Engineering streptavidin and its target ligands with both infinite binding affinity and reversible binding capability

Fogen, Dawson

http://hdl.handle.net/11023/2088
master thesis

University of Calgary graduate students retain copyright ownership and moral rights for their thesis. You may use this material in any way that is permitted by the Copyright Act or through licensing that has been assigned to the document. For uses that are not allowable under copyright legislation or licensing, you are required to seek permission.

Downloaded from PRISM: https://prism.ucalgary.ca
Engineering streptavidin and its target ligands with both infinite binding affinity and reversible binding capability

by

Dawson Herman Fogen

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTERS OF SCIENCE

DEPARTMENT OF BIOLOGICAL SCIENCES
CALGARY, ALBERTA
JANUARY 2015

© Dawson Herman Fogen 2015
Abstract

For development of reusable biosensor chips, bioreactors, protein arrays, and matrices for affinity purification of proteins, it would be ideal for streptavidin to have extremely tight binding to its target ligands (biotin and its binding tags) and at the same time to retain the feature of reversible binding capability. To achieve this objective, a streptavidin mutein (SAVSBPM32) was engineered based on a previously engineered streptavidin variant (SAVSBPM18) that can bind both biotin and its binding tag in a reversible manner. Cysteine residues were placed in strategic positions in both the streptavidin mutein and its binding tags. Disulfide bond formation allows immobilization of tagged proteins to streptavidin. Incubation with biotin in the presence of reducing agents allows stripping off tagged proteins from streptavidin.
Acknowledgments

I would like to thank my thesis committee, Dr. Wong, Dr. Ng, and Dr. Gedamu for their supervision and support during the project. Additionally, I would like to thank Dr. Joenel Alcantara for his role in the thesis examination committee. I am grateful for the guidance and support given by Dr. Wong and Dr. Wu. I am further indebted to Dr. Wu for providing me with cells containing the proposed constructs. I am thankful to Dr. Ng for use of the crystallized structure of the streptavidin-SBP tag complex, support with protein modeling programs, size-exclusion chromatography and binding kinetics experiments. Finally, I would like to thank my parents Dale and Karen, my brother Dillon, and my wife Paige for their understanding, love, and support.

This research is supported by Discovery Grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) to Dr. Wong and Graduate Teaching Assistantships from the University of Calgary.
To my wife Paige,

Thank you
Table of Contents

Abstract ...................................................................................................................... ii
Acknowledgments ...................................................................................................... iii
Dedication .................................................................................................................. iv
Table of Contents ...................................................................................................... v
List of Tables ............................................................................................................. xii
List of Figures .......................................................................................................... xiv
Abbreviations .......................................................................................................... xvii

Chapter one: Introduction .......................................................................................... 1

1.1 Overview ............................................................................................................. 1

1.2 The structure of streptavidin ............................................................................... 2

1.3 Streptavidin homologs and their function ......................................................... 4

1.4 Host cells used for the production of streptavidin ............................................. 6

1.5 Binding of biotin to Streptavidin ....................................................................... 7

1.6 Methods for biotinylation of molecules ............................................................. 11

1.6.1 Challenges in the release of biotinylated biomolecules .................................... 13

1.7 Streptavidin binding peptides (SBP) and the binding of the SBP tag to
streptavidin ................................................................................................................ 14

1.7.1 Binding interactions of the SBP tag central α-helix to streptavidin ............... 17

1.7.2 Binding interactions of the SBP tag N-Terminal GHVV peptide to
streptavidin ............................................................................................................... 18

1.7.3 Binding interactions of the SBP tag C-Terminal HPQGQ peptide to
streptavidin ............................................................................................................... 18
1.7.4 Binding interaction of the SBP tag dipeptide linkers to streptavidin ................. 19
1.7.5 The major drawback of the streptavidin – SBP tag system .................................. 19

1.8 Current trends in engineering: Reusable streptavidin technologies .................... 20
1.8.1 The use of biotin analogs for reusable streptavidin systems .............................. 20
1.8.2 Engineering monomeric and tetrameric streptavidin variants for reusable systems .......................................................................................................................... 22

1.9 Current trends in engineering: Streptavidin variants with higher binding affinity .......................................................................................................................................................................................... 25
1.9.1 Engineering higher binding affinity with non-covalent interactions .................. 26
1.9.2 Engineering infinite binding affinity with intermolecular covalent bond formation .......................................................................................................................................................................................... 27

1.10 The need to develop a tight binding system with reversible binding capability .......................................................................................................................................................................................... 29

1.11 Research objectives ............................................................................................................. 31

Chapter 2: Materials and Methods ....................................................................................... 35

2.1 Modeling of Streptavidin peptide complexes ................................................................. 35
2.1.1 Modeling SAVSBPM18-SBP tag complex ................................................................. 35
2.1.2 Modeling SAVSBPM32-SBPA18C tag complex ........................................................ 37
2.1.3 Modeling SAVSBPM19-CPFB(-2) tag complex ......................................................... 37
2.1.4 Modeling SAVSBPM32-CPFB(-2) tag complex ........................................................ 39
2.1.5 Modeling SAVSBPM96-SBPA18C tag complex ......................................................... 41
2.1.6 Modeling SAVSBPM96-CPFB(-2) tag complex ......................................................... 41
2.1.7 Modeling SAVSBPM80- biotin complex .................................................................. 41

2.2 General procedures .................................................................................................................. 42
2.2.1 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS PAGE) .... 42
2.3 Bacterial strains .............................................................................................................44
  2.3.1 Escherichia coli .............................................................................................................44
  2.3.2 Bacillus subtilis .............................................................................................................44

2.4 Plasmid vectors ..............................................................................................................44

2.5 Vector construction and transformation .........................................................................49
  2.5.1 SAVSBPM18 .....................................................................................................................49
  2.5.2 SAVSBPM19 .....................................................................................................................49
  2.5.3 SAVSBPM32 .....................................................................................................................49
  2.5.4 SAVSBPM32F ....................................................................................................................49
  2.5.5 SAVSBPM96 .....................................................................................................................50
  2.5.6 SAVSBPM80 .....................................................................................................................50
  2.5.7 BLA-L-SBP ......................................................................................................................50
  2.5.8 BLA-L-SBPA18C ..............................................................................................................50
  2.5.9 MBP-L-SBPA18CG19A ....................................................................................................51
  2.5.10 BLA-L-CPFB(-2) ..........................................................................................................51

2.6 Cell preparation for long-term storage of bacteria ..........................................................51
  2.6.1 Bacillus subtilis ..............................................................................................................51
  2.6.2 Escherichia coli ..............................................................................................................52

2.7 Cell culture conditions for protein production ...............................................................52
  2.7.1 Bacillus subtilis ..............................................................................................................52
  2.7.2 Escherichia coli ..............................................................................................................53

2.8 Concentration of proteins ..............................................................................................53
  2.8.1 Bacillus subtilis ..............................................................................................................53
  2.8.2 Escherichia coli ..............................................................................................................54

2.9 Protein purification ..........................................................................................................54
2.9.1  SAVSBPM18 purification by affinity chromatography ........................................... 55
2.9.2  SAVSBPM32F purification by affinity chromatography ........................................... 56
2.9.3  BLA-L-SBP and BLA-L-SBPA18C purification by Ion exchange chromatography ................................................................. 56
2.9.4  MBP-L-SBPA18CG19A purification by affinity chromatography ............................. 57
2.9.5  BLA-L-CPFB(-2) purification ............................................................................. 58

2.10  Quantification of protein ..................................................................................... 61
  2.10.1  Semi-quantitative measurements of crude proteins ............................................. 61
  2.10.2  Quantitative measurements of purified proteins ................................................. 61

2.11  Protein binding characterization .......................................................................... 62
  2.11.1  Disulfide bond formation .................................................................................. 62
  2.11.2  Analysis of protein complex formation ............................................................... 63
  2.11.3  Complex composition ....................................................................................... 65
  2.11.4  Binding kinetics ................................................................................................ 67

2.12  Applications: Affinity and affinity thiol coupling chromatography ................. 69
  2.12.1  Preparation of affinity chromatography columns ............................................... 69
  2.12.2  Coupling efficiency of variants to Affi-gel 15 media ........................................... 70
  2.12.3  Purification of BLA-L-SBP by SAVSBPM18 affinity column chromatography. 71
  2.12.4  Purification of BLA-L-SBP by SAVSBPM32 affinity column chromatography. 72
  2.12.5  Purification of BLA-L-SBPA18C by SAVSBPM18 affinity column chromatography ......................................................................................................................... 73
  2.12.6  Purification of BLA-L-SBPA18C by SAVSBPM32 affinity thiol coupling column chromatography ............................................................................................................. 73
  2.12.7  Purification of BLA-L-SBPA18C by SAVSBPM32 affinity thiol coupling column chromatography under various wash conditions .................................................. 74
Chapter 3: A cysteine containing streptavidin and SBP tag variant:

SAVSBPM32 and SBPA18C................................................................. 77

3.1 Rational design of the streptavidin SAVSBPM32-SBPA18C system .............. 77
  3.1.1 Cysteine containing streptavidin variant SAVSBPM32 and the cysteine
  containing SBP tag variant SBPA18C.................................................. 77
  3.1.2 Design of negative controls for cysteine containing streptavidin and
  ligand tag...................................................................................... 84

3.2 Protein production and purification............................................................... 85

3.3 Intermolecular disulfide bond formation...................................................... 90

3.4 Analyses of protein complex formation......................................................... 94
  3.4.1 Kinetics of disulfide bond formation.................................................. 94
  3.4.2 Binding of the cysteine containing SBP tag to SAVSBPM32F................... 97
  3.4.3 Compositions of different streptavidin (SAVSBPM32F)-SBP tag (BLA-L-
  SBPA18C) complexes analyzed by a modified two dimensional polyacrylamide gel
  electrophoresis................................................................................. 103
  3.4.4 Another composition analysis of different streptavidin (SAVSBPM32F)-SBP tag
  (BLA-L-SBPA18C) complexes using the zinc imidazole reverse staining approach ... 106

3.5 Analysis of the SAV-SBP tag complexes by native polyacrylamide
  electrophoresis.................................................................................. 109

3.6 Applications of SAVSBPM32 to purify SBP and SBPA18C tagged BLA by affinity
  and affinity thiol coupling chromatography........................................... 111

3.7 Binding kinetics of SAVSBPM18 to BLA-L-SBPA18C and SAVSBPM32 to
  BLA-L-SBP....................................................................................... 118
3.8 Purification of BLA-L-SBPA18C using a SAVSBPM32 affi-gel column under different wash conditions ................................................................. 124

3.8.1 Purification with the inclusion of a nonionic detergent in the wash conditions ................................................................................................................................. 124

3.8.2 Purification under high ionic strength wash conditions ............................................ 125

3.8.3 Purification with the inclusion of biotin in the washing buffer .................................. 127

3.8.4 Purification using a high salt and biotin containing washing buffer ......................... 128

3.9 Potential improvements of the SAVSBPM32 and SBPA18C system ...................... 128

Chapter 4: A cysteine containing streptavidin and a PFB tag variant:

SAVSBPM32 and CPFB(-2) .................................................................................... 135

4.1 Rational design of the streptavidin CPFB tag system .............................................. 135

4.1.1 Cysteine containing streptavidin variant SAVSBPM19 and the cysteine containing PFB variant CPFB(-2) ........................................................................................................... 135

4.1.2 SAVSBPM32 and BLA-L-CPFB(-2) ........................................................................ 138

4.1.3 Negative controls for cysteine containing streptavidin and ligand ...................... 138

4.2 Protein production and purification ......................................................................... 138

4.3 Intermolecular disulfide bond formation of BLA-L-CPFB(-2) to the streptavidin variants .............................................................................................................................. 143

4.4 Tetrameric state of SAVSBPM18, SAVSBPM19 and SAVSBPM32 ...................... 145

4.5 Binding kinetics of SAVSBPM32 to BSA(B) .......................................................... 150

Chapter 5: Discussion .............................................................................................. 154

5.1 Binding stoichiometry of SBPA18C tagged BLA to SAVSBPM32 .................... 154

5.2 Interpretation of the binding kinetics data ............................................................... 158

5.3 Advantages of the SAVSBM32 system .................................................................... 160
5.4 Possible reasons for the inefficient saturation of SAVSBPM32 with two SBPA18C tags ................................................................. 162

5.5 Limitation of the CPFB(-2) SAVBPM32 system and potential solutions .... 163

5.6 Conclusion .................................................................................................................. 167

References .......................................................................................................................... 172

Appendix A: Annotated Pymol scripts ............................................................................. 208

Appendix B: Diagrams of plasmid vectors ....................................................................... 223
List of Tables

Table 2.1 Modeling of streptavidin peptide complexes. ................................................. 36

Table 2.2. Variations of torsion angles within the biotinylated nonapeptide from 
various streptavidin-peptide tag modeled complexes. ........................................... 42

Table 2.3. Properties of *E. coli* BL21(DE3) host strain ........................................ 45

Table 2.4. Properties of *E. coli* DH5α host strain .................................................. 46

Table 2.5 Properties of *B. subtilis* WB800 host strain .......................................... 47

Table 3.1 Location of cysteine replacements predicted by Disulfide by Design to form 
intermolecular disulfide bonds .................................................................................... 78

Table 3.2 Interactions of the SBP tag residues predicted to form intermolecular 
disulfide bonds to streptavidin. ................................................................................ 81

Table 3.3 Streptavidin variants nomenclature and mutations. .................................. 82

Table 3.4 Sequences of the ligand tags and variants .................................................. 83

Table 3.5 Physical parameters of streptavidin and the studied variants.................... 100

Table 3.6 Kinetic parameters for the interactions between streptavidin variants and 
the SBP ligands. ........................................................................................................... 123

Table 3.7 Predicted secondary structure of the SBP tag and variants ....................... 130
Table 4.1 Streptavidin variants nomenclature and mutations. ............................................... 139

Table 4.2 Sequences of the ligand tags and variants. ......................................................... 140

Table 4.3. Kinetic parameters for the interactions between streptavidin variants and the biotin ligands. ............................................................................................................. 153

Table 5.1 Capabilities of streptavidin and various muteins ................................................. 169
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Structure of streptavidin</td>
<td>3</td>
</tr>
<tr>
<td>1.2</td>
<td>Structure of biotin</td>
<td>8</td>
</tr>
<tr>
<td>1.3</td>
<td>Biotin binding to streptavidin</td>
<td>10</td>
</tr>
<tr>
<td>1.4</td>
<td>SBP-Tag binding to streptavidin</td>
<td>16</td>
</tr>
<tr>
<td>1.6</td>
<td>Structure of the <em>E. coli</em> TEM-1 β-lactamase (BLA) reporter protein</td>
<td>33</td>
</tr>
<tr>
<td>2.1</td>
<td>Biotinylated nonapeptides</td>
<td>38</td>
</tr>
<tr>
<td>3.1</td>
<td>Engineering of the streptavidin variants SAVSBPM32, SAVSBPM96 and the SBPA18C peptide tag</td>
<td>79</td>
</tr>
<tr>
<td>3.2</td>
<td>Purification of SAVSBPM18, SAVSBPM32, SAVSBPM96 and SAVSBPM32F</td>
<td>86</td>
</tr>
<tr>
<td>3.3</td>
<td>Purification of BLA-SBP and BLA-SBPA18C.</td>
<td>88</td>
</tr>
<tr>
<td>3.4</td>
<td>Purification of BLA-CPFB.</td>
<td>89</td>
</tr>
<tr>
<td>3.6</td>
<td>Disulfide bond formation between various SAV muteins and cysteine containing tags.</td>
<td>93</td>
</tr>
<tr>
<td>3.7</td>
<td>Carboxymethylation of cysteine residues</td>
<td>96</td>
</tr>
<tr>
<td>3.8</td>
<td>Semi-native polyacrylamide gel electrophoresis</td>
<td>99</td>
</tr>
<tr>
<td>3.9</td>
<td>Modified two dimensional polyacrylamide gel electrophoresis</td>
<td>105</td>
</tr>
<tr>
<td>3.10</td>
<td>Zinc imidazole staining for complex extraction and analysis</td>
<td>107</td>
</tr>
</tbody>
</table>
Figure 3.11 Native polyacrylamide gel electrophoresis ....................................................... 110

Figure 3.12 Purification of BLA-L-SBP from culture supernatant using SAVSBPM18 or SAVSBPM32-affigels .............................................................................................................. 114

Figure 3.13 Purification of BLA-L-SBP4A18C using SAVSBPM18 or SAVSBPM32- affigels ...................................................................................................................................... 116

Figure 3.14 Non-linear curve fitting of SAVSBPM32-BLA-L-SBP and SAVSBPM18- BLA-L-SBP4A18C sensorgram interaction data .................................................................................................................. 121

Figure 3.16 Purification of BLA-SBP4A18C using SAVSBPM32-affigel and various wash buffers .................................................................................................................................. 126

Figure 3.17 Local structure prediction profile of the SBP and SBP4A18C tag .......... 131

Figure 3.18 Purification of MBP-L-SBP4A18CG19A using His-Select and SAVSBPM32- affigels ........................................................................................................................................... 134

Figure 4.1 Model of the PFB peptide .................................................................................. 136

Figure 4.2 Modeling of CPFB(-2) and engineered streptavidin variants SAVSBPM19 and SAVSBPM32 .......................................................................................................................... 137

Figure 4.3 Purification of biotinylated BLA-L-CPFB(-2) ................................................... 142

Figure 4.4 Disulfide bond formation between various SAV muteins and cysteine containing CPFB(-2) tag ......................................................................................................................... 144

Figure 4.5 Tetrameric state of SAVSBPM18, SAVSBPM19 and SAVSBPM32 ............ 147

Figure 4.6 Model of the N118C mutation in loop 7-8 of streptavidin ...................... 149
Figure 4.7. Determination of the kinetic parameters of the interaction between SAVSBPM32-BSA(B) ........................................................................................................................................ 152

Figure 5.1 Two conformations of SBP streptavidin binding ....................................................... 157

Figure 5.2 Modeling the intramolecular disulfide bond of SAVSBPM80................................. 166

Figure 5.3 Schematic for the binding and elution of BLA-L-SBPA18C and BLA-L-CPFBl(-2) from SAVSBPM32......................................................................................................................... 171
List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>Anchor domain of A-Kinase anchor protein</td>
</tr>
<tr>
<td>APS</td>
<td>Ammonium Persulfate</td>
</tr>
<tr>
<td>AR2G</td>
<td>Amine Reactive Second-Generation Biosensor</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>BCCP</td>
<td>Biotin Carboxyl Carrier Protein</td>
</tr>
<tr>
<td>BirA</td>
<td>Biotin ligase from <em>Escherichia coli</em></td>
</tr>
<tr>
<td>BLA</td>
<td>TEM-1 β-lactamase from <em>Escherichia coli</em></td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic Local Alignment Search Tool</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>BSA(B)</td>
<td>Biotinylated bovine serum albumin</td>
</tr>
<tr>
<td>CBD</td>
<td>Cellulose-binding domain</td>
</tr>
<tr>
<td>Chi^2 (χ^2)</td>
<td>Chi-squared</td>
</tr>
<tr>
<td>CM</td>
<td>Carboxymethyl</td>
</tr>
<tr>
<td>CPFB</td>
<td>Cysteine containing peptide tag for biotinylation</td>
</tr>
<tr>
<td>DEAE</td>
<td>Diethylaminoethanol</td>
</tr>
<tr>
<td>DhaA</td>
<td><em>Rhodococcus</em> dehalogenase</td>
</tr>
<tr>
<td>doF</td>
<td>Degrees of freedom</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>EDC</td>
<td>1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>HCl</td>
<td>Hydrogen chloride</td>
</tr>
<tr>
<td>IMAC</td>
<td>Immobilized metal ion affinity chromatography</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>$K_d$</td>
<td>Dissociation constant</td>
</tr>
<tr>
<td>kDa</td>
<td>Kilodaltons</td>
</tr>
<tr>
<td>$K_{off}$</td>
<td>Off-rate</td>
</tr>
<tr>
<td>$K_{on}$</td>
<td>On-rate</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani</td>
</tr>
<tr>
<td>M2D</td>
<td>Modified two dimensional</td>
</tr>
<tr>
<td>MBP</td>
<td>Maltose-Binding Protein</td>
</tr>
<tr>
<td>ME</td>
<td>β- Mercaptoethanol</td>
</tr>
<tr>
<td>Mr</td>
<td>Relative molecular mass</td>
</tr>
<tr>
<td>Mw</td>
<td>Molecular weight</td>
</tr>
<tr>
<td>MWCO</td>
<td>Molecular weight cut off</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein Data Bank</td>
</tr>
<tr>
<td>PFB</td>
<td>Peptide tag for biotinylation</td>
</tr>
<tr>
<td>PMSF</td>
<td>Phenylmethanesulfonylfluoride</td>
</tr>
<tr>
<td>pse</td>
<td>PyMol Session</td>
</tr>
<tr>
<td>$R^2$ (R$^2$)</td>
<td>Coefficient of determination</td>
</tr>
<tr>
<td>RII</td>
<td>Type II regulatory subunit of camp-dependent protein kinase</td>
</tr>
<tr>
<td>s-native</td>
<td>Semi-native</td>
</tr>
<tr>
<td>s-NHS</td>
<td>Sulfo-n-hydroxysulfosuccinimide</td>
</tr>
<tr>
<td>SA</td>
<td>Succinic acid</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>SAV</td>
<td>Streptavidin</td>
</tr>
<tr>
<td>SBP</td>
<td>Streptavidin binding peptide</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>StrA</td>
<td>Sortase A transpeptidase</td>
</tr>
<tr>
<td>TBAB</td>
<td>Tryptose blood agar base</td>
</tr>
<tr>
<td>TBS</td>
<td>Tris(hydroxymethyl)aminomethane-buffered saline</td>
</tr>
<tr>
<td>TCEP</td>
<td>Tris(2-carboxyethyl)phosphine</td>
</tr>
<tr>
<td>TEMED</td>
<td>Tetramethylethylenediamine</td>
</tr>
<tr>
<td>Tm</td>
<td>Melting temperature</td>
</tr>
<tr>
<td>Tris</td>
<td>Tris(hydroxymethyl)aminomethane</td>
</tr>
<tr>
<td>Tween-20</td>
<td>Polysorbate 20</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>wt</td>
<td>Wild type</td>
</tr>
</tbody>
</table>
Chapter one: Introduction

1.1 Overview

Streptavidin (SAV) is a 60-kDa homotetramer protein isolated from the bacterium *Streptomyces avidinii* (Chaiet & Wolf, 1964) that can bind up to four biotin molecules with extraordinarily high binding affinity ($K_d = 10^{-14}$ M) (Green, 1990). An affinity that is magnitudes larger compared to common dissociation constants seen in biological systems, such as between enzymes and substrates ($10^{-3}$ - $10^{-5}$ M) or antibodies and antigens ($10^{-7}$ - $10^{-11}$ M) (Bayer & Wilchek, 1990). Besides the high binding affinity to biotin, tetrameric streptavidin is also highly thermostable (Tm of 73°C for apo-SAV and 112°C for SAV-biotin) and resistance to denaturing agents and extreme pH conditions (Christopher M. Dundas, Demonte, & Park, 2013). The ability to biotinylate biomolecules of interest has resulted in a powerful technology that can immobilize those molecules to streptavidin. Exploiting this ability has led to streptavidin being one of the most widely used proteins in biotechnology with applications spanning across both the academic and industrial sectors (Kim, Lee, & Kim, 2005). These applications included the development of biosensor/protein chips, drug delivery systems, protein arrays, affinity cytochemistry procedures, and novel gene therapy approaches (George et al., 2006; O'Shannessy, Brigham-Burke, & Peck, 1992; Schetters, 1999; Stayton et al., 1999; Wilchek & Bayer, 1990). A recent estimate of streptavidin sales around 2 billion dollars a year (Nogueira et al., 2014) highlights the importance and value associated with this technology. The high demand by many users in different fields makes the system ideal.
for further improvement. Currently, the majority of these advancements are achieved through genetic engineering (Laitinen, Nordlund, Hytönen, & Kulomaa, 2007).

This chapter presents an overview of the principles, advantages, and limitations associated with this technology. The current engineering trends, applications and the major objectives of this work are also outlined.

1.2 The structure of streptavidin

The 60-kDa streptavidin protein is composed of four identical subunits. Each subunit consists of 159 residues, of which 120 are ordered (residues 13-133) and termed “core streptavidin” (Argarana, Kuntz, Birken, Axel, & Cantor, 1986; Bayer, Ben-Hur, Hiller, & Wilchek, 1989) (Figure 1.1, panel A). Subunits contain a network of hydrogen bonds and are structured as an eight-stranded antiparallel β-sheet (Figure 1.1, panel B). The β-sheet is connected by staggered hydrogen bonds and several protruding hairpin loops, such that the first and last sheet are hydrogen bonded together forming a slightly flattened barrel structure with aromatic and polar residues lining an open binding pocket (Hendrickson et al., 1989; P. C. Weber, Ohlendorf, Wendoloski, & Salemme, 1989). The complementary shape of the subunit surfaces and extensive van der Waals interactions favour very stable dimer formation of two streptavidin monomers (A/B and C/D) (Figure 1.1, panel C). Each subunit within a dimer has a small area that interacts with another dimer through hydrogen bonds and van der Waals interactions to form the tetrameric quaternary structure of streptavidin (Figure 1.1, panel D) (Hendrickson et al., 1989; P. C. Weber et al., 1989).
**Figure 1.1 Structure of streptavidin**

(A) Amino acid sequence of wild type streptavidin subunit. Residues associated with core streptavidin are underlined. (B) Streptavidin monomer. (C) Dimer formation of two streptavidin monomers. (D) Dimer of dimers forming tetrameric streptavidin. Models were generated using PyMOL with the PDB entry 4J06 as a starting file (Barrette-Ng et al., 2013).
The remaining residues on the C- and N-termini of each subunit (residues 1-12 and 132-159) contribute to the full-length version of streptavidin. These termini are disordered, flexible and prone to proteolytic degradation (Bayer et al., 1989; P. C. Weber et al., 1989). The presence of these termini in the mature protein does not affect overall folding and is desirable for soluble streptavidin production via secretion (Pähler, Hendrickson, Kolks, Argaraña, & Cantor, 1987; Sau-Ching Wu, Hassan Qureshi, & Wong, 2002). The C-terminal peptide has been shown to decrease ligand binding, (Takeshi et al., 1995), likely by increasing steric hindrance to the binding pocket (Isolde Le Trong, Humbert, Ward, & Stenkamp, 2006).

1.3 Streptavidin homologs and their function

Streptavidin homologs have been identified in many different organisms. These include xenavidin (frog), tamavidin (fungi), bradavidin and rhizavidin (bacteria) and avidin from chicken and several avian species (Ahlroth et al., 2000; Helppolainen et al., 2007; Hertz & Sebrell, 1942; Määttä et al., 2009; Nordlund, Hytönen, Laitinen, & Kulomaa, 2005; Takakura et al., 2009; Tuohimaa et al., 1989). Two of these homologs are of particular interest. At the primary structure level, the amino acid sequence from rhizavidin (produced from *Rhizobium etli*) is approximately 20-30% identical to that of streptavidin and other homologs (Helppolainen et al., 2007). Much of the conserved residues are involved with biotin binding, a trend that is ubiquitous among the other homologs (Helppolainen et al., 2007; Määttä et al., 2009; Meir et al., 2009). What makes rhizavidin distinct is that currently, it is the only homolog found in nature to occur as a
dimer (Helppolainen et al., 2007). The ability for rhizavidin to bind biotin strongly as a dimer has already provided new insight into different engineering strategies for the development of monomeric streptavidin (Helppolainen et al., 2007; Lim, Huang, Pralle, & Park, 2013).

The second interesting streptavidin homolog is avidin, isolated from chicken egg white. This protein has the highest binding affinity to biotin ($10^{-15}$ M) (Green, 1990). It is also worth noting that avidin was the first biotin binding protein described in the literature (Gyorgy, 1954). The term streptavidin was based on finding an avidin like protein in *Streptomyces* (*strept*-avidin) (Chaiet & Wolf, 1964). Similar to other homologs, avidin and streptavidin share approximately 30% sequence identity, while the major structural elements and critical binding pocket functional groups are conserved (Livnah, Bayer, Wilchek, & Sussman, 1993). A striking difference between streptavidin and avidin is the carbohydrate content of the two proteins. Close to 10% of the total molecular weight of avidin is contributed by carbohydrates. In contrast, streptavidin does not contain any carbohydrates (Bruch & White, 1982; Korpela, 1984). This high carbohydrate content in avidin contributes to non-specific binding and has limited its uses in biotechnology (Green, 1990). Modified, non-glycosylated derivatives of avidin, NeutraLite Avidin and the recombinant NeutraLite Avidin, have been described to address the non-specific binding limitation (Marttila et al., 2000).

The diverse range of organisms described to have streptavidin homologs suggests that more streptavidin like proteins are likely to be found. Finding and characterizing these proteins are important since they may include unique properties that could aid in future research and technology advancement. To support this possibility, the BLAST
bioinformatics tool was used to search the Uniport database for proteins sequences similar to streptavidin (Altschul et al., 1997). The search revealed 109 such sequences, which correspond closely to the over 104 statistically significant similar streptavidin like sequences reported by others (Dundas, Demonte, & Park, 2013).

Biotin (known as vitamin H) is required by all organism for cell growth and synthesized by plants and many prokaryotes (Chapman-Smith & Jr, 1999; Peterson & Peterson, 1945). Within cells, biotin plays roles in gene regulation, cell signalling and biosynthesis of fatty acids, amino acid catabolism and gluconeogenesis (Dakshinamurti, Chalifour, & Bhullar, 1985; Pacheco-Alvarez, Solórzano-Vargas, & Del Río, 2002; Polyak & Chapman-Smith, 2013; Zempleni, Wijeratne, & Hassan, 2009). Although the biological role of streptavidin and its homologs is not fully understood, the high affinity to biotin has led to the suggestion that these proteins act as antimicrobial agents by depleting the availability of the essential vitamin to other microorganisms (Peterson & Peterson, 1945; Tuohimaa et al., 1989; P. C. Weber et al., 1989).

1.4 Host cells used for the production of streptavidin

Since cloning and the original nucleotide sequence of the streptavidin gene were reported (Argarana et al., 1986), a variety of host cells and procedures have been used for streptavidin production. These include production by secretion from the natural host *Streptomyces avidinii* (Aldwin, Toso, Goodson, & Hunter, 1990; Cazin Jr., Suter, &Butler, 1988), the bacterial recombinant host, *Bacillus subtilis* (Nagarajan, Ramaley, Albertson, & Chen, 1993; Sau-Ching Wu, Hassan Qureshi, et al., 2002; Sau-Ching Wu &
Wong, 2002) and more recently, the eukaryotic host, *Pichia pastoris* (Casteluber et al., 2012; Nogueira et al., 2014). Apart from these host, streptavidin can be produced intracellularly by *Escherichia coli* in either soluble or insoluble form (Gallizia et al., 1998; T Sano & Cantor, 1990).

Each host system has different strengths and weaknesses for producing streptavidin. Insoluble streptavidin produced from *Escherichia coli* with a production yield of \( \approx 65 \text{ mg L}^{-1} \), requires *in vitro* refolding and binds 3.5-3.9 molecules of biotin per tetramer (Sano & Cantor, 1990). Soluble streptavidin produced by *Escherichia coli* with a production yield of \( \approx 70 \text{ mg L}^{-1} \) (Gallizia et al., 1998), does not require re-folding of the protein, and has a reduced number of free binding sites, due to the trapping of biotin in some binding pockets. Although *Pichia pastoris* has a high production yield of streptavidin (\( \approx 650 \text{ mg L}^{-1} \)), which does not require denaturing and refolding steps and can bind 3.2 molecules biotin per tetramer, the overall process takes 10 days (Nogueira et al., 2014).

In contrast, *Bacillus subtilis* (using the WB800BIO strain) can produce soluble streptavidin via secretion at a level of \( \approx 50 \text{ mg L}^{-1} \) in 15-20 hours. The functionally purified streptavidin can bind 3.3 molecules of biotin per tetramer (Wu & Wong, 2002).

### 1.5 Binding of biotin to Streptavidin

Biotin contains a valeric acid side chain attached to a tetrahydrothiophene ring, fused to a ureido ring (Figure 1.2) (Vigneaud, 1942).
Figure 1.2 Structure of biotin

Biotin also known as vitamin H is required for the growth of all cells. The molecule contains a valeric acid chain attached to a tetrahydrothiophene ring, fused to a ureido ring. SMILES: C1C2C(C(S1)CCCCC(=O)O)NC(=O)N2
Binding of biotin to streptavidin involves the displacement of several water molecules bound to the aromatic and polar residues within the binding pocket. Biotin is buried within the binding pocket by ordering the previously disordered residues 45-53, which together form the loop between β-strands three and four of streptavidin (loop_{3-4}). The flexibility of Loop_{3-4} makes this loop function as an entry gate for biotin. With biotin binding, interactions between biotin and the loop result in ordering the loop in a closed conformation with biotin trapped in the binding pocket (Figure 1.3). The formation of a salt bridge interaction between Glu-51 and Arg-84 and a hydrogen bond between residue Ser-45 (loop_{3-4}) and the ureido amino group contributes to the stabilization of the loop into the closed position (Figure 1.3, panel B) (Klumb, Chu, & Stayton, 1998; P. C. Weber et al., 1989).

The second factor that contributes to the exceptionally high biotin binding affinity to streptavidin is the formation of an extensive hydrogen bonding network between biotin and the polar residues in the biotin binding pocket. The ureido ring of biotin forms five hydrogen bonds, three at the ureido oxygen (with Asn-23, Ser-27 and Tyr-43) and one at each of the ureido amino groups (with Ser-45 and Asp-128) (Figure 1.3, panel C). The tetrahydrothiophene ring of biotin forms a hydrogen bond through the thiophene sulfur (with Thr-90). Additional hydrogen bonds are formed between the valeric acid side chain of biotin with Ser-88 of streptavidin (Figure 1.3, panel C) (Klumb et al., 1998; Livnah et al., 1993; Miyamoto & Kollman, 1993).

The third factor for the exceptionally tight binding between streptavidin and biotin is that streptavidin has Leu-25, Leu-110, Val-47 Trp-79, Trp-92 and Trp-108 to form a hydrophobic pocket to interact with biotin.
Figure 1.3 Biotin binding to streptavidin

(A) Residues in the Loop$_{3-4}$ region of streptavidin in the open conformation. (B) Conformational change in Loop$_{3-4}$ when biotin binds the pocket of streptavidin. This binding triggers the Loop$_{3-4}$ region to close over top of the binding site. (C) Hydrogen bonds formed between biotin and streptavidin. Residues involved in hydrogen bonds are shown in cyan with the bond shown in black. Numbers to the sides of the residues indicate the residue number. (D) Hydrophobic interactions between streptavidin and biotin. Residues involved in hydrophobic interactions are shown in yellow (pink indicates residues from the neighboring streptavidin subunit). Numbers beside the residues indicate the residue number. Models were generated using PyMOL with the PDB entry 3RY2 as a starting file (Le Trong et al., 2011).
Two residues, Trp-120 and Lys-121, from the adjacent subunit also contribute to the completion of the biotin binding pocket through both hydrophobic and electrostatic interactions with biotin (Figure 1.3, panel D) (Hendrickson et al., 1989; Miyamoto & Kollman, 1993; Takeshi Sano, Vajda, Reznik, Smith, & Cantor, 1996; P. C. Weber et al., 1989). With the exception of the valeryl carboxyl oxygens which are partially solvent accessible, biotin is effectively buried within the pocket (Figure 1.3, panel D) (Miyamoto & Kollman, 1993; P. C. Weber et al., 1989).

