University of Oxford

Thesis submitted for the degree

Doctor of Philosophy in Structural Biology

Structural Approaches Towards Preventing Host Cell Cytoadhesion and Invasion by the Malaria Parasite



Natalie Michaela Ena Barber

Brasenose College

Professor Matthew Higgins

Abstract

Malaria, a disease caused by *Plasmodia* parasites and transmitted by mosquitoes, causes the deaths of 500,000 people each year. Despite advances in malarial therapeutics, there is still no effective vaccine. Potential targets for vaccines are those involved in erythrocyte invasion or those expressed by parasite-infected erythrocytes that lead them to cytoadhere to endothelial surfaces. The three targets discussed in this thesis are 1) *Plasmodium falciparum* erythrocyte membrane protein 1 (*Pf*EMP1) involved in binding to EPCR, 2) *Plasmodium vivax* Duffy binding protein (*Pv*DBP) required for *P. vivax* invasion into erythrocytes, and 3) a *Pf*EMP1 called Var2CSA, which mediates infected erythrocyte cytoadherence to the placenta. Recently, structure-based vaccine design has successfully been developed. The structures of targets can be used to elucidate the molecular mechanisms by which they interact with the human receptors, as well as to identify epitopes bound by inhibitory antibodies. These regions can then be introduced onto smaller scaffold proteins, thereby focusing the immune response to produce antibodies that block the ability of the parasite protein to interact with its human receptors. Here, I demonstrate, through X-ray crystallography and binding studies, that functional sites and epitopes can be introduced onto smaller proteins whilst still retaining the function of the full-length domains. However, immunisation experiments with small immunogens that contain just the receptor binding site of a *Pf*EMP1 suggested that inhibitory epitopes do not necessarily overlap with the functional binding sites. This was supported by solving the structure of an invasion-blocking antibody in complex with *Pv*DBP, which identified that the epitope was distal to the receptor binding site. Additionally, the challenge to structurally characterise the multidomain binding site and epitopes of Var2CSA, by X-ray crystallography and cryoEM, is highlighted. Overall, this thesis discusses how the structural characterisation of vaccine candidates against malaria can be used to guide future immunogen design.

Acknowledgements

I'm very grateful to the ever-changing, ever-growing Higgins group for providing a fun and focused work environment. I would first like to thank my supervisor, Professor Matthew Higgins, who has been an invaluable source of wisdom and advice throughout my DPhil. From crystal freezing to planning experiments, Matt was always available to guide me when I needed it, and I am very grateful. I would especially like to mention Harriet, Kate and Tom, who all made my time in lab very enjoyable, as well as being sounding boards for ideas and giving excellent advice. In addition, I would like to thank Jon Bartlett for keeping the lab running smoothly and helping my S2 cells to thrive. I'm also grateful to the three students that I had over my DPhil: Georgina Benn, James Nugent and Mateusz Kotowski.

I would also like to thank our collaborators, both from Simon Draper's group (The Jenner Institute, University of Oxford) and Thomas Lavstsen's group (University of Copenhagen). Both groups carried out the immunisations, subsequent antibody isolation and characterisations, providing me with multiple inhibitory monoclonal antibodies to perform structural studies. In addition, I would like to thank David Staunton for help and advice in Biophysical Suite, as well as Ed Lowe, who was always around to help with the trickiest crystallography problems. Moreover, thanks to Justin Deme and Susan Lea, who helped with cryo-EM data processing.

I'd like to thank all of my friends for all of the fun-filled times we've had over the past four years. In particular, I'd like to thank Phil and August for being my Biochemistry buddies from the beginning, and keeping me somewhat sane! Last but not least I would like to thank my family, in particular my Mum, who has always been there for me, supporting me through the toughest of times and without whom none of this would have been possible.

Contents

Abstract i

Acknowledgements ii

Contents iii

List of Figures xi

List of Tables xviii

Abbreviations xx

1 General Introduction 1

1.1 An Introduction to Malaria 1

1.1.1 History 1

1.1.2 The Disease 2

1.1.3 The *Plasmodium* parasite 3

1.1.4 *Plasmodium falciparum* life cycle 4

1.1.5 Differences in *P. vivax* Life Cycle compared to *P. falciparum* . 8

1.2 Disease and Pathology 12

1.2.1 Malaria Pathology and Symptoms 12

1.2.2 Natural Acquired Immunity 14

1.2.3 Polymorphisms, allelic switching, and redundancy 15

1.2.4 Genetic Resistance 16

1.3 Current Interventions and Treatments 16

1.3.1 Vector Control 16

1.3.2 Drugs 18

1.3.3 Vaccine Candidates 19

1.4 Structure-guided Immunogen Design 21
1.5 Thesis Outline 23

2 Materials and Methods **25**

2.1 Cloning 25
2.1.1 Polymerase Chain Reaction 25
2.1.2 Restriction Enzyme Digests 25
2.1.3 Ligation 26
2.1.4 Transformation of ligation reactions 26
2.1.5 Site Directed Mutagenesis 27
2.1.6 Gibson Assembly Cloning 27
2.1.7 Agarose Gel Electrophoresis 28
2.2 Protein Expression 28
2.2.1 Bacterial Expression in *Escherichia coli* cells 28
2.2.2 S2 expression 30
2.2.3 Lepidopteran cell/baculovirus expression 32
2.2.4 Human embryonic kidney (HEK) cells expression 33
2.3 Protein Analysis 35
2.3.1 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE) 35
2.3.2 Western Blot 36
2.4 Protein Purification 37
2.4.1 Sonication 37
2.4.2 Tangential flow filtration 37
2.4.3 Ni-NTA affinity Chromatography 38
2.4.4 Insoluble preparations 38
2.4.5 TEV protease cleavage 40
2.4.6 Size exclusion chromatography 40
2.4.7 Protein Deglycosylation 41

2.4.8 Protein Biotinylation 41
2.4.9 Protein Methylation 42
2.4.10 Trimming of the CSA Polysaccharide 42
2.5 Antibody Purification 43
2.5.1 Monoclonal Antibody Isotyping 43
2.5.2 Monoclonal Antibody Fragment Production 43
2.6 Biophyscial Methods 45
2.6.1 Circular Dichroism 45
2.6.2 Surface Plasmon Resonance 46
2.7 Crystallographic Methods 48
2.7.1 Background 48
2.7.2 Protein Crystallisation 49
2.7.3 Harvesting and Cryo-Protection of Crystals 50
2.7.4 Data Collection 50
2.7.5 Data Processing and Refinement 51
2.7.6 Solving the Phase Problem Using Molecular Replacement . . . 53
2.7.7 Model Rebuilding and Refinement 54
2.8 Electron Cryo-Microscopy Methods 55
2.8.1 Grid Preparation 56
2.8.2 Grid Screening 59
2.8.3 Data Acquisition 60
2.8.4 Data Processing 60
2.9 Immunogen Design 62
2.9.1 Protein Repair One Stop Shop (PROSS) 62

3 Targeting Cytoadherence in Severe Malaria by Blocking EPCR-binding *Pf*EMP1s 64

3.1 Introduction to *Pf*EMP1s 64

3.1.1 Antigenic Variation 65
3.1.2 Architecture 65
3.1.3 Endothelial Binding Partners 68
3.1.4 The Interaction of *Pf*EMP1 with EPCR 70
3.1.5 Design of an Immunogen to Block *Pf*EMP1:EPCR Interaction 75
3.2 CIDRα1 Immunogen Characterisation 81
3.2.1 Project Aims 81
3.2.2 Declaration 81
3.2.3 Designing Panel of Immunogen Bundles to Represent Sequence Diversity 81
3.2.4 Bundle Expression and Purification 83
3.2.5 EPCR Expression and Purification 87
3.2.6 Assessing the Fold of the Bundle Immunogens 88
3.2.7 Binding of Bundles to Recombinant EPCR 90
3.2.8 Structural Basis of the Bundle:EPCR Interaction 94
3.2.9 Analysis of the Crystal Structures 105
3.2.10 Bundles as Immunogens 109
3.2.11 Declaration 109
3.2.12 Discussion 123
3.3 Structural Studies of Mouse mAbs Against CIDRα1s 128
3.3.1 Project Aim 128
3.3.2 Declaration 128
3.3.3 Monoclonal Antibody Fab Fragment Generation 129
3.3.4 CIDRα1 Cloning, Expression and Purification 130
3.3.5 Assessing the Fold of the CIDRα1 Domains 133
3.3.6 SPR of CIDRα1 Binding to EPCR 134
3.3.7 Binding of the CIDRα1 to mAbs 136

3.3.8 Competition SPR of CIDRα1 binding to EPCR in presence of inhibitory mAb . . . 137
3.3.9 Attempting to Crystallise the CIDRα1:Fab complexes . . . 140
3.3.10 Design of Potentially More Crystallisable CIDRα1 Constructs 146
3.3.11 Discussion . . . 156

4 Targeting Invasion in *Plasmodium vivax* Malaria via *Pv*DBP 160

4.1 Reticulocyte Invasion by *P. vivax* . . . 160
4.1.1 *Plasmodium vivax* Duffy Binding Protein . . . 163
4.1.2 DARC . . . 165
4.1.3 DARC:*Pv*DBP Interaction . . . 166
4.1.4 *Pv*DBP-RII as a Vaccine Candidate . . . 169
4.1.5 Pre-clinical and Clinical Trials Using *Pv*DBP-RII . . . 170
4.1.6 Studies of Inhibitory Monoclonal Antibodies Against *Pv*DBP-RII 170
4.1.7 Identification of DB9 as an Inhibitory Human mAb . . . 172
4.2 Structural Characterisation of *Pv*DBP-RII:DB9Fab . . . 174
4.2.1 Project Aims . . . 174
4.2.2 Declaration . . . 174
4.2.3 Expression and Purification of *Pv*DBP-RII . . . 174
4.2.4 Preparation of DB9 Fab Fragment . . . 176
4.2.5 Assessing the Binding of *Pv*DBP-RII to DB9 . . . 177
4.2.6 Crystallisation of the *Pv*DBP-RII:DB9Fab Complex . . . 178
4.2.7 Data Collection and Solution of the *Pv*DBP-RII:DB9-Fab Complex . . . 179
4.2.8 Analysis of the PvDBP-RII:DB9Fab structure . . . 181
4.2.9 Discussion . . . 186
4.3 Triple Complex Formation . . . 188
4.3.1 Rational . . . 188

4.3.2 Project Aims . . . 188
4.3.3 Declaration . . . 189
4.3.4 DARC Ectodomain Cloning, Expression and Purification . . . 189
4.3.5 Crystallisation of DARC:PvDBP-RII:DB9Fab Triple Complex 191
4.3.6 Discussion of the Triple Complex Attempt . . . 196
4.4 Structure-Guided Immunogen Design . . . 198
4.4.1 Rational . . . 198
4.4.2 Project Aims . . . 198
4.4.3 Declaration . . . 198
4.4.4 Design of Subdomain3-based Immunogens . . . 199
4.4.5 Cloning, Expression and Purification of Designed Immunogens 209
4.4.6 Assessing the Fold of the Immunogens . . . 212
4.4.7 Binding of the Immunogens to DB9 . . . 214
4.4.8 Structural Validation of the Designed Immunogens . . . 216
4.4.9 Immunisations of Rabbits with Designed Immunogens . . . 223
4.4.10 Discussion of Immunogen Design . . . 232

5 Structural Studies of Var2CSA 235

5.1 Introduction to PAM and Var2CSA . . . 235
5.1.1 Pregnancy-associated Malaria . . . 235
5.1.2 Var2CSA, an unusual *Pf*EMP1 . . . 236
5.1.3 Var2CSA Binding to the Carbohydrate CSA . . . 237
5.1.4 Structural Investigations of Var2CSA Domains . . . 238
5.1.5 Sequence Conservation of Var2CSA . . . 239
5.1.6 Var2CSA as a Vaccine . . . 240
5.1.7 Multimeric IgM Binding to Var2CSA . . . 242
5.1.8 Initial Structural Studies of Full-length Var2CSA . . . 243
5.2 Structural Studies of Full Length Var2CSA . . . 244

5.2.1 Rational . . . 244
5.2.2 Project Aims . . . 244
5.2.3 Declaration . . . 245
5.2.4 Var2CSA construct expression and purification . . . 245
5.2.5 Attempts to Crystallise Var2CSA . . . 246
5.2.6 Attempting Full-length Var2CSA Structure Determination by CryoEM . . . 251
5.2.7 Ongoing and Future Directions . . . 267
5.3 Discussion . . . 279

6 Concluding Remarks **282**
6.1 Summary of Main Findings from this Thesis . . . 282
6.1.1 Characterisation of Immunogens Against EPCR-binding CIDRα1 Domains . . . 282
6.1.2 The Molecular Characterisation of the Epitope of an Inhibitory Human mAb Against *P. vivax* Invasion and Subsequent Epitope-Focused Vaccine Design . . . 284
6.1.3 Towards the Structure of Var2CSA . . . 285
6.2 Towards a Vaccine Against Malaria . . . 286

A Appendix **288**
A.1 CIDRα1 and Bundle Sequences and Primers . . . 288
A.1.1 Sequences of CIDRα1s on which Bundles were Designed . . . 288
A.1.2 GeneString Sequences for Bundles . . . 290
A.1.3 Primers for Cloning Bundles into pET15bcon . . . 292
A.1.4 Protein Sequences for Bundles . . . 293
A.2 Sequences and Primers for CIDRα1 Constructs . . . 294
A.2.1 Sequences for CIDRα1 Constructs . . . 294

A.2.2 Primers for CIDRα1 Constructs 300
A.2.3 Primers for CIDRα1 Minus Beta, Minus Loop Constructs . . . 301
A.3 Sequences and Primers for *Pv*DBP-RII Project 302
A.3.1 DBP Sequence 302
A.3.2 DARC Construct 302
A.3.3 Immunogen Primers 303
A.4 Gene Identifiers for DBP Sequences in PROSS Alignment 303
A.5 Var2CSA New Constructs Primers 304

Publications **305**

Bibliography **306**

List of Figures

1.1 **Malaria Cases Between 2010 and 2017** . . . 3

1.2 **World map showing the incidence of *P. falciparum* compared to *P. vivax*** . . . 4

1.3 ***Plasmodium* parasite life cycle** . . . 5

1.4 **Features of the *P. vivax* life cycle compared to *P. falciparum*** 9

2.1 **pET15bcon vector map** . . . 29

2.2 **pExpreS^2-1 vector map** . . . 31

2.3 **pAcGP67A vector map** . . . 33

2.4 **pHLsec vector map** . . . 35

3.1 **Architecture of *Pf*EMP1 Proteins** . . . 66

3.2 **Domain Cassettes Associated with Severe Malaria** . . . 67

3.3 ***Pf*EMP1 Endothelial Binding Partners** . . . 69

3.4 **Phylogenetic Tree Showing the Different Subclasses of CIDRα1 Domains** . . . 71

3.5 **Proposed Model of EPCR-binding *Pf*EMP1 proteins Modulating the Immune Response** . . . 72

