 209-218 are arranged in a random coil which tracks around the PHP domain along a groove in its surface\[^{105}\] (Figure 1.7b). In addition, NMR studies have shown that the linker region between the two domains of $\epsilon$ remains flexible even when $\epsilon$ is bound to $\alpha$ and $\theta$\[^{124}\], suggesting that the N-terminal exonuclease domain may be able to move quite freely even when it is contained within the Pol III core complex.

Despite the structural information for both $\alpha$ and $\epsilon$ now available, the details of how the Pol III holoenzyme switches between polymerisation and proofreading modes, with the primer
terminus occupying the active sites of α and ε respectively, remains unclear. It is known that the interaction of ε with α markedly stimulates its exonuclease activity, which is probably the result of α assisting with recognising when an error has been made, as well as the interaction positioning ε correctly for excision to occur. The flexibility of the exonuclease domain of ε may play an important role in the mechanics of the switch from polymerisation to proofreading activity.

Interestingly, the ε subunit has been observed to make a substantial contribution to the processivity of replication in vitro – if all components of the holoenzyme except ε are present, processivity is markedly reduced. This observation seems to suggest that that ε, in addition to its proofreading role, may also possess an important, though less well-defined, structural role. The work described in Chapters 5-7 of this thesis forms part of an investigation in our laboratory which has discovered such a role for ε; further details will be discussed in depth in these chapters.

**The θ subunit**

The θ subunit, as already noted, is the most mysterious of the Pol III holoenzyme subunits, and its function in replication is essentially unknown, though it is known to play a role in stabilising the structure of ε. Expression of ε in the presence of θ allows production of a soluble ε-θ complex, and protects ε from heat-induced denaturation. This stabilisation of the ε structure by θ has also been noted to produce a functional effect, slightly increasing the proofreading activity of ε. The θ subunit itself is a 76 amino acid, 8.8 kDa protein, consisting of a compact bundle of three α-helices. Structures of the N-terminal domain of ε in complex with θ show that the θ-binding site on ε is a considerable distance (≈15 Å) away from the exonuclease active site, suggesting that the stabilising effect of θ upon ε is
exerted indirectly, via other parts of the ε structure. The helical motifs present in θ also hint at the possibility of a DNA-binding function, though no such interaction has been detected\(^{112}\).

1.1.3.2 The β\(_2\) sliding clamp

The β subunit was discovered to be the primary contributor to the extremely high processivity of the holoenzyme in early functional assays which also demonstrated that this contribution is a consequence of an interaction between β and the Pol III core\(^ {113}\). Subsequently, β, which is the 41 kDa product of the dnaN gene\(^ {116}\), was shown to exist as a tightly associated dimer, β\(_2\)\(^ {117}\). The crystal structure of the β\(_2\) dimer, when solved, was illuminating – the dimer forms a ring with a central hole large enough to accommodate double-stranded DNA\(^ {117-119}\) (Figure 5.1). By encircling DNA but retaining the ability to slide freely upon it\(^ {82}\), β\(_2\) creates a topological link between the DNA and those proteins with which it associates. This link enables it to confer processivity to the core polymerase – it acts as a tether, ensuring that any dissociation of the core from the template remains transient.

As well as its interaction with the Pol III core, β\(_2\) binds to all four of the other prokaryotic polymerases, Pol I\(^ {140}\), Pol II\(^ {141,142}\), Pol IV\(^ {143,144}\) and Pol V\(^ {145}\), and to other proteins including the δ subunit of the clamp loader complex\(^ {146}\), DNA ligase\(^ {140}\), MutS\(^ {140}\) and MutL\(^ {147}\). These interactions are all mediated via a common hydrophobic binding pocket on β\(_2\)\(^ {111,148-151}\), and are indicative of its role as a hub for organising proteins involved in repair (MutS and MutL), translesion synthesis (Pol II, IV and V) and Okazaki fragment processing (Pol I, ligase), in addition to its role as a processivity factor in replication.

The newly uncovered interaction between β\(_2\) and the ε subunit of the Pol III core will be the subject of Chapters 5-7 of this thesis. More details regarding the structure and interactions of β\(_2\) can be found in the introductions to these chapters (see especially section 5.1).
1.1.3.3 The clamp loader complex

The β clamp, being a closed ring, is not capable of loading itself onto DNA: this is accomplished by the clamp loader complex, a subassembly of the Pol III holoenzyme which, with the assistance of ATP, assembles β around DNA[82] (Figure 1.8a). The clamp loader complex from E. coli, also referred to as the DnaX complex, is a heptameric assembly containing the products of five genes: dnaX (γ and τ)[152]; holA (δ)[153]; holB (δ′)[153]; holC (χ)[154] and holD (ψ)[154], with an overall stoichiometry of τγδδ′χψ[155], or, possibly, τδδ′χψ[156]. The dnaX gene directs the production of two proteins, γ and τ. The τ subunit is the 71 kDa full-length product of the dnaX gene, and consists of five domains (I-V)[157]. The 47 kDa γ subunit is a truncated version of τ, containing domains I-III of τ, and is the result of a programmed ribosomal frameshift during translation of dnaX, resulting in early termination[158,159] (Figure 1.9). Structurally, γ, τ (domains I-III), δ and δ′ each adopt the same three domain fold, common to members of the AAA+ family of ATPases[148,160-163], but DNA-dependent ATPase activity within the clamp loader complex is confined to the DnaX subunits, γ and τ[164].

An understanding of the mechanism of the clamp loader has been assisted by the determination of some of the aspects of its structure. In vitro, complexes with varying ratios of γ and τ can be assembled, all of which are active in clamp loading[164]. It is the γδδ′ version for which crystal structures have been obtained, both in the absence[162] and presence[165] of DNA. However, the presence of τ, with its unique C-terminal region, is essential for replisome organisation in the physiological complex, as will be described later. The structures reveal that the γδδ′ portion of the clamp loader complex forms a heteropentameric ring, mediated by interactions between domain III from each subunit which form a ‘collar’ supporting the less tightly associated ATPase regions[162] (Figure 1.8 a,b). Within the ring, δ and δ′ are positioned adjacent to each
Figure 1.8. The clamp loader complex. (a) Schematic diagram of the proposed mechanism of action of the clamp loader complex. ATP binding induces a conformational change in the complex which enables it to bind to the open \( \beta \), clamp and to primer-template DNA. ATP hydrolysis then triggers the release of the complex from both DNA and \( \beta \), leaving the clamp assembled on DNA. Part (a) reproduced from reference 165. (b) Crystal structure of the \( \gamma \delta \delta' \) complex. Subunits are coloured as in (a). The front view of the complex (left) emphasizes the tight interactions formed between the collar domains (domain 3) of each subunit, relative to the more loosely associated ATPase domains. The top view of the complex (right) highlights the ring-shaped nature of the \( \gamma \delta \delta' \) complex. The region of \( \delta \) implicated in binding the \( \beta \) clamp is indicated in yellow. Figures produced using PyMol\textsuperscript{160} and PDB coordinates 1JR3\textsuperscript{162}. (c) Crystal structure of the \( \gamma \delta \delta' \) complex bound to primer-template DNA, showing the spiral shape adopted by the complex. Part (c) reproduced from reference 165. (d) Crystal structure of the complex between domain 1 of \( \delta \) (residues 1-140) and a single monomer of \( \beta \). The region of \( \delta \) which contacts \( \beta \) is shown in yellow. Two residues important to binding, L73 and F74, are represented as sticks, and are shown to protrude into a cleft on the surface of \( \beta \). Figure constructed using PyMol and PDB coordinates 1JQL\textsuperscript{148}. 
Figure 1.9. The dnaX gene produces two protein products. The γ subunit, the product of a frameshift mutation, consists of 431 residues and contains three domains, which provide AAA+ ATPase activity (domains I and II) and allow oligomerisation (domain III). The full-length, 643 amino acid product of the gene, τ, shares domains I-III with γ but includes an additional two C-terminal domains, which allow binding to DnaB and the α subunit (domains IV and V, respectively). The positions of domain boundaries are indicated by the numbers above the figure. Adapted from reference 2.

other, with the three γ subunits closing the ring on the opposing side. Crosslinking studies suggest that, in the native complex, τ replaces the γ subunits at the positions marked B and C in Figure 1.8, such that the sole γ subunit within the complex is positioned adjacent to δ′.

The three ATP molecules which bind to the clamp loader complex bind within the subunit interfaces to a conserved arginine residue. This binding results in structural changes within the complex, allowing ATP binding and hydrolysis to be coupled to loading of β2. The δ subunit is responsible for binding directly to the β2 clamp and trapping it in an ‘open’ conformation, in which one of the dimer interfaces is disrupted. In this interaction, a conserved peptide motif from domain I of δ is inserted into the common hydrophobic protein binding pocket on the C-terminal face of β2, which is utilised by its many binding partners (Figure 1.8d; see also section 5.1). However, whilst binding between isolated δ and β2 is able to occur in the absence of ATP, the intact clamp loader complex requires ATP (binding, but not hydrolysis) in order to bind to β2. The presumed reason for this effect is that the δ...
subunit is occluded within the clamp loader complex (most likely by the δ′ subunit\textsuperscript{[162,168]}), and that ATP binding induces a conformational change in the complex which exposes the β\textsubscript{2}-binding motif of δ, allowing it to interact with β\textsubscript{2}\textsuperscript{[146,162,168,169]}. More recent work suggests that this mechanism may in fact be more complex, and that ATP binding may cause conformational changes which involve the clamp loader as a whole adopting a conformation with an overall shape complementary to the surface of the β\textsubscript{1} clamp\textsuperscript{[170]}.

The ATP-bound form of the clamp loader complex also has a high affinity for primed template DNA, resulting in the formation of a ternary clamp loader-β\textsubscript{2}-DNA complex\textsuperscript{[169,171]}. The structure of the eukaryotic clamp loader from yeast, replication factor-C (RFC), in complex with ATPγS (a poorly hydrolysable ATP analogue) and its cognate sliding clamp, proliferating cell nuclear antigen (PCNA), suggested the basis for recognition of primed DNA during this step\textsuperscript{[172]}. In this structure, the ATPγS-bound RFC adopts a spiral conformation which matches the pitch of the DNA double-helix. The RFC-PCNA complex therefore possesses a means for distinguishing between the 3′ and 5′ ends of primed DNA, and, in a proposed model, would thread onto correctly primed DNA, much like a screw-cap\textsuperscript{[172]}. This mechanism would position DNA within the central channel of the clamp, as well as ensure it was orientated correctly for elongation to occur. The structure of the \textit{E. coli} γδδ′ complex bound to primer-template DNA reveals that it too adopts a spiral conformation when bound to DNA, and supports a screw-cap type model for DNA recognition\textsuperscript{[165]} (Figure 1.8c).

In the final stage of clamp loading, ATP hydrolysis by the DnaX subunits (triggered by DNA binding\textsuperscript{[173]}) results in another conformational change in the clamp loader which lowers its affinity for both DNA and β\textsubscript{2}, resulting in its release, and closure of β\textsubscript{2} around DNA\textsuperscript{[168,174-176]} (Figure 1.8a). Interactions of β\textsubscript{2} with the ssDNA template may also play a role in displacing the clamp loader from the clamp\textsuperscript{[119]}. Release of the clamp loader is required in order for the
polymerase to take up its position on the newly loaded $\beta_2$ clamp, as $\delta$ and $\alpha$ compete for binding to the same face of $\beta_2$\[177\]. It appears that ATP hydrolysis, and thus, completion of clamp loading, is triggered preferentially when the clamp loader binds to the 3’ end of a primer-template junction in the correct orientation for elongation to occur, providing a further means of ensuring clamps are assembled onto the correct sites in the DNA template\[175,178\].

As well as acting in a coordinated manner to load the $\beta_2$ clamp, certain subunits of the clamp loader subassembly possess additional functions within the replisome. Whilst the three N-terminal domains of the $\tau$ subunit are identical to the $\gamma$ subunit, its unique C-terminus, containing an additional two domains\[157\] allows it to take on a role as a central replisome organiser and stabiliser\[18\] (Figure 1.9). The C-terminal domain of $\tau$, domain V, interacts with the $\alpha$ subunit of the Pol III core\[108,157,179,180\]. The two copies of $\tau$ in the holoenzyme therefore essentially dimerize the leading and lagging strand Pol III cores (Figure 1.2), tethering the lagging strand polymerase to the advancing leading strand machinery and enabling it to cycle rapidly from one Okazaki fragment to the next with the required speed (an Okazaki fragment is synthesised every 1-2 seconds)\[38,83,181\]. The interaction of $\tau$ with $\alpha$ means that $\tau$ also serves as a bridge between the polymerase core and the $\chi$ and $\psi$ subunits (Figure 1.2), thereby mediating the $\chi$-SSB interaction (described below)\[182,183\]. Mutational studies have shown that $\tau$ contributes to replication fidelity, likely as a consequence of the $\alpha$-$\tau$ interaction, although the specific mechanism for this contribution remains unclear\[184\].

The association between $\tau$ and $\alpha$ has recently been demonstrated to serve another purpose – the clamp loader, if it contains $\tau$, can chaperone the Pol III core onto newly assembled $\beta_2$ clamps\[185\]. Essentially, the $\alpha$-$\tau$ interaction increases the local concentration of Pol III core, greatly increasing the efficiency of initiation complex formation. In the proposed mechanism
for this pathway, the clamp loader remains bound to the template DNA adjacent to the clamp until it is replaced by the Pol III core, thus preventing the competitive binding of SSB, which inhibits formation of initiation complexes in the absence of $\tau$. Recent studies have suggested that the replisome may in fact contain three copies of $\tau$, and therefore three polymerase cores, with the third polymerase kept in ‘reserve’ for use in cases of polymerase stalling or inactivation, or to enable more efficient cycling between Okazaki fragments on the lagging strand\textsuperscript{[156,186]}. Although such trimeric holoenzymes are functional in \textit{in vitro} assays\textsuperscript{[186]}, and mutant \textit{E. coli} cells containing only the $\tau$ product of \textit{dnaX} are viable\textsuperscript{[152]}, the relevance of a $\tau$-containing holoenzyme requires further investigation.

The $\tau$ subunit also interacts with the helicase, DnaB, \textit{via} domain IV, its other $\tau$-specific domain\textsuperscript{[187,188]}. This interaction was originally thought to result in a large stimulation of the speed of an otherwise sluggish helicase, allowing it to progress at rates consistent with the speed of the native replisome\textsuperscript{[188]}. However, more recent evidence suggests that DnaB unwinds duplex DNA much more rapidly than previously thought, and that the observed effects of $\tau$ binding result mostly from DnaB being more stably attached to the replication fork in its presence – \textit{i.e.} by increasing the processivity of the helicase\textsuperscript{[68,78]}. In any case, the $\tau$-DnaB contact is important for coordinating the activities of the holoenzyme and the primosome, and likely also serves a reciprocal function as an additional source of stability for the polymerase core on the DNA template (beyond the $\beta_2$-core interaction)\textsuperscript{[18]}. Finally, $\tau$ appears to contribute to the processivity of replication by preventing the untimely unloading of $\beta_2$ clamps. In the absence of $\tau$, exogenous $\gamma\delta\delta'$ complex can competitively bind to $\beta_2$ during replication, resulting in premature unloading\textsuperscript{[168]}. This, in turn, results in distributive leading strand synthesis which is sensitive to the concentrations of both $\beta$, and $\gamma\delta\delta'$\textsuperscript{[189]}. The presence of $\tau$ within the clamp loader complex eliminates this effect by
protecting $\beta_2$ from being unloaded, although the precise basis for this protection remains to be determined\[^{189}\].

The remaining accessory subunits of the clamp loader complex, $\chi$ and $\psi$, bind tightly to one another\[^{190}\], forming an elongated heterodimer for which a crystal structure has been solved\[^{191}\]. The 26 residues at the N-terminus of $\psi$ are disordered\[^{191,192}\] and bind to domain III of $\gamma$, serving as the sole contact between the $\chi\psi$ dimer and the other subunits of the clamp loader complex\[^{191}\] (Figure 1.2b). Whilst $\chi$ and $\psi$ do not participate directly in clamp loading, they do contribute to the architecture and function of the replisome in a number of ways. The presence of $\psi$ strengthens the interactions between $\gamma$ and $\delta'$ and between $\tau$ and $\delta$\[^{190}\], thus facilitating the assembly of the clamp loader complex. The structure of residues 2-28 from $\psi$ bound to the $\gamma_3\delta\delta'$ clamp loader complex in the presence of DNA\[^{165}\] reveals that the $\psi$ peptide makes contacts with the ‘collar’ domains of each of the three $\gamma$ subunits, and that this interaction appears to promote the asymmetric spiral conformation of the complex which binds with high affinity to DNA\[^{165}\]. This structural information provides an explanation for biochemical data showing that $\psi$ stimulates the interaction of the clamp loader complex with DNA by stabilising a conformational state of the clamp loader with high DNA affinity\[^{194}\].

The $\chi$ subunit also interacts directly with single-stranded DNA-binding protein (SSB)\[^{182,183,195-197}\] which coats the single-stranded DNA generated during replication of the lagging strand. This interaction, in which the C-terminal tail of SSB binds to a conserved surface on $\chi$\[^{195}\] is responsible for overcoming the salt-dependent defects in replication elongation which are observed in the absence of $\chi$\[^{182,183}\], or in the presence of $\chi$ mutants with an impaired ability to interact with SSB\[^{195}\]. An interaction between $\chi$ and SSB is also required for helicase-independent strand displacement synthesis in vitro\[^{198}\]. These observations are attributed to the
fact that the $\chi$-SSB contact serves as a physical link between the polymerase and the SSB-coated template DNA (mediated by an $\alpha$-$\tau$-$\psi$-$\chi$ bridge; see Figure 6.3), thereby providing stability to the replicase which becomes important in the presence of elevated salt, where other interactions are weakened\cite{182,183,195,197,198}. The $\chi$ subunit has also been proposed to play a role in removing DnaG primase from its RNA primers by competitively displacing DnaG from SSB\cite{199}, allowing assembly of the clamp and polymerase. However, recent evidence shows that SSB-binding deficient $\chi$ mutants do not significantly compromise lagging strand replication\cite{195}, suggesting that the primary role of the $\chi$-SSB interaction is in promoting replisome stability during processive replication, rather than in primer handoff.

1.1.4 Dynamics of DNA replication

It is helpful at this point to visualise the overall sequence of events which occurs as the *E. coli* replication fork progresses (Figure 1.10). As described in section 1.1.2, the interaction between the helicase and the primase is responsible for regulating the frequency with which primers are laid down on the lagging strand (every 1-2 s), and therefore the length of Okazaki fragments\cite{72-74}. The clamp loader complex loads a $\beta_2$ clamp onto each new primer, allowing the Pol III core to associate and begin DNA synthesis (Figure 1.10). Upon completion of each Okazaki fragment, the polymerase needs to dissociate from $\beta_2$ and recycle to the newly loaded clamp on the next primer\cite{200}. This cycling is assisted by the $\tau$ subunit of the clamp loader complex, which chaperones the core to the next available $\beta_2$ clamp\cite{185}. Despite the very strong interaction between $\alpha$ and $\beta_2$, there is a very short time available for cycling (< 0.1 s)\cite{201}. This indicates that some form of ‘processivity switch’ which increases the rate of dissociation of the polymerase, must operate. The precise nature of this switch remains a subject of contention\cite{3,112,201-204}, and will not be discussed here. As the polymerase cycles from one primer to the next, $\beta_2$ clamps remain on the lagging strand, where they accumulate, allowing
Figure 1.10. Dynamics at the *E. coli* replication fork. (a) DnaG associates periodically with DnaB to synthesise short RNA primers on the lagging strand. Continuous unwinding of the parental DNA as the lagging strand polymerase synthesises a downstream Okazaki fragment produces a loop of single-stranded DNA, which is bound by SSB. (b) DnaG dissociates from the newly synthesised primer and is replaced by the $\beta_2$ clamp, which is loaded with the assistance of the clamp loader complex. (c and d) The $\tau$ subunit of the clamp loader complex binds to the polymerase core and chaperones it from the completed Okazaki fragment to the newly loaded $\beta_2$ clamp (leaving the original $\beta_2$ clamp associated with the previous Okazaki fragment), where it begins synthesis of a new Okazaki fragment. This cycle is repeated $\sim$2000 times per replication fork in each round of replication. Reproduced from reference 2.

them to mediate later events involved in Okazaki fragment processing\[38,140,200\]. Given that there are $\sim$300 copies of $\beta_2$ per cell\[116\] and $\sim$4000 Okazaki fragments are synthesised in a cycle of *E. coli* replication\[205\], clamps must also be reused. The clamp loader complex and the
free δ subunit are able to unload \( \beta_2 \)\(^{[177,206]} \), and given that DNA synthesis has been observed in the absence of free \( \beta_2 \) clamps\(^{[205]} \), a direct recycling mechanism must exist.

1.1.5 Single-stranded DNA-binding protein (SSB)

The single-stranded DNA-binding protein (SSB) in *E. coli* was first identified in 1972\(^{[207]} \), and was initially named ‘DNA-unwinding protein’ for its ability to lower the melting temperature of double-stranded DNA \(^{[207]} \), before being designated SSB\(^{[208]} \), in accordance with the name of the gene from which it is expressed, *ssb*\(^{[208,210]} \). Several early studies established that SSB was a tetramer of identical subunits which bound preferentially and with high affinity to single-stranded DNA (and much more weakly to double-stranded DNA), and that this binding occurred cooperatively and with little or no sequence specificity\(^{[207,211-217]} \).

The early characterisation of SSB mutants established that SSB is essential for cell survival\(^{[208,209]} \), and it has since been shown to be required for all aspects of DNA metabolism – replication, recombination and repair\(^{[218,219]} \). One aspect of this requirement stems from the role of SSB in stabilising single-stranded DNA (ssDNA). In its double-stranded form, DNA possesses a degree of natural protection against attack from nucleases and other threats to its integrity. However, ssDNA – essential as an intermediate in DNA metabolism – is more vulnerable, and one important function of SSB is as a protective agent against such threats\(^{[220]} \). In addition, SSB melts out regions of secondary structure, such as hairpins or loops, which are liable to form in ssDNA by self-association and act as barriers to Pol III progression or to the action of other proteins\(^{[207,212,211,221]} \). In a related function, it acts to prevent ssDNA strands from reannealing as they are produced by the action of helicases: this feature is thought to be the basis of the observed stimulation of DnaB helicase by SSB\(^{[214]} \) as well as contributing to its stimulation of PriA, the helicase involved in replication restart at collapsed replication forks\(^{[223]} \). Finally, SSB has also been observed to promote renaturation of denatured DNA
under certain solution conditions\textsuperscript{[218]}, although this aspect of its activity and its potential significance are not well understood.

Despite the obvious importance of its role as an ssDNA stabiliser, SSB is more than a passive protective coating. It is an intriguing protein which plays a central role in mediating DNA metabolism, acting as a platform for the assembly and, potentially, organisation of numerous replication, recombination and repair proteins. The importance of single-stranded DNA binding proteins is reflected by the fact that they are found, in various forms, as essential proteins in organisms from all three domains of life, as well as in mitochondria and viruses\textsuperscript{[218,221-227]}. Though sharing certain common features, most notably an oligonucleotide-binding (OB) fold\textsuperscript{[115,223,228]} with which to bind ssDNA, SSBs from the various domains have evolved to possess distinct quaternary structures. The majority of bacterial SSBs, for which \textit{E. coli} SSB serves as the archetype, are homotetramers\textsuperscript{[229-233]} as is the human mitochondrial SSB, which shares a high degree of structural similarity with \textit{E. coli} SSB\textsuperscript{[237,234]}. Exceptions to the tetrameric bacterial SSBs are found in the \textit{Deinococcus}\textsuperscript{[235]} and \textit{Thermus}\textsuperscript{[236,237]} genera, which encode homodimeric SSBs containing two OB-folds per monomer. The eukaryotic equivalent of SSB, replication protein A (RPA), has been characterised in organisms ranging from yeast to humans and is universally heterotrimeric\textsuperscript{[238,239]}. The bacteriophages T4 and T7 encode the monomeric gene 32 protein\textsuperscript{[240]} and dimeric gene 2.5 protein\textsuperscript{[241,242]} respectively, whilst archaeal SSBs are diverse, and may be monomeric, dimeric, or trimeric\textsuperscript{[226,243-245]}.

1.1.5.1 SSB structure

\textit{E. coli} SSB is an 18843 Da protein of 177 amino acids\textsuperscript{[210,246]}. It is stable as a tetramer over a wide range of solution condition\textsuperscript{[247,248]}, and at SSB concentrations as low as 20 nM\textsuperscript{[249]}, and it is the tetrameric form which binds to ssDNA\textsuperscript{[250]}. Early studies of various SSB mutants and proteolysis fragments showed that the residues essential for tetramer formation and ssDNA-
Figure 1.11. Domain organisation and structure of SSB. (a) The N-terminal 115 residues of SSB form a structured OB-fold, responsible for tetramerisation and ssDNA-binding. The extreme C-terminal tip, containing a mixture of acidic and hydrophobic residues, mediates the interactions between SSB and its binding partners. The intervening sequence is thought to be flexible. (b) Crystal structure of the SSB homotetramer. Only the OB-fold domains are visible in the structure. Figure produced using PyMol and PDB coordinates 1SRU²⁵⁴.

binding are contained within the N-terminal 115 residues of the sequence (Figure 1.11(a)) – SSB lacking its C-terminal domain retains the ability to bind to ssDNA, and in fact binds with even higher affinity²⁴⁷,²⁵¹,²⁵². The crystal structures of the SSB tetramer, which have been determined for both full-length and C-terminally truncated versions of SSB²²⁹,²⁵³,²⁵⁴, reveal that these first 112 residues are highly structured, and form a classic oligonucleotide-binding (OB) fold – a five-stranded β-barrel capped by an α helix¹¹⁵,²⁵⁸. In contrast, the C-terminal third of the protein is not visible in the crystal structure of full-length SSB²⁵⁴, which, along with its ease of proteolytic removal²⁴⁷ and abundance of glycine and proline residues, suggests that this region of SSB is conformationally dynamic.
The four OB-folds in the SSB tetramer are oriented inwards, forming a structured DNA-binding core with $D_2$ symmetry – essentially a dimer of dimers, with distinct monomer-monomer and dimer-dimer interfaces\[229\] (Figure 1.11b). The association of monomers into dimers is mediated primarily by hydrogen-bonding between β-strands of adjacent monomers, which form a six-stranded antiparallel β-sheet\[229\]. Hydrogen bonding between His55 from one monomer and Asn6 and Leu 83 from the opposing monomer appears particularly important in stabilizing this interaction\[229\]. The well-characterised SSB mutation, ssb-1, replaces His55 with tyrosine\[248\]. Cells carrying the mutation exhibit temperature-sensitive defects in replication, recombination and repair\[208,209,220\], resulting from destabilisation of the SSB-1 tetramer such that it is present primarily as monomers at nonpermissive temperatures\[248,255,256\]. The structure, therefore, explains the basis for this destabilisation. The dimer-dimer interface consists of two six-stranded β-sheets, one from each dimer\[229\]. This interface is hydrophobic in nature at its core, but also contains extensive hydrogen bonding contacts between a number of lysine, tyrosine, glutamine and glutamic acid residues\[229\].

1.1.5.2 SSB-ssDNA interactions

The interaction between SSB and ssDNA is complex. With four potential ssDNA binding sites per tetramer, SSB is able to utilise a number of distinct binding ‘modes’ in vitro, which differ in the extent of contact between SSB and DNA\[257-260\]. The two best characterised binding modes are termed the (SSB)$_{65}$ and (SSB)$_{35}$ modes, so named because they occlude an average of 65 and 35 nucleotides of DNA, respectively\[257\]. The relative stabilities of these binding modes in vitro are highly sensitive to solution conditions – particularly salt concentration, SSB to ssDNA ratio, and the presence of polyamines such as spermine\[249,257,258,261-263\] – and rapid interconversion between modes is possible upon changes in these conditions\[264\]. At low salt concentrations or high SSB to DNA ratios, the (SSB)$_{65}$ mode is formed preferentially, whilst
Figure 1.12. Proposed structures of the (SSB)$_{35}$ and (SSB)$_{65}$ ssDNA binding modes. (a) The (SSB)$_{65}$ mode, in which ~65 nucleotides of ssDNA (in red) wrap around extensively around the SSB tetramer, making contact with all four subunits. (b) The (SSB)$_{35}$ mode, in which ~35 nucleotides make less extensive contacts with the tetramer. (a) and (b) adapted from reference 219. (c) Model for the positive intertetramer cooperativity in the (SSB)$_{35}$ binding mode. In this mode, adjacent tetramers form a continuous tract along the ssDNA.

The (SSB)$_{65}$ binding mode is favoured at monovalent salt concentrations above ~0.2 M$^{[258,264,265]}$. A crystal structure of the structured core of SSB in complex with two molecules of (dC)$_{35}$ has enabled the construction of models of the (SSB)$_{35}$ and (SSB)$_{65}$ binding modes$^{[253]}$ (Figure 1.12). The (SSB)$_{65}$ mode is modelled as a fully wrapped complex, in which ~65 nucleotides of ssDNA interact with all four subunits of the tetramer$^{[253]}$ (Figure 1.12a). In the modelled (SSB)$_{35}$ mode, the tetramer is wrapped less extensively, with ~35 nucleotides making contact with an average of two subunits$^{[253]}$ (Figure 1.12b). Whilst further information is required to gauge how closely these models reflect true SSB-ssDNA binding, both fluorescence quenching$^{[266]}$ and fluorescence resonance energy transfer$^{[264]}$ studies have provided evidence to support their general architecture. The two binding modes also exhibit very different cooperativities in vitro. The (SSB)$_{35}$ mode possesses a high level of intertetramer cooperativity, with tetramers able to bind adjacent to each other to form a continuous protein 'coating' $^{[261,263]}$ (Figure 1.12c). Inter-tetramer cooperativity is much more limited in the (SSB)$_{65}$ mode, with
binding limited to ‘octamers’ consisting of two adjacent tetramers, separated by stretches of SSB-free ssDNA\cite{259,267}. As well as inter-tetramer positive cooperativity, the binding of SSB to ssDNA is characterised by salt-dependent negative cooperativity between subunits within individual SSB tetramers: ssDNA binding to the third and fourth subunits of a tetramer is relatively unfavourable compared to binding to the first and second subunits\cite{266,268}. This negative cooperativity is therefore linked to binding mode preferences at particular salt concentrations.

The significance of the different binding modes \textit{in vivo} remains unclear. It has been proposed that they may be utilised selectively for different purposes – for example, the (SSB)$_{35}$ mode has been proposed to function in DNA replication\cite{265}, with the (SSB)$_{65}$ mode being used selectively for recombination\cite{260}. Recent single-molecule fluorescence microscopy of living \textit{E. coli} cells suggests that the SSB at the replication fork may actually be bound in the (SSB)$_{65}$ mode, based on an apparent stoichiometry of 8 fluorescent SSB tetramers per replication fork (and an estimated 650 nucleotides of free ssDNA)\cite{156}. However, this work did not account for unlabelled SSB present at the fork, meaning that the true SSB stoichiometry may in fact be much higher. Further investigation is clearly required.

\textbf{1.1.5.3 The SSB C-terminus and SSB-protein interactions}  
What, then, is the function of the flexible C-terminal domain of SSB? The extreme C-terminus is a mixture of acidic and hydrophobic residues (Asp-Phe-Asp-Asp-Ile-Pro-Phe in \textit{E. coli}) and is exceptionally highly conserved amongst bacterial SSBs\cite{183,219,269} (Figure 1.13). The C-terminus has long been recognised as being essential to the function of SSB: the well-characterised \textit{ssb-113} mutation, which substitutes the penultimate proline residue with serine, results in a protein which is fully competent in ssDNA-binding, but nonetheless confers temperature-sensitive replication defects, and recombination and repair defects at all t
The sequence of the SSB C-terminus is highly conserved amongst bacteria. The graph depicts the level of conservation of each of the final nine residues (169-177) in the *E. coli* SSB sequence. The percentage sequence identity for each residue was obtained by comparison of 284 bacterial SSB sequences. Figure adapted from reference 269.

Mutant SSB from which the last 26 residues are entirely removed similarly retains the ability to bind ssDNA strongly, but renders cells inviable\[^{251}\]. These effects are now recognised to arise from the fact that the C-terminus is the site through which SSB interacts with (the majority of) its binding partners, of which there are at least 14\[^{219}\]. Despite a lack of sequence similarity, the single-stranded DNA-binding proteins from phage T4 (gp32) and phage T7 (gp2.5) have a similar functional organisation to *E. coli* SSB, both utilising acidic C-termini to mediate their interactions with other proteins\[^{270,271}\]. The function of the intervening flexible sequence (residues ~112-169) between the structured OB-fold domain and the extreme C-terminus is less clear: it may simply act as a ‘spacer’, allowing the extreme C-terminus sufficient mobility to pursue its interactions\[^{251}\]. Nevertheless, additional unique functions for this region cannot be ruled out, with it potentially playing a role in the modulation of ssDNA-binding affinity by the C-terminus (section 3.1).
The proteins which have been identified as interacting with SSB are a broad group, but all are involved in various aspects of DNA metabolism. During DNA replication, SSB is presumed to be primarily associated with the lagging strand, on which ssDNA accumulates transiently and repeatedly during Okazaki fragment synthesis \[^{182,183,272}\]. As described in section 1.1.3.3, SSB interacts with the Pol III holoenzyme via the $\chi$ subunit of the clamp loader complex in an interaction which is apparently important for stabilising the polymerase on the lagging strand template, and, by extension, for the overall stability of the replisome \[^{182,183,195-199}\]. SSB gains a second contact with the lagging strand through its interaction with the replicative primase, DnaG \[^{199,273}\]. Primase, as described previously, associates intermittently with the lagging strand and with the helicase, DnaB, in order to lay down its primers. Subsequently, it is transferred from DnaB to SSB, in an interaction mediated through the SSB C-terminus \[^{273}\], which seems to stabilise primase on the lagging strand. The SSB-DnaG interaction may also have implications for the recognition and loading of the $\beta_2$ clamp onto newly synthesised primers.

In addition to its integral roles in replication, the interactions of SSB with a variety of proteins associated with recombination and repair emphasize its position as an important central coordinator of genome stability. The blocking of DNA replication by UV-induced lesions or other impediments is potentially disastrous to cells, and as such there are multiple pathways for rescuing replication forks which have stalled, and for repairing damage \[^{274}\]. Uncoupled replication as a result of a DNA lesion on one template strand generally results in the generation of large amounts of ssDNA, which is coated by SSB. SSB, therefore, serves as a convenient platform for targeting a variety of rescue proteins to the sites where they are required \[^{219,275}\].

Homologous recombination pathways are important for restoring activity to stalled forks \[^{276-278}\]. One of these, the RecF pathway \[^{279,280}\], involves the activities of several proteins which
interact directly with SSB: RecO, RecQ, RecJ and RecG. RecA, the central recombinase, does not interact directly with SSB, but is profoundly influenced by its presence. RecA catalyses strand exchange between homologous regions of DNA by forming a nucleoprotein filament on ssDNA which is able to invade homologous DNA duplexes. This filament forms in two stages: nucleation – the initial binding of a few RecA molecules to ssDNA which is inhibited on SSB-coated DNA – and filament extension, which is stimulated by SSB. In the RecF pathway, the mediator proteins RecF, RecO and RecR work concertedly to facilitate nucleation of RecA on SSB-coated templates. Of these mediators, only RecO interacts directly with SSB, and does so through the C-terminal eight residues of SSB. Inoue et al. have examined the RecO-SSB interaction in *Thermus thermophilus*, and proposed a mechanism for the RecO-facilitated displacement of SSB. In the model, binding of RecO to the SSB C-terminus alters the mobility of the C-terminus such that it becomes located close to the OB-fold of SSB, thus weakening the SSB-ssDNA interaction and allowing RecO (which also possesses a strong affinity for ssDNA) to bind to the ssDNA and replace SSB, which remains bound to RecO. This model is consistent with findings that the C-terminus of SSB modulates its ssDNA affinity (discussed later), and such a mechanism may prove to be more generally applicable to the question of how SSB allows access of other proteins to ssDNA upon binding to them. The RecO-SSB interaction is also important for the additional function of RecO in annealing SSB-coated ssDNA during recombination processes.

RecQ, a 3′→5′ helicase, and RecJ, a 5′→3′ exonuclease, act together to process stalled replication forks, degrading nascent DNA and converting the blocked fork to a form which permits removal or repair of the blocking lesion. RecQ is also involved in signalling for the induction of the SOS response to DNA damage, whilst RecJ also has roles in mismatch
repair and base excision repair\textsuperscript{[299,300]}. SSB interacts with and stimulates the activities of both RecQ\textsuperscript{[283,284]} and RecJ\textsuperscript{[285]}. A recently uncovered function of RecQ is its ability to cooperate with topoisomerase III to resolve converging replication forks which have stalled, producing separated sister DNA molecules\textsuperscript{[301]}. The cooperation between RecQ and topoisomerase III is mediated by the direct interactions of both proteins with SSB\textsuperscript{[301]}.

One pathway for rescuing stalled replication forks begins with regression/reversal of the damaged fork to form a Holliday junction, which becomes a substrate for subsequent repair and replication restart\textsuperscript{[302,303]}. A second RecF pathway helicase, RecG, is able to catalyse fork regression\textsuperscript{[103-106]}, and is assisted in this task by an interaction with the C-terminus of SSB\textsuperscript{[286]}, which stabilises RecG on ssDNA\textsuperscript{[286,307]}. SSB, via its C-terminus, also interacts with exonuclease I\textsuperscript{[269,108]}, which excises misincorporated nucleotides in mismatch repair\textsuperscript{[109]} and uracil-N-glycosylase (UNG)\textsuperscript{[110,111]}, which is responsible for uracil removal in base excision repair\textsuperscript{[312]}. That these interactions stimulate both ExoI and UNG appears to arise from the fact that SSB serves to efficiently recruit the enzymes to their ssDNA substrates\textsuperscript{[269,108]}. Under certain circumstances, error-protein translesion synthesis (TLS) by specialised DNA polymerases allows replication fork-blocking DNA damage to be bypassed without being repaired\textsuperscript{[311]}. The TLS polymerase Pol V\textsuperscript{[114]} has been shown to rely on recruitment by SSB to catalyse TLS effectively\textsuperscript{[115]}.

Regardless of the means of stalled fork salvage, the final step requires that the replication machinery be reassembled to allow replication to resume. PriA helicase is responsible for initiating replication restart, and serves as a point of assembly for other replication restart proteins\textsuperscript{[116]}. PriA interacts with the SSB C-terminus\textsuperscript{[196,222]}, an interaction which appears to be integral to the replication restart activity of PriA\textsuperscript{[222]}. SSB also interacts with and stimulates Pol
Figure 1.14. The SSB C-terminus possesses similar binding sites on different proteins. The structures of SSB C-terminal peptides (SSB Ct) in complex with (a) exonuclease I (PDB coordinates 3C94\textsuperscript{[269]}) (b) $\chi$ (PDB coordinates 3SXU\textsuperscript{[195]}) and (c) RecO (PDB coordinates 3Q8D\textsuperscript{[282]}). The SSB peptides are shown as sticks, and the protein binding surfaces are coloured blue, red and white to indicate positive, negative and hydrophobic regions of electrostatic potential, respectively. In all three structures, the side chain of the extreme C-terminal phenylalanine residue from SSB inserts into a deep hydrophobic pocket, which in each case is surrounded by a number of basic residues. Selected residues contributing to each interaction are indicated. Figures constructed using PyMol\textsuperscript{[100]}. 
II\textsuperscript{[207,213,215]}, a damage-inducible repair polymerase which is involved in error-free replication restart at stalled forks\textsuperscript{[117,118]}.

Structural information available for several of the SSB-protein interactions – those with $\chi$\textsuperscript{[195]}, DnaG\textsuperscript{[273]}, exonuclease I\textsuperscript{[269]}, RecQ\textsuperscript{[284]}, RecO\textsuperscript{[282]} and UNG\textsuperscript{[311]} – show that many elements of the SSB-binding surface are common to its binding partners, and are governed by the unique combination of residues at the SSB C-terminus. A hydrophobic pocket, in which the hydrophobic tip of the SSB C-terminus binds, is flanked by basic residues likely to form electrostatic contacts with the acidic portion of the SSB tail\textsuperscript{[195,269,284,319]}. The structures of the interfaces of the SSB C-terminus with $\chi$, ExoI and RecO are shown in Figure 1.14. The apparently conserved manner in which SSB binds to other proteins suggests that these interactions could form a potential target for antibacterials. The Keck group have recently identified several small molecules which function as competitive inhibitors of the SSB-ExoI interaction\textsuperscript{[320]}, providing a basis for further work in this area.

Although the most studied function of the SSB C-terminus relates to its role as a mediator of protein-protein interactions, there is evidence that it may also play a role in regulating the ssDNA-binding modes and affinity of SSB. Deletion of the eight C-terminal residues influences the equilibrium between the (SSB)$_{35}$ and (SSB)$_{65}$ binding modes \textit{in vitro}\textsuperscript{[264]}, raising the possibility that sequestration of the C-terminus by protein-protein interactions may be able to exert a similar influence \textit{in vivo}. Interestingly, the C-terminus also appears to have a negative effect on ssDNA affinity – upon its removal, or upon binding to other proteins, the ssDNA affinity of SSB increases\textsuperscript{[196,247,252]}. These effects will be discussed in more detail in section 3.1.

A relatively poorly understood aspect of SSB behaviour is the manner in which it is able to gain mobility on ssDNA, given that it binds with such high affinity. In a process such as replication, which moves extremely rapidly, SSB clearly needs to be either displaced by advancing
proteins, or be able to migrate upon DNA without actual dissociation. Two potential processes were proposed early: ‘rolling’ of SSB tetramers along ssDNA\(^{[2]}\), and direct transfer of tetramers from one ssDNA strand to another\(^{[4]}\). More recently, the potential mechanisms of SSB migration have been explored in more detail\(^{[22-26]}\), and are described in section 4.1.

1.2 Mass Spectrometry for the Study of Noncovalent Protein Interactions

1.2.1 Electrospray ionisation mass spectrometry for analysis of biomolecules

The analytical technique of mass spectrometry (MS) requires that the molecules under consideration be converted to gaseous ions for analysis, such that they can be separated on the basis of their mass-to-charge ratio (\(m/z\)). Traditional methods for ionization tended to result in fragmentation of ions and precluded the ionization of intact biological macromolecules, such as proteins. During the 1980s, electrospray ionization (ESI) was developed as a method for ionizing large biomolecules for MS without fragmentation\(^{[7]}\). Soon after, it was discovered that ESI-MS could even preserve specific noncovalent interactions\(^{[8,9]}\), and the field then expanded rapidly, branching into a diverse range of increasingly large macromolecular assemblies.

The ability of ESI to preserve these interactions arises from its relative gentleness as a method of ionization. The electrospray ionisation process was first described by Dole and colleagues over 40 years ago\(^{[10]}\), and has now been elucidated in detail\(^{[11,12]}\) (Figure 1.15). In the current model for ESI, a solution containing the sample is sprayed through a fine needle to which a potential difference is applied. Each droplet in the resulting spray carries a large number of surface charges, and, as the surface area is decreased by solvent evaporation (generally with the aid of a nebulising gas), these charges experience increasing repulsion. Eventually, this
Figure 1.15. The electrospray ionisation process. In positive ion mode, illustrated here, application of an electric field to the sample-containing capillary causes positively-charged ions to become enriched near the capillary tip. This results in the emission of positively charged droplets from the ‘Taylor cone’ which forms at the end of the capillary. Each droplet contains numerous positively charged molecules, and moves towards the detector, which is negatively charged. The combined effects of solvent evaporation and Coulombic fissions lead to the generation of increasingly small droplets. Eventually, free gas phase ions (represented here as single charges, although proteins generally carry multiple charges) are formed.