The overall binding interaction causes streptavidin subunits to pack more tightly and become slightly flatter. These geometric changes in streptavidin alter the hydrogen-bonded dimer geometry and dimer packing of the tetramer (P. C. Weber et al., 1989) which has been shown to increase protein stability. This increase is demonstrated by a 50% increase in thermostability of streptavidin in the biotin bound state compared to apo-streptavidin (González, Argaraña, & Fidelio, 1999).

1.6 Methods for biotinylation of molecules

The addition of biotin to molecules has limited effects on the function of the target molecules (Wilchek & Bayer, 1990). This is due to the small size of biotin (244.31 g/mol) and the presence of a derivatizable carboxylic group in valeric acid. This group is partially exposed in the biotin bound streptavidin complex.

Attaching biotin to molecules of interest can be done using either chemical or enzymatic means. Chemical biotinylation of proteins in vitro requires simple chemical reagents and mild conditions. For these reactions, the biotin molecule is first linked to a
reactive moiety either directly or indirectly with the addition of a spacer arm. The mechanism of protein coupling to biotin depends on the reactive moiety used. The disadvantage of chemical biotinylation is that the attachment of biotin is both, not very specific and not uniform (Kay, Thai, & Volgina, 2009).

Enzymatic biotinylation is more time consuming and costly but allows for specific control over the addition of biotin to proteins. The *Escherichia coli* biotin ligase BirA, uses ATP as a co-substrate to biotinylate a single specific lysine residue found in the carboxyl carrier protein (BCCP) (Barker & Campbell, 1981; Chapman-Smith & Jr, 1999). Since a fusion protein containing the full-length BCCP would increase steric interference, the minimum recognition sequence required by BirA for biotinylation was studied by screening peptide libraries (Schatz, 1993). From those identified sequences, three were further analyzed, resulting in the 14-residue peptide (GLNDIFEAQKIEWH) which is biotinylated at twice the rate of the natural substrate (Beckett, Kovaleva, & Schatz, 1999). The addition of a single Glu residue to the C-terminal end of the tag leads to the development of the 15-residue AviTag (GLNDIFEAQKIEWHE), which has a higher biotinylation rate and can be included in either the terminal positions (N-terminus or C-terminus), or an internal surface exposed loop of fusion proteins (Beckett et al., 1999; Cull & Schatz, 2000). The AviTag has been shown to be effective for both, *in vivo* and *in vitro* systems (Ashraf, Benson, Payne, Halbleib, & Grøn, 2004; Athavankar & Peterson, 2003; Cull & Schatz, 2000). In our laboratory, another 15-amino acid (LHHILDAQKMVWNHR) peptide tag (Wu, Yeung, Hwang, & Wong, 2002) was selected from the peptide library (Schatz, 1993) for efficient *in vitro* biotinylation by an engineered BirA enzyme. At that time, the AviTag was not commercially available. This
tag has the biotinylation efficiency comparable to that of AviTag and is named PFB (Peptide for Biotinylation) tag.

1.6.1 Challenges in the release of biotinylated biomolecules

As a result of both chemical and enzymatic biotinylation methods, biotin has been attached to a variety of molecules including proteins (Bayer, Skutelsky, & Wilchek, 1979), peptide hormones (Haigler, 1983), lipids (Blankenburg, Meller, Ringsdorf, & Salesse, 1989), nucleic acids (Langer, Waldrop, & Ward, 1981) and nanoparticles (Aslan, Luhrs, & Pérez-Luna, 2004; Chirra, Sexton, Biswal, Hersh, & Hilt, 2011). These biotinylated molecules are used in a wide range of applications including affinity chromatography, cytochemistry and labelling as well as blotting, immunoassays, flow cytometry, biosensor/protein chips, drug delivery and nanoparticles (George et al., 2006; Li, Wong, & Mann, 1999; O’Shannessy et al., 1992; Wilchek & Bayer, 1990).

It is important to note that the applications described for biotinylated molecules rely on the immobilization of biotinylated molecules to streptavidin. Due to the extremely harsh conditions needed to release the captured biotinylated targets, including the presence of 6 M guanidine hydrochloride at pH 1.5, the reusability of such systems or the functional use of released targeted molecules is severely limited (Takeshi Sano, Vajda, & Cantor, 1998). Other methods used to release biotinylated target molecules include the introduction of a cleavable disulfide bond within the biotin linker (Shimkus, Levy, & Herman, 1985) and the use of photocleavable biotin derivatives (Olejnik, Sonar,
Krzymanska-Olejnik, & Rothschild, 1995). Although release of the target molecules is possible with these systems, these procedures are both time consuming and costly.

1.7 Streptavidin binding peptides (SBP) and the binding of the SBP tag to streptavidin

In addition to biotin, streptavidin can also bind strongly to a group of peptide ligands. Development of streptavidin binding peptides that can be engineered into recombinant proteins addresses the above-mentioned limitation. These peptides further broaden the uses of the streptavidin technology to include applications that do not rely on irreversible binding, such as protein purification (Keefe, Wilson, Seelig, & Szostak, 2001). One such peptide, the “Strep-tag” was developed using a genetic fusion peptide library and can be incorporated into the C-terminal end of engineered proteins for purification (Kₐ = 37 µM) (T. G. M. Schmidt, Koepke, Frank, & Skerra, 1996; T. G. M. Schmidt & Skerra, 1993). Spot synthesis assay was later used to develop a variant (Strep-tag II), which can be introduced to either the N- or C-terminal end of the recombinant protein (Voss & Skerra, 1997). This tag has a relatively weak binding to streptavidin (Kₐ = 72 µM) (T. G. M. Schmidt et al., 1996). A mutated streptavidin variant (Strep-Tactin) was developed with an increased binding affinity to the Strep-tag II (Kₐ ≈ 1 µM) (Korndörfer & Skerra, 2002; T. G. Schmidt & Skerra, 2007). The other approach is to develop newer peptide variants with higher affinities. These included the 15 and 9 amino acids long Nano-tag(15) and Nano-tag(9) peptide tags, which as their name implied, have nanomolar affinities to streptavidin (Kₐ = 4.0 and 17 nM, respectively) (Lamla &
Erdmann, 2004). However, this tag has to be located at the N-terminal end of recombinant proteins, which have to be produced from prokaryotes such as *E. coli*. This restriction is imposed because a formylated methionine at the N-terminus is required for tight binding of the nano-tag to streptavidin (Perbandt et al., 2007). In 2001, several streptavidin binding peptides were isolated by screening a peptide library (Wilson, Keefe, & Szostak, 2001). One of these peptides was further modified to generate a 38-amino acid peptide designated SBP for streptavidin binding peptide. This tag shows nanomolar affinity to streptavidin ($K_d \sim 2.5nM$) (Keefe et al., 2001). SBP tagged proteins can be affinity purified in one step with purity superior to that of His-tagged or maltose binding protein tagged proteins (Keefe et al., 2001). Deleting parts of both the N- and C-terminal segments of the SBP tag (residues 1-10 and 35-38) results in a smaller binding tag that has similar binding properties as the parental tag (Barrette-Ng, Wu, Tjia, Wong, & Ng, 2013). This modified tag is termed SBP-TagII.

The SBP tag can be divided into multiple segments, the N-terminal “GHVV” tetrapeptide, a two-amino acid linker, the central $\alpha$-helix, another two-amino acid linker and the C-terminal “HPQGQ” pentapeptide (Figure 1.4, panel A).
The SBP-Tag can be separated into four segments (Panel A) the middle $\alpha$-helix (cyan), the two dipeptide linkers (yellow), the N-terminal GHVV binding peptide (green) and the C-terminal HPQGQ binding peptide (green). Both hydrogen bonding and hydrophobic interactions in adjacent streptavidin subunit binding pockets are involved in binding. (B) The $\alpha$-helix acts as a spacer to correctly position the binding peptides into their respected streptavidin subunits and anchors the peptide down with leucine residues extending from the helix into the hydrophobic portions of streptavidin. The figure in Panel C was modified from Barrette-Ng, Wu, Tjia, Wong & Ng, 2013 and shows the sequence of the SBP-tag as a schematic diagram indicating residue interactions with streptavidin subunits. Numbers above residues indicate the residue position within the tag. Models were generated using PyMOL with the PDB entry 4J06 (Barrette-Ng et al., 2013) as a starting file.
1.7.1 Binding interactions of the SBP tag central $\alpha$-helix to streptavidin

Residues 17-28 of the SBP-tag adopt an $\alpha$-helix structure that spans between streptavidin subunits A/C or B/D (Barrette-Ng et al., 2013). Four Leu residues (residues 17, 21, 24, and 28) reside within the $\alpha$-helix of the SBP tag. All four Leu residues are buried in streptavidin and likely act as “screws” holding the bound peptide onto the surface (Figure 1.4, panel A). To illustrate the interactions between SBP and streptavidin in detail, the SBP tag bound to the A and C subunits of streptavidin is selected as an example. Leu-17 of the tag forms hydrophobic interactions with the streptavidin subunits A and C (residues; Trp-79, Arg-84, Ala-86, and Ser-88 in subunit A and Trp-120 in subunit C). Leu-21 has binding interactions with residues Ser-88, Ser-112, and Lue-110 on streptavidin subunit A. Leu-24 and Leu-28 also interact with Leu-124 of streptavidin subunits A and C, respectively. Ala-18 of the SBP tag has interactions with Ala-86 on streptavidin subunit A. Arg-25, and Arg-27 of the SBP tag interact with streptavidin residues Lys-121, and Ser-112 on streptavidin subunits A and C, respectively. Lys-121 of subunit C interacts with Glu-20 and Gln-23 of the $\alpha$-helix. Trp-120 of streptavidin subunit C also interacts with Glu-20. Lys-121 of subunit A interacts with Arg-25 of the SBP tag (Figure 1.4). Residues Gly-19, Glu-22, and Ala-26 of the $\alpha$-helix are situated away from streptavidin and subsequently do not contribute to binding interactions (Barrette-Ng et al., 2013).
1.7.2 *Binding interactions of the SBP tag N-Terminal GHVV peptide to streptavidin*

Residue His-12 of the N-terminal GHVV sequence forms extensive hydrophobic and hydrogen-bonding interactions deep within the biotin-binding pocket of streptavidin subunit A (with Asn-23, Ser-27, Trp-79, Thr-90, Trp-92, Trp-108, Asp-128) (Figure 1.4). His-14 also binds within the biotin-binding pocket through hydrophobic interactions with Trp-120, Leu-25, Tyr-43, Trp-79, and Tyr-54 of the streptavidin subunit. Although Val-13 does show some hydrophobic interactions with streptavidin (Trp-120, Leu-25, Tyr-43, Trp-79, and Tyr-54), it is unique within the GHVV sequence, as it is not buried within the binding pocket and largely solvent-exposed (Barrette-Ng et al., 2013). Similar to biotin, the residues in the GHVV sequence (as well as Gly-11) also interact with the neighbouring streptavidin subunit (Trp-120) to form additional hydrophobic interactions (Figure 1.4). Situated in an exit channel, Gly-11 has otherwise limited interactions with streptavidin (Asn-23, Gln-24, Leu-25). Subsequent residues past Gly-11 are not involved in binding interactions, and as such were unable to be captured through electron-density maps, which suggests that residues 1-10 of the SBP tag extend away from streptavidin and do not contribute to binding interactions (Barrette-Ng et al., 2013).

1.7.3 *Binding interactions of the SBP tag C-Terminal HPQGQ peptide to streptavidin*

The C-terminal HPQGQ sequence forms a $3_{10}$-helical structure within the binding pocket of streptavidin subunit C and interacts through a large network of hydrogen bonds (residues 79, 92, and 108), and hydrophobic interactions (residues 23, 27, 43, 88, and 90). The hydrophobic Trp-120 residue on neighbouring streptavidin subunits also contributes
to the strong binding of the HPQGQ motif, and Gly-34 likely adds binding energy by tightly packing next to Trp-120 (Barrette-Ng et al., 2013). Residues Gly-34 and Gln-35 protrude out of the binding pocket and interact with streptavidin residues Trp-120 (from the adjacent subunit) and Arg-54, respectively.

### 1.7.4 Binding interaction of the SBP tag dipeptide linkers to streptavidin

Immediately following the α-helix in both directions are two dipeptide linkers. The Glu-15 and Gly-16 dipeptide linker separates the N-terminal GHVV tetrapeptide and the α-helix. This dipeptide correctly positions the GHVV sequence within the binding pocket and also adds binding energy through hydrophobic and hydrogen-bonding interactions with Ser-45, Val-47, and Arg-84 of streptavidin subunit A (Barrette-Ng et al., 2013). Unlike Glu-15 and Gly-16, the dipeptide Glu-29 and His-30 between the α-helix and the C-terminal HPQGQ binding sequence adds virtually no binding energy with the exception of His-30, which has some interactions with Arg-84 of subunit C (Barrette-Ng et al., 2013). This lack of binding energy suggests that the role of the linker residues is for the most part to act as a flexible scaffold allowing freedom for the HPQGQ binding peptide to become correctly positioned within the binding pocket.

### 1.7.5 The major drawback of the streptavidin – SBP tag system

The limitation of the streptavidin SBP tag system is that the elution of SBP tagged proteins bound to streptavidin is done by competitive binding using excess free biotin (Keefe et al., 2001). Free biotin is able to successfully compete with the SBP tag for the
binding pockets in streptavidin, eluting the SBP tag. The streptavidin matrix therefore becomes saturated with tightly bound biotin ($K_d = 10^{-14}$ M), resulting in a non-reusable streptavidin column.

1.8 Current trends in engineering: Reusable streptavidin technologies

Extensive research has been done addressing this single use limitation. These developments can be grouped into general strategies, which both have the same aim of decreasing binding affinity of streptavidin to its ligands.

1.8.1 The use of biotin analogs for reusable streptavidin systems

The first strategy involves the use of biotin analogues as the ligand for binding. In this strategy, a molecule that resembles biotin is used as the ligand. Streptavidin can recognize and bind to these molecules but the subtle differences result in lower binding. The most frequently used biotin analog is the cyclic guanidino biotin analog (2-iminobiotin). The use of this analog relies on the pH dependant forms of 2-iminobiotin. At high pH (>9), 2-iminobiotin is present in its free base form, which has a high binding affinity to streptavidin. In an acidic environment (pH 4), the molecule exists primarily in a protonated form, which has a low affinity to streptavidin (Fudem-Goldin & Orr, 1990; Heney & Orr, 1981). The harsh acidic environment needed for elution limits the application of the system to purify proteins that are known to be resistant to acidic conditions. Furthermore, due to the inability of BirA to recognize this analog,
specific addition of 2-iminobiotin to proteins is not possible. As a result, 2-iminobiotin is not suitable for protein purification other than streptavidin.

A recent example of a biotin analog that does not require harsh elution conditions is the structurally guided design of N3’ Ethyl biotin (Ying & Branchaud, 2011). This molecule was engineered by modifying the ureido amino group of biotin which forms a hydrogen bond with Ser-45 of loop3-4, promoting the closed loop conformation over the binding pocket (Klumb et al., 1998; P. C. Weber et al., 1989; Ying & Branchaud, 2011). Although this strategy works as streptavidin can bind to N3’ Ethyl biotin with lower affinity, this system requires the addition of the unmodified biotin for N3’ Ethyl biotin elution. Similar to the SBP tag system, the addition of biotin results in a non-reusable streptavidin column.

A promising biotin analog that can be used is desthiobiotin. This molecule binds to streptavidin weaker than biotin and can be eluted by competition using excess free desthiobiotin (Hirsch et al., 2002). Previously, the major limitation in the use of desthiobiotin was the lack of an enzymatic, site-specific method for the addition of the molecule to proteins. A solution to this limitation is the use of BirA, which has been shown to recognize and incorporate desthiobiotin into proteins in a site specific manner (Wu & Wong, 2004). However, the efficiency of incorporation of desthiobiotin to biotinylation tags is quite low. This problem has recently been addressed by the successful isolation of a mutated BirA enzyme that can incorporate desthiobiotin to biotinylation tags with a much greater efficiency (Lu, Levy, Kincaid, & Ellington, 2014).

One of the limitations of the use of biotin analogs is that the vast array of commercially available biotin conjugates cannot be used. The second limitation is that
desthiobiotin contains low levels of biotin as the contaminant. These biotin containing elution buffers can gradually poison the streptavidin column. These limitations decrease the potential uses of the designed system.

1.8.2 Engineering monomeric and tetrameric streptavidin variants for reusable systems

The second strategy involves engineering streptavidin variants that have lower biotin binding affinity. One way to decrease biotin binding is by engineering monomeric streptavidin variants. The monomeric streptavidin approach relies on the fact that a fully functional biotin binding pocket requires the contribution of W120 from the neighbouring subunit to form part of a hydrophobic wall in the biotin binding pocket (Laitinen et al., 1999). By breaking the tetramer into monomers, the W120 residue from the neighbouring subunit is no longer available and the biotin binding affinity can be decreased. The disadvantage of this system is that streptavidin monomers are unstable and prone to aggregate (Christopher M. Dundas et al., 2013; Takeshi Sano, Vajda, Smith, & Cantor, 1997). Another disadvantage is that monomeric variants are unable to bind SBP tagged proteins because of the unique binding of the SBP tag across two streptavidin subunits (Barrette-Ng et al., 2013).

Engineering tetrameric streptavidin with reversible biotin binding can solve many of the described limitations for monomeric streptavidin variants. Recently, the streptavidin variant SAVSBPM18 was engineered, which has high binding affinity to the SBP tag and, at the same time, retains reversible binding capability to biotin (Sau-Ching
Wu & Wong, 2013). The SAVSBPM18 mutant achieves these characteristics through a G48T and S27A double mutation. Gly-48 in streptavidin is located in the mobile loop3-4 structure which becomes ordered during biotin binding in the closed position, burying biotin within the pocket (Klumb et al., 1998; P. C. Weber et al., 1989). During SBP tag binding, the mobile loop3-4 remains in an open conformation and therefore the Gly-48 residue of streptavidin does not contribute significantly to SBP tag binding (Barrette-Ng et al., 2013). Replacing Gly-48 with the bulkier Thr-48 residue lowers the ability for streptavidin to form a closed binding pocket, which decreases biotin binding affinity from $10^{-14}$ to $10^{-10}$ M without changing the SBP tag binding affinity (Figure 1.5) (Sau-Ching Wu & Wong, 2013). To further decrease biotin binding, an additional mutation was included. The Ser-27 residue in streptavidin is involved in a strong hydrogen bond with the ureido oxygen of biotin and it has been shown to decrease the binding affinity from $10^{-14}$ to $10^{-12}$ M when mutated to Ala-27 (Klumb et al., 1998; P. C. Weber et al., 1989). Interestingly, this same Ser-27 residue is only weakly involved in SBP tag binding through interactions with the His-27 and Val-14 of the N-terminal GHVV and Gln-33 of the C-terminal HPQGQ binding peptides (Barrette-Ng et al., 2013). The Ser-27-Ala mutation was therefore predicted to not dramatically reduce the SBP tag binding (Figure 1.5) (Wu & Wong, 2013).
Figure 1.5 SAVSBPM18 mutations S27A and G48T

Mutations included in the SAVSBPM18 variant. (A) The three hydrogen bonds formed with the ureido oxygen of biotin. Residues involved in hydrogen bonding with the ureido oxygen are shown in yellow and labeled. Dotted lines indicate hydrogen bonds. (B) The loss of a hydrogen bond caused by the S27A mutation. (C) Residue G48 located in the closed conformation of loop$_{3-4}$. (D) Model for showing the steric hindrance of loop$_{3-4}$ for the closed conformation when G48 is mutated to T48. Models were generated using PyMOL with the PDB entry 3RY2 (Le Trong et al., 2011) as a starting file.
The streptavidin mutant, SAVSBPM18, with both S27A and G48T mutations forms a stable tetramer which has both a decreased affinity for biotin (\(K_d = 1.57 \times 10^{-8}\) M) and a high affinity for the SBP tag (\(K_d = 7.67 \times 10^{-9}\) M) (Wu & Wong, 2013). The advantage of the SAVSBPM18 system is that the lower biotin binding allows the mutein to bind either biotin tagged molecules or SBP tag fusion proteins in a reversible and reusable manner.

1.9 Current trends in engineering: Streptavidin variants with higher binding affinity

Although the binding of streptavidin to biotin is impressive, the assumption that it is essentially irreversible is misleading. Using the reported off-rate (2.4 \(\times\) 10\(^{-6}\) s\(^{-1}\)) of biotin from streptavidin the following equation was used to calculate the binding half-life of the streptavidin-biotin complexes in seconds (\(T_{1/2} = 0.693/K_{off}\)) (Piran & Riordan, 1990). The calculation shows that the half-life of binding is approximately 2.9-3.3 days (Green, 1990; Piran & Riordan, 1990). This binding has been noted to decrease drastically when other molecules are introduced to the system. For example, the attachment of quantum dot nanocrystals has been reported to increase the off-rate to 5 \(\times\) 10\(^{-5}\) s\(^{-1}\) (Swift, Heuff, & Cramb, 2006). For the streptavidin-biotin complex the resulting half-life of the complexes is in the order of hours instead of days. This lower binding is supported by another finding, which observed the dissociation of biotin-labeled \(\alpha\)-bungarotoxin from streptavidin in 2 hours (Bruneau, Sutter, Hume, & Akaaboune, 2005). Additionally, other studies have reported a decrease in binding affinity from \(10^{-14}\) to \(10^{-7}\)
M when streptavidin coated beads were used to bind biotinylated peptides (Buranda, Lopez, Keij, Harris, & Sklar, 1999).

1.9.1  Engineering higher binding affinity with non-covalent interactions

Recently, streptavidin variants have been engineered to have a higher affinity to biotin compared to wild-type streptavidin. One of these systems utilizes the flexible C and N termini of each subunit (residues 1-12 and 132-159) that are removed from core streptavidin. The Stv-13 streptavidin variant was engineered with an even shorter C and N termini (compared to core streptavidin). This variant showed an enhanced ability to bind biotinylated DNA compared to the wild-type core streptavidin (Takeshi et al., 1995). An even more remarkable variant is Traptavidin, which has a slower off-rate (>10 fold) of free biotin compared to the wild-type streptavidin (Chivers et al., 2010). The slower off-rate of Traptavidin is due to the double mutation S52G and R53D. These mutations promote loop3-4 of streptavidin to be in the closed position, which therefore makes the release of bound biotin less likely (Chivers et al., 2010; Chivers, Koner, Lowe, & Howarth, 2011). The drawback of this system is that although it is harder for biotin to escape the binding pocket once bound, it is also harder for the molecule to enter the binding pocket. As a result, the on-rate is decreased two-fold in Traptavidin (Chivers et al., 2011). Although studies on the ability of the SBP tag to bind to Traptavidin have not been reported, it is worth noting that the preferred closed loop conformation of Traptavidin may limit this possibility.
1.9.2 Engineering infinite binding affinity with intermolecular covalent bond formation

Although the previously discussed streptavidin variants (Stv-13 and Traptavidin) have been able to increase binding affinity by non-covalent interactions, an emerging trend in protein engineering is the development of binding systems with infinite binding affinity through intermolecular covalent bond formation. One approach is through the formation of an isopeptide bond. This type of bond occurs when the side chain of an amino acid becomes linked to a protein's backbone through an amide bond. In nature, these bonds have various functions including the posttranslational modifications of microtubules in protists and the covalent linkage of pilin subunits in Gram-positive bacteria (Hendrickx, Budzik, Oh, & Schneewind, 2011; Proft & Baker, 2009; Redeker et al., 1994; Westermann & Weber, 2003). The spontaneous formation of an isopeptide bond in the *Streptococcus pyogenes* pilin Spy0128 (Kang et al., 2007) has been utilized for generating covalent bond forming fusion tags (Zakeri & Howarth, 2010). In this system, a peptide and a small protein domain were engineered from Spy0128, termed SpyTag and SpyCatcher. These fragments can then be used to covalently link fusion proteins (Zakeri et al., 2012; Zakeri & Howarth, 2010). Currently, the SpyTag/SpyCatcher system is limited by the large size of the SpyCatcher (84 amino acids) and a low on-rate between the SpyTag and SpyCatcher ($10^3$ M$^{-1}$ s$^{-1}$) (Veggiani, Zakeri, & Howarth, 2014).

The other method for intermolecular covalent bonding formation is Sortagging. In this method, two proteins, one containing a sortase recognition motif and the other
containing a sortase substrate motif, can be coupled together by the Sortase A transpeptidase (SrtA) enzyme (Strijbis, Spooner, & Ploegh, 2012). One of the limitations in sortagging is the inability to use phosphate-based buffers. This is due to the requirement of calcium in the reaction buffer. Calcium tends to precipitate in phosphate-based buffers (Ritzefeld, 2014). Another enzyme-based coupling system is the commercially available HaloTag (Promega). In the HaloTag system, a protein of interest is fused to the 34 kDa monomeric *Rhodococcus* dehalogenase (DhaA) variant. This variant can then bind and form a covalent bond to molecules containing a chloroalkane linker (Los et al., 2008). Alternatively, native chemical ligation can be used to covalently attach the side chain of a cysteine residue to a thioester through nucleophilic attack (Dawson, Muir, Clark-Lewis, & Kent, 1994). This system however, is limited to reactions with proteins that have an N-terminal cysteine residue (Dawson et al., 1994; Kent, 2009).

Ironically, a common weakness in the above-described infinite binding affinity systems is that the interactions are permanent via covalent bond formation. These systems can only be used for immobilization. A solution to this limitation is the use of disulfide bonds, which allows covalent bonds to be formed in a reversible manner depending on the presence or absence of reducing agents. The Dock-and-lock method (Chang, Rossi, & Goldenberg, 2007) utilizes disulfide bonds and the binding interactions between the anchor domain (AD) of A-Kinase anchor protein and a dimer formed by two type II regulatory subunit (RIIs) of the camp-dependent protein kinase (Baillie, Scott, & Houslay, 2005; Carr et al., 1991; W. Wong & Scott, 2004). In this system the dimerization of two engineered cysteine containing, 44-amino acid RII peptides become
covalently linked by another engineered, double cysteine containing, 17-amino acid AD peptide (Chang et al., 2007; Goldenberg, Rossi, Sharkey, McBride, & Chang, 2008; Rossi, Goldenberg, & Chang, 2012). A similar system uses the binding affinity of biotin to the monomeric streptavidin variant M6 (Wu, Ng, & Wong, 2009). In this system, the cysteine containing M6 variant and a modified cysteine containing PFB fusion tag (CPFB) become covalently linked via the disulfide bond. The bond can then be broken using a reducing agent allowing separation of the complex and reuse of the system (Sau-Ching Wu et al., 2009). Similar to the other monomeric streptavidin variants, M6 is limited by instability, aggregation, and the inability to bind the SBP tag (Barrette-Ng et al., 2013; Dundas et al., 2013; Sau-Ching Wu et al., 2009).

1.10 The need to develop a tight binding system with reversible binding capability

In many situations, strong but reversible interactions are desirable. The first example is protein purification. If the protein of interest can be covalently captured directly or indirectly by the column matrices, stringent but non-denaturing conditions can be applied for washing and eluting the contaminated proteins off from the column. Covalent chromatography by thiol-coupling is an attractive method for achieving this objective (Hillson, 1981). Protein of interest can be collected in high purity under the reversible binding condition (Batista-Viera, Rydén, & Carlsson, 2011; Oscarsson & Porath, 1993). The traditional thiol coupling approach for protein purification has a major drawback. Proteins of interest have no affinity to the matrices. The thiol coupling reaction has to be driven using high concentrations of the thiol containing proteins. The
strategy in this study introduces the affinity based thiol coupling chromatography which addresses this concern.

The second example is the development of a reusable biosensor chip in the BIAcore system. The BIAcore system allows the study of molecular interactions (Murphy, Jason-Moller, & Bruno, 2006; Rich & Myszka, 2006). One of the interacting molecules (molecule A) has to be immobilized to the sensor chip. The other interacting molecules (molecule B) will then be injected into the microfluidic system (Jason-Moller, Murphy, & Bruno, 2006). Association and dissociation can be monitored by the optical unit in the system (Myszka, 1997). Each sensor chip contains a piece of gold film and can only chemically immobilize one interacting molecule (e.g. molecule A). If a different molecule (e.g. molecule C) would be of interest to be immobilized to the sensor chip, another chip will be needed. Although sensor chip with immobilized wild-type streptavidin has been developed, this chip again is suitable for immobilization of one type of biotinylated molecule (O’Shannessy et al., 1992). Since each chip is relatively expensive, it is desirable to have reusable sensor chip available. The tight binding (e.g. via disulfide bond formation) will ensure that the immobilized proteins (e.g. molecule A) will not fall off during the experimental conditions. The reversible binding condition will allow the immobilized proteins to be stripped off and a different protein (e.g. molecule C) to be immobilized. Similar applications can be applied to develop reusable bioreactors, protein chips, nano-particles, quantum dots, magnetic beads and immobilized reagents (e.g. TEV proteases immobilized to beads for cleavage of protein fusions).
1.11 Research objectives

It is my research objective to expand the current streptavidin technology by engineering a system where streptavidin shows an extremely tight binding to its target ligands (i.e. proteins fused to the engineered variants of the PFB and SBP tags) via covalent bond formation while, at the same time, the streptavidin-ligand complexes retain the reversible binding capability. When immobilization is needed, this system functions as a superglue to allow immobilization to take place. When reversible interactions (e.g. affinity purification of proteins) are desirable, this system acts like a molecular velcro which operates in a reversible manner (i.e. in the presence of reducing agent and biotin).

To achieve this objective, structure-guided design methods are used to generate streptavidin muteins based on the SAVSBPM18 mutein with reversible biotin binding (Wu & Wong, 2013). Cysteine residues were placed in strategic positions in the streptavidin muteins (SAVSBPM32 and SAVSBPM19) and also in both the SBP (SBPA18C) and PFB (CPFB) tags. In this system, the cysteine containing SBP or PFB tag and the engineered streptavidin muteins operate in a two-step manner. In the first step, the SBP tag and the biotin moiety in the biotinylated PFB tag offer affinity guided interactions between the tags and streptavidin so that docking can take place. In the second step, the strategically placed cysteine residues in both the tags and streptavidin allows locking (thiol-coupling) to take place. This system offers advantages of both immobilization and reversible binding capability in the same system depending on the operating conditions. To test this system, the cysteine containing SBP and PFB tags were then fused to the C-terminal end of the *E. coli* TEM-1 β-lactamase (BLA) reporter, which
was used as the model protein throughout this work for the following reasons (Figure 1.6). First, a large body of literature is available on the use of BLA and its expression in both prokaryotic and eukaryotic cells (Remy, Ghaddar, & Michnick, 2007). Additionally, BLA has been shown to be exceedingly versatile and retain enzyme activity when used as a reporter for secreted, intracellular or membrane bound proteins (Moore, Davis, & Dev, 1997). Furthermore, the hydrolysis of Nitrocefin by BLA allows the activity to be monitored by a simple and sensitive colourimetric assay (Remy et al., 2007). Apart from the above-mentioned advantages, the enzyme has another characteristic that is important in this study. In BLA, residues 77 and 123 form a disulfide bond (Figure 1.6), which contributes to the thermostability of BLA (Shimizu-Ibuka, Matsuzawa, & Sakai, 2006; Vanhove et al., 1997). Since reducing agents are commonly used throughout this work, BLA can serve as a model system to evaluate whether a disulfide bond containing target protein can be compatible with the use of the cysteine containing SBP or PFB tag in purification and immobilization processes.

With the two streptavidin muteins developed, one of them (SAVSBPM32) can also be applied to purify biotinylated molecules or proteins containing the modified or unmodified SBP tag in a reusable manner.
Figure 1.6 Structure of the *E. coli* TEM-1 β-lactamase (BLA) reporter protein

Structure of the model protein BLA used throughout this work. The disulfide bond between residues C77 and C123 of BLA is shown in yellow. The C-terminus of BLA is colored in red to indicate the placement of the linkers and binding tags used throughout this work. The model was generated using PyMOL with the PDB entry 1BTL (Jelsch, Mourey, Masson, & Samama, 1993) as a starting file.
This project aims to enhance one of the most commonly used protein reagents in biotechnology by increasing its binding affinity (via disulfide bond formation), flexibility, and also by dramatically reducing the costs associated with its use. The potential implications are far reaching and include an array of applications such as the development of reusable biosensor chips, bioreactors, protein arrays, matrices for affinity purification of proteins, and diagnostic kits with superior sensitivity. Surpassing these direct applications, this system will further serve as a valuable building block in the development of increasingly advanced and sophisticated biological materials in the field of synthetic biology.

In this study, the application of SAVSBPM32 for the affinity purification of SBPA18C tagged BLA was demonstrated. Its superior performance illustrates the advantage of this system.
Chapter 2  

Materials and Methods

2.1 Modeling of Streptavidin peptide complexes

Unless otherwise specified, models of the streptavidin peptide complexes and their respective figures were generated with the MacPyMOL: PyMOL molecular graphics system, v1.7.0.3 (Schrödinger, LLC, 2010). Starting files used for modeling were either previously published Protein Data Bank (PDB) files, obtained from RCSB Protein Data Bank (www.rcsb.org), or unpublished PyMol Session (pse) files, obtained from the designing author with permission. Summary of the models generated and starting files used is illustrated in Table 2.1.

2.1.1 Modeling SAVSBPM18-SBP tag complex

The script M18SBP (Appendix 1) generated the reported model of the SAVSBPM18-SBP tag complex, using the unpublished PDB file (r9a.1) of the SAVSBPM18-SBP tag complex (Wu & Wong, 2013) as the starting file. This starting file was obtained from Dr. S.-L. Wong, whom designed the model using the Molegro molecular viewer, version 2.5.0 (Thomsen & Christensen, 2006) software and the published crystal structure of the streptavidin-SBP tag complex (PDB reference, 4jo6) (Barrette-Ng et al., 2013). Briefly, the script loaded the PDB file (r9a.1) into PyMol, then colored the four streptavidin subunits and two SBP tags as chains A (red), B (blue), C (green), D (purple), Y (cyan), and Z (grey), respectively. The resulting model was saved as the pse file M1S.1.
Table 2.1 Modeling of streptavidin peptide complexes.