3.6 **CIDRα1:EPCR Structure** . . . 74

3.7 **CIDRα1.4 EPCR-binding Residues** . . . 76

3.8 **Grafting of Epitope on Helical Scaffold** . . . 78

3.9 **Rounds of Bundle Design** . . . 79

3.10 **Bundle Sequence Alignment** . . . 83

3.11 **Purification of Bundle Constructs** . . . 85

3.12 **Purification of Bundle Constructs Continued** . . . 86

3.13 **Purification of EPCR** . . . 87

3.14 **Deglycosylation of EPCR** . . . 88

3.15 **CD Spectra of Bundles** . . . 89
3.16 **CD melting curves of Bundles at 220nm** . . . 90
3.17 **SPR curves showing binding of Bundles to EPCR** . . . 91
3.18 **Bundle1.7:EPCR and Bundle1.8:EPCR Attempted Complex Formation** . . . 93
3.19 **Bundle:EPCR Complex Formation** . . . 95
3.20 **Methylation of Bundle1.4** . . . 96
3.21 **Silver stain of Bundle1.4:EPCR crystals** . . . 97
3.22 **Asymmetric Unit and Crystal Packing in Bundle1.4:EPCR Crystal** . . . 100
3.23 **Ramachandran Plot of Bundle1.4:EPCR Structure** . . . 101
3.24 **Ramachandran Plots for Bundle:EPCR Structures** . . . 104
3.25 **Comparison of Bundle1.4 to Rosetta Model** . . . 106
3.26 **Comparison of Bundle1.4:EPCR and CIDRα1.4:EPCR Structures** . . . 107
3.27 **Comparison of Different Bundles** . . . 108
3.28 **Binding of Bundle1.4 to Antibodies from Immunised Rats or from Humans** . . . 110
3.29 **Quantification of Binding by Bundle1.4 to Antibodies from Immunised Rats** . . . 111
3.30 **Binding by Bundle1.4 to Antibodies from Humans in Malaria Endemic Area** . . . 113
3.31 **Immunogenicity of Bundle1.4** . . . 114
3.32 **Cross-reactivity of Antibodies Generated from Immunisation with Bundle1.4** . . . 116
3.33 **SpyCatcher:SpyTag Coupling** . . . 118
3.34 **Immunogenicity and Cross-reactivity of Antibodies Generated from Immunisation with Bundle1.4-VLP** . . . 119

3.35 **Immunogenicity of a Mixture of Bundles Attached to VLPs** . 120
3.36 **Cross-Reactivity from Immunisations with all Bundles attached to VLP** . . . 122
3.37 **EPCR-Binding Site Sequence for the Different Bundles** . . . 124
3.38 **Interactions with EPCR of Residues Mutated in Bundle1.7** . 125
3.39 **Ficin Cleavage of Mouse mAbs** . . . 130
3.40 **An Alignment of the CIDRα1 Sequences Compared to already Crystallised Constructs** . . . 131
3.41 **CIDRα1 Purification** . . . 133
3.42 **CD Spectra of CIDRα1 Domains** . . . 134
3.43 **Binding of CIDRα1 Domains to EPCR** . . . 135
3.44 **Binding of the CIDRα1 domains to Antibodies** . . . 137
3.45 **EPCR Binding Competition Between CIDRα1 domains both in Presence and Absence of mAbs** . . . 139
3.46 **CIDRα1:Fab Complex Purification** . . . 141
3.47 **HB3var03:Anker Microcrystals** . . . 142
3.48 **IT4var20:Jan Initial Crystals and Optimisation** . . . 143
3.49 **Anisotropy in IT4var20:Jan Dataset** . . . 145
3.50 **Overview of the minus β and minus loop constructs** . . . 147
3.51 **Linker length in Minus Loop Design** . . . 148
3.52 **Summary of HB3var03 and IT4var20 Constructs** . . . 149
3.53 **Purification of IT4var20$_{-loop}$, HB3var03$_{-loop}$, IT4var20$_{-loop,\beta}$ and HB3var03$_{-loop,\beta}$** . . . 151
3.54 **CD of Minus Loop and/or Beta Constructs** . . . 152
3.55 **SPR curves showing binding of CIDRα1 constructs binding to EPCR** . . . 153
3.56 **SPR curves showing binding of CIDRα1 constructs binding to their mAb** . . . 155

3.57 **Purification of IT4var20$_{-loop,\beta}$:JanFab Complex** 156

4.1 **Comparison of Red Blood Cell Invasion Pathways in *P. falciparum* and *P.vivax*** 161

4.2 ***DBP* Domain Architecture** 164

4.3 ***Pv*DBP-RII Structure** 164

4.4 **Schematic of DARC Architecture** 165

4.5 **Heterotetramer of *Pv*DBP-RII:DARC$_{19-30}$** 167

4.6 **Electron Density Map for the DARC Fragment** 168

4.7 **Epitopes of Inhibitory Mouse mAbs** 171

4.8 **Inhibition Analysis of the Panel of Human mAbs** 173

4.9 ***Pv*DBP-RII Purification** 176

4.10 **DB9Fab Purification** 177

4.11 **Binding of *Pv*DBP-RII to DB9 by SPR** 178

4.12 ***Pv*DBP-RII:DB9Fab Complex Formation** 178

4.13 **Ramachandran Plot of *Pv*DBP:DB9 Structure** 181

4.14 **Structure of *Pv*DBP-RII bound to DB9 Fab** 182

4.15 **Sequence logo representing sequence conservation across *Pv*DBP-RII** 185

4.16 **Overlay of *Pv*DBP-RII:DB9Fab Structure on the *Pv*DBP-RII:DARC Heterotetramer** 186

4.17 **Comparison of DB9 Epitope with Inhibitory Mouse mAbs** . 187

4.18 **Deglycosylation of DARC$_{1-60}$ Ectodomain** 190

4.19 **Purification of DARC$_{1-60}$ Ectodomain** 191

4.20 ***Pv*DBP-RII:DB9Fab:DARC Complex Formation** 191

4.21 **Crystals contain *Pv*DBP-RII:DB9Fab:DARC Complex** 192

4.22 ***Pv*DBP-RII:DB9Fab:DARC Crystal in Loop and Diffraction Image** 193

4.23 **Limited Proteolysis Screen for Crystal Optimisation** 195
4.24 **SDS-PAGE Gels of Ni-NTA Purification of New DARC Constructs** 196
4.25 **Overview of Immunogen Design Strategies** 200
4.26 **Mutations Introduced at Subdomain 3:Subdomain 2 Interface** 201
4.27 **An Overview of the PROSS Workflow** 203
4.28 **A Comparison of the Structure of *Pv*DBP-RII and Other Known DBL domains** 205
4.29 **"Fixed" residues not allowed to mutate in PROSS algorithm** 206
4.30 **Positions of PROSS-predicted stabilising mutations** 207
4.31 **Positions of PROSS-predicted stabilising mutations** 208
4.32 **Sequence Alignment of Proposed Immunogen Constructs** . . 209
4.33 **Insoluble Expression of Subdomain 3** 210
4.34 **Soluble Expression of Interface and PROSS** 211
4.35 **Purification of Interface and PROSS** 212
4.36 **CD Spectra of Immunogens** 213
4.37 **Immunogen Binding to DB9** 215
4.38 **Subdomain3:DB9Fab Complex Formation and Crystallisation** 216
4.39 **Ramachandran Plot for the Subdomain3:DB9Fab Structure** . 219
4.40 **Overlay of the Subdomain3:DB9Fab Structure with the Subdomain3 from the previous *Pv*DBP-RII:DB9Fab Structure** . 220
4.41 **Interface:DB9Fab Complex Formation** 221
4.42 **Interface:DB9Fab Crystal Contains Both Components** 221
4.43 **Subdomain3:DB9Fab Crystal Contacts in Interface Region** . 222
4.44 **Where to put the SpyTag? Display of Immunogens on VLP** . 225
4.45 **SpyTag-maleimide interaction with InterfaceR59C** 226
4.46 **Purification of Interface R59C** 227

4.47 **Interface R59C conjugation with Malemide Spytag Linker before with Spy-catcher VLP** . . . 228
4.48 **Purification of the Interface-Conjugated VLP** . . . 229
4.49 **SPR binding curve of Interface-coupled VLP and uncoupled VLP to DB9** . . . 230
4.50 **Immunisation Results Using Designed Immunogens** . . . 231

5.1 **Architecture of Var2CSA** . . . 237
5.2 **Solved Structures of Var2CSA DBL Domains** . . . 239
5.3 **mAbs against Var2CSA** . . . 241
5.4 **Negative Stain of NTS-DBL6** . . . 243
5.5 **Purification of NTS-DBL6 and NTS-DBL5 Constructs** . . . 246
5.6 **Chondroitin Sulphate A Architecture** . . . 247
5.7 **Complex formation of Var2CSA with CSA Before and After Chondroitinase Treatment** . . . 248
5.8 **Formation of PAM1.4 Fab** . . . 249
5.9 **Formation of NTS-DBL6:PAM1.4 Fab and NTS-DBL5:PAM1.4 Fab Complexes** . . . 250
5.10 **Representative Micrographs Comparing Non-methylated vs Methylated NTS-DBL6 Samples** . . . 252
5.11 **Representative Micrographs for NTS-DBL6 plus PAM1.4Fab vs NTS-DBL5 plus PAM1.4Fab** . . . 253
5.12 **2D Reconstructions for NTS-DBL6 vs NTS-DBL5** . . . 255
5.13 **3D Reconstructions for NTS-DBL6 vs NTS-DBL5** . . . 256
5.14 **DBL domain size in reference to Var2CSA CryoEM Density** 256
5.15 **3D Reconstructions for NTS-DBL6 Plus and Minus the Addition of PAM1.4Fab** . . . 258

5.16 **2D Classification for NTS-DBL5 with PAM1.4Fab Plus and Minus CSA** . . . 261
5.17 **Representative Micrographs of Grids Prepared with DDM Detergent in the Sample Buffer** . . . 262
5.18 **Representative Micrographs of Grids Prepared with Different Detergents in the Sample Buffer** . . . 264
5.19 **Representative Micrograph for NTS-DBL6 plus PAM1.4 on Graphene Oxide Grids** . . . 265
5.20 **2D Classification for NTS-DBL6 plus PAM1.4 on Graphene Oxide Grids** . . . 265
5.21 **3D Reconstructions for NTS-DBL6 plus PAM1.4 on Graphene Oxide Grids** . . . 266
5.22 **Limited Proteolysis of NTS-DBL5** . . . 268
5.23 **In-gel Trypsin Digest Mass Spectrometry of NTS-DBL5 and AspN Fragment** . . . 269
5.24 **Native Mass Spectrometry of NTS-DBL5 Digested with AspN** 270
5.25 **AspN Digestion of NTS-DBL5** . . . 271
5.26 **AspN Digestion of NTS-DBL6** . . . 272
5.27 **In-gel Trypsin Digest Mass Spectrometry of NTS-DBL6 and AspN Fragments** . . . 273
5.28 **Free Cysteine in DBL6ϵ Domain** . . . 274
5.29 **Boundaries of New Constructs** . . . 277
5.30 **Expression Trials for New Var2CSA Constructs** . . . 278
5.31 **Nanobodies against Var2CSA** . . . 278
5.32 **Expression and Purification of Individual DBL Domains** . . . 279

List of Tables

2.1 **Restriction Enzymes for Cloning into Different Vectors** 26

3.1 **Comparison of CIDRα1.4 Binding to EPCR** 77

3.2 **CIDRα1 domains on which each subclass of Bundle was Designed** 82

3.3 **Melting Temperatures of the Different Bundles** 89

3.4 **Kinetic Analysis of Bundles Binding to EPCR** 92

3.5 **Comparison of Bundle1.4 and CIDRα1.4 Binding to EPCR** . 92

3.6 **Data collection statistics for the Bundle1.4:EPCR complex** . 98

3.7 **Refinement statistics of the Bundle1.4:EPCR complex** 101

3.8 **Data Collection Statistics for the Other Bundle:EPCR Complexes** 103

3.9 **Data Refinement Statistics for the Bundle:EPCR Complexes** 105

3.10 **Ability of Natural Antibody Pool to Inhibit EPCR Binding to Different CIDRα1 Domains** 112

3.11 **Name of Inhibitory Mouse mAb from Immunisations with Different CIDRα1s** 129

3.12 **Name of construct chosen to represent each CIDRα1 Subtype** 130

3.13 **Kinetic Binding Analysis of CIDRα1 Domains Binding to EPCR** 136

3.14 **Summary of CIDRα1:Fab Crystal Plates Set Up** 140

3.15 **Data collection statistics for IT4var20:Jan Fab** 144

3.16 **Kinetic Binding Analysis of CIDRα1 Deletion Constructs Binding to EPCR** 154

4.1 **Kinetic Analysis of *Pv*DBP-RII Binding to DB9** 177

4.2 **Data collection statistics for the *Pv*DBP-RII:DB9Fab Structure** 180
4.3 **Refinement statistics of the *Pv*DBP:DB9 complex** 181
4.4 **Interaction Table for *Pv*DBP-RII binding to DB9** 183
4.5 **Space Group and Unit Cell Parameters for *Pv*DBP-RII:DB9Fab: DARC Complex** 193
4.6 **Melting Temperatures of Immunogen Constructs** 214
4.7 **Kinetic Analysis of Immunogens Binding to DB9** 215
4.8 **Data collection statistics for the Subdomain3:DB9Fab Structure** 217
4.9 **Refinement statistics of the Subdomain3:DB9Fab Complex** . 218

5.1 **Affinity of Recombinant DBL Constructs for CSPG** 238
5.2 **Optimised Grid Preparation and Vitrobot Freezing Parameters** 252

Abbreviations

Ab	antibody
AMA1	apical membrane antigen 1
APC	activated protein C
ATS	acidic terminal segment
BAP	biotin acceptor peptide
CCD	charge coupled device
CD	circular dichroism
CDR	complementarity determining region
CIDR domains	cysteine-rich interdomain region domains
cryoEM	electron cryo-microscopy
CSA	chondroitin sulphate A
CSP	circumsporozoite protein
CV	column volumes
DARC	Duffy antigen receptor for chemokines
DBL domain	Duffy binding-like domain
DC	domain cassettes
DDM	dodecylmaltoside
DED	direct electron detector
DMAB	dimethylaminoborane
DTT	dithiothreitol
Dvl-2	dishevelled-2
EBA	erythrocyte binding antigen
EBL domain	erythrocyte binding-like domain
ECM	extracellular matrix
ELISA	enzyme-linked immunosorbent assay

EPCR	endothelial protein C receptor
ESI	electrospray ionisation
FCS	fetal calf serum
GAG	glycosaminoglycan
GIA	growth inhibition assay
GPCR	G-protein coupled receptor
HDX	hydrogen-deuterium exchange
HEK	human embryonic kidney
HRP	horseradish peroxidase
IPTG	isopropyl β-D-1-thiogalactopyranoside
iRBC	infected red blood cell
IRS	indoor residual spraying
ITNs	insecticide-treated nets
LB	Lysogeny broth
LDAO	lauryldimethylamine N-oxide
LLG	log likelihood gain
LMNG	lauryl maltose neopentyl glycol
mAb	monoclonal antibody
MHC	major histocompatibility complex
MR	molecular replacement
MSP1	merozoite surface protein 1
NEB	New England Biolabs
OD_{600}	optical density at 600 nm
PAR-1	protease activated receptor 1
PBS	phosphate buffered saline
PC	protein C
PCR	polymerase chain reaction

*Pf*CRT	*Plasmodium falciparum* chloroquine resistance transporter
*Pf*EMP1	*Plasmodium falciparum* erythrocyte membrane protein 1
PROSS	Protein Repair One Stop Shop
*Pv*DBP	*Plasmodium vivax* Duffy Binding Protein
*Pv*DBP-RII	*Plasmodium vivax* Duffy Binding Protein Region II
RBC	red blood cell
RH5	reticulocyte binding protein homologue 5
RMSD	root mean square deviation
RTS	repeats and T-cell epitopes
RON	rhoptry neck
RSV	respiratory syncytial virus
RT	room temperature (or run-through from sera affinity columns)
RTS	central repeats and T cell epitopes
SAXS	small angle x-ray scattering
SDM	site-directed mutagenesis
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
SF9 cells	*Spodoptera frugiperda* 9 cells
SPR	surface plasmon resonance
S2 cells	Schneider cells
TAE	tris-acetate-EDTA
Tm	melting temperature
UDM	undecyl-β-D-maltoside
VLP	virus-like particle
WHO	world health organisation
2xYT	2x Yeast-Tryptone broth

Chapter 1

General Introduction

1.1 An Introduction to Malaria

1.1.1 History

Malaria is an ancient disease that has plagued humanity over the millennia. Remnants of sequences of a malarial antigen have been discovered in the mummified remains of Egyptians dating from as far back as 3200 BC [1]. Ancient writings have also chronicled the disease throughout history from clay tablets in Mesopotamia to Indian writings from 1500 BC, which referred to malaria as the "king of diseases", as well as in literature by both Homer (around 750 BC) and Aristotle (384-322 BC).