Repulsion leads to ‘Coulombic fission’ of the droplets to produce smaller droplets, which subsequently experience repeated evaporation and fission. Once the droplets are sufficiently small (containing a single protein or complex), evaporation of the residual solvent results in transfer of the residual charges to the protein and yields gaseous ions. This ‘charge residue model’ for ion production has been the subject of contention, but it is now generally accepted to apply to production of macromolecular ions by ESI\cite{333,334}. The details of the process are considerably more complex than described here, and are covered in depth elsewhere\cite{331-333,335,336}.

The analysis of large proteins and noncovalent assemblies often benefits from the use of a miniaturised form of ESI-MS known as nanoelectrospray ionisation (nanoESI-MS)\cite{337-339}. 

---

43
NanoESI utilises small capillaries (with an inner tip diameter of \( \sim 1 \, \mu m \)) into which 1-2 \( \mu L \) of sample is directly loaded\(^{117}\). The droplets produced in nanoESI are at least an order of magnitude smaller than those generated by conventional electrospray\(^{117,118}\), which is thought to result in milder and more efficient solvent evaporation, and better preserve noncovalent interactions\(^{340,341}\). As a consequence, nanoESI can effectively increase the sensitivity for detection of noncovalently bound complexes\(^{119}\). The small sample requirement of nanoESI is also advantageous for analysis of multiprotein assemblies, which are often only available in small quantities\(^{119}\). Whilst the majority of MS studies of biomolecular assemblies employ nanoESI, in the following discussion ESI-MS will be used as a general term to cover both ESI and nanoESI methods.

Gaseous protein ions generated by ESI carry multiple charges, and within a protein sample, a number of different ‘charge states’, each carrying a different number of charges, are formed\(^{127,135}\). Protein mass spectrometry is usually performed in positive ion mode, with the basic sites exposed on the surface of the protein becoming protonated during the ESI process\(^{134,162}\). The mass spectrum of a protein therefore contains a distribution or ‘envelope’ of peaks at discrete mass-to-charge (\( m/z \)) values, occurring at \([M+nH]^{+}\) (where \( M \) is the mass of the protein and \( n \) is the number of charges), with adjacent peaks differing from each other by a single charge\(^{127,135}\). These can be deconvoluted to obtain the mass of the protein\(^{135,141}\).

The distribution of charge states observed in the mass spectrum reflects the number of accessible basic sites in a protein, which in turn is directly related to the conformation of the protein\(^{342,344-346}\). The solution conditions from which a protein or noncovalent complex is electrosprayed therefore profoundly affect the appearance of the spectrum. The first protein ESI-MS spectra were obtained under acidic conditions, to facilitate ion formation and volatilization\(^{127}\). Under such conditions, or in the presence of organic solvents, proteins adopt
Figure 1.16. Effect of pH on protein folding detected by ESI-MS. ESI-mass spectra of cytochrome c as the pH is decreased from 6.4 (top) to 2.3 (bottom). At near-neutral pH, a small distribution of ions centred at a relatively high m/z ratio is observed. A broader distribution of ions centred at a lower m/z develops as the pH becomes more acidic. The higher m/z ions carry fewer charges, indicating a folded protein with less solvent-accessible surface area. Under acidic conditions, the protein becomes unfolded, leading to a series of highly charged ions. Figure adapted from reference 348.

an unfolded conformation, allowing a large number of charges to be accommodated on the surface of the protein, and a broad distribution of relatively low m/z ions are observed [342, 346-348] (Figure 1.16). These conditions, however, are clearly not compatible with the maintenance of
noncovalent complexes. The use of volatile aqueous buffers at near-neutral pH (ammonium acetate is commonly used\cite{349}) allows protein folding and noncovalent interactions to be maintained whilst still facilitating formation of ions (with NH$_4^+$ ions providing the positive charges)\cite{334}. The more compact conformation of folded proteins causes a smaller number of charges to be acquired, and native protein spectra typically contain a narrower distribution of charge states centred at higher m/z than those of their unfolded counterparts\cite{342,346,347,350} (Figure 1.16). Of course, every sample has its own particular requirements, and solution conditions must be optimised carefully for individual cases. It should also be noted that charge state distributions may be affected by variations in the conditions within the mass spectrometer, as well as by solution conformations.

Although the development of ESI and nanoESI was revolutionary in allowing the generation of gas phase ions from large proteins and protein assemblies, the transmission and mass analysis of these large biomolecular species has also undergone a number of developments since early studies. The most common instruments currently used for analysis of large biomolecules contain a quadrupole mass analyser coupled orthogonally with time of flight (ToF)\cite{349,351,352}. Figure 1.17 shows the arrangement of a typical quadrupole time of flight (Q-ToF) instrument.

ToF analysers are ideally suited to detection of macromolecules because of their high resolution and theoretically unlimited m/z range\cite{353}. The preceding quadrupole allows for tandem mass spectrometry (MS/MS) experiments to be carried out – in MS/MS, ions at a specified m/z can be selected (by the quadrupole in a Q-ToF instrument) and subjected to dissociation by collisions with an inert gas before being analysed (collision-induced dissociation, or CID)\cite{354}. Such controlled dissociation can be useful for elucidating details about the stoichiometry or subunit architecture of oligomeric complexes (see later). Conventional quadrupole cells are limited to operation up to ~3000-4000 m/z\cite{351}, but
Figure 1.17. Arrangement of a quadrupole-time-of-flight (Q-ToF) mass spectrometer used for the transmission of proteins and protein complexes. The quadrupole region of the instrument allows for selection and collision-induced dissociation of ions at specific $m/z$ ratios. The time-of-flight analyser, in which ions are separated by reflection in a flight tube before being detected, has a theoretically unlimited mass range. The path of ions through the mass spectrometer is indicated by the dotted line.

Modifications have made it possible for operation at much higher mass ranges (up to 32000 $m/z$), allowing tandem MS experiments to be performed on much larger complexes\cite{351,355,356}.

Another significant step in the analysis of large noncovalent complexes came with the recognition that increasing the pressures within the early regions of the mass spectrometer could improve transmission of large ions through the instrument and result in enhanced detection\cite{357-359}. This ‘collisional cooling’ or ‘collisional focusing’ arises from the additional collisions with inert gas molecules experienced by the analyte ions, which dissipates some of their translational energy, constraining them to a more focussed trajectory through the subsequent stages of the spectrometer\cite{360,361}. Use of collisional cooling methods is essential for gaining high-quality ESI-MS spectra of large proteins or biomolecular complexes, and can
allow for the maintenance of intact noncovalent complexes that would otherwise be lost during flight through the mass spectrometer\cite{158}.

### 1.2.2 Applications of ESI-MS to intact protein complexes and stoichiometry determination

The earliest examples of noncovalent protein complexes analysed by ESI-MS included receptor-ligand complexes\cite{128}, lysozyme in complex with substrate and product oligosaccharides\cite{362}, haem-bound myoglobin\cite{129}, and, in the first example of a multisubunit interaction, dimeric HIV-1 protease bound to an inhibitor\cite{363}. Subsequently, a large number of examples accumulated in the literature as the advantages of ESI-MS for protein analysis became apparent – its speed, high sensitivity and directness, and in particular its ability to directly elucidate the subunit stoichiometry of oligomeric proteins. Many studies have used ESI-MS to clarify or confirm the stoichiometry of proteins for which it was previously uncertain. In an early example, the enzyme 4-oxalocrotonate tautomerase (4-OT), which had been shown by X-ray crystallography to be hexameric after some contention about its oligomeric structure, was confirmed as a hexamer by ESI-MS\cite{364}. Furthermore, deletion of a catalytically important amino acid resulted in loss of the hexameric form and observation of only monomers, indicating its importance to hexamer assembly\cite{364}. During this period of development, the stoichiometries of a variety of other proteins, including the tetrameric proteins avidin\cite{365}, pyruvate kinase\cite{366}, concanavalin A\cite{365,367}, and haemoglobin\cite{365,367}, as well as many others, were also successfully probed by ESI-MS, proving the utility of the technique for such studies. An additional benefit comes from the fact that multiple stoichiometries within a sample can be observed in a single spectrum, allowing the effects of ligand binding or other modifications upon oligomeric states to be probed very directly. This was shown elegantly in a study of *E. coli* citrate synthase, which was shown by ESI-MS to exist in equilibrium between dimeric and
hexameric forms[^168]. Binding of NADH, an allosteric inhibitor of the enzyme, was shown to occur selectively to the hexamer and alter the equilibrium such that the hexameric form was favoured[^168].

With the advances in instrumentation and methodology described in section 1.2.1, noncovalent assemblies of increasing size, complexity and variety became accessible to analysis by ESI-MS. Large complexes which have been successfully analysed include the 14 subunit GroEL chaperonin complex[^358,359], octameric vanillyl-alcohol oxidase[^369], ribosomes[^370,371], virus capsids[^372,373], bacteriophage portal assemblies[^374] and proteasome complexes[^360,375]. The size of these complexes approaches or extends well into the mega-Dalton range, yet the resolution possible with ToF detection allows relatively accurate mass measurement (and therefore stoichiometry determination) even at very high masses.

Selective dissociation of subunits, either by manipulation of solution conditions or by controlled dissociation within the mass spectrometer (e.g. using collision-induced dissociation and tandem MS) can be utilised to gain additional information about the stoichiometry of large complexes. Interestingly, CID of oligomeric assemblies does not usually result in symmetrical dissociation – rather, monomeric units are ejected sequentially from the larger complex, with the dissociated monomeric ions acquiring a disproportionately large portion of the overall charge[^359,364,370,376-379]. This asymmetric mode of dissociation has been shown to arise from the unfolding of a monomer within the complex, which then dissociates, taking with it a large number of charges due to its conformation[^377,378]. Another consequence of CID is that the oligomers remaining after monomer dissociation have generally been stripped of the residual solvent ions and water molecules which can remain associated with complexes in the gas phase and cause peak broadening[^359]. This can allow the determination of stoichiometries even if the native spectrum is not well-resolved[^359]. These features have been exploited to examine the
stoichiometries of heterogenous assemblies. The small heat-shock protein αB-crystallin, which forms a polydisperse mixture of oligomers, gives rise to overlapping peaks from different oligomeric species in the native mass spectrum\textsuperscript{[176]}. By isolating a single charge state and subjecting it to CID, monomers were able to be ejected sequentially, leaving ‘stripped’ oligomers with reduced charge, which could be resolved into the various oligomeric forms\textsuperscript{[176]}. A similar approach was used to clarify the stoichiometry of the stalk complex of ribosomes from several thermophilic bacteria\textsuperscript{[156]} and of the \textit{Bacillus subtilis} TRAP (tryptophan RNA-binding attenuation protein) complex\textsuperscript{[155]}. Tandem MS can also be used to probe the architecture of large complexes containing a number of different subunits. Within a complex, subunits on the exterior will tend to be held more weakly than those in the core of the structure, and will therefore dissociate more readily\textsuperscript{[179]}. This principle has been used to identify peripheral subunits in the 19S proteasome lid complex\textsuperscript{[180]}, to show that the intact 20S proteasome loses subunits in a manner consistent with expectations based on its structure\textsuperscript{[175]}, and to probe the subunit arrangement of the eukaryotic initiation factor 3\textsuperscript{[381]}. Therefore, when combined with other techniques such as electron microscopy and homology modelling, MS methods have the potential to provide valuable contributions towards mapping the organisation of such complexes\textsuperscript{[382,383]}. Alongside these developments, there has been substantial interest in the question of how well noncovalent interactions in the gas phase reflect those in solution\textsuperscript{[384,385]}. It seems likely that transfer into the gas phase would affect different types of noncovalent interactions in different ways – some studies suggest that electrostatic interactions are strengthened in the gas phase\textsuperscript{[159]} whilst hydrophobic interactions, generally thought to rely on the presence of water, are weakened or lost\textsuperscript{[186]}. However, a recent study of protein-fatty acid complexes suggests that hydrophobic forces may in fact be retained in the gas phase\textsuperscript{[187,188]}. The replisomal complex
which is held together primarily by hydrophobic forces, has also been successfully observed by ESI-MS. Although changes in the strength of noncovalent interactions in the gas phase could be expected to result in observation of complexes which are artificially strengthened or weakened, most noncovalent complexes utilise a combination of electrostatic, hydrophobic and hydrogen-bonding interactions, which may modulate the overall effect. The fact that subunit stoichiometries measured by MS are generally consistent with those obtained using other methods is an indication that interactions between subunits seems to be generally well-preserved. The question of whether protein conformations at the level of folding are maintained in the gas phase remains something of an open question, although there is increasing evidence to show that, in many cases, they can be. Support comes from the ability of viruses and enzymes, collected using ‘soft-landing’ techniques after being electrosprayed, to retain their infectivity and activity, respectively. Furthermore, microscopy of GroEL tetradecamers collected after being separated by nanoESI-MS revealed that the complex appeared to retain its overall topology.

The emergence of ion-mobility mass spectrometry (IM-MS) as an analytical method has greatly benefitted investigations of gas phase protein conformations. In a traditional ion-mobility separation, ions are directed by an electric field through a gas-filled cell of a given length, and the time taken for them to drift through the cell is measured. This allows a separation on the basis of cross-sectional area: larger ions, which undergo more collisions with the gas molecules, have longer drift times; vice versa for more compact ions. A number of variations on this concept exist: in travelling-wave ion-mobility spectrometry (TWIMS), waves of voltage are pulsed sequentially along a series of stacked rings: higher mobility (i.e. smaller cross-section) ions are able to ‘surf’ the waves and are transmitted more rapidly; lower-mobility ions fall behind the waves and travel more slowly. The coupling of ion-mobility with mass spectrometry allows the conformations of ions to be analysed.
simultaneously with their mass-to-charge ratio. In the commercially available Synapt IM-MS instrument (Waters)\cite{396}, a travelling-wave ion-mobility cell is coupled to a Q-TOF mass spectrometer – ions arriving from the TWIMS region are analysed by ToF-MS as they arrive, thus providing two dimensional information about both their mass and shape.

The application of IM-MS to exploring the topology of protein complexes has provided convincing evidence that elements of solution-phase topology are able to be maintained in the gas phase\cite{397}. Measurements of the collisional cross-sections of several small proteins by IM-MS have shown good correlation with the cross-sections predicted by other methods\cite{398-400}. Examples of larger complexes for which IM-MS analysis has shown a preservation of native structure include the ring-shaped TRAP complex\cite{401}, the 20S proteasome\cite{375}, and the GroEL and GroEL-GroES complexes\cite{402}.

1.2.3 Applications of ESI-MS to protein-nucleic acid interactions

Although a large amount of attention has been focussed on the application of ESI-MS to large multiprotein assemblies, a less prominent but nonetheless important contribution of ESI-MS has been made in the study of noncovalent protein-nucleic acid interactions. Mass spectrometry, with its characteristic speed and sensitivity, can provide information about several aspects of these interactions, and there are a relatively small but varied number of examples\cite{403-412} which have taken advantage of the method (reviewed in \cite{413-415}). As it does for multisubunit protein stoichiometries, ESI-MS offers the ability to directly observe protein-nucleic acid binding stoichiometries. Protein-nucleic acid complexes for which ESI-MS has been used to probe stoichiometry include the single-stranded DNA-binding protein from bacteriophage f1 (gene V protein) with ssDNA oligonucleotides\cite{403}, E. coli RNase E with RNA analogues\cite{412} and HIV-1 nucleocapsid protein and its RNA $\psi$-recognition sequence\cite{411}. Often, specific DNA-binding interactions can stabilise a particular oligomeric form of a protein: again,
ESI-MS is able to show such stabilisation very directly – this has been illustrated for the
dimeric vitamin D receptor[407] and a dimeric leucine zipper peptide[408]. In an ESI-MS study of
the bacteriophage λ integrase protein (λ-Int)[410], Kamadurai et al. analysed the pattern of
charge states in the absence of DNA and observed three distributions, assigned to unfolded,
folded and dimeric forms of the protein. In the presence of DNA, the charge state distribution
Corresponding to the unfolded species was abolished, leading to the conclusion that DNA
Stabilised the native fold of λ-Int[410]. This study highlights the great advantage of MS in
Allowing the observation of multiple conformational populations simultaneously.

One concern with the use of ESI-MS for examining protein-nucleic acid interactions is that
They typically include a large contribution from electrostatic forces, which, as described
Earlier, may be strengthened in the gas phase. This raises the question of whether the
Complexes observed by ESI-MS consist of specific interactions, or are the result of nonspecific
Interactions formed within the mass spectrometer[350,414,415]. Evidence from several studies
Suggests that, in many cases, the gas phase interactions observed by ESI-MS are specific. The
Transcription factor PU.1 interacts with a specific dsDNA recognition motif – when mixtures
Of the DNA-binding region of PU.1 and dsDNA were subjected to ESI-MS, complexes were
Only observed when the target sequence was present, indicating a specific interaction[404].

Similar specificity requirements have been shown for the replication terminator protein from
E. coli, Tus, and its specific DNA sequence Ter[409] and the E. coli trp repressor, TrpR, and its
Specific operator DNA[406]. Whilst the overemphasis of electrostatic interactions in the gas
Phase can present a problem for the determination of absolute dissociation constants for
Protein-nucleic acid interactions by MS[409], it can nonetheless be effective for determining
Relative binding strengths, within, for example, a series of mutants. This approach has been
Used to probe the effects of Tus point mutations and Ter sequence alterations in the Tus-Ter
interaction\textsuperscript{[409]}, and to rank the affinity of three different RNA hairpins from the HIV-1 ψ-recognition element for the HIV-1 nucleocapsid protein\textsuperscript{[411]}.

1.2.4 ESI-MS for the study of macromolecular dynamics

Protein complexes are often highly dynamic in nature, and the rapid acquisition times in MS makes it ideally suited to the investigation of dynamic processes in macromolecular complexes, as changes in complex composition can easily be monitored in real-time\textsuperscript{[416-418]} (Figure 1.18). The separative power of mass spectrometry also allows the identification of multiple intermediates and products resulting from dynamic pathways. This is particularly useful for monitoring assembly and disassembly pathways of protein complexes, which may contain a number of intermediates. ESI-MS has been used to follow the assembly and disassembly pathways of the MtGimC chaperone hexamer\textsuperscript{[419]}, the mechanism of bacterial pilus assembly\textsuperscript{[420]}, the mechanism of assembly of the 20S proteasome\textsuperscript{[421]}, the kinetics of haemoglobin subunit disassembly and unfolding\textsuperscript{[422]}, the pathways for assembly of amyloid-β aggregates, which are implicated in a number of diseases\textsuperscript{[423]}, and the thermally-induced dissociation of a small heat shock protein from wheat\textsuperscript{[424]}, among other examples. In addition, ESI-MS has proved to have great utility for following subunit exchange processes in multisubunit proteins. Subunit exchange in the \textit{E. coli} SSB tetramer as monitored by ESI-MS will be described in Chapter 3 of this thesis, and the use of ESI-MS for monitoring subunit exchange is reviewed in detail in section 3.1.
Figure 1.18. Mass spectrometry for measuring protein dynamics. (a) Time-resolved information can be obtained from MS in various ways. For fast reactions, continuous monitoring of a single sample is often useful. MS measurements taken at equilibrium can reveal the relative quantities of products, reactants, and intermediates. Reactions occurring over a more extended timescale can be monitored by analysis of samples taken periodically. (b) Various dynamic processes are accessible to a mass spectrometry approach. MS has the ability to measure mass, charge, intensity and drift time (when combined with ion-mobility spectrometry) simultaneously. It can therefore be applied to study association (and dissociation) of protein complexes, changes in complex stoichiometry, protein folding and conformations, and subunit exchange (not shown). Figure reproduced from reference 417.
1.3 AIMS AND SCOPE OF THIS THESIS

The work described in this thesis was directed towards two broad aims: to use ESI-MS to characterise aspects of the SSB tetramer and SSB-ssDNA binding; and to probe the strength and functional characteristics of the newly identified ε-β₂ interaction.

Despite extensive characterisation of SSB, certain aspects of its behaviour remain poorly understood. The potential for the C-terminal tail of SSB to interact with the DNA-binding OB-fold has recently been proposed[252], but as yet no direct evidence for such an interaction has been presented. In addition, the mechanisms by which SSB is able to move rapidly upon ssDNA during various DNA metabolic processes are still to be fully elucidated. The development of conditions for analysing the SSB tetramer by ESI-MS allowed ESI-MS to be applied to the measurement of SSB subunit exchange rates (Chapter 3). Comparison of exchange rates in the presence and absence of the SSB C-terminus, and of ssDNA, allowed the effects of the C-terminus and ssDNA-binding upon tetramer stability to be probed (Chapter 3). Given the complexity of SSB binding to ssDNA, it represents an ideal complex for addressing concerns about whether protein-ssDNA complexes in the gas phase reflect their solution phase counterparts. To address this, ESI-MS was used to examine complexes of SSB with ssDNA formed under specific solution conditions, and the observed complexes compared to those predicted from solution studies (Chapter 4).

The interaction between the β₂ sliding clamp and the ε exonuclease subunit was recently identified in our laboratory based on an observed requirement for ε in in vitro strand displacement assays. Several questions about the ε-β₂ interaction remained unanswered when this work was commenced. Was the interaction mediated through the common protein binding clefts on β₂, and, if so, were both protein binding clefts in the β₂ dimer required? Did the interaction play a role in DNA synthesis reactions other than the specialised strand
displacement reaction? What was the strength of the interaction? Answering these questions required the construction of β dimers in which the putative ε-binding cleft was disrupted in either one or both subunits. Construction of these dimers is described in Chapter 5. Functional DNA replication assays using these hybrid dimers, as well as wild-type dimers, were used to probe the nature of the ε-β interaction (Chapter 6). Surface plasmon resonance was used to measure the strength of the ε-β interaction using a series of ε peptides (Chapter 7).
Chapter 2: **GENERAL MATERIALS AND METHODS**

### 2.1 MATERIALS

MilliQ water (Millipore, Bedford, USA) was used in all experiments. Ammonium acetate, ammonium sulfate, acetic acid, ammonia, acetonitrile, isopropanol, formic acid, methanol and hydrochloric acid were from Ajax Finechem (Sydney, Australia). Tryptone, yeast extract and agar were from Bacto Laboratories (Sydney, Australia). DTT was from Astral Scientific (Sydney, Australia). All other reagents were purchased from Sigma-Aldrich (St. Louis, USA) unless otherwise specified.

### 2.2 PLASMIDS, STRAINS AND GROWTH MEDIA

#### 2.2.1 Plasmid Vectors

Plasmids described in Chapter 5 were constructed using a number of parental plasmid vectors. Vectors pND706, pND707 and pCL476\(^{425}\) all contain the bacteriophage \(\lambda\) promoters \(p_R\) and \(p_L\) arranged in tandem upstream of a strong ribosome binding site (RBS). The \(bla\) gene confers ampicillin resistance, and transcription control is enabled by the presence of the \(c\nu857\) gene, encoding the \(\lambda\) repressor, which represses transcription from \(p_R\) and \(p_L\) at low growth temperatures (30°C)\(^{426}\). A temperature shift to 42°C inactivates the thermolabile repressor and allows overproduction. In pND706 and pND707, a unique \(NdeI\) site (CATATG) downstream of the RBS contains an ATG start codon, enabling in-frame gene fusion. In pCL476, the \(NdeI\) site is preceded by sequence coding for six His residues, allowing production of N-terminally His\(_6\)-tagged gene products\(^{425}\).
2.2.2 Bacterial strains

During plasmid construction, *E. coli* strain AN1459 (F *ilv* *thr* *leu* hsdR recA srlA::Tn10) was used as a host. For protein expression, plasmids were transformed (section 2.3.4) into *E. coli* strain BL21(λDE3)recA [F *ompT* (*lon*) hsdS<sub>B</sub> (r<sub>R</sub> m<sub>R</sub>) recA srlA::Tn10] [428,429].

2.2.3 LB medium

Unless otherwise specified, *E. coli* strains were grown in LB (lysogeny broth) medium containing 10 g/L tryptone, 5 g/L yeast extract and 5 g/L NaCl, supplemented with 50 mg/L thymine and 100-200 mg/L ampicillin. LB agar plates were of the same composition but contained 10 g/L agar. For expression of genes under the control of the tandem bacteriophage λ <i>p</i><sub>R</sub> and <i>p</i><sub>L</sub> promoters, strains were grown at 30°C then shifted to 42°C to induce expression.

2.2.4 Minimal medium for production of 15N-labelled proteins

Uniformly 15N-labelled proteins were expressed in minimal medium, containing 100 mM K<sub>2</sub>HPO<sub>4</sub>/NaH<sub>2</sub>PO<sub>4</sub>, pH 7.1, 1 mM MgSO<sub>4</sub>, 40 mM D(+) glucose, 1 µg/mL thiamine, trace salts, 100 µg/mL ampicillin, and 1 g/L 15NH<sub>4</sub>Cl (Cambridge Isotope Laboratories) as the sole source of nitrogen. Cells were first grown to stationary phase in LB medium (section 2.2.1), then harvested by centrifugation (4000 × g; 15 min), washed twice with minimal medium, and resuspended in minimal medium before growing and inducing expression as described.

2.3 MOLECULAR GENETIC PROCEDURES

2.3.1 Plasmid preparation

Small scale preparations of plasmids from *E. coli* cells were obtained by overnight growth of the appropriate strain on LB agar (section 2.2.3) followed by extraction and purification with
the QIAprep spin Miniprep kit (Qiagen). Plasmid preparations were subsequently used for transformation of competent cells or restriction digestion, as required.

2.3.2 Restriction digestion and ligation of plasmids

All restriction endonucleases along with the appropriate buffers were supplied by New England Biolabs. Plasmid digestions were carried out at 37°C for 2 h, and products were 5'-dephosphorylated by addition of shrimp alkaline phosphatase (Fermentas) followed by incubation for a further 1 h at 37°C. Fragments were separated by agarose gel electrophoresis and isolated from the gels as described in section 2.3.3. Ligation reactions with purified DNA fragments were carried out with T4 DNA ligase (Fermentas) according to the manufacturer’s directions.

2.3.3 Agarose gel electrophoresis and isolation of DNA fragments

DNA fragments were separated by agarose gel electrophoresis on gels containing 0.8-1.2% (w/v) agarose (Invitrogen). For analytical electrophoresis, gels were prepared and run in 0.5×TBE buffer$^{4}$ (45 mM Tris-borate, 1 mM EDTA). For preparative separation of fragments, gels were prepared and run in 1×TAE buffer$^{4}$ (40 mM Tris-acetate, 1 mM EDTA). Ethidium bromide was present in the gels and running buffer at 0.5 µg/mL. Samples were mixed with DNA loading dye (Fermentas) before loading such that the final concentrations of components were 1.7 mM Tris-HCl (pH 7.6), 0.005% (w/v) bromophenol blue, 0.005% (w/v) xylene cyanol, 10% (v/v) glycerol and 10 mM EDTA. Gels were run at 50 V until the desired resolution of fragments was achieved. DNA fragments were visualised using the UV transilluminator of a GelDoc XR+ system (BioRad), or, for preparative gels, under a small UV lamp. To isolate fragments from gels, individual bands were excised and DNA recovered using the commercially available QIAquick gel extraction kit (Qiagen).
2.3.4 Transformation of competent *E. coli* cells

Electrocompetent *E. coli* cells were prepared and transformed with plasmid DNA essentially as described\(^{[412]}\). Typically, 1-3 µL of DNA were mixed with 100 µL of electrocompetent cells, transferred to a chilled MicroPulser electroporation cuvette (BioRad), and electropulsed at 2.5 kV using a MicroPulser electroporator (BioRad). After transformation, 1 mL of LB medium (section 2.2.3) was immediately added, and the cells recovered at 30°C for 1 h before plating on LB agar (containing ampicillin for selection) and incubating at 30°C overnight.

2.3.5 Screening of transformants by colony PCR

Transformants were first screened using colony PCR to confirm incorporation of DNA inserts. Colony PCR reactions contained cells taken directly from a transformant colony along with 0.4 µM each of a forward and reverse primer, 0.25 mM of each dNTP, 67 mM Tris-HCl, pH 8.8, 10 mM NaCl, 16 mM (NH\(_4\))\(_2\)SO\(_4\), 10 mM KCl, 2.5 mM MgCl\(_2\), and 2.5 units *Taq* DNA polymerase (Bioline). PCR cycle conditions were: initial denaturing step at 94°C for 2 min; 30 cycles of denaturation at 94°C for 30 s, annealing at 50°C for 30 s, polymerization at 72°C for 60 s; and a final 5 min extension step at 72°C. Reaction products were analysed by agarose gel electrophoresis (section 2.3.3). Plasmids were extracted from colonies which gave rise to bands of the predicted insert size and sequenced (section 2.3.6) to confirm successful construction.

2.3.6 Nucleotide sequence determination

Isolated plasmid DNA was sequenced using a commercially available BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems) and appropriate primers external to the gene of interest. PCR cycle conditions were: 96°C for 2 min; 30 cycles of 96°C for 30 s, 50°C for 15 s, 60°C for 4 min. Unincorporated dye terminators were removed from reactions using a
DyeEx 2.0 spin kit (Qiagen) and reactions were analysed on an ABI 3130 Genetic Analyzer (Applied Biosystems).

### 2.4 Protein Purification and Quantification

#### 2.4.1 General purification procedures

All protein purification steps, unless otherwise specified, were carried out at 4°C. Fast protein liquid chromatography (FPLC) was carried out with an ÄKTA purifier system in conjunction with UNICORN control system software (GE Healthcare). Buffers for chromatography were filtered through 0.2 µm filters before use. All dialysis steps used Spectra/Por dialysis tubing (Spectrum Laboratories) with an appropriate molecular weight cut-off (MWCO).

#### 2.4.2 SDS-polyacrylamide gel electrophoresis

Denaturing SDS-polyacrylamide gel electrophoresis (SDS-PAGE) was used to assess expression and purification of proteins. NuPAGE 4-12% Bis-Tris gels (Invitrogen) were run in an Xcell4 Midi-Cell apparatus (Invitrogen) in SDS-PAGE running buffer (50 mM MES, 50 mM Tris base, 1 mM EDTA, 0.1% (w/v) SDS). Protein samples were mixed with SDS loading buffer (300 mM Tris base, 50 mM DTT, 2% (w/v) SDS, 10 % (v/v) glycerol and 0.03% (w/v) bromophenol blue) and heated to 95°C for 3 min before loading. Gels were electrophoresed at 180 V for ~1 hour. Following electrophoresis, gels were stained with Coomassie staining solution (40% (v/v) methanol, 10% (v/v) acetic acid, 0.3% (w/v) Coomassie brilliant blue) for 20 min, then destained in solution containing 25% (v/v) methanol, 10% (v/v) acetic acid. Mark12 protein standards (Invitrogen) were run on each gel and used to estimate protein size.
2.4.3 Determination of protein concentrations

Protein concentrations were determined by measuring absorbance at 280 nm ($A_{280}$). The molar extinction coefficients of proteins at 280 nm ($\varepsilon_{280}$) were either literature values, if available, or calculated from their amino acid compositions as described\cite{433}.

2.5 Oligonucleotides

All oligonucleotides were obtained from Geneworks (Adelaide, Australia). Oligonucleotides were resuspended in TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA) to a concentration of 100 µM, then stored at -20°C until use. Oligonucleotide concentrations were determined by measuring absorbance at 260 nm ($A_{260}$). The molecular weights and molar extinction coefficients (at 260 nm) of oligonucleotides used in mass spectrometry and surface plasmon resonance experiments, as supplied by the manufacturer, are shown in Table 2.1.

Table 2.1. Properties of oligonucleotides used in this thesis.

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Molecular Weight (Da)</th>
<th>$\varepsilon_{260}$ (M$^{-1}$cm$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(dT)$_{35}$</td>
<td>10585</td>
<td>284100</td>
</tr>
<tr>
<td>(dT)$_{70}$</td>
<td>21232</td>
<td>567600</td>
</tr>
<tr>
<td>Bio-(dT)$_{35}$</td>
<td>11154</td>
<td>284100</td>
</tr>
<tr>
<td>Bio-(dT)$_{70}$</td>
<td>21801</td>
<td>567600</td>
</tr>
</tbody>
</table>

*Bio-(dT)$_{35}$ and Bio-(dT)$_{70}$ were biotinylated on their 5’ termini.*
2.6 Mass Spectrometry

2.6.1 Preparation of proteins and oligonucleotides for mass spectrometry

Immediately prior to use in mass spectrometry experiments, proteins and oligonucleotides were dialysed overnight at 4°C against four 2 L changes of the appropriate buffer (details specified for individual experiments).

2.6.2 Mass spectrometry conditions and specifications

All mass spectrometry described in this thesis was performed using either a factory modified quadrupole time-of-flight (Q-ToF) Ultima mass spectrometer (m/z 32000), or a Synapt Q-Tof/ion mobility mass spectrometer (both Waters, UK). All spectra were obtained in positive ion mode, using a nanoelectrospray (nanoESI) source. NanoESI needles for sample loading were prepared by pulling borosilicate glass capillaries (1.0 mm outer diameter, 0.78 mm inner diameter; Harvard Apparatus) with a Flaming/Brown P-97 micropipette puller (Sutter Instruments). A conductive gold coating was applied to the needles using an Emitech K500x sputter coater (EM technologiess). Prior to sample loading, the pulled tips were cut with tweezers, creating a final tip inner diameter of 1-5 µm.

Typically, approximately 3 µL of sample was inserted into capillaries for analysis. Prior to analysis of samples, external calibration was performed using a solution of caesium iodide (10 mg/mL in 70% isopropanol). Conditions within the mass spectrometers were optimised for preservation of noncovalent interactions. Typical parameters for each instrument are given in Table 2.2: some variation in these parameters was required for specific experiments to enable optimal transmission of complexes. Spectra were collected and processed using MassLynx version 4.1 (Waters). Depending on sample concentration, between 20 and 150 acquisitions were combined to obtain a spectrum with the required intensity for meaningful data analysis.
Combined spectra were minimally smoothed using a Savitzky-Golay algorithm. Masses were calculated within MassLynx by manual selection of adjacent ions in a series.

**Table 2.2.** Typical instrumental parameters used for nanoESI-MS.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Q-ToF <em>Ultima</em></th>
<th>Synapt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capillary voltage (kV)</td>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>Sampling cone (V)</td>
<td>80-120</td>
<td>120-150</td>
</tr>
<tr>
<td>Scan time (s)</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>ToF (kV)</td>
<td>9.1</td>
<td>9.1</td>
</tr>
<tr>
<td>MCP detector voltage (V)</td>
<td>1700</td>
<td>1800</td>
</tr>
<tr>
<td>Source Temperature (°C)</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Desolvation Temperature (°C)</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Backing pressure (mbar)</td>
<td>1.5-2.5</td>
<td>3-4</td>
</tr>
</tbody>
</table>
Chapter 3 : **SUBUNIT DYNAMICS IN *E. coli* SINGLE-STRANDED DNA-BINDING PROTEIN (SSB)**

3.1 **INTRODUCTION**

3.1.1 **Potential interactions between the SSB C-terminus and OB-fold**

The C-terminal tail of SSB has been well-characterised as the site of interaction with the many protein binding partners of SSB, including the clamp loader subunit $\chi$ \[^{[182,183,195]}\], exonuclease I[^{[269]}], PriA[^{[222]}], and several others (see section 1.1.3). The residues of SSB implicated in these interactions are contained within the predominantly negatively charged extreme C-terminus of 8-10 amino acids, with the sequence PMDFDDDIPF. However, little is known about the structure of the C-terminus — it is not observed in any of the available SSB crystal structures and is thought to be intrinsically disordered[^{[214]}].

Several pieces of recent evidence suggest that, in the absence of ssDNA, the C-terminal tails may be sequestered by the ssDNA-binding sites of SSB, being displaced by ssDNA-binding to allow protein-protein interactions to occur. Deletion of the C-terminal 8 amino acids has been shown (by fluorescence quenching) to increase the affinity of SSB for ssDNA at moderate salt concentrations[^{[212]}]. Using isothermal titration calorimetry (ITC), two of the binding partners of SSB, PriA and the $\chi$ subunit of DNA pol III holoenzyme, were found to interact with the isolated C-terminal tail with increased affinity relative to full-length SSB[^{[196]}]. In addition, prebinding of oligonucleotides to intact SSB resulted in a significant increase in its affinity for both PriA[^{[196]}] and $\chi$[^{[196,197]}]. In the case of the PriA-SSB complex, greater increases in affinity, as well as stoichiometry, were seen as the length of the prebound oligonucleotide was increased (from (dT)$_{15}$ to (dT)$_{70}$). Finally, deletion of the 42 C-terminal amino acids of SSB has been
Figure 3.1. Model for the proposed interaction between the SSB C-terminus and the structured OB-fold domain. Occupation of the OB-folds of the SSB tetramer (blue) by the C-termini (green) has an inhibitory effect on ssDNA binding. Upon binding to ssDNA, the C-termini are displaced and are free to interact with other proteins. The flexible region between the OB-fold and extreme C-terminus is represented in grey.

shown to affect the transition between the (SSB)$_{65}$ and (SSB)$_{35}$ binding modes, shifting the equilibrium in favour of the (SSB)$_{35}$ binding mode$^{[264]}$.

These behaviours could be explained by a situation in which the C-termini are able to interact with the ssDNA binding sites of SSB in the absence of ssDNA (Figure 3.1). Such a mechanism could regulate both the ssDNA binding and protein-protein interaction activities of SSB: binding of ssDNA to SSB would displace the C-terminus from the ssDNA binding site, freeing it to interact with its binding partners. This would account for the inhibition of ssDNA-binding by the C-terminus$^{[252]}$, and for the apparent inhibitory effect of the SSB core on interactions of the C-terminus with its binding partners$^{[196]}$. In this model, the more SSB subunits that are occupied by ssDNA, the more C-terminal tails should be available for other interactions. This provides an explanation for the increased efficiency of the PriA-SSB interaction when SSB is bound to (dT)$_{70}$, which occupies all four subunits, compared to when SSB is bound to (dT)$_{35}$, which only occupies two subunits$^{[196]}$.  


The ssDNA binding site of SSB has a net positive charge, and any interaction with the acidic C-terminus is likely to have a large electrostatic component\cite{252}. Such an interaction would be more favourable in low salt conditions, and would explain why the inhibitory effect of the C-terminus on ssDNA affinity can be discerned at moderate salt concentrations (100-200 mM NaCl) but is not observed at higher salt concentrations\cite{252}.

These recent observations are consistent with early observations that cooperative binding of SSB to ssDNA increased the susceptibility of its C-terminus to proteolysis\cite{247}. In these proteolytic studies, truncated SSB also appeared to form less stable tetramers, with a greater proportion of monomers observed by gel filtration relative to full-length SSB\cite{247}. However, this observation has not been investigated further with regard to a possible role for the SSB C-terminus in stabilising the tetrameric interactions, and the N-terminus is generally considered the sole contributor to tetramer formation.

Effects of an acidic C-terminus upon DNA binding are not limited to \textit{E. coli} SSB, and indeed a model in which a competitive C-terminus can modulate ssDNA binding was first proposed for the single-stranded DNA-binding protein from bacteriophage T4 (gene 32 protein)\cite{434,435}. T4 gene 32 protein and the ssDNA-binding protein from bacteriophage T7 (gp2.5) share a similar functional domain arrangement with \textit{E. coli} SSB – their N-termini contain OB-folds which bind to ssDNA\cite{240,242,436}, and they have flexible, acidic C-termini required for protein-protein interactions\cite{270,271,434,437-442}. Removal of the C-terminal tails of T4 gene 32 protein\cite{441,445} or T7 gp2.5\cite{436,439-441} results in proteins with increased ssDNA-binding affinity. Consequently, models in which the C-termini act as negatively charged analogues of ssDNA, and can compete with ssDNA for the ssDNA binding sites, have been proposed for both proteins\cite{242,271,441,445}.

Chemical cross-linking experiments with T7 gp2.5 have shown that a peptide corresponding to the C-terminal tail of gp2.5 can cross-link with the core of the protein, but not when
ssDNA is bound\cite{441}. Furthermore, this cross-linking prevents ssDNA-binding, which, along with NMR data, indicates that the C-terminus indeed competes with ssDNA for binding to the same site within the N-terminal domain\cite{441}. In its native form, T7 gp2.5 is a dimer in solution\cite{241}, and its C-terminus has also been shown to play a role in stabilising the dimeric interaction, as a C-terminally truncated version of gp2.5 (gp2.5-Δ21C) exists as a monomer in solution\cite{270}. The crystal structure of gp2.5 lacking 26 residues from its C-terminus\cite{242} suggests that the C-termini within the dimer are disordered and protrude outwards from the core (as suggested for \textit{E. coli} SSB\cite{254}). The location of the truncated tails in the structure led to the proposal of a model in which the C-terminus of one monomer could bind across the dimer interface to the ssDNA binding site of the other monomer, thus stabilising the dimer\cite{242,442}. However, as is the case for \textit{E. coli} SSB and T4 gene 32 protein, definitive structural information for the C-terminus is lacking.

It seems that modulation of DNA-binding affinity by acidic C-terminal tails may be an evolutionarily conserved mechanism in proteins which bind to DNA without sequence specificity\cite{441}. Limitation of DNA-binding by the C-terminus may be important for these proteins to ensure that they retain sufficient mobility upon DNA during dynamic processes. Effectively, the C-terminus could act as an ‘electrostatic shield’ for ssDNA-binding proteins, preventing them from binding in an uncontrolled manner to negatively charged sites in cells, but allowing for coverage of exposed ssDNA as well as removal when required\cite{441}. As the C-termini of these proteins also mediate their protein-protein interactions, this competition for the ssDNA-binding site could also act as a switch, whereby ssDNA-binding displaces the C-terminal tails from the ssDNA-binding sites, making them available to interact with other proteins\cite{252,441} (see Figure 3.1).
3.1.2 Methods for studying subunit exchange in oligomeric proteins

Many multisubunit protein assemblies are far from static in their quaternary structure – rather, they are dynamic entities and are able to undergo spontaneous subunit exchange, either with themselves or with other similar proteins. Characterising the properties of this exchange can be instructive in several ways. Subunit exchange is used as a regulatory mechanism for some proteins\cite{446-450}, the nature of which may be illuminated by studying the details of exchange. Measuring rates of subunit exchange is one way of estimating the stability of oligomeric complexes. Subunit exchange rates for mutated or truncated versions of a protein may reveal details about the effects of such modifications upon the stability of its oligomeric state(s), which may, in turn, provide support for proposed structural features or insight into disease states related to oligomerisation.

As a dynamic reaction which may result in multiple products, subunit exchange can be challenging to study. Fluorescence resonance energy transfer (FRET) has been widely applied to the study of subunit exchange in multisubunit proteins\cite{451-458}. In this technique, the proteins to be examined are labelled with either donor or acceptor fluorescent probes. Subunit exchange upon mixing the differentially-labelled populations results in, on average, greater proximity between the donor and acceptor fluorophores, which can be measured in real-time as an increased FRET signal (i.e. quenched donor fluorescence and increased acceptor fluorescence). However, the presence of fluorescent labels may influence subunit exchange, and the signal obtained represents only the overall reaction – FRET is unable to directly identify the stoichiometries or abundances of individual species resulting from exchange.

Crosslinking assays\cite{251,452,459,460} and chromatography\cite{461} have also been used as methods for following subunit exchange. To examine subunit exchange by crosslinking, oligomers with distinct electrophoretic mobilities are required (which, in the case of homooligomers, requires
the use of a tag). Mixtures of these oligomers are incubated for varying lengths of time before being ‘captured’ by covalent crosslinking (either disulfide or chemically mediated). Heterooligomers resulting from subunit exchange can be detected electrophoretically, as bands of intermediate mobility. This approach has been previously used to examine subunit exchange between various C-terminally truncated SSBs, as well as to show that *E. coli* SSB is unable to form heterotetramers with human mitochondrial SSB\(^{251}\). A drawback of this method is that quantification of exchange products is generally inexact – as such, it is better suited to qualitative investigations rather than measurements of subunit exchange rates.