<table>
<thead>
<tr>
<th>Model generated</th>
<th>Starting file(s)</th>
<th>Script used</th>
<th>Resulting file</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAVSBPM18-SBP</td>
<td>r9a.1</td>
<td>M18SBP</td>
<td>M1S.1</td>
</tr>
<tr>
<td>SAVSBPM32-SBPA18C</td>
<td>r9a_submit_A86_A18C</td>
<td>M32A18C</td>
<td>M2S.1</td>
</tr>
<tr>
<td>SAVSBPM19-CPFB(-2)</td>
<td>New6pept</td>
<td>M19(-2)</td>
<td>M3P.1</td>
</tr>
<tr>
<td>SAVSBPM32-CPFB(-2)</td>
<td>M2S.1, M3P.1</td>
<td>M32(-2)</td>
<td>M2P.1</td>
</tr>
<tr>
<td>SAVSBPM96-SBPA18C</td>
<td>M2S.1</td>
<td>M96A18C</td>
<td>M9A.1</td>
</tr>
<tr>
<td>SAVSBPM96-CPFB(-2)</td>
<td>M3P.1</td>
<td>M96(-2)</td>
<td>M9P.1</td>
</tr>
<tr>
<td>SAVSBPM80-Biotin</td>
<td>3ry1, 3ry2</td>
<td>M80B</td>
<td>M80B.1</td>
</tr>
</tbody>
</table>
2.1.2 *Modeling SAVSBPM32-SBPA18C tag complex*

The script M32A18C (Appendix 1) generated the reported model of the SAVSPM32-SBPA18C tag complex, using the unpublished PDB file (r9a_submit_A86C_A18C) of the SAVSBPM32-SBPA18C tag complex (designed by and obtained from Dr. S.-L. Wong) as the starting file. Briefly, the script first loaded the PDB file (r9a_submit_A86C_A18C) into PyMol, then colored and labeled the chains as described in section 2.1.1. Following those actions the streptavidin-tag disulfide bond (C86 and C18, respectively) was characterized by bond length (Sγ-Sγ) and angle (Cβ-Sγ-Sγ). The script then mutated streptavidin residues S27 and G48 to A27 and T48, respectfully. The resulting model was saved as the pse file M2S.1.

2.1.3 *Modeling SAVSBPM19-CPFB(-2) tag complex*

The script M19(-2) (Appendix 1) generated the reported model of the SAVSBPM19-CPFB(-2) tag complex, using unpublished PDB file (new6pept), generated by Dr. Kenneth Ng and obtained from Dr. S.-L. Wong as the starting file. This file contains a biotinylated lysine nonapeptide (Figure 2.1) bound to tetrameric streptavidin. Briefly, the script first loaded the file into PyMol, then labeled and colored streptavidin as described in section 2.1.1. The biotinylated nonapeptide labeled chain E was then colored cyan. Residues N118, S27 and G48 of streptavidin were mutated to Cys, Ala, and Thr, respectively. The residue -2 of the nonapeptide was then mutated to Cys and the torsion angles adjusted (Figure 2.1).
**Figure 2.1. Biotinylated nonapeptides.**

Sequences, residue numbers, and designated labels of the biotinylated nonapeptide and cysteine containing variants (A). + or - sign in the designated residue label indicates position of the residue from the biotinylated lysine residue (designated as 0). (B and C) residue -2 of the modeled nonapeptide with dihedral angles phi (ϕ), psi (ψ), and omega (Ω) defined by the atoms C-N-Cα-C, N-Cα-C-N and Cα-C-N-Cα, respectively. Models in panels B and C were generated using PyMOL with the unpublished PDB file new6pept (designed by Dr. Kenneth Ng and obtained from Dr. S.-L. Wong with permission) as a starting file.
Following these actions the streptavidin-tag disulfide bond (C118 and C-2, respectively) was generated and characterized by bond length (Sγ-Sγ) and angle (Cβ-Sγ-Sγ). The resulting model was saved as the pse file M3P.1.pse

2.1.4 Modeling SAVSBPM32-CPFB(-2) tag complex

The script M32(-2) (Appendix 1) generated the reported model of SAVSBPM32-CPFB(-2) tag complex, using the modeled SAVSBPM32-SBPA18C tag complex (M2S.1) and SAVSBPM19-CPFB(-2) tag complex (M3P.1) (sections 2.1.2 and 2.1.3, respectively) as starting files. Briefly, the script first extracted both the tetrameric SAVSBPM32 structure from the M2S.1 file and the biotinylated nonapeptide bound to SAVSBPM19 subunit B from the M3P.1 file. The SAVSBPM19 subunit B was then aligned with subunit B of the tetrameric SAVSBPM32 and subsequently deleted. The remaining structures were merged as one object, which contained the biotinylated nonapeptide bound to subunit B of tetrameric SAVSBPM32. Torsion angles in the biotinylated lysine side chain and backbone of the nonapeptide were then varied (Table 2.2) to generate a reasonable model for intermolecular disulfide bond formation between C86 of SAVSBPM32 and C-2 of the biotinylated nonapeptide tag. The disulfide bond was then generated and characterized by bond length (Sγ-Sγ) and angle (Cβ-Sγ-Sγ). The resulting model was saved as the pse file M2P.1.pse.
Table 2.2. Variations of torsion angles within the biotinylated nonapeptide from various streptavidin-peptide tag modeled complexes.

<table>
<thead>
<tr>
<th>Residue</th>
<th>Dihedral angle</th>
<th>Atoms involved</th>
<th>Angle in M19-P* complex</th>
<th>Angle in M32-P* complex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysine</td>
<td>$\chi_3$</td>
<td>$C_\varepsilon-C_\delta-C_\gamma-C_\beta$</td>
<td>-178</td>
<td>-160</td>
</tr>
<tr>
<td>Lysine</td>
<td>$\chi_2$</td>
<td>$C_\delta-C_\gamma-C_\beta-C_\alpha$</td>
<td>169</td>
<td>163</td>
</tr>
<tr>
<td>Lysine</td>
<td>$\chi_1$</td>
<td>$C_\gamma-C_\beta-C_\alpha-N$</td>
<td>55.2</td>
<td>-118</td>
</tr>
<tr>
<td>Lysine</td>
<td>$\phi$</td>
<td>$C_\beta-C_\alpha-N-C$</td>
<td>-148</td>
<td>119</td>
</tr>
<tr>
<td>Lysine</td>
<td>$\omega$</td>
<td>$C_\alpha-N-C-C_\alpha$</td>
<td>83.8</td>
<td>80.0</td>
</tr>
<tr>
<td>Cysteine</td>
<td>$\psi$</td>
<td>$N-C-C_\alpha-N$</td>
<td>-174</td>
<td>-117</td>
</tr>
</tbody>
</table>

*M19-P and M32-P designate the modeled SAVSBPM19-CPFB(-2) tag and SAVSBPM32-CPFB(-2) tag complexes, respectively.
2.1.5 **Modeling SAVSBPM96-SBPA18C tag complex**

The script M96A18C (Appendix 1) generated the reported model of SAVSBPM96-SBPA18C tag complex, using the pse file (M2S.1) of the modeled SAVSBPM32-SBPA18C tag complex (section 2.1.2) as the starting file. Briefly, the script first loaded the M2S.1 file to PyMol, and then mutated residues C86 and A63 of each SAVSBPM32 subunit to A86 and C63, respectively. The resulting file was saved as the pse file M9A.1.

2.1.6 **Modeling SAVSBPM96-CPFB(-2) tag complex**

The script M96(-2) (Appendix 1) generated the reported model of SAVSBPM96-CPFB(-2) tag complex, using the pse file (M3P.1) of the modeled SAVSBPM19-CPFB(-2) tag complex (section 2.1.3) as the starting file. Briefly, the script first loaded the M3P.1 file to PyMol, and then mutated residues C118 and A63 of each SAVSBPM19 subunit to N118 and C63, respectively. The resulting file was saved as the pse file M9P.1.

2.1.7 **Modeling SAVSBPM80-biotin complex**

The script M80B (Appendix 1) generated the reported model of SAVSBPM80-biotin complex, using the PDB files (3ry1 and 3ry2) of the crystallized structures of apo-streptavidin and streptavidin-biotin complexes, respectively, (Le Trong et al., 2011) as starting files. Briefly, the script first retrieved the PDB files from www.rcsb.org and
loaded them into the PyMol program (internet connection is required for this step of the
script). Once loaded, the structure of 3ry2 (containing two streptavidin subunits with
bound biotin) was modeled as a tetramer by duplicating the object and aligning both
copies to the apo-streptavidin (3ry1). Residues N49, A86, and W120 of both the apo and
biotin bound streptavidin subunits were then mutated to C49, C86, and A120,
respectively. The N-C-Cα-Cβ and C-Cα-Cβ-γ dihedral angles of the C49 residues in
biotin bound streptavidin was then set to -165 and -94.5, respectively. Following this, the
intramolecular Sγ-Sγ bond was generated between the C49 and C86 residues of the biotin
bound streptavidin subunits and the Sγ-Sγ bond distance and Cβ-Sγ-Sγ angle was
calculated. The resulting file was saved as the pse file M80B.1.

2.2 General procedures

2.2.1 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS PAGE)

SDS-PAGE was performed with the mini-PROTEAN Tetra electrophoresis
system (BioRad) according to the Laemmli method (Laemmli, 1970). Unless specified
otherwise, the gel was constructed in an 8 cm x 7.3 cm gel cassette with a 1.0 mm spacer
and contained a 12% polyacrylamide resolving gel [375 mM Tris-HCl pH 8.8, 12%
acrylamide, 0.4% bis-acrylamide, 0.05% ammonium persulfate (APS), 0.067%
tetramethylethylenediamine (TEMED), 0.1% Sodium dodecyl sulfate (SDS)] and a 4%
polyacrylamide stacking gel (124 mM Tris-HCl pH 6.8, 4% acrylamide, 0.133% bis-
acrylamide, 0.05% APS, 0.099% TEMED, 0.1% SDS). Each sample was prepared by
mixing 12 μl of the sample with 4 μl 5X loading buffer (250 mM Tris-HCl pH 6.8, 10%
SDS, 30% glycerol, 5% β-mercaptoethanol, 0.02% bromophenol blue). Each sample was then subjected to a 5-minute incubation in a boiling water bath prior to loading into the gel. Electrophoresis was done at 4°C for 1 hour under a constant ampere of 30 milliamperes in SDS running buffer (0.3% Tris, 1.44% glycine, 0.1% SDS, pH 8.3).

Following electrophoresis proteins were simultaneously fixed to gel and stained with a solution containing 5 parts methanol, 1 part acetic acid and 4 parts H₂O (v/v) with 2.5% Coomassie brilliant blue R250 at 23°C for 6-16 hours with agitation (De St. Groth, Webster, & Datyner, 1963; Meyer & Lamberts, 1965). The gel was then destained with a solution containing 1 part methanol, 1 part acetic acid and 8 parts H₂O (v/v) at 23°C for 6-16 hours with agitation (Cleveland, Fischer, Kirschner, & Laemmli, 1977). The resulting gel was then photographed and digitalized using the BioRad Gel Doc 2000 system and software (Quantity One 4.0.3; BioRad). Apparent molecular mass of proteins was estimated based on electrophoretic mobilities (Shapiro, Viñuela, & V. Maizel Jr., 1967; K. Weber & Osborn, 1969) compared to the molecular markers. For SDS-PAGE the low-molecular weight standard proteins (GE Healthcare) were used as markers, which includes phosphorylase b (97,412) (Titani et al., 1977), albumin (66,430) (Hirayama, Akashi, Furuya, & Fukuhara, 1990), ovalbumin (45,000) (Yao, Dey, Pastor, & Wilkins, 1995), carbonic anhydrase (29,000) (Yao et al., 1995), trypsin inhibitor (20,100) (Koide & Ikenaka, 1973) and α-lactalbumin (14,437) (Brew, Vanaman, & Hill, 1967).
2.3 Bacterial strains

2.3.1 Escherichia coli

*Escherichia coli* BL21(DE3) (Table 2.3), the recombinant protein T7 Expression strain (Jeong et al., 2009; Studier & Moffatt, 1986) was used throughout this study as the bacterial expression host for MBP-L-SBPA18CG19A recombinant protein.

*Escherichia coli* DH5α (Table 2.4), the highly efficient transformation and vector propagation strain (Casali, 2003; Hanahan, Jessee, & Bloom, 1991) was used throughout this study as the bacterial host for recombinant DNA cloning.

2.3.2 Bacillus subtilis

*Bacillus subtilis* WB800 (Table 2.5), the engineered 8-protease deficient strain (Wu, Yeung, et al., 2002) was used throughout this study as the bacterial expression host for SAVSBPM18, SAVSBPM32, SAVSBPM32F, SAVSBPM19, SAVSBPM96, BLA-L-SBP, BLA-L-SBPA18C, and BLA-L-CPFB(-2) proteins.

2.4 Plasmid vectors

The *B. subtilis* vector pSSAV (Wu, Qureshi, et al., 2002) (Appendix 2) was used as the parental bacterial expression plasmid for construction of specific expression vectors for producing the SAVSBPM18 (Wu & Wong, 2013), SAVSBPM32, SAVSBPM32F, SAVSBPM96, and SAVSBPM19 proteins.
Table 2.3. Properties of *E. coli* BL21(DE3) host strain  
(Berlyn, 1998; Casali, 2003; Jeong et al., 2009; Studier & Moffatt, 1986)

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Description</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>ompT</em></td>
<td>Inactivates outer membrane protease</td>
<td>Reduces protein degradation</td>
</tr>
<tr>
<td><em>gal</em></td>
<td>Inactivates galactose metabolism</td>
<td>Inhibits use of galactose as a carbon source</td>
</tr>
<tr>
<td><em>dcm</em></td>
<td>Inactivates DNA cytosine methylase</td>
<td>Increases efficiency of cloning</td>
</tr>
<tr>
<td>Δ<em>hsdS</em></td>
<td>Inactivates Eco site recognition</td>
<td>Increases efficiency of cloning</td>
</tr>
<tr>
<td>(DE3)</td>
<td>Insertion of the prophage DE3 containing lavUV5 promoter</td>
<td>Inducible expression of targeted genes</td>
</tr>
<tr>
<td></td>
<td>and T7 RNA polymerase</td>
<td></td>
</tr>
</tbody>
</table>
Table 2.4. Properties of *E. coli* DH5α host strain
(Berlyn, 1998; Casali, 2003; Peters, Thate, & Craig, 2003)

<table>
<thead>
<tr>
<th>Relevant genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textbf{deOR endA1 gyrA96 hsdR17 Δ(lac)U169 recA1 relA1 supE44 thi-1 Φ80lacZΔM15}</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Description</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textbf{deOR}</td>
<td>Inactivates \textit{deo} operon regulatory gene</td>
<td>Improves large plasmids replication</td>
</tr>
<tr>
<td>\textbf{endA1}</td>
<td>Inactivates nonspecific endonuclease activity</td>
<td>Increases quality of plasmid DNA</td>
</tr>
<tr>
<td>\textbf{gyrA96}</td>
<td>Point mutation in DNA Gyrase</td>
<td>Introduces resistance to nalidixic acid</td>
</tr>
<tr>
<td>\textbf{hsdR17}</td>
<td>Inactivates Eco endonuclease</td>
<td>Increases cloning efficiency</td>
</tr>
<tr>
<td>\textbf{Δ(lac)U169}</td>
<td>Deletes lactose genes (y kd-bo350)</td>
<td>Blocks lactose utilization</td>
</tr>
<tr>
<td>\textbf{recA1}</td>
<td>Inactivates homologous recombination recombinant</td>
<td>Increases cloning efficiency</td>
</tr>
<tr>
<td>\textbf{relA1}</td>
<td>Inactivates stringent response</td>
<td>Increases RNA synthesis</td>
</tr>
<tr>
<td>\textbf{supE44}</td>
<td>Inactivates supE44 tRNAs</td>
<td>Suppresses amber mutations</td>
</tr>
<tr>
<td>\textbf{thi-1}</td>
<td>Inactivates hydroxyethylthiazole synthesis</td>
<td>Requires thiamine for growth</td>
</tr>
<tr>
<td>\textbf{Φ80lacZΔM15}</td>
<td>Insertion the prophage Φ80 containing a partially deleted β-Galactosidase gene</td>
<td>Allows selection of lacZ cells containing recombinant plasmids</td>
</tr>
</tbody>
</table>
Table 2.5 Properties of *B. subtilis* WB800 host strain
(Kodama et al., 2012; Wu, Yeung, et al., 2002; X.-C. Wu, Lee, Tran, & Wong, 1991)

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Description</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>nprE</em></td>
<td>Inactivates extracellular neutral protease E</td>
<td>Decreases degradation of secreted proteins</td>
</tr>
<tr>
<td><em>aprE</em></td>
<td>Inactivates extracellular protease subtilisin</td>
<td>Decreases degradation of secreted proteins</td>
</tr>
<tr>
<td><em>epr</em></td>
<td>Inactivates minor extracellular serine protease epr</td>
<td>Decreases degradation of secreted proteins</td>
</tr>
<tr>
<td><em>bpr</em></td>
<td>Inactivates extracellular protease bacillopeptidase</td>
<td>Decreases degradation of secreted proteins</td>
</tr>
<tr>
<td><em>mpr::ble</em></td>
<td>Bleomycin marker insertion inactivates extracellular protease metalloprotease</td>
<td>Decreases degradation of secreted proteins, introduces bleomycin resistance</td>
</tr>
<tr>
<td><em>nprB::bsr</em></td>
<td>Blasticidin S marker insertion inactivates extracellular neutral protease</td>
<td>Decreases degradation of secreted proteins, introduces blasticidin S resistance</td>
</tr>
<tr>
<td><em>Δvpr</em></td>
<td>Deletion of minor extracellular protease vpr</td>
<td>Decreases degradation of secreted proteins</td>
</tr>
<tr>
<td><em>wprA::hyg</em></td>
<td>Hygromycin marker insertion inactivates cell wall bound protease wprA</td>
<td>Decreases degradation of secreted proteins, introduces hygromycin resistance</td>
</tr>
</tbody>
</table>
The pSSAV plasmid was previously engineered based on the pUB110 expression vector (Gryczan, Contente, & Dubnau, 1978; McKenzie, Hoshino, Tanaka, & Sueoka, 1986) and contains a P43 promoter for constitutive expression (Wang & Doi, 1984), a levansucrase (SacB) signal peptide for protein secretion (S.-L. Wong, 1989) (Steinmetz, Le Coq, Aymerich, Gonzy-Tréboul, & Gay, 1985) and a kanamycin resistance marker for selection (Sadaie, Burtis, & Doi, 1980).

The *B. subtilis* vector pWB980 (Wu & Wong, 1999) (Appendix 2) was used as the parental bacterial expression plasmid for construction of specific expression vectors for producing the BLA-L-SBP (Barrette-Ng et al., 2013; Chen et al., 2008), BLA-L-SBPA18C and BLA-L-CPFB(-2) proteins. The pWB980 plasmid contains the same properties as the pSSAV vector, except the orientation of the P43 promoter, SacB peptide and multiple cloning site are reversed (Wu & Wong, 1999).

The pET19b bacterial vector was used as the inducible expression plasmid for construction of specific expression vectors for producing the His-tagged MBP-L-SBPA18CG19A protein. The pET19b vector contains a polylinker with a N-terminal His tag sequence for purification of recombinant proteins by immobilized metal ion affinity chromatography (IMAC) (Belew & Porath, 1990; Hemdan & Porath, 1985), a T7 promoter and terminator for highly specific transcription and transcription termination by T7 RNA polymerase (Rosenberg et al., 1987), a *lac* operator and *lacI* gene encoding the repressor protein for inducible production of recombinant proteins (Busby & Ebright, 1999; Gilbert & Maxam, 1973).
2.5 Vector construction and transformation

2.5.1 SAVSBPM18

Plasmid pSSAVSBPM18 (Appendix 2), which includes a synthetic gene encoding the streptavidin mutein SAVSBPM18 was previously constructed and transformed into *Bacillus subtilis* WB800 cells generating WB800[pSSAVSBPM18] by Dr. S.-C. Wu (Wu & Wong, 2013).

2.5.2 SAVSBPM19

Plasmid pSSAVSBPM19 (Appendix 2), which includes the DNA fragment encoding the streptavidin mutein SAVSBPM19 was constructed and transformed into *Bacillus subtilis* WB800 cells generating WB800[pSSAVSBPM19] by Dr. S.-C. Wu.

2.5.3 SAVSBPM32

Plasmid pSSAVSBPM32 (Appendix 2), which includes the DNA fragment encoding the streptavidin mutein SAVSBPM32 was constructed and transformed into *Bacillus subtilis* WB800 cells generating WB800[pSSAVSBPM32] by Dr. S.-C. Wu.

2.5.4 SAVSBPM32F

Plasmid pSSAVSBPM32F (Appendix 2), which includes the DNA fragment encoding the streptavidin mutein SAVSBPM32F was constructed and transformed into *Bacillus subtilis* WB800 cells generating WB800[pSSAVSBPM32F] by Dr. S.-C. Wu.
2.5.5 SAVSBPM96

Plasmid pSSAVSBPM96 (Appendix 2), which includes the DNA fragment encoding the streptavidin mutein SAVSBPM96 was constructed and transformed into *Bacillus subtilis* WB800 cells generating WB800[pSSAVSBPM96] by Dr. S.-C. Wu.

2.5.6 SAVSBPM80

Plasmid pSSAVSBPM80 (Appendix 2), which includes the DNA fragment encoding the streptavidin mutein SAVSBPM80 was constructed and transformed into *Bacillus subtilis* WB800 cells generating WB800[pSSAVSBPM80] by Dr. S.-C. Wu.

2.5.7 BLA-L-SBP

Plasmid pWB980-BLA-L-SBP (Appendix 2), which includes the DNA fragment encoding the fusion protein BLA-L-SBP was previously constructed and transformed into *Bacillus subtilis* WB800 cells generating WB800[pWB980-BLA-L-SBP] by Dr. S.-C. Wu (Wu & Wong, 2013).

2.5.8 BLA-L-SBPA18C

Plasmid pWB980-BLA-L-SBPA18C (Appendix 2), which includes the DNA fragment encoding the fusion protein BLA-L-SBPA18C was constructed and transformed
into *Bacillus subtilis* WB800 cells generating WB800[pWB980-BLA-L-SBPA18C] by Dr. S.-C. Wu.

2.5.9 **MBP-L-SBPA18CG19A**

Plasmid pET19b-MBP-L-SBPA18CG19A (Appendix 2), which includes the DNA fragment encoding the fusion protein MBP-L-SBPA18CG19A was constructed and transformed into *Escherichia coli* BL21(D3) cells generating BL21(D3)[pET19b-MBP-L-SBPA18CG19A] by Dr. S.-C. Wu.

2.5.10 **BLA-L-CPFB(-2)**

Plasmid pWB980-BLA-L-CPFB(-2) (Appendix 2), which includes the DNA fragment encoding the fusion protein BLA-L-CPFB(-2) was constructed and transformed into *Bacillus subtilis* WB800 cells generating WB800[pWB980-BLA-L-CPFB(-2)] by Dr. S.-C. Wu.

2.6 **Cell preparation for long-term storage of bacteria**

2.6.1 **Bacillus subtilis**

*B. subtilis* cells carrying the described plasmids (Section 2.5) was streaked onto Tryptose Blood Agar Base plates (Casman, 1942) (30 g L\(^{-1}\) TBAB, Difco\textsuperscript{TM}) containing 10 mg L\(^{-1}\) kanamycin (Wu & Wong, 1999) and incubated at 30°C for 16 hours. Cells
from the resulting bacterial colonies were transferred into 1 ml Luria-Bertani (LB) broth (10 g L⁻¹ tryptone, 5 g L⁻¹ yeast extract, 10 g L⁻¹ NaCl, pH 7.0) (Bertani, 1951; Sezonov, Joseleau-Petit, & D’Ari, 2007) containing 50% glycerol (Calcott, Calvert, & Draper, 1978; Hubálek, 2003). The mixture was then vortexed and the resulting high cell density homogenous stock solution was stored at -80°C throughout this study.

2.6.2 *Escherichia coli*

Generation of bacterial stocks containing *E. coli* cells carrying the described plasmids (Section 2.5) was done using the same methods described for the preparation of long-term storage stocks of *B. subtilis* (Section 2.6.1) with modifications. The cells transferred to the 50% glycerol LB broth solution were obtained from an LB agar streak plate (15 g L⁻¹ agar) containing appropriate antibiotics (75 mg L⁻¹ ampicillin) which was incubated at 37°C for a 16 hours.

2.7 Cell culture conditions for protein production

2.7.1 *Bacillus subtilis*

Production of proteins from *B. subtilis* cells was done by streaking cells from a frozen LB-glycerol stock (Hanahan et al., 1991) onto a TBAB plate (Difco™) containing 10 mg L⁻¹ kanamycin, and incubating the plate for 16 hours at 30°C. A single colony from the resulting plate was then cultured for 16 hours with 250 rpm agitation at 30°C in 200 ml super-rich medium (25 g L⁻¹ tryptose, 20 g L⁻¹ yeast extracts, 3 g L⁻¹
dipotassium hydrogen orthophosphate) (Halling, Sanchez-Anzaldo, Fukuda, Doi, & Meares, 1977) containing 10 µg ml\(^{-1}\) kanamycin.

### 2.7.2 *Escherichia coli*

Production of proteins from *E. coli* cells was done by streaking cells from frozen LB-glycerol stock (Hanahan et al., 1991) onto LB plates containing 75 mg L\(^{-1}\) ampicillin, then incubating the plate for 16 hours at 37°C. A single colony from the resulting plate was then cultured for 16 hours with 250 rpm agitation at 37°C in 200 ml LB medium containing 75 mg L\(^{-1}\) of ampicillin. A subculture was then generated by inoculating 100 ml of fresh LB medium with 2 ml of the previously cultured cells and incubating the new culture at 37°C for 3 hours with 250 rpm agitation. Production of recombinant protein was then induced by the addition of 0.2 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) into the culture (Marbach & Bettenbrock, 2012) and the induced culture was further incubated for 4 hours with 250 rpm agitation at 25°C.

### 2.8 Concentration of proteins

#### 2.8.1 *Bacillus subtilis*

*B. subtilis* cells were removed from the culture medium by centrifugation at 10,000g for 15 minutes and decanting the supernatant (Wu, Qureshi, et al., 2002). Secreted proteins within the supernatant were concentrated by either direct use of Amicon ultra-15 centrifugal filters (10,000 MWCO, Millipore) (Wu & Wong, 2013), or
by protein precipitation with ammonium sulfate to 75% saturation at 4°C for 16 hours (Qureshi & Wong, 2002).

2.8.2 *Escherichia coli*

Harvesting of *E. coli* cells in the culture medium was done by centrifugation at 10,000 g for 15 minutes and discarding the supernatant. The cell pellet was then resuspended in 2 ml of physiological buffered saline (PBS, 0.1 M sodium phosphate, 0.15 M sodium chloride, pH 7.5) and lysed in two steps. In the first step, cells were enzymatically digested by lysozyme (50 µg ml⁻¹) (Andrews & Asenjo, 1987) during a 10 minute incubation at 37°C. Cells in the mixture were then further lysed in the second step by sonication (Doula, 1977; Harrison, 1991). Sonication of each sample was done using 5-7 cycles of short 10 second sonication burst followed by 1 minute cooling on ice (Shrestha, Michaele Holland, & Charles Bundy, 2012). The lysed cell mixture was then centrifuged at 10,000 g for 15 minutes and the supernatant containing soluble proteins was collected.

2.9 Protein purification

Unless otherwise specified, prior to protein purification, supernatant samples were dialyzed in physiological buffered saline buffer (PBS, 0.1 M sodium phosphate, 0.15 M sodium chloride, pH 7.5) using Spectra/Por® 3 Dialysis Membrane Tubing, MWCO: 3,500 (product number: 132 720) purchased from Spectrum Laboratories. The term “PBS” in subsequent text refers to the above mentioned buffer unless specified otherwise.
2.9.1 SAVSBPM18 purification by affinity chromatography

SAVSBPM18 was purified using a modified version of the biotin-agarose gravity flow affinity chromatography methods previously reported for the purification of various streptavidin muteins including SAVSBM18 (Chilkoti, Schwartz, Smith, Long, & Stayton, 1995; O’Sullivan et al., 2012; Wu et al., 2009; Wu & Wong, 2005, 2013). Briefly, tris(2-carboxyethyl)phosphine (TCEP, Sigma) was added to 1 ml WB800[pSSAVSBPM18] culture supernatant at a final concentration of 2 mM and incubated for 30 minutes at 23°C. The sample was then loaded into a biotin-agarose (Sigma, Canada) column, which had previously been equilibrated with 4 column volumes PBS buffer. Once the supernatant had fully entered the column the ends of the column were capped and the mixture was incubated for 30 minutes at 23°C to maximize binding. Following the incubation, unbound proteins were removed by washing the column with 4 column volumes of PBS buffer containing 2 mM TCEP. Bound proteins were then eluted with 6 column volumes of PBS buffer containing 5 mM d-biotin (Sigma). Analysis of the collected samples was done by SDS-PAGE (Section 2.2) and fractions identified to contain purified SAVSBPM18 were pooled together and concentrated on Amicon ultra-15 centrifugal filters (10,000 MWCO, Millipore). Excess biotin in the purified samples was then removed by dialysis in PBS buffer. The same approach was used to purify SAVSBPM96, SAVSBPM19, and SAVSBPM32 using WB800[pSSAVSBPM96], WB800[pSSAVSBPM19], and WB800[pSSAVSBPM32] culture supernatants, respectively.
2.9.2 SAVSBPM32F purification by affinity chromatography

SAVSBPM32F was purified by affinity chromatography using the same methods described for the purification of SAVSBPM18 (Section 2.9.1), with two modifications. WB800[pSSAVSBPM32] culture supernatant was used and 1 mM of the serine protease inhibitor phenylmethylsulfonyl fluoride (PMSF) was added to the culture supernatant and purified protein sample. The addition of PMSF limited proteolytic degradation of the FLAG tag (Deutscher, 1990).

2.9.3 BLA-L-SBP and BLA-L-SBPA18C purification by Ion exchange chromatography

BLA-L-SBP produced and concentrated from WB800[pWB980-BLA-L-SBP] as previously described (Sections 2.7.1 and 2.8.1) was buffer exchanged to low salt Tris-buffered saline [TBS, 15 mM tris(hydroxymethyl)aminomethane, 15 mM sodium chloride, pH 7.5] by dialysis. Based on the theoretical isoelectric point (pI) of BLA-L-SBP (pI 5.37) (Gasteiger et al., 2005), BLA-L-SBP at pH 7.5 is negatively charged and will therefore bind to DEAE Sepharose, a weak anion exchange chromatography platform (GE Healthcare). A flow rate of 3 ml min$^{-1}$ was used to load 10 ml of dialyzed supernatant into an 18 ml bed volume DEAE-Sepharose Fast Flow column (26 mm column diameter), which had previously been equilibrated with 4 column volumes of low salt TBS buffer. Unbound proteins were then removed by washing the column with 4 column volumes of low salt TBS buffer. Bound proteins were then eluted using a 120 ml
linear NaCl gradient (15-330 mM) in TBS buffer at a flow rate of 4 ml minute⁻¹.
Throughout the purification samples were collected in 4.5 ml fractions and proteins were analyzed by SDS-PAGE (Section 2.2.1). Fractions containing BLA-L-SBP at >95% purity were pooled together, concentrated on Amicon ultra-15 centrifugal filters (10,000 MWCO, Millipore) and dialyzed in PBS buffer. The same approach was used to purify BLA-SBPA18C from the culture supernatant of WB800[pWB980-BLA-L-SBPA18C].

2.9.4 MBP-L-SBPA18CG19A purification by affinity chromatography

MBP-L-SBPA18CG19A produced and concentrated from BL21(D3)[pET19b-MBP-L-SBPA18CG19A] as described in (Sections 2.7.2 and 2.8.2) was buffer exchanged to H-S equilibration buffer (50 mM sodium phosphate, 300 mM sodium chloride, 10 mM imidazole, pH 8.0) by centrifugation using Amicon ultra-15 centrifugal filters (10,000 MWCO, Millipore) and purified on His-Select Nickel HC Affinity Gel (Sigma) column according to manufacturers instructions with modifications. A gravity flow column was packed with 1 ml bed volume of the affinity gel, which was then washed in two column volumes H₂O, followed by 5 column volumes 6 M guanidine HCL (pH 7.5), and an additional 2 column volumes of H₂O. Nickel ions were then stripped from the gel with 5 column volumes of 0.1 M (Ethylenedinitrilo)tetaacetic acid (EDTA) (EMD Chemicals) pH 7.5 (Porath & Olin, 1983). The column was then washed with 2 column volumes of H₂O and fresh nickel ions were immobilized to the column by passing 2 column volumes of 10 mg ml⁻¹ nickel sulfate (ICN Biochemicals) over the
column. The recharged column was then washed with 2 column volumes of H$_2$O and equilibrated in 3 column volumes of H-S equilibration buffer.

Crude protein extract containing MBP-L-SBPA18CG19A was clarified by passing 4 ml of the sample through a cellulose acetate membrane syringe filter (0.45 µm pore size) (SARSTEDT), which was then reduced by the addition of 2.5 mM Dithiothreitol (DTT), and loaded to the column. Unbound and non-specifically bound protein was removed by washing the column with 13 column volumes H-S equilibration buffer (Gaberc-Porekar & Menart, 2001). The remaining protein bound to the column was eluted in 6 column volumes of the same buffer containing 250 mM imidazole (Porath, 1992). Fractions containing MBP-L-SBPA18CG19A at ≥95% purity were identified by SDS-PAGE (Section 2.2.1), pooled together, concentrated and buffer exchanged to PBS on Amicon ultra-15 centrifugal filters (10,000 MWCO, Millipore).

2.9.5 **BLA-L-CPFB(-2) purification**

2.9.5.1 **BLA-L-CPFB(-2) purification by Ion exchange chromatography**

BLA-L-CPFB(-2) was partially purified using the same methods described for the purification of BLA-L-SBP (Section 2.9.3), except that WB800[pWB980-BLA-L-CPFB(-2)] culture supernatant was used. SDS-PAGE was then used to identify fractions containing partially purified proteins (Section 2.2.1), which were then pooled together, concentrated on Amicon ultra-15 centrifugal filters (10,000 MWCO, Millipore) and dialyzed in low salt SA buffer (17 mM succinic acid, 15 mM sodium chloride, pH 4.25). Based on the theoretical isoelectric point (pI) of BLA-L-CPFB(-2) (pI 5.74), (Gasteiger
et al., 2005) BLA-L-CPFB(-2) at pH 4.25 is positively charged and therefore will bind to the CM Sepharose Fast Flow weak cation exchange chromatography platform (GE Healthcare). A flow rate of 3 ml minute\(^{-1}\) was used to load 10 ml of dialyzed partially pure protein into an 18 ml bed volume CM Sepharose Fast Flow column (26 mm column diameter), which had previously been equilibrated with 4 column volumes low salt SA buffer. Unbound proteins were then removed by washing the column with 4 column volumes low salt SA buffer. Bound proteins were then eluted using a 120 ml linear NaCl gradient (15-330 mM) in SA buffer (4 ml minute\(^{-1}\) flow rate). Throughout the purification samples were collected in 4.5 ml fractions and analyzed by SDS-PAGE (Section 2.2.1). Fractions containing BLA-L-CPFB(-2) protein at >95% purity were pooled together, concentrated on Amicon ultra-15 centrifugal filters (10,000 MWCO, Millipore) and dialyzed in PBS buffer.