Until late in the 19th century it was believed that evil vapours (or miasmas) from swamps caused this fatal disease, and as a result the term 'malaria' arose, from medieval Italian 'mala aria' meaning 'bad air'. However, in the age of microbe discovery (Louis Pasteur's "germ theory") in 1880, Charles Louis Alphonse Laveran, a French army doctor, observed crescent-shaped bodies in the blood of malaria patients, which disappeared when patients were given quinine [2]. This was the first evidence that malaria was in fact caused by a parasite, and in the following years the main human-infective species of malaria-causing parasite were also observed and characterised. The link to the mosquito vector was made almost two decades later, by Ronald Ross, who observed pigment-containing bodies in mosquitoes which had

fed on malarial-infected patients [3]. Ross went on to demonstrate using the avian parasite *Plasmodium relictum* that malaria parasites were transmitted by mosquitoes, which were classified as female *Anopheles* mosquitoes by an Italian scientist named Grassi [4].

1.1.2 The Disease

Despite its long history, even in the 20th century, malaria was responsible for the deaths of between 150 million and 300 million people, which represented 2-5% of all deaths during that period [5]. Today, malaria is still a global burden. According to the World Health Organisation (WHO), an estimated 3.4 billion people in 91 countries and territories are at risk of being infected with malaria and developing disease, with 1.1 billion at high risk (high risk being defined as >1 in 1000 chance of getting malaria in a year) [6]. The latest figures show that more than 435,000 people die each year and there are nearly 220 million cases annually. The burden, in 2017, was heaviest in the WHO African Region, where an estimated 92% of all malaria deaths occurred, and in children aged under 5 years, who accounted for more than two thirds of all deaths [6]. Worryingly, despite an unprecedented period of success in global malaria control over the last two decades, progress appears to have stalled in the last couple of years **(Figure 1.1)**.

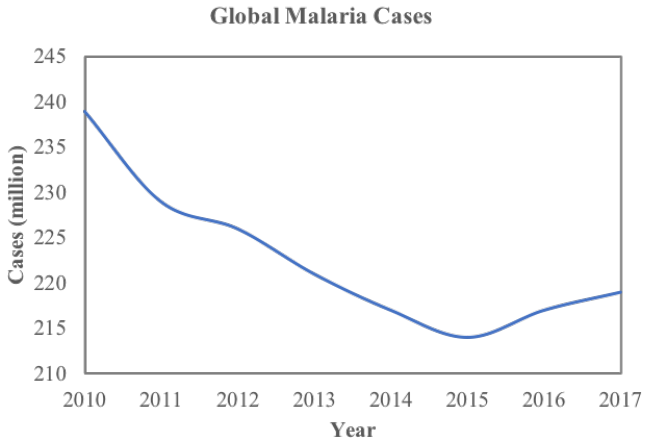


Figure 1.1: Malaria Cases Between 2010 and 2017 (Data obtained from the WHO [6])

1.1.3 The *Plasmodium* parasite

Malaria is caused by protozoans, of genus *Plasmodia*, which belong to the phylum Apicomplexa. These are characterised by a series of organelles that are specialised for host cell invasion, called the apical complex. There are over 50 different species of *Plasmodium*. However, only five species of *Plasmodium* are known to infect humans: *P. falciparum*, *P. vivax* [4], *P. ovale* [7], *P.malariae* [8], and *P. knowlesi* [9]. Out of these five, *P. falciparum* is the most deadly, accounting for >80% of cases worldwide **(Figure 1.2)**. *P. falciparum* infection is particularly fatal in children under 5 years of age, often developing into severe disease, characterised by fever and anaemia, as well as neurological symptoms that can lead to coma and death. Although *P. falciparum* causes the most fatalities, the most widespread species of malarial parasite is *P. vivax*, which mostly predominates South-East Asia and the Americas [10, 11]. Indeed, in 2017, *P. vivax* was the predominant parasite in the Americas, accountable for 74% of all malaria cases [6]. Moreover, *P. vivax* was responsible for 37% and 31% of cases in South East Asia Region and Eastern Mediterranean Region, respectively [6] **(Figure 1.2)**. There is even growing evidence that *P. vivax* is more prevalent in Africa than once thought [12]. As a result of the spread and the increasing prevalence

of severe disease caused by *P. vivax* [13], more research is now being targeted on this previously thought "benign" species of malaria [14].

Together, *P. falciparum* and *P. vivax* are responsible for over 90% of malaria cases. The other human-infective malarial species, *P. ovale*, *P.malariae*, and *P. knowlesi* cause fewer and milder infections which rarely progress into clinical disease, and are often found in mixed infections with *P. falciparum* and *P. vivax* [6]. As a result, much research has focused on *P. falciparum* and *P. vivax* as the main two causative agents of malaria.

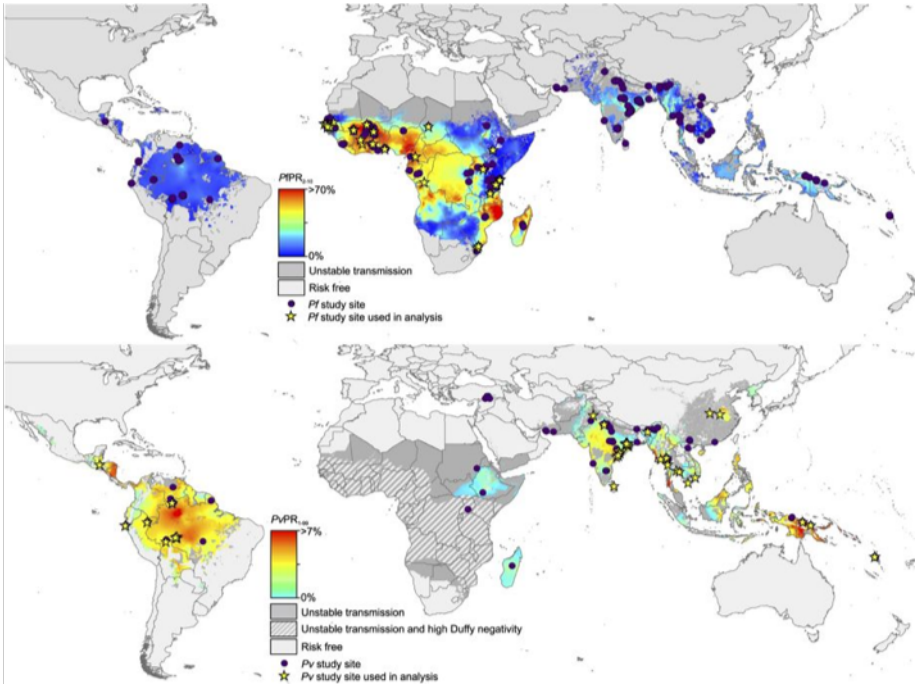


Figure 1.2: World map showing the incidence of *Plasmodium falciparum* (top) compared to *Plasmodium vivax* (bottom) This data represents the annual point estimates of malaria incident in 2010 and is adapted from [15].

1.1.4 *Plasmodium falciparum* life cycle

Plasmodia have a complex life cycle, which involves stages both in the *Anopheles* mosquito and in the human host **(Figure 1.3)**. The life cycle can be split into four main stages: exo-erythrocytic cycle, erythrocytic cycle, gametocytogenesis, and the

sporogonic cycle. These stages of the life cycle for *P. falciparum* are described in more detail in the following sections. During these stages, the *Plasmodia* adopt multiple different morphologies, each specialised to the function of the parasite at that particular point in the life cycle.

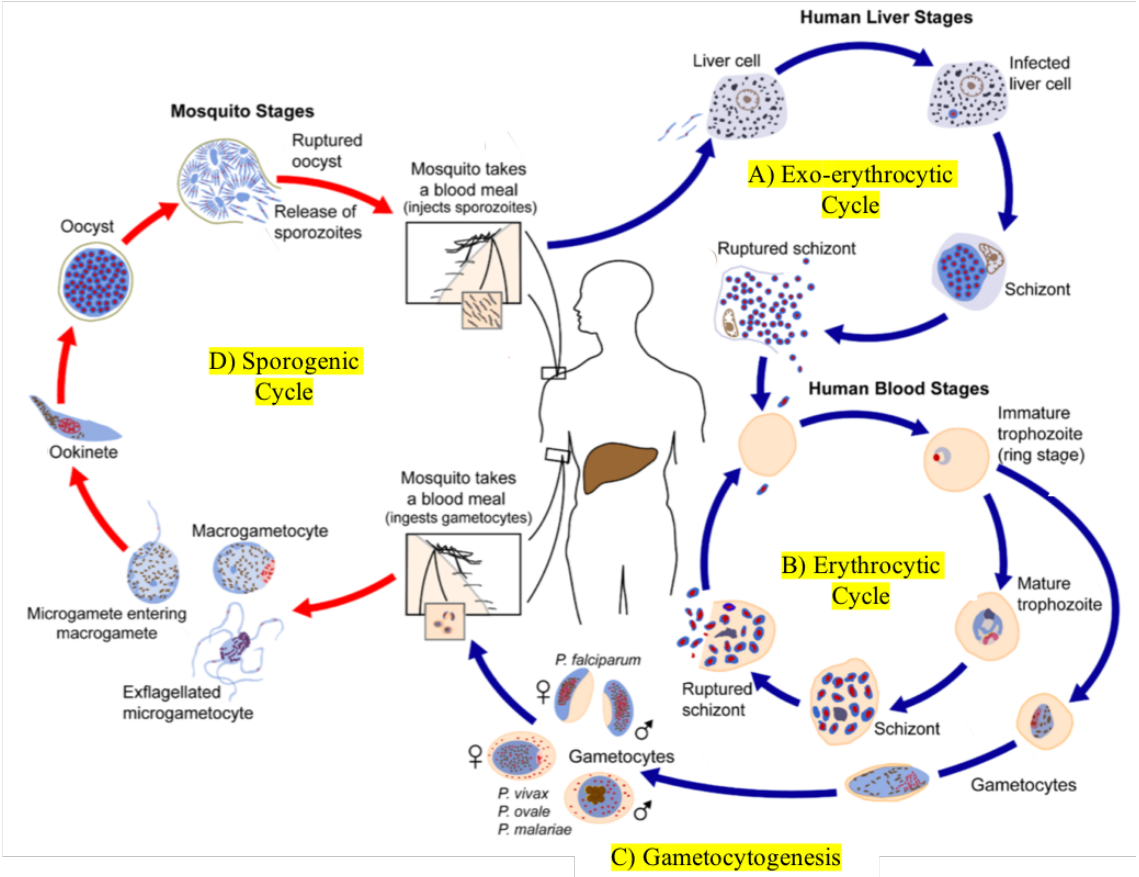


Figure 1.3: *Plasmodium* parasite life cycle Highlighting the four main stages of the life the cycle. Adapted from the Centers for Disease Control and Prevention (CD) at https://www.cdc.gov/malaria/about/biology/.

1.1.4.1 Exo-erythrocytic cycle

When an *Anopheles* mosquito takes its blood-meal and bites a human host it injects *Plasmodium* sporozoites into the skin [16]. The sporozoites are highly motile and elongated, allowing them to make their way to the bloodstream, either directly into capillaries or through the lymphatic system. Once in the bloodstream, the

sporozoites migrate quickly to the liver [17]. In the liver the parasite multiplies and this stage is known as the exo-erythrocytic cycle **(Figure 1.3 A)**.

The localisation of the sporozoites to the liver sinusoid is mediated by the circumsporozoite protein (CSP) coat binding to the heparan sulphate proteoglycans in the extra-cellular matrix (ECM) of the hepatocytes [18]. This initial contact is followed by either invasive or non-invasive movement through resident macrophages (Kupffer cells) or endothelial cells, before reaching the hepatocytes. The sporozoites may traverse multiple hepatocytes by a "brute-force" approach [19], before developing an internalised parasitophorous vacuole in the final hepatocyte invasion step [20]. In this vacuole the sporozoites replicate asexually, developing into schizonts which contain around two thousand parasites in the merozoite form. Merozoites are non-motile, specialised cells for initiating new sexual or asexual cycles of development. The mature merozoites bud away from the hepatocytes in membrane structures called merosomes. Since these merosomes are packaged in non-immunogenic host cell membranes, it is hypothesised that this enables them to pass safely back into the bloodstream without detection by immune cells [21]. In the pulmonary capillaries, the merosomes are disrupted, resulting in the release of thousands of merozoites into the bloodstream [22].

1.1.4.2 Erythrocytic cycle

In the bloodstream, the merozoites go through repeated cycles of erythrocyte invasion, followed by replication inside the erythrocytes before cell rupture and the release of more merozoites into the bloodstream; this is known as the erythrocytic cycle **(Figure 1.3 B)**.

Free merozoites invade erythrocytes very quickly (within 30 seconds) in an ordered sequence of events [23, 24]. The merozoite makes initial contact with the red blood cell (RBC) via low affinity but specific interactions, which stimulate the parasite

actomyosin deformation of the host cell [25]. This process is not fully understood but some interactions have been characterised, for example, merozoite surface protein 1 (MSP1) binding to Band 3 on the RBC surface [26, 27]. After the initial contact and the deformation of the erythrocyte, the parasite cell re-orients itself so that the apical complex end of the merozoite abuts the erythrocyte membrane. At this point, higher affinity interactions take place involving proteins in the *Pf*RH and EBL (erythrocyte binding-like) families [28]. An irreversible tight junction is then formed by the interactions between parasite proteins, primarily through AMA1 (apical membrane antigen 1) and the RON (rhoptry neck) complex [29]. This sequence of interactions, leads to intracellular signalling both within the parasite and the erythrocyte, mostly mediated by Ca^{2+} influx. In the merozoite, this signalling cascade leads to release of proteins and lipids from the rhoptries, providing a parasitophorous membrane in which the parasite invades the RBC [30]. This active invasion process involves the actomyosin motor [31], and after the merozoite enters the RBC, both the parasitophorous and the erythrocyte membranes are sealed [32].

Once inside the erythrocyte, the parasite (now in trophozoite stage specialised for growth and adsorption of nutrients from the host) asexually replicates and matures into 16-32 schizonts. Once fully mature, both the parasitophorous vacuole membrane and the erythrocyte membrane are disrupted, resulting in the release of the merozoites into the bloodstream [33]. These merozoites can now access new erythrocytes for invasion.