Chromatographic investigations of subunit exchange rely upon differences in charge or size between the subunits being exchanged. Again, however, this may require modification of native proteins. For example, ion exchange chromatography was used to study subunit exchange within homotetrameric transthyretin (TTR)\(^{461}\). To do this, a negatively charged FLAG tag (DYKDDDDKDYKDDDDK) was attached to the protein, such that tagged and native TTR tetramers, as well as heterotetramers arising from exchange, were able to be separated on an ion exchange column\(^{461}\). If there is a sufficiently large size difference between the proteins, size exclusion chromatography can be used to separate and quantify the products of subunit exchange\(^{462}\). However, the relatively large size difference required to obtain sufficient resolution between oligomers of different subunit composition limits the usefulness of this approach. In addition, the long analysis time involved in chromatographic methods makes them somewhat impractical for studying the rates of subunit exchange, except in extremely slow cases.

The rapid analysis possible with mass spectrometry makes it an excellent technique for following subunit exchange in real time. For monitoring relatively fast subunit exchange reactions, samples have been continuously electrosprayed from sources designed with thermal
Figure 3.2. Monitoring subunit exchange by mass spectrometry. Hypothetical subunit exchange experiment in which two dimeric proteins consisting of subunits differing in mass are mixed. Samples are removed from the mixture periodically and analysed by ESI-MS. The progression of subunit exchange over time can be monitored by examining the relative abundance of ions from unexchanged and exchanged species in each spectrum. These hypothetical spectra show the changes in ion abundance for a selected charge state.

control[448]. Automated sampling procedures have also been developed, allowing rapid and reproducible sampling at the initial stages of a reaction[449]. However, for slower reactions, manual sampling at successive timepoints provides sufficiently time-resolved data, and allows analysis over long periods, even when sample volumes are small. Taking discrete samples avoids potential electrochemical changes in the solution which can occur with prolonged electrospraying[463].

Mass spectrometry also offers the capacity to directly identify subunit stoichiometry and visualise the relative amounts of multiple subunit exchange products (Figure 3.2). This capacity provides a unique level of detail about a subunit exchange process, as opposed to other methods which only allow observation of the average progression of exchange. For
homooligomeric proteins, uniform isotopic labelling (for example, with $^{15}$N) provides a means of introducing a mass difference, such that ‘heavy’ (labelled) and ‘light’ (unlabelled) subunits can be resolved, but generally does not alter the ionisation efficiency/response factor of the protein[464-468]. An isotopic labelling approach also avoids the need for chemical labelling of proteins, which is generally required for other methods of following subunit exchange, such as FRET. The ability to observe subunit exchange between protein subunits in their native state, providing that MS-compatible conditions can be found, is a major benefit of a mass spectrometry approach to subunit exchange. Despite the advantages of mass spectrometry for investigating subunit exchange in multisubunit proteins, there are still relatively few studies utilising the technique for this purpose, although the number is steadily increasing. In one of the earliest applications, mass spectrometry was used to follow the exchange of subunits between the dimeric histone-like HU proteins from *Bacillus stearothermophilus* and *Bacillus subtilis*[469]. The formation of heterodimers occurred very slowly, correlating well with expectations from other techniques, and establishing mass spectrometry as a viable means for monitoring the kinetics of such reactions[469].

Several studies have used mass spectrometry approaches to examine subunit exchange in the small heat shock proteins (sHSPs) – a diverse class of chaperone proteins which form large multisubunit assemblies. The dodecameric plant sHSPs $P_b$HSP18.1 and $T_a$HSP16.9, from pea and wheat respectively, are closely homologous and their subunits are able to interact with one another. Sobott *et al.*[450] used nanoESI-MS to follow subunit exchange between $P_b$HSP18.1 and $T_a$HSP16.9. Whilst the equilibrium distribution of subunit exchange products suggested that exchange occurred statistically, without preference for any particular subunit stoichiometry, examination of the rates of formation of various heterododecamers revealed a kinetic preference for complexes containing an even number of each type of subunit. It was concluded that the mechanism of subunit exchange in these sHSPs involved the exchange of dimeric
intermediates, rather than individual monomeric units\textsuperscript{[450]}. Using nanoESI-MS in conjunction with automated sampling, another pair of dodecameric plant sHSPs, HSP18.1 and HSP17.6 from \textit{Arabidopsis thaliana}, were found to rapidly exchange dimeric units in a similar manner\textsuperscript{[449]}. Subunit exchange between diverse sHSPs expressed in the same cellular compartments is likely to be important to their function as chaperones, perhaps allowing modulation of substrate specificity and activity\textsuperscript{[449,450]}.

\(\alpha\)-Crystallin, a sHSP with a role in maintaining the transparency of the mammalian lens, is a polydisperse assembly consisting of two different subunits, \(\alpha\)A-crystallin and \(\alpha\)B-crystallin\textsuperscript{[448]}. Measurement of subunit exchange between assemblies of \(\alpha\)A and \(\alpha\)B subunits by nanoESI-MS revealed that deletion of five amino acids from the C-terminus of \(\alpha\)A-crystallin dramatically reduced the rate of subunit exchange, but did not compromise chaperone activity\textsuperscript{[448]}. The C-terminal tails of \(\alpha\)A-crystallin are highly flexible and are proposed to act as ‘keys’ for facilitating subunit exchange, which in turn is suggested to provide a means of optimising the ratio of \(\alpha\)A and \(\alpha\)B subunits in \(\alpha\)-crystallin for particular cellular requirements\textsuperscript{[448]}.

The convenience of an isotopic labelling strategy for examining subunit exchange by mass spectrometry has been demonstrated for the homotetrameric protein transthyretin (TTR), which, as a 56 kDa dimer-of-dimers, is similar in size and subunit organisation to SSB\textsuperscript{[465,466]}. TTR transports thyroxine in cerebrospinal fluid and retinol in plasma, and its disassembly and misfolding is associated with development of amyloid diseases. Subunit exchange in a TTR mutant, L55P, implicated in amyloid fibril formation, was found to proceed much more rapidly than for wild-type TTR, and exhibited an enhanced preference for exchange of dimeric intermediates\textsuperscript{[466]}. In further nanoESI-MS studies of TTR, it was shown that formation of the retinol transport complex, in which retinol is bound to TTR along with retinol-binding protein (RBP), dramatically reduced the rate of TTR subunit exchange\textsuperscript{[465]}. Both of these
studies shed light on factors involved in the stabilisation and destabilisation of TTR, with implications for development of strategies to overcome amyloid diseases.

Also associated with the formation of amyloid fibrils is β₂-microglobulin, which can accumulate during haemodialysis, leading to dialysis-related amyloidosis. Smith et al.\textsuperscript{[467]} used ESI-MS to investigate the transient oligomeric species of β₂-microglobulin present during the lag phase of amyloid fibril formation, and found populations of monomers, dimers, trimers and tetramers. When the subunit exchange properties of these oligomers were examined, it was found that increased oligomeric size corresponded to decreased subunit dynamics: all dimers fully exchanged within the dead time of MS sampling, whilst trimers and tetramers displayed biphasic exchange kinetics – certain charge states, apparently corresponding to a very dynamic population, showed rapid subunit exchange, whilst other charge states, corresponding to a more stable population, showed slower subunit exchange. For the trimer, ion-mobility mass spectrometry suggested that these populations possessed distinct conformations, and allowed proposal of a model for progressive oligomerisation at the beginning of β₂-microglobulin fibril formation\textsuperscript{[467]}.

The influence of ligand binding upon subunit exchange was probed for the homodimeric glucosamine-6 phosphate synthase (GlmS) enzyme from \textit{E. coli}\textsuperscript{[464]}. This study also used isotopic labelling to provide a mass difference and enable the observation of heterodimers by mass spectrometry. The free enzyme exhibited very slow subunit exchange kinetics (with a half-time of \(~20\) hours), but upon binding of fructose-6 phosphate or glucosamine-6 phosphate, the respective substrate and product of the enzyme, subunit exchange was completely abolished\textsuperscript{[464]}. This observation was in accord with crystallographic studies suggesting that binding of a sugar to GlmS induces a structural rearrangement within its C-terminus, whereby it interacts with the adjacent GlmS subunit, thus stabilising the dimer\textsuperscript{[464]}.
The reaction catalysed by GlmS also involves the conversion of glutamine to glutamate, both of which were found to increase the rate of GlmS subunit exchange, in contrast to the sugars. This effect was also attributed to a conformational change, in this case causing destabilisation of the dimer interface.[464]

Subunit exchange in the tetrameric tumour suppressor p53 was followed by ESI-MS and found to exhibit slow exchange kinetics.[447] Given that wild-type p53 is inherently unstable and denatures at a rate faster than its measured oligomerisation rate (a stabilised mutant was used to examine subunit exchange), the slow exchange equilibrium has been proposed as part of a regulatory mechanism for p53.[447] Furthermore, the addition of response element DNA to mixtures of wild-type and mutant p53 tetramers which had undergone subunit exchange completely abolished further exchange, allowing relative DNA-binding affinities of the hybrid tetramers to be probed[447].

Finally, in an application to a particularly large system, subunit exchange in hepatitis B virus (HBV) capsids was examined by mass spectrometry.[468] HBV capsids assemble from either 90 or 120 dimeric building blocks, and the truncated versions used in this experiment had molecular weights of 3 and 4 MDa, respectively. The larger capsid was found to undergo immeasurably slow subunit exchange with an excess of 15N-labelled dimeric units, with no exchange after 3 months of incubation. Very slow and partial subunit exchange was observed for the smaller capsid over the same period. Tandem mass spectrometry was used to detect exchange, by subjecting the capsids to collision-induced dissociation (CID; section 1.2.1) and examining the relative quantities of unlabelled and 15N-labelled monomers ejected[468].

Each of these studies, which cover a diverse range of proteins, highlight the utility of mass spectrometry for observing subunit exchange in multimeric proteins, and its potential for
uncovering mechanistic features of the exchange process via detailed examination of the kinetics of exchange, which may not be accessible using other methods.

The salt-dependence of SSB warrants investigation of the effects of salt on previously unexplored aspects of its behaviour, such as the subunit exchange dynamics. The aims of this chapter are: (1) to optimise conditions for observing the SSB tetramer by electrospray ionisation mass spectrometry, and to use these conditions to analyse the subunit exchange rate of the SSB tetramer with its $^{15}$N-labelled counterpart; (2) to investigate whether an eight residue C-terminal truncation has a measurable effect on the SSB subunit exchange rate, and (3) to observe the effects of ssDNA-binding on SSB subunit exchange.
3.2 MATERIALS AND METHODS

3.2.1 Buffers
Buffers used in this chapter were: lysis buffer (50 mM Tris-HCl, pH 8.0, 1 mM EDTA, 2 mM DTT, 2 M NaCl, 20 mM spermidine, 10% (w/v) sucrose); buffer T_{50} (50 mM Tris-HCl, pH 8.0, 1 mM EDTA); buffer T_{70} (70 mM Tris-HCl, pH 8.0, 1 mM EDTA); phosphocellulose buffer (20 mM sodium phosphate, pH 6.9, 1 mM EDTA, 1 mM DTT, 10% (v/v) glycerol); SSB storage buffer (25 mM Tris-HCl, pH 8.0, 1 mM EDTA, 300 mM NaCl, 10% (v/v) glycerol); SSBΔC8 storage buffer (50 mM Tris-HCl, pH 8.0, 1 mM EDTA, 1 mM DTT, 500 mM NaCl, 30% (v/v) glycerol).

3.2.2 Overproduction and purification of SSB
SSB was overexpressed from plasmid pND73\cite{426}, containing the ssb gene under the control of the thermoinducible bacteriophage λ P_r and P_l promotors. Strain BL21(DE3)/pND73 was grown at 30°C, with shaking, in LB medium (section 2.2.3) containing 100 µg/mL ampicillin to an A_{600} of ~0.7. The temperature was shifted to 42°C to induce overexpression, and incubation was continued for a further 3 hours. Cells were collected by centrifugation (11000 \times g; 7 min) and stored at -80°C.

After thawing, cells (1.55 g from 1 L of culture) were resuspended in lysis buffer (15 mL/g of cells) by stirring at 4°C for 1 hour. Cells were lysed by passing through a French press twice (12000 psi), and cell debris was removed by centrifugation (40 min at 38000 \times g). Streptomycin sulfate (10% (w/v) in H_{2}O) was added dropwise to the soluble fraction to a final concentration of 13 mg/mL, and the mixture stirred at 4°C for 30 min. The mixture was centrifuged to remove precipitated nucleic acids and ribosomes (80 min at 34000 \times g) before ammonium sulfate (0.5 g/mL in buffer T_{50}) was added to the supernatant to give a final

78
ammonium sulfate concentration of 0.279 g/mL. After stirring for 30 min, precipitated proteins were collected by centrifugation (30 min at \(38000 \times g\)), and resuspended in buffer \(T_{so}+300\) mM NaCl (10 mL), before being dialysed against two 3 L changes of buffer \(T_{so}\). Wild-type SSB is highly insoluble in 70 mM Tris-HCl (Prof. N. Dixon, unpublished observations) and its selective precipitation during this dialysis step affords significant purification. The dialysed sample, containing insoluble SSB, was centrifuged at \(38000 \times g\) for 30 min, and the pellet resuspended in buffer \(T_{so}+300\) mM NaCl (5 mL). The suspension was clarified by centrifuging (36000 \(\times g\); 20 min), and the supernatant was added dropwise to 60 mL of Toyopearl DEAE-650M anion-exchange resin (Tosoh Bioscience) which had been equilibrated in buffer \(T_{so}+30\) mM NaCl. The resin was stirred for 90 min at 4°C, then packed into a column (2.5 cm I.D.) and allowed to settle by gravity, before attaching to the AKTA system. The column was washed with 2 column volumes of buffer \(T_{so}+30\) mM NaCl (1 mL/min), then a linear gradient of NaCl (470 mL) from 30-600 mM in buffer \(T_{so}\) was applied. Fractions containing pure SSB, which eluted in a single peak at \(\sim 165\) mM NaCl, were pooled and dialysed against two 3 L changes of SSB storage buffer, yielding 52 mg of SSB at a concentration of 1.7 mg/mL. Aliquots were frozen at -80°C. To confirm the mass of the purified SSB, a sample was dialysed into 0.1% (v/v) formic acid to facilitate unfolding and denaturation and analysed by nanoESI-MS. The measured mass of 18840 Da was in reasonable agreement with the monomeric mass of SSB calculated from its amino acid sequence (18843.8 Da).

3.2.3 Overproduction and purification of \(^{15}\text{N}\)-labelled SSB

Uniformly \(^{15}\text{N}\)-labelled SSB (\(^{15}\text{N}\)-SSB) was overexpressed and purified as described in section 3.2.2 for unlabelled SSB, with the following exceptions. Strain BL21(DE3)/pND73 was grown in minimal medium (section 2.2.4) containing \(^{15}\text{NH}_4\)Cl (1g/L) as the sole nitrogen
source, to ensure uniform incorporation of $^{15}$N. One litre of culture generated 1.8 g of cells, which, upon purification, yielded 80 mg of $^{15}$N-SSB at a concentration of 2.4 mg/mL. The nanoESI mass spectrum of denatured $^{15}$N-SSB in 0.1% (v/v) formic acid gave a measured mass of 19085 Da. SSB contains 245 nitrogen atoms (per monomer), so a monomeric mass 245 Da greater than that of unlabelled SSB would be expected if $^{15}$N was incorporated with 100% efficiency. The mass obtained from the spectrum, 19085 Da, is 242 Da larger than the sequence-based mass of SSB, allowing calculation of an average labelling efficiency of approximately 98.8%.

### 3.2.4 Overproduction and purification of SSBΔC8 and $^{15}$N-labelled SSBΔC8

Plasmid pAL1379\(^{273}\) contains the gene for SSB without its eight C-terminal amino acids (ssbΔc8) under control of the thermoinducible p$_{R}$ and p$_{L}$ promoters. *E. coli* strain BL21(λDE3)/pAL1379 was grown, with shaking, in either LB medium (section 2.2.3) for production of SSBΔC8, or minimal medium (section 2.2.4) containing 1 g/L $^{15}$NH$_4$Cl as the sole nitrogen source for production of $^{15}$N-SSBΔC8. Upon reaching an A$_{600}$ of ~0.7, protein expression was induced by rapid heat shift to 42°C, followed by incubation at 42°C for a further 4 hours. Cells were harvested by centrifugation ($11000 \times g; 7$ min), and frozen and stored at -80°C.

Unlabelled SSBΔC8 was purified by Mr Yao Wang, using the following procedure. Thawed cells containing SSBΔC8 were resuspended in lysis buffer (15 mL/g of cells; buffer as described in section 3.2.1 but containing 500 mM rather than 2 M NaCl) before being lysed by passing twice through a French press (12000 psi). The lysate was clarified by centrifugation ($38000 \times g; 30$ min), and solid ammonium sulfate (0.15 g/mL) was added to the supernatant. After stirring at 4°C for 40 minutes, the mixture was centrifuged ($38000 \times g; 30$ min), and
the pellet was resuspended in 15 mL of phosphocellulose buffer. The resuspension was added dropwise to 50 mL of P11 phosphocellulose resin (Whatman) which had been equilibrated in the same buffer, and the mixture was stirred at 4°C for 30 minutes. After packing into a column (2.5 cm I.D.), the resin was washed with one column volume of phosphocellulose buffer (1 mL/min), then a linear gradient of NaCl (390 mL) from 0-1400 mM in phosphocellulose buffer was applied. SSBΔC8 did not bind to the resin, and was eluted in the flow through. Fractions containing partially pure SSBΔC8 were pooled, and proteins were precipitated by addition of solid ammonium sulfate (0.38 g/mL). Precipitated protein was collected by centrifugation (38000 × g; 30 min) and resuspended in SSBΔC8 storage buffer (1 mL), before being dialysed against 2 × 3 L changes of buffer T70 to selectively precipitate SSBΔC8. Pure SSBΔC8 was collected by centrifuging (15000 × g; 10 min) and resuspended in 4 mL of SSBΔC8 storage buffer, yielding 17 mg of SSBΔC8 at 4.3 mg/mL. The sample was frozen at -80°C until use.

NanoESI-MS of SSBΔC8 in 0.1% (v/v) formic acid returned a mass of 17877 Da, agreeing well with the calculated monomeric mass of 17878.8 Da. Despite the absence of a step designed to remove DNA, no DNA was detected in the native nanoESI mass spectra of SSBΔC8 (see section 3.3.3).

To purify 15N-SSBΔC8, thawed cells (5.25 g from 2 L of culture) were resuspended in lysis buffer (15 mL/g of cells) by stirring at 4°C for 1 hour. Cells were lysed by passing twice through a French press (12000 psi), and cell debris was removed by centrifugation (38000 × g; 40 min). Streptomycin sulfate (10% (w/v) in H2O) was added to the supernatant to a final concentration of 13 mg/mL, and the mixture stirred at 4°C for 30 min to allow precipitation of ribosomes and nucleic acids. Following centrifugation (35000 × g, 80 min), ammonium sulfate (0.5 g/mL in buffer T50) was added to the supernatant to a final concentration of 0.279 g/mL, and the mixture was stirred for 30 min at 4°C. Precipitated proteins were collected by
centrifugation (38000 × g; 40 min), then resuspended in 30 mL of buffer T50+300 mM NaCl and dialysed against 4 L of buffer T50+30 mM NaCl. After clarification by centrifugation (38000 × g; 30 min), this fraction was applied at a flow rate of 1 mL/min to a column (2.5×12.5 cm) of Toyopearl DEAE-650M resin, which had been equilibrated in Buffer T50+30 mM NaCl. The column was washed with 100 mL of buffer T50+30 mM NaCl, and a linear gradient (500 mL) of 30-600 mM NaCl in buffer T50 was applied. Fractions containing partially pure 15N-SSBΔC8 were pooled and clarified by centrifuging at 38000 × g for 20 min. The supernatant was dialysed against two changes of 3 L of buffer T50+30 mM NaCl. A precipitate containing pure 15N-SSBΔC8 developed during dialysis, which was collected by centrifugation and resuspended in buffer T50+500 mM NaCl (5 mL). The resuspended sample, containing 23 mg of 15N-SSBΔC8 at a concentration of 4.55 mg/mL, was frozen at -80°C.

NanoESI-MS analysis of a sample after dialysis into 0.1% (v/v) formic acid, 50% (v/v) acetonitrile gave a mass of 18111 Da for the denatured 15N-SSBΔC8 monomer, corresponding to a mass difference of 232 Da from the unlabelled protein. As SSBΔC8 contains 237 nitrogen atoms, this represents an average labelling efficiency of 97.9%.

3.2.5 Subunit exchange reactions
To monitor subunit exchange, equimolar quantities (as judged by absorbance at 280 nm) of SSB and 15N-SSB were mixed (to give total protein concentrations, as tetramers, of ~0.5 µM in 1 M NH₄OAc and ~3 µM in 10 mM NH₄OAc) and placed at 30°C to allow subunit exchange to occur. Aliquots (3 µL) were withdrawn immediately after mixing, then at regular timepoints throughout the exchange reaction, and subjected to nanoESI-MS analysis. The time between sampling and analysis was <1 min. Subunit exchange of ssDNA-bound SSB was monitored in the same way, but SSB and 15N-SSB were first bound to oligonucleotides (separately) to form complexes by mixing the required amounts of protein and ssDNA. The
complexes were analysed by nanoESI-MS to confirm their identity and stoichiometry before the unlabelled and labelled complexes were mixed. Sampling was continued for a shorter time (~6 hours) than for the non-DNA bound proteins (~10 hours) as deterioration in the spectra of SSB-oligonucleotide complexes was evident after longer periods.

3.2.6 Kinetic analysis of subunit exchange

A native mass spectrum of an equimolar mixture of SSB and 15N-SSB showed that each protein generated ions of very similar intensity, and indicated that the isotopic labelling of 15N-SSB did not alter its ionisation efficiency. It was therefore presumed that heterotetramers arising from subunit exchange between SSB and 15N-SSB would also possess essentially identical ionisation efficiencies, and the abundances of all species were able to be compared directly. The relative abundance of each homo- and heterotetramer was calculated from the intensities of its most abundant ions (15+ and 16+ for spectra recorded in 1 M ammonium acetate, or 14+, 15+ and 16+ for spectra recorded in 10 mM ammonium acetate) and expressed as a percentage of the total abundance for those ions. These data were plotted versus time using Prism version 5.0 (GraphPad Software), and fitted to either a single-exponential model or a two-phase exponential model as specified in the results. The final values were constrained to 12.5% (for homotetramers), 50% (for 3:1 and 1:3 heterotetramers) and 37.5% (for 2:2 heterotetramers).
3.3 RESULTS AND DISCUSSION

3.3.1 Optimisation of conditions for mass spectrometry of SSB proteins

In order to examine subunit exchange in the SSB tetramer by nanoESI-MS, it was first necessary to establish conditions under which SSB was both soluble and amenable to mass spectrometry, whilst preserving its native tetrameric state. Solubility trials conducted by dialysing SSB against a range of concentrations of ammonium acetate (pH 7.2) resulted in heavy precipitation of SSB under most conditions, which in turn resulted in either soluble SSB concentrations too low to acquire mass spectra, or poor quality spectra. Eventually, 1 M ammonium acetate, pH 7.2, was found to give an acceptable balance between solubility (although it was still relatively low, i.e. <1 µM) and mass spectral quality. Further investigation showed that dialysing SSB into 10 mM ammonium acetate (pH 7.2) afforded even higher SSB solubilities (up to ~3 µM) as well as good spectral quality. Finding both high and low salt conditions in which mass spectrometry of SSB could be carried out effectively was fortuitous, given its known salt-dependent behaviour, as it allowed comparison of exchange properties under different conditions, as will be described.

The nanoESI mass spectrum of SSB in 1 M ammonium acetate, pH 7.2, is shown in Figure 3.3a. Under these solution conditions, and with the conditions for mass spectrometry also optimised (section 2.6), the spectrum was exclusively populated by a distribution of ions ranging between m/z 4400 and 5400, corresponding to tetrameric SSB with a mass of 75431.8 Da (compared to a calculated sequence-based mass for the tetramer of 75375.2 Da\textsuperscript{[470]}). The peak width at half height of the 15+ ion was 5 Thoms. The difference between the calculated and measured mass can likely be attributed to the presence of water molecules or other ions trapped within the native tetrameric structure\textsuperscript{[470]}. Given that SSB has been observed to maintain its tetrameric structure at concentrations as low as 30 nM (tetramer)\textsuperscript{[248,255,259]}, it
was expected that essentially all the SSB would retain its tetrameric structure under these conditions. Despite the low SSB concentrations attainable, the spectral quality was high. The mass spectrum of SSB (~2 µM) in 10 mM ammonium acetate, pH 7.2 (Figure 3.3b) possessed a very similar charge state distribution to the spectrum in 1 M ammonium acetate. Again, the tetramer was the only species observed when mass spectrometry conditions were optimised, with a calculated mass of 75398.7 Da. The higher SSB concentrations achievable with this solution condition afforded higher ion abundances which were important for later experiments involving oligonucleotide binding (section 3.3.4).

The 15N-labelled form of SSB was subjected to nanoESI-MS analysis under the same conditions. The nanoESI mass spectra of 15N-SSB in 1 M and 10 mM ammonium acetate are shown in Figures 3.3c and 3.3d, respectively. As observed for unlabelled SSB, the ions in these spectra
were exclusively those of the native tetramer, corresponding to masses of 76370.3 Da (in 1 M ammonium acetate) and 76378.4 Da (in 10 mM ammonium acetate). The calculated mass of the $^{15}$N-SSB tetramer, based on a labelling efficiency of 98.8% (section 3.2.3), is 76343.2 Da, so these measured values were in good agreement. The difference in mass between unlabelled SSB and $^{15}$N-SSB tetramers is almost 1000 Da: readily allowing resolution of the tetramers themselves, as well as resolution of all three potential heterotetramers arising as a result of subunit exchange. In addition, the ionisation efficiencies and charge state distributions of unlabelled SSB and $^{15}$N-SSB were essentially identical under equivalent conditions, which was important for quantifying subunit exchange.

3.3.2 Subunit exchange between SSB and $^{15}$N-SSB

Subunit exchange between SSB and $^{15}$N-SSB has the potential to form three different heterotetrameric species, each with a unique mass: $(SSB)_3(^{15}N$-$SSB)_1$ (denoting a tetramer containing three unlabelled SSB subunits and one $^{15}$N-labelled subunit), $(SSB)_2(^{15}N$-$SSB)_2$, and $(SSB)_1(^{15}N$-$SSB)_3$. For brevity, these will be referred to as 3:1, 2:2 and 1:3 heterotetramers, respectively, and the SSB and $^{15}$N-SSB homotetramers denoted 4:0 and 0:4, respectively. Given the known salt-dependence of SSB in many of its activities, the subunit exchange properties of SSB were investigated in both the high and low ammonium acetate concentrations described in the previous section.

In an initial experiment, equimolar amounts of unlabelled and $^{15}$N-labelled SSB in 1 M ammonium acetate were mixed and placed at 4°C, with aliquots periodically removed and subjected to nanoESI-MS. However, subunit exchange was so slow at this temperature that heterotetrameric peaks were barely detectable 10 hours after mixing (not shown). By increasing the temperature to 30°C, the rate of subunit exchange was sufficiently increased that it could be monitored over a period of hours. Raising the temperature further, to 37°C,
Figure 3.4. Subunit exchange between SSB and $^{15}$N-SSB monitored by nanoESI-MS. (a) In 10 mM ammonium acetate, pH 7.2 and (b) in 1 M ammonium acetate, pH 7.2. Equimolar quantities of each protein were mixed and placed at 30°C, with aliquots withdrawn at the times indicated and analysed by nanoESI-MS. Total protein concentrations were ~3 µM in reactions performed in 10 mM ammonium acetate and ~0.5 µM in reactions performed in 1 M ammonium acetate. Dotted lines through the highest intensity charge state show the identity of each tetramer. The other charge states shown contain analogous distributions.

resulted in poorly resolved nanoESI mass spectra which deteriorated rapidly (not shown).

Subunit exchange was monitored by acquiring nanoESI mass spectra immediately after mixing equimolar SSB and $^{15}$N-SSB, then every hour (or in some cases, more frequently) for 10 hours during incubation of the mixture at 30°C. Figure 3 shows the most abundant charge states in
the spectra acquired directly after mixing, and after 120, 360, and 540 minutes of incubation, in either 10 mM ammonium acetate (Fig 3.4a) or 1 M ammonium acetate (Fig 3.4b). The spectra acquired immediately after mixing, labelled ‘0 min’ in Figure 3.4, consisted solely of ions corresponding to homotetrameric SSB and 15N-SSB, indicating that no appreciable subunit exchange occurred within the ‘dead time’ of mixing and acquisition (~1 min). With increasing time, a slow decay in the abundance of both homotetramers, and simultaneous increase in abundance of the three hybrid tetramers resulting from subunit exchange, was observed at both ammonium acetate concentrations. After 540 minutes of exchange, the dominant species in the spectra under both conditions was the 2:2 heterotetramer consisting of 2 ‘heavy’ and 2 ‘light’ subunits.

Thus, under these conditions, SSB undergoes spontaneous exchange of subunits, but over a relatively extended time period. This slow rate of subunit exchange, at both ammonium acetate concentrations, reflects the strength of the subunit interactions in the SSB tetramer. After 10 hours of exchange, the distribution of ion abundances for each tetrameric species (4:0, 3:1, 2:2, 1:3, 0:4) appeared to be approaching the binomial 1:4:6:4:1 distribution expected at equilibrium for a statistical exchange of subunits. However, closer examination of the kinetics of exchange revealed some differences between the two salt concentrations. In 1 M ammonium acetate, the disappearance of homotetramers over time was well-described by a first-order exponential decay model (Figure 3.5, open squares), with a rate constant of 0.00344 ± 0.00005 min⁻¹ (Table 3.1), corresponding to a half-time for exchange of ~200 minutes. The exponential fit was constrained to plateau at 12.5% to reflect its expected value at equilibrium based on a binomial distribution. In 10 mM ammonium acetate, the rate of decrease in homotetramer abundance was not well described by a first-order exponential – it could, however, be fit to a two-phase exponential decay (Figure 3.6, open circles), with fast
Figure 3.5. SSB subunit exchange rates in 1 M ammonium acetate. Relative homotetramer intensities were calculated from nanoESI-MS spectra at each timepoint by dividing the total intensity arising from homotetrameric ions by the total intensity of all ions (for the most intense charge states). Open squares (□) show the progression of subunit exchange between SSB and 15N-SSB, fit to a first order exponential decay equation. Filled squares (■) show the progression of subunit exchange between SSBΔC8 and 15N- SSBΔC8, fit to a two-phase exponential decay equation. Rate constants are shown in Table 3.1.

and slow rate constants of $0.0126 \pm 0.0017 \text{ min}^{-1}$ and $0.00224 \pm 0.00023 \text{ min}^{-1}$, respectively (Table 3.1). Direct comparison of the plots showing the loss of homotetramer abundance shows that the overall rates of exchange in 1 M and 10 mM ammonium acetate were not substantially different, however the biphasic exchange kinetics displayed in 10 mM ammonium acetate indicated that there were underlying differences in the exchange process at high and low salt concentration.
Relative homotetramer intensities were calculated from nanoESI-MS spectra at each timepoint by dividing the total intensity arising from homotetrameric ions by the total intensity of all ions (for the most intense charge states). Open circles (○) show the progression of subunit exchange between SSB and 15N-SSB. Filled circles (●) show the progression of subunit exchange between SSBΔC8 and 15N-SSBΔC8. Both sets of data were fit to two-phase exponential decay equations. Rate constants are shown in Table 3.1.

Table 3.1. Rate constants for decay of homotetrameric species during SSB subunit exchange. For the constants derived from two-phase exponential fits, percentages in brackets refer to the proportion of the overall rate contributed by each rate constant.

<table>
<thead>
<tr>
<th></th>
<th>Full-length SSB</th>
<th>SSBΔC8</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1 M NH₄OAc</strong></td>
<td>( k = 0.00344 \pm 0.00005 \text{ min}^{-1} )</td>
<td>( k_{fast} = 0.00669 \pm 0.00162 \text{ min}^{-1} ) (69.9 ± 22.4%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( k_{fast} = 0.0126 \pm 0.0017 \text{ min}^{-1} ) (43.6 ± 5.6%)</td>
</tr>
<tr>
<td><strong>10 mM NH₄OAc</strong></td>
<td>( k_{fast} = 0.00224 \pm 0.00023 \text{ min}^{-1} ) (56.4 ± 7.2%)</td>
<td>( k_{slow} = 0.00353 \pm 0.00031 \text{ min}^{-1} ) (21.7 ± 0.5%)</td>
</tr>
</tbody>
</table>
Also notable in all spectra is the absence of any monomeric or dimeric species, suggesting that
the rate of exchange is limited by tetramer dissociation: the intermediate species involved in
exchange re-associate much faster than the tetramer dissociates. Such a situation is to be
expected for an oligomer with a driving force for assembly as strong as the SSB tetramer. To
further examine this, it was desirable to follow exchange over a range of initial SSB
concentrations. The low SSB concentrations achievable in 1 M ammonium acetate precluded
any further dilution, but in 10 mM ammonium acetate, it was possible to observe subunit
exchange over a modest (four-fold) SSB concentration range. Figure 3.7 shows the rates of loss
of homotetramer intensity over 10 hours with initial total homotetramer concentrations of 3
µM, 1.5 µM and 0.75 µM. The subunit exchange rates were essentially indistinguishable,
providing further evidence that tetramer dissociation determines the overall rate of exchange
(if reassociation rates of intermediates were a factor, concentration dependence would be
expected). Given the solubility profile of SSB, it was not possible to achieve concentrations
substantially higher than 3 µM. However, the relatively low concentrations minimise the
potential for non-specific associations of protein subunits in the ionisation source of the mass
spectrometer.

One way of explaining the salt-dependent behaviour of the subunit exchange rates involves
consideration of the potential contribution of the C-terminal tail of SSB to subunit exchange.
As described in the introduction, recent evidence has suggested that the C-terminus is capable,
under some conditions, of occupying the DNA-binding site\cite{196,252}. Such an interaction can be
visualized in several ways. Within the tetramer, the C-termini of each subunit could fold back
and occupy the OB-fold of the same subunit (Figure 3.8a), or, alternatively, could bind in trans
to the DNA-binding site of an adjacent subunit. The latter configuration could occur in
multiple ways (Figure 3.8b illustrates two possibilities), and would be expected to contribute
some additional stability to the tetramer. Therefore, the two distinct subunit exchange rates
Figure 3.7. The SSB subunit exchange rate is independent of protein concentration. The progression of homotetramer loss is plotted against time for the exchange of SSB and $^{15}$N-SSB in 10 mM ammonium acetate, with protein concentrations of 3 µM (blue circles), 1.5 µM (green squares) or 0.75 µM (pink triangles). Each data set was fit to a two phase exponential decay equation.

observed in 10 mM ammonium acetate could potentially arise from two distinct populations of tetramers with their C-termini bound in different ways. The single rate observed in 1 M ammonium acetate could reflect the disruption of the C-terminus-ssDNA site interaction which is likely in high salt, resulting in a single population of tetramers which are not stabilised by such an interaction. If this is the case, it should be noted that the rate of exchange observed in 1 M ammonium acetate was similar to the slower of the two rates in 10 mM ammonium acetate. This initially seems at odds with the prediction above. However, hydrophobic interactions, which are strengthened by increasing salt concentration$^{[471,472]}$, make a significant contribution to SSB tetramer formation$^{[229]}$. Stabilisation of these interactions in 1 M ammonium acetate relative to 10 mM ammonium acetate could therefore account for the
similar exchange rates. The existence of two or more populations of SSB tetramers in 10 mM ammonium acetate, defined by the behaviour of their C-terminal tails, would suggest that truncation of the C-terminus should give rise to a single population, with exchange kinetics similar to the more rapidly exchanging full-length population due to a loss of stability. To test this hypothesis, the truncated SSB mutant SSBΔC8 was produced. SSBΔC8 is missing the eight C-terminal residues from the C-terminus of SSB – the region implicated in most SSB-protein interactions as well as the potential interaction with the ssDNA-binding site. If such an interaction is causing stabilisation of the SSB tetramer at low salt concentrations, deletion of this region would be expected to eliminate the biphasic subunit exchange kinetics, and a single, faster subunit exchange rate would be expected. Experiments monitoring subunit exchange in the truncated mutant are described in the next section.
3.3.3 Subunit exchange between SSBΔC8 and 15N-SSBΔC8

SSBΔC8 and its 15N-labelled counterpart were exclusively tetrameric when subjected to nanoESI-MS under the same conditions as those used for the full-length proteins, and like the full-length proteins, displayed extremely similar ionisation behaviour (Figure 3.9). Masses obtained for SSBΔC8 (calculated mass 71515.2 Da) were 71555.1 and 71537.1 Da, in 1 M and 10 mM ammonium acetate, respectively (Figure 3.9 a, b). The masses obtained for 15N-SSBΔC8 (calculated mass 72444 Da) were 72476.7 and 72471.3 Da, in 1 M and 10 mM ammonium acetate, respectively (Figure 3.9 c, d). Subunit exchange between SSBΔC8 and 15N-SSBΔC8 was analysed as for the full-length proteins, in both 1 M and 10 mM ammonium acetate, and selected spectra from timepoints during the subunit exchange reactions are shown in Figure 3.10. It was immediately obvious from examination of these spectra that, while subunit exchange between the truncated proteins in 1 M ammonium acetate appeared to proceed at a relatively similar rate as that observed for the full-length SSBs, exchange in 10 mM ammonium acetate proceeded much more rapidly, appearing almost complete after just 2 hours.

Further analysis of the SSBΔC8 exchange rates revealed some interesting results. Although the rate of exchange in 1 M ammonium acetate showed little overall difference to that of its full-length counterpart under the same conditions (Figure 3.5; Table 3.1), the rate was no longer well described by a simple single exponential model, suggesting that the interpretation of a single population giving rise to a first-order rate of exchange represents an oversimplification of the situation. Consistent with this (and contrary to the original prediction), the rate of exchange in 10 mM ammonium acetate also retained biphasic kinetics even in the absence of
Figure 3.9. Native nanoESI mass spectra of SSBΔC8 and 15N-SSBΔC8 tetramers. (a) SSBΔC8, after dialysis into 1 M ammonium acetate, pH 7.2 (b) SSBΔC8, after dialysis into 10 mM ammonium acetate, pH 7.2 (c) 15N-SSBΔC8 after dialysis into 1 M ammonium acetate, pH 7.2 (d) 15N-SSBΔC8 after dialysis into 10 mM ammonium acetate, pH 7.2. The measured masses of the SSBΔC8 tetramer were 71555.1 Da and 71537.1 Da, in 1 M and 10 mM ammonium acetate, respectively (theoretical mass 71515.2 Da). The measured masses of the 15N-SSBΔC8 tetramer were 72476.7 Da and 72471.3 Da, in 1 M and 10 mM ammonium acetate, respectively (theoretical mass 72444 Da).

the C-terminus, with fast and slow rate constants of 0.0260 ± 0.0009 min⁻¹ and 0.00353 ± 0.00031 min⁻¹, respectively (Figure 3.6; Table 3.1). The rate constant for the fast phase was approximately two-fold faster than the corresponding rate constant for the full-length proteins, and, importantly, made a much larger contribution to the overall rate (78% vs 43%; see Table 3.1). A visual comparison of the truncated and full-length exchange rates allows an appreciation of the distinctly faster overall rate of exchange between the truncated variants (Figure 3.6). Thus, the C-terminal tail of SSB undoubtedly stabilises the SSB tetramer with respect to exchange at low salt, but the persistence of biphasic kinetics in its absence suggests that the mechanism by which this occurs is more complex than suggested in our simple stabilised/non-stabilised model above. Nonetheless, the fact that the stabilisation is selectively
Figure 3.10. Subunit exchange between SSBAC8 and $^{15}$N-SSBAC8 monitored by nanoESI-MS. (a) In 10 mM ammonium acetate, pH 7.2, (b) in 1 M ammonium acetate, pH 7.2. Equimolar quantities of each protein were mixed and placed at 30°C, with aliquots withdrawn at the times indicated and analysed by nanoESI-MS. Total protein concentrations were ~3 µM in reactions performed in 10 mM ammonium acetate and ~0.5 µM in reactions performed in 1 M ammonium acetate. Dotted lines through the highest intensity charge state show the identity of each tetramer. The other charge states shown contain analogous distributions.

observed in low salt supports the hypothesis that it is a result of electrostatic contacts between the C-terminus and the more positively charged OB-fold core. The actual range of potential C-terminus/OB-fold interactions possible within the SSB tetramer makes it likely that several ‘modes’ of binding may contribute to the observed stabilisation: our data do not allow us to
resolve these multiple possibilities. The non-equivalence of the two dimer interfaces in SSB (see Figure 1.11) adds further complexity: stabilising contacts made by the C-terminus may occur preferentially across one interface, changing their relative probability of dissociation. It is likely that a combination of these factors contribute to the complexity of the observed subunit exchange kinetics. Nonetheless, when taken together with the recent evidence that the SSB C-terminus is able to interact with the OB-fold domain of SSB, this salt-dependent modulation of the subunit exchange rate by the C-terminal residues of SSB provides strong evidence that the C-terminus of SSB, via an electrostatic interaction with the more positively charged SSB core, exerts a moderately stabilising effect on the SSB tetramer. The fact that subunit exchange does proceed essentially to completion under all conditions, even in the presence of the proposed stabilising interactions, suggests that these interactions are likely to be relatively transient. Given that each subunit of the tetramer is identical, it is probable that interactions of the C-termini with the OB-folds of both their own and adjacent subunits occur, dependent on the range of movements explored by the flexible linker region of the C-terminal domain. A similar mechanism has been proposed for the dimeric SSB protein from T7 bacteriophage, gp 2.5. The C-terminus of gp2.5, like that of E. coli SSB, is highly acidic, and has been shown to compete with ssDNA for the DNA-binding surface of the protein\textsuperscript{[441]}. Furthermore, this interaction has been proposed to occur in trans between subunits and stabilise the dimer\textsuperscript{[242]}.

### 3.3.4 Subunit exchange between SSB and $^{15}$N-SSB in the presence of ssDNA

Having established that SSB tetramers were able to undergo slow subunit exchange, it was desirable to examine the effect of DNA-binding upon this exchange. Given the complexity with which SSB binds to DNA (section 1.1.5.2), it was first necessary to establish that it was possible to observe SSB-DNA complexes by ESI-MS which were consistent with those
Figure 3.11. SSB-oligonucleotide complexes observed by nanoESI-MS. (a) Mixing an excess of (dT)$_{70}$ with SSB results in a 1:1 SSB:(dT)$_{70}$ complex. The measured mass of this complex was 96651 Da (theoretical mass 96604 Da). (b) An excess of (dT)$_{35}$ mixed with SSB gives rise to a 1:2 SSB:(dT)$_{35}$ complex. The measured mass of this complex was 96654 Da (theoretical mass 96542 Da). Ions marked with black circles (●) arise from a small amount of 1:1 SSB:(dT)$_{35}$ complex. (c) Mixing equimolar quantities of SSB and (dT)$_{35}$ allows observation of a 1:1 SSB:(dT)$_{35}$ complex. The measured mass of this complex was 85997 Da (theoretical mass 85957 Da). All mixtures were in 10 mM ammonium acetate, pH 7.2.
expected from solution studies. Fluorescence quenching studies have shown that, maximally, four molecules of (dT)$_{16}$, two molecules of (dT)$_{15}$, or one molecule of (dT)$_{70}$ are able to bind to a single SSB tetramer$^{[266,268]}$. It should also be noted that most previous studies of SSB-ssDNA interactions have used poly(dT) sequences, due to an observed preference of SSB to bind to such sequences$^{[259]}$. For consistency, (dT) homooligonucleotides have been used for all SSB binding studies described in this thesis.