### 2.9.5.2 BLA-L-CPFB(-2) biotinylation

BLA-L-CPFB(-2) was biotinylated \textit{in vitro} using the engineered \textit{E. coli} biotin ligase (CBD-BirA-His) and reported methods for Staphylokinase–PFB biotinylation (Wu, Yeung, et al., 2002) with modifications. Briefly, unbiotinylated BLA-L-CPFB(-2) produced and concentrated from WB800[pWB980-BLA-L-CPFB(-2)] (Sections 2.7.1 and 2.8.1) was first purified using the ion exchange chromatography methods previously described (Sections 2.9.3 and 2.9.5.1) and dialyzed in 10 mM TBS, (pH 8.0). Biotinylation was done at 30°C for 16 hours in 1 ml reactions containing 500 µg BLA-L-CPFB(-2), 6 µg purified \textit{E. coli} CBD-BirA-His, 50 mM bicine pH 8.3, 10 mM adenosine...
5'-triphosphate (ATP), 10 mM magnesium acetate, and 50 µM d-biotin. Subsequently CBD-BirA-His was removed from the mixture by batch purification using chitin affinity chromatography (Watanabe et al., 1994; Wu, Yeung, et al., 2002). This was done by adding 200 µl chitin beads (New England BioLabs), previously washed with bicine pH 8.3 to the sample and incubating the mixture at 23°C for 20 minutes. Chitin beads and bound CBD-BirA-His was then removed by centrifuging the mixture at 15,000 g for 5 minutes and decanting the supernatant. Excess biotin was then removed from supernatant by dialysis in PBS.

2.9.5.3 Biotinylated BLA-L-CPFB(-2) purification by affinity chromatography

Biotinylated BLA-L-CPFB(-2) was separated from unbiotinylated proteins by loaded the sample into a 1 ml bed volume Monomeric Avidin Agarose (Pierce, Rockford, IL) column, which had previously been equilibrated with 4 column volumes PBS (pH 7.0), 3 column volumes elution buffer (2 mM d-biotin, PBS, pH 7.0), 6 column volumes regeneration buffer (0.1 M glycine, pH 2.8) and equilibrated with 2 column volumes PBS (Green & Toms, 1973; Henrikson, Allen, & Maloy, 1979). Binding of biotinylated BLA-L-CPFB(-2) to the column was maximized by capping the column and incubating the sample for 1 hour at 23°C. Unbound protein was then removed by washing the column with 6 column volumes PBS. Bound biotinylated protein was then eluted with 6 column volumes elution buffer (Henrikson et al., 1979; Kohanski & Daniel Lane, 1990; Wu, Yeung, et al., 2002). Fractions containing purified biotinylated BLA-L-CPFB(-2) was identified by SDS-PAGE (Section 2.2.1), pooled together, concentrated on Amicon
ultra-15 centrifugal filters (10,000 MWCO, Millipore) and buffer exchanged to PBS by dialysis.

2.10 Quantification of protein

2.10.1 Semi-quantitative measurements of crude proteins

Concentrations of BLA-L-SBP, BLA-L-SBPA18C and BLA-L-CFPB(-2) in crude supernatant samples was semi-quantified by standard curves generated on SDS-PAGE. Varying amounts of crude supernatant and known concentrations of BSA (New England BioLabs) were analyzed by SDS-PAGE following the procedure previously described (Section 2.2.1) and resulting band intensities of each samples were analyzed using the ImageJ software (Abramoff, Magalhães, & Ram, 2004; Schneider, Rasband, & Eliceiri, 2012). Band intensities of the known BSA concentrations were used to generate the standard curve, which then allowed for semi-quantitative measurements of recombinant protein concentration in non-purified samples.

2.10.2 Quantitative measurements of purified proteins

Quantification of purified proteins was done using spectrophotometric methods (Layne, 1957). The UV absorption of the samples was measured at 280 nm on a NanoDrop 1000 spectrophotometer (Thermo Scientific) and used in combination with the calculated extinction coefficients to determine concentrations (Gasteiger et al., 2005; Gill & von Hippel, 1989).
2.11 Protein binding characterization

2.11.1 Disulfide bond formation

2.11.1.1 Reducing and non-reducing SDS-PAGE electrophoresis

The ability for the studied proteins to form intermolecular disulfide bonds was tested using reducing and non-reducing SDS-PAGE gels (Wu et al., 2009). Briefly, three types of reactions were prepared in PBS buffer and designated reaction types A, B, and C. Reaction types A contained 1 µM of the tetrameric streptavidin muteins. Reaction types B contained 4 µM of SBP or CPFB tagged recombinant proteins and reaction types C types contained both 1 µM of the streptavidin variant tetramer and 4 µM of the ligand containing recombinant protein. Reactions were then reduced in the presence of Immobilized TCEP (Pierce, Rockford, IL), prepared prior to use by centrifuging a slurry volume which was equal to the sample volume for 1 minute at 1,000 g and discarding the storage buffer supernatant. Reactions were reduced for 1 hour at 23°C with occasional agitation then separated from the reductant by centrifuging the mixture at 1000 g for 1 minute and removing the supernatant. Disulfide bonds were then allowed to form throughout a final 1 hour incubation at 23°C. Following the final incubation, samples from each reaction were analyzed on SDS-PAGE using the methods described in (Section 2.2.1). Additionally, samples from each reaction were also analyzed on non-reducing SDS-PAGE using the same procedures described in Section 2.2.1, except that no reducing agent was included in the sample loading buffer.
2.11.1.2 Carboxymethylation of cysteine residues

Carboxymethylation of reactions in the absence of a reducing agent was used to modify only cysteine residues not currently involved in disulfide bonds (Hansen & Winther, 2009). This was done using a modified Aitken and Learmonth method (Aitken & Learmonth, 2002). Briefly, 150 µl of 2 µM SAVSBPM32 tetramer and 150 µl of 4 µM BLA-L-SBPA18C was reduced separately in the presence of immobilized TCEP using the previously described methods (Section 2.11.1.1). Following reduction, aliquots containing 10 µl of each reduced protein were mixed into 14 tubes and incubated at 23°C. 5 µl of 5X carboxymethylation buffer (250 mM iodoacetamide, 250 mM Tris-HCl, pH 8.0) was then added to individual tubes after 0, 2, 4, 6, 8, 20, 15, 20, 30, 40, 50, 60, 90 and 120 minutes had passed from the time SAVSBPM32 and BLA-L-SBPA18C were initial mixed. After the addition of the carboxymethylation buffer, the mixture was incubated at 23°C for 15 minutes and then stored on ice. Once all the reactions were finished, the samples were then analyzed SDS-PAGE using the reducing and non-reducing methods previously described (Section 2.11.1.1).

2.11.2 Analysis of protein complex formation

2.11.2.1 Native polyacrylamide gel electrophoresis

Samples containing 5 µM of the tested streptavidin variants and various concentrations of the ligand fusion proteins (0-20 µM) were prepared in PBS buffer
containing 1 mM TCEP. Native PAGE was performed using the same methods described for SDS-PAGE (Section 2.2.1) with the following modifications. The resolving gel contained 18% acrylamide and 0.6% bis acrylamide. SDS was not included in the resolving gel, stacking gel, loading buffer or running buffer. Samples were also not subject to incubation in boiling water bath and electrophoresis was done at 4°C for 21 hours under a constant current of 17 milliamperes.

2.11.2.2 2.11.2.2 Semi-native polyacrylamide gel electrophoresis

Samples containing 5 μM of SAVSBPM32F tetramer and various concentrations of BLA-L-SBPA18C (0 - 150 μM) were prepared using methods described in (Section 2.11.1.1). Semi-native polyacrylamide gel electrophoresis (s-native PAGE) was then performed using the same methods as described for non-reducing SDS-PAGE (Section 2.2.1) with modifications. The resolving gel contained 8% acrylamide and 0.267% bis acrylamide. SDS was not included in the resolving gel, stacking gel and loading buffer (0.1% SDS was included in the running buffer). Samples were not subjected to incubation in boiling water bath and electrophoresis was done at 4°C for 1 hour under a constant current of 30 milliamperes.
2.11.3 Complex composition

2.11.3.1 Modified two dimensional polyacrylamide gel electrophoresis

The separation of protein complexes and their components in distinct dimensions was done using a modified two dimensional polyacrylamide gel electrophoresis (M2D-PAGE). This procedure was done in two steps loosely based on O’Farrell methods (O’Farrell, 1975) with major modifications. Briefly, reduction and bond formation of a sample containing 5 µM tetrameric SAVSBPM32F and 15 µM BLA-L-SBPA18C were done using previously described methods (Section 2.11.1.1). The mixture was then loaded to an s-native PAGE gel and electrophoresis was carried out accordingly (Section 2.12.2.2). Following electrophoresis, proteins were reduced within the gel by first removing the gel from the cassette and incubating the free gel in 50 ml running buffer containing 10% β-mercaptoethanol for 10 minutes at 23°C with slight agitation. The gel was then rinsed with H₂O and washed in running buffer for 1 minute at 23°C with slight agitation. Following the reduction and wash steps, the gel was then inserted back into the gel cassette at a 90° angle compared to the first dimension and electrophoresis was done for 30 minutes at 4°C with a constant current of 35 milliamps. Subsequent staining, destaining and digitalization of the resulting gel was then done according to SDS-PAGE methods (Section 2.2.1).
2.11.3.2 Zinc imidazole reverse staining, complex separation, extraction and analysis

A sample containing 5 µM SAVSBPM32F and 20 µM BLA-L-SBPA18C was prepared and separated on s-native PAGE (Section 2.12.2.2). Following electrophoresis, proteins were visualized using the zinc imidazole reverse staining method (Fernandez-Patron et al., 1998). Briefly, the gel was first soaked for 10-minutes with agitation in 30 mM imidazole, 60 mM zinc, pH 5.3, followed by a 30-second rinse in H₂O and a 5-minute incubation with agitation in 50 mM Tris-HCl, pH 8.8. Distinct protein bands were excised using a scalpel and the proteins within each gel slices were extracted using passive diffusion methods (Hardy et al., 1996; Hardy, Pupo, Santana, Guerra, & Castellanos-Serra, 1998) with slight modification. Zinc ions were first removed from the excised bands by washing the gel pieces twice in PBS containing 100 mM EDTA (EMD Chemicals). EDTA was then removed by washing the gel slices twice in PBS. Gel slices were then crushed into small pieces, which were then soaked in 200 µl PBS. The gel PBS mixture was then incubated at 23°C for 2 hours under constant vortexing, followed by a 16 hour incubation at 30°C without agitation. The mixture was then centrifuged for 15 minutes at 10,000 g and the supernatant containing extracted protein was collected. Samples were then concentrated using a centrifugal evaporator and analyzed using the reducing and non-reducing SDS-PAGE procedure previously described (Section 2.11.1.1).
2.11.4 Binding kinetics

Each binding kinetic experiment was done using the Blitz™ label free protein assay system (ForteBio) with the Amine Reactive 2nd Generation (AR2G) biosensors (ForteBio). Biosensors were hydrated in H₂O for 10 minutes prior to use, unless otherwise specified, each step was performed at 23°C with a shake rate of 2,200 rpm in a 0.5 ml microcentrifuge tube containing 300 µl of sample. The BLITz Pro software (version 1.1.0.25) (ForteBio) advanced kinetics module was used for programming experimental steps and data acquisition.

2.11.4.1 Preparation of biosensors

2.11.4.1.1 AR2G-BLA-L-SBPA18C biosensor construction

Generation of each AR2G-BLA-L-SBPA18C biosensor was done individually by coupling BLA-L-SBPA18C to hydrated biosensors as follows. The biosensor was first placed in PBS buffer for 60 seconds, equilibrated in acetate buffer (10 mM acetate, pH 5.0) for 300 seconds and activated with 1-Ethyl-3-[3-dimethylaminopropyl] carbodiimide hydrochloride (EDC) and N-hydroxysulfosuccinimide (s-NHS) (Staros, Wright, & Swingle, 1986) buffer (20 mM EDC, 10 mM s-NHS, 10 mM acetate, pH 5) for 300 seconds. Purified BLA-L-SBPA18C was then coupled to the sensors at 40 µg ml⁻¹ for 600 seconds in acetate buffer. Quenching of the remaining uncoupled active sites was done in 1 M ethanolamine (Frederix et al., 2004) at pH 8.5 for 300 seconds and the resulting AR2G-BLA-L-SBPA18C biosensor was then washed in PBS for 60 seconds. The same approach was applied to generate AR2G-BLA-L-SBP and AR2G-BSA(B)
sensor chips using purified BLA-L-SBP and Biotinylated BSA (BSA(B)), (BioVision) with one exception. For the preparation the AR2G-BSA(B) sensor chips, 80 µg ml\(^{-1}\) of BSA(B) was used during the coupling step.

2.11.4.2 **Association and dissociation measurements**

2.11.4.2.1 **Association and dissociation of SAVSBPM18 to AR2G-BLA-L-SBPA18C biosensor**

The initial baseline for binding was measured by placing the biosensor in PBS buffer containing 0.005% Surfactant P20 (GE Healthcare), and 1 mM TCEP, for 60 seconds. Association and dissociation measurements of SAVSBPM18 were obtained by placing the biosensor in 157 nM purified SAVSBPM18 in the previously mentioned TCEP, P20, PBS buffer for 400 seconds followed by the same buffer alone for 800 seconds. The same procedure was then done using 117.8, 78.5, 39.25 and 0 nM SAVSBPM18 with newly generated AR2G-BLA-L-SBPA18C biosensors.

2.11.4.2.2 **Association and dissociation of SAVSBPM32 to AR2G-BLA-L-SBP biosensors**

The binding kinetics of SAVSBPM32 to AR2G-BLA-L-SBP biosensors was done using the same procedures as the binding kinetics of SAVSBPM18 to AR2G-BLA-L-SBPA18C (Section 2.11.4.2.1), except that AR2G-BLA-L-SBP biosensors and purified SAVSBPM32 were used.
2.11.4.2.3  Association and dissociation of SAVSBPM32 to AR2G-BSA(B) biosensors

The binding kinetics of SAVSBPM32 to AR2G-BSA(B) biosensors was done using the same procedures as the binding kinetics of SAVSBPM18 to AR2G-BLA-L-SBPA18C (Section 2.11.4.2.1), except that AR2G-BSA(B) biosensors and purified SAVSBPM32 were used.

2.12  Applications: Affinity and affinity thiol coupling chromatography

2.12.1  Preparation of affinity chromatography columns

2.12.1.1  Preparation of SAVSBPM18 affinity chromatography column

Purified SAVSBPM18 was immobilized to Affi-gel 15 activated affinity media (BioRad) at 1.5 mg protein per ml gel by coupling methods formally reported for SAVSBPM18 (Wu & Wong, 2013). Briefly, Affi-gel 15 was first washed with 3 bed volumes of cooled (4°C) deionized water by applying a vacuum to a Buchner funnel containing Whatman™ qualitative filter paper (7cm diameter, 8µm pore size) (GE Healthcare). Spontaneous coupling of the ligand to the matrix was then achieved by mixing 1.5 mg protein in PBS buffer (2.4 ml of 625 µg ml⁻¹ SAVSBPM18 cooled to 4°C) with 1 ml of the washed matrix slurry and incubating the mixture for 16 hours at 4°C (Cuatrecasas & Parikh, 1972). Following coupling incubation a small aliquot of the protein solution was collected and used to semi-quantitatively calculate coupling efficiency (Section 2.12.2.1). Blocking of the remaining active esters in the affinity media was done by adding 0.1 ml of 1 M ethanolamine-HCL (pH 8) to the affinity gel.
(Frederix et al, 2004) and incubating the mixture for 1 hour at 23°C with agitation. The mixture was then transferred to an empty gravity flow chromatography column, washed with 5 ml S-Elution buffer (PBS, 2 mM TCEP, 5 mM d-Biotin) and equilibrated with 5 ml PBS buffer.

2.12.1.2 Preparation of SAVSBPM32 affinity chromatography column

Purified SAVSBPM32 was immobilized to Affi-gel 15 activated affinity media (BioRad) using the same methods described for the coupling of SAVSBPM18 (Section 2.12.1.1) with a slight modification. Coupling of SAVSBM32 was done at concentrations of 2 mg protein per ml gel using 0.6 ml Affi-gel 15 activated affinity media and 1.14 ml of 1090 µg ml⁻¹ SAVSBPM32 cooled to 4°C.

2.12.2 Coupling efficiency of variants to Affi-gel 15 media

2.12.2.1 Coupling efficiency of SAVSBPM18 to Affi-gel 15 media

Coupling efficiency of SAVSBPM18 to Affi-gel 15 was calculated using the standard curve method described in (Section 2.10.2) with modification. The known concentration of pre-coupled SAVSBPM18 was used to generate the standard curve and calculate Post-coupled protein concentrations. The difference between the post and pre-coupled SAVSBPM18 concentrations indicated coupling efficiency and was used to estimate the total amount of protein coupled to the affi-gel 15 activated affinity media.
2.12.2.2 Coupling efficiency of SAVSBPM32 to Affi-gel 15 media

Coupling efficiency of SAVSBPM32 to Affi-gel 15 was calculated using the same method described for calculating coupling efficiency of SAVSBPM18 to Affi-gel 15 (Section 2.12.2.1), except pre and post-coupled SAVSBPM32 was used.

2.12.3 Purification of BLA-L-SBP by SAVSBPM18 affinity column chromatography

Purification of BLA-L-SBP by SAVSBPM18 affinity chromatography was done for both overloaded (150% the theoretical binding capacity of the column) and non-overloaded (20% the theoretical binding capacity of the column) conditions using a modified version of previously published methods [Sau-Ching Wu & Wong, 2013]. Briefly, The amount (µg) of BLA-L-SBPA18C loaded to the column was based on the theoretical binding capacity of the matrix. This was estimated on the assumption that one BLA-L-SBPA18C molecule binds across two subunits in the SAVSBPM32 tetramer. 1 µg of SAVSBPM18 dimer (Mr = 33,037.8) will bind 1.062 µg of BLA-L-SBP (Mr = 35,101.7). Crude BLA-L-SBP was dialyzed in PBS and the concentration was semi-quantified (Section 2.10.1). Samples were then reduced and separated from immobilized TCEP as described in previously (Section 2.11.1.1). For the non-overloaded purification condition an aliquot containing 142 µg of reduced BLA-L-SBP was loaded into the 1 ml SAVSBPM18 column. Binding was maximized by capping the column and incubating the mixture for 1 hour at 23°C. Following the incubation, unbound protein was removed by washing the column with 6 column volumes PBS. Bound protein was then eluted in 6
column volumes S-Elution buffer. The first column volume of S-Elution buffer was incubated in the column for 30 minutes prior to elution. The column was regenerated with 10 column volumes PBS (Wu & Wong, 2013). Fractions containing purified BLA-L-SBPA18C were identified by SDS-PAGE, pooled together, concentrated on Amicon ultra-15 centrifugal filters (10,000 MWCO, Millipore) and dialyzed in PBS buffer. The same procedure was used for the overloaded condition with the exception that 1,065 µg of reduced BLA-L-SBP was loaded into the regenerated SAVSBPM18 column.

2.12.4 Purification of BLA-L-SBP by SAVSBPM32 affinity column chromatography

Purification of BLA-L-SBP by SAVSBPM32 affinity chromatography for both overloaded and non-overloaded conditions was done using the same methods described for the purification of BLA-L-SBP by SAVSBPM18 affinity column chromatography (Section 2.12.3), except that 176 and 1,320 µg of reduced BLA-L-SBP supernatant was loaded into a 0.6 ml SAVSBPM32 column. The amount (µg) of BLA-L-SBP loaded into the column was based on the theoretical binding capacity of the matrix. This was estimated on the assumption that one BLA-L-SBP molecule binds across two subunits in the SAVSBPM32 tetramer. 1 µg of SAVSBPM32 dimer (Mr = 33,102) will bind 1.06 µg of BLA-L-SBP (Mr = 35,101.7).
2.12.5 Purification of BLA-L-SBPA18C by SAVSBPM18 affinity column chromatography

Purification of BLA-L-SBPA18C by SAVSBPM18 affinity chromatography for both overloaded and non-overloaded conditions was done using the same methods described for the purification of BLA-L-SBP by SAVSBPM18 affinity column chromatography (Section 2.12.3), except that 142 and 1,065 µg of reduced BLA-L-SBPA18C supernatant was loaded into the regenerated SAVSBPM18 column. The amount (µg) of BLA-L-SBPA18C loaded to the column was based on the theoretical binding capacity of the matrix. This was estimated on the assumption that one BLA-L-SBPA18C molecule binds across two subunits in the SAVSBPM18 tetramer. 1 µg of SAVSBPM18 dimer (Mr = 33,037.8) will bind 1.063 µg of BLA-L-SBPA18C (Mr = 35,113.8).

2.12.6 Purification of BLA-L-SBPA18C by SAVSBPM32 affinity thiol coupling column chromatography

Purification of BLA-L-SBPA18C by SAVSBPM32 affinity thiol coupling chromatography for both overloaded and non-overloaded conditions was done using the same methods described for the purification of BLA-L-SBP by SAVSBPM18 affinity column chromatography (Section 2.12.3), except that 176 and 1320 µg of reduced BLA-L-SBPA18C supernatant was loaded into the regenerated SAVSBPM32 column. The amount (µg) of BLA-L-SBPA18C loaded to the column was based on the theoretical binding capacity of the matrix. This was estimated on the assumption that one BLA-L-
SBPA18C molecule binds across two subunits in the SAVSBPM32 tetramer. 1 µg of SAVSBPM32 dimer (Mr = 33,102) will bind 1.061 µg of BLA-L-SBPA18C (Mr = 35,113.8).

2.12.7 Purification of BLA-L-SBPA18C by SAVSBPM32 affinity thiol coupling column chromatography under various wash conditions

2.12.7.1 Purification of BLA-L-SBPA18C by SAVSBPM32 affinity thiol coupling column chromatography under nonionic detergent wash conditions

The purification procedure was done using the overloaded condition and procedure described in Section 2.12.3, with one modification. The wash buffer used contained 2% Tween-20 in PBS.

2.12.7.2 Purification of BLA-L-SBPA18C by SAVSBPM32 affinity thiol coupling column chromatography under high ionic strength wash conditions

The purification procedure was done using the overloaded condition and procedure described in Section 2.12.3.4, with one modification. The wash buffer used contained 300 mM KCl.
2.12.7.3 *Purification of BLA-L-SBPA18C by SAVSBPM32 affinity thiol coupling column chromatography under competitive binding wash conditions*

The purification procedure was done using the overloaded condition and procedure described in Section 2.12.3.4, with one modification. The wash buffer used contained 5 mM d-biotin.

2.12.7.4 *Purification of BLA-L-SBPA18C by SAVSBPM32 affinity thiol coupling column chromatography under high ionic strength and competitive binding wash conditions*

The purification procedure was done using the overloaded condition and procedure described in Section 2.12.3.4 with two modifications. First, the wash buffer used contained both 5 mM d-biotin and 300 mM KCl. Additionally, once the first column volume of wash buffer entered the column was capped and incubated at 23°C for 10 minutes prior to collection of the wash fractions.

2.12.8 *Purification of MBP-L-SBPA18CG19A by SAVSBPM32 affinity thiol coupling column chromatography*

MBP-L-SBPA18CG19A previously purified by HIS-Select Nickel HC Affinity chromatography (Section 2.9.8) was buffer exchanged to PBS by dialysis and re-purified by SAVSBPM32 affinity thiol coupling column chromatography using modified methods.
described for the purification of BLA-L-SBPA18C by SAVSBPM32 affinity thiol coupling column chromatography (Section 2.12.3.4). Briefly, 2 mM TCEP (Sigma) was first added to a 500 µl aliquot of the buffer exchanged sample, which was then loaded onto the regenerated SAVSBPM32 affinity column. Binding was then maximized by capping the column followed by a 1 hour incubation at 23°C. Following the incubation TCEP and unbound protein was removed by washing the column with 7 column volumes PBS. Sufficient time was then allowed for the formation of intermolecular disulfide bonds between the protein and column by capping and incubating the column at 23°C for 2.5 hours. Non-covalently bound proteins were then eluted in 7 column volumes PBS containing 5 mM d-biotin. The column was then incubated at 23°C for 30 minutes in the presence of PBS containing 5 mM d-biotin and 2 mM TCEP and then eluted using 7 ml of the same buffer. Fractions containing pure MBP-L-SBPA18CG19A were then identified by SDS-PAGE (Section 2.2.1).
Chapter 3 A cysteine containing streptavidin and SBP tag variant: SAVSBPM32 and SBPA18C

3.1 Rational design of the streptavidin SAVSBPM32-SBPA18C system

3.1.1 Cysteine containing streptavidin variant SAVSBPM32 and the cysteine containing SBP tag variant SBPA18C

It has been observed that disulfide bond formation in proteins favours a C$_\beta$-S$_\gamma$-S$_\gamma$ bond angle and S$_\gamma$-S$_\gamma$ bond length of 104° and 2.04 Å, respectively (Alan A. Dombkowski, Sultana, & Craig, 2014; Petersen, Jonson, & Petersen, 1999). To locate positions where cysteine replacement would satisfy these constraints, the “Disulfide by Design” program (A. A. Dombkowski, 2003) was used to probe the recently crystallized structure of the SBP-Tag-streptavidin complex (Barrette-Ng et al., 2013) (Table 3.1). Due to the steric and geometric constraints required for disulfide bond formation (Figure 3.1, panel A), it is not surprising that the suggested residues of the SBP tag for mutagenesis are those residues already involved in binding interactions with SAV (Barrette-Ng et al., 2013). In order to limit the loss of binding interactions caused by the mutations, the ideal candidates in the SBP tag should have less interactions with streptavidin. In this program, the calculated energy reflects the potential for the formation of disulfide bond with the lowest one to have the highest possibility to form a disulfide bond. Ala-86 of SAV and Ala-18 of the SBP tag were chosen for mutagenesis (Figure 3.1) as they are the best candidates for three reasons. First, this pair has the highest probability to form a disulfide bond (Table 3.1).
Table 3.1 Location of cysteine replacements predicted by Disulfide by Design to form intermolecular disulfide bonds (A. A. Dombkowski, 2003).

<table>
<thead>
<tr>
<th>SAV residue</th>
<th>SBP residue</th>
<th>Bond</th>
<th>X3*</th>
<th>Energy**</th>
<th>Σ B-Factors***</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chain: A.A</td>
<td>Chain: A.A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A: ALA86</td>
<td>Y: ALA18</td>
<td>-107.68</td>
<td>1.49</td>
<td></td>
<td>64.19</td>
</tr>
<tr>
<td>B: ALA86</td>
<td>Z: ALA18</td>
<td>-109.49</td>
<td>1.77</td>
<td></td>
<td>67.64</td>
</tr>
<tr>
<td>B: GLN24</td>
<td>Z: GLY11</td>
<td>113.79</td>
<td>3.67</td>
<td></td>
<td>85.72</td>
</tr>
<tr>
<td>A: SER45</td>
<td>Y: ALA14</td>
<td>83.49</td>
<td>5.19</td>
<td></td>
<td>78.98</td>
</tr>
<tr>
<td>B: SER45</td>
<td>Z: GU15</td>
<td>115.9</td>
<td>6.44</td>
<td></td>
<td>89.61</td>
</tr>
<tr>
<td>B: SER45</td>
<td>Z: VAL14</td>
<td>85.4</td>
<td>7.77</td>
<td></td>
<td>83.98</td>
</tr>
</tbody>
</table>

* Estimated X3 torsion angle

** Calculated bond energy

*** Sum of the temperature factors associated with the two residues
Figure 3.1 Engineering of the streptavidin variants SAVSBPM32, SAVSBPM96 and the SBPA18C peptide tag.

(A) The geometric and steric requirements of disulfide bond formation. (B) Model of the SBP tag bound to SAVSBPM18 with residues for cysteine replacement (A86 and A18) shown in black. (C) Model of the covalently linked SBPA18C and SAVSBPM32 complex with cysteine residues shown in black and the disulfide bond shown in yellow. (D) Cysteine containing streptavidin control SAVSBPM96. Mutations resulting in the reversible binding characteristics of SAVSBPM18 (G48T and S27A) are labeled and depicted in pale green and brown respectively (panel B, C and D). Models were generated using PyMOL with PDB entry 4J06 (Barrette-Ng et al., 2013) as the starting file.
Second, A18 in SBP does not interact extensively with many other residues (Table 3.2) in streptavidin (Barrette-Ng et al., 2013). Furthermore, the interaction between A18 in SBP with A86 in SAV is suggested to be relatively weak using the ligand energy inspector module in the Molegro Viewer program which predicts the strength of molecular interactions (Thomsen & Christensen, 2006). Third, A86 does not play a vital role in interacting with biotin. The A86C mutation is unlikely to weaken the biotin binding affinity in the mutated streptavidin. This feature is important because biotin should still be an effective competitor to elute the SBP tagged proteins off from the streptavidin mutein matrices. To examine the potential for disulfide bond formation, A18C and A86C were introduced to the SBP tag and SAVSBPM18, respectively, to develop the SBPA18C tag and SAVSBPM32 (Tables 3.3 and 3.4). The SBPA18C tag was fused to the C-terminal end of the *E. coli* TEM-1 β-lactamase (BLA) reporter. To ensure the projection of the SBPA18C tag away from BLA, a 19-amino-acid linker (IDPAGTSPSTPEGPSTPSN) was introduced between BLA and the tag. Finally, an additional Ala residue was included at the C-terminal end of the SBP tag. The original SBP tag ends with a proline residue. In *E. coli*, proteins with proline at their C-terminals are known to be unstable because of the marking of these proteins for degradation by the addition of the ssrA tag (Hayes, Bose, & Sauer, 2002). Although it is not sure whether the same feature is applicable to *B. subtilis*, it is safer to end a protein with alanine than proline. The other option is to delete the C-terminal proline. However, it is advantageous to have a proline located near the C-terminus which may minimize the sequential degradation of proteins from the C-terminus by carboxypeptidase.
Table 3.2 Interactions of the SBP tag residues predicted to form intermolecular disulfide bonds to streptavidin (Barrette-Ng et al., 2013).

<table>
<thead>
<tr>
<th>SBP residue</th>
<th>Streptavidin interactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>V14</td>
<td>S27, Y43, E44, S45, N49, S52, W120</td>
</tr>
<tr>
<td>A18</td>
<td>A86</td>
</tr>
<tr>
<td>G11</td>
<td>N23, Q24, L25, W120</td>
</tr>
<tr>
<td>E15</td>
<td>S45, V47, R84</td>
</tr>
<tr>
<td>Streptavidin variant</td>
<td>Mutations</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>SAVSBPM18</td>
<td>G48T, S27A</td>
</tr>
<tr>
<td>SAVSBPM32</td>
<td>G48T, S27A, A86C</td>
</tr>
<tr>
<td>SAVSBPM32F</td>
<td>G48T, S27A, A86C, DYKDDDDK insertion</td>
</tr>
<tr>
<td>SAVSBPM96</td>
<td>G48T, S27A, A63C</td>
</tr>
</tbody>
</table>
Table 3.4 Sequences of the ligand tags and variants.

<table>
<thead>
<tr>
<th>Ligand tag</th>
<th>Amino acid sequence*</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBP</td>
<td>MDEKTGWRGGHVVEGLAGELEQLRLEHHPQGQREP</td>
</tr>
<tr>
<td>SBPA18C</td>
<td>MDEKTGWRGGHVVEGLCGELEQLRLEHHPQGQREP**A</td>
</tr>
<tr>
<td>PFB</td>
<td>LHHILDAQKMWVNHRE</td>
</tr>
<tr>
<td>CPFB(-2)</td>
<td>LHHILDCQKMWVNHRE</td>
</tr>
</tbody>
</table>

* Bolded and underlined letters indicate mutated residues

---

83
The expected molecular masses of BLA-L-SBPA18C and SAVSBPM32 are ≈ 35 kDa and 16.5 kDa/subunit, respectively. A variation of the SAVSBPM32 mutein termed SAVSBPM32F was also engineered, which included a 9-amino-acid insertion consisting of a serine residue followed by a FLAG tag (DYKDDDDK) between residues Glu-14 and Ala-15 of SAVSBPM32 (Table 3.3). This version of streptavidin mutein offers better separation of streptavidin from SBPA18C tagged BLA and other complexes during the analysis of the covalently bounded complex formation via native gel electrophoresis.

### 3.1.2 Design of negative controls for cysteine containing streptavidin and ligand tag

Disulfide bond formation can occur spontaneously if the two cysteine containing proteins are present at high concentrations under the oxidizing environment. These two cysteine containing proteins do not require affinity to interact with each other. This is known as the regular thiol coupling process. In contrast, in this study, disulfide bond formation between SAVSBPM32 and BLA-L-SBPA18C is an affinity-guided thiol coupling process directed by SBP. The efficiency for disulfide bond formation should be much higher even at low concentrations of SAVSBPM32 and BLA-L-SBPA18C. In order to study and confirm that the formation of intermolecular disulfide bonds is a result of the SBP guided affinity thiol coupling instead of the regular thiol coupling, the following controls were generated. For the streptavidin control, a surface exposed residue, A63, which is located far away from the SBP binding pocket of SAVSBPM18 was replaced by cysteine and the resulting mutein is termed SAVSBPM96 (Figure 3.1, panel D). Non-biotinylated BLA-L-CPFB(-2) was used as the ligand control, as it has an
accessible cysteine residue and no binding affinity to streptavidin (Wu et al., 2009) (Table 3.4).

3.2 Protein production and purification

SAVSBPM18, SAVSBPM96, SAVSBPM32 and SAVSBPM32F were produced by secretion from Bacillus subtilis WB800 cells containing their respective plasmids (Appendix 2) with a production yield of ~10 mg L^{-1} and affinity purified on biotin agarose (section 2.9). The majority of SAV variant proteins were eluted in the first 1-3 fractions (~ 90% recovery) with no other contaminants visible by SDS-PAGE with Coomassie blue staining (Figure 3.2). The apparent molecular mass of the variants was estimated to be approximately 17 kDa based on their electrophoretic mobilities (Shapiro et al., 1967; K. Weber & Osborn, 1969) in reference to the molecular markers. This is in close agreement with the expected molecular mass of 16.5 kDa, with the exception of SAVSBPM32F which has an expected molecular mass of 17.6 kDa. Elution fractions of SAVSBPM18, SAVSBPM96, SAVSBPM32 and SAVSBPM32F were then collected, dialyzed, and quantified using methods described in section 2.10.