1.1.4.3 Gametocytogenesis

After completing the erythrocytic cycle two or three times, some merozoites invade a new erythrocyte and instead of developing into schizonts, they develop into male and female gametocytes; this process is known as gametocytogenesis **(Figure 1.3 C)** [34]. The gametocytes develop in erythrocytes found in the bone marrow and the spleen.

The signal that leads to development of gametocytes instead of schizonts is unknown, however, it is thought that lack of nutrients, accumulation of parasitic debris and development of immunity may all play a role [35, 36]. Once the gametocytes have matured, these crescent shaped cells are released into the peripheral blood, ready to be ingested by feeding mosquitoes.

1.1.4.4 Sporogonic cycle

The multiplication of the parasite within the mosquito is known as the sporogenic cycle **(Figure 1.3 D)**. After a blood meal, when in the stomach of the mosquito, the male gametocytes penetrate the female gametocytes leading to the formation of zygotes. The zygotes become more motile and elongated developing into forms known as ookinetes, which can invade the mid-gut wall of the mosquito [37]. This invasion process has been shown to involve the same proteins required by the sporozoites in the exo-erythrocytic phase to invade the hepatocytes, and so the invasion mechanism is thought to be similar [38]. It is here that they develop into oocysts, which grow and rupture, releasing sporozoites that make their way to the mosquito's salivary glands [39]. Injection of the sporozoites into a new human host then perpetuates the malaria life cycle.

1.1.5 Differences in *P. vivax* Life Cycle compared to *P. falciparum*

The life cycle of the *P. vivax* parasite is very similar to that of *P. falciparum*, in that the complex cycle involves stages both in the mosquito and human host. However, there are some significant differences between these two species, which affect the virulence and global distribution of the disease. These differences therefore also impact the treatments and the way in which the parasites are targeted by potential

vaccines. The main differences in life cycle are described in the following sections and summarised in **Figure 1.4**.

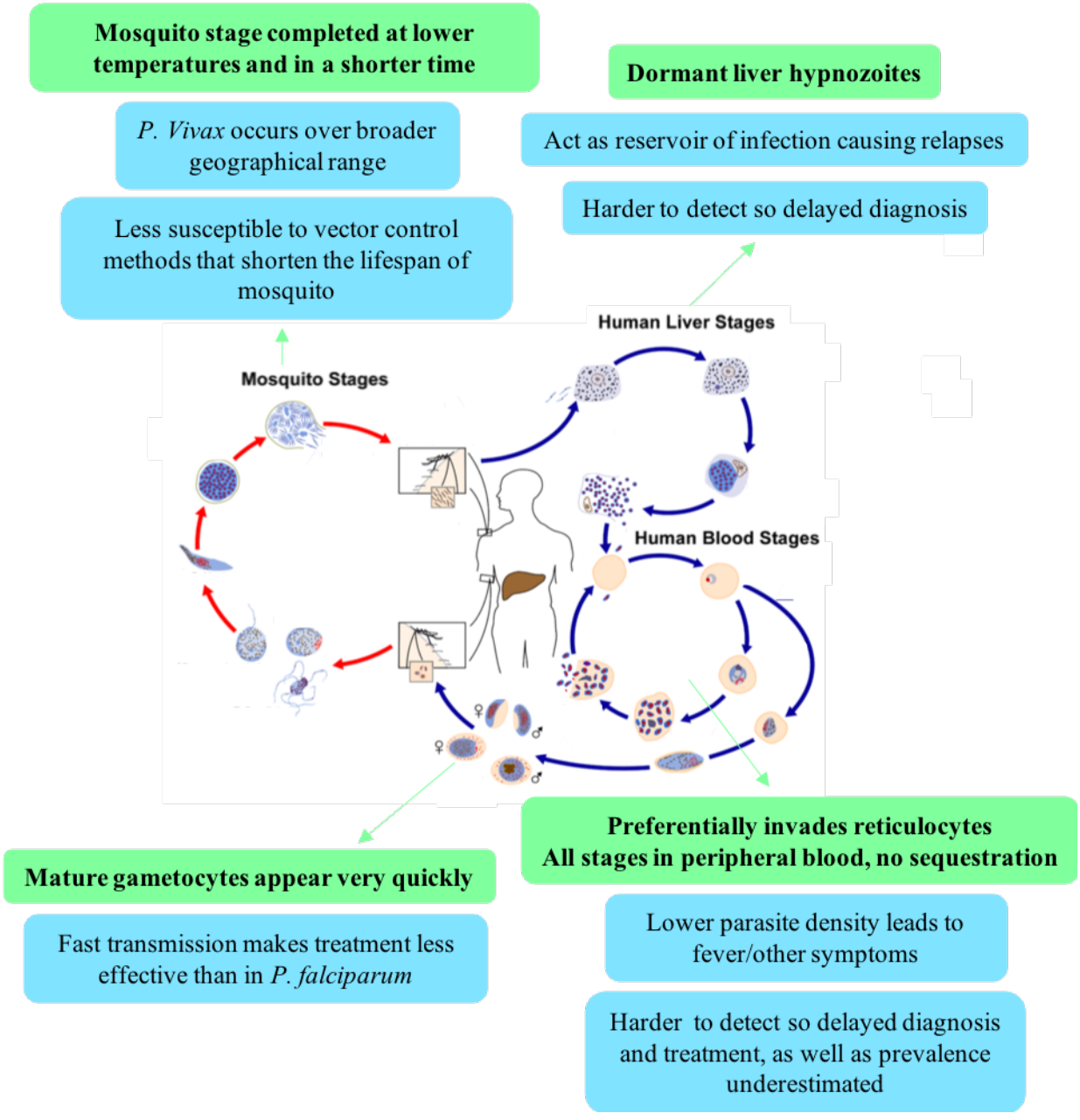


Figure 1.4: Features of the *Plasmodium vivax* life cycle The differences between the *Plasmodium vivax* and *Plasmodium falciparum* life cycles higlighted in the green boxes. The consequences of these features on *P. vivax* malaria and treatment are summarised in the blue boxes.

1.1.5.1 Mosquito Stages

The *P. vivax* parasite can complete its development within the mosquito at lower temperatures (16°C average compared to 18°C) and more quickly (8-16 days compared to 9-22 days until transmission) than *P. falciparum* [40]. This feature of *P. vivax* contributes to the widespread incidence of *P. vivax* malaria into more temperate regions compared to *P. falciparum* infections, which only occur in the tropics [41] **(Figure 1.2)**.

1.1.5.2 Formation of dormant liver hypnozoites

In *P. vivax*, in the exo-erythrocytic cycle, the maturation of the sporozoites into schizonts in the liver may be delayed for up to 1 to 2 years [42]. In this dormant phase of the life cycle, the parasites are known as hypnozoites and they exist intracellularly, and so are protected from the immune system and can remain undetected for months/years. This feature of the life cycle provides an additional challenge since when inside the liver cells it is very hard to target the parasite with drugs/vaccines. This quiescent phase of the life cycle also means that there are often relapses, weeks, months or years post primary infection, which are not seen in *P. falciparum* infections.

This ability to remain in the body for long periods of time undetected also contributes to the ability of *P. vivax* infections to survive in colder climates. The parasite can remain dormant and undetected in the cold winter months when transmission is not possible as the mosquitoes cannot survive but then when the temperature increases, proliferation and transmission can begin again, further explaining the spread of *P. vivax* infection into more temperate climates.

1.1.5.3 Specific Invasion of Reticulocytes

Another difference in the life cycle of *P. vivax* compared to *P. falciparum* is that during the asexual blood stage, the merozoites do not invade all RBCs. They specifically select immature reticulocytes in the bone marrow [43, 44]. Although *P. falciparum* also preferentially targets reticulocytes over mature erythrocytes [45, 46], they can target all ages of RBCs circulating in the blood, allowing fast invasion and subsequent growth in this asexual cycle. It is thought that the advantage of targeting immature RBCs is that erythrocytes have a limited life span and so rejecting them reduces the chance of invading a senescent cell, which would lead to early splenic clearance [47]. Therefore, invading reticulocytes allows extended time for asexual development compared to erythrocytes [48, 49]. In addition, the reticulocyte is not yet discoid and so the increased flexibility of these infected young cells allow them to pass unhindered through splenic fenestrations and capillaries more easily. Making infected reticulocytes more resistant to splenic clearance than more mature erythrocytes [47, 50].

Another consequence of *P. vivax* selectively invading reticulocytes, is that the erythrocytic part of the life cycle occurs in the peripheral blood. Since reticulocytes make up 0.5-1.5% of all circulating erythrocytes, this means that the overall parasite density in circulating blood is low. As a result, *P.vivax* is capable of transmission at lower levels of parasites in the blood than *P. falciparum* [51] and also causes symptoms, such as fever, at much lower blood parasitemia. Consequently, *P. vivax* is often not detected in blood until the infection is already causing symptoms. This means it is hard to detect *P. vivax* infection early and since *P. vivax* is more transmissible at lower parasite densities, there is fast transmission throughout populations. Moreover, the inability of current detection methods to detect the lower concentrations of *P. vivax* in the blood means the incidence rate is probably underestimated. In addition, this increased specificity of *P. vivax* RBC invasion may contribute to the decreased

virulence of *P. vivax* compared to *P. falciparum*.

It is also noteworthy, that the preference for reticulocytes has so far delayed the development of a continuous *in vitro* culture of *P. vivax* parasites, which has hindered the ability to test potential vaccines/probe the invasion process in *P. vivax*. Currently *P. knowlesi* transgenic lines or freshly collected parasites from patients in *P. vivax* endemic regions are used for invasion challenge studies.

1.1.5.4 Fast transmission

Moreover, compared to *P. falciparum*, the gametocytes develop very quickly in the life cycle and circulate early in infection (after 3 days compared to 10-14 days in *P. falciparum*), which allows transmission before patients feel ill and seek treatment [52] **(Figure 1.4)**. This enables *P. vivax* to spread quickly throughout a population [53]. Moreover, *P. vivax* gametocytes are short-lived (3-4 days, compared to up to weeks for *P. falciparum*) in the human circulation, and so may make the sexual stage of the life cycle harder to target as they are only exposed in circulation for short periods of time.

1.2 Disease and Pathology

1.2.1 Malaria Pathology and Symptoms

Malaria is often divided into two different categories: uncomplicated or severe [54]. Uncomplicated malaria is characterised by symptoms associated with a minor systemic viral illness. Comprising of headaches, fatigue, abdominal discomfort, muscle and joint aches, followed by fever, chills, perspiration and vomiting [55]. If treated correctly and promptly, full recovery from uncomplicated malaria is expected. However, if the parasite burden is allowed to increase, the patient may develop severe

malaria, which if left untreated is fatal in the majority of cases. Around 1% of uncomplicated malarial infections develop into severe disease and this progression may take as little as a few hours. Severe malaria is characterised by symptoms including impaired consciousness/coma, severe anaemia, metabolic acidosis, acute respiratory distress syndrome, hypotension, and acute kidney injury [55–58].

The pathophysiology leading to these conditions are not well understood. It is thought that most of clinical manifestations of malaria are primarily due to the high parasite load and the infection of RBCs, leading to sequestration, less flow, and lower density of healthy RBCs [59, 60]. There are three main mechanisms of RBC sequestration: the binding of infected red blood cells (iRBCs) to endothelial receptors (cytoadhesion), to uninfected RBCs (rosetting), or to other iRBCs (platelet-induced clumping). In addition to restricting flow and causing inflammation in tissues, the sequestration of the iRBCs to the endothelium enables these infected cells to avoid destruction by the immune system and splenic clearance, as well as avoiding detection in the peripheral blood [61, 62]. In addition, many of the symptoms of malaria, like fever, nausea and headaches, are the result of the inflammatory response caused by the innate immune system, which has been stimulated by the parasites and their products at the rupture of the iRBC sequestered in tissues [63–65].

Most fatal cases that occur in *P. falciparum* infections are due to cerebral malaria [6]. Cerebral malaria, which is prevalent mostly in African children under the age of five, is characterised by loss of consciousness and/or coma and is fatal in ∼15% of cases [66–68]. Those that do survive, often have permanent neurological damage [69]. The exact mechanism of how cerebral malaria manifests is not well understood. However, the sequestration of iRBC in the brain endothelium through interactions between parasite proteins, like *P. falciparum* erythrocyte membrane protein 1 (*Pf*EMP1), and various endothelial receptors have been strongly implicated [70]. (The role of *Pf*EMP1s in sequestration and disease is discussed in more detail in **Section 3.1**).

The main cause of coma is not known, but it is thought that iRBC sequestration causes microvascular congestion that leads to severe endothelial damage in the brain tissue. This results in disruption of the vessel wall, resulting in myelin and axonal damage and subsequent breakdown of the blood-brain barrier [71, 72]. Moreover, the release of high levels of pro-inflammatory cytokines, caused by the cytoadherence of iRBCs to the brain endothelium, has been implicated in the pathogenesis of cerebral malaria [73–76]. Due to the iRBC sequestration and consequent inflammation, often retinopathy is used as a diagnostic tool of severe malaria. The vascular damage visualised in the retina correlates with the severity of microvascular brain damage and can be used both as a predictor of death and of neurological damage that survivors might incur [77–80].

As mentioned, whilst most adults living in endemic regions have acquired some protective immunity against malaria, pregnancy causes severe symptoms in women and significant risk to their unborn babies [81]. This type of malaria has been attributed to the sequestration of iRBC in the placenta via the binding of the *Pf*EMP1 variant, Var2CSA, to chondroitin sulphate A (CSA) on the endothelial lining [82]. (The role of Var2CSA is discussed in more detail in **Section 5.1**).

1.2.2 Natural Acquired Immunity

The first evidence that the presence of antibodies conferred protection to malarial infections was gained in the 1960s: purified antibodies from malarial immune individuals were passively transferred to non-immune children, and resulted in a 99% reduction in parasitemia in four days [83]. It has since been shown that different types of immunity to malaria can be gained: 1) anti-disease immunity, conferring protection against clinical disease, which affects the risk and extent of morbidity associated with a given parasite density; 2) anti-parasite immunity, which confers protection against parasitemia by keeping parasite densities low; and 3) premunition,

which provides protection against new infections by maintaining low-level and asymptomatic parasitemia [84, 85]. After a few symptomatic infections, young children can gain immunity to some of the most severe aspects of malaria (anti-disease immunity) [86]. This explains the rise and rapid decline in the presence of severe malaria and deaths in the first five years of life in malarial endemic areas [87]. However, immunity to the clinical symptoms of the disease (i.e. when parasitemia levels are kept so low that they are undetectable and prevent the onset of disease symptoms - anti-parasite immunity) takes multiple years of repeated exposure and infections to develop [88, 89].

1.2.3 Polymorphisms, allelic switching, and redundancy

Natural acquired immunity is difficult to obtain and requires repeated exposure to the disease, mainly due to mechanisms the parasite has evolved in order to evade immune detection. One problem that the immune system has to overcome is that there are multiple different strains of parasite, each expressing different polymorphic homologues. Therefore the development of adaptive immunity against all the strains requires multiple infections to produce antibodies that target the different highly polymorphic antigens. Moreover, often targets of the immune system, such as *Pf*EMP1s, RIFINs and STEVORs, are members of large gene families whose expression can be switched from one allele to another, this process is called allelic switching. Once the immune system has detected and produced antibodies that recognise a particular protein, allelic switching occurs, where another protein of the same family is expressed. The circulating antibodies do not recognise this protein and so the parasite can evade the immune system. This is possible because the different variants of the same family contain highly polymorphic sequences, meaning that antibodies do not recognise and bind other variants of the same protein [25, 90]. Furthermore, the presence of multiple invasion pathways and functional redundancy

of the invasion ligands of the merozoites also help evasion of the immune system [25, 91]. These features not only make immunity to malaria hard to achieve, but also the prospect of designing a vaccine that generates strain-transcendent antibodies that cannot be evaded by antigenic switching or redundant pathways a daunting task.