When SSB was mixed with an excess of (dT)$_{70}$ or (dT)$_{35}$ in 10 mM ammonium acetate, complexes containing a maximum of one (dT)$_{70}$ strand (Figure 3.11a), or two (dT)$_{35}$ strands (Figure 3.11b) per SSB tetramer were observed in the nanoESI mass spectra. These maximum stoichiometries were also observed when the mixtures were formed in 1 M ammonium acetate (not shown). Being able to observe stoichiometries matching those predicted is an indication that SSB-ssDNA complexes formed in solution are well preserved in the gas phase. By decreasing the ratio of (dT)$_{35}$ to SSB, a 1:1 SSB:(dT)$_{35}$ complex was also able to be observed (Figure 3.11c). The 1:1 SSB complexes of SSB with (dT)$_{70}$ and (dT)$_{35}$ have commonly been used as models for the (SSB)$_{65}$ and (SSB)$_{35}$ binding modes, respectively$^{[196,264,266,268,322,325]}$.

To analyse SSB subunit exchange in the presence of single-stranded DNA, SSB and $^{15}$N-SSB were premixed (separately) with either (dT)$_{70}$ or (dT)$_{35}$ to form complexes, which were then mixed in equimolar quantities and incubated at 30°C, analogous to the protein-only experiments (section 3.3.2). The mixture was sampled periodically and analysed by nanoESI-MS for evidence of subunit exchange. A shorter total incubation time of 6 hours was used, as it was difficult to maintain spectral quality after longer periods in some cases. Three complexes were examined: a 1:1 complex of SSB/$^{15}$N-SSB with (dT)$_{70}$, a 1:1 complex of SSB/$^{15}$N-SSB with (dT)$_{35}$, and a 1:2 complex with (dT)$_{35}$ (two (dT)$_{35}$ strands bound to SSB/$^{15}$N-SSB). Before
Figure 3.12. Monitoring subunit exchange between SSB and 15N-SSB bound to oligonucleotides. Preformed complexes of (a) 1:1 SSB/15N-SSB with (dT)$_{70}$ (b) 1:1 SSB/15N-SSB with (dT)$_{35}$ and (c) 1:2 SSB/15N-SSB with (dT)$_{35}$ were mixed and incubated at 30°C (in 10 mM ammonium acetate, pH 7.2). The total protein concentration in each case was $\sim$3 µM. For each mixture, nanoESI-MS spectra were recorded immediately after mixing, then periodically for 360 minutes. Only the initial (‘0 min’) and final (‘360 min’) spectra are
shown. The dotted lines designate the identity of the complexes for a single charge state in each spectrum. The other charge states are similarly assigned. The measured mass of the 1:1 $^{15}$N-SSB:(dT)$_{70}$ complex was 97665 Da (theoretical mass 97572 Da). The measured mass of the 1:1 $^{15}$N-SSB:(dT)$_{35}$ complex was 86989 Da (theoretical mass 86925 Da). The measured mass of the 1:2 $^{15}$N-SSB:(dT)$_{35}$ complex was 97626 Da (theoretical mass 97510 Da). For masses of unlabelled SSB complexes, refer to Figure 3.11.

mixing the SSB and $^{15}$N-SSB complexes, they were analysed separately by nanoESI-MS to confirm their stoichiometry and the absence of free protein.

Upon mixing equimolar quantities of the 1:1 SSB:(dT)$_{70}$ complex and the 1:1 $^{15}$N-SSB:(dT)$_{70}$ complex (in 10 mM ammonium acetate), both complexes could be clearly observed in the spectrum (Figure 3.12a, bottom spectrum). Interestingly, the spectrum after 6 hours at 30°C remained essentially unchanged from the spectrum obtained immediately after mixing (Figure 3.12a, top spectrum). This lack of observed subunit exchange peaks was in direct contrast with the extensive exchange seen between the free proteins after a comparable period (Figure 3.4a). Spectra deteriorated after further incubation, but these results were nonetheless sufficient to show a dramatic inhibition of subunit exchange when SSB is bound to (dT)$_{70}$. The extensive wrapping of SSB by (dT)$_{70}$ in the (SSB)$_{65}$ binding mode (Figure 1.12a) provides an intuitive explanation for this observed inhibition of subunit exchange by (dT)$_{70}$ binding. The high affinity of the SSB-(dT)$_{70}$ interaction, combined with the relatively slow rate of subunit exchange observed in the free tetramer, could be predicted to prevent dissociation of the complex for long enough for exchange to occur. However, a similar absence of subunit exchange was observed with both 1:1 and 1:2 complexes of SSB and $^{15}$N-SSB with (dT)$_{35}$. In both cases, the starting complexes (Figures 3.12b and 3.12c, bottom spectra) remained unchanged after 6 hours at 30°C (Figure 3.12b and 3.12c, top spectra). In the 1:1 SSB:(dT)$_{35}$ complex, which is presumed to utilise the (SSB)$_{35}$ binding mode, only two of the subunits of the tetramer are expected to be occupied by DNA (Figure 1.12b). It was therefore anticipated
that the tetramer would be freer to undergo subunit exchange than when it is fully wrapped by (dT)₇₀. Nonetheless, the inhibition of subunit exchange was just as complete as it was with the more highly occluded complexes.

The stabilising effect of ssDNA-binding on the SSB tetramer is much greater than that provided by the C-terminus, reflecting the relative weakness of the proposed C-terminus-SSB interaction relative to that of the SSB-ssDNA interaction. The model in Figure 1.12b shows that the path of the ssDNA in the (SSB)₃₅ mode is predicted to traverse the ‘dimer-dimer’ (rather than the monomer-monomer) interface of the tetramer. The fact that binding of a single (dT)₁₅ molecule is sufficient to completely abolish exchange suggests that the mechanism of SSB subunit exchange very likely proceeds via dissociation to dimers across this interface as a first step, with further dissociations only possible after this has occurred. It therefore also rules out a mechanism of exchange whereby individual subunits of SSB are able to be removed from the tetramer and exchange in an isolated manner.

DNA-binding has been observed to affect the subunit exchange properties of several DNA-binding proteins, especially those with gene regulatory functions. A nucleoid-associated protein, Fis [462], the DNA-binding/bending protein TF1 [473] and the transcription factors C/EBP [474] and Fos/Jun [475] are dimeric complexes which exhibit rapid subunit exchange [452, 474, 475] that is dramatically slowed by binding to their specific (and in some cases, nonspecific) DNA sequences. DNA-binding has also been shown to have a destabilising effect on some oligomeric proteins, such as cAMP receptor protein (CRP), a dimer which undergoes slow subunit exchange in the absence of DNA, but greatly accelerated subunit exchange upon binding to nonspecific DNA [476].
3.4 CONCLUSIONS

This work has demonstrated that the C-terminus of SSB exerts a modest, salt-dependent effect on the stability of the *E. coli* SSB tetramer, suggesting for the first time that the C-terminus may play a role in SSB tetramer formation. NanoESI-MS was used as a sensitive and direct method to follow subunit exchange in the SSB tetramer in real time under high and low salt solution conditions. The rate of exchange under low salt conditions was dependent on the presence of the C-terminal eight residues of SSB: removal of the C-terminus increased the rate of exchange, indicating some destabilisation of the tetramer. Previously, the SSBΔC8 tetramer was thought to form equally as efficiently as the tetramer comprised of the full-length protein, and the C-terminus was assumed to be extraneous to tetramer stability. This work therefore shows how examination of subunit exchange processes can expose differences between very stable oligomeric proteins which otherwise appear to be equivalent. It is proposed that the increased stability provided to the SSB tetramer by its C-terminal tails results from intersubunit interactions between these tails and the OB-folds of adjacent subunits. Understanding the behaviour of the SSB C-terminus is of critical importance, as its interactions with a variety of binding partners form a central part of the cellular role of SSB. Interactions between the C-terminus and the OB-fold domain of SSB may prove to be important to the regulation of these interactions, or indeed to the regulation of the SSB-ssDNA interaction. Further work will be required to probe the precise nature of the behaviour of the C-terminus and of any such intersubunit contacts. This work has also shown the utility of mass spectrometry for probing the effect of ssDNA-binding upon SSB subunit exchange. Binding of oligonucleotides to the SSB tetramer was shown to substantially inhibit subunit exchange such that it was not observable within the timescale examined. Even one molecule of (dT)$_{35}$, which occludes only two subunits of the tetramer, was sufficient to exert this inhibition, suggesting that subunit exchange proceeds via dissociation of the dimer-dimer interface as a first step.
Chapter 4: **EXAMINATION OF SSB-ssDNA COMPLEXES BY MASS SPECTROMETRY AND SURFACE PLASMON RESONANCE**

### 4.1 Introduction

The role of SSB as a ubiquitous mediator of processes involving exposed ssDNA leads to a question: how does a protein with such a high affinity for DNA binding achieve mobility on DNA? For example, during replication, SSB needs to be rapidly displaced as the replication machinery progresses (at \( \sim 1000 \) nucleotides/second)\(^7\), and re-bind to newly-formed lagging strand ssDNA. In recombinational repair, SSB initially acts to remove secondary structure from ssDNA, facilitating RecA filament formation\(^{288,477}\) – it therefore needs to be progressively removed from ssDNA as the RecA filament advances, and transferred to the displaced ssDNA strand formed during strand exchange\(^{478}\).

Given these requirements, along with the stability of SSB-ssDNA complexes\(^{322}\), it has been predicted for some time that SSB may be recycled from one site to another, without fully dissociating from DNA at any point. Such a system could potentially account for the speed with which SSB is evidently displaced by the action of other proteins, as well as ensuring that newly exposed sites are coated efficiently. One idea, first proposed by Römer *et al.* in 1984\(^{321}\), was that SSB tetramers may be able to ‘roll’ or ‘slide’ along ssDNA strands by progressive unwrapping/rewrapping of the tetramers by DNA. Recent single-molecule FRET studies have shown that SSB is able to diffuse along ssDNA\(^{324}\) using a ‘reptation’ mechanism in which small ssDNA bulges are propagated around the SSB tetramer\(^{136}\) (Figure 4.1). This diffusion has been shown to facilitate the formation of RecA filaments\(^{324}\), suggesting that it may potentially play a role in SSB repositioning during recombination events *in vivo*\(^{124,136}\). A second mechanism, whereby intact SSB tetramers can rapidly transfer directly between
Figure 4.1. SSB can diffuse along ssDNA by reptation. In this model, dissociation of a small length of DNA from the SSB tetramer can result formation of a small bulge (~3 nt). Propagation of this bulge around the SSB tetramer via a random walk results in a net movement of the DNA relative to SSB. The blue asterisk represents the position of a single nucleotide. Figure reproduced from reference 326.

distinct ssDNA sites (not necessarily sequentially adjacent), without passing through a free SSB intermediate, has also been proposed as a prospective mechanism for SSB recycling\cite{214,325}. The kinetics of this ‘direct transfer’ mechanism have been examined in detail in vitro by Kozlov and Lohman\cite{325}. Using stopped-flow fluorescence quenching, transfer of SSB from ‘donor’ to ‘acceptor’ ssDNA molecules was found to occur via the transient formation of ternary intermediates containing both ssDNA molecules bound simultaneously to a single SSB tetramer (Figure 4.2a). Subsequent dissociation of the ‘donor’ strand resulted in SSB being transferred to the acceptor molecule without ever passing through a free SSB intermediate\cite{325}. The identity of the initial SSB-ssDNA complex was shown to be crucial: the rate of transfer was highly dependent on whether or not there were unoccupied binding sites available on the SSB tetramer to accept the second strand\cite{325}. Therefore, it was concluded that transfer occurs most efficiently when the initial complex is in the (SSB)$_{15}$ mode, in which two subunits are unoccupied (see Figure 1.12). As the (SSB)$_{15}$ mode has been proposed to function in replication\cite{265}, the direct transfer mechanism has been proposed to represent a potential
Figure 4.2. Direct transfer of SSB tetramers between oligonucleotides. (a) Transfer from a donor to an acceptor oligonucleotide occurs via an intermediate ternary complex in which both ssDNA strands are simultaneously bound to the tetramer. In the example here, a large excess of unlabelled (dT)$_{70}$ is added to a 1:1 complex of SSB with fluorescently labelled (dT)$_{70}$ (SSB:(dT)$_{70}^*$). After passing through a transient intermediate state in which both the fluorescent and unlabelled (dT)$_{70}$ strands are bound, the fluorescent (dT)$_{70}$ dissociates, leaving a 1:1 complex of SSB with the unlabelled (dT)$_{70}$. (b) Such a mechanism could potentially be used during DNA replication for SSB transfer on the lagging strand. In this model, as the advancing lagging strand polymerase fills an Okazaki fragment, SSB tetramers from the lagging strand ssDNA ahead of the polymerase (in green) could be transferred to the newly generated lagging strand ssDNA. Several replication fork components are omitted for clarity (see Figure 1.2). Based on figures from reference 325.
mechanism operating during replication (Figure 4.2b). In such a mechanism, SSB, bound in the (SSB)$_{15}$ mode to the lagging strand, could be transferred from the site of a completed Okazaki fragment to the newly generated lagging strand DNA via direct transfer. Given the high level of inter-tetramer cooperativity in the (SSB)$_{15}$ mode\textsuperscript{[265]}, this transfer could potentially occur in a cooperative manner, with a tract of adjacent SSB tetramers being transferred simultaneously.

As described in Chapter 3 (section 3.3.4), mass spectrometry was used successfully to observe SSB-oligonucleotide complexes with stoichiometries consistent with those observed in solution phase studies\textsuperscript{[266,268]}. In this chapter, the capability of ESI-MS for examining SSB-ssDNA complexes was probed further by examining direct transfer of SSB between oligonucleotides, and comparing the results to those predicted from previous work\textsuperscript{[325]}. In addition, surface plasmon resonance (SPR) was used to examine direct transfer and the results were compared with those obtained using mass spectrometry. Surface plasmon resonance experiments form the basis of Chapter 7 of this thesis, and a description of SPR principles and instrumentation is presented in section 7.1.
4.2 **MATERIALS AND METHODS:**

4.2.1 **Proteins and oligonucleotides**

SSB was purified as described in section 3.2.2 and stored at -80°C, then dialysed against the specified buffer before use. The single-stranded deoxyribooligonucleotides (dT)$_{15}$ and (dT)$_{20}$, along with their 5′-biotinylated counterparts, were obtained HPLC-purified from Geneworks (Adelaide, Australia), and dialysed into the appropriate buffer prior to use in individual experiments.

4.2.2 **Mass spectrometry competition experiments.**

All SSB-oligonucleotide complexes were formed by first dialysing SSB and oligonucleotides separately into either 10 mM or 1 M ammonium acetate, pH 7.2, then mixing appropriate ratios of SSB and oligonucleotides to form specific complexes (as specified in the text). Complexes were analysed by nanoESI-MS using the conditions described in section 2.6.2. For competition experiments, initial complexes were challenged by addition of an equimolar quantity of a second oligonucleotide, then immediately re-analysed by nanoESI-MS.

4.2.3 **Surface plasmon resonance competition experiments.**

Surface plasmon resonance experiments to examine direct transfer of SSB between DNA strands were conducted using a Biacore T100 instrument (GE Healthcare, UK) equipped with a SA (streptavidin-coated) sensor chip. The streptavidin allows essentially irreversible attachment of biotinylated molecules to the surface of the sensor chip. All experiments were carried out at 20°C with a flow rate of 5 µl/min. The chip was prepared by injecting 70% (v/v) glycerol for 1 min, followed by three 1 min injections of 1M NaCl, 50mM NaOH. All experiments were carried out in Biacore buffer (10mM HEPES, pH 7.4, 150 mM NaCl, 0.05% (vol/vol) surfactant P20 (GE Healthcare, UK)). Oligonucleotides (5′-biotinylated
(dT)_{70} and (dT)_{15} were diluted to 5 nM in Biacore buffer and immobilized onto separate flow cells of the sensor chip to levels of ~200 RU and ~25 RU, respectively. SSB was dialysed into Biacore buffer and diluted to 90 nM before being injected over the immobilised oligonucleotides for 5 minutes, which was sufficient to saturate the immobilised oligonucleotides. After allowing the bound SSB to stabilize, (dT)_{70} (1 \mu M) was injected for 15 minutes. Once all SSB was removed (as judged by a return of the signal to the baseline), SSB (90 nM) was again injected for 5 minutes, allowed to stabilise, then (dT)_{15} (1 \mu M) was injected for 10 minutes. An unmodified flow cell (treated with the preparatory glycerol and NaCl/NaOH injections but without immobilised oligonucleotide) was used as a reference surface and subtracted from all signals.
4.3 RESULTS AND DISCUSSION:

4.3.1 Conversion between SSB ssDNA-binding modes observed by mass spectrometry

Chapter 3 described the formation and nanoESI-MS analysis of SSB-oligonucleotide complexes containing one molecule of (dT)$_{70}$, and either one or two molecules of (dT)$_{35}$ (section 3.3.4). These stoichiometries were consistent with those predicted from solution studies, suggesting that the DNA-binding properties of SSB were able to be preserved in the gas phase. To further investigate the capabilities of mass spectrometry for examining the DNA-binding modes, conditions were manipulated to observe conversion between the two major binding modes: the (SSB)$_{65}$ and the (SSB)$_{35}$ modes. Generally, 1:1 complexes of SSB with (dT)$_{70}$ are considered to reflect binding in the (SSB)$_{65}$ mode, in which the ssDNA wraps completely around the SSB tetramer.$^{[264,325]}$ However, addition of a ~2-fold excess of SSB to (dT)$_{70}$ (in 10 mM ammonium acetate) was able to ‘force’ formation of a complex containing two SSB tetramers bound to a single (dT)$_{70}$ strand, observed directly by nanoESI-MS (Figure 4.3a). For such a complex to form implies that it utilises the (SSB)$_{35}$ binding mode$^{[265,325]}$: only an average of 35 nucleotides are available to interact with each tetramer (which, based on the (SSB)$_{45}$ model (see Figure 1.12), each use an average of two subunits for this interaction). This 2:1 complex formed readily in 10 mM ammonium acetate (pH 7.2), but was not observed when SSB and (dT)$_{70}$ were mixed in 1 M ammonium acetate. This is consistent with previous work$^{[265,325]}$: whilst low ionic strengths permit formation of a cooperative 2:1 SSB:(dT)$_{70}$ complex, higher salt concentrations preclude this cooperativity. Addition of (dT)$_{70}$ to the 2:1 SSB:(dT)$_{70}$ complex resulted in conversion back to the 1:1 complex, presumably in the (SSB)$_{65}$ binding mode (Figure 4.3b), whilst further addition of SSB allowed the 2:1 complex to re-form (Figure 4.3c), demonstrating the reversible nature of the binding mode conversions in response to
Figure 4.3. Conversion between the DNA-binding modes of SSB can be observed by nanoESI-MS. (a) A complex of two SSB tetramers bound to a single strand of (dT)$_{70}$, formed by addition of a two-fold excess of SSB to (dT)$_{70}$ in 10 mM ammonium acetate, pH 7.2. This complex is assumed to utilise the (SSB)$_{15}$ binding mode. The measured mass of this complex was 172331 Da (theoretical mass 171976 Da). (b) After addition of (dT)$_{70}$ to the complex in (a), reversion to a 1:1 SSB:(dT)$_{70}$ complex (and the (SSB)$_{15}$ binding mode) was observed in the spectrum. The measured mass of this complex was 96685 Da (theoretical mass 96604 Da). (c) Further addition of SSB to the complex in (b) allowed the (SSB)$_{15}$ mode complex to re-form. The peaks marked with filled circles correspond to a small excess of unbound SSB.
changes in the SSB: (dT)$_{30}$ ratio. Rapid and reversible conversion between the DNA-binding modes of SSB in response to changes in salt concentration or SSB:DNA ratio is a well-documented feature of these modes\cite{264}. These experiments, therefore, provide further evidence that the SSB-ssDNA complexes observed by nanoESI-MS are genuine reflections of the solution phase complexes.

4.3.2 Transfer of SSB between ssDNA oligonucleotides detected by mass spectrometry

Having established that SSB-oligonucleotide complexes utilising multiple binding modes could be successfully observed by nanoESI-MS, ‘direct transfer’ of SSB tetramers between oligonucleotides (section 4.1) was probed by challenging preformed SSB-oligonucleotide complexes with ‘acceptor’ oligonucleotides, and using nanoESI-MS to analyse the resulting mixtures. Although the kinetics of transfer, which occurs in seconds\cite{325}, are too rapid to observe using this method, examining the products of various challenge reactions was anticipated to provide insight into how well SSB-ssDNA gas-phase complexes could be compared with previous solution phase results.

Complexes of SSB with various oligonucleotides were formed, observed by nanoESI-MS, then challenged with an equimolar quantity of either an oligonucleotide of a different length, or an oligonucleotide of the same length which had been 5’-biotinylated to provide a mass difference. NanoESI-MS of the mixture immediately after addition of the second oligonucleotide revealed whether (a) the spectrum remained unchanged, with SSB still bound to the original oligonucleotide; (b) the spectrum contained only complexes of SSB with the added oligonucleotide, indicating a complete transfer onto the challenging ssDNA, or (c) the spectrum contained peaks corresponding to complexes of both oligonucleotides with SSB, demonstrating a partial transfer, or partitioning of SSB between the two ssDNA strands.
Figure 4.4. Transfer of SSB from a 1:1 SSB:(dT)$_{70}$ complex onto added (dT)$_{35}$ does not occur. Complexes of SSB with (dT)$_{70}$ were formed in (a) 10 mM ammonium acetate, pH 7.2, and (b) 1 M ammonium acetate, pH 7.2, then challenged by addition of equimolar (dT)$_{35}$. No changes in the spectra were observed following (dT)$_{35}$ addition (compare top and bottom spectra). The measured masses of the 1:1 SSB:(dT)$_{70}$ complexes were 96759 Da and 96675 Da in 10 mM and 1 M ammonium acetate, respectively (theoretical mass 96604 Da).

Owing to the highly salt-dependent nature of direct transfer$^{325}$, experiments were performed in both low (10 mM ammonium acetate) and high (1 M ammonium acetate) salt solutions.

A 1:1 complex of SSB with (dT)$_{70}$ in 10 mM ammonium acetate, when challenged by addition of equimolar (dT)$_{35}$, was not perturbed – no transfer onto the added (dT)$_{35}$ was observed (Figure 4.4a). This was also the case in 1 M ammonium acetate (Figure 4.4b), as well as when greater excesses of (dT)$_{35}$ were added. When the same complex was challenged with 5’-biotinylated (dT)$_{70}$ [Bio-(dT)$_{70}$] in 10 mM ammonium acetate, no transfer onto Bio-(dT)$_{70}$ was observed (Figure 4.5a). However, when the complex was formed in 1 M ammonium acetate,
Figure 4.5. Transfer of SSB from a 1:1 SSB:(dT)$_{70}$ complex onto added Bio-(dT)$_{70}$ can occur at elevated salt concentration. (a) The nanoESI-mass spectrum of a 1:1 SSB:(dT)$_{70}$ complex in 10 mM ammonium acetate (top) was unchanged after addition of equimolar Bio-(dT)$_{70}$ (bottom). The measured mass of this complex was 96761 Da (theoretical mass 96604 Da). (b) Addition of equimolar Bio-(dT)$_{70}$ to the same 1:1 SSB:(dT)$_{70}$ complex in 1 M ammonium acetate (top) resulted in observation of approximately equal quantities of 1:1 complexes of SSB with both (dT)$_{70}$ and Bio-(dT)$_{70}$ (bottom). The masses of the (dT)$_{70}$ and Bio-(dT)$_{70}$ complexes were measured to be 96718 Da (theoretical mass 96604 Da) and 97219 Da (theoretical mass 97173 Da) respectively.

Challenge by addition of Bio-(dT)$_{70}$ resulted in 1:1 complexes of SSB with both DNA strands being observed in the spectra (Figure 4.5b), indicating partial transfer of SSB onto Bio-(dT)$_{70}$.

By mixing a two-fold excess of SSB with (dT)$_{70}$ in 10 mM ammonium acetate, a complex consisting of two SSB tetramers cooperatively bound to a single (dT)$_{70}$ strand [2:1 SSB:(dT)$_{70}$] was formed, as described in section 4.3.1 (Figure 4.6, top spectrum). Upon addition of equimolar Bio-(dT)$_{70}$ to this complex, 1:1 complexes of SSB with both (dT)$_{70}$ and Bio-(dT)$_{70}$ in approximately equal intensity were observed in the mass spectrum (Figure 4.6, bottom spectrum).
Figure 4.6. SSB bound to (dT)\textsubscript{70} in the (SSB)\textsubscript{35} mode can be transferred to added Bio-(dT)\textsubscript{70}. Addition of equimolar Bio-(dT)\textsubscript{70} to the complex containing two SSB tetramers bound to one (dT)\textsubscript{70} molecule (top) resulted in formation of 1:1 complexes of SSB with both (dT)\textsubscript{70} and Bio-(dT)\textsubscript{70}. All components were in 10 mM ammonium acetate. The measured mass of the 2:1 SSB:(dT)\textsubscript{70} complex was 172171 Da (theoretical mass 171976 Da). The measured masses of the 1:1 (dT)\textsubscript{70} and Bio-(dT)\textsubscript{70} complexes were 96715 Da (theoretical mass 96604 Da) and 97265 Da (theoretical mass 97193 Da) respectively.

These results are consistent with expectations based on the binding modes predicted to be utilised in each case. The 1:1 SSB:(dT)\textsubscript{70} complex, assumed to adopt the fully-wrapped (SSB)\textsubscript{65} binding mode, lacks accessible binding sites for the challenging oligonucleotide to bind to, and as such transfer becomes unfeasible (Figure 4.4 \textit{a,b}, 4.5\textit{a}). In contrast, to accommodate two SSB tetramers on one strand of (dT)\textsubscript{70}, the 2:1 SSB:(dT)\textsubscript{70} complex needs to occupy the (SSB)\textsubscript{35} binding mode, leaving two unoccupied binding sites on each tetramer. This permits binding of Bio-(dT)\textsubscript{70}, and, because there is no longer an excess of SSB over DNA, 1:1 complexes, in the
Figure 4.7. Transfer of SSB from complexes with (dT)$_{35}$ onto added (dT)$_{70}$ or Bio-(dT)$_{70}$ readily occurs. Complexes of SSB with one molecule of (dT)$_{35}$ in (a) 10 mM ammonium acetate and (b) 1 M ammonium acetate were completely converted to 1:1 SSB:(dT)$_{70}$ complexes after addition of equimolar (dT)$_{70}$. (c) In 10 mM ammonium acetate,
SSB from complexes with two (dT)$_{35}$ strands was completely transferred onto added equimolar Bio-(dT)$_{70}$. The measured masses of the 1:1 SSB:(dT)$_{35}$ complexes were 86087 Da and 86051 Da in 10 mM and 1 M ammonium acetate, respectively (theoretical mass 85957 Da). The measured mass of the 1:2 SSB:(dT)$_{15}$ complex in (c) was 96775 Da (theoretical mass 96542 Da). See Figures 4.4 and 4.6 for masses of the (dT)$_{70}$ and Bio-(dT)$_{70}$ complexes.

(SSB)$_{65}$ mode, are formed (Figure 4.6). This transition has been shown previously to occur via a ternary intermediate in which both strands bind to the two SSB tetramers, followed by a rapid rearrangement to form the 1:1 complexes$^{[325]}$. The formation of approximately equal amounts of each complex indicates a similar affinity of SSB for (dT)$_{70}$ and Bio-(dT)$_{70}$. The ability of SSB to partially transfer from the 1:1 SSB:(dT)$_{70}$ complex to Bio-(dT)$_{70}$ at higher ionic strength (Figure 4.5$b$) is probably due to transient partial unwrapping of the (dT)$_{70}$ from around the tetramer$^{[323,325]}$, resulting in a small population of tetramers with unoccupied binding sites, and the opportunity for binding of a second DNA strand. However, challenging with (dT)$_{15}$, even at elevated salt concentrations, does not result in any detectable transfer onto (dT)$_{15}$ (Figure 4.4). This is the case even when (dT)$_{35}$ is added in excess, and after extended periods of time (not shown). Presumably, binding of (dT)$_{15}$ to the unoccupied sites to form a ternary (dT)$_{70}$:SSB:(dT)$_{15}$ complex is able to occur, but because of the higher affinity of SSB for (dT)$_{70}$, (dT)$_{35}$ dissociates preferentially, resulting in a return to the original 1:1 SSB:(dT)$_{70}$ complex.

Challenge of the 1:1 complex of SSB with (dT)$_{35}$ by addition of (dT)$_{70}$ resulted in complete transfer of SSB onto (dT)$_{70}$, with none of the original SSB-(dT)$_{15}$ complex remaining in the spectrum (regardless of the ionic strength of the solution) (Figure 4.7 $a,b$). This is again consistent with the (SSB)$_{35}$ binding mode providing the necessary ‘free’ binding sites for (dT)$_{70}$ to be able to form an intermediate ternary complex required for direct transfer. The higher affinity of SSB for (dT)$_{70}$ than for (dT)$_{35}$ is then able to drive the transfer to completion.
Similarly, even when a 1:2 SSB:(dT)$_{35}$ complex (with all four binding sites occupied by two separate (dT)$_{15}$ strands) was challenged with Bio-(dT)$_{70}$, complete transfer of SSB to Bio-(dT)$_{70}$ was observed (Figure 4.7c), despite the lack of free binding sites in the original complex. This result can be explained by the high degree of negative cooperativity which exists between the subunits of the SSB tetramer.$^{266,268}$ Binding of a second (dT)$_{35}$ molecule to an SSB tetramer is much more difficult than binding of the initial (dT)$_{15}$, with the consequence that the second strand remains weakly bound, enabling it to be displaced by the longer, higher affinity strand (dT)$_{70}$.

4.3.3 Transfer of SSB between ssDNA oligonucleotides detected by surface plasmon resonance

Surface plasmon resonance was also used to examine the direct transfer of SSB tetramers between ssDNA molecules. When Bio-(dT)$_{70}$ was immobilised on the surface of a streptavidin-coated sensor chip and SSB injected over it, SSB binding could be detected as a rapid increase in response (Figure 4.8), which was maintained stably for long periods. The magnitude of the response was consistent with an approximately 1:1 SSB:(dT)$_{70}$ stoichiometry. Despite the large excess of SSB available to bind to the immobilised (dT)$_{70}$, the salt concentration (150 mM NaCl) was too high to permit stable formation of the complex containing two SSB tetramers in the (SSB)$_{15}$ mode.

Flowing free (unbiotinylated) (dT)$_{70}$ over this complex resulted in an immediate and rapid drop in response (Figure 4.8), corresponding to transfer of SSB from the immobilised (dT)$_{70}$ to the free (dT)$_{70}$. After re-binding SSB to the immobilised (dT)$_{70}$, the complex was again challenged by flowing unbiotinylated (dT)$_{15}$ over the surface. A minor decrease in response was observed upon injection of (dT)$_{15}$ (Figure 4.8), but this decrease was very small in comparison to that seen with the previous injection of (dT)$_{70}$.
Figure 4.8. Transfer of SSB tetramers from immobilised (dT)$_{70}$ detected by surface plasmon resonance. Biotinylated (dT)$_{70}$ was immobilised on the surface of a streptavidin-coated sensor chip, and saturated by SSB such that a stable complex was formed. Free (dT)$_{70}$ (1 µM) was injected over the complex at 5 µL/min until a stable signal close to the original baseline was obtained. The immobilised (dT)$_{70}$ was again saturated with SSB, then free (dT)$_{35}$ (1 µM) injected at 5 µL/min.
These SPR results are consistent with those obtained by mass spectrometry. The relatively complete transfer from Bio-(dT)$_{70}$ to free (dT)$_{70}$ (compared to the partial transfer observed by ESI-MS) can be attributed to the large, continuous excess of free (dT)$_{70}$ supplied by injection over the SPR sensor surface. The lack of transfer from (dT)$_{35}$ to (dT)$_{15}$, even with a large excess of (dT)$_{35}$, is as expected: (dT)$_{35}$ is unable to insinuate into the fully wrapped complex. The small decrease in response is potentially due to a small population of SSB tetramers bound in the (SSB)$_{15}$ mode to the immobilised Bio-(dT)$_{70}$ (2 tetramers per strand), which could allow some transfer onto (dT)$_{35}$. Alternatively, the excess (dT)$_{35}$ may be capable of competing for the small population of SSB tetramers which are bound in the (SSB)$_{65}$ mode but which have unbound subunits due to transient unwrapping processes.

With Bio-(dT)$_{35}$ immobilised on the SPR sensor chip, SSB was also able to bind and form a stable complex on the surface (Figure 4.9). Upon injection of either free (dT)$_{70}$ or (dT)$_{35}$, the response dropped rapidly to the level observed before SSB binding, indicating complete transfer of SSB onto both lengths of free oligonucleotide. Comparison of the sensorgrams in Figures 4.8 and 4.9 shows that, with SSB bound to immobilised Bio-(dT)$_{35}$, the transfer rate onto both lengths of ssDNA was faster, with a steeper decrease in signal upon injection of free oligonucleotide, than the rate of transfer onto free (dT)$_{70}$ from immobilised Bio-(dT)$_{70}$. This, again, reflects that transfer is able to occur more readily when binding sites are available, as in the (SSB)$_{15}$ binding mode.
Figure 4.9. Transfer of SSB tetramers from immobilised (dT)$_{35}$ detected by surface plasmon resonance. Biotinylated (dT)$_{35}$ was immobilised on the surface of a streptavidin-coated sensor chip and saturated with SSB, then free (dT)$_{70}$ (1 µM) was injected over the surface at 5 µL/min. After again saturating the immobilised (dT)$_{35}$ with SSB, free (dT)$_{35}$ was injected at 5 µL/min.
4.4 conclusions:

A discrete sampling approach to mass spectrometry such as that used in this work is not suited to following rapid kinetic processes with precision. However, the directness of mass spectrometry – stoichiometries are generally able to be observed unambiguously – is valuable for observing the start and end points of reactions, such as the SSB direct transfer reactions examined in this work. Techniques which offer kinetic accuracy, such as fluorescence quenching or FRET, are less direct and are able to give an average picture of the molecules under consideration: therefore, mass spectrometry can provide a complementary method for confirming the results of other studies.

The consistency of the results obtained here using nanoESI-MS with those obtained using fluorescence quenching\cite{325} shows that MS provides an effective method for directly examining SSB-ssDNA complexes. Taken together with the previous evidence, our results support the superior ability of the (SSB)$_{11}$ mode, which is proposed to function in DNA replication\cite{265} to facilitate direct transfer of SSB to other oligonucleotides. Whether such a mode of transfer is physiologically relevant awaits further investigation. The complexities of the SSB binding make such investigations experimentally challenging, but the idea of cooperative transfer of adjacent SSB tetramers during DNA replication is intuitively appealing.

This work also represents the first MS observation of the DNA-binding modes of SSB. One constraint of mass spectrometry for the study of intact protein complexes is that the conditions required to obtain optimal spectra are sometimes quite different from solution conditions commonly used in solution studies, which may hinder comparisons. Here, we have shown that the behaviour of SSB and SSB-ssDNA complexes in low (10 mM) and high (1 M) ammonium acetate solutions reflects that observed in previous solution studies using traditional buffer conditions and salts, and that the complexes are able to be preserved in the gas phase. Transfer
of intact SSB tetramers between discrete oligonucleotides was shown to depend, as expected, on the identity of the initial complex, with complexes in the (SSB)$_{15}$ mode undergoing transfer much more readily. Ultimately, a combination of techniques is likely to shed light on the detailed mechanisms by which SSB achieves its dynamics on ssDNA.
Chapter 5: PRODUCTION OF MODIFIED $\beta_2$ CLAMPS FOR
EXAMINATION OF THE $\varepsilon$-$\beta_2$ INTERACTION

5.1 INTRODUCTION

The $\beta_2$ sliding clamp, the product of the dnaN gene in E. coli\[136\], was introduced briefly in section 1.1.3.2 as an integral component of the DNA replication machinery. The dnaN gene encodes a 366 amino acid protein with a molecular mass of 40586 Da, which forms the highly stable $\beta$ homodimer ($\beta_2$)\[82,479\]. The clamp is loaded onto DNA with the assistance of the clamp loader (DnaX) complex (via a direct interaction with the $\delta$ subunit\[146\]), such that it encircles DNA whilst retaining the ability to slide along it\[82,480\]. Once loaded, it acts as a tether for the catalytic core of the replicative polymerase (Pol III) by interacting directly with the $\alpha$ subunit\[82\], ensuring that it remains linked to DNA during replication and enabling rapid and highly processive (>50 kb) DNA synthesis\[82\].

The structure of the clamp is intimately related to its function. The crystal structures of $\beta_2$ alone\[17,138\] and in complex with primed DNA\[139\] have been solved and reveal it to be a ring-shaped dimer with a central cavity approximately 35 Å in diameter, large enough to accommodate dsDNA (Figure 5.1a, b). Within the dimer, the two identical subunits are arranged in a head-to-tail fashion, with the C-terminal end of each subunit forming an interface with the N-terminal end of the opposite subunit\[137\]. As a consequence of its head-to-tail organisation, the $\beta$ dimer possesses two-fold rotational symmetry, as well as two distinct faces – both N-termini are found on one face, whilst both C-termini are located on the other. A total of twelve $\alpha$ helices form a lining for the inside surface of the ring, supported on the exterior by six 8-stranded $\beta$ sheets\[137\]. The inner surface of the clamp has a significant positive charge\[480\], and electrostatic interactions between this basic interior and the negatively charged
Figure 5.1. The crystal structure of the β₂ sliding clamp. (a) The clamp is a ring-shaped homodimer, lined by α-helices, with continuous β-sheets forming the outer edge. The central channel has a diameter of ~35 Å, large enough to accommodate dsDNA. This structure was generated using PyMol and the PDB coordinates 1MMI[138], and shows the C-terminal face of the dimer. The residues forming the protein binding site on each subunit are represented as green spheres. (b) Schematic organisation of the three domains comprising each monomer, which are oriented head-to-tail within the dimer. The protein binding sites are located between domains II and III. (c) Side view of the crystal structure of β₂ bound to double-stranded DNA, generated using the PDB coordinates 3BEP[139]. The two strands of the DNA duplex are represented as blue and cyan spheres. The DNA passes through the centre of the clamp at an angle of approximately 22° from the axis of rotation of the dimer.

DNA backbone are thought to provide a driving force for closure of the clamp around DNA during clamp loading, and increase the stability of the clamp on DNA[138]. The highly symmetrical nature of the ring is a result of the three almost identical domains which comprise each subunit, each of which contribute two helices to the lining of the ring[137] (the domain organisation of the dimer is represented schematically in Figure 5.1b). The external β sheets
form a continuous scaffold around the clamp, and hydrogen bonding between adjacent β strands in the scaffold contributes substantially to stabilisation of the two interfaces between the monomers[137]. Additional stability is provided by a small hydrophobic core formed by packing of Phe-106 and Ile-278 from one subunit against Ile-272 and Leu-273 from the other (an I272A/L273A β mutant forms a stable monomer[481]). Surrounding this core, electrostatic attractions between the positively charged N-terminal surface and the negatively charged C-terminus (involving up to six specific ion pairs) further strengthen the dimer[117].

The co-crystal structure of β2 in complex with a short primer-template DNA (a 10/14-mer with a 5’ ssDNA extension)[139] showed that the β dimer can bind directly to DNA (this was contrary to predictions based on its ability to slide along DNA). Direct contacts between the clamp and DNA have subsequently been found to be required for the viability of E. coli, and probably serve a role in clamp loading[482]. In the crystal structure, the duplex portion of the DNA passes through the clamp at the rather steep angle of 22° from the rotational axis of the ring, allowing it to form specific contacts with several residues (Figure 5.1c)[139]. In addition, the ssDNA template strand appears to bind within a hydrophobic cleft on the C-terminal face of the clamp (coincident with the protein-binding cleft which will be described below). A β2-ssDNA interaction may serve to keep β2 in place at a primed site following the departure of the clamp loader, and before the arrival of Pol III. Presumably, upon binding to Pol III, this contact is disrupted, along with other specific contacts, allowing the clamp to slide freely along DNA[139]. Water molecules lining the central channel of the ring, which are observed in the high-resolution crystal structure of β2[118], probably also facilitate its mobility on DNA.

Although it was first identified as a processivity factor for chromosomal DNA replication, the β2 clamp is now known to play a central role in multiple processes, interacting with a variety of proteins involved in other aspects of DNA metabolism. In addition to its interactions with
the α subunit of DNA Pol III[82] and the δ subunit of the clamp loader complex[146], which are required for its role in replication, β₂ also interacts with (and stimulates the activity of) the other four known prokaryotic DNA polymerases: Pol I[140], Pol II[141,142], Pol IV[141,144] and Pol V[145]. Pol I is involved in Okazaki fragment maturation, along with DNA ligase, which has also been shown to interact with β₂[140]. Pols II, IV and V are error-prone polymerases involved in translesion synthesis (TLS) at sites of DNA damage (reviewed in [183]). Furthermore, interactions of β₂ with Hda, a DnaA-related protein involved in regulating the inactivation of DnaA[484], as well as two essential proteins in mismatch repair, MutS[140,147] and MutL[147], have been demonstrated.

This large number of interactions raised the question of whether the various β₂-binding proteins shared a common β₂-binding motif, as had been identified in the binding partners of PCNA, the eukaryotic counterpart of β₂[485,486]. A bioinformatic analysis of a number of eubacterial β₂-binding protein families revealed that variations of a conserved pentapeptide motif, with the consensus sequence QL(S/D)LF, were present within the sequences of most members of these families[110]. Peptides containing the consensus sequence, or variants from individual proteins, were shown experimentally to interact with E. coli β₂, and were able to compete for binding with known β₂-binding proteins[110,487]. It was thus proposed as a universal eubacterial clamp binding motif (CBM)[110]. A hexameric consensus sequence, based on the β₂-binding motif from the Hda protein, QLSLPL, has also been proposed[484,487]. The discovery of a conserved CBM suggested that a similarly conserved region of β₂ might be involved in mediating its protein-protein interactions. It had already been shown that the α subunit of Pol III and the δ subunit of the clamp loader complex competed for binding to β₂, suggesting that they bound to a common (or at least overlapping) site on the clamp[146,177].
The crystal structure of β (as a monomer) in complex with the δ subunit of the clamp loader complex (see Figure 1.8d) provided the first insight into the structural details of a β clamp-protein interaction\textsuperscript{[148]}. Subsequent crystal structures of β in complex with the ‘little finger’ domain of Pol IV\textsuperscript{[150]}, the C-terminal 16 residues of Pol IV\textsuperscript{[149]} and the clamp binding peptides from Pol II and the α subunit of Pol III\textsuperscript{[151]} revealed a common site of interaction on the clamp: a conserved hydrophobic region on the C-terminal face of β\textsubscript{2}, between domains II and III (Figure 5.1a). All of these proteins contain peptide motifs reminiscent of the consensus CBM. In δ, this motif occurs towards the N-terminus (residues 69-74) and consists of the sequence QAMSLF. The Leu-73 and Phe-74 residues from δ protrude into a deep hydrophobic pocket on β\textsubscript{2}, formed by Leu-155, Leu-177, Pro-242, Val-247, Val-360 and Met-362\textsuperscript{[148]} (Figure 5.2a). In Pol IV, the clamp binding motif resides at the extreme C-terminus (residues 346-351) and has the sequence QLVLGL. The two leucines from the LGL tripeptide at the C-terminus of Pol IV (Leu-349 and Leu-351) insert into the same pocket as the Leu-73 and Phe-74 residues from δ\textsuperscript{[149,150]}. Adjacent to this hydrophobic pocket in β\textsubscript{2} is a shallower hydrophobic groove, formed by residues His-175, Asn-320, Tyr-323, Val-344, Met-362, Pro-363 and Met-364\textsuperscript{[149,150]}, with which the preceding residues in the motifs from δ and Pol IV (QAMS and QLV, respectively) interact.