BLA-L-SBP and BLA-L-SBPA18C were produced by secretion from Bacillus subtilis WB800 cells that contained their respective plasmids (Appendix 2) with a production yield of ~10 mg L^{-1} and purified by ion exchange chromatography using DEAE Sepharose Fast Flow (GE Healthcare) (section 2.9.6).
Figure 3.2 Purification of SAVSBPM18, SAVSBPM32, SAVSBPM96 and SAVSBPM32F.

Secreted SAVSBPM18 (A), SAVSBPM32 (B), SAVSBPM96 (C) and SAVSBPM32F (D) produced from WB800[pSAVSBPM18], WB800[pSAVSBPM32], WB800[pSAVSBPM96] and WB800[pSAVSBPM32F] respectively were affinity purified using biotin-agarose. Fractions were collected and analyzed by SDS-PAGE with Coomassie blue staining. M: molecular weight markers. Numbers shown on the left represent the molecular mass expressed in kDa; S: culture supernatant; FT: flow-through fraction; W1-W4: wash fractions; E1–E6: elution fractions. Arrowhead marks the position of the SAV muteins.
The majority of the recombinant proteins were eluted over 10 fractions. However, most of those fractions were unusable due to the presence of other proteins which were visible by SDS-PAGE stained with Coomassie blue (Figure 3.3). The apparent molecular mass of the recombinant proteins was estimated to be approximately 36 kDa, which is in close agreement with the expected molecular mass of 35 kDa. Fractions 12-14 (Figure 3.3, panel A) and 11-12 (Figure 3.3, panel B) collected during the purification of BLA-L-SBP and BLA-L-SBPA18C were estimated to have greater than 95% purity. These fractions were collected, dialyzed and quantified using methods described in section 2.10.

As a model for the cysteine containing ligand tag CPFB(-2), a 18-amino-acid linker (PAGTSPSTSGGSTGSGS) was inserted between BLA and the CPFB(-2) ligand tag. The resulting protein termed BLA-L-CPFB(-2) (Wu et al., 2009) was produced by secretion from *Bacillus subtilis* WB800 cells containing the appropriate plasmid (Appendix 2) and semi-purified by ion exchange chromatography using a DEAE Sepharose Fast Flow (GE Healthcare) column following methods described in section 2.9.3. Fractions were then analyzed on SDS-PAGE with Coomassie blue staining. Fractions 6-11 (Figure 3.4, panel A) containing partially purified protein were collected and further purified by the CM Sepharose Fast Flow weak cation exchange chromatography (GE Healthcare). The apparent molecular mass of the recombinant proteins was estimated to be approximately 33 kDa, which is in close agreement with the expected molecular mass of 32.5 kDa. Fractions 4-7 were estimated to have greater than 95% purity (Figure 3.4, panel B). They were collected, dialyzed, and quantified using methods described in section 2.10.
Figure 3.3 Purification of BLA-SBP and BLA-SBPA18C.

Secreted BLA-L-SBP (A) and BLA-L-SBPA18C (B) produced from WB800(pBLA-SBP) and WB800(pBLA-SBPA18C), respectively. Culture supernatant was purified by DEAE Sepharose fast-flow chromatography using a salt gradient. Fractions were collected and analyzed by SDS-PAGE with Coomassie blue staining. M: molecular weight markers. Numbers shown on the left and right represent the molecular mass expressed in kDa; S: culture supernatant; FT: flow-through; F₁-F₂₆: fractions. Arrowhead marks the position of BLA-L-SBP and BLA-L-SBPA18C.
Figure 3.4 Purification of BLA-CPFB.

Secreted BLA-CPFB produced from WB800(pBLA-CPFB). (A) Culture supernatant was purified by DEAE Sepharose fast-flow chromatography using a salt gradient. Fractions were collected and analyzed by SDS-PAGE with Coomassie blue staining. (B) Semi-purified fractions containing BLA-CPFB were purified by CM Sepharose fast-flow chromatography using a salt gradient. M: molecular weight markers. Numbers shown on the left represent the molecular mass expressed in kDa; S: culture supernatant; DF: collected DEAE fractions; FT: flow-through; F_1-F_{12}: fractions. Arrowheads mark the position of BLA-CPFB.
It should be noted that throughout this work the purified BLA-L-CPFB(-2) had a tendency to precipitate if stored at -20°C and as a result the purified protein was stored at +4°C and used within one month of purification.

3.3 Intermolecular disulfide bond formation

Analysis of disulfide bond formation between BLA-L-SBPA18C and SAVSBPM32 was done utilizing the reducing and non-reducing SDS-PAGE procedure described in section 2.11.1.1. Briefly, three types of samples were prepared which included 1 μM of the tetrameric streptavidin muteins, 4 μM of the ligand variants, and a mixture of 1 μM of the tetrameric streptavidin muteins and 4 μM of the ligand variants. Reduction and subsequent disulfide bond formation was then done according to section 2.11.1.1 methods.

Since each subunit in the streptavidin tetramer and each of the fusion protein ligand tags contains a cysteine, an array of potential complexes could exist (Figure 3.5). In order to distinguish between these complexes, the mixtures were separated into two samples of equal volumes, mixed with SDS-PAGE loading dye, and denatured by a 5-minute incubation in boiling water in the presence or absence of the reducing agent β-mercaptoethanol. Denaturing the samples in the presence of a reducing agent decreased the types of complexes present in the sample, resulting in either streptavidin subunits (~16.5 kDa) or the ligand tag containing BLA fusion proteins (~35 kDa) (Figure 3.5). These samples act as a control to ensure that the proteins added were free from contaminants and to ensure that each reaction contained similar amounts of protein.
Figure 3.5 Schematic of reducing and non-reducing SDS-PAGE electrophoresis. Two BLA-L-SBPA18C tags (shown in orange and grey) covalently linked by disulfide bonds (shown as a dashed line) to tetrameric streptavidin (shown in blue). One set of samples is exposed to a reducing agent and denatured resulting in streptavidin subunits (16.5 kDa) and BLA-L-SBPA18C (35 kDa). The second set of samples is denatured in the absence of a reducing agent resulting in disulfide bonded heterodimeric BLA-L-SBPA18C-SAVSBPM32 (51.5 kDa) and streptavidin subunits (16.5 kDa).
The second set of samples that were denatured in the absence of a reducing agent has 5 predicted complexes that include streptavidin subunits alone (~16.5 kDa), streptavidin subunits covalently linked to other streptavidin subunits (~33 kDa), streptavidin subunits covalently linked to ligand tag containing BLA fusion proteins (~51.5 kDa), BLA-L-SBPA18C and dimers of BLA-L-SBPA18C (~35 and 70 kDa, respectively) (Figure 3.5). These complexes were separated and analyzed by SDS-PAGE with Coomassie blue staining. The results show that SAVSBPM32 was able to effectively form intermolecular disulfide bonds with BLA containing the SBPA18C tag (Figure 3.6, lane 6). A small amount of SAVSBPM32 covalently linked to the non-biotinylated CPFB(-2) tag was also observed (Figure 3.6, panel B, lane 7), indicating that the vast majority of thiol coupling seen in the reactions was not caused simply by the availability of accessible cysteine residues. This conclusion is further supported by the inability of SAVSBPM96 to form covalent bonds with BLA-L-SBPA18C (Figure 3.6, panel B, lane 5). Since the cysteine residue of SAVSBPM96 is outside of the binding pocket and on the accessible surface, disulfide bond formation would be expected if thiol coupling was a result of the high concentration of accessible cysteine rather than the affinity of the tag. Both results indicate that effective affinity thiol coupling between streptavidin and the ligand tag requires both the SAVSBPM32 variant and SBPA18C tag.
Figure 3.6 Disulfide bond formation between various SAV muteins and cysteine containing tags.

Reduced SAVSBPM32 (M32) and reduced SAVSBPM96 (M96) were mixed with either reduced BLA-SBPA18C (A18C), reduced unbiotinylated BLA-CPFB(-2) (-2) or buffer containing no protein. The ratio of SAV to BLA fusion proteins was 1 µM tetrameric SAV to 4 µM fusion protein. +ME and –ME indicate the presence and absence of β-mercaptoethanol in the sample loading buffer. Arrowheads to the right of the gels indicate approximate size of protein species. Arrowhead labeled 16.5 marks the position of SAV subunits with an apparent molecular mass 17 kDa (lanes 1, 2, 5, 6 and 7 panels A and B). Arrowhead labeled 35 marks the position of monomeric BLA-L-SBPA18C with an apparent molecular mass of 36 kDa (lanes 4, 5 & 6 panel A and B), monomeric non-biotinylated BLA-L-CPFB(-2) with an apparent molecular mass of 33 kDa (lanes 3 and 7 panels A and B) and disulfide bonded homodimeric SAVSBPM96 with an apparent molecular mass of 36 kDa (lane 1 panel B). Arrowhead labeled 51.5 marks the position of disulfide bonded heterodimeric BLA-L-SBPA18C-SAVSBPM32 with an apparent molecular mass of 48 kDa (lane 6 panel B) and disulfide bonded heterodimeric non-biotinylated BLA-L-CPFB(-2)-SAVSBPM32 with an apparent molecular mass of 45 kDa (lane 7 panel B). Arrowhead labeled 70 marks the position of disulfide bonded homodimeric BLA-L-SBPA18C with an apparent molecular mass of 70 kDa (lanes 4 and 5 panel B) and disulfide bonded homodimeric unbiotinylated BLA-L-CPFB(-2) with an apparent molecular mass of 66 kDa (lanes 3 & 7 panel B). Lane numbers are shown at the bottom of the gels. Protein molecular weight markers are included in Lane labeled MW. Numbers shown on the left represent the molecular mass expressed in kDa. Samples containing SAVSBPM96, SAVSBPM32, unbiotinylated BLA-L-CPFB(-2) and BLA-L-SBPA18C were loaded to lanes 1, 2, 3 and 4 respectively. Samples containing a mixture of SAVSBPM96:BLA-L-SBPA18C, SAVSBPM32:BLA-L-SBPA18C and SAVSBPM32:unbiotinylated BLA-L-CPFB(-2) were loaded to lanes 5, 6 and 7 respectively. Lane B represents an empty lane separating samples that contain β-mercaptoethanol to the samples that do not contain β-mercaptoethanol. Numbers above the gels represent protein concentration. Panel C indicates both the expected and apparent molecular mass for each species.
It is interesting to note that SAVSBPM96 was observed to form disulfide bonded homodimers but SAVSBPM32 was not (Figure 3.6, panel B, lanes 1 and 2). This is likely due to the difference in the position of the cysteine residue in the two variants. As previously mentioned, the cysteine in SAVSBPM96 is located on a surface exposed area that is outside of the SBP binding pocket. It would be easier to form a disulfide bond between SAVSBPM96 in a concentration dependant manner. This self dimerization may be avoided in SAVSPM32 as the cysteine in this mutant is less accessible and pointed towards the inside of the binding pocket of each subunit (Figure 3.1).

3.4 Analyses of protein complex formation

3.4.1 Kinetics of disulfide bond formation

The first step involved in studying the formation of the protein complex was to gain a general understanding of the time required for the BLA-L-SBPA18C to bind to SAVSBPM32 and form the disulfide bond. This was done using the modified Aitken and Learmonth carboxymethylation method described in section 2.11.1.2 (Aitken & Learmonth, 2002). Briefly, tetrameric SAVSBPM32 (2 µM) and BLA-L-SBPA18C (4 µM) were reduced separately in immobilized TCEP as previously described. The reduced BLA-L-SBPA18C and SAVSBPM32 were then mixed together at a final concentration of 1 µM tetrameric SAVSBPM32 and 2 µM BLA-L-SBPA18C, separated into 14 tubes and incubated at 23°C. At different time points (0-120 minutes), the reactions were stopped by the addition of a carboxymethylation buffer. The iodoacetamide within this buffer effectively stops the formation of disulfide bonds by covalently modifying the
thiol group of any cysteine residues not already involved in disulfide bonds (Hansen & Winther, 2009). The time required for binding and disulfide bond formation between SAVSBPM32 and BLA-L-SBPA18C was then analyzed using the same reducing and non-reducing SDS PAGE procedure previously described. The band intensities of the disulfide bonded streptavidin subunits covalently linked to BLA-L-SBPA18C (~51.5 kDa) was measured using the ImageJ software (Abramoff et al., 2004; Schneider et al., 2012) as described in section 2.10. Results indicate that at concentrations of 1 µM tetrameric SAVSBPM32 and 2 µM BLA-L-SBPA18C, disulfide bonds formed quickly, with 50% being formed within the first 5 minutes and over 90% reaction completion after 20 minutes (Figure 3.7). Therefore, the previous incubation time used of 1 hour was continued throughout the remainder of this work unless specified otherwise.
Figure 3.7 Carboxymethylation of cysteine residues

Measurements of disulfide bond formation over time in reactions containing 1 µM tetrameric SAVSBPM32 and 2 µM BLA-L-SBPA18C by (A) reducing and (B) non-reducing SDS-PAGE with Coomassie blue staining. +ME and –ME indicate the presence and absence of β-mercaptoethanol in the sample loading buffer. Numbers shown to the left of panels A and B indicate molecular mass of the marker proteins (lane MW). Numbers on the bottom of the gels indicate the time in minutes reactions were allowed to occur before Carboxymethylation of free cysteine residues inhibited formation of further heterodimeric protein. (C) Apparent and expected molecular mass of SAVSBPM32 (M32), BLA-L-SBPA18C (A18C), and disulfide bonded M32 (subunits) to A18C. (D) Intensity of the bands seen at 51.5 kDa were graphed as a function of time and curve fitting was done using SciDAVis (http://scidavis.sourceforge.net/).
3.4.2 Binding of the cysteine containing SBP tag to SAVSBPM32F

The next step in the analysis of protein complex formation was to study the stoichiometry and composition of complexes formed. This was done using the Semi-native polyacrylamide gel electrophoresis (s-native PAGE) methods described in section 2.11.2.2. Briefly, s-native PAGE was designed as a variation from the classical native PAGE methods in that 0.1% SDS is included in the running electrophoresis buffer. The addition of 0.1% SDS into the running buffer is strong enough to dissociate any non-covalently bound SBP tags from streptavidin without denaturing the streptavidin tetramer. This characteristic of s-native PAGE was demonstrated by mixing 5 µM of SAVSBPM18, SAVSBPM19, SAVSBPM32 and wild type core streptavidin (wtSAV, residues 13-139) with excess BLA-L-SBP (50 µM). The mixtures were then subjected to a reducing environment using immobilized TCEP, followed by the removal of the immobilized TCEP beads, and incubated at 23°C for 1 hour before running on an 8% polyacrylamide, s-native PAGE.

For this experiment, the SBP tag was used since it should have the highest non-covalent binding affinity to streptavidin. Additionally, even though the SBP tag does not contain a cysteine residue, the mixtures were initially subjected to and subsequently removed from a reducing environment using immobilized TCEP. This was done to ensure that any covalent linkage observed was not caused by the reporter protein BLA in case the natural disulfide bond within BLA is only partially formed. Results visualized by Coomassie blue staining show that the presence of the detergent SDS at 0.1% final concentration in the running buffer used during s-native PAGE is sufficient to dissociate
any BLA-L-SBP bound to streptavidin and its muteins without denaturing the streptavidin tetramers (Figure 3.8, panel A). Furthermore, these results indicate that any band shifts seen in subsequent experiments using s-native PAGE are a result of covalently linkage.

Further examination of the gel shows that wtSAV migrates further through the gel compared to SAVSBPM18, SAVSBPM32, and SAVSBPM96 (Figure 3.8, panel A). This difference in migration is expected as wtSAV is composed of residues 13-139 of the full length streptavidin (i.e. the core form SAV). This results in the protein having a mass 22% less than that of the mutants (12.9 and 16.5 kDa subunit$^{-1}$ respectively) (Table 3.5). Additionally, the charge to mass ratio of wtSAV is greater than that of the mutants which would also contribute to the faster migration of wtSAV observed (Gasteiger et al., 2005). In order to increase the separation of different protein complexes, the variant SAVSBPM32F was used. SAVSBPM32F which includes the FLAG tag insertion (DYKDDDDK) has a charge to mass ratio for each subunit that is 79% greater than SAVSBPM32.
Figure 3.8 Semi-native polyacrylamide gel electrophoresis

(A) 5 µM of reduced tetrameric wild type streptavidin (Wt), SAVSBPM18 (M18) and SAVSBPM32F (M32F) alone (lanes 1, 3, and 5, respectively) or as mixtures containing reduced BLA-L-SBP at a molar ratio of 10:1 (SBP:SAV) (lanes 2, 4, and 6, respectively). Reduced BLA-L-SBPA18C (A18C) mixed with 5 µM reduced tetrameric SAVSBPM32F (M32F) at molar ratios ranging from 0.43:1 - 5.3:1 (B) and 30:1 (C). Samples were separated over 1 hour at constant current of 30 milliamperes on an 8% semi-native polyacrylamide gel. Arrowheads c, f, g, and e marks the position of tetrameric SAVSBPM32F, SAVSBPM18, wtSAV, and BLA fusion proteins respectively (panels A, B and C). Arrowhead d marks the position of homodimeric disulfide bonded BLA-L-SBPA18C (panels B and C). Arrowheads a$_1$ and a$_2$ marks the position of the doublet termed “complex a$_1$ a$_2$” and Arrowhead b marks the position of the complex termed “complex b” (panels B and C).
Table 3.5 Physical parameters of streptavidin and the studied variants

<table>
<thead>
<tr>
<th>Protein*</th>
<th>Number of amino acids</th>
<th>Mw (kDa)</th>
<th>Theoretical PI</th>
<th>Charge at pH 8.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>wtSAV</td>
<td>123</td>
<td>12.9</td>
<td>5.78</td>
<td>-3.4</td>
</tr>
<tr>
<td>M18</td>
<td>159</td>
<td>16.5</td>
<td>6.53</td>
<td>-2.7</td>
</tr>
<tr>
<td>M32</td>
<td>159</td>
<td>16.5</td>
<td>6.53</td>
<td>-3.4</td>
</tr>
<tr>
<td>M32F</td>
<td>168</td>
<td>17.6</td>
<td>5.28</td>
<td>-6.5</td>
</tr>
</tbody>
</table>

* Parameters are reported for based on one streptavidin subunit
It is additionally predicted that the two positively charges residues (Lys) will favour SDS binding (Tal, Silberstein, & Nusser, 1985). This combination of traits should allow further migration of SAVSBPM32F in s-native PAGE and result in better resolution of the BLA-L-SBPA18C-SAVSBPM32F protein complexes (Table 3.5). To study these complexes, samples were prepared containing 5 µM of SAVSBPM32F tetramer in the presence of various concentrations of BLA-L-SBPA18C (0 - 150 µM). Samples were then subjected to s-native PAGE using the same method previously described. Results show 6 distinct bands termed and labeled by increasing migration through the gel as a₁, a₂, b, c, d, and e (Figure 3.8). Lanes 1 and 14 contain 5 µM SAVSBPM32F and 26.2 µM BLA-L-SBPA18C, respectively. Bands c, d and e are composed of tetrameric SAVSBPM32F, disulfide bonded homodimeric BLA-L-SBPA18C and BLA-L-SBPA18C, respectively. Since the concentration of SAVSBPM32F shown in lane 1 is kept constant throughout the reactions, any band shifts seen in subsequent reactions were not caused by the self-crosslinking of SAVSBPM32F. Additionally, since the concentration of BLA-L-SBPA18C shown in lane 14 was the highest ligand concentration used throughout the titration experiment, band shift associated with self-crosslinking of BLA-L-SBPA18C has been accounted for (Figure 3.8, panel B, lane 14). When SAVSBPM32F and BLA-L-SBPA18C were mixed together, the first visible band shift occurs when the molar ratio of BLA-L-SBPA18C is 1.3 times higher than that of tetrameric SAVSBPM32F (Figure 3.8, panel B, lane 3). Since the ratio of the ligand to SAVSBPM32F is close to 1 ligands per tetramer, the
logical composition of complex b would be a tetrameric SAVSBPM32F bound to a single BLA-L-SBPA18C.

When the concentration of BLA-L-SBPA18C is increased to 1.7 times that of SAVSBPM32F, complex b becomes more pronounced and complexes a₁ and a₂ become apparent (panel B, lane 5). This suggests that bands a₁ and a₂ contain protein complexes in which more than 1 BLA-L-SBPA18C is bound to each tetrameric SAVSBPM32F. As the concentration of BLA-L-SBPA18C continues to increase up to 5.2 times that of the SAVSBPM32F, the intensity of bands a₁ and a₂ increases linearly, where the intensity of band b only increases until the ratio of 4.8 is reached (Figure 3.8, panel B). The intensity of band b slightly decreases when the concentration of the ligand tag is increased from 4.8 to 5.2 times that of the SAVSBPM32F (Figure 3.8, panel B, lanes 12 and 13) and drastically decreases when the ratio of 30 to 1 is used (Figure 3.8, panel C). This observation is consistent with the idea that band b contains protein complexes composed of tetrameric SAVSBPM32F covalently linked to a single BLA-L-SBPA18C and bands a₁ and a₂ have more than one BLA-L-SBPA18C covalently linked per tetrameric SAVSBPM32F. This interpretation is consistent with the crystal structure of the SBP tag-streptavidin complex (Barrette-Ng et al., 2013) which also shows that two SBP tags can bind to one streptavidin tetramer. The cause of the doublet (a₁, a₂) is unclear.

It is also of interest to note that when the concentration of BLA-L-SBPA18C is 5.2 times greater than that of tetrameric SAVSBPM32F, bands c, d, and e are still present (Figure 3.8). The presence of these bands indicates that even when BAL-L-SBPA18C is in excess, a portion of the SAVSBPM32F population is still unable to covalently link to
the ligand. This may indicate that some of the SAVSBPM32F tetramers are “dead”
protein and unable to bind. This was tested by further increasing the ratio of BLA-L-
SBPA18C to 30 times that of SAVSBPM32F. Results show that the amount of tetrameric
SAVABPM32F (non-complexed form) was decreased to the level that was almost
undetectable by Coomassie blue staining (Figure 3.8, panel C). These results indicate that
almost all of the SAVSBPM32F proteins are capable of binding and covalently linking to
SBPA18C tags. One of the major limitations of s-native PAGE is that in order to interpret
the results assumptions must be made of the individual components that make up the
distinct protein complexes. One assumption is that although, the data reflects one and two
BLA-L-SBPA18C proteins coupled per tetrameric SAVSBPM32F, the argument could
be made that since each SAVSBPM32F subunit contains an accessible cysteine residue
any combination from 1-4 BLA-L-SBPA18C could occur.

3.4.3 Compositions of different streptavidin (SAVSBPM32F)-SBP tag (BLA-L-
SBPA18C) complexes analyzed by a modified two dimensional polyacrylamide
gel electrophoresis

In order to gain a better understanding of the individual components, which make
up these complexes, a novel two-dimensional gel electrophoresis method was
developed (section 2.11.3.1). In this procedure, protein complexes are separated in
the first dimension by s-native PAGE (Figure 3.9 panel A). The second dimension is
then done at a 90° angle compared to the first dimension with reduced proteins in
the presence of 0.1% SDS. This second step separates the individual components of
the already separated protein complexes (Figure 3.9, panel A). Once the gels are stained, the intensities of each component in the complexes are measured and the ratios of SAVSBPM32F and BLA-L-SBPA18C can be determined.

Results from this experiment indicate that the ratios of BLA-L-SBPA18C to SAVSBPM32F are equivalent in the doublet \(a_1 = 1.367\) and \(a_2 = 1.405\) and twice as large compared to the ratio found in complex b \(b = 0.705\) (Figure 3.9, panel B and D). These results are reported as ratios of BLA-L-SBPA18C to SAVSBPM32F for the following reasons. Direct comparison of BLA-L-SBPA18C to SAVSBMP32F was avoided due to complications that could arise from the folded nature of SAVSBPM32F in s-native PAGE and the differential binding of Coomassie blue for proteins depending on the basic amino acid composition (Tal et al., 1985). Using ratios avoided these potential problems since one molecule of BLA-L-SBPA18C or SAVSBPM32F will stain the same as another molecule of that same protein, regardless of which complex it originated from.

The results not only provide strong evidence that SAVSBPM32F can bind and covalently link to two SBPA18C-tags, but that the composition of the doublet \(a_1\) and \(a_2\) is the same. It should be noted that although the data supports the claim of one and two BLA-L-SBPA18C molecules bound per tetramer, the same ratios would be expected if instead two and four BLA-L-SBPA18C molecules were bound per tetramer.
Figure 3.9 Modified two dimensional polyacrylamide gel electrophoresis

Distinct components of the complexes formed in reactions containing SAVSBPM32F (M32F) and BLA-L-SBPAA18C (A18C) are analyzed using modified two dimensional polyacrylamide electrophoresis (M2D-PAGE) with Coomassie blue staining. (A) Schematic and experimental pictures of complex separation based in the first dimension of complexes and component separation in the second dimension. 1D, 2D, and arrows indicate order of the dimensions and current direction. +ME and -ME indicate the presence and absence of β-mercaptoethanol. Arrowhead a1, a2, and b marks the position of complex a1, a2, and b, respectively. Arrowheads c and e marks the position of tetrameric SAVSBPM32F and BLA-L-SBPAA18C, respectively. (B) Band intensity for each the components making up complexes a1, a2, b, and c were graphed. Intensities shown in orange consist of BLA-L-SBPAA18C and intensities shown in blue correspond to SAVSBPM32F. Curve fitting of the data was done on SciDAVis (http://scidavis.sourceforge.net/) using Gaussian functions (C) where A = area under the peak, xc = center of the peak, w = width of the peak, yo = offset of the y axis. χ²/dof = chi-squared distribution and R² = the coefficient of distribution. (D) Relative amount of each protein in complexes a1, a2, and b were calculated by dividing the area of the peaks corresponding to BLA-L-SBPAA18C by the area of the SAVSBPM32F peak for that complex.
Another composition analysis of different streptavidin (SAVSBPBM32F)-SBP tag (BLA-L-SBPA18C) complexes using the zinc imidazole reverse staining approach

To address the possibility that the data obtained from the modified two-dimensional gel electrophoresis method could equally describe protein complexes containing two and four tags per tetramer (Figure 3.10, panel A and B), the following experiment was conducted. A reaction containing SAVSBPM32F (5 µM) and BLA-L-SBPA18C (20 µM) was prepared and subjected to s-native PAGE as previously described (section 2.11.2.2). Staining of the gel was done using the zinc imidazole reverse staining method (Fernandez-Patron et al., 1998) which allowed for the separation and recovery of complexes a₁, a₂, and c (Figure 3.10, panel C) using the modified passive diffusion methods (Hardy et al., 1996, 1998) described in section 2.11.3.2. Samples from each separated protein complex were then analyzed using reducing and non-reducing SDS-PAGE procedures.

Results show the presence of monomeric SAVSBPM32F subunits in each of the extracted protein complexes (Figure 3.10, panel D). The presence of SAVSBPM32F subunits not covalently linked to BLA-L-SBPA18C in the non-reducing SDS PAGE is only consistent with the model of one and two tags coupled per tetramer. If the complexes were instead composed of two and four tags, the proteins extracted from bands a₁ and a₂ should consist of SAVSBPM32F with each subunit covalently linked to BLA-L-SBPA18C.
Figure 3.10 Zinc imidazole staining for complex extraction and analysis

Schematic of reducing and non-reducing SDS-PAGE electrophoresis for two and four BLA-L-SBPA18C tags binding to SAVSBPM32 (A and B, respectively). BLA-L-SBPA18C tags (shown in orange and grey) covalently linked by disulfide bonds (shown as a dashed line) to tetrameric streptavidin (shown in blue). One set of samples is exposed to a reducing agent and denatured resulting in streptavidin subunits (16.5 kDa) and BLA-L-SBPA18C (35 kDa). The second set of samples is denatured in the absence of a reducing agent resulting in disulfide bonded heterodimeric BLA-L-SBPA18C-SAVBPM32 (51.5 kDa) (A and B). The boiled non-reduced samples also produces streptavidin subunits (16.5 kDa) except when four tags are bound (A and B). (C) Complexes from reactions containing SAVSBPM32 (M32) and BLA-L-SBPA18C were separated by s-native PAGE and detected using zinc imidazole staining. (D) The protein complexes $a_1$, $a_2$, and $b$ were removed from the gel using a scalpel, extracted from the gel slices by diffusion and analyzed on reducing and non-reducing SDS PAGE with Coomassie blue staining. +ME and –ME indicate the presence and absence of β-mercaptoethanol in the sample loading buffer. Numbers shown to the left of the gels in panel D indicate molecular mass of the marker proteins. Arrowheads A18C marks the position of BLA-L-SBPA18C (Mw 35 kDa). Arrowheads M32 and M32-A18C mark the
position of SAVSBPM32 subunits (Mw = 16.5 kDa) and heterodimeric disulfide bonded SAVSBPM32-BLA-L-SBPA18C (Mw = 51.5 kDa), respectively.
3.5 Analysis of the SAV-SBP tag complexes by native polyacrylamide electrophoresis

Since the s-native PAGE method selectively dissociates non-covalently bound ligand tags from streptavidin (section 3.4.2), it is of interest to also study the complexes by native polyacrylamide electrophoresis (native PAGE). Previously it has been reported that the C-terminal region in streptavidin negatively interferes ligand binding through steric hindrance (Takeshi et al., 1995). Additionally, shortening both termini has been shown to increase biotin binding (Takeshi et al., 1995). In SAVSBPM32F, the FLAG tag was inserted into the flexible N-terminal region increasing its length. It is unclear if increasing the length of the termini in the SAVSBPM32F variant decreases binding. Therefore in order to avoid this possibility the SAVSBPM32 variant which does not contain the FLAG tag insertion was used for native PAGE experiments. Results using this variant would allow better comparisons to SAVSBPM18, which also does not contain the FLAG tag insertion. Reactions containing 5 µM tetrameric SAVSBPM32 were mixed with 0-17.5 µM BLA-L-SBPA18C or BLA-L-SBP. The reducing agent TCEP was then added to the reactions (1 mM final concentration) and the samples were analyzed by native PAGE following section 2.11.2 methods.

Results show band shifts in the BLA-L-SBPA18C and SAVSBPM32 reactions but not in the BLA-L-SBP and SAVSBPM32 reactions (Figure 3.11). Since the reducing agent TCEP was added to the samples it was expected that the BLA-L-SBPA18C would show equal or lower binding ability to SAVSBPM32 compared to BLA-L-SBP.
Figure 3.11 Native polyacrylamide gel electrophoresis

Reactions containing 5 µM tetrameric SAVSBPM32 mixed with 0-17.5 µM BLA-L-SBPA18C (A) or BLA-L-SBP (B) in the presence of a reducing agent (1 mM TCEP). Samples were analyzed using an 18% polyacrylamide native PAGE with Coomassie blue staining. Arrowheads c and e mark the position of tetrameric SAVSBPM32 and either BLA-L-SBPA18C (A) or BLA-L-SBP (B). Arrowheads a and b mark the position of BLA-L-SBPA18C-SAVSBPM32 bound complexes.
One explanation as to why BLA-L-SBPA18C and not BLA-L-SBP showed binding to SAVSBPM32 is that TCEP is negatively charged in alkaline solution and during electrophoresis quickly migrates towards the anode leaving the proteins in a non-reducing environment (Shaw & Riederer, 2003). As previous results identified (section 3.4.1) disulfide bonds quickly form in the absence of a reducing agent and therefore the band shifts seen in Figure 3.11 (panel A), likely contains covalently linked BLA-L-SBPA18C and SAVSBPM32. Since BLA-L-SBP does not covalently link to SAVSBPM32 and shows an absence of band shifts in the native PAGE gels it is assumed that complexes were once formed but gradually dissociate throughout the electrophoresis procedure. Because of this possibility, conclusions based on native PAGE experiments are uncertain and more sensitive binding experiments are needed.

### 3.6 Applications of SAVSBPM32 to purify SBP and SBPA18C tagged BLA by affinity and affinity thiol coupling chromatography

One application of the SAVSBPM32 system that was tested is the ability to purify SBPA18C tagged proteins by affinity thiol coupling chromatography. The ability of SAVSBPM32 to purify fusion proteins containing the unmodified SBP tag was also tested. It has previously been reported that leakage of bound protein occurs when overloading a SAVSBPM18-affigel column with BLA-L-SBP (Wu & Wong, 2013). Throughout this study, leakage will refer to proteins that are eluted from the column when only the wash buffer is present. Proteins in both the flow through and initial wash (W1) fractions are not considered as leakage. The flow-through fraction was excluded
since an excess amount of SBP tagged protein was loaded. The first wash fraction is also excluded due to the difficulty in obtaining “pure” W1 fractions with confidence that they do not contain any of the flow through that has not yet passed through the column.

Purified SAVSBPM32 and SAVSBPM18 were coupled to affi-gel 15 activated affinity media (BIO-RAD) and used throughout the purification studies. Culture supernatant from the model proteins, BLA-L-SBP and BLA-L-SBPA18C were first reduced and then loaded onto each matrix for both overloaded and non-overloaded conditions (150% and 20% the theoretical column capacity, respectively). The theoretical binding capacity of each was calculated individually for each column based on the amount of protein coupled to the affi-gel matrix and the assumption that a tetrameric streptavidin can bind two ligand tags (Wu & Wong, 2013). Results show that both SAVSBPM32 and SAVSBPM18-affigel matrices were able to purify BLA-L-SBP and BLA-L-SBPA18C in a single step to levels where no other contaminates were detectable (by Coomassie blue staining) (Figure 3.12 and 3.13). In the overloaded condition, the highest recovery with the least leakage was obtained using SAVSBPM32-affigel for the purification of BLA-L-SBPA18C. Protein recovery from this system was >50% higher and showed a 70% decrease in leakage compared to any other combinations tested (Figure 3.13, panel D). Additionally, both SAVSBPM32 and SAVSBPM18-affigel matrices showed a similar ability for purifying BLA-L-SBP. The recovery and leakage observed when the SAVSBPM18-affigel matrix was used for the purification of BLA-L-SBP in both conditions corresponds closely to those previously reported using the same combination (Sau-Ching Wu & Wong, 2013). Notably, the lowest recovery of protein was observed during the purification of BLA-L-SBPA18C by the SAVSBPM18-affigel
matrix (Figure 3.13, panels A and C). These observations may suggest that the A18C SBP mutation reduces binding affinity more drastically compared to the streptavidin A86C mutation.
**Figure 3.12 Purification of BLA-L-SBP from culture supernatant using SAVSBPM18 or SAVSBPM32-affigel**

Purification of BLA-L-SBP from WB800[pBLA-SBP] culture supernatant using SAVSBPM18 (panels A and C) or SAVSBPM32 (panels B and D) under column non-overloaded (panels A and B) and overloaded (panels C and D) conditions, respectfully. Fractions were collected and analyzed by SDS-PAGE with Coomassie blue staining. M: molecular weight markers. Numbers shown on the left represent the molecular mass expressed in kDa; S: culture supernatant; FT: flow-through; W1-W6: wash Fractions; E1-E4: elution fractions. Arrowheads mark the position of BLA-SBP. Analysis of protein recovery shown in panel E.