1.2.4 Genetic Resistance

Several genetic conditions in the human host can confer resistance to *Plasmodium* infection. For example, it has been known for a long time that sickle-cell anaemia protects individuals against *P. falciparum* severe malaria [92]. The exact mechanism of protection is unknown, however the current leading hypothesis is that the infected sickle cells have a reduced display of the *Pf*EMP1 proteins on their surface, which leads to lower cytoadherence and so reduces the symptoms of the disease [93]. Another example of genetic resistance is in individuals who have glucose-6-phosphate dehydrogenase (G6PD) deficiency. This deficiency reduces the ability of erythrocytes to maintain the oxidation levels that are required for the survival and replication of the *Plasmodium* within the red blood cell [94]. In addition, the absence of the Duffy receptor on erythrocytes in Duffy negative individuals provides protection against *P. vivax* infection [95] (discussed in further detail in **Section 4.1**).

1.3 Current Interventions and Treatments

1.3.1 Vector Control

The spread and distribution of malarial infection is entirely determined by the location and abundance of the *Anopheles* vector. Therefore an essential component of malaria prevention is vector control, which targets the mosquitoes that transmit

the *Plasmodia* parasites. In the past, large-scale spraying of mosquito habitats with dichlorodiphenyltrichloroethane (DTT)-based insecticides has been successful in eradicating mosquitoes. Notably, this huge-scale spraying in the 1940s and 1950s eradicated malaria in the USA and Northern Europe. However, despite its early success, resistance to DTT insecticides emerged in many areas, potentially due to its over-liberal use and so is no longer used at the same scale in the control of mosquitoes [96].

Currently, the WHO advise smaller scale use of insecticides, with the main measures being insecticide-treated nets (ITNs) and indoor residual spraying (IRS) [6]. Since 2000, there has been a huge increase in the delivery of ITNs to malarial endemic countries and as a result the number of children under 5 years of age sleeping under ITNs has increased from $<2\%$ in 2000 to $\sim 61\%$ in 2017. The rapid scale-ups of ITNs has been one of the largest contributors to the drops seen in the incidence rate of malaria since 2000, highlighting their effectiveness [97].

Unfortunately, the main insecticides in use, DTT and pyrethoids, target the same voltage-gated sodium channel and resistance is growing rapidly [98, 99]. During the period 2010-2017, mosquito resistance to at least one out of the main four different types of insectides, was detected in 68/80 malaria-endemic countries [6].

Recently, due to insecticide resistance other measures such as biological vector controls have gained more attention. For example, the fungal biopesticide *Wolbachia* [100], and densoviruses [101] that alter the life cycle of the mosquito have been shown to have the ability to target older mosquitoes. Targeting older mosquitoes is one way to "evolution-proof" the vector control as these mosquitoes would have already mated and so would no longer be under natural selection. Moreover, more studies are looking into the genetic modification and sterilisation of *Anopheles* mosquitoes in an attempt to decrease the vector pool [102].

1.3.2 Drugs

In the 1950s, chloroquine was the main drug used to target malaria, however widespread resistance has put chloroquine mainly out of use. Inside the iRBC, the parasite digests the RBC haemoglobin in a digestive vacoule before converting the oxidised haem into inert haemozoin crystals [103]. Chloroquine targets the haem crystallisation process [104]. Resistance arose from mutations in the chloroquine resistance transporter (*Pf*CRT), an integral protein in the membrane of the digestive vacuole that pumps chloroquine out of the vacuole away from the haemozoin crystals [105].

Currently the main drug in use is artemisinin [6], which also targets the haemoglobin digestion pathway. The free iron produced during haemoglobin digestion, cleaves the endoperoxide bond in the artemisinin, which can then alkylate hundreds of proteins [106]. These alkylated metabolites damage the digestive vacuole membrane and stop the haemozoin crystal formation, as well as having other destructive downstream effects [107–109]. In this way, artemisinins are fast-acting drugs against the bloodstage, rapidly reducing blood parasitemia, as well as acting against the gametocytes preventing further transmission. However, artemisinins are short-lived and so are often used in combination with other drugs that have longer half-lives, such as lumefantrine. Lumefantrine acts by binding to β-hemin, which is produced during the breakdown of haemoglobin, and prevents the haemozoin crystal formation [110]. These combination therapies are called artemisinin-combination therapies (ACTs), with the second component required to clear any remaining parasites after the artemisinin has been metabolised. Worryingly, in the past decade evidence of artemisinin resistance is growing, with parasites being cleared from the blood progressively more slowly [111]. In addition, when ACTs fail, it has been shown that this can be due to resistance to the non-artemisinin component of the ACT. Resistance

to the lumefantrine and amodiaquine partner drugs, both of which target the haem degradation pathway, has been correlated with mutations in the *Pf*CRT, which pumps the drug out of the cell [112]. Other partner drugs include pyrimethamine and sulfadoxine, which are antifolates that inhibit enzymes in the essential folate synthesis pathway [113, 114]. Resistance to these drugs has also been observed, attributed to mutations in the parasite's dihydrofolate reductase and dihydropteroate synthase [115].

Due to increasing resistance observed in most drugs currently in use, there is a requirement to identify new drug candidates to which the parasite will be less likely to develop resistance. In the last decade, drug discovery efforts have centred on high-throughput screening pipelines in the hope that a potential new candidate will be found, which will not have the same problems with resistance development [116].

1.3.3 Vaccine Candidates

The complex life cycle of the *Plasmodium* parasites, in which they adopt multiple different morphologies, together with their ability to present different antigens and evade the immune system, has made the production of a vaccine that offers long-term protection a very difficult challenge, which has yet to be successful. Much research has gone into the development of vaccine candidates against the three main stages of the *Plasmodium* life cycle: transmission-blocking vaccines, pre-erythocytic vaccines and blood-stage (erythrocytic) vaccines [117, 118].

Transmission-blocking vaccines target the gametocytogenesis stage of the life cycle **(Section 1.1.4.3)**. The vaccines induce the production of antibodies that target antigens on the surface of the gametocytes. Once the gametocytes are taken up by the feeding mosquito, the antibodies bound to their surface prevent fertilisation in

the mosquito mid-gut, or destroy the gametes or ookinetes [119, 120]. These vaccines would not provide immunity to the individual that had been vaccinated but would prevent the transmission of the parasite, via a mosquito, to another person. The main candidates that are currently being investigated in preclinical trials in the search for a transmission-blocking vaccines are the post-fertilisation antigen *Pfs*25 and the pre-fertilisation antigens *Pfs*230 and Pfs48/45 [121, 122].

Pre-erythrocytic vaccine candidates target the malarial parasite at the sporozoite stage, before it invades the red blood cell and symptoms of the disease are seen **(Section 1.1.4.1)**. Any vaccines that target this stage of the life-cycle would provide protection to the vaccinated individual against new infections before they caused any symptoms of the disease. In the 1970s, whole sporozoite vaccines were shown to protect humans from *P. falciparum* infection demonstrating that targeting this part of the life cycle can be effective [123]. Since then most research for this stage of vaccine has focused on CSP, with the most advanced malaria vaccine being RTS,S. This pre-erythocytic stage vaccine consists of a virus-like particle (VLP) which displays the hepatitis B surface antigen fused with *P. falciparum* CSP fragment containing its central repeats and T cell epitopes (RTS). This vaccine has completed Phase III clinical efficacy trials, in which it prevented 39% of cases of malaria and only 29% of cases of severe malaria over a four year period [124]. Generally, the WHO has a target of 90% efficacy so this is not ideal, although this is our best vaccine candidate to date.

As discussed in **Section 1.2.1**, most clinical manifestations of malaria occur in the erythrocytic cycle. Blood stage vaccines target either the merozoites directly, before they invade the next RBC or they target *Plasmodia* proteins that have been displayed on the surface of the iRBC. Currently, in *P. falciparum*, the main merozoite targets are MSP1, AMA1 and reticulocyte binding protein homologue 5 (RH5). The first-developed AMA1 vaccines that target merozoite invasion, which were promising

in initial trials, did not provide protection in controlled human malarial infected individuals [125]. However, more recently the use of the complex of AMA1 with another protein RON2, which is also found at the merozoite and erythrocyte invasion junction was tested. The AMA1-RON2 complex used as a vaccine in *Aotus* monkeys, was more effective than the AMA1 alone [126]. Therefore, it is hoped that the complex may enhance the immunogenicity of some conserved AMA1 epitopes and could form the basis of a vaccine in the future. In addition, almost a decade ago it was shown that vaccines that target the highly conserved *P. falciparum* RH5 produce inhibitory antibodies that block merozoite invasion of erythrocytes and prevent the blood-stage malarial infection [127]. Subsequent immunisations in *Aotus* monkeys, showed that the *Pf*RH5 vaccines induced antibodies that conferred protection against a *P. falciparum* challenge [128]. These *Pf*RH5-based vaccines are currently in clinical trials in humans [129].

It is evident that there is a plethora of research ongoing in the quest to determine conserved antigens that can be targeted using vaccines. However there is still a long journey ahead with many potential vaccine candidates that have yet to be tested, and the quest to produce a vaccine that effectively targets multiple strains of *Plasmodia* with high efficacy has yet to be developed.

1.4 Structure-guided Immunogen Design

Until a decade ago, there was a very limited set of antigen design and presentation methods available to vaccine engineers. More recently, there has been an increase in development of epitope-focused vaccine design strategies. These methods use information about the structural epitopes targeted by protective antibodies to design new immunogens that will then induce these protective antibody responses [130, 131]. In order to encapsulate the structure of the epitope and stabilise it in a conformation

that leads to the production of the desirable antibodies, it is often transplanted onto a scaffold protein. Previously, scaffold proteins of known and similar structure to the epitope were used, with the side chain and/or backbone constraints of the epitope transplanted in methods called "side-chain grafting" and "backbone grafting" [132, 133]. In some cases, where the epitope was comprised of commonly found structural elements like helices, this was successful. However, immunogens developed by these methods often induced the production of antibodies that were not inhibitory [134–136]. This was thought to be due to the structure of the scaffold proteins being too rigid and predetermined, so that the conformation of the epitope could not be completely mimicked. More recently, computational methods have been developed that allow the *de novo* folding and design of scaffold proteins around a "fixed" functional epitope that is held rigid [137]. This method is called Fold from loops, and has been used to successfully recreate an RSV (respiratory syncytial virus) epitope on a small scaffold. The advantage of this method is that it can force the scaffold to accommodate the epitope, rather than the epitope changing to accommodate the scaffold. The scaffold is then mutated to stabilise the artificial conformation. This approach holds great promise, as long as a) there is structural information available about a broadly neutralising epitope and b) a suitable starting scaffold that allows grafting of the inhibitory epitope can be found [138]. In the last few years, the increased isolation and purification of broadly neutralising antibodies in the Influenza and HIV-1 fields has sped up vaccine design efforts [139]. A greater understanding of the mechanisms and structural interactions of inhibitory monoclonal antibodies against the different malarial stages would greatly aid future malarial structure-based vaccine design [140].

In malaria, as previously described, targeting the erythrocytic stage of the *Plasmodium* life cycle could block tissue sequestration, which is associated with clinical malaria and severe diseases symptoms. However, one main obstacle in the devel-

opment of effective vaccines thus far has been the high degree of polymorphisms and molecular redundancy in parasite proteins expressed at this stage [141]. One of the most advanced malarial vaccine targets from this erythrocytic stage is based on the merozoite protein RH5 [142–144]. RH5 binds to basigin on the erythrocyte surface, as a triple complex with two other parasite proteins, and is involved in the erythrocyte invasion process. The isolation and structural characterisation of human mAbs from vaccination with RH5, have allowed inhibitory and non-inhibitory epitopes to be elucidated [24, 145]. Surprisingly, an antibody with no neutralising effect itself, potentiates the effects of inhibitory antibodies indirectly. It is thought that this may happen by slowing down the invasion process and therefore increasing the time available for the inhibitory antibodies to bind [24]. This knowledge could lead to a vaccine designed with multiple inhibitory as well as potentiating epitopes, to target erythrocyte invasion in the future.

This demonstrates that the more structural information gathered about the binding of human antibodies to their malarial targets, the more knowledge there is of which epitopes to include in designed vaccines. Therefore, for future structure-guided vaccine design against the erythrocytic stage, the most conserved, non-redundant, and protective epitopes need to be identified and characterised. With prime candidates for proteins that contain these likely to be involved in erythrocyte invasion, as well as those proteins that mediate iRBC cytoadhesion.

1.5 Thesis Outline

This thesis is divided into three distinct chapters, focusing on the structural-based design of potential novel blood-stage vaccine candidates both in *P. falciparum* and in *P. vivax*. A brief summary of each chapter is provided below.

- **Chapter 3: Targeting Cytoadherence in Severe Malaria by Blocking**

EPCR-binding *Pf*EMP1s

  - The design of a panel of immunogens, based on the EPCR-binding site, using the previously solved structure of EPCR in complex with the minimal binding portion of *Pf*EMP1.
  - The characterisation of the immunogen structures, function and ability to induce the production of antibodies in rats that blocked binding to EPCR and subsequent cytoadherence.
  - The structural studies of the minimal EPCR-binding domain of *Pf*EMP1 in complex with mice inhibitory monoclonal antibodies (mAbs) to identify inhibitory epitopes that could be used for future vaccine design.

- **Chapter 4: Targeting Invasion in *P. vivax* Malaria via *Pv*DBP (Duffy Binding Protein)**
  - The structural elucidation of the functional domain of *Pv*DBP in complex with a broadly-neutralising human mAb.
  - The design of new immunogens that were based on the inhibitory epitope using the solved structure.
  - Functional, structural and immunogenic characterisation of the designed immunogens.
- **Chapter 5: Targeting Cytoadherence in Pregnancy-Associated Malaria via Var2CSA**
  - Progress made so far, ongoing experiments and future directions in the structural studies of the full-length var2CSA, an unusual *Pf*EMP1 required for the cytoadhesion to the placenta.

Chapter 2

Materials and Methods

2.1 Cloning

2.1.1 Polymerase Chain Reaction

In cloning for all constructs, the polymerase chain reaction (PCR) was carried out using Phusion Flash PCR Master Mix (ThermoFisher Scientific) on a MJ Mini Personal Thermal Cycler (Bio-Rad, USA), as per the manufacturer's protocol. A typical PCR cycle involves an initial denaturation step at 98°C for 10 s, followed by 30 cycles of 1) denaturation step of 98°C for 1 s, 2) annealing step of melting temperature (Tm) + 5°C for 5 s , then 3) an extension step of 15 s/1kb at 72°C. The PCR reaction ends with a final extension of 72°C for 1 minute.

2.1.2 Restriction Enzyme Digests

New England Biolabs (NEB) restriction enzymes were used as indicated in **Table 2.1** to clone constructs into different vectors. In each reaction, 1 μl of each restriction enzyme was incubated with 3 μl of 10x CutSmart Buffer (NEB) and 25 μl of template at 37°C for at least 2 hours. Cut DNA was separated from other components using the PCR purification protocol (QIAGEN) for small inserts or the QIAquick gel extraction kit (QIAGEN) for vectors, according to the manufacturer's manuals.