Pol I, Pol II, the α subunit from Pol III, Pol V and MutL (or peptides corresponding to their clamp binding motifs) have all been shown to bind to the same hydrophobic region on β\textsubscript{2} utilised by δ and Pol IV and to adopt very similar conformations\textsuperscript{[111,140,147,151]}. The peptides from Pol II and Pol IV are shown superimposed within the β binding site in Figure 5.2b. Residues contributing to the Pol II peptide-β interaction (which are representative of those involved in all CBM-β interactions) are shown schematically in Figure 5.2c. The existence of this common interaction surface suggests that various binding partners compete with each
Figure 5.2. A hydrophobic cleft in β mediates its protein-protein interactions. (a) The interface between β and the δ subunit of the clamp loader complex. Residues 69-74 of δ (QAMSLF) are shown as green sticks, and the β binding surface is coloured blue, red and white to indicate positive, negative and hydrophobic regions of electrostatic potential, respectively. The Leu-73 and Phe-74 residues from δ insert into a hydrophobic pocket on β. Residues lining the pocket are labelled. Figure produced using PyMol and the PDB coordinates 1JQL. (b) Conservation of the protein binding site on β. Clamp binding peptides from Pol II (GQLGLF; blue), Pol III (C-terminal peptide EQVELEFD; green) and Pol IV (RQLVLGL; pink) superimposed within the β binding site. Residues comprising the β binding pocket are coloured according to their sequence conservation (based on alignment of β sequences from 42 bacterial species), ranging from red (90% conservation) to yellow (50% conservation). The two ‘subsites’ which comprise the β binding site are circled. All three polymerase peptides adopt similar conformations within the binding pocket. (c) Schematic representation of the residues comprising the β binding site, based on the structure of the Pol II clamp binding peptide, GQLGLF, bound to β. Main-chain atoms of the Pol II peptide are represented by circles (C: green; Ca: grey; N: blue; O: red). Residues from β (ovals representing side-chain interactions, rectangles representing main-chain interactions) are coloured as in (b). Parts (b) and (c) reproduced from reference 151.
other for access to the clamp. How such competition is regulated, such that the appropriate protein for a task is associated with β₂ at the correct time, is still the subject of investigation. Of course, since the β₂ dimer consists of two identical subunits, it also contains two identical hydrophobic clefts, both presumably able to interact with other proteins (depending on steric constraints). Using a mutated β₂ dimer in which one of the clefts was disrupted by a C-terminal truncation, it was shown that a single cleft was sufficient to interact normally with both α and a version of the clamp loader complex (γδδ'χψ)⁴⁷⁹. In a functional sense, a single cleft was also sufficient to enable loading of the clamp onto DNA by the clamp loader, and the hemi-mutant dimer was also able to stimulate Pol III replication in vitro to an extent comparable to the wild-type dimer⁴⁷⁹.

Given these observations, the role of the second protein-binding cleft in the β₂ dimer has been the subject of speculation. It has been suggested that multiple polymerases may be able to bind the clamp simultaneously, gaining access to DNA as required⁴⁸⁸. In this so-called ‘toolbelt’ model, the replicative polymerase, Pol III, could occupy the binding site on one β₂ subunit during the elongation phase of DNA replication, whilst a translesion polymerase such as Pol IV could simultaneously occupy the binding site on the second subunit, facilitating rapid takeover from Pol III upon encountering a DNA lesion⁴⁸⁸. Simultaneous binding of Pol III and Pol IV by the clamp has been demonstrated in vitro⁴⁸⁸, but whether such a mechanism actually operates in vivo has not been established, with recent work showing that a single protein-binding cleft on β₂ is sufficient for Pol III-Pol IV switching⁴⁸⁹, and that Pol IV disrupts the association of Pol III with the β₂ clamp when it takes over at a lesion⁴⁹⁰.

It should be noted that, in addition to the common hydrophobic cleft on β₂ utilised by the clamp binding motifs of its binding partners, it is likely that most of these partners also possesses additional, unique contact sites with β₂, which assist in regulating their interactions.
with the clamp\textsuperscript{[491,492]}. This has been demonstrated for several partners, including \( \alpha \)\textsuperscript{[493]}, \( \delta \)\textsuperscript{[148]}, Pol IV\textsuperscript{[150]} and Pol V\textsuperscript{[493]}. In the \( \beta_2 \)-Pol IV structure, Pol IV also contacts the edge of the \( \beta_2 \) clamp, interacting with the beta sheet spanning the dimer interface\textsuperscript{[150]}. This contact has been shown to be essential for Pol III-Pol IV switching \textit{in vitro}\textsuperscript{[489]}. The structure of the \( \beta \)-\( \delta \) complex shows that an interaction between an \( \alpha \) helix in \( \delta \) (preceding its CBM) with part of domain III of \( \beta \) results in substantial rearrangements in both proteins, consistent with the clamp-opening function of \( \delta \)\textsuperscript{[148]}. Although structural information for the interaction between \( \beta \), and \( \alpha \) is not available, \( \alpha \) contains two potential clamp-binding motifs, an ‘internal’ site (residues 920-924, QADMF)\textsuperscript{[110,111]}, and a C-terminal site (residues 1154-1159, QVELEF)\textsuperscript{[111,112]}, both of which are able to bind to the hydrophobic cleft on \( \beta \)\textsuperscript{[151,487]}. The relative importance of these two sites for the \( \beta_2 \)-\( \alpha \) interaction has been the subject of dispute, but the internal site has been shown to be absolutely required for processive replication, which can occur in the absence of the C-terminal site\textsuperscript{[113]}. The function of the C-terminal site, which also interacts with the \( \tau \) subunit of the clamp loader complex\textsuperscript{[112]}, remains unclear, but it may be involved in removal of the polymerase from DNA upon completion of Okazaki fragments\textsuperscript{[111,113]}.

Although the detailed processes underlying the management of various polymerases and other proteins by the \( \beta_2 \) clamp remain to be fully elucidated, it is clear that they are regulated in a complex way, involving the coordination of multiple protein-protein and protein-DNA interactions. The fact that \( \beta_2 \) is a central attachment point for many proteins involved in DNA replication, repair and damage tolerance, its lack of sequence homology with the eukaryotic clamp, PCNA, and the conserved nature of its interactions also suggest that it could be a promising pharmacological target for antibiotic development\textsuperscript{[111,494,495]}. Several small-molecule compounds which bind within the hydrophobic cleft of \( \beta_2 \) and inhibit its interactions with other proteins have been described\textsuperscript{[111,494,495]}, and represent leads for future development.
The existence of a previously undescribed interaction between $\beta_2$ and the proofreading exonuclease subunit of the Pol III core, $\varepsilon$, has recently been discovered in our laboratory by Dr Slobodan Jergic. This discovery (the basis of which will be discussed in detail in Chapter 6) provided the motive for the present chapter, which describes the preparation of various $\beta_2$ constructs produced to probe aspects of this newly identified $\beta_2$-$\varepsilon$ interaction. The general aim of this work was to produce clamps with one of their two hydrophobic protein-binding clefts disrupted, and examine the effects of this disruption upon specially designed $\varepsilon$-dependent replication assays (the results of these assays are presented in Chapter 6).

As $\beta_2$ is a homodimer produced from a single gene, $dnaN$, the production of heterodimers presents a challenge. Two strategies towards the production of such dimers will be described in this chapter: firstly, the production of a ‘single-chain’ $\beta_2$ construct, in which both subunits are expressed within a single polypeptide, theoretically allowing mutation of one subunit at a time; secondly, the formation and purification of mixed dimers using subunit exchange. One approach to the creation of mixed dimers has been described previously. The current work utilised a different strategy, requiring knowledge of the subunit exchange properties of the $\beta_2$ dimer. Consequently, a nanoESI-MS investigation of $\beta_2$ subunit exchange was undertaken, and is also presented in this chapter.
5.2 MATERIALS AND METHODS

5.2.1 Oligonucleotide primers

414: 5’-AAAAAATGAAATTTACCCTAGAGAGCATTTATTAAAAACGC-3’

(The NdeI site, which includes a start codon (in italics), is underlined).

415: 5’-AAAAACGCCGCTACGCTCAGTCTCATTGGCAGACAACATAAGCC-3’

(The MluI site is underlined).

416: 5’-AAAGGATCGCCATGAAATTTACCCTAGAGAGCATTTATTAAAAACCG-3’

(The BamHI site is underlined).

417: 5’-TTGAATTCTTGACGTCTCATGACAGACATGAAATTTAAAAACGC-3’

(The EcoRI site and the reverse complement of the TAA stop codon are underlined and italicised, respectively).

562: 5’-TTTTTTTTTTTTACATGAAATTTACCCTAGAGAGGAGC-3’

(The NdeI site, with the start codon italicised, is underlined)

563: 5’-TTTTGAATTCTTGACACATAAGCCCGCTCCTGG-3’

(The EcoRI site and the reverse complement of the TAA stop codon are underlined and italicised, respectively).

5.2.2 Construction of plasmids

For general methods related to plasmid construction, refer to section 2.3. Specific vectors and promoters are described in section 2.2. Unless otherwise specified, all plasmids were hosted in
\textit{E. coli} strain AN1459 (section 2.2.2) during construction, and were sequenced using appropriate external primers to confirm the integrity of inserted genes (section 2.3.6).

\textit{pYW1480}: This plasmid, containing two copies of the \textit{dnaN} gene (which encodes \(\beta\)) cloned in tandem, was designed to express a single protein (‘single-chain \(\beta_\text{2}\)’) containing both subunits of the \(\beta_2\) dimer, separated from each other by a flexible linker (SGSGTRESGSIGSG) and was constructed in collaboration with Mr Yao Wang. Plasmid pND262\textsuperscript{[138]} contains the complete \textit{E. coli} \textit{dnaN} gene, and was used as a template for amplification of the two \textit{dnaN} genes by PCR. The \textit{dnaN} fragment to be expressed at the N-terminus of the fusion construct, designated \(\beta_N\), was amplified with the 5´-primer 414 (section 5.2.1), designed to incorporate a start codon into an \textit{NdeI} restriction site, and the 3´-primer 415 (section 5.2.1), designed to replace the natural TAA stop codon of \textit{dnaN} with the codons for the first six amino acids of the linker sequence (SGSGTR), which include an \textit{MluI} restriction site. For amplification of the \textit{dnaN} gene to be expressed at the C-terminus of the linked \(\beta\) protein (\(\beta_C\)), the 5´-primer 416 (section 5.2.1) was used to incorporate a \textit{BamHI} restriction site within the codons for the final three residues of the linker sequence (GSG) prior to the start codon for \textit{dnaN}, whilst the 3´-primer 417 (section 5.2.1) was designed to create an \textit{EcoRI} site immediately after the TAA stop codon for \textit{dnaN}. PCR conditions for amplification of both \(\beta_N\) and \(\beta_C\) were: initial denaturing at 95°C for 2 min; 30 cycles of denaturing at 95°C for 30 s, annealing at 55°C for 30 s, extension at 72°C for 1 min; final extension at 72°C for 5 min. Plasmid pCM1444, which is derived from vector pND706 (section 2.2.1), was used as a vector for construction of pYW1480. pCM1444 contains the gene for the photoswitchable fluorescent protein mKikGR\textsuperscript{[497]} between \textit{NdeI} and \textit{MluI} sites, upstream of the \textit{ssb} gene, which is flanked by \textit{BamHI} and \textit{EcoRI} sites. The two genes are separated by a short linker sequence. To construct pYW1480, the \(\beta_C\) PCR fragment was digested with \textit{BamHI} and \textit{EcoRI}, and purified using a QIAquick PCR purification kit (Qiagen).
pCM1444 was digested with the same enzymes and isolated from the excised ssb fragment by electrophoretic separation in an agarose gel (1%) followed by gel extraction (section 2.3.3). Ligation of the βc fragment into the vector gave plasmid pYW1479 which was then digested with NdeI and MluI, and isolated from the excised mKikGR fragment by electrophoretic separation and gel extraction, as above. The βN fragment, which was isolated after digestion with the same enzymes, was inserted between the restriction sites to give the 6526 bp plasmid pYW1480, containing two dnaN genes (separated by a 42 bp linker sequence coding for the amino acid sequence SGSGTRESGSIGSG) under transcriptional control of the tandem bacteriophage λpR and pL promoters.

pYW1451: Plasmid pKO1274[179] contains the sequence coding for MAGLNDIFEAQKIEWHEH in the vector pETMCSI[498], upstream of unique NdeI and EcoRI sites. This sequence serves as a recognition target for biotin ligase[499], and allows specific post-translational biotinylation (which occurs at the lysine residue in bold) at the N-terminus of proteins inserted between the NdeI and EcoRI restriction sites. Full-length dnaN was amplified from pND262 by PCR, using the 5´-primer 414 (section 5.2.1) to incorporate a start codon within an NdeI site and the 3´-primer 417 (section 5.2.1) to create an EcoRI site immediately following the TAA stop codon of dnaN. PCR conditions were as described above for the βN and βC fragments. Following digestion by NdeI and EcoRI, the dnaN fragment was inserted between the NdeI and EcoRI sites of pKO1274 (directly following the biotinylation sequence) to produce pYW1451, in which overproduction of N-terminally biotinylated β (bio-β) is under transcriptional control of the phage T7 ϕ10 promoter.

pCM1502: This plasmid, in which overproduction of bio-β is under transcriptional control of the tandem bacteriophage λpR and pL promoters, was constructed by ligating the XbaI-NcoI fragment from pYW1451 (containing dnaN along with the N-terminal biotinylation sequence)
between the same restriction sites in vector pND707 (section 2.2.1), resulting in the 5410 bp plasmid pCM1502.

*pCM1513*: Plasmid pYW1480 was digested with NdeI and EcoRI and the fragment containing the sequence encoding single-chain β₂, was isolated by electrophoresis and gel extraction. The fragment was inserted between the NdeI and EcoRI sites of pCM1502, and, after ligation, transformed into SURE2 supercompetent *E. coli* cells (Stratagene). The resulting 6550 bp plasmid pCM1513 (Figure 5.3), contains the single-chain β₂ sequence (preceded by the biotin ligase recognition sequence) downstream of the λ pR and pL promoters.

**Figure 5.3. Plasmid pCM1513.** Two *dnaN* genes are arranged in tandem, such that the plasmid directs overproduction of N-terminally biotinylated single-chain β under control of the tandem bacteriophage λ promoters.
pCM1503: Plasmid pYW1451 was digested with NdeI and EcoRI, and the dnaN gene (without the preceding biotinylation sequence) was isolated by electrophoresis and purified from the gel. This fragment was inserted between the NdeI and EcoRI sites of vector pCL476 (section 2.2.1) to produce the 5343 bp plasmid pCM1503 (Figure 5.4), in which the tandem \( \lambda \) promoters direct overproduction of \( \beta \) with an N-terminal 6 \( \times \) His tag (His\(_6\)\( \beta \)).

**Figure 5.4. Plasmid pCM1503.** This plasmid directs overproduction of N-terminally His\(_6\)-tagged \( \beta \) under control of the tandem bacteriophage \( \lambda \) promoters.

pCM1510: Plasmid pYW1480 was digested with NdeI and EcoRI, and the single-chain \( \beta \) fragment was isolated and inserted into vector pCL476 (section 2.2.1), which had been digested with the same enzymes. The 6483 bp plasmid pCM1510 (Figure 5.5) contains the single-chain \( \beta \) sequence with an N-terminal 6 \( \times \) His tag (His6-single-chain \( \beta \)) under transcriptional control of the tandem \( \lambda \) promoters.
Figure 5.5. Plasmid pCM1510. This plasmid directs overexpression of N-terminally His$_6$-tagged single chain β under control of the λ promoters $p_R$ and $p_L$.

$pCM1531$: Plasmid pCM1531 was constructed for expression of the His$_6$-tagged deletion mutant (His$_6$-β$^C$)$_2$, which has the five residues (MPMRL) from the C-terminus removed of β removed. An Ndel-EcoRI fragment encoding residues 1-361 of β was generated by PCR using pND262 as a template. The 5’-primer 562 (section 5.2.1) was used to incorporate the start codon of the dnaN gene into an Ndel site, and the 3’-primer 563 (section 5.2.1) was used to insert a TAA stop codon and EcoRI site immediately following the codon for Val361 in the dnaN sequence. PCR conditions were as described above for other β amplifications. After digestion of the PCR product with Ndel and EcoRI, the dnaN(1-361) fragment was isolated, and inserted between the Ndel and EcoRI sites of pCL476 (section 2.2.1). The 5328 bp plasmid pCM1531 (Figure 5.6) directs the expression of β$^C$ with an N-terminal 6 × His tag (His$_6$-β$^C$), under transcriptional control of the tandem λ promoters.
Figure 5.6. Plasmid pCM1531. This plasmid was constructed to allow expression of an N-terminally His₆-tagged β mutant (βC) with its 5 C-terminal residues (362-366; MPMRL) removed, under control of the tandem λ promoters $p_R$ and $p_L$.

5.2.3 Buffers for protein purification

Buffers used for the protein purifications described in this chapter were: Lysis buffer I (50 mM Tris-HCl, pH 7.6, 1 mM EDTA, 2 mM DTT, 20 mM spermidine); Lysis buffer II (50 mM Tris-HCl, pH 7.6, 1 mM EDTA, 500 mM NaCl); Lysis buffer III (50 mM Tris-HCl, pH 7.6, 1 mM EDTA, 2 mM DTT); Buffer A (35 mM Tris-HCl, pH 7.6, 1 mM EDTA, 1 mM DTT, 10% (v/v) glycerol, [NaCl] as specified in the text); Buffer B (35 mM Tris-HCl, pH 7.6, 20 mM imidazole, 1 mM DTT, 10% (v/v) glycerol, 200 mM NaCl); Buffer C (35 mM Tris-HCl, pH 7.6, 10 mM imidazole, 1 mM DTT, 10% (v/v) glycerol, 500 mM NaCl); Buffer D (15 mM sodium phosphate, pH 6.5, 1 mM DTT, 10% (v/v) glycerol, 200 mM NaCl); Buffer E (30 mM Tris-HCl, pH 7.6, 1 mM EDTA, 1 mM DTT, 20% (v/v) glycerol, 200 mM NaCl); Buffer F (50 mM Tris-HCl, pH 7.6, 1 mM EDTA, 5 mM DTT, 20% (v/v) glycerol); Refolding buffer (40 mM Tris-HCl, 1 mM EDTA, 5 mM DTT, 20% (v/v) glycerol, 500 mM NaCl).
NaCl); Buffer G (35 mM Tris-HCl, pH 7.6, 0.5 mM DTT, 35 mM imidazole, 15% (v/v) glycerol, 500 mM NaCl).

5.2.4 Overproduction and purification of His$_6$-single-chain $\beta_2$

*E. coli* strain BL21(λDE3)/pCM1510 was grown at 30°C, with shaking, in 6 L of LB medium (section 2.2.3) supplemented with ampicillin (150 mg/L). Upon reaching an $A_{\infty}$ of ~0.6, expression was induced by rapidly shifting the temperature to 42°C, followed by a further 3 hours of incubation at 42°C. Cells were collected by centrifugation (7 min at $11000 \times g$), and frozen at -80°C. Overexpression of a protein corresponding to the predicted size of His$_6$-single-chain $\beta_2$ (83.3 kDa) was verified by SDS-PAGE (section 2.4.2).

Upon thawing, cells (6 g) were resuspended in Lysis buffer I (15 mL/g of cells) by stirring at 4°C for 1 hour. Two complete protease inhibitor cocktail tablets (Roche) were added to the suspension, before the cells were lysed by passing twice though a French Press (12000 psi). PMSF (100 mM in ethanol) was added dropwise to the lysate to a concentration of 0.75 mM, and after stirring for 30 min, the lysate was centrifuged ($38000 \times g$; 40 min). Proteins were precipitated from the soluble fraction (containing His$_6$-single-chain $\beta_2$) by gradual addition of solid ammonium sulfate (0.32 g/mL), followed by stirring at 4°C for 45 min. Precipitated proteins were collected by centrifugation ($38000 \times g$; 45 min) and resuspended in Buffer A containing 175 mM NaCl (27 mL), then dialysed overnight against three 2 L changes of the same buffer.

After dialysis, the sample was clarified by centrifugation ($15000 \times g$; 15 min), and the supernatant (28 mL) was loaded at 1 mL/min onto a column (2.5×12.5 cm) of Toyopearl DEAE-650M resin which had been equilibrated in Buffer A + 175 mM NaCl. Fractions containing His$_6$- single-chain $\beta_2$, which did not bind to the column, were pooled and dialysed
against three changes of 2 L of Buffer B. The dialysed sample was again clarified by centrifugation (15000 × g; 15 min) and the supernatant (41 mL) was loaded (1 mL/min) onto a 1 mL HisTrap affinity column (GE Healthcare) charged with Ni²⁺ ions, which had been equilibrated in Buffer B. After washing with 10 column volumes of Buffer B, partially pure His₆-single-chain β₂ was eluted in a broad peak using a linear gradient (15 mL) from 20-350 mM imidazole in Buffer B. Selected fractions from the peak (eluted at ~125 mM imidazole) were pooled and dialysed against three changes of 2 L of Buffer A containing 25 mM NaCl, before being applied at 0.3 mL/min to a 1 mL MonoQ column (GE Healthcare) which had been equilibrated in the same buffer. His₆-single-chain β₂ eluted in a sharp peak at ~220 mM NaCl in a linear gradient (35 mL) of 25-600 mM NaCl. Fractions containing His₆-single-chain β₂ (1.5 mL, containing 0.6 mg of protein), which was about 85% pure as judged by SDS-PAGE, were pooled, frozen in liquid nitrogen and stored at -80°C.

NanoESI-MS analysis of a sample of His₆-single-chain β₂, following dialysis into 0.1% (v/v) formic acid gave a mass of 83327.7 Da (see Figure 5.7), in good agreement with the mass calculated from the amino acid sequence (83329.4 Da).

5.2.5 Overproduction and purification of (His₆-β₂)

*E. coli* strain BL21(λDE3)/pCM1503 was grown at 30°C, with shaking, in 5 L of LB medium (section 2.2.3) supplemented with ampicillin (150 mg/L). Upon reaching an A₆00 of ~0.7, protein overproduction was induced by rapid temperature shift of the cultures to 42°C, after which they were shaken at 42°C for a further 3 hours. Cells were collected by centrifugation (11000 × g; 7 min), and the cell pellets were frozen in liquid nitrogen and stored at -80°C. Overproduction of an approximately 41 kDa protein was confirmed by SDS-PAGE (section 2.4.2).
Thawed cells (6 g) were resuspended in Lysis buffer II (90 mL), by stirring at 4°C for 1 hour. Cells were lysed by passing twice through a French Press (12000 psi), and PMSF was added dropwise to the lysate to a concentration of 0.5 mM. After stirring at 4°C for a further 30 min, the lysate was clarified by centrifugation (35000 × g; 40 min), and the soluble fraction (~90 mL) was dialysed against three changes of 2 L of Buffer C.

After dialysis, the sample (~70 mL) was applied (1 mL/min) in two separate batches to a 1 mL Ni²⁺-charged HisTrap affinity column (GE Healthcare) which had been equilibrated in Buffer C. For each batch, the column was washed with 15 column volumes of the same buffer, then the bound (His₆-β)₂ was eluted with a linear gradient (15 mL) of 10-350 mM imidazole in Buffer C. The fractions containing partially pure (His₆-β)₂, which eluted in a broad peak at approximately 100 mM imidazole, were pooled and dialysed against three changes of 2 L of Buffer D, yielding Fraction II (16 mL). It was found that 200 mM NaCl was required in the buffer at this point, owing to heavy precipitation of the protein at lower NaCl concentrations. Nonetheless, a light precipitate formed during dialysis, so Fraction II was clarified by centrifugation (35000 × g; 30 min) before the supernatant was applied (1.5 mL/min) to a hydroxyapatite column (2.5 × 10 cm) which had been equilibrated in Buffer D. After washing the column with the same buffer for 100 mL, the bound (His₆-β)₂ was eluted with a linear gradient (350 mL) of 15-200 mM sodium phosphate in buffer D. Fractions from the single peak (which eluted at about 140 mM sodium phosphate) contained pure (His₆-β)₂, and were pooled and dialysed against three changes of 2 L of Buffer E, yielding 34.2 mg of (His₆-β)₂ in a volume of 38 mL. Aliquots of pure (His₆-β)₂ were frozen in liquid nitrogen and stored at –80°C.

The molecular weight of the monomeric His₆-β in 0.1 % (v/v) formic acid was measured to be 41540 Da by nanoESI-MS, in excellent agreement with the calculated mass of 41540 Da. Mass
spectrometry of (His₆-β)₂ under native conditions (200 mM NH₄OAc, pH 7.2) provided a dimeric mass of 83087 Da (see Figure 5.8b), also in good agreement with the calculated mass of 83080 Da.

5.2.6 Overproduction and purification of (His₆-β⁵)₂

The deletion mutant (His₆-β⁵)₂ has 5 amino acids (residues 362-366; MPMRL) removed from the extreme C-terminus of each β subunit. An initial attempt to purify (His₆-β⁵)₂, based on the method described for wild-type (His₆-β)₂ (section 5.2.5) showed that the mutant protein was completely insoluble, even in the presence of 500 mM NaCl. Therefore, it was necessary to carry out the purification under denaturing conditions, then refold the protein.

E. coli strain BL21(λDE3)/pCM1531 was grown at 30°C, with shaking, in 2 L of LBT medium containing ampicillin (150 mg/L), until an A₆₀₀ of ~0.6 was reached. Protein overexpression was then induced by rapid temperature shift to 42°C, followed by a further 3 hours of shaking at 42°C. Cells were collected by centrifugation (11000 × g; 7 min), frozen in liquid nitrogen and stored at −80°C. Overexpression of a ~41 kDa protein was confirmed by SDS-PAGE (section 2.4.2).

Upon thawing, cells (2.5 g) were resuspended in Lysis buffer III (38 mL) by stirring at 0°C for 1 hour, then lysed by passing the suspension twice through a French press (12000 psi). The lysate was centrifuged (35000 × g; 40 min) to give an insoluble fraction, containing (His₆-β⁵)₂, then washed several times with lysis buffer to remove contaminating proteins.

In the first washing step, the pellet was resuspended in Lysis buffer III containing 1 M NaCl (40 mL) by stirring at 0°C for 45 min followed by two passes through a French press (12000 psi) to aid homogenisation. The insoluble pellet was collected by centrifugation (35000 × g; 40 min). This washing step was repeated and the pellet collected again, followed by a third
wash in which the pellet was resuspended in Lysis buffer III without NaCl, then collected as in the previous wash steps to produce a final insoluble fraction. This fraction was resuspended in Buffer F (8 mL), giving a homogenous suspension which was frozen in liquid nitrogen and stored at –80°C.

The proteins were unfolded by diluting 1 mL of thawed suspension with Buffer F (4 mL), then adding 7 mL of denaturing buffer (Buffer F + 6 M guanidine hydrochloride), to give a final GuHCl concentration of 3.5 M. The suspension was stirred at 4°C for 45 min until transparent (indicating that denaturation had occurred). This transparent suspension (12 mL) was brought to room temperature and a Complete protease inhibitor cocktail tablet (Roche) was added, before Refolding buffer containing 0.659 M GuHCl (88 mL) was added slowly. The resultant 100 mL volume contained a final GuHCl concentration of 1 M. The volume was deliberately large so that the concentration of His₆-β⁵ was kept low: preliminary trials showed that these conditions gave the best refolding yields (compared to relatively higher concentrations). This fraction was dialysed overnight against two changes of 4 L of Refolding buffer.

After dialysis, the sample was clarified by centrifugation (35000 × g; 40 min), and the supernatant (containing refolded (His₆-β⁵)₂) was dialysed against two changes of 3 L of Buffer G and again clarified by centrifugation (35000 × g; 40 min) before being applied (1 mL/min) to a 1 mL HisTrap column which had been equilibrated in the same buffer. After washing with Buffer G for 15 mL, bound (His₆-β⁵)₂ was eluted from the column using a linear gradient (10 mL) of 35-500 mM imidazole in Buffer G. Fractions containing ~70% pure (His₆-β⁵)₂, which eluted at ~170 mM imidazole, were pooled and dialysed against two changes of 2 L of Buffer E containing 500 mM NaCl, yielding ~1 mg ((His)₆-β⁵)₂ in a volume of 4 mL. Aliquots were frozen in liquid nitrogen and stored at –80°C. Further purification was not carried out at this
point, as the protein was to be used in replication assays (for which this was of sufficient purity) as well as for preparation of mixed $\beta_2$ dimers (see section 5.2.8), which required additional purification. Owing to the small amount of protein purified by this procedure, it was repeated as required, using additional aliquots from the suspension frozen prior to denaturing. Similar yields were obtained each time.

The monomeric molecular mass of the His$_6$-$\beta^C$ measured by nanoESI-MS after dialysis into 0.1% (v/v) formic acid was 40911 Da, corresponding very well with the calculated mass of 40911.7 Da. NanoESI-MS of the native protein dimer in 200 mM ammonium acetate, pH 7.2 was also performed, and is described in section 5.3.3.

5.2.7 Subunit exchange between ($\beta_{\text{WT}}$)$_2$ and (His$_6$-$\beta$)$_2$

Subunit exchange between wild-type $\beta_2$ [denoted ($\beta_{\text{WT}}$)$_2$] (which was expressed and purified by Dr Slobodan Jergic essentially as described by Oakley et al.$^{[138]}$) and (His$_6$-$\beta$)$_2$ was examined under a number of conditions using nanoESI-MS. In the following, all concentrations refer to concentrations of $\beta_2$ as dimer, and all dialysis steps were performed at 4°C. For initial experiments, ($\beta_{\text{WT}}$)$_2$ (20 $\mu$L; 49 $\mu$M) and (His$_6$-$\beta$)$_2$ (50 $\mu$L; 11 $\mu$M) were each diluted to 300 $\mu$L with 200 mM ammonium acetate, pH 7.4, then dialysed separately against four 2 L changes of the same buffer. Equimolar mixtures of the two proteins (~0.8 $\mu$M each in a volume of 100 $\mu$L) were prepared, and placed at either 0°C or 30°C. Subunit exchange was monitored by removing aliquots (3 $\mu$L) from each mixture immediately after mixing, every hour after mixing for 5 hours, then 8 hours and 20 hours after mixing. Analysis of these aliquots by nanoESI-MS was used to show the progression of exchange between the wild-type and His$_6$-tagged $\beta$ subunits.
To explore conditions for promoting and inhibiting subunit exchange, several buffer conditions were tested. Buffers used were: Buffer SE1 (30 mM Tris-HCl, pH 7.6, 1 mM DTT, 1 M NaCl, 10 mM MgCl\textsubscript{2}); Buffer SE2 (30 mM Tris-HCl, pH 7.6, 1 mM DTT, 1 M MgCl\textsubscript{2}, 100 mM NaCl); Buffer SE3 (30 mM Tris-HCl, pH 7.6, 1 mM DTT, 10% (v/v) glycerol, 50 mM NaCl); and Buffer SE4 (30 mM Tris-HCl, pH 7.6, 10 mM MgCl\textsubscript{2}, 50-150 mM NaCl, as specified).

Buffers SE1 and SE2 were intended to provide conditions likely to promote subunit exchange. (\(\beta\textsubscript{WT}\))\textsubscript{2} and (His\textsubscript{6}-\(\beta\))\textsubscript{2} were mixed to give final concentrations of 2 \(\mu\)M each in Buffer SE1 or SE2 (final volumes 1 mL) and placed at 4°C. After 3.5, 8, and 12 hours, 300 \(\mu\)L aliquots were removed and dialysed against three changes of 200 mM ammonium acetate, pH 7.4 for a total of 12 hours, before being analysed by nanoESI-MS.

Buffer SE3 was used to examine a potential purification condition for isolation of hybrid \(\beta\textsubscript{2}\) dimers. (\(\beta\textsubscript{WT}\))\textsubscript{2} and (His\textsubscript{6}-\(\beta\))\textsubscript{2} were mixed and diluted in Buffer SE3 to a concentration of 3 \(\mu\)M each (final volume 300 \(\mu\)L). The mixture was dialysed against four 1 L changes of the same buffer over a period of 12 hours, before dialysis against three changes of 200 mM ammonium acetate, pH 7.4 for a further 12 hours, followed by nanoESI-MS analysis.

Buffer SE4 was used to mimic conditions of the replication assay in which the hybrid \(\beta\textsubscript{2}\) dimers were to be used (section 6.2). (\(\beta\textsubscript{WT}\))\textsubscript{2} and (His\textsubscript{6}-\(\beta\))\textsubscript{2} were mixed and diluted to 2 \(\mu\)M in Buffer SE4 containing either 50 mM, 100 mM or 150 mM NaCl (final volumes 500 \(\mu\)L). The mixtures were placed at 30°C for 30 minutes (to replicate the assay), then dialysed into 200 mM ammonium acetate, pH 7.4, and analysed by nanoESI-MS as for the previous mixtures.
5.2.8 Preparation of heterodimeric His₆-β<sub>WT</sub>/β<sub>WT</sub> and His₆-β<sup>C</sup>/β<sub>WT</sub>

**His₆-β<sub>WT</sub>/β<sub>WT</sub>**

An 8-fold molar excess of (β<sub>WT</sub>)<sub>2</sub> was mixed with (His₆-β)<sub>2</sub>, and the mixture was dialysed against 3 changes of 2 L of Buffer SE1 (section 5.2.7) containing 10% (v/v) glycerol for 12 hours at 4°C to allow subunit exchange, before being dialysed against 3 changes of 2 L of isolation buffer (30 mM Tris-HCl, pH 7.6, 30 mM NaCl, 0.5 mM DTT, 10% (v/v) glycerol, 30 mM imidazole) for a further 12 hours at 4°C. The sample was loaded onto a 1 mL Ni<sup>2+</sup>-charged HisTrap column (GE Healthcare) which had been equilibrated in the same buffer, then washed with 10 mL of the same buffer before being eluted with a linear gradient (10 mL) of 30 – 500 mM imidazole in isolation buffer. Fractions containing His₆-β/β<sub>WT</sub>, which eluted at about 110 mM imidazole, were immediately pooled and dialysed into 30 mM Tris-HCl, pH 7.6, 1 mM EDTA, 1 mM DTT, 30% (v/v) glycerol, 30 mM NaCl, giving a final yield of ~1 mg of His₆-β<sub>WT</sub>/β<sub>WT</sub> in a total volume of 0.9 mL (13.4 µM). Aliquots were frozen in liquid nitrogen and stored at -80°C until immediately before use.

**His₆-β<sup>C</sup>/β<sub>WT</sub> hemi-mutant**

A 15-fold molar excess of (β<sub>WT</sub>)<sub>2</sub> was added to (His₆-β<sup>C</sup>)<sub>2</sub>, and the mixture was dialysed against 3 changes of 2 L of Buffer SE1 (section 5.2.7) for a total of 12 hours at 4°C, before being dialysed for a further 12 hours against 3 changes of 2 L of isolation buffer (as above but containing 40 mM imidazole). The sample was loaded onto and eluted from a 1 mL Ni<sup>2+</sup>-charged HisTrap column (GE Healthcare) as described above for the His₆-β<sub>WT</sub>/β<sub>WT</sub> dimer. Fractions containing the His₆-β<sup>C</sup>/β<sub>WT</sub> hemi-mutant, which eluted at ~100 mM imidazole, were immediately pooled and dialysed into 30 mM Tris-HCl, 1 mM EDTA, 1 mM DTT, 30% (v/v) glycerol, 50 mM NaCl, yielding ~0.57 mg of His₆-β<sup>C</sup>/β<sub>WT</sub> in a volume of 3 mL (~2.3 µM). Aliquots were frozen in liquid nitrogen and stored at -80°C until use.
To minimise the possibility of undesired subunit exchange when preparing both heterodimers, all purification and dialysis steps (subsequent to the initial exchange) were carried out at 4°C and as rapidly as possible. Samples of each purified mixed dimer were dialysed into 200 mM ammonium acetate, pH 7.2, and analysed by nanoESI-MS (see section 5.3.4 for results).
5.3 RESULTS AND DISCUSSION

5.3.1 Problems with single-chain β₂ constructs

The initial motivation for constructing a fused β₂ dimer, with both subunits expressed in a single polypeptide, was to be able to probe the newly characterised ε-β₂ interaction (see section 6.1) by mutating only one of the two subunits at a time. Such an interaction could potentially be probed by using surface plasmon resonance (SPR), and as such it was desirable to produce a biotinylated form of the fusion protein, which would be convenient to attach to a streptavidin-coated SPR chip. An initial attempt to insert the single-chain β₂ sequence (from pYW1480; section 5.2.2) into the biotinylation vector pKO1272[179] was unsuccessful. In this vector, the biotinylation sequence and protein of interest are under transcriptional control of the phage T7 φ10 promoter. Consequently, the biotinylation sequence (along with the dnaN gene) was cloned into the pND707 vector, thus transferring transcriptional control of the gene to the tightly regulated tandem bacteriophage λ p₆ and p₇ promoters (section 2.2.1). The resulting plasmid, pCM1502 (section 5.2.2), allows genes to be inserted between NdeI and EcoRI/NcoI sites downstream of the biotinylation sequence, analogous to pKO1272.

Despite this transfer, insertion of the single-chain β₂ sequence into pCM1502 proved to be problematic. In several attempts, using the E. coli strain AN1459 (section 2.2.2) as a host, inserts corresponding in length to a single copy of the dnaN gene were obtained, as well as other products which, upon sequencing, gave sequences corresponding to parts of the E. coli chromosome. Repetitive sequences are often problematic to clone, as they exhibit a propensity to undergo unwanted recombination or deletion, even in the absence of RecA[1600-504]. SURE 2 E. coli cells (Stratagene) are a commercially available strain (SURE: Stop Unwanted Rearrangement Events) which are restriction minus (mcrA mcrCB mcrF mrr hsdR),
Figure 5.7. Positive ion nanoESI-mass spectrum of denatured His$_6$-single chain β$_2$.
The protein was dialysed into 0.1% (v/v) formic acid before analysis. The measured mass of 83327.7 Da is consistent with the calculated mass of the fusion protein (83329.4 Da), showing that the protein remained intact after purification.

endonuclease (endA) deficient, and deficient in recombination (recB recJ). These deficiencies inhibit pathways which can lead to rearrangement and deletion events in unstable DNA sequences. When ligation products were transformed into SURE 2 cells, plasmid pCM1513, containing the complete biotinylated single-chain β$_2$ sequence, was able to be isolated (although several transformants containing rearrangements and deletions were also obtained). However, overproduction of bio-single-chain β$_2$ from this plasmid was not successful (in either BL21(DE3) or SURE cells), indicating that the construct is unstable even in a strain specifically designed to encourage stability.

As an alternative, a plasmid encoding an N-terminally His$_6$-tagged version of single chain β$_2$ was constructed. This was achieved readily in AN1459, without the problems associated with construction of the biotinylated construct. The plasmid, pCM1510, enabled good overproduction of His$_6$-single-chain β$_2$, which was subsequently able to be purified, albeit in a small quantity. The nanoESI-mass spectrum of the purified protein in 0.1% formic acid (Figure
5.7) showed that the full-length protein was intact, giving a mass in very close agreement with the predicted sequence-based mass.

To test the ability of His$_6$-single-chain $\beta_2$ to bind to the Pol III core, $\alpha\varepsilon\theta$, preliminary surface plasmon resonance experiments (see Chapter 7 for further detail on SPR) were performed (not shown). His$_6$-single-chain $\beta_2$ was immobilised on the surface of a BiACore nitritotriacetic acid (NTA)-coated sensor chip which had been saturated with Ni$^{2+}$ ions. A preassembled $\alpha\varepsilon\theta$ complex, where $\alpha_t$ and $\varepsilon_t$ designate mutants of $\alpha$ and $\varepsilon$ with stronger $\beta$-binding sites (described in sections 6.1 and 6.2.1) was injected over the immobilised protein to test for binding. Given the large mass of $\alpha\varepsilon\theta$ (~166 kDa) compared to the mass of His$_6$-single-chain $\beta_2$ (83 kDa), a large binding response would be expected. However, the response was barely detectable, indicating a possible binding defect in (His)$_6$-single-chain $\beta_2$. For comparison, an equivalent amount of (His$_6$-$\beta_2$)$_2$ was immobilised on a separate channel of the same Ni$^{2+}$/NTA chip, and $\alpha\varepsilon\theta$ flowed over it. A much larger response was observed, indicating that the single-chain variant was deficient in binding to the core polymerase. Furthermore, addition of His$_6$-single-chain $\beta_2$ to a $\beta_2$-dependent DNA replication assay (see Chapter 6) showed that this protein was functionally deficient (not shown), as it was not able to substitute successfully for wild-type $\beta_2$. Given these pieces of evidence, the single-chain $\beta_2$ strategy was abandoned. Interestingly, the use of a single-chain $\beta_2$ protein has been described previously [505]. In this work, the fused dimer was described as retaining approximately 50% of the activity of wild-type $\beta_2$, but was able to be effectively loaded onto DNA by the clamp loader complex, suggesting that it was able to interact productively with the $\delta$ subunit[501]. This version of the construct contained the two dnaN sequences separated by an intervening sequence ((SG)$_7$) similar to that used in our work (SGSGTRESGSGSG); however, the N-terminus, in addition to a $10\times$His tag, was attached to a protein kinase sequence. It is possible that the (His)$_6$-single-
chain $\beta$, produced in our work, which was hindered by solubility issues, may not have folded properly, thus inhibiting its activity and ability to bind to other proteins. The additional N-terminal sequence on the previously described construct[505] may have aided protein folding. Nonetheless, it is interesting to note the significant loss of activity described for the fused dimer, suggesting that its interactions with the core of Pol III are indeed compromised.

5.3.2 Subunit exchange in the wild-type $\beta_2$ dimer

The lack of success with the single-chain $\beta$, constructs necessitated a change in strategy to create $\beta_2$ clamps containing a single mutated binding site: as an alternative, mixed dimers containing one wild-type subunit and one mutant subunit were created. Although these dimers did not offer some of the utilities of the single-chain proteins, such as the ability to withstand regeneration in SPR experiments, they nonetheless provided an effective alternative for use in the replication assays. The production of such $\beta_2$ heterodimers has been described previously[479]. The previous approach involved expression of two forms of $\beta$ (myc-tagged wild-type $\beta$ and the His$_6$-tagged deletion mutant $\beta^C$, described further in section 5.3.3) within the same $E. coli$ cell, using a dual expression vector. The three different dimers formed within the cell (the wild-type and mutant homodimers along with the heterodimer) were separated by affinity chromatography on a cobalt-charged column, and the presence of both subunits in the isolated heterodimer was confirmed in western blot and SPR analyses with anti-myc and anti-Penta-His antibodies[479].

Whilst this approach was effective, we have developed a somewhat simpler approach, in which separately purified dimers are mixed, their subunits allowed to exchange, and the heterodimeric products isolated by affinity chromatography. The use of mass spectrometry to directly confirm the identity of the purified heterodimers meant that it was not necessary to
label the wild-type $\beta_2$ incorporated into the mixed dimers. The success of this method was dependent on an understanding of the subunit exchange behaviour of the $\beta_2$ dimer. For the production of mixed dimers to be plausible, conditions needed to be found under which extensive (preferably complete) subunit exchange was able to occur. In addition, solution conditions which minimised subunit exchange were required, such that heterodimers formed by exchange could then be separated from unexchanged dimers without further exchange taking place and hindering the purification. Finally, an exploration of the conditions in the replication assays in which the $\beta_2$ heterodimers were to be used (see Chapter 6) was required, to gauge the extent of subunit exchange to be expected during the course of the assays.

Subunit exchange between $\beta_2$ dimers has been investigated previously, but it has not been characterised in depth. As part of a study to determine whether the $\beta_2$ dimer is monomerized during clamp loading, Stewart et al.\textsuperscript{[481]} used an immunoprecipitation assay to follow spontaneous subunit exchange between two distinct $\beta_2$ species. Both species incorporated a protein kinase sequence and haemagglutinin (HA) epitope at their N-termini. In one construct ($^{\text{ha}}\beta_2$), the HA epitope was functional, and able to be immunoprecipitated with a HA antibody. In the other ($^{32}\text{P}-\beta_2$), the HA epitope was mutated to be nonfunctional, and the kinase sequence was phosphorylated with $^{32}\text{P}$. After mixing the two $\beta_2$ dimers at $37^\circ\text{C}$ for varying time periods, HA antibody-conjugated beads were added, and after immunoprecipitation, their radioactivity was quantified. Only $^{32}\text{P}-\beta/^{\text{ha}}\beta$ heterodimers resulting from subunit exchange were detected (as any radioactivity must come from $^{32}\text{P}-\beta$ subunits, which needs to be associated with $^{\text{ha}}\beta$ in order to be precipitated). The accumulation of radioactivity over time revealed that subunit exchange was slow, approaching equilibrium after 8 hours and with a half-time of about 2 hours\textsuperscript{[481]}. This study, whilst providing evidence that the $\beta_2$ dimer is able to undergo slow spontaneous subunit exchange, was limited – there was no means of
observing the original homodimers, or the extent of exchange as a proportion of the total protein originally mixed. Therefore, further information was required to allow sufficient understanding of exchange for construction and purification of mixed dimers.