*Both theoretical binding capacity and theoretical maximum percent recovery is expressed in terms of the amounts (µg) of BLA-L-SBP that can bind to the matrix. This*
value is estimated with the assumption that one BLA-L-SBP molecule binds across two subunits in either the SAVSBPM18 or SAVSBPM32 tetramer. 1 µg of SAVSBPM18 dimer (Mr = 33,037.8) or 1 µg of SAVSBPM32 dimer (Mr = 33,102) will bind 1.062 and 1.060 µg of BLA-L-SBP (Mr = 35,101.7), respectfully.
Figure 3.13 Purification of BLA-L-SBPA18C using SAVSBPM18 or SAVSBPM32-affigel

Purification of BLA-SBPA18C from WB800[pBLA-SBPA18C] culture supernatant using SAVSBPM18 (panels A and C) or SAVSBPM32 (panels B and D) under column non-overloaded (panels A and B) and overloaded (panels C and D) conditions. Fractions were collected and analyzed by SDS-PAGE with Coomassie blue staining. M: molecular weight markers. Numbers shown on the left represent the molecular mass expressed in kDa; S: culture supernatant; FT: flow-through; W₁-W₆: wash Fractions; E₁-E₄: elution fractions. Arrowheads mark the position of BLA-SBPA18C. Analysis of protein recovery shown in panel E.

*Both theoretical binding capacity and theoretical maximum percent recovery is expressed in terms of the amounts (µg) of BLA-SBPA18C that can bind to the matrix. This value is estimated with the assumption that one BLA-SBPA18C molecule binds
across two subunits in either the SAVSBPM18 or SAVSBPM32 tetramer. 1 µg of SAVSBPM18 dimer (Mr = 33,037.8) or 1 µg of SAVSBPM32 dimer (Mr = 33,102) will bind 1.063 and 1.061 µg of BLA-SBPA18C (Mr = 35,133.8), respectively.
3.7  **Binding kinetics of SAVSBPM18 to BLA-L-SBPA18C and SAVSBPM32 to BLA-L-SBP**

The binding kinetics of purified SAVSBPM18 and SAVSBPM32 to BLA-L-SBPA18C and BLA-L-SBP fusion protein immobilized to Amine Reactive 2nd Generation (AR2G) biosensors (ForteBio) was tested using the Blitz™ label free protein assay system (ForteBio) (Concepcion et al., 2009) following section 2.11.4 methods. SAVSBPM32 and SAVSBPM18 concentrations of 157 nM, 117.8 nM, 78.5 nM, and 39.25 nM were analyzed on AR2G-BLA-L-SBP and AR2G-BLA-L-SBPA18C biosensors, respectively. Sensorgram data was plotted on dRU/dt Versus RU plot. T1 was calculated from the plot using the equation

\[ Y_{\text{max}} \left(1 - e^{-\frac{(x - X_0)}{T1}}\right) \]

where \( Y_{\text{max}} \) is the predicted maximum Y value, T1 is the predicted time for the signal to reach \( \frac{1}{2} Y_{\text{max}} \), x is the x-axis value for the data point used and \( X_0 \) is the initial value data x-axis value.

\( K_s \) was then calculated using the equation

\[ K_s = (1/T1)/C \]

Where \( K_s \) is the rate constant, and C is the analyte concentration (M).

\( K_{\text{on}} \) was calculated using the \( K_s \) Versus C plot and the equation

\[ K_s = K_{\text{on}}C + K_{\text{off}} \]

Where \( K_{\text{on}} \) is the slope of the plot and \( K_{\text{off}} \) is the Y-intercept.

\( K_{\text{off}} \) was determined using primary sensorgram data and the equation

\[ \ln(R_0/R_t) \text{ Versus } T-T_0 \]
Where $R_0$ is the initial signal, $R_t$ is the signal at time $t$, $T$ is the time and $T_0$ is the initial time. Data plotted was the average of three replicates and $K_{off}$ is equal to the slope. $K_d$ was calculated using the equation $K_d = K_{off}/K_{on}$.

Results indicate a $K_{on}$ value for SAVSBPM18 binding to BLA-L-SBPA18C and SAVSBPM32 binding to BLA-L-SBP of $3.58 \times 10^4$ and $3.14 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$, respectively (Figure 3.14 and 3.15). The calculated $K_{off}$ value for SAVSBPM18 binding to BLA-L-SBPA18C was $3.30 \times 10^{-3} \text{ s}^{-1}$ and the calculated $K_{off}$ value for SAVSBPM32 binding to BLA-L-SBP was $3.17 \times 10^{-4} \text{ s}^{-1}$ (Figure 3.14 and 3.15). The calculated $K_d$ value for SAVSBPM18 binding to BLA-L-SBPA18C and SAVSBPM32 binding to BLA-L-SBP was $9.20 \times 10^{-8} \text{ M}$ and $1.01 \times 10^{-8} \text{ M}$, respectively (Table 3.6). Additional results studying the binding interactions for SAVSBPM18 to BLA-L-SBP showed $K_{on}$, $K_{off}$, and $K_d$ values of $3.32 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$, $3.41 \times 10^{-4} \text{ s}^{-1}$, and $1.03 \times 10^{-8} \text{ M}$, respectively (Figure 3.15, panel B and Table 3.6). Relatively speaking, the results indicate that the A18C mutation in the SBPA18C tag drastically reduces its binding affinity to SAVSBPM18 by increasing the off-rate. These results correlate well with earlier binding study data (section 3.6) and further support the idea that the A18C SBP mutation reduces binding affinity more drastically compared to the streptavidin A86C mutation.
### A

**SBP M32**

![Graph showing binding over time with different concentrations](image)

### B

**Ymax*(1-exp(-(x-X0/T1))**

<table>
<thead>
<tr>
<th>Curve</th>
<th>157 nM</th>
<th>117.8 nM</th>
<th>78.5 nM</th>
<th>39.25 nM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data points</td>
<td>0-311.8</td>
<td>0-311.8</td>
<td>0-311.8</td>
<td>0-311.8</td>
</tr>
<tr>
<td>Ymax</td>
<td>0.295</td>
<td>0.331</td>
<td>0.372</td>
<td>0.295</td>
</tr>
<tr>
<td>T1</td>
<td>155</td>
<td>193</td>
<td>206</td>
<td>408</td>
</tr>
<tr>
<td>Chi^2/doF</td>
<td>2.60E-05</td>
<td>9.42E-06</td>
<td>5.65E-06</td>
<td>6.73E-06</td>
</tr>
<tr>
<td>R^2</td>
<td>0.995</td>
<td>0.998</td>
<td>0.999</td>
<td>0.997</td>
</tr>
</tbody>
</table>

### C

**A18C M18**

![Graph showing binding over time with different concentrations](image)
Figure 3.14 Non-linear curve fitting of SAVSBPM32-BLA-L-SBP and SAVSBPM18-BLA-L-SBPA18C sensorgram interaction data

Curve fitting of the sensorgram data for the binding interactions of immobilized BLA-L-SBP with the analyte SAVSBPM32 (A) and immobilized BLA-L-SBPA18C with the analyte SAVSBPM18 (C) was done on SciDAVis (http://scidavis.sourceforge.net) using equation shown in panel B and D. Data points = the range of points used in the curve fitting function, numbers correspond the X value of those data points. Ymax = the predicted maximum Y value. T1 is the predicted time for the signal to reach \( \frac{1}{2} \) Ymax. \( \chi^2/dof \) = chi-squared distribution. \( R^2 \) = the coefficient of distribution. Linearized data (T1), panels B and D were used to calculate \( k_{on} \).
Figure 3.15 Determination of the kinetic parameters of the interaction between SAVSBPM32-BLA-L-SBP and SAVSBPM18-SBPA18C.
BLA-L-SBP and BLA-SBPA18C were immobilized to biosensor chips and SAVSBPM32 and SAVSBPM18 serves as the analyte. (A) Linearized data from sensorgrams for the determination of the on-rate (slope of the plots). (B) Linearized data from sensorgrams for the determination of off-rate (slope of the plots). Data plotted for M18:A18C and M32:SBP in panel B is the average of three replicates. Error bars indicate ± SD. Data plotted for M18:SBP is from one trial (B).
Table 3.6 Kinetic parameters for the interactions between streptavidin variants and the SBP ligands.

<table>
<thead>
<tr>
<th>Analyte</th>
<th>Ligand</th>
<th>On-rate (M^{-1} s^{-1})</th>
<th>Off-rate (s^{-1})</th>
<th>K_d (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>wtSAV</td>
<td>BLA-L-SBP</td>
<td>9.5x10^4</td>
<td>8.16x10^{-4}</td>
<td>8.59x10^{-9}</td>
</tr>
<tr>
<td>M18</td>
<td>BLA-L-SBP</td>
<td>3.32x10^4</td>
<td>3.41x10^{-4}</td>
<td>1.03x10^{-8}</td>
</tr>
<tr>
<td>M32</td>
<td>BLA-L-SBP</td>
<td>3.14x10^4</td>
<td>3.17x10^{-4}</td>
<td>1.01x10^{-8}</td>
</tr>
<tr>
<td>M18</td>
<td>BLA-L-SBPA18C</td>
<td>3.58x10^4</td>
<td>3.30x10^{-3}</td>
<td>9.20x10^{-8}</td>
</tr>
</tbody>
</table>

*Values reported by (Wu & Wong, 2013)
3.8 Purification of BLA-L-SBPA18C using a SAVSBPM32 affi-gel column under different wash conditions

Leakage of BLA-L-SBPA18C was observed throughout all of the overloaded column purification experiments (section 3.6). It would be ideal to identify simple parameters needed to quickly and effectively remove any non-covalently bound proteins from the column. This would be especially important if the system is used for immobilization of proteins to sensor chips. In these types of experiments leaking of the supposedly immobilized SBPA18C tagged ligands from the chip surface would affect the validity of data obtained and limit the usefulness of the system. Purification procedures were done using the same overloaded methods previously described with the exception that various buffers containing different additives were used during the washing steps.

3.8.1 Purification with the inclusion of a nonionic detergent in the wash conditions

The nonionic detergent polysorbate 20 (Tween-20), which is able to dissociate proteins from nitrocellulose (at 0.05%) (Gershoni & Palade, 1983; Hoffman & Jump, 1986) and commonly used (up to 1%) to decrease nonspecific hydrophobic interactions in other affinity purification methods (Bornhorst & Falke, 2000) was tested for its ability to strip the non-specifically bound proteins from the SAVSBPM32 column in overloaded conditions. Purification procedure was the same as previously described for overloaded conditions with the exception that 2% Tween-20 was included in the wash buffer. Results indicate that the addition of Tween-20 was not able to significantly lower the leakage seen suggesting that leakage is either not caused by nonspecific hydrophobic
interactions or that those interactions are stronger than the parameters tested (Figure 3.16, panel A).

### 3.8.2 Purification under high ionic strength wash conditions

The metal halide salt, potassium chloride (KCl) was used to test nonspecific ionic interactions in the overloaded purification of BLA-L-SBPA18C from SAVSBPM32. A concentration of 300 mM KCl was previously reported as a component in the buffer used for the purification of the unmodified SBP-tag on streptavidin (Keefe et al., 2001). During this experiment, 300 mM KCl was included in the wash buffer. All other aspects of the purification procedure including the sample loading buffer did not contain the added KCl. These steps were carried out using the overloaded conditions described in section 2.12.7.2. Results indicate a decrease in leakage by 33% was achieved when using 300 mM KCl compared to the previous PBS buffer (Figure 3.16, panel B). This decrease in leakage is likely due to the ability of salt to suppress some nonspecific ionic interactions. These non-specifically bound proteins would quickly be eluted from the column into the flow through and first wash fractions. Although these results suggest that there may be some nonspecific ionic interactions occurring, the use of 300 mM KCl was not able to fully eliminate leakage and therefore suggested that other variables are still unaccounted for.
Figure 3.16 Purification of BLA-SBP A18C using SAVSBPM32-affigel and various wash buffers

Purification of BLA-SBP A18C from WB800[pBLA-SBP A18C] culture supernatant using SAVSBPM32 under overloaded conditions. BLA-SBP A18C bound to the column was washed with PBS buffer containing either 2% Tween 20 (A), 0.3 M KCl (B), 5 mM Biotin (C), or 5 mM biotin, 0.3 M KCl with a brief incubation (D). Fractions were collected and analyzed by SDS-PAGE with Coomassie blue staining. M: molecular weight markers. Numbers shown on the left represent the molecular mass expressed in kDa; S: culture supernatant; FT: flow-through; W1-W6: wash Fractions; E1-E4: elution fractions. Arrowheads mark the position of BLA-SBP. (E) Calculated amounts of BLA-L-SBP A18C found in leakage and elution fractions shown.

* Values reported in panel E for Leakage and Elution are expressed in terms of the % of total protein added.

<table>
<thead>
<tr>
<th>Type of wash</th>
<th>Theoretical binding capacity (ug)</th>
<th>Amount of BLA-L-SBP A18C loaded (ug)</th>
<th>Percent capacity</th>
<th>Recovery of elution's (%)</th>
<th>Leakage W2-W6 (%)</th>
<th>Recovery of W2-E4 (%)</th>
<th>Theoretical max recovery (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2% Tween 20</td>
<td>880</td>
<td>1320</td>
<td>150</td>
<td>45</td>
<td>7</td>
<td>52</td>
<td>67</td>
</tr>
<tr>
<td>0.3 M KCl</td>
<td>880</td>
<td>1320</td>
<td>150</td>
<td>50</td>
<td>4</td>
<td>54</td>
<td>67</td>
</tr>
<tr>
<td>5 mM Biotin</td>
<td>880</td>
<td>1320</td>
<td>150</td>
<td>51</td>
<td>3</td>
<td>54</td>
<td>67</td>
</tr>
<tr>
<td>5 mM Biotin + 0.3 M KCl</td>
<td>880</td>
<td>1320</td>
<td>150</td>
<td>47</td>
<td>1</td>
<td>48</td>
<td>67</td>
</tr>
</tbody>
</table>
3.8.3 Purification with the inclusion of biotin in the washing buffer

Throughout the literature, the elution of the SBP-tag from streptavidin and its relevant muteins has been done by competitive binding with 2-5mM biotin (Barrette-Ng et al., 2013; Keefe et al., 2001; Wu & Wong, 2013). Biotin was therefore used at the highest concentration reported (5 mM) to test these effects. This approach can be tested by taking advantage of the covalently linked BLA-L-SBPA18C, which will not be washed off from the column in the absence of a reducing agent. The purification procedure was done using the overloaded conditions with the exception that 5 mM d-biotin was included in the wash buffer (section 2.12.7.3). Results show that biotin was slightly more effective than KCl in decreasing leakage. Only 3% of the total amount of protein initially loaded could be detected in the leakage fractions (Figure 3.16, panel C). If the proteins bound to the column non-specifically via interaction located outside of the biotin binding pockets in streptavidin, the addition of free biotin would not cause a decrease in leakage. Results indicate that this was not the case and the addition of free biotin had the largest impact in terms of decreasing leakage (Figure 3.16). This data may suggest that the decreased binding affinity of the A18C mutation is caused by a conformational change in the α-helix. The “GHVV” and “HPQGQ” binding peptides of the SBP tag may be able to bind to the binding pockets of streptavidin even if a “kink” is formed in the SBP tag α-helix. This “kink” could then displace the cysteine residue from the position needed for the formation of the disulfide bond to C86 of streptavidin.
3.8.4 *Purification using a high salt and biotin containing washing buffer*

The purification procedure was done using the overloaded conditions with the exception that 5 mM d-biotin and 300 mM KCl was included in the wash buffer (section 2.12.7.4). The sample loading buffer (PBS) did not contain biotin or KCl. Once the sample was loaded to the column one column volume of the wash buffer was added. The top and bottom of the column was then capped and incubated at 23°C for 10 minutes prior collecting the fraction and continuing purification.

Figure 3.16, panel D shows only 1% leakage occurring after the first wash fraction. No observable leakage was seen after the second wash fraction. As a result these optimized conditions are suggested to be used for this system. It is particularly important for biosensor studies since only covalently linked proteins immobilized to the chip surface should function as binders for the binding study.

3.9 *Potential improvements of the SAVSBPM32 and SBPA18C system*

Current results indicate that the largest decrease in binding affinity is caused by the A18C mutation in the SBPA18C tag (section 3.6). Additionally, the results show that the lower binding affinity of this tag is a result of an increased off-rate from streptavidin (section 3.7).

Although Ala-18 of the SBP tag does interact with Ala-86 of streptavidin, losses in these interactions are not likely to drastically decrease binding. In terms of interactions with streptavidin, Ala-18 is neighboured by Leu-17, which is buried in the hydrophobic binding pocket of streptavidin (residues Trp-79, Arg-84, Ala-86 and Ser-88). Flanking
the opposite side of Ala-18 are Leu-21, Leu-24, and Leu-28 which also show hydrophobic interactions with streptavidin (Barrette-Ng et al., 2013). The loss of the relatively insignificant interactions Ala-18 has with streptavidin (compared to its neighbouring residues) does not explain the decrease in binding observed. One explanation may be that the decrease in binding is due to a larger conformational change caused by the mutation. The location of Ala-18 is in the α-helix (residues 17-28) of the SBP tag, which spans across the subunits acting as a linker or spacer for the GHVV and HPQGQ binding peptides (Barrette-Ng et al., 2013). Mutating alanine, an amino acid with the highest helix propensity (0 kcal mol$^{-1}$), with a low helix propensity cysteine (0.68 kcal mol$^{-1}$) (Pace & Scholtz, 1998) may introduce a “kink” in the α-helix. This kink could interfere with the binding of the SBPA18C tag to streptavidin by positioning the Cys-18 residue of the SBPA18C tag away from the Cys-86 residue of SAVSBPM32 decreasing intermolecular disulfide bond formation.

The web servers PROTEUS2 and PEP-FOLD were used to predict secondary structure of the SBP tag with and without the mutation (Table 3.7 and Figure 3.17) (Maupetit, Derreumaux, & Tuffery, 2009; Maupetit, Derreumaux, & Tufféry, 2010; Montgomerie et al., 2008; Thévenet et al., 2012). Additionally, new mutations were probed using PROTEUS2 for the ability to strengthen the α-helix. Results predict that the α-helix will become weakened by the A18C mutation (Table 3.8 and Figure 3.17). However, with a G19A mutation to increase the chances of helix formation, the predicted “kink” induced effect caused by the A18C mutation may be minimized. The G19A mutation is ideal both because it is neighbouring A18 and because it does not play a role in binding interactions with streptavidin residues (Barrette-Ng et al., 2013).
Table 3.7 Predicted secondary structure of residues 15-30 of the SBP tag and variants (Montgomerie et al., 2008).

<table>
<thead>
<tr>
<th>Tag</th>
<th>Predicted secondary structure*</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBP</td>
<td>15 EGLAGELEQLRARLEH 30 CCHHHHHHHHHHCC</td>
</tr>
<tr>
<td>SBPA18C</td>
<td>15 EGLCGELEQLRARLEH 30 CCCCCCCCCCCCCC</td>
</tr>
<tr>
<td>SBPA18CG19A</td>
<td>15 EGLCAELEQLRARLEH 30 CCCCCCCCCCCCCC</td>
</tr>
</tbody>
</table>

*Predicted secondary structure of residues 15-19 reported were obtained from the PROTEUS Structure Prediction Server 2.0. Hand C indicate alpha helix and coil structures, respectively.
Figure 3.17 Local structure prediction profile of the SBP and SBPA18C tag

Graphical representation of the probabilities residues in the SBP tag (A) and SBPA18C tag (B) will form helical (red), extended (green), or coil (blue) structure. Figures were generated on the PEP-FOLD structure prediction server (http://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD/).
For testing the new SBP tag containing both the A18C and G19A mutations the model protein Maltose-binding protein (MBP) was used instead of the previously used BLA. The main reason MBP was chosen as a model protein compared to BLA is that it lacks a disulfide bond. Using a model protein without a disulfide bond would help interpret if the doublet previously seen during s-native PAGE experiments was caused by the incomplete formation of disulfide bonds in the BLA population (Figure 3.8, panel A, band e). The recombinant fusion protein termed MBP-L-SBPA18CG19A was generated to have MBP fused to the C-terminal SBPA18CG19A tag with a 19-amino-acid linker (IDPAGTSPSTPEGPSTPSN) separating the SBPA18CG19A tag and the MBP (Appendix 2). At the N-terminal end, a His tag was also added. MBP-L-SBPA18CG19A was produced by induction in E. coli BL21(D3) cells and purified on His-Select Nickel HC affinity gel (Sigma) (section 2.9.8).

Figure 3.18, panel A show that although leakage of the recombinant protein occurred throughout the wash fractions, MBP-L-SBPA18CG19 was able to be purified to >95% in a single step. Elution fractions were then collected, dialysed, reduced and loaded to the SAVSBPM32-affigel affinity column. Figure 3.18, panel B shows that MBP-L-SBPA18CG19A loaded to the column leaked during the wash fractions and was unable to effectively bind to the SAVSBPM32-affigel column. The inability for MBP-L-SBPA18CG19A to bind to SAVSBPM32-affigel column may be caused by the inability of the mutated tag to adopt the predicted conformation. PROTEUS2 predicted that the $\alpha$-helix in the SBPA18CG19A variant would extend from residue 17-28 of the tag, which is the desired residues and length. Interestingly, PROTEUS2 also predicted that the $\alpha$-helix in the unmodified SBP tag would extend from residues 18-28 of the tag, slightly shorter
then the correct conformation. This difference is likely due binding interactions the SBP tag has with streptavidin promoting the formation of the helix that were not taken into account for our predictions.
Figure 3.18 Purification of MBP-L-SBPA18CG19A using His-Select and SAVSBPM32-affigel

(A) Purification of MBP-L-SBPA18CG19A by His-Select Nickel Affinity column. (B) E1-E3 collected, dialyzed in PBS and applied to SAVSBPM32-affigel. Fractions were collected and analyzed by SDS-PAGE with Coomassie blue staining. M: molecular weight markers. Numbers shown on the left represent the molecular mass expressed in kDa; S: culture supernatant; HPS: purified fraction from His-select Nickel Column; FT: flow-through; W1-W6: wash Fractions; E1-E4: elution fractions. Arrowheads mark the position of MBP-L-SBPA18CG.
4.1 Rational design of the streptavidin CPFB tag system

4.1.1 Cysteine containing streptavidin variant SAVSBPM19 and the cysteine containing PFB variant CPFB(-2)

Previously, it has been shown that the monomeric streptavidin variant M6 is able to effectively form disulfide bonds with the cysteine containing biotinylation tag variant CPFB(±2) through an N118C mutation in M6 (Wu et al., 2009). The CPFB(±2) tag is a variation of the previously described PFB tag that has a Cys residue located either two residues downstream or upstream of a lysine residue which can be enzymatically biotinylated (Cull & Schatz, 2000; Wu, Yeung, et al., 2002) (Figure 4.1, panels A, B, and C). Since these mutations were shown to form disulfide bonds in the previous system, it was logical to use them in this study. The same cysteine mutation used in M6 (N118C) was therefore combined with the SAVSBPM18 to create the variant SAVSBPM19 (Table 4.1). The mature form of the recombinant CPFB(-2) fusion proteins termed BLA-L-CPFB(-2) (~ 32.5 kDa) was also generated. This fusion protein includes a 20-amino-acid linker (IDPAGTSPSTSGGSTSGSGS) between the CPFB(-2) tag and BLA (Table 4.2). The combination of the N118C mutation in SAVSBPM19 and the A-2C mutation in CPFB(-2) is predicted to allow the formation of intermolecular disulfide bonds upon binding (Figure 4.2).
Figure 4.1 Model of the PFB peptide

(A) Sequence of the PFB tag, ± numbers below indicate the residues position upstream or downstream form the Lysine residue (0). (B) Model of the unbiotinylated PFB tag with Lys at position 0 is indicated by K0 in cyan. ±2 residue locations are labeled and colored in magenta (C) Model of the PFB tag after biotinylation of K0. (D) Schematic showing the rotation of the PFB tag with K0 as the axis. Models were generated using PyMOL with the unpublished PDB file new6pept (designed by Dr. Kenneth Ng and obtained from Dr. S.-L. Wong with permission) as a starting file.
Figure 4.2 Modeling of CPFB(-2) and engineered streptavidin variants
SAVS BPM19 and SAVSBPM32

(A) Wild type streptavidin and PFB residues in orange previously shown to form disulfide bonds when mutated to Cys in monomeric M6 and CPFB(-2) (Wu et al., 2009).
(B) Resulting mutations used to generate the SAVSBPM19 and CPFB(-2) variants with the predicted disulfide bond shown in yellow. (C) Position of PFB after rotation around the biotinylated lysine (cyan). Wild type streptavidin and PFB residues (orange) predicted to form disulfide bonds (A-2 and A86) after mutation to Cys. (D) Resulting mutations used to generate SAVSBPM32 and CPFB(-2) variants with the predicted disulfide bond shown in yellow. Models were generated using PyMOL with the unpublished PDB file new6pept (designed by Dr. Kenneth Ng and obtained from Dr. S.-L. Wong with permission) as a starting file.
4.1.2  SAVSBPM32 and BLA-L-CPFB(-2)

Interestingly, the cysteine residues located in either the +2 or -2 positions showed similar disulfide bond formation ability with M6 (Wu et al., 2009). This indicates that once biotin binds to the pocket of M6, the tag can rotate around the biotinylated lysine until the cysteine is correctly positioned for disulfide bond formation (Figure 4.1, panel D). As a result of this rotation, the CPFB(-2) tag is predicted to be able to form disulfide bonds with other residues of streptavidin that fall within the radius of the rotating cysteine. One interesting residue of streptavidin that falls within this radius A86. Therefore, SAVSBPM32 (Table 4.1) which can form disulfide bonds with SBPA18C tags, is also predicted to be able to form disulfide bonds with BLA-L-CPFB(-2) (Figure 4.2).

4.1.3  Negative controls for cysteine containing streptavidin and ligand

The cysteine containing streptavidin control (SAVSBPM96) used in Chapter 3 was also used throughout these experiments (Table 4.1). Additionally the non-biotinylated CPFB(-2) also used in Chapter 3 was again used throughout these experiments as a control (Table 4.2).

4.2  Protein production and purification

SAVSBPM32, SAVSBPM96, and unbiotinylated BLA-L-CFB(-2) used in these studies were from aliquots obtained during Chapter 3 experiments.
Table 4.1 Streptavidin variants nomenclature and mutations.

<table>
<thead>
<tr>
<th>Streptavidin variant</th>
<th>Mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAVSBPM18</td>
<td>G48T, S27A</td>
</tr>
<tr>
<td>SAVSBPM32</td>
<td>G48T, S27A, A86C</td>
</tr>
<tr>
<td>SAVSBPM32F</td>
<td>G48T, S27A, A86C, DYKDDDDK insertion</td>
</tr>
<tr>
<td>SAVSBPM96</td>
<td>G48T, S27A, A63C</td>
</tr>
<tr>
<td>SAVSBPM19</td>
<td>G48T, S27A, N118C</td>
</tr>
</tbody>
</table>
Table 4.2 Sequences of the ligand tags and variants.

<table>
<thead>
<tr>
<th>Ligand tag</th>
<th>Amino acid sequence*</th>
</tr>
</thead>
<tbody>
<tr>
<td>PFB</td>
<td>LHHILDAQKMVWNHR</td>
</tr>
<tr>
<td>CPFB(-2)</td>
<td>LHHILDCQKMVWNHR</td>
</tr>
</tbody>
</table>

*Bolded and underlined letters indicate mutated residues
SAVSBP19 was similarly produced by secretion from *Bacillus subtilis* WB800 cells and affinity purified on biotin agarose following methods described in section 2.9.

Biotinylation of BLA-L-CPFB(-2) done overnight at 30°C using the engineered BirA enzyme and purified on Monomeric Avidin Agarose showed a low recovery of ≈ 20%, (n = 2) (Figure 4.3, panel A). The flow through was then collected and reloaded to the column, which showed no binding. The results indicate that the low recovery observed was a result from inefficient biotinylation by the BirA enzyme (Figure 4.3, panel B). Increased biotinylation with a higher recovery (≈ 80% recovery, n = 2) was achieved by first semi-purifying BLA-L-CPFB(-2) on a DEAE ion exchange column prior to biotinylation (section 2.9.9) (Figure 4.3, panel C). Although recovery of biotinylated BLA-L-CPFB(-2) was increased, precipitation of the biotinylated protein was extensive. As a result, purified biotinylated protein was stored in +4°C and used within several days of purification. The apparent molecular mass of the biotinylated BLA-L-CPFB(-2) was estimated to be approximately 33 kDa which is in close agreement with the expected molecular mass of 32.5 kDa. Elution fractions E₁-E₆ were collected, dialyzed, and quantified using section 2.10 methods.
Figure 4.3 Purification of biotinylated BLA-L-CPFB(-2)

Secreted BLA-L-CPFB(-2) produced from pWB980-BLA-L-CPFB(-2) and biotinylated using the engineered BirA enzyme, affinity purified on Monomeric Avidin agarose. (A) Purification of BLA-L-CPFB(-2) after biotinylation of crude sample. (B) Re-purification of flow through (FT) collected from panel A. (C) Purification of BLA-L-CPFB(-2) after biotinylation of semi purified BLA-L-CPFB(-2). Fractions were collected and analyzed by SDS-PAGE with Coomassie blue staining. Mw: molecular weight markers. Numbers shown on the left represent the molecular mass expressed in kDa; S: Supernatant; FT: flow-through fraction; W1-W4: wash fractions; E1–E7: elution fractions. Arrowheads marks the position of BLA-L-CPFB(-2) (Mw = 32.5 kDa).
4.3 Intermolecular disulfide bond formation of BLA-L-CPFB(-2) to the streptavidin variants

Reduced proteins were mixed together at a molar ratio of 1 µM tetrameric streptavidin variant (SAVSBPM19, SAVSBPM96, SAVSBPM32, or SAVSBPM32F) to 1 µM biotinylated fusion protein (BLA-L-CPFB(-2)) using the same conditions and incubation times as described in section 3.3. Disulfide bond formation was then analyzed utilizing the reducing and non-reducing SDS-PAGE procedure described in section 2.11.1.1. Results show that SAVSBPM32, SAVSBPM32F and SAVSBPM19 can equally form disulfide bonds with the biotinylated BLA-L-CPFB(-2) protein (Figure 4.4). Additionally, SAVSBPM96 is unable to form disulfide bonds with the biotinylated ligand tag fusion protein (Figure 4.4, lane 4). This indicates that the disulfide bond formation observed is a result of affinity thiol coupling.

It is notable that disulfide bonded homodimeric biotinylated BLA-L-CPFB(-2) was apparent in the reaction containing SAVSBPM19, but only slightly visible in the reaction only containing biotinylated BLA-L-CPFB(-2) (Figure 4.4, lanes 4 and 5). Although the difference in the presence of disulfide bonded homodimeric biotinylated BLA-L-CPFB(-2) within the reaction containing BLA-L-CPFB(-2) (lane 5) compared to the reaction containing both BLA-L-CPFB(-2) and SAVSBPM96 (lane 4) seems contradictory, the following explanation is provided. First, it is assumed that formation of these CPFB(-2) tagged homodimers is a result of unspecific and concentration dependant thiol coupling.
Figure 4.4 Disulfide bond formation between various SAV muteins and cysteine containing CPFB(-2) tag

Reduced SAVSBPM32 (M32), SAVSBPM32F (M32F), SAVSBPM19 (M19) and SAVSBPM96 (M96) were mixed with reduced biotinylated BLA-L-CPFB(-2) at a molar ratio of 1 µM tetrameric SAV to 1 µM fusion protein. +ME and −ME indicate the presence and absence of β-mercaptoethanol in the sample loading buffer. Arrowheads to the right of the gels indicate approximate size of protein species. Arrowhead labeled 17 marks the position of SAVSBPM32, SAVSBPM19 and SAVSBPM96 (lanes 1, 3 and 4 panels A and B) with an apparently molecular mass of 17 kDa and SAVSBPM32F (lane 2 panels A and B) with an apparent molecular mass of 22 kDa. Arrowhead labeled 32 marks the position of biotinylated BLA-L-CPFB(-2) with an apparent molecular mass of 32 kDa (lanes 1-5 panels A and B). Arrowhead labeled 46 marks the position of disulfide bonded heterodimeric biotinylated BLA-L-CPFB(-2)-SAVSBPM32, disulfide bonded heterodimeric biotinylated BLA-L-CPFB(-2)-SAVSBPM32F and disulfide bonded heterodimeric biotinylated BLA-L-CPFB(-2)-SAVSBPM19 (lanes 1, 2 and 3 panel B) with an apparent molecular mass of 45, 47 and 45 kDa respectively. Arrowhead labeled 52 marks the position of disulfide bonded homodimeric biotinylated BLA-L-CPDB(-2) (lanes 1-5 panel B) with an apparent molecular mass of 52 kDa. Lane numbers are shown at the bottom of the gels. Numbers to the left of the gels represent the molecular mass expressed in kDa. Numbers above the gels represent protein concentration. Panel C indicates both the expected and apparent molecular mass for each species. The faint 17 kDa band observed in lane 5 panel A is M96 protein that flowed over from lane 4.
Second, it is assumed that SAVSBPM19 still posses affinity to the biotinylated tag without the ability to form covalent bonds. If both assumptions are correct then the SAVSBPM19 could act as a scaffold, increasing the local concentration of biotinylated BLA-L-CPFB(-2) around the binding pocket. This increased local concentration would result in more non-specific thiol coupling between two tags compared to the reaction where only BLA-L-CPFB(-2) is present.

4.4 Tetrameric state of SAVSBPM18, SAVSBPM19 and SAVSBPM32

Throughout this work the ability to generate sufficient amounts of soluble biotinylated BLA-L-CPFB(-2) was problematic. Three key factors contributed to this problem. First, the semi-purification of the unbiotinylated BLA-L-CPFB(-2) was required for effective biotinylation by the BirA enzyme. This added purification step by ion exchange chromatography results in a lower protein yield. Second, complete biotinylation of all CPFB(-2) tags present in the sample was not achieved (Figure 4.3), again resulting in lower protein yield. Finally, even when the purification of biotinylation of BLA-L-CPFB(-2) was achieved these proteins tended to precipitate in solution.