Vector	Forward RE	Reverse RE
pET15bcon	BamHI	NheI
pExpreS^2-1	EcoRI/XbaI	NotII
pHLsec	BamHI/AgeI	KpnI

Table 2.1: Restriction Enzymes for Cloning into Different Vectors

2.1.3 Ligation

Ligation of cut inserts and vectors was carried out using T4 DNA Ligase. Normally, 1 μl cut vector (around 100ng) with 7 μl cut insert was incubated with 1 μl of 10x DNA T4 Ligase Buffer and 1 μl T4 Ligase Enzyme at 25°C for 16 hours at room temperature (RT).

2.1.4 Transformation of ligation reactions

50 μl of subcloning competent DH5a cells (NEB) were incubated on ice with 5 μl of ligated DNA for 20 minutes. The cells were then heat-shocked for 45 seconds at 42°C. 250 μl of Lysogeny broth (LB) medium was then added to the cells and incubated for 1 hour in the 37°C shaking incubator before plating on appropriate antibiotic-selective agar plates and left overnight at 37°C. Colonies that grew were then picked and incubated in 10 ml LB medium with the appropriate antibiotic overnight at 37°C. The plasmid DNA was extracted by centrifugation at 10,000 x g for 5 minutes and then by following the manufacturer's protocol of the QIAprep Spin Miniprep kit (Qiagen). DNA sequencing was performed using the Sanger Sequencing service by Source Bioscience using appropriate sequencing primers.

2.1.5 Site Directed Mutagenesis

Site Directed Mutagenesis (SDM) was carried out by designing a pair of complementary primers that flank the site of mutagenesis by at least 15 base pairs, complementary to the vector. These primers were used to amplify the whole vector with the mutated DNA sequence by PCR as described in **Section 2.1.1**. This amplified DNA was then treated with DpnI (NEB) for 16 hours at 37°C to digest the methylated template DNA. The remaining PCR product was transformed into XL-10 Gold Ultracompetent cells (Agilent Technologies), as per manufacturer's protocol and plated on agar plates with the appropriate antibiotic. These plates were incubated for 16 hours at 37°C. Colonies that grew were individually picked and grown in 10 ml of LB overnight and the vector DNA was extracted from these bacteria using a QIAprep miniprep columns (QIAGEN). The DNA was sequenced using the Sanger Sequencing service (Source Bioscience).

2.1.6 Gibson Assembly Cloning

This method of cloning uses an exonuclease-based method to assemble different strands of DNA, allowing large fragments of DNA to be joined together in the correct order. Primers were designed according to the guidelines that came with the Gibson Assembly® Master Mix - Assembly (NEB) and the reaction carried out as per manufacturer's protocol. The master mix contains a 5' exonuclease, which generates 5' long overhangs on the blunt ended PCR products, which allows the complementary PCR fragments to align and anneal. A polymerase then fills in the gaps of the annealed single strand regions, and a ligase enzyme seals the nicks. In short, 100 ng of linearised vector was mixed with 5x molar of the insert and the Gibson Assembly Master Mix (2x) for 1 hour at 50°C. 2 μL of the Gibson reaction was then transformed into NEB 5-alpha Competent *E. coli* cells, which were provided

in the kit.

2.1.7 Agarose Gel Electrophoresis

DNA samples were visualised by gel electrophoresis on a 1% agarose gel made with Tris-acetate-EDTA (TAE) buffer, with SYBR Safe (Invitrogen, 10,000X). Samples were loaded with Gel Loading Dye, Purple (6X) (NEB). Both 1 kb and 100 kb molecular markers (NEB) were used to determine fragment size. Gels were run at 100 V for 30 minutes and DNA was visualised with 302 nm UV-B light. Gel extraction was performed using a QIAquick Gel extraction kit (Qiagen) and PCR purification was performed using the QIAquick PCR purification kit (Qiagen), both according to manufacturer's protocol.

2.2 Protein Expression

2.2.1 Bacterial Expression in *Escherichia coli* cells

Escherichia coli is the simplest and quickest expression system in use by the lab. There are different strains of *E. coli* that were used during this PhD, mostly BL21-DE3 (NEB) unless otherwise stated. All strains made use of the T7 promoter, controlled by the lac operator, which allowed inducible expression with isopropyl β-D-1-thiogalactopyranoside (IPTG) of the recombinant protein of interest. Constructs for expression in BL21-DE3 cells were cloned into pET15b converted to contain an N-terminal TEV-cleavable His-tag **(Figure 2.1)**.

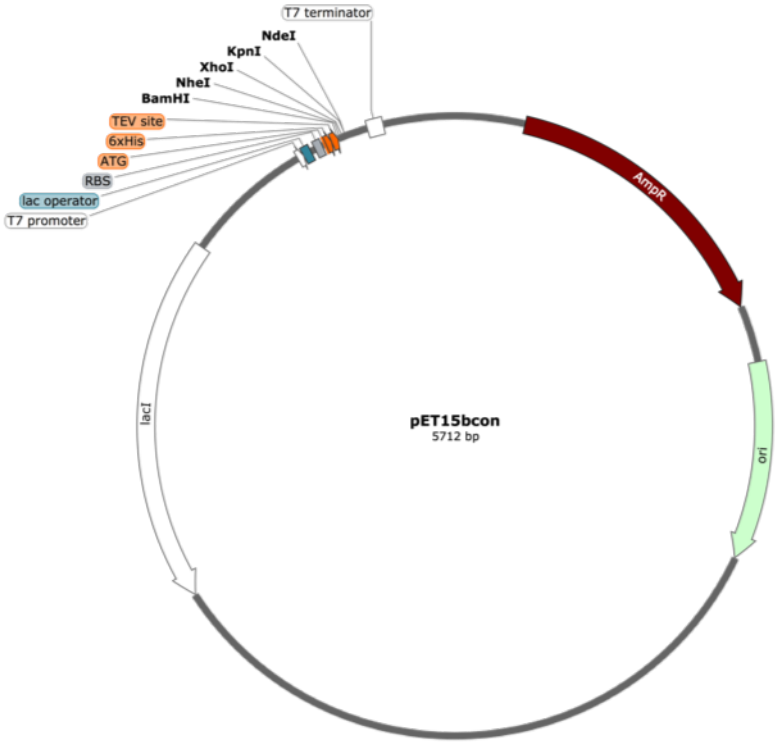


Figure 2.1: pET15bcon vector map The pET15b plasmid was originally obtained from Novagen (EMD Millipore). The multiple cloning site of this plasmid was modified by Professor Matthew Higgins to contain an N-terminal 6-His tag after the start codon (ATG) followed by a TEV site. The pET15bcon plasmid contains an Ampicillin resistance gene (AmpR) which allows for selection of cells which contain the pET15bcon plasmid in the presence of Ampicillin.

2.2.1.1 Large Scale Expression

To produce large amounts of protein for biophysical analysis, structural studies and immunisations, typically 6-12 L of media was used. Either freshly transformed cells or sterile glycerol stocks were used to innoculate 100 ml LB with the appropriate antibiotic, which were grown overnight at 37°C. The following day 6 ml of the overnight culture was diluted into each autoclaved litre of 2 x Yeast-Tryptone Broth (2xYT, Melford). The cells were grown in a shaking incubator at 37°C until an optical density at 600 nm (OD_{600}) between 0.6 and 0.8 was reached, when they were induced with a final concentration of 1 mM IPTG. The temperature was

then reduced, normally to 25°C, (unless otherwise stated and the cells were either harvested after 3 hours or overnight. The cells were harvested by centrifugation at 5,000 x g for 20 minutes in a Beckman Coulter Avanti J-20 centrifuge (JLA-8.1000 rotor). The cell pellets were then stored in the freezer at -20°C until ready for protein purification.

2.2.2 S2 expression

2.2.2.1 Maintenance of S2 cells

Schneider cells (S2 cells) were derived from a primary culture of late culture *Drosophila melanogaster* embryos (Schneider). Untransfected S2 cells were maintained in EX-CELL® 420 Serum-Free Medium for Insect Cells (Sigma-Aldrich), supplemented with 10% FCS (fetal calf serum, Gibco), 100 U penicillin and 0.1 mg/ml streptomycin (Sigma). These cells were split every 3-4 days to 5 x 10^6 cells/ml. Stable S2 cells were maintained in media that was not supplemented with FCS and were split every 3-4 days to 8 x 10^6 cells/ml.

2.2.2.2 Transfection to form stable S2 cell lines

Constructs for expression in S2 cells were cloned into pExpresS^2-1 vector **(Figure 2.2)**. The day before transfection, the S2 cells were spun down and resuspended to a density of 5 x 10^6 cells/ml in fresh FCS-containing medium. On the day of transfection the cells were spun down again and resuspended to a density of 2 x 10^6 cells/ml in serum-free media. The stress of spinning the cells down two days in a row is thought to increase the transfection efficiency, as well as ensure that number of cells in each transfection is consistent as cell debris will be removed. For each transfection, 50 μl of ExpresS^{2}Insect-TRx5 (ExpresS^{2}ion Biotechnologies) was added to 5 ml of cells (at 2 x 10^6 cells/ml), followed by 12.5 μg DNA. These mixtures

were transferred to 25 cm^2 canted neck cell culture flasks (Corning®) and incubated at 25°C. Four hours post transfection, 12.5% FCS was added to each flask, and 24 hours post transfection, antibiotic selection was started by the addition of 1.5 mg/ml zeocin. Over three weeks, a polyclonal cell line was selected for, replacing 2 or 3 ml of media with fresh media containing 2 mg/ml zeocin twice a week.

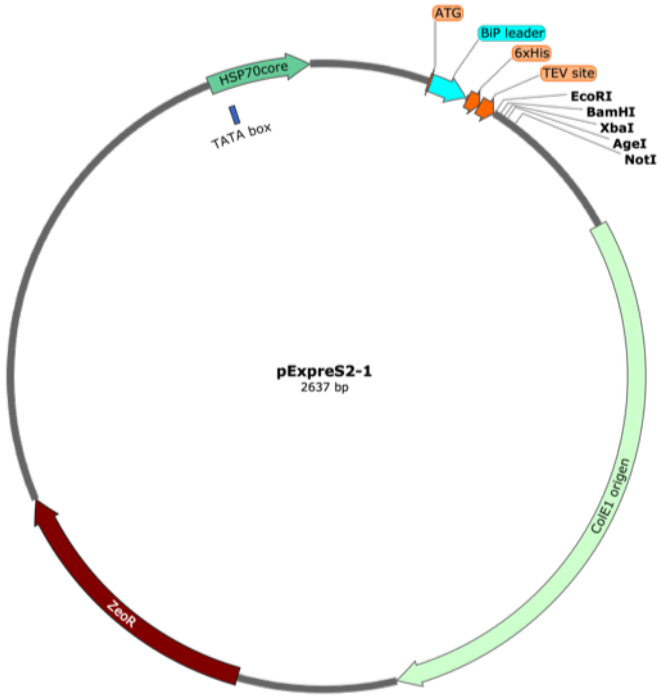


Figure 2.2: pExpreS^2-1 vector map pExpreS^2-1 plasmid was obtained from ExpreS2ion Biotechnologies. The multiple cloning site of this plasmid was modified by Professor Matthew Higgins to contain an N-terminal BiP leader sequence to direct the recombinant fusion protein through the secretory pathway of S2 cells into the culture medium. In addition a 6-His tag after the start codon (ATG) followed by a TEV site was inserted. The pExpreS^2-1 plasmid contains an Zeocin resistance gene (ZeoR) which allows for selection of cells which contain the pExpreS^2-1 plasmid in the presence of Zeocin.

2.2.2.3 Protein expression in S2 cells

Once the stable S2 cell lines had been established, they were expanded in EX-CELL® 420 medium without zeocin or FCS. For expression trials, the cells were expanded

to 25 ml. Once protein expression had been confirmed, usually by Western Blot, the cell lines could be expanded further (usually up to ∼8L). Cells were usually harvested four days after their final split and the supernatant retained for protein purification.

2.2.3 Lepidopteran cell/baculovirus expression

2.2.3.1 Co-transfection of DNA

Constructs for expression in the lepidopteran cell/baculovirus expression system were cloned into a converted transfer vector, pAcGP67A **(Figure 2.3)**. The converted pAcGP67A contained an N-terminal His6-tag, TEV cleavage site and a multiple cloning site (MCS). Co-transfection of sterilised recombinant DNA plasmid into *Spodoptera frugiperda 9* (SF9) cells (Invitrogen) was performed using a BD Baculogold trasnfection Kit (BD Biosciences), according to the manufacterer's protocol. Supernatant consisting of viral particles were used for plaque assays, from which viral stocks were amplified in three steps. The final viral stock (P3) could then be used to inoculate the insect lepidopteran cells.

2.2.3.2 Expression of recombinant protein in High-Five cells

High throughput expression of soluble protein was performed in High-Five (*Trichopulsia ni*) cells. These cells were inoculated with P3 viral supernatant at a ratio of 0.2 pfu/cell (50 ml of P3 viral supernatant per 500 ml of High-Five cells at 1 x 10^6 cells/ml). The cells were incubated at 27°C for 4 days, before the cell supernatant was harvested by centrifugation (5000 rpm, 20 minutes, 4°C).

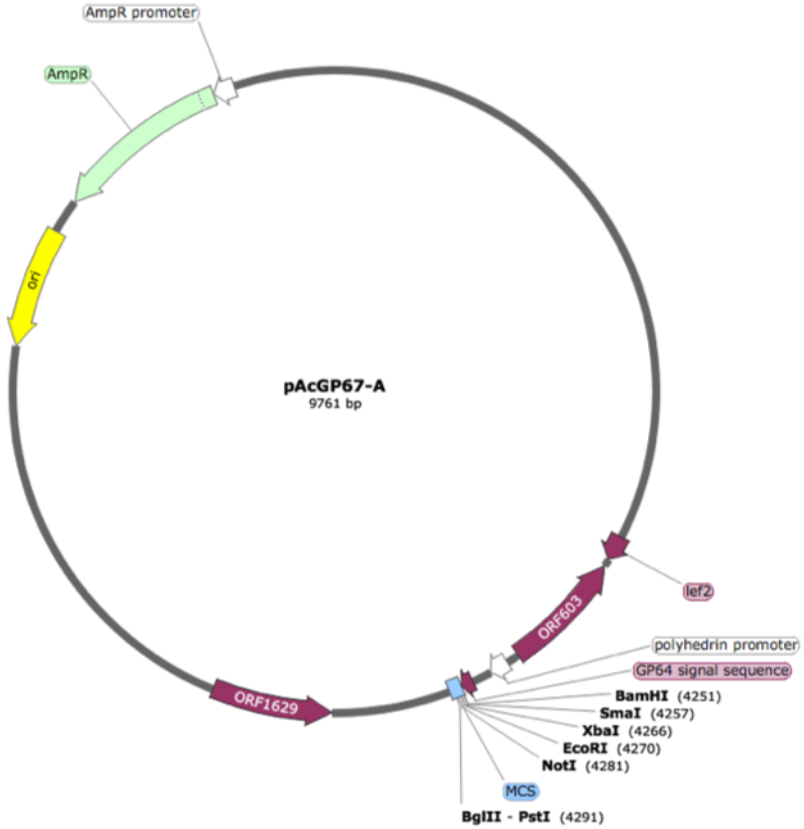


Figure 2.3: pAcGP67A vector map pAcGP67A originated from BD biosciences. The multiple cloning site contains an N-terminal GP64 leader sequence to direct the recombinant fusion protein through the secretory pathway of lepidpoteran cells into the culture medium. The pAcGP67A plasmid contains an Ampicillin resistance gene (AmpR) which allows for selection of cells which contain the pAcGP67A plasmid in the presence of Ampicillin.