In this work, nanoESI-MS was used to monitor subunit exchange, as described for the SSB tetramer in Chapter 3. This method allows a much more direct determination of the progression of the exchange reaction than the immunoprecipitation assay described above. Rather than using isotopic labelling to provide the mass difference between the subunits (as done in the SSB experiments), exchange between wild-type, untagged $\beta_2$ (subunit mass 40586.6 Da), and wild-type His$_6$-tagged $\beta_2$ (subunit mass 41540.6 Da) was explored.

5.3.2.1 $\beta_2$ subunit exchange is minimal under native mass spectrometry conditions

The solution conditions to be investigated contained non-volatile buffers and salts, and thus were not compatible with ESI-MS analysis. Consequently, the approach taken was to place a mixture of the untagged and His$_6$-tagged $\beta_2$ into the solution condition of interest (either by dialysis or direct dilution with the solution), followed by dialysis into an MS compatible buffer after a specified time period. Of course, this approach does not allow accurate measurement of subunit exchange rates, but, for the purpose of these investigations, the ability to measure the extent of exchange after a defined period of time was sufficient. For this strategy to be effective, subunit exchange needs to be essentially ‘frozen’ during the final dialysis step, such that the eventual result reflects only the exchange which occurred under the initial solution conditions. Therefore, initial experiments to examine subunit exchange in ammonium acetate solutions were carried out.
Figure 5.8. Positive ion nanoESI mass spectra of (a) β₂ and (b) (His₆-β)₂. Each protein was dialysed into 200 mM ammonium acetate, pH 7.2 before analysis. The measured mass of the β₂ dimer was 81182 Da (theoretical mass 81173 Da), and that of the (His₆-β)₂ dimer was 83087 Da (theoretical mass 83080 Da).

NanoESI mass spectra of the β₂ and (His₆-β)₂ dimers in 200 mM ammonium acetate, pH 7.2 are shown in Figure 5.8. Both proteins gave spectra which were well-resolved and of high quality. These conditions were thus chosen for an initial investigation of subunit exchange. An equimolar mixture of (β₂) and (His₆-β)₂ (0.8 µM each, as dimers) in 200 mM ammonium acetate, pH 7.2, was placed on ice and aliquots were analysed immediately after mixing, then periodically for 8 hours. A final sample was analysed after 20 hours on ice. Little change was seen in the spectra acquired over the first 8 hours (Figure 5.9a) – only ions corresponding to the two homodimeric species were present. The only evidence for any subunit exchange was seen after 20 hours, when heterodimeric β/His₆-β ions of very low abundance could be discerned in the spectra (Figure 5.9a; top spectrum). Such a minimal amount of subunit exchange suggests that, provided that dialysis is carried out at low temperatures, 200 mM ammonium acetate is an appropriate buffer in which to analyse the results of subunit exchange occurring under other solution conditions.

For comparison, an equimolar mixture of the same composition was placed at 30°C and sampled over the same period as the mixture on ice. At this temperature, subunit exchange
Figure 5.9. Subunit exchange between $\beta_2$ and $(\text{His}_6-\beta)_2$. Equimolar $\beta_2$ and $(\text{His}_6-\beta)_2$ were mixed in 200 mM ammonium acetate, pH 7.2, and placed either (a) on ice or (b) at 30°C, with aliquots withdrawn after the indicated times and analysed by positive ion nanoESI-MS. The measured mass of the $\beta$/His$_6$-$\beta$ dimer was 82194 Da (theoretical mass 82126 Da). See Figure 5.8 for masses of $\beta_2$ and $(\text{His}_6-\beta)_2$.

Proceeded much more quickly (Figure 5.9b), with heterodimeric ions observed in the spectrum after only 2 hours. After 20 hours, the heterodimeric $\beta$/His$_6$-$\beta$ ions accounted for $\sim$50% of the total ion abundance, indicating that subunit exchange between $\beta_2$ and $(\text{His}_6-\beta)_2$ had proceeded essentially to completion (for an initially equimolar mixture of dimers, the expected equilibrium distribution of ions after complete exchange is 25% for each homodimer and 50% for the heterodimer). The increase in subunit exchange with increased temperature indicates that, to minimise $\beta$, subunit exchange in 200 mM ammonium acetate, it is important
for mixtures to be maintained at low temperatures. Nonetheless, even at this elevated temperature, subunit exchange is relatively slow, reflecting the stability of the β₂ dimer.

5.3.2.2 Development of solution conditions for formation and purification of mixed β₂ dimers

Given the strength of the β₂ dimer, it was important to find a way of promoting subunit exchange using sufficiently gentle solution conditions, and preferably without using elevated temperatures, in order to protect the structural integrity of the proteins for subsequent experiments. The noncovalent interactions holding the dimer together are a combination of hydrophobic interactions, hydrogen-bonding and ion pairs. As high salt concentrations tend to destabilise hydrogen-bonding and electrostatic interactions, subunit exchange in two high salt buffers – Buffer SE1 (section 5.2.7), containing 1 M NaCl, and Buffer SE2 (section 5.2.7), containing 1 M MgCl₂ – was investigated. As very little has been reported in the literature about the subunit exchange behaviour of β₂, it was desirable to investigate exchange in the presence of common salts (rather than dialysing directly into high ammonium acetate).

To compare subunit exchange in the two buffers and examine the time-dependence of exchange, equimolar mixtures of β₂ and (His₆-β₂)₂ were diluted separately with Buffer SE1 and Buffer SE2 such that the final concentration of each dimer was 2 µM. The mixtures were placed at 4°C, and aliquots were removed after 3.5, 8 or 12 hours, then dialysed against 200 mM ammonium acetate, pH 7.2, for 12 hours before being analysed by nanoESI-MS. This analysis showed that all samples had undergone a significant amount of subunit exchange, with each spectrum dominated by heterodimeric His₆-β/β ions (Figure 5.10 a, b). Although the rates of exchange in each buffer were similar, at each timepoint, the sample which had been treated with Buffer SE1 (high NaCl concentration), had a slightly higher proportion of heterodimers than the sample which had been treated with Buffer SE2 (high MgCl₂).
Figure 5.10. Subunit exchange between $\beta_2$ and (His$_6$-$\beta_2$) at high salt concentrations. Equimolar mixtures of $\beta_2$ and His$_6$-$\beta_2$ were diluted with either (a) Buffer SE1 (30 mM Tris-HCl, pH 7.6, 1 mM DTT, 1 M NaCl, 10 mM MgCl$_2$); or (b) Buffer SE2 (30 mM Tris-HCl, pH 7.6, 1 mM DTT, 1 M MgCl$_2$, 100 mM NaCl), then placed at 4°C. Aliquots were withdrawn from each mixture after 3.5, 8 and 12 hours, and dialysed separately against 200 mM ammonium acetate, pH 7.2 before analysis by positive ion nanoESI-MS. See Figures 5.8 and 5.9 for dimeric masses.

In both buffers, increased treatment time led to a slight increase in subunit exchange, but most exchange was complete after 3.5 hours. This was in contrast to the relatively small amount of exchange seen after 3 hours at 30°C in 200 mM ammonium acetate (Figure 5.9b), and indicated that elevated salt concentrations dramatically increase the subunit exchange rate between $\beta_2$ and (His$_6$-$\beta_2$). Given the slightly more efficient subunit exchange
Figure 5.11. Subunit exchange between $\beta_2$ and $(\text{His}_6-\beta)_2$ at low salt concentrations. An equimolar mixture of $\beta_2$ and $(\text{His}_6-\beta)_2$ was dialysed against Buffer SE3 (30 mM Tris-HCl, pH 7.6, 1 mM DTT, 10% (v/v) glycerol, 50 mM NaCl) for 12 hours, then against 200 mM ammonium acetate, pH 7.2, before being analysed by positive ion nanoESI-MS. See Figures 5.8 and 5.9 for dimeric masses.

observed in Buffer SE1, it was chosen as the buffer with which to promote exchange of monomers when producing hybrid mutant dimers (section 5.3.4).

Having established that high salt conditions were effective for promoting subunit exchange of $\beta_2$, it seemed likely that low salt concentrations would have the opposite effect, and act to stabilise the dimer. The low salt Buffer SE3 (section 5.2.7) contained 50 mM NaCl, which was included to ensure protein solubility. The spectrum in Figure 5.11 was obtained after dialysis of an equimolar mixture of $\beta_2$ and $(\text{His}_6-\beta)_2$ against Buffer SE3 for 12 hours at $4^\circ$C, followed by dialysis against 200 mM ammonium acetate. Predominant in this spectrum were the original homodimeric species, with only two ions from heterodimer present at very low abundance. This minimal amount of exchange further emphasised the strong salt-dependence of subunit exchange in the $\beta_2$ dimer, and, importantly, indicated that working with low salt
concentrations should allow the isolation of mixed dimers. Taken together, the results described above were used to develop a method for production of mixed $\beta_2$ dimers in which the required combination of subunits were dialysed against a buffer containing 1 M NaCl to allow complete exchange of subunits, then dialysed into a buffer containing 50 mM (or lower) NaCl in order to inhibit further exchange during subsequent purification steps.

5.3.2.3 $\beta_2$ subunit exchange under replication assay conditions

As the main use of the hybrid $\beta_2$ dimers was to be in DNA replication assays, it was important to gauge the extent of subunit exchange expected to occur under the conditions used in these assays. For a hemi-mutant dimer containing one wild-type and one mutant subunit, significant subunit exchange during an assay could be problematic, as wild-type dimers resulting from exchange could obscure the effects of the hemi-mutant. The replication assay conditions were simulated by diluting equimolar mixtures of $\beta_2$ and (His$_6$-$\beta_2$)$_2$ with Buffer SE4 (section 5.2.7) containing either 50 mM, 100 mM or 150 mM NaCl, then placing the diluted mixtures at 30°C for 30 minutes. After dialysis against 200 mM ammonium acetate, each sample was analysed by nanoESI-MS (Figure 5.12). All three conditions showed evidence of a limited amount of subunit exchange – the spectra were dominated by unexchanged dimeric species, but small amounts of heterodimer could be clearly observed. Again, the effect of salt concentration was apparent, with the heterodimeric ions increasing slightly in abundance as the NaCl concentration was increased from 50 mM to 150 mM. These results suggest that, provided that the NaCl concentration is limited to around 100 mM, subunit exchange during the course of replication assays is not likely to be extensive. However, the potential for a small amount of exchange should be considered.
Figure 5.12. Subunit exchange between $\beta_2$ and (His$_6$-$\beta$)$_2$ under replication assay conditions. Equimolar mixtures of $\beta_2$ and (His$_6$-$\beta$)$_2$ were diluted with Buffer SE4 (30 mM Tris-HCl, pH 7.6, 10 mM MgCl$_2$) containing either (a) 50 mM, (b) 100 mM, or (c) 150 mM NaCl. Each mixture was placed at 30°C for 30 minutes, then dialysed into 200 mM ammonium acetate, pH 7.2 for analysis by positive ion nanoESI-MS. See Figures 5.8 and 5.9 for dimeric masses.
5.3.3 The truncated mutant $\beta^C$ forms a stable dimer

The five residues at the extreme C-terminus of the $\beta$ sequence (Met-362, Pro-363, Met-364, Arg-365, Leu-366) lie close to, and partially overlap, the hydrophobic cleft on the clamp largely responsible for interactions with its binding partners (Figure 5.13a). Deletion of these five residues is known to result in disruption of the cleft: the truncated mutant, denoted $\beta^C$, has been constructed previously and was shown to be deficient in its interaction with the $\gamma\delta\epsilon\chi\psi$ clamp loader complex, as well as the isolated $\delta$ subunit and the $\alpha$ subunit of Pol III$^{[479]}$. Computer modelling of the mutant suggested that the loss of these last five residues (MPMRL) would cause partial collapse of the hydrophobic cleft, but leave the overall tertiary structure of the protein intact$^{[479]}$. Indeed, despite its binding deficiencies, the circular dichroism spectrum of the $\beta^C$ mutant was found to be essentially identical to that of wild-type $\beta$ (Figure 5.13b), indicating that it retained its ability to fold normally$^{[479]}$.

As will be discussed in more detail in Chapter 6, the putative $\beta_2$-binding site on $\epsilon$ consists of a clamp binding motif similar to those in $\alpha$ and $\delta$ (and to the consensus sequence). Therefore, the proposed $\beta_2$-$\epsilon$ interaction was considered likely to involve the same hydrophobic cleft on $\beta_2$ utilised in its other interactions. If this is the case, the $\beta^C$ deletion mutant should be similarly impaired in its interaction with $\epsilon$, and it was this mutant that was therefore chosen for incorporation into replication assays to explore the function of the $\beta_2$-$\epsilon$ interaction, both as a homodimer and in a heterodimer with wild-type $\beta$ (such that only one of the two binding sites is attenuated). During purification of (His$_6$-$\beta^C$)$_2$ (section 5.2.6), solubility problems were encountered which meant that it was necessary to use a denaturing and refolding method for purification. Therefore, to confirm that purified (His$_6$-$\beta^C$)$_2$ had maintained its dimeric quaternary structure, it was dialysed against 200 mM ammonium acetate, pH 7.2, and
Figure 5.13. The truncated mutant $\beta^c$ clamp folds normally, but has a disrupted protein-binding cleft. (a) $\beta^c$ is missing the C-terminal five residues (362-MPMRL-366), represented as blue sticks in the figure. These residues are located adjacent to, and partially overlap, the hydrophobic protein-binding cleft (grey spheres) resulting in a loss of protein-binding competency in the mutant clamp. Figure constructed using PyMol and the PDB coordinates 1MMI[138]. (b) Truncation of the C-terminal five residues does not affect the overall tertiary structure of the clamp, as judged by the similarity of the circular dichroism spectra of full-length ($\beta^+$) and truncated $\beta$ ($\beta^c$). Figure reproduced from reference 479.

analysed by nanoESI-MS. The resulting spectrum (Figure 5.14) allowed calculation of a mass of 81820 Da, in good agreement with the calculated molecular mass of the dimer (81823 Da). In addition, no monomeric ions were present in the spectrum, indicating that the dimer apparently retains its stability. Importantly, comparison of the spectrum of (His$_6$-$\beta^c$)$_2$ in Figure 5.14 with that of wild-type (His$_6$-$\beta$)$_2$ (Figure 5.8b) reveals that the charge state distributions of both proteins are very similar, providing good evidence that the mutant protein is folded correctly.
Figure 5.14. (His$_6$-β$^C$)$_2$ is a stable dimer. The positive ion nanoESI-mass spectrum of (His$_6$-β$^C$)$_2$ in 200 mM ammonium acetate, pH 7.2. The measured mass of the dimer was 81820 Da (theoretical mass 81823 Da).

5.3.4 Production and purification of mixed β$_2$ dimers using insights gained from subunit exchange.

In initial attempts to create and purify mixed β$_2$ dimers, a double affinity approach was used. His$_6$-tagged β$_2$ (either wild-type or β$^C$) was mixed with N-terminally biotinylated wild-type β$_2$ and allowed to undergo subunit exchange in the high salt Buffer SE1. The mixture was then dialysed into a low salt buffer to inhibit exchange and passed through two sequential columns: a Ni$^{2+}$ affinity column, followed by a column containing a monomeric avidin resin designed to allow reversible binding of biotinylated proteins. Only heterodimers containing both the His$_6$-tagged and biotinylated subunits were able to be retained by both columns. Unfortunately, the retention of biotinylated proteins by the avidin resin was, under the conditions tested, very weak, resulting in low yields of isolated hybrid dimers.

To create and purify hybrid β$_2$ dimers in higher yields, a large molar excess of untagged wild-type β$_2$ was mixed with (His$_6$-β$_{WT}$)$_2$ or (His$_6$-β$^C$)$_2$, such that complete subunit exchange
Figure 5.15. Production of hybrid $\beta_2$ dimers. A large excess of untagged wild-type $\beta_2$ was mixed with His$_6$-tagged ($\beta^c$), or wild-type $\beta_2$ in a high salt buffer, allowing incorporation of the majority of His$_6$-tagged subunits into hybrid dimers. After transfer into a low salt buffer, heterodimers were isolated from excess untagged $\beta_2$ by passing the mixture through a Ni$^{2+}$ affinity column.

resulted in essentially all the His$_6$-tagged subunits being incorporated into heterodimers. This allowed a single Ni$^{2+}$ affinity chromatography step to be used for removal of the excess untagged $\beta_2$, minimising purification time (and thus unwanted subunit exchange) and maximising heterodimer yields (Figure 5.15).
Figure 5.16. NanoESI-MS of $\beta$ heterodimers following purification. (a) $\beta_{\text{WT}}$/His$_c^{-}\beta_{\text{WT}}$, (b) $\beta_{\text{WT}}$/His$_c^{-}\beta^C$. Each dimer was dialysed into 200 mM ammonium acetate, pH 7.2 before analysis. See Figures 5.8 and 5.9 for masses of proteins in (a). The measured mass of the (His$_c^{-}\beta^C$)$_2$ dimer was 81822 Da (theoretical mass 81823 Da) and that of the His$_c^{-}\beta^C$/\(\beta_{\text{WT}}\) dimer was 81498 Da (theoretical mass 81496 Da).

Two heterodimers were eventually produced: a His$_c^{-}\beta^C$/\(\beta_{\text{WT}}\) hemi-mutant, as well as a His$_c^{-}\beta_{\text{WT}}$/\(\beta_{\text{WT}}\) dimer for use as a control in replication assays to verify that the presence of the His$_c^{-}$ tag combined with the method of production of the mixed dimers did not affect activity. To produce the His$_c^{-}\beta_{\text{WT}}$/\(\beta_{\text{WT}}\) dimer, an eight-fold molar excess of (\(\beta_{\text{WT}}\))$_2$ was mixed with (His$_c^{-}\beta_{\text{WT}}$)$_2$, whilst for the His$_c^{-}\beta^C$/\(\beta_{\text{WT}}\) hemi-mutant, a fifteen-fold molar excess of (\(\beta_{\text{WT}}\))$_2$ was used. The details of the final purification procedure used for each dimer can be found in section 5.2.8. Following isolation, samples of each mixed dimer were immediately dialysed into 200 mM ammonium acetate, pH 7.2 for analysis by nanoESI-MS in order to confirm their identity and lack of contamination by homodimers. Figure 5.16a shows that, for the purified His$_c^{-}\beta_{\text{WT}}$/\(\beta_{\text{WT}}\) dimer, the heterodimer was essentially the only species present in the mass spectrum, with minimal contamination from homodimers. The spectrum of the purified hemi-mutant His$_c^{-}\beta^C$/\(\beta_{\text{WT}}\) (Figure 5.16b) shows that, importantly, all (\(\beta_{\text{WT}}\))$_2$ was removed, and the predominant ions were those corresponding to heterodimeric His$_c^{-}\beta^C$/\(\beta_{\text{WT}}\). A relatively small
amount of \((\text{His}_6\beta)^2\) was still present, but this is preferable to contamination by the wild-type homodimer, which would be more likely to confound results in replication assays by obscuring potential loss-of-function by the hemi-mutant.
5.4 CONCLUSIONS

Attempts to create β2 clamps containing a single competent protein binding site by constructing a fusion protein with both subunits contained in a single polypeptide chain, were problematic. The His₆-single-chain β2 which was eventually produced was found to be deficient in binding to the Pol III αεθ core, even with α and ε subunits engineered to interact strongly with β2. This fusion construct was also unable to support replication in an in vitro assay, providing further evidence of its inactivity. Whether this inactivity was an inherent property of the fused dimer (resulting, perhaps, from its lack of flexibility in choosing which dimer interface to open at), or whether it stemmed from improper protein folding during purification, is unclear, but fortunately, another avenue of hemi-mutant production was available. As predicted by its stability, the rate of subunit exchange in the β2 dimer was found to be slow, but highly dependent on salt concentration and temperature. At low salt concentrations and low temperatures, subunit exchange was barely detectable, but at high salt concentrations (or higher temperatures), subunit exchange could proceed to completion over a period of several hours. This knowledge was harnessed to create a heterodimeric form of β2 in which one of the subunits contained a disrupted hydrophobic protein-binding cleft. This hemi-mutant is intended as a tool for investigating the interaction between β2 and the ε subunit of Pol III, which will be described in the following chapter. This method of production of mixed β2 dimers, as well as the characterisation of β2 subunit exchange described here, should prove useful for future studies investigating the function of the clefts in the β2 clamp.
Chapter 6  :  FUNCTIONAL INVESTIGATIONS OF THE E. coli ε-β₂ INTERACTION

Acknowledgement of contributions from others: Work described in this chapter was carried out with and under the supervision of Dr Slobodan Jergic. Dr Jergic was responsible for experimental design and the original development of replication assays. He also provided extensive assistance with analysis of results.

6.1  INTRODUCTION

6.1.1 DNA replication assays as a means for dissecting the functional properties of polymerase subunits

DNA replication in E. coli has been extensively studied: the current view of how the multisubunit Pol III holoenzyme accomplishes the task of replicating the E. coli chromosome was described in detail in section 1.1. In vitro replication assays have played an important role in elucidating the functions of the various subunits, and have provided information about many of the protein-protein interactions which exist (often transiently) within the replisome. In such assays, the Pol III holoenzyme (or selected subunits) is assembled on primed DNA, and the amount of DNA synthesis quantified after a specified period, generally by measuring the incorporation of radioactive nucleotides. Whilst modern biophysical techniques (surface plasmon resonance, FRET, etc) have great utility for characterising isolated protein-protein interactions, functional assays remain extremely useful for gaining a more complete picture of the effects of adding, removing or mutating a component upon replication as a whole.

The simplest replication assays are performed on a circular, single-stranded DNA template to which an oligonucleotide primer has been hybridised (Figure 6.1a). Upon addition of the Pol III holoenzyme (or its isolated subunits), the polymerase is able to assemble on the 3’ end of the primer and synthesise new DNA around the template, stopping once it reaches the duplex
Figure 6.1. A simple DNA replication assay. (a) A circular single-stranded DNA template is primed with a short complementary oligonucleotide (red). (b) Upon addition of components required for replication, the $\beta_2$ clamp and $\alpha$ polymerase are assembled at the 3' end of the primer, allowing synthesis around the circular template to begin. (c) Replication to the 5' end of the primer fills the circular template with new DNA (blue), transforming it to double-stranded DNA.

DNA at the 5' end of the primer\[506\] (Figure 6.1 b, c). Simple strand extension reactions of this type can be reconstituted with just the $\beta_2$ clamp, the $\alpha$ polymerase, and the minimal $\gamma_\delta\delta'$ clamp loader complex, which is sufficient to load $\beta_2$ and enable processive synthesis.

In the presence of DnaB helicase, Pol III holenzyme assembled onto a double-stranded, nicked template with a 5' ‘tail’ can support an efficient rolling circle (or ‘strand-displacement’) replication reaction, in which DnaB is loaded onto the single-stranded tail and unwinds the duplex ahead of the advancing polymerase\[507\] (Figure 6.2a). This results in displacement of the non-template strand and extrusion of an increasingly long single-stranded tail as replication proceeds around the template. In the presence of DnaG primase, synthesis of Okazaki fragments occurs on the extruded strand, such that the rolling circle assay simulates the coupled leading and lagging strand synthesis which occurs at true replication forks\[59,508\].

Although less well-characterised, helicase-independent strand displacement synthesis has also been observed on 5' tailed templates, suggesting that the Pol III holenzyme possesses intrinsic
Figure 6.2. Strand displacement rolling-circle DNA replication assays. (a) DnaB helicase-dependent strand displacement. DnaB helicase can be loaded onto a short single-stranded 5’ DNA tail on an otherwise double-stranded circular template. Along with the other components required for DNA synthesis, DnaB can unwind the double-stranded template, allowing further synthesis around the template and extrusion of an increasingly long single-stranded tail. (b) A helicase-independent rolling-circle assay. Relative to a helicase-containing assay, this assay requires the full clamp loader complex (τδδ’γψ), SSB, and an elevated concentration of dNTPs. Beginning with a circular template with a partially hybridised primer and short single-stranded 5’ tail, synthesis proceeds around the template until it is fully double-stranded (circle-filling), then proceeds to displace the original primer, extruding the 5’ tail as in (a). SSB presumably binds to the single-stranded region as it is produced.

strand displacement activity\(^{198,509}\) (Figure 6.2b). This strand displacement reaction is significantly weaker than that observed in the presence of DnaB\(^{509}\), and is also slower and less processive than replication on single-stranded templates. Its requirements have recently been investigated in the McHenry laboratory\(^{198}\). In contrast to single-stranded synthesis, strand
Figure 6.3. Proposed model for stabilisation of helicase-independent strand displacement replication. The interaction of between α and SSB, mediated via a τ-χ-ψ ‘bridge’ is proposed to stabilise α onto the template sufficiently for strand displacement to occur. The SSB coats the single-stranded 5´ tail which is present in the assay. Based on a figure from reference 198.

displacement was found to be highly dependent on the presence of full-length single-stranded DNA binding protein (SSB), as well as the τ, χ and ψ subunits of the clamp loader complex (in addition to requiring the remaining Pol III subunits)\[198\]. It also relied on much higher dNTP concentrations than required for single-stranded template replication\[198\]. These stringent protein requirements led to the proposal that SSB, bound to the displaced strand, forms an α-
τ-ψχ-SSB protein bridge which stabilises the Pol III core on the template sufficiently for strand displacement to occur (Figure 6.3).

6.1.2 A strand displacement reaction with an absolute dependence on ε

A similar helicase-independent strand displacement assay was developed recently in our laboratory by Dr Slobodan Jergic[496]. This assay employs a primed single-stranded circular phage M13 template (section 6.2.2), such that primer extension synthesis around the single-stranded template occurs prior to the commencement of strand displacement/rolling circle synthesis. In contrast to the relatively low salt conditions (typically <40 mM MgCl₂ or NaCl) traditionally used in replication assays, these assays employed a salt concentration of 130 mM NaCl, more closely approximating physiological conditions. The β₂ clamp was also added to the assay at a higher concentration (200 nM) than has been used conventionally in replication assays (1-30 nM), more consistent with the cellular β₂ concentration which is estimated to be about 250 nM[16]. Despite these modifications, strand displacement activity in the assay exhibited a requirement for SSB, high dNTP concentrations, and the presence of τ and ψχ in the clamp loader complex, consistent with the previous findings. Figure 6.4 shows the products of a representative assay over time following separation on an agarose gel and staining with a dye able to bind to both single- and double-stranded DNA. The primed M13 template migrates fastest through the gel, whilst double-stranded ‘filled circles’ resulting from synthesis around the single-stranded template (i.e. non strand displacement products) are observed as a less mobile, well-defined band. The products of strand displacement synthesis are present as numerous discrete bands at larger sizes than the filled circles, the average size of which increases with time. The reason for such a banding pattern is presumably related to the relative fallibility of this form of synthesis[198] – certain difficult structures in the template result in a
population of polymerases being unable to proceed, resulting in either prolonged stalling or falling off the template.

An examination of the indispensability of the various subunits to this strand displacement assay resulted in a surprising discovery – strand displacement synthesis depended absolutely on the presence of the proofreading exonuclease subunit of the Pol III core, ε (Figure 6.5). This was unexpected, as a replication reaction with an absolute requirement for ε has not been described previously. Studies of the function of ε have concentrated on its exonuclease activity (see section 1.1.3.1), which should not be essential for replication in vitro, so the basis for its requirement in this strand displacement reaction warranted further examination.
Figure 6.5. Helicase-independent strand displacement is dependent on the presence of the ε subunit. Each assay contained β, τ, δδ′χψ, SSB, dNTPs, ATP, primed template, and preassembled αθ or individual core components as indicated. Assays were performed for 20 min at 30°C. In the absence of ε, α was able to fill the single-stranded template, but no strand displacement synthesis was observed. Addition of the θ subunit did not remedy this, but addition of ε (even in the absence of θ) resulted in robust strand displacement (assay performed by Dr Slobodan Jergic).

To determine whether the exonuclease activity of ε was responsible for its contribution to strand displacement, several mutants of ε were constructed and examined for their ability to support strand displacement. The exonuclease active site of ε resides in its structured N-terminal domain\cite{122}, which contains the binding sites for the two metal ions (Mg\(^{2+}\) or Mn\(^{2+}\)) required for exonuclease function\cite{121}. Proofreading-deficient ε mutants have been described previously\cite{510}. The Asp12 and Glu14 residues form part of the metal binding sites, and ε mutants with D12A or D12A/E14A substitutions lack the ability to coordinate metal ions, and
consequently, lack exonuclease activity\textsuperscript{510}. Both mutants were able to support strand displacement synthesis (Figure 6.6), suggesting that the $\varepsilon$ contribution to strand displacement synthesis was not a consequence of its exonuclease activity\textsuperscript{496}. For further confirmation, the structured exonuclease domain of $\varepsilon$ (residues 1-180) was completely removed, and the remaining C-terminal region (residues 181-243) fused to the C-terminus of ubiquitin (as a solubility tag for the intrinsically insoluble C-terminal domain of $\varepsilon$). The resulting fusion protein maintained the ability to support strand displacement (Figure 6.6) – evidence that the prevalent contribution by the $\varepsilon$ subunit to strand displacement was not connected, either functionally or structurally, to its proofreading domain.

Whilst $\varepsilon$ has not generally been recognised for any function beyond its proofreading role, several studies have hinted at other potential activities. The processivity of the $\alpha$ subunit on single-stranded templates was shown to be significantly stimulated by the addition of $\varepsilon$\textsuperscript{130}, resulting in the suggestion that $\varepsilon$ may assist in stabilising the polymerase upon template DNA. Notably, genetic studies in Salmonella bacteria have shown that disruption of the dnaQ gene, which encodes $\varepsilon$, results in not only a mutator phenotype, as would be expected due to the loss of proofreading function, but also severely restricts growth\textsuperscript{89,511,512}. The growth defect, along with the fact that it could be suppressed by specific mutations in the dnaE ($\alpha$) gene, suggested that $\varepsilon$ was critical for successful replication via a function quite separate from its proofreading role. It was speculated that $\varepsilon$ could have an allosteric effect upon $\alpha$ that increased the affinity of $\alpha$ for $\beta_2$\textsuperscript{512}, or that $\varepsilon$ itself could serve as an attachment point for other holoenzyme subunits\textsuperscript{511}, but the actual basis of its non-proofreading role has remained unknown since these studies were carried out over 20 years ago.
Figure 6.6. The contribution of ε to strand displacement synthesis is independent of its proofreading exonuclease activity. Each assay contained α, β, τ, δ, γ, ψ, SSB, dNTPs, ATP, primed template, and wild-type or mutant ε as indicated. Assays were performed at 30°C for the times indicated. The proofreading-deficient ε mutants ε_{D12A} and ε_{D12A/E14A}, as well as a fusion protein consisting of only the (non-proofreading) C-terminal domain of ε fused to ubiquitin were able to support strand displacement synthesis (assay performed by Dr Slobodan Jergic).
6.1.3 Discovery of an $\varepsilon$-$\beta_2$ interaction

Since the C-terminal domain of $\varepsilon$ (which is able to support the strand displacement reaction) is the region responsible for its interaction with the N-terminal PHP domain of $\alpha^{[122,123]}$, the contribution of $\varepsilon$ to strand displacement, and to replication more generally, could potentially arise from enhanced stability of $\alpha$ on the template as a consequence of its interaction with $\varepsilon$. However, inspection of the amino acid sequences of $\varepsilon$ subunits from a number of bacteria revealed a second possibility. A moderately conserved sequence, reminiscent of the penta- or hexapeptide motifs responsible for $\beta_2$ clamp binding in a variety of proteins (refer to section 5.1), was found in the $\varepsilon$ sequences, immediately following the structured N-terminal exonuclease domain (Figure 6.7). The *E. coli* sequence, QTSMAF (residues 182-187) resembles the hexameric consensus clamp binding motif, $Qxx\Phi x\Phi$ (where $x$ is any residue; $\Phi$ is a hydrophobic residue), and was therefore proposed as a potential site of interaction between $\varepsilon$ and the $\beta_2$ clamp$^{[496]}$.

The $\beta_2$ binding strengths of a number of synthetic peptide variations of the clamp binding motifs have been characterised$^{[487]}$, allowing the effects of modifications to native clamp binding sequences upon their cognate protein-$\beta_2$ interactions to be predicted. Based on this knowledge, two $\varepsilon$ mutants, designated $\varepsilon_L$ and $\varepsilon_Q$, were produced to test whether varying the putative $\beta_2$ binding site in $\varepsilon$ resulted in an altered ability to support strand displacement synthesis$^{[496]}$. In $\varepsilon_L$, the native QTSMAF sequence was replaced by QLSLPL, a known high-affinity clamp binding motif$^{[487]}$, predicted to bind more strongly to the $\beta_2$ clamp. In $\varepsilon_Q$, the highly conserved first residue of the motif (Q182), which plays an important role in all clamp binding motifs, was substituted by alanine (ATSMAF) to produce a mutant with a predicted weakened $\beta_2$ binding affinity.
Figure 6.7. A potential $\beta_2$ binding site in $\varepsilon$ based on sequence analysis. (a) Domain organisation of the full-length $\varepsilon$ protein. Exonuclease activity resides in the N-terminal domain (residues 1-180), which is connected to the $\alpha$-binding C-terminal domain (residues 209-243) by a flexible linker (residues 181-208). (b) Position of the putative clamp binding motif in $\varepsilon$ (red) relative to the structured N-terminal domain. A sequence alignment of the C-terminal regions of $\varepsilon$ from several bacterial species is shown. The penta- or hexapeptide motifs immediately following the structured N-terminal domain (residues 182-187 in *E. coli*) resemble the previously identified clamp binding motifs which are responsible for $\beta_2$-binding in a number of proteins. Alignment reproduced from reference 496.

The use of either of these mutants in place of wild-type $\varepsilon$ ($\varepsilon_{WT}$) in the strand displacement assay described above gave results which did indeed reflect their predicted $\beta_2$ binding strengths – $\varepsilon_L$ was able to contribute to strand displacement synthesis to a significantly greater extent than $\varepsilon_{WT}$ (Figure 6.8), whilst $\varepsilon_Q$ was completely unable to support strand displacement$^{[496]}$ (Figure 6.8). Such well-defined effects were strong evidence for the existence of a functionally relevant $\varepsilon$-$\beta_2$ interaction, under the conditions of these assays.
Figure 6.8. Variations in the putative $\beta_3$ binding site of $\varepsilon$ affect strand displacement synthesis according to their predicted $\beta_3$ binding strength. Each assay contained $\beta_2$, $\tau_3$$\delta$$\delta'$$\chi$$\psi$, SSB, dNTPs, ATP, primed template, and preassembled $\alpha\varepsilon\theta$ core containing either wild-type $\varepsilon$, $\varepsilon_Q$ (weakened proposed $\beta_3$-binding site) or $\varepsilon_L$ (strengthened putative $\beta_3$-binding site) as indicated. Assays were performed at 30°C for the times indicated. Strand displacement synthesis was completely abolished in the assay containing $\varepsilon_Q$, but was more efficient (relative to that observed with wild-type $\varepsilon$) in the assay containing $\varepsilon_L$ (assays performed by Dr Slobodan Jergic).

A physical interaction between $\varepsilon$ and $\beta_3$ was subsequently confirmed using mass spectrometry and surface plasmon resonance\cite{496}. The SPR experiments demonstrating the $\varepsilon$-$\beta_3$ interaction will be described in Chapter 7 of this thesis. The discovery of a previously unidentified interaction within such a well-characterised system as the Pol III holoenzyme was an exciting one, but also raised a number of questions with regard to how an $\varepsilon$-$\beta_3$ interaction would operate during replication, and whether the interaction played a role in replication reactions other than helicase-independent strand displacement, which is significantly different from ‘normal’ replication and has an unclear physiological function.

This chapter will describe further functional assays designed to probe the $\varepsilon$-$\beta_3$ interaction. Specifically, experiments designed to test whether both binding clefts in the $\beta_3$ dimer are
required for the interaction to be functional will be described, as well as experiments designed
to probe whether the $\varepsilon$-$\beta_1$ interaction contributes to replication on single-stranded templates
in addition to its requirement during strand displacement replication. Finally, a model for the
$\alpha\varepsilon\theta$-$\beta_1$ complex which incorporates this newly characterised interaction will be described, and
potential mechanisms regulating the interaction will be discussed.
6.2 MATERIALS AND METHODS

6.2.1 Proteins

The methods of production and purification of a truncated form of the $\beta_2$ clamp missing the 5 extreme C-terminal residues, $(\text{His}_6-\beta)^5$, as well as two mixed $\beta_2$ clamp dimers, $\text{His}_6-\beta^5/\beta_{WT}$ and $\text{His}_6-\beta_{WT}/\beta_{WT}$, were described in Chapter 5. Wild-type $\varepsilon$, and the $\varepsilon$ variants $\varepsilon_Q$ and $\varepsilon_L$ are insoluble, and were overproduced and purified by Dr Slobodan Jergic, using a denaturing and refolding protocol as described previously\cite{88}, followed by chromatography on a Super-Q column (Toyopearl). SSB was expressed and purified by Dr Jergic, essentially as described in section 3.2.2. All other Pol III holoenzyme subunits and subassemblies used in the assays described in this chapter were produced and purified by Dr Slobodan Jergic, using methods described previously: $\alpha$\cite{487}, $\theta$\cite{121,513}, wild-type $\beta_2 [(\beta_{WT})_2]$\cite{138}, $\delta$ and $\delta'$\cite{487}, $\gamma$ and $\chi$\cite{192}, $\tau$\cite{78}, and $\psi$ (refolded in the presence of $\chi$)\cite{78}. The $\alpha$ mutants $\alpha\Delta 7$, in which the 7 C-terminal residues are deleted, and $\alpha_L$, in which the internal $\beta$-binding motif (residues 920-924) is modified from QADMF to QLDLF were purified as for wild-type $\alpha$\cite{487}. The wild-type core subassembly, $\alpha\varepsilon\theta$, and those containing the various $\alpha$ or $\varepsilon$ mutants ($\alpha\Delta 7\varepsilon\theta$, $\alpha_7\varepsilon\theta$, $\alpha\varepsilon_Q\theta$ and $\alpha\varepsilon_L\theta$) were prepared from the individually purified subunits as described previously\cite{78}, then isolated chromatographically using a Super-Q column. The clamp loader complexes $\tau\delta\delta'\chi\psi$ and $\gamma_3\delta\delta'$ were also preassembled from their constituent subunits and purified as described previously\cite{78}.

6.2.2 Oligonucleotides and DNA templates

The 5'-tailed primer-template DNA for the replication assays described in this chapter was prepared by Dr Slobodan Jergic. An oligonucleotide primer with the sequence 5’-TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTGACCCCCGGTTGATAATCAG
Figure 6.9. The primed M13 template DNA. A 69 nucleotide primer was annealed to a single-stranded M13 circular template such that a 5' tail (36 x T) remained unhybridised.

AAAAGCCCCA-3' was annealed to a circular, single-stranded wild-type M13 DNA template (6407 nucleotides) such that the 36 nucleotides highlighted in bold remained unhybridised and formed a 5' tail (Figure 6.9). A 30-fold molar excess of the primer was mixed with M13 ssDNA (35 nM) in a buffer of final composition 30 mM Tris-HCl, pH 7.6, 15 mM MgCl₂, 130 mM NaCl and 0.1 mM EDTA. The mixture was incubated at 55°C for 10 minutes, then slowly cooled over 8 hours to allow hybridisation before being stored at -20°C until use.
6.2.3 Buffers

Replication assay buffer was prepared as a 2x concentrated solution containing 50 mM Tris-HCl, pH 7.6, 20 mM MgCl₂, and 260 mM NaCl. Appropriate quantities of this buffer were added to replication assays such that the final concentration of each component was 25 mM Tris-HCl, pH 7.6, 10 mM MgCl₂, and 130 mM NaCl.

Quenching buffer consisted of 200 mM EDTA, 0.08% (w/v) bromophenol blue, 0.08% (w/v) xylene-cyanol, 10% (v/v) glycerol and 2% (w/v) SDS. It was added to replication assay reactions to halt synthesis after a specified time, and to allow loading of reactions onto an agarose gel for analysis.

6.2.4 DNA replication assays

Replication assays were carried out in collaboration with Dr Slobodan Jergic. Assays designed to facilitate primer extension synthesis (single-stranded template filling) as well as subsequent strand displacement rolling circle synthesis were performed in a volume of 13 µL containing 2.5 nM 5’-tailed M13 primed-template DNA, 90 nM preassembled αεθ (or an equivalent quantity of preassembled αεθ or αεθ), 30 nM preassembled τδδχψ clamp loader complex, 200 nM βclamp (or an equivalent quantity of modified βclamp dimer, as indicated), 750 nM SSB, 2 mM dNTPs (Bioline), 1 mM ATP and 10 mM DTT in replication assay buffer (section 6.2.3). For assays designed to probe the effects of the βclamp concentration on strand displacement synthesis (described in section 6.3.2), a range of βclamp concentrations were added as indicated, whilst the concentrations of all other components were unchanged.

Weakened primer extension assays designed to assess single-stranded replication (described in section 6.3.3) contained the same components, with the following modifications: (1) Preassembled αεθ (or αεθ, αΔ7εθ, αεθ or αεθ, as indicated) was present at 150 nM, (2)
preassembled $\gamma_0\delta\delta'$ minimal clamp loader complex (40 nM) was used in place of the full
$\tau_3\delta\delta'\chi\psi$ complex, and (3) SSB was omitted.

For all replication assays, all components except template DNA were mixed, and placed on ice
for 5 minutes. Replication was initiated by addition of template DNA, then reactions were
placed at 30°C to allow replication to proceed, generally for 20 minutes unless otherwise
indicated. After the specified time, reactions were quenched by addition of 11 $\mu$L of quenching
buffer (section 6.2.3) which provided a final EDTA concentration of ~100 mM. For
timecourse assays, reactions were performed in a larger volume (with equivalent
concentrations of all components), and 13 $\mu$L aliquots were sequentially removed and
quenched after the indicated times.

For analysis, reactions were loaded onto 0.8% agarose gels in a running buffer of 180 mM
Tris-borate and 4 mM EDTA (2x TBE buffer)\cite{431}. For each gel, Gene Ruler 1 kb Plus DNA
ladder (Fermentas) was used as a size reference, and a sample of M13 DNA template was also
loaded as a size control. Gels were run at 45 V for 180 minutes, then stained with SYBR Gold
nucleic acid gel stain (Invitrogen) according to the manufacturer’s directions. SYBR Gold is a
fluorescent cyanine dye capable of detecting both single-stranded and double-stranded DNA
with a high level of sensitivity\cite{514}. After staining, gels were illuminated with 302 nm UV light
and photographed using a GelDoc XR+ imaging system (BioRad). The extent of DNA
synthesis for individual reactions was assessed qualitatively, based on visual inspection of the
stained DNA bands.
6.3 RESULTS AND DISCUSSION

6.3.1 Both protein binding clefs in the β₂ dimer are required for ε-dependent strand displacement DNA synthesis

The work described in section 6.1 established that the ε subunit is absolutely required for strand displacement synthesis to occur (under high salt, high [dNTP] conditions). Furthermore, a mutant version of ε (εL) containing an artificially strengthened putative β₂-binding site (QLSLPL) was able to stimulate strand displacement to a greater extent, whilst the εQ mutant, containing a weakened putative site (ATSMAF) was unable to support strand displacement synthesis, suggesting that the requirement for ε was indeed based on a need for an ε-β₂ interaction⁴⁹⁶.