As a consequence of the low concentration of biotinylated BLA-L-CPFB(-2) acquired, detailed s-native PAGE titration curves comparing SAVSBPM19 and SAVSBPM32 binding to the biotinylated BLA-L-CPFB(-2) was not obtained. As a result, other comparisons between SAVSBPM19 and SAVSBPM32 were used to reveal potential benefits of one variant over the other. One such comparison tested was the ability of each variant to retain in a tetrameric state under ionic detergent induced protein
denaturation conditions. Samples containing SAVSBPM18, SAVSBPM19, and SAVSBPM32 at 200 ng µl⁻¹ were prepared and analyzed on SDS-PAGE using three conditions. First, in the presence of 0.1% SDS, without the addition of biotin or incubation of the samples in boiling water. Second, in the presence of 0.1% SDS and 2.5 mM biotin, without incubating the samples in boiling water. Third, in the presence of 0.1% SDS, 2.5 mM biotin and incubating the samples in boiling water for 5 minutes prior to electrophoresis.

Figure 4.5 shows monomeric subunits with an apparent molecular mass of ≈ 19 kDa, which is in close agreement with the expected molecular mass of 16.5 kDa. The non-boiled tetrameric versions have an apparent molecular mass of ≈ 100 kDa which is notably larger than the expected molecular mass of 66 kDa. The difference in observed and expected molecular mass of the tetramers is explained by the folded nature of the tetramers which would allow less SDS to bind resulting in less negative charges being added causing a slower migration through the gel compared to denatured subunits.

Results also show no notable difference in the amount of monomeric streptavidin seen for samples containing SAVSBPM18 and SAVSBPM32 in the presence of SDS with or without biotin (≈ 2% total protein as monomers) (Figure 4.5). In the presence of SDS samples containing SAVSBPM19 showed an average amount of monomers approximately 10 fold higher than the averages seen for SAVSBPM18 and SAVSBPM32.
Figure 4.5 Tetrameric state of SAVSBPM18, SAVSBPM19 and SAVSBPM32

Tetrameric and monomeric states of SAVSBPM18, SAVSBPM19 and SAVSBPM32 at 200 ng µl\(^{-1}\) analyzed by SDS-PAGE with coomassie blue staining. Lines under panels A, B, and C indicate samples incubated at 23°C for 30 minutes in the presence of (A) 0.1% SDS, (B) 0.1% SDS, 2.5 mM biotin, or (C) 0.1% SDS, 2.5 mM biotin followed by a 5 minute incubation in a boiling water bath. (E) Amount of monomeric protein detected as a percent of the total protein present for samples incubated in the presence of 0.1% SDS with or without biotin (+ or – Biotin, respectively). Plotted data was obtained from three replicates.
The amount of monomers observed in SAVSBPM19 samples was decreased by an average of 58% when biotin was present, indicating that biotin binding to SAVSBPM19 can increase subunit interactions, which leads to better stability of tetrameric structure (Figure 4.5, panel B). All samples dissociated into monomers when boiled even in the presence of biotin.

The large decrease in tetrameric stability seen in SAVSBPM19 may be due to the location of residue 118 in loop 7-8 of streptavidin. Also in loop 7-8 is Trp-120, this residue is known to be critical in both forming strong biotin binding, and stabilizing the tetramer through subunit-subunit interactions (Chilkoti, Tan, & Stayton, 1995; O’Sullivan et al., 2012; Qureshi, Yeung, Wu, & Wong, 2001; T. Sano & Cantor, 1995). Interestingly, the PyMol molecular graphics system shows that within loop 7-8, residues 115-122 form a α-helix structure (Figure 4.6). If the Asn-118-Cys mutation destabilizes this structure the critical Trp-120 interactions could also be lost, causing the decreased tetrameric stability observed for SAVSBPM19. Although it is currently unclear if the N118C mutation truly destabilized the α-helix, the less favourable α helix propensity of Cys compared to Asn (0.65 and 0.68 kcal mol$^{-1}$, respectively) (Pace & Scholtz, 1998) and the loss of a polar contact between the side chains of Asn-118 and Thr-115 (Figure 4.6) provide some support to this idea.
Helical structure formed by residues 115-122 of streptavidin loop 7-8. (A) Polar contacts (shown as dotted lines) of Asn-118 (yellow) to the side chain of Thr-115 and the main chains of Lys-121 and Thr-115. (B) Polar contacts (shown as dotted lines) of Cys-118 (yellow) to the side chain of Thr-115 and the main chains of Lys-121 and Thr-115. Location of the critical Trp-120 residue (red) involved in subunit-subunit interactions and biotin binding of the neighboring subunit (A and B). Models were generated using PyMOL with the PDB entry 3RY1 (Le Trong et al., 2011) as the starting file.
4.5 Binding kinetics of SAVSBPM32 to BSA(B)

The binding kinetics of purified SAVSBPM32 to biotinylated Bovine Serum Albumin (BSA(B), Biotin-LC-BSA, BioVision) immobilized to Amine Reactive 2nd Generation (AR2G) biosensors (ForteBio) was tested using the Blitz™ label free protein assay system (ForteBio) (Concepcion et al., 2009) following section 2.11.4 methods. SAVSBPM32 concentrations of 157 nM, 117.8 nM, 78.5 nM, and 39.25 nM were analyzed on AR2G-BSA(B) biosensors. $K_{on}$, $K_{off}$, and $K_d$ were calculated using the same methods described in section 3.7.

Results indicate $K_{on}$, $K_{off}$, and $K_d$ values of SAVSBPM32 binding to BSA(B) of $4.12 \times 10^4 \text{ M}^{-1} \text{s}^{-1}$, $5.58 \times 10^{-4} \text{ s}^{-1}$, $1.36 \times 10^{-8} \text{ M}$, respectively (Figure 4.7 Table 4.3).
A

\[ y = 0.000558x \]
\[ R^2 = 0.985040 \]

B

\[ \text{Binding (nm)} \]
\[ \text{Time (s)} \]

C

<table>
<thead>
<tr>
<th>Curve</th>
<th>157 nM</th>
<th>117.8 nM</th>
<th>78.5 nM</th>
<th>39.25 nM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data points</td>
<td>0-200</td>
<td>0-250</td>
<td>50-250</td>
<td>0-250</td>
</tr>
<tr>
<td>Ymax</td>
<td>0.224</td>
<td>0.236</td>
<td>0.162</td>
<td>0.171</td>
</tr>
<tr>
<td>T1</td>
<td>46.4</td>
<td>48.9</td>
<td>50.9</td>
<td>60.8</td>
</tr>
<tr>
<td>( \chi^2/\text{DOF} )</td>
<td>6.35E-05</td>
<td>7.94E-06</td>
<td>1.02E-06</td>
<td>3.73E-06</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.978</td>
<td>0.973</td>
<td>0.957</td>
<td>0.978</td>
</tr>
</tbody>
</table>
Figure 4.7. Determination of the kinetic parameters of the interaction between SAVSBPM32-BSA(B)

Biotinylated BSA was immobilized to biosensor chips and SAVSBPM32 served as the analyte. (A) Linearized data from sensorgrams for the determination of off-rate (slope of the plots). Data plotted in panel A is the average of three replicates. Error bars indicate ± SD. (B) Curve fitting of the sensorgram data for on-rate. (C) Curve fitting was done on SciDAVis (http://scidavis.sourceforge.net) with the equation shown in panel C. Data points = the range of points used in the curve fitting function, numbers correspond the X value of those data points. Ymax = the predicted maximum Y value. T1 is the predicted time for the signal to reach ½ Ymax. \( \chi^2/\text{doF} \) = chi-squared distribution. \( R^2 \) = the coefficient of distribution. (D) Linearized data from sensorgrams for the determination of the on-rate (slope of the plots).
Table 4.3. Kinetic parameters for the interactions between streptavidin variants and the biotin ligands.

<table>
<thead>
<tr>
<th>Analyte</th>
<th>Ligand</th>
<th>On-rate (M⁻¹ s⁻¹)</th>
<th>Off-rate (s⁻¹)</th>
<th>K_d (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>wtSAV</td>
<td>MBP(B)*</td>
<td>5.13x10⁶****</td>
<td>2x10⁻⁷***</td>
<td>4x10⁻¹⁴****</td>
</tr>
<tr>
<td>M18</td>
<td>MBP(B)*</td>
<td>2.3x10⁵***</td>
<td>2.65x10⁻³***</td>
<td>1.15x10⁻⁸***</td>
</tr>
<tr>
<td>M32</td>
<td>BSA(B)**</td>
<td>4.12x10⁴</td>
<td>5.56x10⁻⁴</td>
<td>1.35x10⁻⁸</td>
</tr>
</tbody>
</table>

*Biotinylated maltose binding protein

** Biotinylated bovine serum albumin

***Values reported by (Wu & Wong, 2013)

**** Values reported by (Qureshi & Wong, 2002)

***** Values reported by (Green, 1990)
Chapter 5   Discussion

5.1 Binding stoichiometry of SBPA18C tagged BLA to SAVSBPM32

A major portion of this work involved studying the structure of the complexes $a_1$ and $a_2$, which formed the doublet observed during the binding experiments of BLA-L-SBPA18C to SAVSBPM32F (Figure 3.8). Results from the titration curve shown in Figure 3.8 illustrate that the intensity of the band $a_1$ relative to $a_2$ remains constant throughout the experiment (both increase similarly), suggesting that both $a_1$ and $a_2$ contain the same ratio of BLA-L-SBPA18C coupled to SAVSBPM32F. The novel modified two-dimensional polyacrylamide gel electrophoresis method described in section 2.11.3.1 successfully showed that both $a_1$ and $a_2$ contain the same amount of bound ligand tagged fusion proteins per SAVSBPM32F tetramer (Figure 3.9, panel A). Furthermore, the use of this method combined with non-linear curve fitting suggested that the stoichiometric ratio of these complexes is two BLA-L-SBP18C tagged proteins per streptavidin tetramer (Figure 3.9, panel B and D). The suggested quaternary structure for both complexes was then confirmed using the zinc imidazole staining for complex extraction and analysis methods described in section 3.4.4 (Figure 3.10).

Although the stoichiometry of the complexes was solved, the cause of the two distinct migration patterns observed for complexes containing two BLA-L-SBPA18C proteins linked per SAVSBPM32F tetramer (Figure 3.8) remains unclear. One possible explanation involves the reporter protein, BLA used in this study, which as previously mentioned (Section 1.11) contains a disulfide bond between residues 77 and 123.
As outlined in section 2.11.2.2, samples were first exposed to a reducing environment, which would break the disulfide bonds in the sample. The samples were then exposed to a non-reducing environment, allowing the formation of disulfide bonds and analysed by s-native PAGE. It was assumed that all reduced disulfide bonds between residues 77 and 123 of BLA would reform during the non-reducing exposure step. If this assumption is incorrect and instead, incomplete disulfide bond formation in BLA occurs, two populations of BLA fusion proteins are predicted to exist. These populations would differ only in the presence or absence of the disulfide bond between residues 77 and 123. The reduced form of BLA has been shown to migrate slower during electrophoresis compared to its oxidized counterpart (Kobayashi et al., 1997; Pollitt & Zalkin, 1983). Incomplete disulfide bond formation of BLA residues 77 and 123 would predict multiple migration patterns for complexes containing two BLA-L-SBPA18C proteins linked per SAVSBPM32F tetramer.

The presence of two bands is seen in the BLA fusion protein (Figure 3.8, bands labeled e) initially were thought to support this idea. More detailed predictions made by the described explanation were not supported by the data obtained in Figure 3.8. First, nearly all the BLA-L-SBPA18C protein seen in Figure 3.8 (panel e, lane 14) migrates as a single band. This suggests that if two BLA conformations are present, one population is far more abundant than the other. The similar intensities of bands a₁ and a₂ observed throughout the experiment do not support this observation. Specifically, if the proposed explanation were correct than in order to obtain similar intensities of bands a₁ and a₂, all of the lower band seen in band e should be consumed. Comparison of band e in lanes 14
and 15 of this figure show that this is not the case. The proposed explanation also predicts multiple bands for complexes containing BLA-L-SBPA18C, BLA-L-SBPA18C dimer, tetrameric SAVSBPM32F with one BLA-L-SBPA18C protein linked, and tetrameric SAVSBPM32F with two BLA-L-SBPA18C proteins linked. Data from Figure 3.8 does not support these predictions with the exception of two bands seen in the position corresponding to BLA-L-SBPA18C (Figure 3.8, bands labeled e). It should be noted that protein degradation would also support the presence of the second lower band seen for bands labeled e in Figure 3.8.

Alternatively, the difference in migration ability of complexes $a_1$ and $a_2$ could be explained by variations in the quaternary structure of the complexes. Due to the unique binding of the SBP tag to streptavidin, two distinct conformation of complexes containing two SBP tags bound to streptavidin has previously been suggested (Barrette-Ng et al., 2013). One of these conformations, observed in the crystallized structure of the SBP tag streptavidin complex (Barrette-Ng et al., 2013), occurs when The C and N-terminal ends of one SBP tag bind to subunits C and A of streptavidin, respectively. In this conformation, the C- and N-terminal ends of the second SBP tag bind to subunits D and B on the opposite face of streptavidin (Figure 5.1). The second conformation occurs when one of the SBP tags binds in the opposite orientation, such that the subunits which were previously bound to the N- and C-terminal ends of the SBP tag are reversed (Barrette-Ng et al., 2013) (Figure 5.1). This explanation correctly predicts the dimer observed when two tags are bound to streptavidin (Figure 5.1) and the singlet when one tag and a dimer of tags are formed.
Figure 5.1 Two conformations of SBP streptavidin binding

(A) Streptavidin subunits A (yellow), B (red), C (blue) and D (green) with two bound SBP tags Y and Z (grey and orange). N-terminal ends of the SBP tags are shown in cyan. Two conformations predicted to occur for SBP binding. Top of panel A shows N-terminal end of SBP Y bound to streptavidin subunit A and the N-terminal end of SBP Z bound to subunit D. Bottom of panel A shows the second conformation where one peptide binds in the opposite orientation, here SBP Z is rotated so that the N-terminal end binds to subunit B. (B) Structural differences of the two predicted conformations of binding when using the full length BLA-L-SBP fusion protein. Models were generated using PyMOL with PDB entry 4J06 (Barrette-Ng et al., 2013) as the starting file. Models of the BLA-L-SBP fusion proteins and subsequent binding (B) were designed by and obtained from Dr. Kenneth Ng with permission.
5.2 Interpretation of the binding kinetics data

Some of the data collected from the BLITz experiments show conflicting results compared to the published literature. Specifically, the dissociation of bound complex of BLA-L-SBP to SAVSBPM32 and SAVSBPM18 occurred at approximately half the speed expected when compared to the previously reported off-rate of BLA-L-SBP SAVSBPM18 binding (Wu & Wong, 2013) (Table 5.1). A similar trend is observed when comparing the off-rate of SAVSBPM32 to the immobilized biotinylated BSA and the off-rate previously reported for SAVSBPM18 and biotinylated maltose binding protein (Wu & Wong, 2013) (Table 4.3). One reason that may explain the slow off-rate in SAVSBPM32 is that the biotinylated BSA used in this work has three biotins linked per molecule. The reported values of SAVSBPM18 binding to biotinylated protein used mono-biotinylated MBP (Wu & Wong, 2013). The SAVSBPM32-BSA(B) binding tests may then have been biased by the avidity effect of streptavidin to bind to multiple biotins within the same BSA protein or multiple BSA proteins. However, this reason is unable to explain the slower off-rate of SAVSBPM32 from immobilized BLA-L-SBP.

It is likely that the BLITz system has a greater degree of rebinding during the dissociation steps. In the Blacore system, a buffer is constantly flowed over the sensor chip. Any bound protein that dissociates is then expected to be removed from the system and unable to rebind. The BLITz system does not have a strategy to minimize rebinding (Abdiche, Malashock, Pinkerton, & Pons, 2008).
Table 5.1 Comparison of BLITz and BIAcore kinetic data for the binding of the SBP tag to streptavidin

<table>
<thead>
<tr>
<th>System</th>
<th>Analyte</th>
<th>Ligand</th>
<th>On-rate (M⁻¹ s⁻¹)</th>
<th>Off-rate (s⁻¹)</th>
<th>K_d (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIAcore</td>
<td>M18</td>
<td>SBP*</td>
<td>4.53x10⁴**</td>
<td>7.11x10⁻⁴**</td>
<td>1.57x10⁻⁸**</td>
</tr>
<tr>
<td>BLITz</td>
<td>M18</td>
<td>SBP*</td>
<td>3.32x10⁴</td>
<td>3.34x10⁻⁴</td>
<td>1.03x10⁻⁸</td>
</tr>
</tbody>
</table>

*Indicates BLA-L-SBP
** Data from Wu & Wong, 2013
From the collected data the true binding affinity of SAVSBPM32 to biotin or unmodified SBP tag is currently not clear. However, the ability to elucidate the relative changes in binding affinities by comparing BLITz experiments shows more promise.

The data shows similar binding affinities of the SBP tag to SAVSBPM32 and SAVSBPM18 (Figure 3.18), which correlates well with the similar binding of BLA-L-SBP to SAVSBPM32 and SAVSBPM18 observed during the purification experiments (Figure 3.13). The BLITz data also suggest that BLA-L-SBPA18C has a drastically faster dissociation rate from SAVSBPM18 compared to unmodified BLA-L-SBP (Figure 3.18). Again, this correlates well to the purification experiments, which showed the weakest binding when BLA-L-SBPA18C was purified on the SAVSBPM18 matrix (Figure 3.13).

5.3 Advantages of the SAVSBM32 system

Most engineered streptavidin systems are limited in their flexibility. Some systems offer reversible binding and are reusable, while others specialize in immobilization applications. As a result, experiments can require multiple types of streptavidin variants which can be time consuming and costly. The SAVSBPM32 system allows for users to choose immobilization and reversibility, it is arguably the most flexible streptavidin technology currently reported. Furthermore, the SAVSBPMM32 system can also be applied to purify proteins tagged with the unmodified SBP tag or chemically biotinylated biomolecules. The reusability of this system can reduce the cost associated with its use. The SAVSBP32-affi gel column utilized throughout this work was used over 20 times
and over a four-month period. No noticeable loss in binding has been observed yet. The column was easily regenerated by a simple wash step in PBS buffer.

One of the limitations in previous streptavidin systems is that the matrix cannot distinguish between biotinylated proteins of interest and naturally biotinylated proteins from the cell. As a result *Escherichia coli* and *Bacillus subtilis* expression systems can be contaminated with biotin carboxyl carrier proteins and biotin carboxylase proteins (Cronan Jr. & Reed, 2000; Marini, Li, Gardiol, Cronan, & de Mendoza, 1995). The ability to capture proteins of interest from crude samples via affinity thiol coupling easily avoids this complication by allowing the use of a biotin wash step to remove non-covalently linked biotinylated proteins. Furthermore, the dual binding modes (covalent and non-covalent) of the modified tags forms the basis for a highly flexible and modifiable system. This results in the ability to generate samples of unmatched purity in a single purification step. For example, once the crude sample is loaded to the column, cysteine containing contaminants (such as glutathione) can be removed through a TCEP wash step. Disulfide bonds between the modified tags and SAVSBPM32 are then allowed to form. The user can then remove any other non-specifically and non-covalently bound contaminants from the column using wash buffers that do not contain a reducing agent.

For use in immobilization applications such as biosensors, the covalent bond between the modified tags and SAVSBPM32 will form the basis to generate reliable data, where the same biosensor can be reused by a simple wash step.

In addition to the advantageous described in experimental applications of the SAVSBPM32 system, another potential advantage should be noted. A recently rapidly
emerging field of research is in the use of plant-expression hosts for the production of industrial proteins at low costs (Ma, Drake, & Christou, 2003). Over the past 25 years a variety of useful proteins including antibodies, enzymes and vaccines have been successfully produced in plants (Hiatt, Cafferkey, & Bowdish, 1989; Hood et al., 2003; Mason, Lam, & Arntzen, 1992). Transgenic maize has been used for the production of avidin, which is estimated to reduce the cost of production by ten fold. The tight binding of avidin to biotin with the plant is suggested to be the cause for the sterile male phenotype observed in the transgenic maize (Hood et al., 1997). If produced in plants, the reversible biotin binding characteristic of SAVSBPM32 may eliminate the sterile male phenotype issue while further reducing the cost of associated with the use of streptavidin.

5.4 Possible reasons for the inefficient saturation of SAVSBPM32 with two SBPA18C tags

A recurring question that has been presented throughout this work is the inability to drive all SAVSBPM32F tetramers to covalently link to two SBPA18C tags (section 3.4.2). The decreased affinity to streptavidin caused by the A18C mutation in SBPA18C is likely a factor contributing to this observation (section 3.6). Attempts at increasing the binding of the modified SBP tag through the introduction of a G19A mutation were unsuccessful. It is unclear if the results were caused by changing the model protein from BLA to MBP or if the G19A mutation inhibited binding. One possible explanation is that G19A may increase the length of the α-helix beyond the expected length resulting in the projection of the N-terminal binding sequence away from the binding pocket. Using
alternative residues with lower helical propensities in the G19 position could be used as a way to test this explanation.

An alternative approach that could be used to address this inability to saturate the streptavidin would be to introduce a cysteine residue downstream from the SBP tag. A cysteine residue should also be introduced to streptavidin outside of the SBP tag binding pocket. With these conditions the SBP tag binding should not be affected which may result in the ability to saturate the streptavidin with two covalently attached SBP tags.

It should be noted that other factors could also play a role in the inability to drive the reaction to completion. One such factor could be that streptavidin variants tested were full-length proteins containing both the N-terminal and C-terminal flexible tails. In core streptavidin (residues 13-139), these flexible regions are cleaved and result in higher ligand binding affinity (Le Trong et al., 2006; Takeshi et al., 1995). It would be interesting to see whether production of core SAVSBPM32 would increase the amount of SAVSBM32 to bind to two SBP tags covalently.

5.5 Limitation of the CPFB(-2) SAVBPM32 system and potential solutions

Effective affinity guided thiol coupling between streptavidin and the PFB tag was achieved using either the engineered SAVSBPM32 or SAVSBPM19 variant and CPFB(-2) tag. However, caution should be taken when interpreting the results, which suggest that the SAVSBPM32 variant is more stable as a tetramer when exposed to SDS, compared to SAVSBPM19 (section 4.4). All three trials reported originated from the same stock solution of protein. It is unclear if the differences seen in tetramer stability are
specific to those stock solutions or indicate a larger overall difference in the proteins in general. For this reason the sample size is considered to be one (n = 1) and data points indicate independent replicates (Cumming, Fidler, & Vaux, 2007). Experiments should be done using newly purified SAVSBM19 and SAVSBPM32 to test if the reported results can be replicated.

Although the implications of SAVSBPM32 in both the academic and industrial sectors are truly impressive, the system currently requires intermolecular disulfide bond formation to achieve ultra-tight binding. The vast array of commercially available biotinylated molecules generated through chemical biotinylation therefore cannot be immobilized to SAVSBPM32. The ability to immobilize any biotinylated molecule without the need to use CPFB tag would significantly extend the flexibility of the system. One approach to address this concern is to identify residues within streptavidin that could form an intramolecular disulfide bond upon biotin binding, which would then trap the molecule in the binding pocket. In this system the disulfide bond would only be formed between streptavidin residues, the biotinylated molecule would not be required to contain a cysteine residue. The Disulfide by Design program was used to probe the apo and biotin bound streptavidin structures using the PDB files 3RY1 and 3RY2 (Le Trong et al., 2011) for residues which are predicted to form disulfide bonds in the biotin bound complex but not the apo streptavidin complex (A. A. Dombkowski, 2003; Le Trong et al., 2011). This study identified the potential double mutation, A86C and N49C. The A86C mutation has already been shown throughout this work to be able to covalently link CPFB(-2) and the SBPA18C tag. The suggested N49C mutation is located in loop3-4 of streptavidin. The predicted result is that once biotin binds to streptavidin, the loop closes over the binding
pocket and becomes “locked” via the intramolecular disulfide bond formation and minimizing the chance for biotin to dissociate from the binding pocket (Figure 5.2). In order to make the biotin binding reversible at least one more mutation is needed to reduce the biotin binding capability in the biotin binding pocket of streptavidin. The additional mutation should be carefully selected since any changes in the overall conformation of the binding pocket could interfere with the formation of the intramolecular disulfide bond. The resulting mutein may have the potential to allow the capture of any biotinylated biomolecules (no longer required to be proteins) in a reversible and reusable manner. Additionally, the mutein may retain the ability to form ultra-tight binding to the SBPA18C tag through the formation of the same intermolecular disulfide bond described throughout this work.
Figure 5.2 Modeling the intramolecular disulfide bond of SAVSBPM80

(A) Streptavidin variant with N49C and A86C mutations shown in black. Biotin is shown as a stick model in white. (B) Predicted intramolecular disulfide bond formed between C49 and C86. Panel C and show a surface representation of the protein “locking” biotin (spheres) within the pocket. (D) Completed binding pocket of the tetramer. Models were generated using PyMOL with the PDB files 3RY1 and 3RY2 (Le Trong et al., 2011) used as a starting file.
5.6 Conclusion

Currently, the streptavidin technologies engineered to date are predominantly focused on developing systems that can perform one specific type of task very well. These characteristics are usually either, reversible binding, reusability, or irreversible binding. The limitations of these technologies is the need for multiple systems which can be time consuming and costly. On the contrary, the SAVSBM32 system was engineered based on flexibility.

The SAVSBPM32 variant contains the reversible biotin binding characteristic seen in the biotin analogue systems (2-imonobiotin, N3’ Ethyl biotin, and desthiobiotin), without the disadvantages of irreversibly saturating the streptavidin or requiring harsh elution conditions (Fudem-Goldin & Orr, 1990; Heney & Orr, 1981; Hirsch et al., 2002; Ying & Branchaud, 2011). Unlike the monomeric streptavidin systems, which are prone to aggregation and cannot bind the SBP tag (Barrette-Ng, Wu, Tjia, Wong, & Ng, 2013; Christopher M. Dundas, Demonte, & Park, 2013; Laitinen et al., 1999; Sano, Vajda, Smith, & Cantor, 1997; Sano et al., 1997), SAVSBPM32 is soluble and able to bind the SBP tag.

The SAVSBPM32 system also provides the ability to obtain ultra tight irreversible binding to the engineered CPFB(-2) and SABA18C tags through intermolecular disulfide bond formation. The ability to covalently bind the tags results in immobilization ability that far surpasses those found in the wtSAV, Stv-13, and Traptavidin systems (Chivers et al., 2010; Chivers, Koner, Lowe, & Howarth, 2011; Green, 1990; Piran & Riordan, 1990; Takeshi, Pandori, Xiaomin, Smith, & Cantor,
1995). The immobilization ability of SAVSBPM32 is only comparable to other covalently linked systems, such as the SpyTag, HaloTag, or sortagging (Los et al., 2008; Strijbis, Spooner, & Ploegh, 2012; Veggiani, Zakeri, & Howarth, 2014). The advantage of the SAVSBPM32 system over those is that the covalent bond is reversible, allowing the removal of ligand tags and reuse of the system.

The ability to develop reusable biosensor chips, bioreactors, protein arrays and matrices for affinity purification of proteins can greatly reduce the cost associated with the use of streptavidin technologies. Additionally, the SAVSBPM32 variant retains binding ability to the unmodified SBP tag and biotinylated molecules, making the system exceptionally versatile compared to the currently available streptavidin systems (Table 5.2 and Figure 5.3).

One function missing in SAVSBPM32 is the ability to bind non-cysteine containing biotinylated molecules in both irreversible and reversible manners. This function can potentially be addressed by further engineering the SAVSBPM32 system.
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>wtSAV</th>
<th>M18*</th>
<th>M6**</th>
<th>M32***</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reversibly bind SBP or A18C</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>Immobilize SBP or A18C</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>Reversibly bind PFB or CPFB(-2)</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Immobilize PFB or CPFB(-2)</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Reversibly bind biotinylated molecules</td>
<td>✗</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Immobilize biotinylated molecules</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Reusable</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

*Reported in Wu and Wong, 2013
**Reported in Wu et al., 2009
***Reported in this study
Figure 5.3 Schematic for the binding and elution of BLA-L-SBPA18C and BLA-L-CPFB(-2) from SAVSBPM32

(A) Schematic of the fusion proteins BLA-L-SBPA18C (panel A1 – D1) and BLA-L-CPFB(-2) (panel A2 – D2) and immobilized SAVSBPM32 (shown in blue). (B) Affinity of SAV to biotin and the SBP tag allows for docking and locking (disulfide bond) of the engineered tags to SAVSBPM32. (C) The disulfide bond is broken by the addition of a reducing agent and the fusion proteins are removed from SAVSBPM32 by competition with free biotin. (D) The system is reused by washing the saturated SAVSBPM32 with buffer, which removes biotin from the biotin binding pockets of SAVSBPM32.
REFERENCES


doi:10.1107/S0907444913002576


staining technique. *ELECTROPHORESIS, 19*(14), 2398–2406.
doi:10.1002/elps.1150191407

Frederix, F., Bonroy, K., Reekmans, G., Laureyn, W., Campitelli, A., Abramov, M. A.,
Maes, G. (2004). Reduced nonspecific adsorption on covalently immobilized
protein surfaces using poly(ethylene oxide) containing blocking agents.
*Journal of Biochemical and Biophysical Methods, 58*(1), 67–74.
doi:10.1016/S0165-022X(03)00150-7

columns. In Meir Wilchek and Edward A. Bayer (Ed.), *Methods in Enzymology*

Gaberc-Porekar, V., & Menart, V. (2001). Perspectives of immobilized-metal affinity
chromatography. *Journal of Biochemical and Biophysical Methods, 49*(1), 335–
360.

Gallizia, A., de Lalla, C., Nardone, E., Santambrogio, P., Brandazza, A., Sidoli, A., &
Streptavidin in *Escherichia coli*. *Protein Expression and Purification, 14*(2),

Protein identification and analysis tools on the ExPASy server. In *The
proteomics protocols handbook* (pp. 571–607). Springer. Retrieved from
http://link.springer.com/10.1385/1-59259-890-0:571


Miyamoto, S., & Kollman, P. A. (1993). Absolute and relative binding free energy calculations of the interaction of biotin and its analogs with streptavidin
doi:10.1002/prot.340160303


doi:10.1002/0471142301.ps1914s45


stability and higher accessibility to biotinylated macromolecules. *Journal of Biological Chemistry, 270*(47), 28204–28209. doi:10.1074/jbc.270.47.28204


77-Cys-123 disulfide bond for the folding of TEM-1 β-lactamase.  

*Biochemical Journal, 321*(2), 413–417.


Protein Expression and Purification, 24(3), 348–356.
doi:10.1006/prep.2001.1582

doi:10.1002/prot.22446


Proteins or Biotinylated Proteins. *PLoS ONE, 8*(7), e69530.
doi:10.1371/journal.pone.0069530


A.1.1. Script M18SBP

#Step 1. Load pdb file (r9a.1) of the modeled SAVSBPM18-SBP tag complex (Dr. S.-L. Wong).
load r9a.1.pdb

#Step 2. Show object in cartoon mode.
show cartoon, (chain A, chain B, chain C, chain D, chain Y, chain Z)

#Step 3. Hide object shown in lines and remove water molecules.
hide lines
remove solvent

#Step 4. Show amino acid sequence of file.
set seq_view, 1

#Step 5. Color SAV subunits A, B, C, and D as red, blue, green and purple, respectively.
color red, (chain A)
color blue, (chain B)
color green, (chain C)
color purple, (chain D)

#Step 6. Color SBP tag Y and SBP tag Z as cyan and grey, respectively.
color cyan, (chain Y)
color grey, (chain Z)

#Step 7. Save file titled M1S.1 in pse format.
save M1S.1.pse, /r9a.1
A.1.2. Script M32A18C

#Step 1. Load pdb file (r9a_submit_A86C_A18C) of the modeled SAVM32-SBP A18C tag complex (Dr. S.-L. Wong).
load r9a_submit_A86C_A18C.pdb

#Step 2. Show object in cartoon mode.
show cartoon, (chain A, chain B, chain C, chain D, chain Y, chain Z)

#Step 3. Hide object shown in lines and remove water molecules.
hide lines
remove solvent

#Step 4. Show amino acid sequence of file.
set seq_view, 1

#Step 5. Color SAV subunits A, B, C, and D as red, blue, green and purple, respectively.
color red, (chain A)
color blue, (chain B)
color green, (chain C)
color purple, (chain D)

#Step 6. Color SBP tag Y and SBP tag Z as cyan and grey, respectively.
color cyan, (chain Y)
color grey, (chain Z)

#Step 7. Show disulfide bonds as sticks.
show sticks, (cys/ca+cb+sg)

#Step 8. Show the $\gamma_\alpha$-$\gamma_\gamma$ bond distance in angstroms and calculate $\text{C}_\beta$-$\gamma_\alpha$-$\gamma_\gamma$ bond angle.
distance A/86/SG, Y/18/SG
get_angle A/86/CB, A/86/SG, Y/18/SG

Step 9. Mutate residues S27 and G48 of SAV subunits A, B, C, and D to A27 and T48 respectively.

cmd.wizard("mutagenesis")
cmd.do("refresh_wizard")
cmd.get_wizard().set_mode("ALA")
cmd.get_wizard().do_select("/r9a_submit_A86C_A18C//A/27/ ")
cmd.get_wizard().apply()
cmd.get_wizard().do_select("/r9a_submit_A86C_A18C//B/27/ ")
cmd.get_wizard().apply()
cmd.get_wizard().do_select("/r9a_submit_A86C_A18C//C/27/ ")
cmd.get_wizard().apply()
cmd.get_wizard().do_select("/r9a_submit_A86C_A18C//D/27/ ")
cmd.get_wizard().apply()
cmd.wizard("mutagenesis")
cmd.do("refresh_wizard")
cmd.get_wizard().set_mode("THR")
cmd.get_wizard().do_select("/r9a_submit_A86C_A18C//A/48/ ")
cmd.get_wizard().apply()
cmd.get_wizard().do_select("/r9a_submit_A86C_A18C//B/48/ ")
cmd.get_wizard().apply()
cmd.get_wizard().do_select("/r9a_submit_A86C_A18C//C/48/ ")
cmd.get_wizard().apply()
cmd.get_wizard().do_select("/r9a_submit_A86C_A18C//D/48/ ")

cmd.get_wizard().apply()

#Step 10. Save file titled M2S.1 in pse format.

save M2S.1.pse, /r9a.1_submit_A86C_A18C
A.1.3. *Script M19(-2)*

#Step 1. Load pdb file (*new6pept*) of the modeled SAVM19-CPFB(-2) tag complex (Dr. Ng).

load new6pept.pdb

#Step 2. Show amino acid sequence of file.

set seq_view, 1

#Step 3. Hide the object.

hide everything

#Step 4. Show SAV subunits and the biotinylated nonapeptide in cartoon and stick mode, respectively.

show cartoon, (chain A, chain B, chain C, chain D)
show sticks, chain E

#Step 5. Color SAV subunits A, B, C, and D as red, blue, green and purple, respectively.

color red, (chain A)
color blue, (chain B)
color green, (chain C)
color purple, (chain D)

#Step 6. Color the biotinylated nonapeptide as cyan.

color cyan, (chain E)

#Step 7. Mutate residue N118 of SAV subunit B and S-2 of nonapeptide to cysteine.

cmd.wizard("mutagenesis")
cmd.do("refresh_wizard")
cmd.get_wizard().set_mode("CYS")
cmd.get_wizard().do_select("/new6pept//B/ASN'318/")

cmd.frame(3)

cmd.get_wizard().apply()

cmd.get_wizard().do_select("/new6pept//E/SER'3/")

cmd.frame(1)

cmd.get_wizard().apply()

#Step 8. Create disulfide bond shown as sticks.

bond E/3/SG, B/318/SG

show sticks, (/new6pept//E/3/, /new6pept//B/318/)

#Step 9. Designate region of nonapeptide (biotin) and SAV subunits that are not flexible.

select btn1, E/5/C5+N4+C3+N1+C2+C8+S7+C6+C8+S7+C6+C8+C10

select btn2, E/5/C11+C12+C13+C14+O15+Z+O9+NZ

protect (chain A, chain B, chain C, chain D, btn1, btn2 )

#Step 10. Rotate residues of nonapeptide around the C-Cα-N-C bond of residue 4

(biotinylated lysine chain E).