2.2.4 Human embryonic kidney (HEK) cells expression

During this DPhil, the Expi293F^{TM} system was used for HEK cell expression (ThermoFisher Scientific). This system originates from a 293 permanent cell line established from primary embryonal human kidney transformed with human adenovirus type 5 DNA, which activates viral promoters increasing recombinant protein expression. The newer Expi293^{TM} system has been optimised to increase growth rate and protein yield compared to the original FreeStyle^{TM} 293-F cells.

2.2.4.1 HEK cell maintenance

The Expi^{TM}293F cells were maintained in suspension culture in 125 ml Corning Erlenmeyer Shaker Flasks at 37°C, 8%CO_2 in the air, shaking at 125 rpm. The Expi^{TM}293F cells grew quickly reaching between 3 x 10^6 and 5 x 10^6 cells/ml every 2-3 days. They were then split to 0.5 x 10^6 in Expi293^{TM} Expression Medium.

2.2.4.2 Transient transfection of HEK cells

The plasmid pHLsec was used for transient HEK cell transfections **(Figure 2.4)**. For the Expi293F^{TM} cells, the ExpiFextamine^{TM} 293 Transfection Kit (ThermoFischer Scientific) was used for transient transfection. In short, for each 100 ml transfection, on the day before transfection around 100 ml of cells were passaged to $-$ 3 x 10^6 cells/mL. After 24 hr, the cells were diluted in Expi293^{TM} Expression Medium to 3 x 10^6 cells/mL. For each transfection 100 μg DNA was diluted into a total volume of 5 ml Opti-MEM I medium. In a separate tube, 267 μL of ExpiFectamine 293 Reagent was diluted to a total volume of 5 ml in Opti-MEM I medium. After an incubation of 5 min, these two tubes were combined to form a total volume of 10 ml containing both the ExpiFectamine 293 Reagent and the DNA. These were left for 30 minutes to allow the DNA-ExpiFectamine Reagent complex to form. After this incubation period, the 10 ml DNA-ExpiFectamine^{TM} 293 Reagent complex was added to each 100 ml cells. Approximately 16-18 hours post transfection, 500 μL of ExpiFectamine 293 Transfection Enhancer 1 and 5 ml of ExpiFectamine 293 Transfection Enhancer 2 was added to each flask. After 3 or 4 days post-transfection, the cells were spun down by centrifugation (5000 rpm, 20 minutes, 4°C) and the supernatant collected, filtered and kept at -20°C, until the protein was purified.

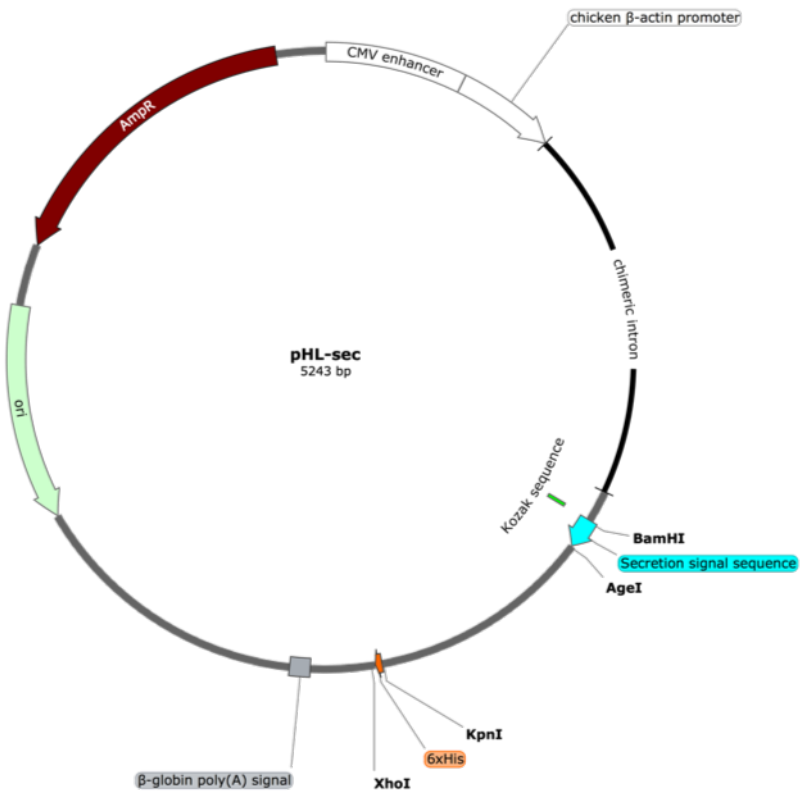


Figure 2.4: pHLsec vector map The pHLsec plasmid was originally kindly gifted from Edith Yvonne Jones. The multiple cloning site contains an N-terminal signal sequence to direct the recombinant fusion protein through the secretory pathway of HEK cells into the culture medium. In addition, a C-terminal 6-His tag for purification was included. The pHLsec plasmid contains an Ampicillin resistance gene (AmpR) which allows for selection of cells which contain the pHLsec plasmid in the presence of Ampicillin.

2.3 Protein Analysis

2.3.1 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

Protein samples were mixed with 2 x loading buffer (0.02 M Tris-HCl, pH 8; 10% SDS (w/v); 20% glycerol; 0.05% bromophenol blue) and boiled at 95°C for 5 minutes to denature and coat proteins in the negatively charged sodium dodecyl sulphate

(SDS). The samples were then loaded into hand-cast gels (with varying % acrylamide depending on the size of protein and desired separation) and run at 200 mV for 40 minutes. The running buffer for the gel electrophoresis consisted of 30 mM Tris pH 7.5, 192 mM glycine, and 0.1% SDS. The gels were stained with a fast, Coomassie-based stain: InstantBlue^{TM} (Expedeon).

If the quantities of protein run on the gel were too low to detect with the traditional coomassie staining, the gel could be silver stained instead. This was achieved using the Silver Stain Plus^{TM} Kit (Bio-Rad Laboratories Inc.) carried out as per the manufacturer's protocol.

2.3.2 Western Blot

For Western blotting, a polyacrylamide gel was run as above, however instead of staining the gel, the protein from the gel was transferred onto an Immobilon PVDF transfer membrane (Merck Millipore) using a semi-dry Western blotting system (Bio-Rad Laboratories Inc.) for 1 hour at 50 mA. Prior to transfer, the PVDF membrane was rinsed in 100% methanol, before the membrane, gel and blotting paper were soaked in transfer buffer: 50 mM Tris pH 9.2, 40 mM glycine, and 20% methanol. After transfer was complete, the membrane was incubated with 20 ml Pierce protein-free blocking buffer (Life Technologies, UK) for 1 hour to block any sites on the membrane not bound to protein. The penta-His-HRP (horseradish peroxidase) conjugate IgG antibody (Qiagen; 1:5000 dilution), which detects any protein with a histidine-tag, was then incubated with the membrane in 20 ml blocking buffer. The membrane was then further washed Tris buffered saline (TBS; 20 mM Tris-HCl pH 8, 500 mM NaCl) with added Tween-20 to a final concentration of 0.2% for at least one hour, renewing the buffer four times. Finally, the blot was developed using ECL Prime Western Blotting Detection Kit (Amersham), with bands visualised using an Odyssey Fc Dual Mode Imaging System (LI-COR). The coloured marker

was detected by exposure of the membrane to 700 nM light and the histidine-tagged protein bands were detected by using a chemiluminesence detector.

2.4 Protein Purification

2.4.1 Sonication

E. coli cell lysis was achieved using a Misonix S4000 (Qsonica) sonicator. Typically the harvested cells were resuspended in 20 mM Tris pH 8, 300 mM NaCl, 15 mM imidazole, 0.05% TRITON X-100, 10% glycerol (lysis buffer) and lysed at a power of 70 W, 3 seconds on and 7 seconds off for a total time of 3 minutes 30 seconds. The whole cell lysate was centrifuged at 50,000 x g for 30 minutes at 4°C.

2.4.2 Tangential flow filtration

For soluble, secreted proteins expressed in the S2, baculovirus and HEK cell expression systems, the cell supernatant was harvested by centrifugation at 5,000 x g for 30 minutes. The cell supernatant was filtered using Steritop-GP 0.44μM filter units (Millipore). The filtered supernatant was then buffered exchanged, to remove additives in the media that can prevent protein binding to Ni-NTA, using a Centramate^{TM} tangential flow filtration (TFF) system (Pall Corp., USA). The TFF machine was firstly equilibrated with 2 L of suitable buffer (typically 20 mM Tris pH8, 150 mM NaCl and 20 mM imidazole) before up to 3 L of cell supernatant were loaded onto the TFF at $\backsim$ 5 mL/min. The system was then washed with at least 1 L of buffer and then the supernatant eluted in 150 ml in the new buffer. In this way, as well as performing buffer exchange, the TFF also concentrates the sample 10-20 fold.

2.4.3 Ni-NTA affinity Chromatography

Soluble His-tagged proteins from *E. coli* or eukaryotic cell supernatants were purified using Ni-NTA beads (QIAGEN). Ni-NTA uses the chelating ligand nitrilotriacetic acid (NTA), which is coupled to cross-linked 6% agarose resin. The Ni-NTA beads were first equilibrated with 30 column volumes (CVs) of nickel binding buffer (20 mM Tris pH8, 300 mM NaCl, 20 mM imidazole), before the soluble protein was incubated with the beads and the flow-through was collected via gravity flow. The beads were then washed with 50 CVs of nickel binding buffer to remove any non-specific binding and the wash collected. Finally, the His-tagged proteins were eluted from the Ni-NTA beads using the nickel elution buffer (20 mM Tris pH 8, 300 mM NaCl, 200 mM imidazole).

2.4.4 Insoluble preparations

Some *E .coli* proteins were expressed in the insoluble inclusion bodies rather than solubly and to purify these proteins, the lysate pellet was used.

For most proteins in this thesis the following methods was used. The lysate pellet was washed with 1 ml of lysis buffer and resuspended in 6 M guanidine, 20 mM Tris pH 8, 300 mM NaCl, 20 mM imidazole, 10 mM β-mercaptoethanol (unfolding buffer). This was left unfolding for at least 2 hours rolling at room temperature, before centrifugation at 50,000 x g at 4°C for 30 minutes. The supernatant was then incubated on Ni-NTA beads, which had been pre-equilibrated in unfolding buffer. The Ni-NTA beads with the bound protein were washed with 10 CVs of 6 M guanidine, 20 mM Tris pH 8, 300mM NaCl, 20 mM imidazole with the addition of a glutathione redox buffer (300μM oxidised glutathione, 3 mM reduced glutathione) which allowed the shuffling of the disulphide bonds as the chaotropic agent was removed. To refold the bound His-tagged protein, a decreasing step-wise 10% gradient of guanidine was

applied over the column until no guanidine was present. At each step, the bound protein was incubated with each buffer for 10 minutes to allow time for the protein to refold. The refolded protein was then eluted using the nickel elution buffer (20 mM Tris pH 8, 300 mM NaCl, 200 mM imidazole).

2.4.4.1 *Pv*DBP-RII Purification

For some proteins, like *Pv*DBP-RII, purifying the cell pellet further before solubilisation in guanidine resulted in a higher yield of refolded protein. The inclusion body preparation consisted of multiple cycles off sonication, centrifugation and resuspension of the pellet in the next buffer. The lysate pellet was resuspended in a high detergent lysis buffer (20 mM Tris pH 8, 300 mM NaCl, 15 mM imidazole, 1% TRITON X-100, 10% glycerol) to remove solubilised membrane components and proteins. After sonication and centrifugation the pellet was then resuspended in a high salt lysis buffer (20 mM Tris pH 8, 1 M NaCl, 15 mM imidazole, 0.05% TRITON X-100, 10% glycerol) to increase the osmolarity and further solubilise other proteins expressed inside the cell. This pellet was resuspended in normal lysis buffer (20 mM Tris pH 8, 300 mM NaCl, 15 mM imidazole, 0.05% TRITON X-100, 10% glycerol) as a final wash step. After another round of sonication and centrifugation, the final pellet was resolubilised in 6 M guanidine, 20 mM Tris pH 8, 300 mM NaCl, 20 mM imidazole, 10 mM β-mercaptoethanol for at least 2 hours at room temperature. After centrifugation at 22,000 rpm for 30 min, the supernatant was filtered through a 0,22 μm membrane (Millipore) and was refolded by dropwise adding to 2L flash dilution buffer (400 mM L-arginine, 10 mM EDTA, 0.1 mM phenylmethanesulfonylfluoride, 3 mM reduced glutathione, 0.3 mM oxidized glutathione, pH 8.0). After being stirred at 4°C for >36 hours, the precipitate was removed by spinning at 15,000 x g. The refolded protein was buffer exchanged using the TFF and concentrated to 100ml in 20 mM Tris, 300mM NaCl, 20mM imidazole pH 8.0. This was then loaded on a

NiNTA column overnight. Protein was eluted in 20 mM Tris-HCl pH 8, 300 mM NaCl, 200 mM imidazole.

2.4.5 TEV protease cleavage

A laboratory TEV construct was expressed in BL21-DE3 and purified using it's His-tag on a NiNTA column. In order to remove the His-tag from tagged proteins, the elution fractions from Ni-NTA were pooled and dialysed into phosphate buffered saline (PBS). TEV protease was added at a weight-to-weight ratio of 1:10 (TEV:His-tagged protein), together with 300 μM oxidised glutathione and 3 μM reduced glutathione. This mixture was incubated overnight at room temperature, before being applied over Ni-NTA beads (pre-equilibrated in nickel binding buffer) to separate the un-tagged protein of interest. The flow-through was collected since this contained the un-tagged recombinant protein, whereas the elution contained any uncleaved His-tagged protein and the His-tagged TEV protease.

2.4.6 Size exclusion chromatography

Size exclusion chromatography (SEC) was used as the final step in the purification protocol, in order to separate monodisperse protein from other contaminants, including aggregates. Samples were concentrated to at least 2 ml using Amicon Ultra-15 Centrifugal Filters (Millipore) with an appropriate molecular weight cut-off. Before loading onto the SEC column, the sample was centrifuged at 10,000 x g for 2 minutes to remove any aggregates or precipitate. SEC was carried out using either an *Äkta* pure or an *Äkta* purifier (GE Healthcare UK). The concentrated sample was loaded on an appropriate column: HiLoad Superdex 75 10/30 or 16/60 (GE Healthcare), HiLoad Superdex 200 10/30 or 16/60 (GE Healthcare) or Superose 6 Increase 10/30 (GE Healthcare) depending on the size of the protein of interest. The fractions where

there was an A_{280} peak were analysed by SDS-PAGE and those judged to contain pure protein were pooled.

2.4.7 Protein Deglycosylation

Recombinant proteins expressed in the eukaryotic expression systems were often glycosylated, which could hinder crystallisation. In order to deglycosylate proteins expressed in S2 and the lepidopteran stable cell lines, these proteins were buffer exchanged into 20 mM MES pH 6.5, 150 mM NaCl. EndoHf (NEB) and EndoF3 (made in the lab) were added in a 1:100 molar ratio (endoglycosidase:recombinant protein) and incubated overnight at room temperature. For proteins expressed in HEK cells, the protein was buffer exchanged into PBS and PNGaseF (NEB) was added in a 1:100 molar ratio to the recombinant protein and incubated overnight at room temperature.