The existence of an ε-β₂ contact during strand displacement suggests that both hydrophobic protein-binding clefs in the β₂ dimer (see section 5.1) would need to be occupied simultaneously during strand displacement – one by the internal clamp binding motif of α; the other by the proposed clamp binding motif (CBM) in ε. In order to investigate whether this was indeed the case, a β₂ clamp containing only one competent binding pocket was required. The production of such a clamp, the heterodimeric His₆βC/βWT, was described in Chapter 5. The βC subunit of this hemi-mutant, in which the final five residues from the C-terminus of β are absent (see section 5.3.3) is known to be impaired in its interactions with α and δ⁴⁷⁹. Consequently, βC was anticipated to exhibit a similarly impaired interaction with ε, since the ε-β₂ interaction is predicted to be mediated via a similar CBM/hydrophobic pocket contact to those involved in the α-β₂ and δ-β₂ interactions.
Therefore, to examine the influence of the second binding pocket upon strand displacement, the His$_6$-β$^C$/β$_{WT}$ heterodimer was added to the coupled primer extension-strand displacement rolling circle assay (section 6.2.4), and the resulting DNA synthesis after 20 minutes compared to that observed with the fully competent (β$_{WT}$)$_2$ clamp, as well as to synthesis in the absence of β$_2$ or in the presence of the double mutant (His$_6$-β$^C$)$_2$ (Figure 6.10a). The bands visible in this gel are labelled in the same manner as those presented in the introduction to this chapter (see Figure 6.4): the original primed template ssDNA and the dsDNA produced by filling of the circle can be clearly distinguished from one another, and additional DNA resulting from strand displacement synthesis is also easily identifiable.

In the absence of the β$_2$ clamp, DNA synthesis around the circle was able to occur (the most intense band corresponds to the filled circular template), but it was evidently quite inefficient, with a significant amount of unfilled single-stranded template remaining after 20 minutes (Figure 6.10a). There were no products larger than the filled circle, indicating that, as expected in the absence of β$_2$, strand displacement synthesis was not possible. Addition of the (His$_6$-β$^C$)$_2$ dimer gave a very similar result (Figure 6.10a). The inefficient DNA synthesis carried out by αεθ in the absence of β$_2$ indicates that this synthesis was likely to be distributive – without a clamp, the αεθ core would continually dissociate from and reassociate with the DNA template, resulting in the observed low efficiency synthesis. The similar low efficiency seen with the (His$_6$-β$^C$)$_2$ double mutant supported previous evidence$^{[479]}$ that the β$^C$ deletion mutation does indeed destroy (or significantly compromise) its hydrophobic protein-binding cleft, rendering it unable to interact with α, δ or ε. As such, it is unable to confer processivity and/or contribute stability to the Pol III core.

When the His$_6$-β$^C$/β$_{WT}$ hemi-mutant was present in the assay, the circle-filling reaction was much more efficient, with no detectable template remaining after 20 minutes (Figure 6.10a).
Figure 6.10. Both protein binding sites in the β₂ dimer are required to support strand displacement synthesis. (a) Products observed after 20 minutes of synthesis in assays containing various β₂ clamp constructs, or lacking β₂. Each assay contained αεθ, τ₃δδ’χψ, SSB, dNTPs, ATP, primed template, and wild-type, mutant or no β₂, as indicated. (b) Assays contained the same components as those in (a), but were quenched progressively after the times indicated. The His₆-βWT/βWT dimer was used as a positive control. This assay confirms that the single protein binding site in the His-βC/βWT dimer, whilst allowing very efficient circle-filling, is not sufficient to allow strand displacement. The βWT/His₆-βWT dimer, constructed in the same way as His-βC/βWT but containing two intact protein binding sites, was able to support extensive strand displacement synthesis. All assays were carried out at 30°C.
However, more importantly, no products larger than the full dsDNA circle were synthesised, indicating that, with only a single competent binding site, the $\beta_2$ dimer is incapable of stimulating strand displacement. This was in contrast to the assay containing wild-type $\beta_2$, which showed evidence of significant strand displacement synthesis, indicated by the presence of numerous bands larger than the filled circle (Figure 6.10a).

These results are entirely consistent with the proposal that it is the $\varepsilon$-$\beta_2$ interaction which is responsible for providing sufficient stability to drive the strand displacement reaction, and that this interaction needs to occur simultaneously with the $\alpha$-$\beta_2$ interaction, thus requiring concurrent occupation of both hydrophobic clefts in the $\beta_2$ clamp. The presence of a single intact protein-binding site in the His$_6$-$\beta^{C}/\beta^{WT}$ hemi-mutant allows interaction of the clamp with the $\alpha$ subunit of $\alpha\varepsilon\theta$, stimulating processivity. However, the lack of strand displacement synthesis observed with this mutant indicates that the replication fork does not possess the high level of stability required for strand displacement, which is explained by a loss of the $\varepsilon$-$\beta_2$ interaction. That the $\beta^{C}$ mutant subunit is able to disrupt this interaction provides evidence that $\varepsilon$ does indeed interact with $\beta_2$ via the same hydrophobic cleft utilised by many of its binding partners, as hypothesised.

To allow the DNA synthesis efficiencies to be distinguished more clearly, the progression of the same assay over time was examined. After initiation of the assay, samples were removed and quenched at a number of time points, as indicated in Figure 6.10b. This time, the His$_6$-$\beta^{WT}/\beta^{WT}$ heterodimer, prepared in the same way as the His$_6$-$\beta^{C}/\beta^{WT}$ hemi-mutant (section 5.3.4), was used as a positive control (in place of $\beta^{WT}$) to verify that the neither the His$_6$ tag nor the method of heterodimer preparation impaired the activity of the dimer in this assay. The time course further emphasizes the increased efficiency with which $\alpha\varepsilon\theta$ is able to fill the
template in the assay containing His$_6$β$^C$/β$_{WT}$ compared to the assays containing (His$_6$−β$^C$)$_2$ or completely lacking β$_2$ (compare Figure 6.10b, top left, with Figure 6.10b, bottom). When His$_6$β$^C$/β$_{WT}$ was present, circle filling was essentially complete after ~1 minute, but despite this, no strand displacement synthesis was able to proceed due to the lack of an ε-β$_2$ contact.

With no β$_2$ clamp, or with the (His$_6$-β$^C$)$_2$ dimer, the rate of synthesis was dramatically slower, with circle filling not complete even after 20 minutes (consistent with the result in Figure 6.10a). In the presence of the control dimer His$_6$β$_{WT}$/β$_{WT}$, extensive strand displacement was observed, comparable with that observed for native (β$_{WT}$)$_2$ (Figure 6.10b, top right). This provided confirmation that the heterodimeric His$_6$β$_{WT}$/β$_{WT}$ preparation retained close to full ability to stimulate strand displacement, and thus that the method used in this work for creation of mixed β$_2$ dimers had no significant effect on activity.

### 6.3.2 Dependence of strand displacement synthesis on β$_2$ concentration

The concentration of all β$_2$ variants used in the assays described above was relatively high (200 nM) to reflect the concentration of β$_2$ thought to be present within the bacterial cell (~250 nM$^{[16]}$). Previous replication assays have generally utilised much lower concentrations of β$_2$ (between 1-30 nM), and it was therefore of interest to observe the effects of a range of β$_2$ concentrations on strand displacement synthesis under our conditions. In the analysis of strand displacement synthesis carried out by Yuan and McHenry$^{[198]}$, concentrations of β$_2$ similar to those required for standard processive replication on single-stranded templates were found to be sufficient to support strand displacement (~10-20 nM$^{[198]}$).

Figure 6.11 shows the results of the coupled primer extension-strand displacement assay containing β$_2$ at a range of concentrations from 0-400 nM. No strand displacement synthesis was observed with up to 50 nM β$_2$, but addition of 100 nM β$_2$ resulted in extensive strand
displacement, which was not stimulated further by higher concentrations of $\beta_2$. Repeating the experiment gave reproducible results. This result is interesting in two ways – that such a high concentration of $\beta_2$ is required to support strand displacement, and that there is a relatively small concentration range over which the transition from no strand displacement to full strand displacement synthesis occurs. The result is in quite striking contrast to the previous finding (under lower salt conditions) that $\beta_2$ concentrations of 10-20 nM were sufficient to support strand displacement on a similar circular flap-primed template$^{[198]}$. It therefore appears that the requirement for a high $\beta_2$ concentration is linked to the physiological conditions used in this work to expose the contribution of the $\varepsilon$-$\beta_2$ interaction. The reasons for such a requirement are unclear, as are the reasons that a twofold increase in $\beta_2$ concentration is able to produce such a dramatic stimulation of strand displacement. Further experiments will be required to clarify these observations.
Figure 6.12 Strand displacement is observed at a very high concentration of His-β/C/β WT. Each assay contained αεθ, τδθχψ, SSB, dNTPs, ATP, primed template, and either wild-type β2 or His-β/C/β WT at the concentrations indicated. At a concentration of 400 nM, His-β/C/β WT appeared to be able to support a small amount of strand displacement synthesis (for comparison, the lane on the far right shows the strand displacement products resulting from 200 nM wild-type β2). Reactions were performed for 20 minutes at 30°C.

Given this unusual concentration dependence, it was also important to examine the behaviour of the hemi-mutant over the same range of concentrations (Figure 6.12). If the conclusion drawn in section 6.3.1 is correct, and strand displacement synthesis is indeed dependent on ε interacting with the second binding site on β2, then strand displacement should not occur, regardless of the His-n-β/C/β WT concentration. Surprisingly, although there was no evidence of strand displacement at hemi-mutant concentrations up to 200 nM (consistent with the results in Figure 6.10), at the highest concentration used, 400 nM, a small amount of strand displacement synthesis was observed (Figure 6.12). One explanation for this is that the disrupted hydrophobic cleft in the His-n-β/C/β WT dimer retains a weakened ability to interact with ε, and thus, at a high enough concentration, allows a small amount of strand displacement synthesis.
synthesis. However, given that even the native $\epsilon$-$\beta_2$ interaction is very weak (see Chapter 7), this seems unlikely. A second explanation is that the synthesis could result from the formation of wild-type $\beta_2$ dimers by subunit exchange. Based on the previous assay (Figure 6.11), a $\beta_2$ concentration between 50 and 100 nM would be required to achieve this degree of strand displacement. Therefore, at a concentration of 400 nM, subunit exchange between less than 25% of the His$_6$-$\beta^C$/$\beta_{WT}$ dimers could provide sufficient ($\beta_{WT}$), for some strand displacement synthesis to occur. The results presented in section 5.3.2.3 showed that $\beta_2$ does undergo some subunit exchange after being subjected to conditions similar to those in these assays (the actual salt concentration in the assays, 130 mM NaCl, is closest to the highest concentration, 150 mM, used in these trials). Additionally, the presence of other components in the assay mixture (for example, the clamp loader) could potentially stimulate subunit exchange. It thus seems reasonable that this could be the basis for the strand displacement observed at high His$_6$-$\beta^C$/$\beta_{WT}$ concentrations.

6.3.3 The $\epsilon$-$\beta_2$ interaction is also required for ‘normal’ DNA replication

The results obtained using the coupled primer extension-strand displacement assay showed definitively that an interaction between $\epsilon$ and the $\beta_2$ clamp is required for strand displacement synthesis to occur. However, the biological significance of the strand displacement reaction is uncertain, and it was thus of interest to examine whether the $\epsilon$-$\beta_2$ interaction also played a role in ‘normal’ primer extension synthesis.

In the presence of the full $\tau_3\delta\delta'\chi\psi$ clamp loader complex as well as SSB, $\epsilon$ is not required for $\alpha$ to be able to perform primer extension synthesis – $\alpha$ alone is able to efficiently fill the circular template whilst being stabilised by its interactions with $\beta_2$ and $\tau$. However, in an assay performed by Dr Slobodan Jergic, it was found that when the minimal clamp loader complex
\(\gamma,\delta,\delta'\) was used (i.e. in the absence of \(\tau\), and in the absence of SSB, \(\alpha\) was no longer able to fill the circle alone, and required the addition of \(\varepsilon\) to regain primer extension activity. This was attributed to an overall weakened stability of \(\alpha\) on the template due to the absence of an \(\alpha-\tau\) interaction, as well as to the presence of inhibitory secondary structures in the DNA template which are unable to be easily removed in the absence of SSB.

The ability of \(\varepsilon\) to restore DNA synthesis activity to the weakened primer extension assay could be a result of the \(\varepsilon-\beta\) contact providing additional stability to rescue synthesis. However, it is also possible that the interaction of \(\varepsilon\) with the N-terminus of \(\alpha\) is able to sufficiently stabilise \(\alpha\) upon the DNA and allow synthesis. To distinguish between these possibilities, the assay was performed using preassembled Pol III cores containing the three variations of \(\varepsilon\) described in section 6.1: \(\varepsilon_Q\) (weakened clamp binding motif ATSMAF), \(\varepsilon_n\) (strengthened clamp binding motif QLSLPL) and \(\varepsilon_{\text{WT}}\) (native clamp binding motif QTSMAF), as well as with \(\alpha\) alone (Figure 6.13). Once again, each assay was sampled at multiple time points to allow differences in synthesis rates to be resolved. As observed previously, \(\alpha\) alone was unable to perform DNA synthesis under these conditions (Figure 6.13, top left), although a small amount of weak synthesis was observed after 40 minutes (as a small smudge above the band corresponding to unfilled template). The faint band that migrates between the 500 and 750 bp markers in this assay is due to excess primer in the M13 primer-template DNA preparation. In assays containing active \(\varepsilon\), this band is not observed, presumably due to its digestion by the exonuclease action of \(\varepsilon\) (Figure 6.13).

In contrast to \(\alpha\) alone, the \(\alpha\varepsilon_{\text{WT}}\theta\) core was able to perform primer extension synthesis and fill the circle, but this synthesis was slow, remaining incomplete after 40 minutes (Figure 6.13, top right). Several intermediate bands can also be seen between the unfilled template and the
Figure 6.13. Variations in the predicted $\beta_2$-binding motif in $\varepsilon$ also affect primer extension DNA replication. Each assay contained $\gamma_3\delta\delta'$, $\beta_2$, dNTPs, ATP, primed template, and $\alpha$ or preassembled $\alpha\varepsilon\theta$ core complexes containing either wild-type or mutant $\varepsilon$, as indicated. Use of the minimal clamp loader complex, $\gamma_3\delta\delta'$, and omission of SSB provides conditions under which strand displacement synthesis does not occur, and replication is limited to filling of the circular template. The positions of the initial primer-template and filled template are indicated. Assays were performed at 30°C for the times indicated.
filled circle, which likely indicate regions of secondary structure in the template that the polymerase overcomes with lower efficiency.

The results obtained with the $\alpha \varepsilon Q \theta$ and $\alpha \varepsilon L \theta$ cores were significant. The core containing $\varepsilon Q$ gave a very similar result to that obtained with $\alpha$ alone – it was essentially unable to support DNA synthesis (Figure 6.13, bottom left). In the presence of $\varepsilon L$, synthesis around the template was observed, and, importantly, this synthesis was more efficient than with wild-type $\varepsilon$, with filled circles observed after 2 minutes, and circle filling essentially complete after 40 minutes (Figure 6.13, bottom right). The enhanced synthesis efficiency stimulated by $\varepsilon L$, which has been modified to facilitate a stronger interaction with $\beta_2$, and the complete lack of synthesis with $\varepsilon Q$, which has been modified such that its clamp binding site is weaker, indicates clearly that it is the $\varepsilon \beta_2$ interaction which is responsible for enabling synthesis to occur in this ‘destabilised’ assay. The contribution of the interaction, masked by stronger interactions in the presence of the full complement of Pol III holoenzyme subunits, has been exposed by the conditions of this assay, in which the lack of an $\alpha - \tau$ contact, the lack of SSB, and the relatively high ionic strength (130 mM NaCl) provide ‘difficult’ conditions for replication. The same ‘difficult’ assay conditions (i.e. $\gamma \delta \delta$ complex, no SSB) were then used to probe the effects of mutating the $\beta$-binding sites in $\alpha$, in the presence of either ($\beta_{WT}$)$_2$ or the His$_6$-$\beta^C$/\$\beta_{WT}$ hemi-mutant. Again, three variations of preassembled cores were used, containing either wild-type $\alpha$, $\alpha \Delta 7$ (from which the C-terminal 7 residues of $\alpha$, which comprise its proposed ‘external’ clamp binding motif[111,112], are absent), or $\alpha_i$ (in which the ‘internal’ clamp binding motif in $\alpha_i^{[113]}$ has been mutated from QADMF to QLDLF to strengthen its interaction with $\beta_2$). The wild-type $\varepsilon$ subunit was used in each case. In the presence of wild-type $\beta_2$, $\alpha \Delta 7 \varepsilon \theta$ was able to synthesise around the circle with a very similar efficiency to $\alpha \varepsilon \theta$ (Figure 6.14). Close examination of the assay containing the $\alpha_i \varepsilon \theta$ core (Figure 6.14) revealed that it stimulated
Figure 6.14. Processive replication is mediated by the internal clamp binding motif in $\alpha$. Each assay contained $\gamma, \delta'$, wild-type $\beta$, dNTPs, ATP, primed template, and preassembled $\alpha\varepsilon\theta$ core complexes containing either wild-type or mutant $\alpha$, as indicated. Removal of the external clamp binding motif in $\alpha$ by deletion of the C-terminal 7 residues ($\alpha\Delta7$) did not affect the efficiency of replication, with $\alpha\varepsilon\theta$ cores containing $\alpha\Delta7$ and wild-type $\alpha$ showing almost identical amounts of synthesis. In contrast, mutation of the internal $\beta_2$ binding site in $\alpha$ to a motif known to interact more strongly with $\beta_2$ ($\alpha_L$) resulted in slightly more extensive synthesis, suggesting that it is the internal and not the external motif in $\alpha$ which interacts with $\beta_2$ during processive replication. Assays were performed at 30°C for the times indicated.

synthesis slightly – a larger proportion of fully replicated products were visible at each individual timepoint when compared to the results with $\alpha\varepsilon\theta$ and $\alpha\Delta7\varepsilon\theta$.

The increased synthesis efficiency achieved by $\alpha_L\varepsilon\theta$ can presumably be attributed to the slightly stronger interaction between $\alpha_L$ and $\beta_2$, whilst the near-identical results obtained with $\alpha\varepsilon\theta$ and $\alpha\Delta7\varepsilon\theta$ indicate that, in the context of this assay (which does not contain $\tau$), the external $\beta_2$-binding site of $\alpha$ (the last 7 residues) does not make any observable contribution to stabilising the polymerase core onto the template. This is consistent with previous work which found that only the internal $\beta$ binding site in $\alpha$ was required for processive replication$^{113}$, and provides further evidence that it is indeed the internal clamp binding motif which occupies $\beta_2$. 

197
during ‘normal’ DNA synthesis. Taken together with the effects mediated by the ε mutants (Figure 6.13), these results suggest that during primer extension, $\beta_2$ is occupied simultaneously by ε and the internal clamp binding motif of α.

Based on these results alone, the function of the ε-$\beta_2$ interaction during this assay remains ambiguous. It may facilitate more efficient synthesis by providing additional stability to the α-$\beta_2$ complex (as is suggested by the results in the strand displacement assays) (Figure 6.15a). Alternatively, the interaction could contribute to efficiency by acting as a ‘shield’, of sorts – in the absence of τ, $\beta_2$ is vulnerable to being unloaded by the $\gamma_3\delta\delta'$ complex[189]. Thus, by occupying the second hydrophobic cleft on $\beta_2$, ε may prevent incursion by the $\delta$ subunit of the $\gamma_3\delta\delta'$ complex (which, under these conditions would exist free in solution), and therefore inhibit unloading of the clamp (Figure 6.15b). If this was the case, the equivalent assay containing the hemi-mutant His$_6$-$\beta^{C}/\beta^{WT}$ would be expected to show similar or even more efficient synthesis: without a second binding cleft, the clamp should not be susceptible to $\delta$-binding and subsequent unloading, and the ε-$\beta_2$ reaction should not be required. On the other hand, if the function of the ε-$\beta_2$ interaction is to stabilise the core onto the template (via $\beta_2$) sufficiently for replication to occur, then His$_6$-$\beta^{C}/\beta^{WT}$ should be unable to support synthesis (similar to α alone in Figure 6.13).
Figure 6.15. Potential roles for the ε-β₂ contact during replication. (a) A stabilising role. The ε-β₂ interaction provides an additional point of contact between the polymerase core and the β₂ clamp (alongside the α-β₂ interaction), which could be expected to result in the polymerase gaining stability on the template. (b) A sequestering role. In the absence of an ε-β₂ interaction, the second binding site in β₂ may remain exposed, allowing binding of other proteins, such as the δ subunit of the clamp loader complex, which could unload the clamp prematurely and cause synthesis to be disrupted.
Figure 6.16. The absence of a second protein binding cleft in β2 results in weakened synthesis. Assays were performed under the same conditions and with the same components as those in Figure 6.14, with the exception that wild-type β2 was replaced by the His-βC/βWT hemi-mutant. Primer extension synthesis was relatively weak in the presence of all α mutants, but was still able to occur. This synthesis is attributed to the formation of active wild-type β2 dimers via subunit exchange.

The results of the His6-βC/βWT assays, which are shown in Figure 6.16, were therefore surprising: synthesis was not abolished, but was significantly weakened (for each α variant) compared to the assays containing wild-type β2. This indicates that the ε-β2 interaction provides a contribution to synthesis efficiency beyond simply preventing clamp unloading, and therefore suggests that its role is a stabilising one. However, it is puzzling that some synthesis is still observed. Again, this may be the result of subunit exchange forming active β2 dimers.

The high concentration of His6-βC/βWT in the assays (200 nM) means that a relatively small amount of subunit exchange could produce a concentration of (βWT)2 sufficient to support primer extension synthesis. Alternatively, the synthesis observed could indicate that the ε-β2 interaction is not absolutely required for replication under these conditions, but that it does enhance replication efficiency by providing a stabilising force.
Figure 6.17. The ε-β₂ interaction plays a stabilising role in replication. Each assay contained γ₂δ₂', dNTPs, ATP, primed template, His-β²/β⁻WT hemi-mutant, and α or preassembled αεθ core complexes containing either wild-type or mutant ε, as indicated (comparable to those in Figure 6.13, which were performed with wild-type β₂). A potential role for the ε-β₂ interaction could be sequestration of the second binding site on the β₂ dimer, thus protecting β₂ from unloading by free clamp loader complex and allowing processive synthesis. However, these assays show that the mutant dimer His-β²'/β⁻WT, which contains only a single protein binding site, showed no evidence of synthesis in the presence of α alone or αεθ, suggesting that the ε-β₂ interaction is required for stability, and does not simply protect against β₂ unloading. The synthesis observed in the presence of αεWTθ is attributed to the formation of active wild-type β₂ dimers via subunit exchange between His-β²'/β⁻WT dimers.

Despite the ambiguity in these results, they do provide evidence that the ε-β₂ interaction provides stability during primer extension replication, and that both binding sites in the β₂ dimer are required to enable the interaction to occur whilst β₂ is also bound to α. To verify further that the interaction of ε with β₂ does not act to block access of the clamp loader complex to β₂, the results of the assays containing α alone or αεθ with the hemi-mutant Hisε-β²'/β⁻WT were examined. It was shown earlier in this section that, in the presence of wild-type β₂, neither α alone nor αεθ were able to perform primer extension and fill the circle (Figure 6.13). If this was due to unloading of the clamp mediated by its unoccupied second binding
site, then the hemi-mutant would be expected to allow synthesis to occur. As shown in Figure 6.17, this was not the case – there was no evidence of synthesis by either $\alpha$ alone or by $\alpha\varepsilon_\theta\theta$.

Since this work was carried out, experiments examining DnaB (helicase)-dependent leading strand synthesis at the single-molecule level\textsuperscript{[78]} have been performed by Dr Slobodan Jergic, Mr Nick Horan and Mr Mohamed Salem\textsuperscript{[496]}. These single-molecule assays measured the rate and processivity of synthesis by individually assembled Pol III holoenzymes containing $\alpha$ alone, or $\alpha\varepsilon\theta$ cores with either $\varepsilon_{\text{wt}}$, $\varepsilon_Q$, $\varepsilon_L$ or the proofreading deficient D12A/E14A $\varepsilon$ mutant. Comparison of the processivities of the various holoenzymes revealed that both wild-type $\varepsilon$ and the D12A/E14A mutant stimulated processivity compared to $\alpha$ alone. The core containing $\varepsilon_L$ produced even higher processivities than those observed with $\varepsilon_{\text{wt}}$, whilst $\varepsilon_Q$ resulted in a more modest stimulation of processivity. These results provide further confirmation that the $\varepsilon$-$\beta_2$ interaction plays a role in stabilising the Pol III holoenzyme on its template DNA during normal replication.

6.3.4 A model for the $\varepsilon$-$\beta_2$ interaction during replication

The results described above demonstrate that, as well as being essential for helicase-independent strand displacement synthesis under a particular set of conditions, the $\varepsilon$-$\beta_2$ interaction is also relevant in a primer extension assay analogous to leading-strand replication. In both strand displacement and primer extension, it appears that both binding sites in the $\beta_2$ dimer need to be occupied simultaneously by $\alpha$ and $\varepsilon$. This raises the question of whether such an arrangement is structurally plausible. Although a complete structure for the $\alpha\varepsilon\theta$-$\beta_2$-DNA complex is not available, structures of many of the components have been solved, allowing construction of a model for how the overall $\alpha\varepsilon\theta$-$\beta_2$-DNA complex may look, as well as how the $\varepsilon$-$\beta_2$ interaction could be modulated to facilitate various events at the replication fork\textsuperscript{[496]}.  

202
Figure 6.18. Models for the $\alpha$-$\beta_2$ complex in the absence (a) and presence (b) of DNA. (a) In the absence of DNA, the $\alpha$ subunit adopts an ‘open’ conformation. In this model, in which $\alpha$ occupies one of the protein binding clefts (yellow spheres) on the $\beta_2$ clamp, its PHP domain (the site of the $\alpha$-$\epsilon$ interaction) is placed far away from the second binding site on $\beta_2$, which would make an $\epsilon$-$\beta_2$ interaction improbable in this conformation. (b) The rotation of the fingers domain of $\alpha$ upon DNA binding brings the PHP domain (and therefore $\epsilon$) into close proximity to the second binding site on $\beta_2$. Thus, conformational changes in $\alpha$ could potentially regulate the $\epsilon$-$\beta_2$ interaction. Models constructed by Dr Aaron Oakley (University of Wollongong), based on the structures of full-length Thermus aquaticus $\alpha$ alone\cite{98} and in complex with DNA\cite{99}, as well as the structures of $\beta_2$ binding peptides in complex with $\beta_2$ \cite{151}.

As described in section 1.1.3.1, the crystal structure of a truncated mutant of E. coli $\alpha$ (residues 1-917) which lacks the C-terminal domain has been solved\cite{97}, as have the structures of the full length $\alpha$ subunit from Thermus aquaticus (which shares 39% sequence identity and a high degree of similarity with E. coli $\alpha$\cite{98}) both in the absence of DNA\cite{98} and bound to primer-template DNA\cite{99}. The combined evidence from these structures reveals that $\alpha$ undergoes a number of significant conformational changes upon binding to DNA. In the largest of these, the section of the fingers domain containing the internal $\beta_2$ binding site (refer to section 1.1.3.1 for an explanation of $\alpha$ domain organisation) rotates inward by approximately 20°,
closing the ‘hand’ around the newly synthesised duplex DNA and allowing the fingers domain as a whole to make extensive contacts with the duplex\(^{97,99}\).

By combining information from these \(\alpha\) structures with the structure of the \(\beta\) clamp bound to DNA\(^{119}\), models of the \(\alpha-\beta\) complex in the absence and presence of DNA can be constructed (Figure 6.18). In these models, the internal clamp binding motif from \(\alpha\) is positioned within the binding pocket on \(\beta\). As well as optimising its own interactions with DNA, the transition in \(\alpha\) from an open (Figure 6.18a) to a relatively closed (Figure 6.18b) conformation also has the effect of bringing the PHP domain of \(\alpha\) (its site of interaction with \(\varepsilon\)\(^{104}\)) into closer proximity to the unoccupied second binding site on the \(\beta\) dimer. Based on these models, the putative \(\alpha\varepsilon\theta-\beta\) complex would likely position \(\varepsilon\) and \(\beta\) too far from each other to form an interaction, but the conformational change in \(\alpha\) upon binding to DNA could facilitate the simultaneous interaction of the clamp binding motifs from both \(\alpha\) and \(\varepsilon\) with the two binding clefts from \(\beta\).

The crystal structure of the C-terminus of \(\varepsilon\) (residues 209-243) in complex with the PHP domain of \(\alpha\) (residues 2-270)\(^{105}\), along with the structure of the N-terminal domain of \(\varepsilon\) in complex with the \(\theta\) subunit\(^{113}\) allowed a more complete model of the full \(E.\ coli\ \alpha\varepsilon\theta-\beta\)-DNA complex to be constructed (Figure 6.19). In this model, which was constructed by Dr Thomas Huber (Research School of Chemistry, Australian National University), the internal \(\beta\) binding site from \(\alpha\) and the presumed clamp binding motif from \(\varepsilon\) are positioned within the two binding pockets from \(\beta\). This allows placement of the structured N-terminal domain of \(\varepsilon\), which immediately precedes the clamp binding motif, and of \(\theta\), which interacts with this structured region of \(\varepsilon\). The model shows one way in which the gap between the region of \(\varepsilon\) bound to \(\beta\) (residues 182-187) and its \(\alpha\)-bound C-terminus (residues 209-243) could be
Figure 6.19. Proposed model for the αεθ-β₂-DNA complex. The same model, which was constructed by Dr Thomas Huber (ANU) is shown here from three perspectives. In the model, α (in tan) interacts with one of the binding sites in β₂ (in blue) via its internal β₂ binding site. The C-terminus of ε (in green) binds to the PHP domain of α, and is connected to the reminder of ε by a long flexible segment (residues 188-208). The proposed β₂-binding peptide in ε (QTSMAF) is shown occupying the second binding site on β₂. The θ subunit and duplex DNA are also shown, in pink and orange respectively. The model was created based on the structure of β₂ in complex with DNA[139], the structure of residues 2-917 of E. coli α[97], the structure of full-length α from Thermus aquaticus[98], the structure of θ in complex with the N-terminal domain of ε[133], and the structure of the C-terminal domain of ε (residues 209-243) in complex with the PHP domain of α (residues 2-270). Adapted from reference 105.
bridged by the intervening residues, (residues 188-208) which are known to be flexible[124]. Of course, the models shown in Figures 6.18 and 6.19 rely on structural information obtained from another species (T. aquaticus), as well as speculation as to the relative positions and orientations of the various subunits, and will require modification as further structural information becomes available. Nonetheless, they provide insight into how ε and α may be able to contact β₂ simultaneously, as well as how conformational changes in α may be able to exert control over the ε-β₂ interaction.

Such control would be important for replication in vivo, during which rearrangements of the replication fork components are frequently required. The model presented here proposes that the ε-β₂ interaction is intact during processive replication, providing additional stabilisation of the Pol III core on template DNA. However, with an intact ε-β₂ interaction, both binding sites within β₂ are occluded, and there are multiple situations which require that other proteins (e.g. translesion or repair polymerases) gain access to one of the binding sites. A conformational switch mediated by α could potentially enable this – a site which causes the polymerase to stall (for example, a lesion which requires an alternative polymerase for bypass synthesis), could induce α to ‘open’, thus disrupting the ε-β₂ contact and allowing other proteins (which may have already established secondary contacts with other parts of the clamp) to bind to the vacated hydrophobic cleft on β₂ and perform their function.

The polymerase also stalls upon completion of an Okazaki fragment on the lagging strand, when it meets the 5’ end of the previous Okazaki fragment. A similar conformational change in α could potentially break the ε-β₂ interaction and allow association of Pol I and DNA ligase with β₂ to seal the gap between the fragments. Finally, ε-mediated proofreading (upon incorporation of a mismatched nucleotide) also requires substantial rearrangement of the replication fork, such that the DNA is transferred from α to the exonuclease active site of ε. It
is tempting to speculate that, upon this transfer, the $\varepsilon\beta_2$ contact may be maintained (at the expense of the $\alpha\beta_2$ interaction) and may play a similar stabilising role during proofreading as it does during replication. Whether this is indeed the case requires further investigation, but the demonstrated importance of the $\varepsilon\beta_1$ interaction to DNA replication is nonetheless an intriguing new discovery in such a well-characterised system.
6.4 CONCLUSIONS

The work described in this chapter builds on recent work carried out by and under the direction of Dr Slobodan Jergic in our laboratory which discovered a requirement for an interaction between the replicative exonuclease, ε, and the sliding processivity clamp, β₂, to enable helicase-independent strand displacement synthesis in an in vitro assay under relatively high ionic strength conditions, similar to those occurring in vivo. Using a hemi-mutant clamp in which only one of the two protein binding clefts in the β₂ dimer was intact, it was shown that both protein binding sites are required for this ε-dependent strand displacement synthesis to occur, further strengthening the evidence that the ε-dependence is indeed the result of a requirement for an ε-β₂ interaction, and suggesting that this interaction needs to occur whilst the α subunit is simultaneously occupying the other protein binding cleft on β₂. Under these conditions, primer extension synthesis around the single-stranded template is still observed in the absence of ε, demonstrating that, whilst the interaction of α with β₂ (and with τ) is sufficient to allow processive replication along an unimpeded template, the ε-β₂ interaction is likely to impart additional stability to the replication fork, enabling the more demanding strand displacement reaction. This reaction was also shown here to be dependent on an unexpectedly high β₂ concentration, the reason for which will require further investigation.

Whilst the unusual requirements for the strand displacement reaction may indeed reflect an interaction geometry utilised by the replication fork for certain situations in vivo, its physiological significance remains unclear. By omitting certain protein components (τ, χ, ψ and SSB) and maintaining a relatively high salt concentration, simple strand extension synthesis was weakened to such an extent that it also became dependent on ε. It was shown in this work that, as in the strand displacement assay, this dependence reflected a need for an ε-β₂ interaction, and that, therefore, the interaction appears to play a stabilising role in ordinary
leading strand DNA replication. This interaction also appeared to benefit from the presence of two intact binding sites in $\beta_2$, although assessing whether both sites were definitively required was hindered by what was suspected to be subunit exchange between hemi-mutant $\beta_2$ clamps, resulting in the formation of fully competent wild-type dimers. Therefore, such heterodimeric $\beta_2$ complexes (which have been described and used previously in the literature) should be utilised with caution and awareness of the potential for subunit exchange, particularly in the presence of elevated salt.

More generally, this work also highlights the extent to which the contribution of a single protein-protein interaction may be obscured in the presence of a large network of other interactions, such as those which exist between the replisomal subunits. The existence of the $\varepsilon-\beta_2$ interaction was only exposed under conditions of moderate salt concentration, which results in a relative weakening of other protein-protein and protein-DNA interactions. Knowledge gained about this new interaction can now help to form a more complete picture of the geometry of an active bacterial replication fork during processive polymerisation. The properties of the $\varepsilon-\beta_2$ contact during dynamic events at the fork such as polymerase switching and proofreading remain to be illuminated by future experiments.
Chapter 7: The $\beta_2$ Clamp Interacts Directly with the Predicted Clamp Binding Motif in $\varepsilon$

Acknowledgement of contributions from others: Work described in this chapter was carried out with and under the supervision of Dr Slobodan Jergic. Dr Jergic was responsible for experimental design and gave extensive assistance with analysis of results. Mass spectrometry work described in section 7.1 was performed by Mr Nick Horan and Dr Thitima Urathamakul.

7.1 Introduction

Functional evidence for the interaction between the replicative exonuclease, $\varepsilon$, and the sliding processivity clamp, $\beta_2$, was presented in the previous chapter. The combined evidence from these functional assays supports the hypothesis that a direct $\varepsilon$-$\beta_2$ interaction contributes to replication by stabilising the polymerase on the template DNA. However, as with any protein-protein interaction, it is desirable to also establish that a direct physical interaction between the proteins can be observed.

Previous work carried out in our laboratory using gel filtration chromatography and mass spectrometry provided evidence of a direct $\varepsilon$-$\beta_2$ interaction, and suggests that this interaction is mediated by the proposed clamp binding motif in $\varepsilon$, residues 182-187 (see Figure 6.7). These studies were based on comparison of the behaviours of the three $\varepsilon$ variants $\varepsilon_L$, $\varepsilon_{WT}$ and $\varepsilon_Q$, which contain strengthened, native and weakened versions of the proposed clamp binding motif, respectively (section 6.1.3). The $\varepsilon_L$ mutant was able to form a complex with $\beta_2$ that could be observed by gel filtration, whilst similar complexes with $\varepsilon_{WT}$ and $\varepsilon_Q$ were not (see Figure 6.7).

Given that the only difference between the mutants resides in the putative hexapeptide binding motif, this was strong evidence that this motif was involved in mediating the $\varepsilon_L$-$\beta_2$ interaction. This, then, also indicated that the hexapeptide region of $\varepsilon$ proposed to form the clamp binding
motif, despite being immediately adjacent to the structured N-terminal domain (Figure 6.7), is sufficiently exposed to allow it to interact with $\beta_2$. Demonstration of a direct $\varepsilon_i-\beta_2$ interaction does not, however, prove the existence of a direct interaction between $\beta_2$ and wild-type $\varepsilon$.

Mass spectrometry of mixtures of $\beta_2$ with excess quantities of each of the $\varepsilon$ variants provided similar results\textsuperscript{[496]}. Both 1:1 and 2:1 $\varepsilon_i-\beta_2$ complexes were detected in substantially greater abundance than those arising from free $\beta_2$ (Figure 7.1d). On the other hand, the ions corresponding to both $\varepsilon_Q-\beta_2$ and $\varepsilon_{WT}-\beta_2$ complexes were too low in abundance to provide convincing evidence for a $\varepsilon_{WT}-\beta_2$ interaction (Figure 7.1 b,c). However, when excess $\beta_2$ was mixed with preassembled $\alpha\varepsilon\theta$ cores containing the three $\varepsilon$ variants and subjected to ESI-MS, the contribution of the native $\varepsilon-\beta_2$ interaction was revealed, with significantly more $\alpha\varepsilon_{WT}\theta-\beta_2$ complex than $\alpha\varepsilon_Q\theta-\beta_2$ complex (relative to their respective free cores) detected (Figure 7.2 a, b). That this effect was related to the proposed clamp binding motif of $\varepsilon$ was emphasized by the even stronger (essentially stoichiometric) formation of an $\alpha\varepsilon_{WT}\theta-\beta_2$ complex (Figure 7.2c).

The fact that the contribution of the wild-type $\varepsilon-\beta_2$ interaction was only able to be observed by ESI-MS in the presence of the full core subassembly suggests that the isolated $\varepsilon-\beta_2$ interaction is very weak, and that the $\alpha-\beta_2$ interaction is able to facilitate it by bringing $\varepsilon$ and $\beta_2$ into close proximity. Whilst these ESI-MS results provided evidence for an $\varepsilon-\beta_2$ interaction mediated by the proposed $\varepsilon$ clamp binding motif, the direct interaction itself remained to be explicitly demonstrated, and its strength and other binding characteristics (association and dissociation rates) remained to be characterised. In the work described in this chapter, surface plasmon resonance was used to characterise the interactions of $\beta_2$ with the clamp binding peptides from $\varepsilon_i$, $\varepsilon_{WT}$ and $\varepsilon_Q$. 
Figure 7.1. NanoESI-mass spectra of ε-β₂ mixtures. (a) β₂ alone (1 μM). (b) β₂ (1 μM) mixed with ε_{WT} (20 μM). (c) β₂ (1 μM) mixed with ε₂ (20 μM). (d) β₂ (1 μM) mixed with ε₃ (20 μM). Ions arising from excess ε are present at a lower m/z than the range shown. Figure adapted from reference 105.
Figure 7.2. NanoESI-mass spectra of αθ-β₂ mixtures. (a) αεₒθ (1.78 µM) mixed with β₂ (2.81 µM). (b) αε₇θ (1.78 µM) mixed with β₂ (2.81 µM). (c) αε₇θ (1.78 µM) mixed with β₂ (2.81 µM). Ions arising from excess β₂ are outside the m/z range shown. Figure adapted from reference 105.
Biosensors which utilise the principle of surface plasmon resonance (SPR) are now well-established instruments for obtaining details of both the kinetics and affinity of macromolecular interactions. Surface plasmon resonance itself is a phenomenon which occurs when photons from a beam of polarised light are directed onto a metallic interface between two media of different refractive indices (Figure 7.3a). At a particular angle of incidence, total internal reflection (TIR) occurs, and the photons excite surface plasmons (essentially oscillations of free electrons) within the metal, resulting in a sharp minimum in the intensity of the light reflected from the surface\footnote{515} (Figure 7.3b). The angle at which this occurs (sometimes referred to as the ‘SPR angle’) varies with the refractive index of the media on either side of the metallic surface\footnote{516}. In a practical sense, SPR instruments designed to measure biomolecular interactions consist of a polarised laser light source, an optical detection system, and a metallic (generally gold) sensor chip, which is sandwiched between a glass prism (through which the light penetrates) and a microfluidic flow cell system (through which molecules in solution are injected)\footnote{517} (Figure 7.3a). Accumulation of molecules within ~200 nm of the surface of the sensor chip (caused by a binding interaction) results in a localised increase in the refractive index of the solution which is approximately proportional to the mass bound to the sensor surface\footnote{516}. SPR instruments detect this increase and convert it to a signal measured in resonance units (RU), where 1 RU corresponds approximately to a surface protein concentration of 1 pg mm\(^{-2}\)\footnote{518}.

For detection of a simple macromolecular interaction, of the form A+B \(\rightarrow\) AB, one binding partner (the ligand, by convention) is first immobilised upon the sensor chip by injecting it over the surface. Several methods for immobilisation exist: sensor chips are usually coated with a matrix of carboxymethylated dextran or other hydrogel, forming a hydrophilic layer on which further functionalization can be achieved\footnote{516,519}. Direct covalent coupling of
Figure 7.3. Principles of surface plasmon resonance applied to the detection of macromolecular interactions. (a) Schematic organisation of an SPR instrument. Polarised light is directed through a prism onto the surface of a gold sensor chip at a specific angle of incidence, inducing surface plasmon resonance which results in a dip in reflected light intensity. Ligand molecules are immobilised on the opposing surface of the sensor chip, which is coupled to a flow cell. Analytes in solution can be injected over the surface, with interactions causing a change in refractive index of the solution near the surface. (b) Changes in refractive index result in a shift in the angle at which the SPR induced ‘dip’ in reflected intensity is observed (I→II). This shift in angle is recorded as a resonance signal, and plotted as a function of time to produce a sensorgram.
biosensor utilizes surface plasmon resonance (SPR) to detect interaction between immobilised ligand and an analyte in solution. The SPR principle relies on the sensing of small changes in refractive index at the surface of a transducer, typically a gold-coated sensor chip. An analyte in solution is injected over an immobilised ligand, resulting in a time-dependent resonance signal change. Changes in interaction are detected as a change in the resonance units (RU), reflecting the binding of analyte to the ligand. A plateau in the signal indicates the interaction has reached equilibrium. When the injection of analyte ceases, its dissociation from the ligand causes a decrease in signal. The signal vs. time plot generated in an SPR experiment is known as a ‘sensorgram’ (Figure 7.4). As a consequence of its ability to measure biomolecular interactions
in real time, SPR is often used to examine interaction kinetics: association and dissociation rate constants can be derived from the association and dissociation phases of the sensorgram, respectively, by fitting kinetic models [516,522-524]. SPR is also useful for examining the affinities of interactions at equilibrium. Affinities can be determined using either the kinetic rate constants or by measuring equilibrium binding responses (obtained by allowing sensorgrams to reach a plateau) over a range of analyte concentrations spanning the equilibrium dissociation constant ($K_D$) of the interaction [525,526]. Using this approach, affinities in the pico- to micromolar range are readily able to be determined [525,526].