#Step 11. Show Sγ-Sγ bond distance in angstroms and calculate the Cβ-Sγ-Sγ bond angle.

get_angle E/3/SG, B/318/SG, B/318/CB

distance E/3/SG, B/318/SG

#Step 12. Mutate residue S27 and G48 of SAV subunits to A27 and T48.

cmd.wizard("mutagenesis")

cmd.do("refresh_wizard")
cmd.get_wizard().set_mode("ALA")

cmd.get_wizard().do_select("/new6pept//A/27/")

cmd.get_wizard().apply()

cmd.get_wizard().do_select("/new6pept//B/227/")

cmd.get_wizard().apply()

cmd.get_wizard().do_select("/new6pept//C/27/")

cmd.get_wizard().apply()

cmd.get_wizard().do_select("/new6pept//D/227/")

cmd.get_wizard().apply()

cmd.wizard("mutagenesis")

cmd.do("refresh_wizard")

cmd.get_wizard().set_mode("THR")

cmd.get_wizard().do_select("/new6pept//A/48/")

cmd.get_wizard().apply()

cmd.get_wizard().do_select("/new6pept//B/248/")

cmd.get_wizard().apply()

cmd.get_wizard().do_select("/new6pept//C/48/")

cmd.get_wizard().apply()

cmd.get_wizard().do_select("/new6pept//D/248/")

cmd.get_wizard().apply()

#Step 13. Save file titled M2S.1 in pse format.

save M3P.1.pse, /new6pept
**A.1.4. Script M32(-2)**

**Step 1. Load pdb file (M3P.1) of the modeled SAVM19-CPFB(-2) tag complex.**

load M3P.1.pse

**Step 2. Show amino acid sequence of file, extract and save the SAV subunit B and its bound biotinylated nonapeptide tag in pdb file format.**

set seq_view, 1

select M3P.2, (chain B, chain E)

save M3P.2.pdb, M3P.2

**Step 3. Remove all objects and load pse file (M2S.1) of the modeled SAVSBPM32-SBPA18C tag complex.**

delete all

load M2S.1.pse

**Step 4. Remove both SBP tags and water molecules from the M2S.1 file. Load the pdb file (M3P.2) of the SAV subunit B and its bound biotinylated nonapeptide tag**

remove solvent

remove (chain Y, chain Z)

load M3P.2.pdb

**Step 5. Align SAV subunit B of both objects then create an object containing SAV subunits from M2S.1 and the biotinylated nonapeptide tag from M3P.2.**

select M3P.2.B, /M3P.2//B

align (M3P.2.B), (chain B)

remove /M3P.2//B
create M2P.1, (chain A, chain B, chain C, chain D, chain E,)

show sticks, chain E

#Step 6. Designate region of nonapeptide that is not flexible (biotin).

select btn1, E/5/C5+N4+C3+N1+C2+C8+S7+C6+C8+C10
select btn2, E/5/C11+C12+C13+C14+O15+Z+O9+NZ
create btn, (btn1, btn2)
protect (btn)

#Step 7. Rotate the nonapeptide tag around flexible residues.


#Step 8. Create disulfide bond, measure Sγ-Sγ bond distance in angstroms and calculate the Cβ-Sγ-Sγ bond angle.

create M2P.1, (/r9a_submit_A86C_A18C, /M3P.2)
bond /M2P.1//B/CYS’86/SG, /M2P.1//E/CYS’3/SG
found_it//E/D’d//Le, O5-//89
distance /M2P.1//B/CYS’86/SG, /M2P.1//E/CYS’3/SG
get_angle /M2P.1//B/CYS’86/SG, /M2P.1//E/CYS’3/SG, /M2P.1//E/CYS’3/CB

#Step 9. Save file titled M2P.1 in pse format.

save M2P.1.pse, /M2P.1
A.1.5. Script M96A18C

#Step 1. Load pdb file (M2S.1) of the modeled SAVM32-SBPA18C tag complex

load M2S.1.pse

#Step 2. Mutate residues C86 and A63 of SAV subunits to A86 and C63.

cmd.wizard("mutagenesis")
cmd.do("refresh_wizard")
cmd.get_wizard().set_mode("ALA")
cmd.get_wizard().do_select("/r9a_submit_A86C_A18C//A/CYS`86/"")
cmd.get_wizard().apply()
cmd.get_wizard().do_select("/r9a_submit_A86C_A18C//B/CYS`86/"")
cmd.get_wizard().apply()
cmd.get_wizard().do_select("/r9a_submit_A86C_A18C//C/CYS`86/"")
cmd.get_wizard().apply()
cmd.get_wizard().do_select("/r9a_submit_A86C_A18C//D/CYS`86/"")
cmd.get_wizard().apply()
cmd.wizard("mutagenesis")
cmd.do("refresh_wizard")
cmd.get_wizard().set_mode("CYS")
cmd.get_wizard().do_select("/r9a_submit_A86C_A18C//A/ALA`63/"")
cmd.get_wizard().apply()
cmd.get_wizard().do_select("/r9a_submit_A86C_A18C//B/ALA`63/"")
cmd.get_wizard().apply()

cmd.get_wizard().do_select('''/r9a_submit_A86C_A18C//C/ALA`63/ ''')

cmd.get_wizard().apply()

cmd.get_wizard().do_select('''/r9a_submit_A86C_A18C//D/ALA`63/ ''')

cmd.get_wizard().apply()

*Step 3. Save file titled M9A.1 in pse format.*

save M9A.1.pse, /M2S.1
A.1.5. Script M80B

#Step 1. Retrieve pdb files (3ry1 & 3ry2) of the crystallized structures of apo streptavidin and streptavidin-biotin complexes, retrieved from www.rcsb.org.

fetch 3ry1
fetch 3ry2

#Step 2. Display amino acid sequence of objects, duplicate 3ry2 and align both objects with apo-streptavidin.

set seq_view, 1

copy 3ry2.2, 3ry2

align (/3ry2.2//A), (/3ry1//A)
align (/3ry2//A), (/3ry1//C)

#Step 3. Remove water molecules, hide lines and show objects as cartoon.

remove solvent

hide lines

show cartoon

#Step 4. Mutate residues N49 and A86 to Cys..

cmd.wizard("mutagenesis")

cmd.do("refresh_wizard")

cmd.get_wizard().set_mode("CYS")

cmd.get_wizard().do_select("/3ry1//C/ASN`49/")

cmd.frame(3)

cmd.get_wizard().apply()

cmd.get_wizard().do_select("/3ry1//C/ALA`86/"
cmd.frame(3)

cmd.get_wizard().apply()

cmd.get_wizard().do_select("/3ry2//A/ASN`49/")

cmd.frame(3), (DI-L/L/"0.(n))

cmd.get_wizard().apply()

cmd.get_wizard().apply()

cmd.get_wizard().apply()

cmd.get_wizard().do_select("/3ry2//A/ALA`86/")

cmd.frame(3)

cmd.get_wizard().apply()

#Step 5. Designate residues that are flexible (49 and 86).

select flex, (/3ry2//A/49, /3ry2//A/86 )

protect chain A

deprotect (flex)

#Step 6. Rotate atom SG of C49 about the N-C-Cα-Cβ and C-Cα-Cβ-Cγ bonds.

set_dihedral /3ry2//A/ALA`50/N, /3ry2//A/CYS`49/C, /3ry2//A/CYS`49/CA, /3ry2//A/CYS`49/CB, -165

set_dihedral /3ry2//A/ALA`50/C, /3ry2//A/CYS`49/CA, /3ry2//A/CYS`49/CB, /3ry2//A/CYS`49/SG, -94.5,

#Step 7. Create Cγ-Sγ bond shown in stick mode. Show Cγ-Sγ bond distance in angstroms and calculate Cβ-Cγ-Sγ angle.

distance /3ry2//A/CYS`49/SG, /3ry2//A/CYS`86/SG

bond /3ry2//A/CYS`49/SG, /3ry2//A/CYS`86/SG

show sticks, (CYS/CA+CB+SG)
get_angle 3ry2/A/CYS’86/CB, 3ry2/A/CYS’86/SG, 3ry2/A/CYS’49/SG

#Step 8. Show biotin as stick model.

show sticks, 3ry2/A/BTN

#Step 9. Save file titled shh.1 in pse format.

save M80B.1.pse, /all
A.1.6. Script M96(-2)

#Step 1. Load pdb file (M3P.1) of the modeled SAVM19-CPFB(-2) tag complex.

load M3P.1.pse

#Step 2. N118C mutation in SAV sub B. A63C mutation in SAV sub A, B, C, and D.

cmd.wizard("mutagenesis")
cmd.do("refresh_wizard")
cmd.get_wizard().set_mode("ASN")
cmd.get_wizard().do_select("/new6pept//B/CYS\'318/"")
cmd.get_wizard().apply()
cmd.wizard("mutagenesis")
cmd.do("refresh_wizard")
cmd.get_wizard().set_mode("CYS")
cmd.get_wizard().do_select("/new6pept//A/ALA\'63/")
cmd.get_wizard().apply()
cmd.get_wizard().do_select("/new6pept//B/ALA\'263/ ")
cmd.get_wizard().apply()
cmd.get_wizard().do_select("/new6pept//C/ALA\'63/ ")
cmd.get_wizard().apply()
cmd.get_wizard().do_select("/new6pept//D/ALA\'263/ ")
cmd.get_wizard().apply()

#Step 3. Save file titled M9P.1 in pse format.

save M9P.1.pse, new6pep
**APPENDIX B: DIAGRAMS OF PLASMID VECTORS**

**B.1.1. pWB980 cloning vector**

pWB980 is a 3787 bp pWB110-based expression vector containing the P43 constitutive expression promoter (green), SacB secretion signal peptide (cyan) and selection markers for both kanamycin (pink) and bleomycin (red). Locations of unique restriction sites (black) and the replication protein (blue) are also indicated. Arrows designate the direction of transcription for the specified genes.
pWB980-BLA-L-SBP is a 4708 bp pWB980-based expression vector containing the coding sequence of the BLA-L-SBP recombinant protein (orange), which is constitutively expressed by the P43 promoter (green) and secreted by the SacB signal peptide (cyan). Locations of unique restriction sites, replication protein (blue), kanamycin (pink), and bleomycin (red) selection markers are indicated. Arrows designate the direction of transcription for the specified genes.
**B.1.3. pWB980-BLA-L-SBPA18C**

pWB980-BLA-L-SBPA18C is a 4711 bp pWB980-based expression vector containing the coding sequence of the BLA-L-SBPA18C recombinant protein (orange), which is constitutively expressed by the P43 promoter (green) and secreted by the SacB signal peptide (cyan). Locations of unique restriction sites, replication protein (blue), kanamycin (pink), and bleomycin (red) selection markers are indicated. Arrows designate the direction of transcription for the specified genes.
pWB980-1BLA-LCPFB(-2) is a 4631 bp pWB980-based expression vector containing the coding sequence of the BLA-LCPFB(-2) recombinant protein (orange), which is constitutively expressed by the P43 promoter (green) and secreted by the SacB signal peptide (cyan). Locations of unique restriction sites, replication protein (blue), kanamycin (pink), and bleomycin (red) selection markers are indicated. Arrows designate the direction of transcription for the specified genes.
**B.1.5. pSSAV**

pSSAV is a 4792 bp pWB110-based expression vector containing the coding sequence of the synthetic streptavidin gene (orange), which is constitutively expressed by the P43 promoter (green) and secreted by the SacB signal peptide (cyan). The pSSAV plasmid included both kanamycin (pink) and bleomycin (red) selections markers. The locations of unique restriction sites and the replication protein (blue) are indicated. Arrows designate the direction of transcription for the specified genes.
pSSAVSBPM18 is a 4792 bp pSSAV-based expression vector containing the coding sequence of the synthetic streptavidin mutein SAVSBPM18 (orange), which is constitutively expressed by the P43 promoter (green) and secreted by the SacB signal peptide (cyan). Locations of unique restriction sites, replication protein (blue), kanamycin (pink), and bleomycin (red) selection markers are indicated. Arrows designate the direction of transcription for the specified genes.
pSSAVSBPM19 is a 4792 bp pSSAV-based expression vector containing the coding sequence of the synthetic streptavidin mutein SAVSBPM19 (orange), which is constitutively expressed by the P43 promoter (green) and secreted by the SacB signal peptide (cyan). Locations of unique restriction sites, replication protein (blue), kanamycin (pink), and bleomycin (red) selection markers are indicated. Arrows designate the direction of transcription for the specified genes.
pSSAVSBPM32 is a 4792 bp pSSAV-based expression vector containing the coding sequence of the synthetic streptavidin mutein SAVSBPM32 (orange), which is constitutively expressed by the P43 promoter (green) and secreted by the SacB signal peptide (cyan). Locations of unique restriction sites, replication protein (blue), kanamycin (pink), and bleomycin (red) selection markers are indicated. Arrows designate the direction of transcription for the specified genes.
pSSAVSBPM32F is a 4819 bp pSSAV-based expression vector containing the coding sequence of the synthetic streptavidin mutein SAVSBPM32F (orange), which is constitutively expressed by the P43 promoter (green) and secreted by the SacB signal peptide (cyan). Locations of unique restriction sites, replication protein (blue), kanamycin (pink), and bleomycin (red) selection markers are indicated. Arrows designate the direction of transcription for the specified genes.
pSSAVSBPM80 is a 4792 bp pSSAV-based expression vector containing the coding sequence of the synthetic streptavidin mutein SAVSBPM80 (orange), which is constitutively expressed by the P43 promoter (green) and secreted by the SacB signal peptide (cyan). Locations of unique restriction sites, replication protein (blue), kanamycin (pink), and bleomycin (red) selection markers are indicated. Arrows designate the direction of transcription for the specified genes.
pSSAVSBPM96 is a 4792 bp pSSAV-based expression vector containing the coding sequence of the synthetic streptavidin mutein SAVSBPM96 (orange), which is constitutively expressed by the P43 promoter (green) and secreted by the SacB signal peptide (cyan). Locations of unique restriction sites, replication protein (blue), kanamycin (pink), and bleomycin (red) selection markers are indicated. Arrows designate the direction of transcription for the specified genes.
B.1.12. pET19b

pET19b is a 5717 bp expression vector which carries an N-terminal His tagged (grey) cloning site and ampicillin (pink) selection marker. Expression of proteins between the T7 promoter and terminator are controlled by LacI and LacO regulators. Unique restriction (black) sites are shown and arrows designate the direction of transcription for the specified genes.
B.1.13. pET19b-MBP-L-SBPA18CG19A

pET19b-HF-MBP-L-SBPA18CG19A is a 7193 bp expression vector containing the coding sequence for the N-terminal his tagged (grey) HF-MBP-L-SBPA18CG19A recombinant protein (orange) which is expressed by the T7 promoter (green) under the control of the LacI (blue) and LacO (red) regulators. Unique restriction (black) sites and the ampicillin selection marker are shown and arrows designate the direction of transcription for the specified genes.
<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Recognition Sequence</th>
<th>Enzyme</th>
<th>Recognition Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>AanI</td>
<td>TTATAA</td>
<td>AhlI</td>
<td>ACTAGT</td>
</tr>
<tr>
<td>AarI</td>
<td>CACCTGC</td>
<td>Ajil</td>
<td>CACGTC</td>
</tr>
<tr>
<td>AasI</td>
<td>GACNNNNNGTGC</td>
<td>Ajnl</td>
<td>CCWGG</td>
</tr>
<tr>
<td>AatI</td>
<td>AGGCCT</td>
<td>Ajul-1</td>
<td>GAANNNNNNTTGG</td>
</tr>
<tr>
<td>AatII</td>
<td>GACGTC</td>
<td>Ajul-2</td>
<td>GAANNNNNNTTGG</td>
</tr>
<tr>
<td>AbsI</td>
<td>CCTCGAGG</td>
<td>Alel</td>
<td>CACNNNNGTG</td>
</tr>
<tr>
<td>Acc16I</td>
<td>TCGGCA</td>
<td>Alfl-1</td>
<td>GCANNNNNNTGC</td>
</tr>
<tr>
<td>Acc36I</td>
<td>ACCTGC</td>
<td>Alfl-2</td>
<td>GCANNNNNNTGC</td>
</tr>
<tr>
<td>Acc65I</td>
<td>GGTACC</td>
<td>AloI-1</td>
<td>GAAACNNNNNTCC</td>
</tr>
<tr>
<td>AccB1I</td>
<td>GGYRCC</td>
<td>AloI-2</td>
<td>GAAACNNNNNTCC</td>
</tr>
<tr>
<td>AccB7I</td>
<td>CCANNNNNNTGG</td>
<td>AluI</td>
<td>AGCT</td>
</tr>
<tr>
<td>AccBSI</td>
<td>CCGCTC</td>
<td>Alw21I</td>
<td>GWGCWC</td>
</tr>
<tr>
<td>Acc1</td>
<td>GTMKAC</td>
<td>Alw26I</td>
<td>GTCTC</td>
</tr>
<tr>
<td>AccII</td>
<td>CGCG</td>
<td>Alw44I</td>
<td>GTGCAC</td>
</tr>
<tr>
<td>AccIII</td>
<td>TCCGGA</td>
<td>AlwNI</td>
<td>CAGNNNCTG</td>
</tr>
<tr>
<td>AccII</td>
<td>GTAC</td>
<td>Ama87I</td>
<td>CYCGRG</td>
</tr>
<tr>
<td>AclI</td>
<td>AACGTT</td>
<td>ApaI</td>
<td>GGGCCC</td>
</tr>
<tr>
<td>AclWI</td>
<td>GGATC</td>
<td>ApeK1</td>
<td>GCWGC</td>
</tr>
<tr>
<td>Acol</td>
<td>YGGCCCR</td>
<td>Ascl</td>
<td>GGCGCGCC</td>
</tr>
<tr>
<td>AcsI</td>
<td>RAATTY</td>
<td>Asein</td>
<td>ATTAAT</td>
</tr>
<tr>
<td>AcuI</td>
<td>CTGAAG</td>
<td>AsSI</td>
<td>GCATCGC</td>
</tr>
<tr>
<td>AcvI</td>
<td>CACGTG</td>
<td>Asp700I</td>
<td>GAANNNNTTC</td>
</tr>
<tr>
<td>AcyI</td>
<td>GRGCYC</td>
<td>AspA2I</td>
<td>CCTAGG</td>
</tr>
<tr>
<td>Adel</td>
<td>CACNNNGTG</td>
<td>Aspl</td>
<td>GACNNNGTC</td>
</tr>
<tr>
<td>AflI</td>
<td>GTAC</td>
<td>AspLEI</td>
<td>GCNGC</td>
</tr>
<tr>
<td>AfeI</td>
<td>AGCGCT</td>
<td>AspS9I</td>
<td>GGNCC</td>
</tr>
<tr>
<td>AflII</td>
<td>CCNNNNNNNGG</td>
<td>Assl</td>
<td>AGTACT</td>
</tr>
<tr>
<td>AflIII</td>
<td>CTAAAG</td>
<td>AsuC2I</td>
<td>CCSGG</td>
</tr>
<tr>
<td>AflIII</td>
<td>ACRYGT</td>
<td>AsuHPI</td>
<td>GGTGGA</td>
</tr>
<tr>
<td>AgeI</td>
<td>ACCGGT</td>
<td>AsuII</td>
<td>TTCGAA</td>
</tr>
<tr>
<td>AhdI</td>
<td>GACNNNNGTGC</td>
<td>AsuNHI</td>
<td>GCTAGC</td>
</tr>
<tr>
<td>AvaII</td>
<td>GGWCC</td>
<td>BoxI</td>
<td>GACNNNNGT</td>
</tr>
<tr>
<td>Enzyme</td>
<td>Recognition Sequence</td>
<td>Enzyme</td>
<td>Recognition Sequence</td>
</tr>
<tr>
<td>----------</td>
<td>----------------------</td>
<td>----------</td>
<td>----------------------</td>
</tr>
<tr>
<td>AxyI</td>
<td>CCTNAGG</td>
<td>BplII-2</td>
<td>GAGNNNNNCTC</td>
</tr>
<tr>
<td>BaeGI</td>
<td>GKGCMC</td>
<td>BpmI</td>
<td>CTGGGAG</td>
</tr>
<tr>
<td>BaeI-1</td>
<td>ACNNNNGTAYC</td>
<td>Bpu10I</td>
<td>CCTNAGC</td>
</tr>
<tr>
<td>BaeI-2</td>
<td>ACNNNNGTAYC</td>
<td>BpuEII</td>
<td>CTTGAG</td>
</tr>
<tr>
<td>Ball</td>
<td>TGGCCA</td>
<td>BpvUI</td>
<td>CGATCG</td>
</tr>
<tr>
<td>BamHI</td>
<td>GGATCC</td>
<td>BsaAI</td>
<td>YACGTR</td>
</tr>
<tr>
<td>BanII</td>
<td>GRGCYC</td>
<td>BsaBI</td>
<td>GATNNNNATC</td>
</tr>
<tr>
<td>BanIII</td>
<td>ATCGAT</td>
<td>BsaI</td>
<td>GGTCTC</td>
</tr>
<tr>
<td>BarI-1</td>
<td>GAAGNNNNNTAC</td>
<td>BsaJI</td>
<td>CCNNGG</td>
</tr>
<tr>
<td>BarI-2</td>
<td>GAAGNNNNNTAC</td>
<td>BsaMI</td>
<td>GAATGC</td>
</tr>
<tr>
<td>BauI</td>
<td>CACGAG</td>
<td>BsaWI</td>
<td>WCCGGW</td>
</tr>
<tr>
<td>BbsI</td>
<td>GAAGAC</td>
<td>BsaXI-1</td>
<td>ACNNNNNCTCC</td>
</tr>
<tr>
<td>BbuI</td>
<td>GCATGC</td>
<td>BsaXI-2</td>
<td>ACNNNNNCTCC</td>
</tr>
<tr>
<td>BbvCI</td>
<td>CCTCAGC</td>
<td>Bse118I</td>
<td>RCCGGY</td>
</tr>
<tr>
<td>BbvI</td>
<td>GCAGC</td>
<td>Bse1II</td>
<td>ACTGGG</td>
</tr>
<tr>
<td>BccI</td>
<td>CCATC</td>
<td>Bse3DI</td>
<td>GCAATG</td>
</tr>
<tr>
<td>BceAI</td>
<td>ACGGC</td>
<td>BseGI</td>
<td>GGATG</td>
</tr>
<tr>
<td>BcgI-1</td>
<td>GCANNNNNNTCG</td>
<td>BseMII</td>
<td>CTCAG</td>
</tr>
<tr>
<td>BcgI-2</td>
<td>GCANNNNNNTCG</td>
<td>BsePI</td>
<td>GCGCGC</td>
</tr>
<tr>
<td>BciVI</td>
<td>GTATCC</td>
<td>BseRI</td>
<td>GAGGAG</td>
</tr>
<tr>
<td>BclI</td>
<td>TGATCA</td>
<td>BseX3I</td>
<td>CGGCGC</td>
</tr>
<tr>
<td>BdaI-1</td>
<td>TGANNNNNNTCA</td>
<td>BseYI</td>
<td>CCCAGC</td>
</tr>
<tr>
<td>BdaI-2</td>
<td>TGANNNNNNTCA</td>
<td>BsgI</td>
<td>GTGCAG</td>
</tr>
<tr>
<td>BfaI</td>
<td>CTAG</td>
<td>Bsh1285I</td>
<td>CGRYCG</td>
</tr>
<tr>
<td>BfiI</td>
<td>ACTGGG</td>
<td>BshFI</td>
<td>GGCC</td>
</tr>
<tr>
<td>BfmI</td>
<td>CTRYAG</td>
<td>BsiSI</td>
<td>CCGG</td>
</tr>
<tr>
<td>BfuCI</td>
<td>GATC</td>
<td>BsiWI</td>
<td>CGTACG</td>
</tr>
<tr>
<td>BglI</td>
<td>GCCNNNNNGGC</td>
<td>BslFI</td>
<td>GGGAC</td>
</tr>
<tr>
<td>BglII</td>
<td>AGATCT</td>
<td>BsmBI</td>
<td>CGTCTC</td>
</tr>
<tr>
<td>BisI</td>
<td>GCNGC</td>
<td>Bsp1286I</td>
<td>GDGCHC</td>
</tr>
<tr>
<td>Blpl</td>
<td>GCTNAGC</td>
<td>Bsp1407I</td>
<td>TGTACA</td>
</tr>
<tr>
<td>Bme1390I</td>
<td>CCNGG</td>
<td>Bsp19I</td>
<td>CCATGG</td>
</tr>
<tr>
<td>BmlI</td>
<td>GGNNCC</td>
<td>Bsp68I</td>
<td>TCGCGA</td>
</tr>
<tr>
<td>BplI-1</td>
<td>GAGNNNNNCTC</td>
<td>BspHI</td>
<td>TCATGA</td>
</tr>
<tr>
<td>Enzyme</td>
<td>Recognition Sequence</td>
<td>Enzyme</td>
<td>Recognition Sequence</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------</td>
<td>----------</td>
<td>----------------------</td>
</tr>
<tr>
<td>BspQI</td>
<td>GCTCTTC</td>
<td>Eco32I</td>
<td>GATATC</td>
</tr>
<tr>
<td>BssNAI</td>
<td>GTATAC</td>
<td>Eco57MI</td>
<td>CTGRAG</td>
</tr>
<tr>
<td>BssT1I</td>
<td>CCWWGG</td>
<td>EcoRI</td>
<td>GAATTC</td>
</tr>
<tr>
<td>Bst4CI</td>
<td>ACNWGT</td>
<td>EcoT22I</td>
<td>ATGCAT</td>
</tr>
<tr>
<td>Bst6I</td>
<td>CTTTCT</td>
<td>Fall-1</td>
<td>AAGNNNNNCTT</td>
</tr>
<tr>
<td>BstAPI</td>
<td>GCANNNNNNTGC</td>
<td>Fall-2</td>
<td>AAGNNNNNCTT</td>
</tr>
<tr>
<td>BstC8I</td>
<td>GCNNGC</td>
<td>Faul</td>
<td>CCGGC</td>
</tr>
<tr>
<td>BstDEI</td>
<td>CTNAG</td>
<td>FauND1</td>
<td>CATATG</td>
</tr>
<tr>
<td>BstDSI</td>
<td>CCRYGG</td>
<td>Fnu4HI</td>
<td>GCNGC</td>
</tr>
<tr>
<td>BstEI1</td>
<td>GGTNACC</td>
<td>FseI</td>
<td>GCGCCGGC</td>
</tr>
<tr>
<td>BstENI</td>
<td>CCTNNNNNNAGG</td>
<td>FspAI</td>
<td>RTGCGCAY</td>
</tr>
<tr>
<td>BstH2I</td>
<td>RGCGCY</td>
<td>Glal</td>
<td>GCGC</td>
</tr>
<tr>
<td>BstMWI</td>
<td>GCNNNNNNNNGC</td>
<td>Hin4I-1</td>
<td>GAYNNNNNVTC</td>
</tr>
<tr>
<td>BstNSI</td>
<td>RCATGY</td>
<td>Hin4I-2</td>
<td>GAYNNNNNVTC</td>
</tr>
<tr>
<td>BstSNI</td>
<td>TACGTA</td>
<td>HinclII</td>
<td>GTYRAC</td>
</tr>
<tr>
<td>BstX2I</td>
<td>RGATCY</td>
<td>HindIII</td>
<td>AAGCTT</td>
</tr>
<tr>
<td>BstXI</td>
<td>CCANNNNNNTGG</td>
<td>Hinfl</td>
<td>GANTC</td>
</tr>
<tr>
<td>BtgZI</td>
<td>GCGATG</td>
<td>Hpal</td>
<td>GTTAAC</td>
</tr>
<tr>
<td>BtsI</td>
<td>GCAGTG</td>
<td>Hpy188I</td>
<td>TCNGA</td>
</tr>
<tr>
<td>CciNI</td>
<td>GCGGCCGC</td>
<td>Hpy188III</td>
<td>TCNNGA</td>
</tr>
<tr>
<td>Cfr42I</td>
<td>CCGCGG</td>
<td>Hpy8I</td>
<td>GTNNAC</td>
</tr>
<tr>
<td>Cfr9I</td>
<td>CCGCGG</td>
<td>Hpy99I</td>
<td>CGWCG</td>
</tr>
<tr>
<td>CpoI</td>
<td>CGGWCCG</td>
<td>HpyAV</td>
<td>CCTTC</td>
</tr>
<tr>
<td>CseI</td>
<td>GACGC</td>
<td>HpyCH4IV</td>
<td>ACGT</td>
</tr>
<tr>
<td>CspCl-1</td>
<td>CAANNNNNGTGG</td>
<td>HpyCH4V</td>
<td>TGCA</td>
</tr>
<tr>
<td>CspCl-2</td>
<td>CAANNNNNNGTG</td>
<td>Lwel</td>
<td>GCATC</td>
</tr>
<tr>
<td>CviAll</td>
<td>CATG</td>
<td>MabI</td>
<td>ACCWGGT</td>
</tr>
<tr>
<td>CviJ1</td>
<td>RGCY</td>
<td>MaeIII</td>
<td>GTTNAC</td>
</tr>
<tr>
<td>DinI</td>
<td>GGCGCC</td>
<td>MauB1</td>
<td>CGCGCGC</td>
</tr>
<tr>
<td>DpnI</td>
<td>GATC</td>
<td>MboII</td>
<td>GAAGA</td>
</tr>
<tr>
<td>DraI</td>
<td>TTTAAA</td>
<td>Mfel</td>
<td>CAATTG</td>
</tr>
<tr>
<td>DraII</td>
<td>RGGNCCY</td>
<td>MluL</td>
<td>ACGCGT</td>
</tr>
<tr>
<td>EciI</td>
<td>GGCGGA</td>
<td>MlyI</td>
<td>GAGTC</td>
</tr>
<tr>
<td>Ecl136II</td>
<td>GAGCTC</td>
<td>Mmel</td>
<td>TCCRAC</td>
</tr>
<tr>
<td>Enzyme</td>
<td>Recognition Sequence</td>
<td>Enzyme</td>
<td>Recognition Sequence</td>
</tr>
<tr>
<td>----------</td>
<td>----------------------</td>
<td>----------</td>
<td>----------------------</td>
</tr>
<tr>
<td>MnlI</td>
<td>CCTC</td>
<td>SanDI</td>
<td>GGGWCCC</td>
</tr>
<tr>
<td>MreI</td>
<td>CGCCGGCG</td>
<td>SbfI</td>
<td>CCTGCAGG</td>
</tr>
<tr>
<td>MroNI</td>
<td>GCCCGGC</td>
<td>SetI</td>
<td>ASST</td>
</tr>
<tr>
<td>MseI</td>
<td>TTAA</td>
<td>SfiI</td>
<td>GGCCNNNNNGGCC</td>
</tr>
<tr>
<td>MsiII</td>
<td>CAYNNNNRTG</td>
<td>SgrAI</td>
<td>CRCCGGYG</td>
</tr>
<tr>
<td>MspA1I</td>
<td>CMGCKG</td>
<td>SgrDI</td>
<td>CGTCGACG</td>
</tr>
<tr>
<td>MsiI</td>
<td>GTTTAAAC</td>
<td>SmiI</td>
<td>ATTTAAAT</td>
</tr>
<tr>
<td>NmeAII</td>
<td>GCCGAG</td>
<td>SmII</td>
<td>CTYRAG</td>
</tr>
<tr>
<td>NmuCl</td>
<td>GTSAC</td>
<td>SfiI</td>
<td>GCCCGGGGC</td>
</tr>
<tr>
<td>PacI</td>
<td>TTAATTAA</td>
<td>Sse9I</td>
<td>AATT</td>
</tr>
<tr>
<td>PaeR7I</td>
<td>CTCGAG</td>
<td>SspI</td>
<td>AATATT</td>
</tr>
<tr>
<td>PasI</td>
<td>CCCWGGG</td>
<td>TaqI</td>
<td>TCGA</td>
</tr>
<tr>
<td>PciI</td>
<td>ACATGT</td>
<td>TaqII-1</td>
<td>GACCCGA</td>
</tr>
<tr>
<td>PfeI</td>
<td>GAWTC</td>
<td>TaqII-2</td>
<td>CACCCA</td>
</tr>
<tr>
<td>PfoI</td>
<td>TCCNGGA</td>
<td>TatI</td>
<td>WGTACW</td>
</tr>
<tr>
<td>Ppii-1</td>
<td>GAACNNNNNCTC</td>
<td>TauI</td>
<td>GCSGC</td>
</tr>
<tr>
<td>Ppii-2</td>
<td>GAACNNNNNCTC</td>
<td>TscAI</td>
<td>CASTGNN</td>
</tr>
<tr>
<td>PpuMI</td>
<td>RGGWCCY</td>
<td>TsoI</td>
<td>TARCCA</td>
</tr>
<tr>
<td>PspXI</td>
<td>VCTCGAGB</td>
<td>TspDTI</td>
<td>ATGAA</td>
</tr>
<tr>
<td>PsrI-1</td>
<td>GAACNNNNNTAC</td>
<td>TspGW1</td>
<td>ACGGA</td>
</tr>
<tr>
<td>PsrI-2</td>
<td>GAACNNNNNNTAC</td>
<td>Tstl-1</td>
<td>CACNNNNNNTCC</td>
</tr>
<tr>
<td>PstI</td>
<td>CTGCAG</td>
<td>Tstl-2</td>
<td>CACNNNNNNTCC</td>
</tr>
<tr>
<td>PvulII</td>
<td>CAGCTG</td>
<td>Xbal</td>
<td>TCTAGA</td>
</tr>
<tr>
<td>SalI</td>
<td>GTCGAC</td>
<td>XcmI</td>
<td>CCANNNNNNNNTGG</td>
</tr>
</tbody>
</table>

* Non-standard bases
  B = C, G, or T
  H = A, C, or T
  K = G or T
  M = A or C
  N = A, C, G or T
  R = A or G
  S = C or G
  V = A, C, or G
  W = A or T
  Y = C or T