A wide variety of SPR instruments are available, and a number of different configurations are utilised by these systems. Generally, the microfluidic component consists of multiple flow cells which allow samples to be injected independently over different areas of the sensor chip. As well as allowing multiple interactions to be examined simultaneously, such an arrangement also allows an unmodified area of the sensor chip to be used as a reference surface, with any signal generated being subtracted from the ‘active’ signals. The work described in this chapter utilised the ProteOn protein interaction array system (BioRad), which is somewhat different in organisation to the BIAcore system (containing four flow cells) that was used for the experiments described in Chapter 4. The ProteOn is a relatively newly available SPR instrument which allows the simultaneous analysis of up to 36 interactions by creating a two-dimensional array on the surface of the sensor chip [527]. Six parallel flow channels allow up to six ligands to be immobilised in one direction (nominally, the ‘vertical’ direction). Rotation of the sensor chip through $90^\circ$ allows injection of up to six different analytes, or concentrations of the same analyte, perpendicularly (in the ‘horizontal’ direction) across the multiple immobilised ligands [527,528] (Figure 7.5). This instrument possesses obvious advantages in terms of throughput and speed — in conventional SPR arrangements each analyte or analyte concentration needs to be injected sequentially, with dissociation or regeneration time
Figure 7.5. The ProteOn XPR36 system allows analysis of multiple interactions simultaneously. (a) Six flow channels are arranged in parallel, allowing immobilisation of up to six ligands in one direction across the sensor chip surface. (b) After rotation of the sensor chip by 90°, the six channels can be used to inject multiple analytes simultaneously across the immobilised ligands, creating an array of 36 interaction 'spots'. Adapted from reference 528.

between injections. The work described in this chapter, which required a large range of analyte concentrations, was significantly expedited by the multiplex arrangement.

SPR has been used previously to examine the interaction between the $\beta_2$ clamp and a number of short peptides corresponding to variations of the pentameric and hexameric clamp binding motifs found in the binding partners of $\beta_2$. In this work, the biotinylated peptides were immobilised on a streptavidin-coated sensor chip, and free $\beta_2$ flowed over the top. This work established that the variations in sequence in these short motifs appear to govern a hierarchy of interaction affinity. In the work described in this chapter, a similar approach was used to establish conclusively whether the putative clamp binding motif in wild-type $\epsilon$, QTSMAF,
indeed interacts with $\beta_2$, and to determine the strength of this interaction. Biotinylated peptides corresponding to the predicted CBM sequence in wild-type $\varepsilon$ (QTSMAF), as well as the sequences which replace it in the $\varepsilon_i$ mutant (QLSLPL) and the $\varepsilon_o$ mutant (ATSMAF) were all examined to confirm that their relative $\beta_2$-binding affinities corresponded with those predicted.
7.2 MATERIALS AND METHODS

All experiments described in this chapter were carried out under the supervision of Dr Slobodan Jergic, who has extensive experience in and knowledge of surface plasmon resonance experimental design and data interpretation.

7.2.1 Peptides and proteins

Custom peptides were synthesised by Mimotopes (Melbourne, Australia) and supplied in lyophilised form. Peptides were designed which contained the proposed $\beta_2$ binding sequence from $\varepsilon$ (residues 182-187; QTSMAF), or the corresponding peptides from the $\varepsilon$ mutants with artificially strengthened ($\varepsilon_L$; QLSLPL) or weakened ($\varepsilon_Q$; ATSMAF) putative $\beta_2$ binding sequences. Each hexameric sequence was incorporated within a decapeptide, such that the sequences of the three peptides were:

$\varepsilon_{WT}$: GGQTSMAFAV

$\varepsilon_L$: GGQLSLPLAV

$\varepsilon_Q$: GGATSMAFAV

The amino acids flanking the hexapeptide sequences were those found in the full-length $\varepsilon$ sequence, with the exception of the C-terminal valine, which was used in place of the native methionine in this position to improve solubility (as suggested by the manufacturer). Each peptide was N-terminally biotinylated to allow immobilisation on a neutravidin-coated sensor chip (see section 7.2.2), and in addition contained a 6-aminohexanoic acid linker between the biotin tag and the N-terminal glycine to act as a spacer between the sensor chip surface and the peptide itself. Peptides were synthesised from the C-terminus, such that any incompletely
synthesised peptides did not contain biotin and were therefore not immobilised on the SPR sensor chip.

The lyophilised peptides were each dissolved in 100 mM Tris-HCl, pH 7.6 containing 25% (v/v) acetonitrile to a concentration of 1 mM (the ε_{WT} and ε_{Q} peptides were insoluble in the absence of acetonitrile). Immediately prior to immobilisation, each peptide was further diluted to 100 nM in SPR buffer (50 mM Tris-HCl, pH 7.6, 50 mM NaCl, 0.5 mM TCEP, 0.2 mM EDTA, 0.005% (v/v) P20 surfactant).

Wild-type β_{2} was overproduced and purified by Dr Slobodan Jergic as described previously[138]. For surface plasmon resonance experiments, purified β_{2} was dialysed against 50 mM Tris, pH 7.6, 80 mM NaCl, 0.5 mM TCEP and 0.2 mM EDTA, then concentrated to 26.2 mg/mL (323 µM, as dimer) using an Amicon centrifugal filter unit (Millipore). This sample was further diluted with SPR buffer (see above) to 100 µM, then serially diluted with the same buffer to produce the concentration ranges indicated for specific SPR analyses. All concentrations refer to concentrations of β_{2} as a dimer.

### 7.2.2 Surface plasmon resonance

All protein interaction experiments described in this chapter were carried out at 20°C using a ProteOn XPR36 Protein Interaction Array System (BioRad) and a ProteOn NLC sensor chip (BioRad), which possesses a coating of tetrameric neutravidin to allow immobilisation of biotinylated molecules via the extremely strong biotin-neutravidin interaction. The sensor chip surface was activated by three sequential injections of 1 M NaCl, 50 mM NaOH through the 6 flow paths in both the horizontal and vertical directions, followed by two sequential injections of 1 M MgCl_{2} in both directions (all injections were for 40 seconds, at 40 µl/min in the vertical direction and 100 µl/min in the horizontal direction). All subsequent experiments
were carried out in SPR buffer (50 mM Tris-HCl, pH 7.6, 50 mM NaCl, 0.5 mM Tris(2-carboxyethyl)phosphine (TCEP), 0.2 mM EDTA, 0.005% (v/v) P20 surfactant) which was filtered and degassed before use.

The three biotinylated ε peptides (εLpep, εWTpep and εQpep, each at 100 nM) were immobilised on the sensor chip by independent injections through separate flow paths with the chip in the vertical orientation. Injection times of 33 seconds, 42 seconds and 40 seconds (at a flow rate of 50 µl/min) were used for εL-pep, εWT-pep and εQ-pep respectively, resulting in immobilisation levels of ~50 (εLpep), ~70 (εWTpep) and ~60 (εQpep) response units (RU).

The sensor chip was then rotated by 90°, allowing β2 to be injected simultaneously across the immobilised peptides (in the ‘horizontal’ orientation). One flow path in this direction was reserved for injection of SPR buffer (to allow reference subtraction), and the remaining 5 flow paths used to inject β2 at a range of concentrations. Three separate concentration ranges were prepared:

1. 0.0060 µM, 0.012 µM, 0.024 µM, 0.049 µM and 0.098 µM β2.

2. 0.195 µM, 0.390 µM, 0.780 µM, 1.56 µM and 3.12 µM β2.

3. 6.25 µM, 12.5 µM, 25.0 µM, 50.0 µM and 100 µM β2.
Figure 7.6. Experimental setup for surface plasmon resonance experiments. The $\varepsilon_1$, $\varepsilon_{\text{WT}}$ and $\varepsilon_9$ peptides were immobilised across three separate vertical channels on an NLC (neutravidin-coated) sensor chip via biotin-neutravidin interactions. Various concentrations of $\beta_2$ were then flowed over all three immobilised peptides simultaneously.

Each concentration series was injected independently (along with a blank) through the 6 horizontal flow cells at 100 $\mu$L/min for 60 seconds, followed by a dissociation period of 300 seconds. Regeneration of the sensor surface between injections was not required, as dissociation of bound $\beta_2$ from the immobilised $\varepsilon$ peptides was rapid and complete. The resulting sensorgrams were doubly reference subtracted, using the responses from both the blank injections in the horizontal direction and from one of the unmodified flow paths in the vertical direction. A schematic of the experimental setup is presented in Figure 7.6.
7.2.3 Determination of equilibrium binding affinities

All $\varepsilon$ peptide-$\beta_2$ interactions demonstrated extremely rapid binding and dissociation kinetics (see results and discussion), and as such the association and dissociation rates of the interactions were not determined quantitatively. Instead, the equilibrium binding responses over a range of $\beta_2$ concentrations were used to determine binding affinities (for the $\varepsilon_{\text{Lpep}}-\beta_2$ and $\varepsilon_{\text{WTpep}}-\beta_2$ interactions). Equilibrium response values were calculated using the ProteOn Manager software version 2.0 (BioRad) following manual selection of a portion of each set of sensorgrams where binding was judged to have reached equilibrium. These equilibrium responses were then plotted against the $\beta_2$ concentration, and a binding curve was fitted using equation 7.1:

$$R = \frac{R_{\text{max}} [\beta_2]}{K_D + [\beta_2]}$$

(Equation 7.1)

where $R$ is the equilibrium response at a particular $\beta_2$ concentration, $[\beta_2]$, $R_{\text{max}}$ is the theoretical maximum response with 100% saturation of immobilised binding sites, and $K_D$ is the dissociation constant of the interaction (in this context, the $\beta_2$ concentration at which half the maximum response is observed). Equation 7.1 is a form of the Langmuir isotherm, which is commonly applied to the analysis of 1:1 interactions by SPR (see section 7.3.1.1).

In the case of the $\varepsilon_{\text{Qpep}}-\beta_2$ interaction, the observed binding was so weak that only two $\beta_2$ concentrations produced reliable responses, and the approach described above could not be used to fit a binding curve and determine the $R_{\text{max}}$ and $K_D$. Instead, $R_{\text{max}}$ for the $\varepsilon_{\text{Qpep}}-\beta_2$ interaction was estimated using the values for $R_{\text{max}}$ determined for the $\varepsilon_{\text{Lpep}}-\beta_2$ and $\varepsilon_{\text{WTpep}}-\beta_2$ interactions, and the relative amount of each peptide immobilised ($R_{\text{max}}$ is dependent on the number of available binding sites, not the affinity of the interaction). This estimated $R_{\text{max}}$, ...
along with the measured equilibrium response at a single $\beta_2$ concentration, was then used to estimate the $K_D$ according to equation 7.2:

$$K_D = \frac{R_{\text{max}}[\beta_2]}{R} - [\beta_2] \quad \text{(Equation 7.2)}$$

which is obtained by rearrangement of equation 7.1. The estimated values for $R_{\text{max}}$ and $K_D$, along with equation 7.1, were then used to generate calculated responses for the $\varepsilon_{Qpep} - \beta_2$ interaction over a range of $\beta_2$ concentrations (0.006 $\mu$M – 1 000 000 $\mu$M).

To visually compare the binding affinities of all three $\varepsilon$ peptide-$\beta_2$ interactions, the experimentally determined responses (for the $\varepsilon_{Lpep}$ and $\varepsilon_{WTpep}$ interactions), and calculated responses (for the $\varepsilon_{Qpep}$ interaction) were normalised by dividing each response by the $R_{\text{max}}$ for the interaction, then plotted against the $\beta_2$ concentration on a log scale. These values could also be fit to equation 7.1, which rearranges to give equation 7.3:

$$\frac{R}{R_{\text{max}}} = \frac{[\beta_2]}{K_D + [\beta_2]} \quad \text{(Equation 7.3)}$$

Prism version 5.0 (GraphPad software) was used for all plots and curve fitting.
7.3 RESULTS AND DISCUSSION

7.3.1 Direct interactions between $\beta_2$ and $\varepsilon$ peptides can be detected by SPR

7.3.1.1 Interaction between the $\varepsilon_L$ peptide and $\beta_2$

The $\varepsilon_L$ hexapeptide, and its full-length $\varepsilon_L$ counterpart, are based on the clamp binding motif from the Hda protein (a homolog of DnaA), which in *E. coli* has the sequence QLSLPL\(^{[484]}\).

Binding of $\beta_2$ to a decapeptide containing this hexapeptide, flanked by the native Hda sequence, was analysed by SPR in a previous study by Wijffels *et al.*\(^{[487]}\). Although the $\varepsilon_L$ decapeptide in the current work differs in its flanking sequence (which is native to $\varepsilon$) and uses a shorter linker between the peptide and its biotin tag, it serves as a useful standard to assess whether the results acquired here are comparable with those obtained previously.

A range of ten $\beta_2$ concentrations between 0.006 $\mu$M and 3.12 $\mu$M (series 1 and 2 in section 7.2.2) were injected over ~50 RU of $\varepsilon_L$pep immobilised on the surface of a neutravidin-coated sensor chip, resulting in a series of sensorgrams (Figure 7.7). The first notable feature of these sensorgrams was the very fast association and dissociation phases – equilibrium was reached very rapidly upon injection of $\beta_2$, and the response returned to baseline only slightly more slowly after injection was ceased. These rapid on- and off-rates were consistent with those previously observed by Wijffels *et al.* for similar interactions\(^{[487]}\).

For interactions which display such fast kinetics, the dissociation constant of the interaction is best determined by analysis of the equilibrium responses obtained using several concentrations of the mobile reactant\(^{[516,526]}\). Such analysis is relatively simple as it avoids the complications inherent in measuring the rates of association and dissociation. If the immobilised ligand and the analyte interact in a simple 1:1 ratio, fitting of the concentration-dependent equilibrium
Figure 7.7. ProteOn sensorgrams showing binding of \( \beta_2 \) to immobilised \( \varepsilon_L \) peptide. Solutions containing \( \beta_2 \) at various concentrations were injected over \( \sim 50 \) RU of immobilised \( \varepsilon_{\text{Lpep}} \) in two sets: 0 \( \mu \)M, 0.0060 \( \mu \)M, 0.012 \( \mu \)M, 0.024 \( \mu \)M, 0.049 \( \mu \)M and 0.098 \( \mu \)M \( \beta_2 \) were injected simultaneously for 60 seconds, then, after 300 seconds of dissociation, 0 \( \mu \)M, 0.195 \( \mu \)M, 0.390 \( \mu \)M, 0.780 \( \mu \)M, 1.56 \( \mu \)M and 3.12 \( \mu \)M \( \beta_2 \) were injected simultaneously for 60 seconds, followed by 300 seconds of dissociation. Full dissociation time not shown as baseline responses were regained rapidly. Sensorgrams are shown with blank and reference surface responses subtracted. All \( \beta_2 \) injections were at 50 \( \mu \)L/min. Measurements were performed at 20°C.

Responses to equation 7.1 (section 7.2.3) yields the \( K_D \) for the interaction, as well as the \( R_{\text{max}} \) (the response at full saturation of the available binding sites)\[522,526\]. The interactions of \( \beta_2 \) with immobilised peptides measured by SPR were previously found to be well-described by a simple 1:1 binding model\[487\], thus, it was anticipated that the \( \varepsilon_{\text{Lpep}} \beta_2 \), \( \varepsilon_{\text{WTpep}} \beta_2 \) and \( \varepsilon_{\text{Qpep}} \beta_2 \) interactions would be able to be similarly analysed.
Figure 7.8. Binding isotherm for the binding of $\beta_2$ to immobilised $\varepsilon_L$ peptide. Equilibrium responses from the sensorgrams in Figure 7.7 were plotted against the $\beta_2$ concentration and fit to a 1:1 Langmuir binding model using Prism 5.0 (GraphPad).

Fitting of equation 7.1 to the equilibrium responses for the $\varepsilon_{\text{lep}} - \beta_2$ interaction resulted in the binding isotherm shown in Figure 7.8, which indeed fits the data well, particularly at $\beta_2$ concentrations less than 1 $\mu$M. The value for $R_{\text{max}}$ determined from this fit, 184 ± 6 RU, is only slightly higher than the measured response at the highest $\beta_2$ concentration used (3.12 $\mu$M; 174 RU), indicating that this concentration is sufficient to almost saturate the immobilised $\varepsilon_{\text{lep}}$ with $\beta_2$. Accurate determination of binding affinity by SPR requires analyte concentrations in a range which includes concentrations both below and above the $K_D^{[526]}$. The $K_D$ determined from this fit was $0.377 \pm 0.041$ $\mu$M, indicating that the range of $\beta_2$ concentrations used for this analysis was appropriate.
A $K_D$ of 0.377 $\mu$M for the $\epsilon_{\text{lep}}\beta_2$ interaction indicates a relatively high affinity interaction, and is in close agreement with the $K_D$ of $\sim0.4$ $\mu$M determined for the Hda peptide-$\beta_2$ interaction$^{[487]}$. Such consistency is encouraging, and indicates that data from the ProteOn instrument are comparable with data from BIAcore instrumentation (from which the previous value was obtained). However, the detection of an interaction between the $\epsilon_L$ peptide and $\beta_2$, whilst helpful for validating the instrument and conditions, does not prove that $\beta_2$ interacts with the predicted clamp binding motif in $\epsilon$, as the $\epsilon_L$ peptide sequence (QLSLPL) differs significantly from the wild-type CBM (QTSMAF), and was already known to be a strong $\beta_2$ binding sequence. Confirmation of this requires detection of an interaction with the $\epsilon_{\text{WT}}$ peptide, which is predicted to be substantially weaker than the $\epsilon_{\text{lep}}\beta_2$ interaction (Nick Dixon, personal communication).

### 7.3.1.2 There is a direct interaction between the $\epsilon_{\text{WT}}$ peptide and $\beta_2$

The $\beta_2$ concentrations used in the previous section to determine the affinity of the $\epsilon_{\text{lep}}\beta_2$ interaction were not sufficient to give a measurable response when injected over the $\sim70$ RU of immobilised $\epsilon_{\text{WT}}$ peptide. However, a series of higher concentrations (6.25-100 $\mu$M; series 3 in section 7.2.2) was able to produce detectable binding responses (Figure 7.9). Whilst requiring much higher $\beta_2$ concentrations to give comparable responses, the sensorgrams for the $\epsilon_{\text{WT}pep}\beta_2$ interaction possessed the same overall shape as those obtained for the $\epsilon_{\text{lep}}\beta_2$ interaction (Figure 7.7), with extremely fast association and dissociation rates on either side of the injection.
Figure 7.9. ProteOn sensorgrams showing binding of $\beta_2$ to immobilised $\varepsilon_{\text{WT}}$ peptide. Solutions containing $\beta_2$ at 0 $\mu$M, 6.25 $\mu$M, 12.5 $\mu$M, 25 $\mu$M, 50 $\mu$M and 100 $\mu$M were injected simultaneously over ~70 RU of immobilised $\varepsilon_{\text{WTpep}}$ for 60 seconds, followed by 300 seconds of dissociation (only partially shown). Sensorgrams are shown with blank and reference responses subtracted. Injections were at 50 $\mu$L/min and measurements were performed at 20°C.

The highest concentration of $\beta_2$ used in this concentration series was 100 $\mu$M, which is towards the upper limit for analyte concentration in SPR (very high protein concentrations can behave inconsistency and begin to cause binding artifacts). In addition, the concentration of the stock solution of $\beta_2$ was ~300 $\mu$M (section 7.2.1), meaning that it was not possible to raise the concentration significantly above 100 $\mu$M. It was clear, however, that the response obtained with this concentration (106 RU) fell considerably short of saturating the available binding sites: more $\varepsilon_{\text{WTpep}}$ (~70 RU) than $\varepsilon_{\text{Lpep}}$ (~50 RU) was immobilised on the surface, and given that the maximum response depends only on the immobilisation level and molecular weights.
Figure 7.10. Binding isotherm for the binding of $\beta_2$ to immobilised $\varepsilon_{\text{WT}}$ peptide. Equilibrium responses from the sensorgrams in Figure 7.7 were plotted against the $\beta_2$ concentration and fit to a 1:1 Langmuir binding model using Prism 5.0 (GraphPad).

of the interacting species, the $R_{\text{max}}$ for the $\varepsilon_{\text{WTpep}}-\beta_2$ interaction would be expected to be higher than that determined for the $\varepsilon_{\text{Lpep}}-\beta_2$ interaction (184 RU). For equilibrium affinity analyses using SPR, it is desirable to use a range of concentrations which is broad enough that the highest concentration produces a response close to saturation. Although this was not achievable here, the responses were nonetheless able to be fitted to the same 1:1 binding model used to analyse the $\varepsilon_{\text{Lpep}}-\beta_2$ interaction.

The binding isotherm, shown in Figure 7.10, appears to be a very good fit to the data, but the limited concentration range is reflected by the fact that the curve is not approaching a plateau (compare Figure 7.8). The value for $R_{\text{max}}$ determined from this fit, 327 ± 56 RU is reasonable, albeit slightly high (based on the relative amounts of $\varepsilon_{\text{WTpep}}$ and $\varepsilon_{\text{Lpep}}$ immobilised). The $K_D$, 210 ± 49 $\mu$M, indicates an extremely weak interaction. The large errors associated with both
values are a consequence of the small number of concentrations and the unavoidably narrow concentration range used: the highest $\beta_2$ concentration used, 100 $\mu$M, is significantly smaller than the $K_D$ of the interaction, which decreases the reliability of the determined values. Nevertheless, these results make it clear that the interaction is very weak, and are sufficient to give a good estimate of the strength of the interaction.

That the $\varepsilon_{WT} \beta_2$ interaction was detected at all is extremely important, as it is the first direct evidence that the native clamp binding motif in $\varepsilon$ does indeed interact with $\beta_2$. The $K_D$ determined from these results indicates that the interaction is $\sim$500 times weaker than the interaction observed with the $\varepsilon_i$ peptide ($K_D = 0.377 \mu$M), which suggests that the interaction between full length wild-type $\varepsilon$ and $\beta_2$ in vivo is likely to be similarly weak (see section 7.3.2), and is consistent with expectations based on gel filtration and mass spectrometry (section 7.1), as well as the results from functional assays (Chapter 6).

7.3.1.3 The interaction between the $\varepsilon_Q$ peptide and $\beta_2$ is extremely weak

The $\varepsilon_Q$ peptide differs from the wild-type $\varepsilon$ peptide in that the glutamine (Q) residue of the proposed clamp binding motif is replaced by alanine, which is predicted to weaken the clamp binding ability of the motif. Given the weakness of the $\varepsilon_{WT} \beta_2$ interaction, it was anticipated that any interaction with the $\varepsilon_Q$ peptide would probably be difficult, if not impossible, to detect. This prediction was supported by results obtained using the $\varepsilon_Q$ mutant in replication assays (Chapter 6), which showed that it is unable to support strand displacement synthesis (section 6.3.1) or primer extension synthesis under challenging conditions (section 6.3.2).

To confirm whether this was the case, the same range of $\beta_2$ concentrations used to probe the $\varepsilon_{WT}$ peptide interaction (6.25-100 $\mu$M) was injected over immobilised $\varepsilon_{Q\text{pep}}$ (~60 RU). The responses were extremely small (Figure 7.11), but close inspection of the sensorgrams showed
that two of the higher $\beta_2$ concentrations, 25 $\mu$M and 50 $\mu$M, produced detectable responses (the sensorgram resulting from injection of 100 $\mu$M $\beta_2$ was omitted as it gave an anomalously low response). The responses from the two lower $\beta_2$ concentrations (6.25 and 12.5 $\mu$M) were indistinguishable from the instrument noise.

Given that only two reliable responses were obtained, it was not possible to determine the affinity of the $\varepsilon_{Qpep}\beta_2$ interaction by fitting a binding model to the responses as was done for the interactions with the $\varepsilon_L$ and $\varepsilon_{WT}$ peptides. Instead, the values for $R_{\text{max}}$ and the $K_D$ were estimated as described in section 7.2.3. To estimate $R_{\text{max}}$, the $R_{\text{max}}$ values determined for the $\varepsilon_{Lpep}\beta_2$ and $\varepsilon_{WTpep}\beta_2$ interactions were adjusted according to the relative amount of $\varepsilon_{Qpep}$ immobilised on the sensor chip surface:

- 50 RU immobilised $\varepsilon_{Lpep} \rightarrow R_{\text{max}} = 184$ RU
- 60 RU immobilised $\varepsilon_{Qpep} \rightarrow R_{\text{max}} = 224$ RU
- 70 RU immobilised $\varepsilon_{WTpep} \rightarrow R_{\text{max}} = 327$ RU
- 60 RU immobilised $\varepsilon_{Qpep} \rightarrow R_{\text{max}} = 280$ RU

Estimated $R_{\text{max}}$ for $\varepsilon_{Qpep}\beta_2$ interaction (average) = 250 RU

This is a valid method to estimate the $R_{\text{max}}$, as the maximum response should not depend on the affinity of the interaction – it is the response that would be observed if all immobilised binding sites were occupied by analyte ($\beta_2$). In this case, therefore, the $R_{\text{max}}$ should only be dependent on the immobilisation levels of $\varepsilon$ peptide (since the molecular weights of all three peptides are very similar, responses were considered to be comparable in terms of how many molecules were immobilised). This estimated $R_{\text{max}}$ was then used to estimate the $K_D$ (see section 7.2.3), using the experimental response resulting from injection of 50 $\mu$M $\beta_2$ (5 RU) (Figure 7.11). This resulted in an estimated $K_D$ of $\sim$2.5 mM, placing the $\varepsilon_{Qpep}\beta_2$ interaction a
Figure 7.11. ProteOn sensorgrams showing binding of $\beta_2$ to immobilised $\varepsilon_Q$ peptide. Solutions containing $\beta_2$ at 0 $\mu$M, 6.25 $\mu$M, 12.5 $\mu$M, 25 $\mu$M, and 50 $\mu$M were injected simultaneously over ~60 RU of immobilised $\varepsilon_Q$pep for 60 seconds, followed by 300 seconds of dissociation (only partially shown). Sensorgrams are shown with blank and reference responses subtracted. Injections were at 50 $\mu$L/min and measurements were performed at 20°C.

Further order of magnitude weaker than the $\varepsilon_{WTpep}-\beta_2$ interaction. Whilst these values represent very approximate estimations, it is evident from comparison of the sensorgrams in Figures 7.9 and 7.11 that the $\varepsilon_Q$pep-$\beta_2$ interaction is significantly weaker than the $\varepsilon_{WTpep}-\beta_2$ interaction, highlighting the effect that a single amino acid substitution in this clamp binding sequence can produce.
7.3.2 A proposed role for the $\varepsilon$-$\beta_2$ interaction in replication

By normalising the responses (dividing by $R_{\text{max}}$) and plotting the $\beta_2$ concentration on a log scale, the binding curves for all three $\varepsilon$ peptide-$\beta_2$ interactions were able to be plotted on the same graph (Figure 7.12). The dashed line for the $\varepsilon_{\text{Qpep}}$-$\beta_2$ interaction indicates that this curve is based on calculated values, rather than fitted to the experimental data. The two experimental data points are overlaid on the curve.

![Binding curves for the $\varepsilon_{\text{Lpep}}$-$\beta_2$, $\varepsilon_{\text{WTpep}}$-$\beta_2$ and $\varepsilon_{\text{Qpep}}$-$\beta_2$ interactions measured by SPR. Diamonds, solid line: $\varepsilon_{\text{Lpep}}$-$\beta_2$ interactions; circles, solid line: $\varepsilon_{\text{WTpep}}$-$\beta_2$ interaction; squares, dashed line: $\varepsilon_{\text{Qpep}}$-$\beta_2$ interaction. The curves for the $\varepsilon_{\text{Lpep}}$-$\beta_2$ and $\varepsilon_{\text{WTpep}}$-$\beta_2$ interactions represent the fit of a 1:1 binding model to the measured equilibrium response values at a number of $\beta_2$ concentrations. The dashed line for the $\varepsilon_{\text{Qpep}}$-$\beta_2$ interaction represents a calculated binding curve based on a single response value. The points shown (squares) represent the two experimental equilibrium responses measured for this interaction.](image-url)

Figure 7.12. Binding curves for the $\varepsilon_{\text{Lpep}}$-$\beta_2$, $\varepsilon_{\text{WTpep}}$-$\beta_2$ and $\varepsilon_{\text{Qpep}}$-$\beta_2$ interactions measured by SPR. Diamonds, solid line: $\varepsilon_{\text{Lpep}}$-$\beta_2$ interactions; circles, solid line: $\varepsilon_{\text{WTpep}}$-$\beta_2$ interaction; squares, dashed line: $\varepsilon_{\text{Qpep}}$-$\beta_2$ interaction. The curves for the $\varepsilon_{\text{Lpep}}$-$\beta_2$ and $\varepsilon_{\text{WTpep}}$-$\beta_2$ interactions represent the fit of a 1:1 binding model to the measured equilibrium response values at a number of $\beta_2$ concentrations. The dashed line for the $\varepsilon_{\text{Qpep}}$-$\beta_2$ interaction represents a calculated binding curve based on a single response value. The points shown (squares) represent the two experimental equilibrium responses measured for this interaction.
Comparison of the three interactions on this scale highlights the large differences in $\beta_2$ affinity resulting from alterations in the peptide sequence, particularly the greatly reduced affinity of the $\epsilon_{WT}$ peptide compared to the $\epsilon_L$ sequence. These differences in interaction strength reflect the predictions about the three peptides made based on previous analyses of similar peptides and their relative strength in binding to $\beta_2$. Importantly, the $\beta_2$ binding affinities of the three $\epsilon$ peptides, $\epsilon_{Lpep} > \epsilon_{WTpep} > \epsilon_{Qpep}$, also reflect the behaviour of their full length $\epsilon$ counterparts in both mass spectrometry studies and replication assays. This consistency provides confirmation that these observed behaviours are directly related to the strength of the interaction between $\beta_2$ and the hexapeptide sequence comprising residues 182-187 of $\epsilon$.

The most important result here was that a direct interaction was demonstrated between the $\epsilon_{WT}$ peptide and $\beta_2$, as this provides the first direct evidence that the native sequence does indeed serve as a clamp binding motif in $\epsilon$. Thus, $\epsilon$ becomes a new member of the large group of proteins using such sequences to interact with a common region on the $\beta_2$ clamp. The very weak nature of the $\epsilon_{WTpep}-\beta_2$ interaction presumably translates to a similarly weak interaction with the full-length wild-type $\epsilon$ protein. It is of course possible that other parts of $\epsilon$ could also contribute to the $\epsilon-\beta_2$ interaction, but the clamp binding motif itself likely contributes the majority of the binding affinity.

A weak $\epsilon-\beta_2$ interaction makes sense in several ways. Of course, the lack of its discovery until now provided the first clue that the interaction was likely to be weak. A weak $\epsilon-\beta_2$ interaction is also consistent with its proposed role within the advancing polymerase at the replication fork. As discussed in Chapter 6, the $\epsilon-\beta_2$ interaction, if used to assist with maintaining the stability of the polymerase on the DNA template during replication, would need to be disrupted for certain events to be able to occur – for example, to allow binding of translesion
or repair polymerases to β₂ at sites of DNA damage. The low affinity of the interaction, along with its very fast association and dissociation kinetics, would allow it to be readily disrupted when required. It is also important to note the relative strength of the α-β₂ interaction – its $K_D$ of $\sim 0.8 \mu M^{[95,113]}$ (which becomes significantly stronger ($K_D < 5 \text{nM}$) when β₂ is bound to DNA$^{[177]}$), makes it many times stronger than the ε-β₂ interaction. Although a $K_D$ of $\sim 200 \mu M$ is weak enough to be considered functionally irrelevant, it is important to note that in the context of the assembled replication fork in vivo, which provides a scaffold in which ε and β₂ are positioned close to each other, the effective $K_D$ for their interaction is likely to be many times lower than that of the isolated interaction. This effect was seen in the mass spectrometry results described in section 7.1, in which the ε-β₂ interaction was only able to be observed in the presence of α$^{[196]}$. Thus, despite forming a weak isolated contact, the other interactions present at the replication fork allow the ε-β₂ contact to make its own contribution to the stability of the replication fork.
7.4 CONCLUSIONS

Binding analyses using surface plasmon resonance showed, for the first time, that the $\beta_2$ sliding clamp directly interacts with the predicted clamp binding motif in $\epsilon$, QTSMAF. The interaction between $\beta_2$ and the wild-type clamp binding sequence was found to have a $K_D$ of $\sim 200 \mu$M, making it about 500-fold weaker than the equivalent interaction with the known strong $\beta_2$-binding sequence QLSLPL (from the protein Hda), which had a measured $K_D$ of 0.377 $\mu$M. Nonetheless, these results demonstrated that the interaction with the wild-type peptide, though weak, was real and measurable, thus confirming the predictions about the $\epsilon$-$\beta_2$ interaction made on the basis of results from various other techniques.

The low affinity of the $\epsilon$-$\beta_2$ interaction also makes it, as expected, much weaker than the $\alpha$-$\beta_2$ interaction ($K_D \sim 0.8 \mu$M). Although the interaction is probably significantly strengthened by the network of surrounding interactions at the replication fork, particularly the $\alpha$-$\beta_2$ interaction itself, it is nonetheless most likely to be relatively transient, allowing other interactions to take over the binding site on $\beta_2$ when necessary.

This discovery of another protein which uses a short peptide binding motif to interact with the $\beta_2$ sliding clamp adds even further complexity to the questions surrounding the hierarchy and regulation of access of proteins to the clamp during various activities at the replication fork. Further work will be required to gain a full understanding of the timing and management of these transient, but undoubtedly important interactions.
Chapter 8: CONCLUDING REMARKS

The findings described in this thesis contributed to an understanding of macromolecular complexes of the *E. coli* replication machinery: in the first part of the thesis ( Chapters 3-4) mass spectrometry was used to examine the SSB tetramer for the first time, allowing its subunit exchange behaviour to be characterised and revealing that the C-terminus of SSB has a stabilising effect on the tetramer. In the second part of the thesis ( Chapters 5-7), the newly discovered interaction between the β₂, sliding clamp and ε exonuclease was shown to be a weak but direct interaction, and the implications of the interaction for replisome stability were demonstrated.

The C-terminus of SSB has been well characterised as the site through which SSB interacts with many other proteins, but has not previously been considered to play a role in tetramer assembly. The inherent flexibility of the C-terminus of SSB is reflected by the fact that it is not observed in the crystal structure of full-length SSB\[^{254}\]. Recent observations have suggested that the C-terminal tip of SSB may, in the absence of DNA, fold back and interact with the DNA-binding OB-fold region\[^{196,252}\]. By using mass spectrometry to characterise the subunit exchange rate of SSB and of the C-terminally truncated SSBΔC8, this work showed that removal of the last 8 residues has a destabilising effect on the SSB tetramer at low salt concentrations, observed as a more rapid subunit exchange rate. This adds support to the proposed C-terminus-OB-fold interaction, but, furthermore, suggests for the first time that this interaction may occur in an inter-subunit manner, as has been suggested for the dimeric T7 SSB, gp 2.5\[^{242,443}\]. The role of sequestration of the C-terminus by the OB-fold of SSB remains speculative. If SSB does indeed utilise the (SSB)\[^{35}\] mode during DNA replication, as has been suggested\[^{265}\], two OB-folds in the tetramer would remain unoccupied by ssDNA. Interactions of the C-terminus with these unoccupied subunits could potentially modulate the
SSB-ssDNA interaction as replication occurs. The C-terminus-OB-fold interaction could also represent a mechanism by which the interactions between SSB and other proteins can be regulated. For these reasons, gaining knowledge of the behaviour of the SSB C-terminus is critical. Whilst examining the subunit exchange rates of protein mutants is a powerful method for probing the effects of mutations on oligomeric stability, it does not provide structural information. Given the likely transient nature of the C-terminus-OB-fold interaction, structural information may prove difficult to obtain. Collaborators are currently working on NMR studies of SSB to examine the binding of the tail to the OB-fold; whilst this will only provide single subunit information, modelling may allow extrapolation of configurations for inter-subunit interactions. It will be interesting to find out whether it is only the tip of the C-terminus which makes contact with the core of SSB, or whether more extensive contacts involving the flexible ‘spacer’ region of the protein are involved.

The work described in this thesis also represents the first application of mass spectrometry to the complex SSB-ssDNA interaction. Most previous studies of the SSB-ssDNA interaction have utilised methods such as fluorescence quenching and FRET, and have revealed a huge amount of information about the interaction. One area which remains poorly understood is how the SSB tetramer is able to attain mobility on ssDNA, as it is required by the nature of the processes it mediates to be very dynamic. SSB has been shown to be able to slide along ssDNA\cite{324,326} and also to be able to transfer directly between ssDNA strands via a ternary intermediate\cite{325}. Here, mass spectrometry was applied to the direct transfer mechanism to determine whether the expected products of transfer could be observed in the gas phase. The complexes observed by MS correlated very well with expectations based on the observation of transfer in solution\cite{325}. This suggests that SSB binding modes are faithfully transmitted into the gas phase and provides a ‘proof-of-concept’ which opens the way for future studies probing the SSB-DNA interaction by mass spectrometry.
The recent identification of the $\varepsilon\beta_2$ interaction as a novel contact in the *E. coli* replisome was based on the observed requirement for $\varepsilon$ in replication assays designed for strand displacement rolling circle DNA synthesis, and identification of a potential clamp binding motif in the $\varepsilon$ sequence\[496]. It was hypothesised that these assays would require simultaneous occupation of both of the hydrophobic binding clefts in the $\beta_2$ dimer. Testing this required the construction of hybrid $\beta_2$ dimers in which the protein-binding pocket from one subunit was compromised. This was achieved by truncating the C-terminal 5 amino acids from the $\beta$ sequence\[479], then purifying the hybrid dimers created by subunit exchange between the truncated mutant and wild-type $\beta_2$ dimers (Chapter 5). ESI-MS was used to investigate subunit exchange in the $\beta_2$ dimer for this purpose, once again highlighting its utility for efficiently identifying the subunit composition of proteins.

The mutant $\beta_2$ heterodimer containing a single intact binding site was unable to support strand displacement DNA synthesis, demonstrating that the $\varepsilon\beta_2$ and $\alpha\beta_2$ interactions needed to occur concurrently, via opposite sides of the $\beta_2$ dimer. However, given the uncertain physiological relevance of the strand displacement reaction, further work was undertaken to assess whether the $\varepsilon\beta_2$ interaction played a role in ‘normal’ DNA replication. By using challenging conditions including relatively high salt concentration and minimal protein components, a weakened replication assay was developed in which simple primer extension synthesis on a single-stranded template became dependent on $\varepsilon$. It was shown that this synthesis was also dependent on the presence of both intact $\beta_2$ binding clefts, and that the likely role for the $\varepsilon\beta_2$ interaction in this reaction was as a stabilising force at the replication fork. Surface plasmon resonance was used to demonstrate a direct interaction between $\beta_2$ and a peptide corresponding to the proposed clamp binding site in $\varepsilon$ (QLSLPL). The interaction is very weak ($K_D \sim 200 \mu\text{M}$), suggesting that the network of other interactions surrounding it in
the replisome are required for it to gain its functional significance. The existence of a direct \( \varepsilon - \beta_2 \) contact also adds another layer of complexity to the question of how the \( \beta_2 \) clamp manages and coordinates its multiple interactions with the replicative polymerase, translesion polymerases, and, now, the replicative exonuclease.

An understanding of the \( \varepsilon - \beta_2 \) interaction is important for obtaining a more complete picture of the dynamic interactions within the replisome. Discovery of the interaction represents an additional avenue through which replisomal interactions may be strengthened and thereby increases the chances of crystallising and determining the structure of an intact DNA Pol III holoenzyme, which has so far proved elusive due to the transient nature of many of the interactions. Future work using a variety of methods will allow the role of the \( \varepsilon - \beta_2 \) interaction to be probed in more depth. Evidence from single-molecule experiments already carried out suggests a role for the interaction in leading strand synthesis\[^{496}\]. In vivo experiments examining the cellular effects of \( \varepsilon \) mutants with altered \( \beta_2 \)-binding sequences will also provide instructive complementary information.

Both SSB and the \( \beta_2 \) clamp are central players in DNA replication and repair, and in the case of SSB, recombination. Work described in this thesis has contributed to a deeper understanding of the SSB C-terminus by showing that the C-terminal eight residues provide additional stability to the SSB tetramer, suggesting that the C-terminus is able to participate in inter-subunit interactions. This thesis has also provided direct physical evidence for the \( \varepsilon - \beta_2 \) interaction, and showed that, despite its weakness, the interaction plays a role in replisome stability, highlighting the importance of multiple transient and weak interactions to the replisome.
REFERENCES


100. DeLano, W.L. The PyMOL molecular graphics system. Schrodinger LLC


Escherichia coli DNA polymerase III $\beta$ sliding clamp on DNA. Bona fide primer/templates preferentially trigger the $\gamma$ complex to hydrolyze ATP and load the clamp. *J. Biol. Chem.* **278**, 10033-10040


increasing the affinity of τ and γ for δδ′ to a physiologically relevant range. J. Biol. Chem. 270, 29570-29577


203. Li, X., and Marians, K.J. (2000) Two distinct triggers for cycling of the lagging strand polymerase at the replication fork. J. Biol. Chem. 275, 34757-34765


Waksman, G. (2004) The C-terminal domain of full-length E. coli SSB is disordered
even when bound to DNA. *Protein Sci.* 13, 1942-1947
*Escherichia coli* ssb-1 mutant single strand binding protein. *J. Biol. Chem.* 266,
1616-1626
mutant protein bind single-stranded DNA. *J. Mol. Biol.* 217, 63-74
257. Lohman, T.M., and Overman, L.B. (1985) Two binding modes in *Escherichia coli*
single strand binding protein-single stranded DNA complexes. Modulation by NaCl
concentration. *J. Biol. Chem.* 260, 3594-3603
protein forms multiple, distinct complexes with single-stranded DNA. *Biochemistry*
25, 7799-7802
63, 527-570
13, 250-255
DNA binding co-operativity of *Escherichia coli* single strand binding protein. *J. Mol.
Biol.* 187, 603-615
of *Escherichia coli* single strand binding protein-single-stranded DNA complexes.
Cation, anion, pH, and binding density effects. *J. Biol. Chem.* 263, 4629-4640
*Escherichia coli* single-strand binding protein to single-stranded nucleic acids in the
(SSB)₆₅ binding mode. Cation and anion effects and polynucleotide specificity.
*Biochemistry* 27, 456-471
rearrangements between DNA binding modes of *E. coli* SSB protein. *J. Mol. Biol.*
369, 1244-1257
*Escherichia coli* SSB tetramers to single-stranded DNA in the (SSB)₃₅ binding mode.
J. Mol. Biol. 236, 106-123
266. Bujalowski, W., and Lohman, T.M. (1989) Negative co-operativity in *Escherichia coli*
single strand binding protein-oligonucleotide interactions: I. Evidence and a
quantitative model. *J. Mol. Biol.* 207, 249-268
acid interactions: a thermodynamic model for the interactions of *Escherichia coli* single
strand binding protein with single-stranded nucleic acids in the “beaded”, (SSB)₆₅
mode. *J. Mol. Biol.* 195, 897-907
single strand binding protein-oligonucleotide interactions: II. Salt, temperature and
oligonucleotide length effects. *J. Mol. Biol.* 207, 269-288
269. Lu, D., and Keck, J.L. (2008) Structural basis of *Escherichia coli* single-stranded DNA-
binding protein stimulation of exonuclease I. *Proc. Natl. Acad. Sci. USA* 105,
9169-9174


259


264


265


bacteriophage T4-coded gene 32 (helix destabilizing) protein to nucleic acid lattices. *Biophys. J.* **32**, 403-418


Park, M.S., and O'Donnell, M. (2009) The clamp loader assembles the β clamp onto either a 3′ or 5′ primer terminus. J. Biol. Chem. 284, 31473-31483


dextran sensor surfaces by analysis of the specific activity of monoclonal antibodies. *J. Mol. Recognit.* **8**, 125-131