2.4.8 Protein Biotinylation

2.4.8.1 Enzymatic

Proteins cloned with a biotin acceptor peptide (BAP) tag could be biotinylated using the enzyme BirA (Sigma Aldrich). 20μg of BirA was added to 1 mg of protein at 30 μM with 0.3 mM biotin and 5 mM ATP. This was incubated at 25°C overnight then buffer exchanged using a Sephadex G-25 PD10 Desalting column (GE Healthcare, USA) into PBS.

2.4.8.2 Chemical

The *N*-Hydroxysuccinimide (NHS) esters of biotin react efficiently with primary amino groups to form stable amide bonds, and so all lysine side chains can be biotin

labelled in this way. EZ-Link™ Sulfo-NHS-Biotin Reagent (ThermoFisher Scientific) was used as per manufacturer's protocol. In short, the EZ-Link™ Sulfo-NHS-Biotin Reagent was solubilised in water and mixed with the protein to be biotinylated in a non-Tris buffer (20 mM Hepes pH7.5, 150 mM NaCl) for 1 hour at room temperature. Any excess non-reacted biotin and reaction by-products were subsequently removed by SEC.

2.4.9 Protein Methylation

In order to reduce flexibility of lysine side chains to promote crystallisation, often proteins were methylated by the following protocol. The protein sample was concentrated to 5 mg/ml and buffer exchanged by dialysis or using a Sephadex G-25 PD10 Desalting column (GE Healthcare, USA) into PBS. The tube was covered with aluminium foil and per 1 ml of protein 20 μL of 1M dimethylaminoborane (DMAB, Sigma Aldrich) followed by 40 μL of 1M formaldehyde solution (Sigma Aldrich) was added. The aluminium covered tube was then shaken at room temperature in the dark for 2 hours. Another 20 μL of 1M DMAB followed by 40 μL of 1M formaldehyde solution per 1ml was added, before being returned to the shaker in the dark for 2 hours. This step was repeated one more time. Finally, 10 μL of DMAB per 1 ml of protein was added and incubated overnight in the dark at room temperature whilst shaking. The methylated protein was then put on a SEC column run with 20 mM Tris pH8, 150 mM buffer to further purify and buffer exchange to stop the reaction.

2.4.10 Trimming of the CSA Polysaccharide

Chondroitin sulphate A (CSA) was bought as a polysaccharide (Sigma Aldrich). Smaller fragments of the repeating dimer could be produced by incubation with the

enzyme Chonroitinase ABC (Sigma Aldrich). Chondroitinase ABC catalyzes the eliminative degradation of polysaccharides containing (1-4)-β-D-hexosaminyl and (1-3)-β-D-glucuronosyl or (1-3)-α-L-iduronosyl linkages to disaccharides containing 4-deoxy-β-D-gluc-4-enuronosyl groups. The CSA was solubilised in 20 mM Tris pH8, 150 mM NaCl and mixed in a 1.5:1 molar ratio with Var2CSA (CSA:Var2CSA). Chonroitinase ABC was then added at a 1:100 volume ratio and incubated for 2 hours at 37°C. The enzyme and any cleaved and unbound polymers were removed from the Var2CSA:CSA complex by running on a Superose 6 10/30 SEC column (GE Healthcare).

2.5 Antibody Purification

2.5.1 Monoclonal Antibody Isotyping

The isotype of purified mouse mAbs was determined by using Thermo Scientific^{TM} Pierce^{TM} Rapid Mouse Antibody Isotyping Kits as per manufacturer's protocol.

2.5.2 Monoclonal Antibody Fragment Production

For structural characterisation of a protein's interaction with mAbs, the Fab fragment of the antibody was obtained by cleaving the mAb with specific proteases, depending on the origin and isotype of the mAb.

2.5.2.1 Ficin

All mouse antibodies used in this thesis were of isotype IgG1 and were cleaved with Ficin, which degrades mAb into Fab. 1 ml of immobilised Ficin (ThermoFisher

Scientific) settled resin (2 ml of suspension) was packed into a column. The immobilised Ficin was equilibrated with 0.1 M citrate buffer pH6, 5 mM EDTA, 25 mM cysteine (equilibration buffer). The mouse antibody was concentrated to 3 ml and buffer exchanged into the equilibration buffer using a Sephadex G-25 PD10 Desalting column (GE Healthcare). The mouse IgG1 sample was added to the resin and incubated at 37°C overnight. Since Ficin cleaved the Fc domain into small fragments no further purification was required to obtain the Fab fragments.

2.5.2.2 Papain

All human antibodies used in this thesis were also isotype IgG1 and were cleaved with Papain, which results in production of Fab and Fc fragments. 1 ml of immobilised Papain (ThermoFisher Scientific) settled resin was packed into a column. The immobilised papain was equilibrated with 20 mM sodium phosphate pH 7.0, 10 mM EDTA, 20 mM cysteine (Papain Digestion Buffer). The IgG1 was concentrated to 3 ml and buffer exchanged into the Papain Digestion Buffer. The IgG and papain resin were incubated overnight at 37°C maintaining constant mixing. The Papain resin was separated from the digest by applying to a small column and the enzyme was washed with 20 mM sodium phosphate pH7, 10 mM EDTA. The flow-through and wash were combined, and the Fab fragments were separated from the undigested IgG and Fc fragments using an Immobilised Protein A column (Pierce^{TM}) as described in the next section.

2.5.2.3 Protein A Purification

1ml of Protein A agarose resin was packed into a column and equilibrated with 20 mM sodium phosphate pH7 (Protein A equilibration buffer). The mixture from the papain digestion (already in the Protein A equilibration buffer) was incubated on the column for at least 1 hour. The flow-through was collected and the resin washed

with 10CV of Protein A Equilibration Buffer. Since protein A specifically binds to the Fc region of the antibody, the Fab fragments were present in the flow-through and wash. Any undigested mAb and Fc fragments, bound to the protein A column was then eluted in 0.1M glycine pH 2.5. These elutions were quickly neutralised by the addition of 100 μL of 1M Tris pH8 per every 1 ml of elution.

2.6 Biophyscial Methods

2.6.1 Circular Dichroism

2.6.1.1 Background

Circular dichroism (CD) spectroscopy is a form of light absorption spectroscopy that measures the difference in the absorbance of right- and left-circularly polarised light by a protein. CD spectra between 260 nm and 190 nm can be analysed for the different secondary structural types due to the chirality of the molecules held in a particular secondary structure: α-helix, β-sheet and coiled-coil [146, 147]. The stability of the folded protein can be assessed by attaching a water bath to the spectrometer, which allows the affect of increasing temperature on the secondary structure to be assessed.

2.6.1.2 Experiment

The J-815 spectropolarimeter (JASCO, Japan) with attached Pelitier water bath was used for CD experiments. The protein was buffered exchanged into 10 mM phosphate pH 7.5, 150 mM NaF, and concentrated/diluted to a concentration of 0.2 mg/mL. A 300 μl sample was pipetted into a quartz cuvette (Starna Scientific Ltd, UK) for each experiment. Single circular dichroism spectra were measured between

190 nm and 260 nm at 25°C. Four spectra were averaged per sample, subtracting the average measurement of the buffer blank. For thermal melt experiments, CD spectra were measured between 200 nm and 250 nm at 2°C intervals between 20°C and 90°C.

2.6.2 Surface Plasmon Resonance

2.6.2.1 Background

The phenomenon of surface plasmon resonance (SPR) is exploited to analyse macromolecular interactions. A surface plasmon is a "propagating electron density wave occurring at the interface between a metal and dielectric" [148] and travels parallel to the surface. Associated with the plasmon is an electric field, known as the evanescent field, which is perpendicular to the surface and decays exponentially over distance. This field makes the plasmon sensitive to changes in the dielectric constant near the surface and provides the mechanism by which interactions can be probed.

In SPR, a thin gold chip is mounted onto a prism which is then subjected to a beam of plane polarised light. The intensity of the light reflected from the metal surface is measured as a function of the incident angle of light on the chip, which shows a reduction in intensity where the energy is being absorbed by excitation of surface plasmons. The angle with minimum reflection intensity is where resonance occurs and is highly dependent on the dielectric constants of the medium on either side of the gold chip due to the interaction with the evanescent field. The dialectric constant on the prism side of the chip is constant, so any change in the angle are due to an interaction with the other site of the chip.

By coupling a protein to the upper surface of the chip, an interaction with another protein will alter the dielectric constant which will be observed as a change in the angle of minimum reflected intensity. By subtraction of a reference with no coupled

protein, the change in angle is directly proportional to the amount of bound protein. Therefore, this technique allows the study of protein-protein interactions in real time, yielding a sensogram which shows the change in SPR angle (i.e. amount of bound protein) versus time.

2.6.2.2 Experiment

All SPR experiments were carried out in a Biacore T200 instrument (GE Healthcare) with the following SPR buffer: 20 mM Hepes pH7.5, 150 mM NaCl, 0.005% Tween-20. Two different SPR chip systems were used in this thesis: a CM5 chip to which proteins could be amine-coupled, and a CAP chip to which a biotinylated protein could be coupled to the streptavidin surface (both GE Healthcare).

2.6.2.2.1 SPR using a CM5 chip This gold plated chip was activated using the Amine Coupling Kit (GE Healthcare). The protein (usually a mAb) in a low pH buffer (to increase the positive charge) was immobilised to the chip to a total of 1500-2000RU. After immoblisation, any unbound sites on the chip were blocked using ethanolamine. Since a tag is not being used, this allowed the attachment of the protein of interest in multiple different orientations. The whole arrangement meant hopefully every orientation was represented. A concentration series of each analyte starting at 1 μM was injected over the mAb-coated chip for 240 seconds at 30 μL/min, followed by a 300 seconds dissociation period. The chip surface was regenerated between each sample with a 30 second pulse of 10 mM NaOH, that removed any bound analytes.

2.6.2.2.2 SPR using CAP chip This gold-plated chip is coated with a layer of single stranded DNA (ssDNA). In each run, in order to activate the chip, streptavidin attached to complementary ssDNA to that on the chip was flowed over, allowing

the coating of the chip in streptavidin (this was the kit's capture reagent). The biotinylated EPCR was coupled to the streptavidin layer to around 150RU. A concentration series from 1 μM of the second component was flown over the chip at 30 μL/min with a 240 second association time and 300 second dissociation time. The chip was regenerated between each run using a 120 second pulse of 6 M Guanidine, 0.25 M NaOH, which removed the capture reagent, biotinylated protein and analyte.

2.6.2.2.3 Data Analysis The BiAevaluation software was used to fit the kinetic binding of the analyte to the chip at different concentrations in order to determine the association rate constant (k_{on}), dissociation rate constant (k_{off}) and from those the affinity (K_D).

2.7 Crystallographic Methods

2.7.1 Background

In protein crystallography, an incident beam of X-rays is directed at the pure crystalline protein. Most of the beams pass straight through the sample but some are diffracted by the electron shells that surround each atom in the protein molecule. In crystalline latices, the proteins are organised in a regular, repeating array, which allows the signal of each diffracted beam to be amplified and seen as a diffraction spot on a detector. The positions and intensities of these diffraction spots (providing information about frequency and amplitude, respectively) are combined with phase information to produce a three-dimensional map of the electron density within the crystal. The pipeline of the crystallography process, from crystallisation to data refinement, are described in the following sections.

2.7.2 Protein Crystallisation

2.7.2.1 Broad Initial Screens

The first step is forcing the protein out of solution into an ordered crystalline lattice. For all proteins in this thesis, crystallisation was attempted using the sitting-drop vapour diffusion method in MRC 2-well crystallisation plates (Swissci, Hampton Research USA). In each commercial screen, there is 50 μl of reservoir solution in each well and 100 nl of protein is dispensed in each drop, mixed with 100 nl of well solution by a mosquito (TTP Labtech, UK). As the water evaporates from the drop, the concentration of both the protein and the precipitant in the crystallisation solution increase until ideally supersaturation occurs. If suitable nucleation sites are present, then crystals may grow. There are many different factors that influence whether crystals will grow, including the protein itself, protein concentration, temperature, precipitant, buffers, and additives. Therefore, when attempting to crystallise proteins broad commercial screens are set up, which contain different mixtures of different precipitants, buffers, additives, and detergents, all to maximise the chance that the condition in which the protein crystallises is discovered. The initial broad screens generally set up (unless otherwise stated) were JCSG+, Pact Premier, MIDAS, and Morpheus, with ProPlex used for protein complexes (all Molecular Dimensions).

2.7.2.2 Fine Screens and Optimisation

If initial crystals form, then optimisation of the initial crystallisation condition may be required to improve the diffraction-quality of the crystal. Typically, a condition was optimised using the additive screen, Silver Bullets (Hampton Research), in which 50 nl was added to each droplet. Moreover, in cases where overnucleated crystals were seen, initial crystals were harvested, vortexed with a Seed-Bead (Hampton

Research) to generate crystal seeds, which when added to fresh crystal drops to act as nucleation sites for crystal growth. Seeding in this way often resulted in larger (and sometimes better diffracting) crystals. Sometimes, fine screens were set up in which the pH and/or concentration of components in the screen or protein concentration could be systematically varied to try to improve crystal morphology.

2.7.3 Harvesting and Cryo-Protection of Crystals

X-ray radiation is very damaging to protein crystals and results in the production of free radicals that can break the ordered crystal contacts and destroy diffraction potential. To protect the crystal, data is collected at cryogenic temperatures to slow down the diffusion of the damaging free radicals and therefore slow the rate of crystal damage. The crystals are therefore cryocooled on harvesting and stored in liquid nitrogen. Generally, the crystals are "fished" out of their droplet using cryoloops, and transferred into a solution of their mother liquor containing a cryoprotecting substance (normally 25% glycerol), unless there is already a cryoprotectant in the condition. This cryoprotectant causes water to form a glass upon cooling (vitrify), rather than freezing into ice, which could disrupt internal order of the crystal. Once the crystal has been transferred to the cyroprotected condition, it can be flash-cooled in liquid nitrogen and stored until collection.

2.7.4 Data Collection

All data in this thesis was collected at the synchrotron at Diamond Light Source, UK. The loop containing the crystal was automatically loaded onto the goniometer, continuously cooled by a stream of gaseous nitrogen. X-rays with wavelengths $\sim$1Å were directed at the crystal, with data collected at three angles (0°, 45° and 90°) initially to assess the diffraction of the crystal. The diffraction pattern data was

first assessed visually by looking at the diffraction spot quality and resolution to which the crystal diffracted. Following this, autoprocessing programmes, such as Xia2 [149], assessed the space group and the resolution to which, the crystal was predicted to diffract. For crystals that had sharp, clear diffraction spots, predicted to go to a resolution of better than ∼4Å, a full dataset was collected. The programmes EDNA [150] and Mosflm [151] suggested data collection strategies that would allow collection of all the unique reflections based on the space group of the crystal. The data were collected by multiple 0.1° oscillations, collecting all reflections over a wedge determined by the strategy.

2.7.5 Data Processing and Refinement

Processing of diffraction data can be split into four major stages: 1) autoindexing; 2) parameter refinement 3) integration; 4) scaling and merging. All of the datasets in this thesis were automatically processed using the Xia2 pipeline incorporated at the beamline [149].

In the Xia2 pipeline, autoindexing was achieved by Mosflm [151, 152] or XDS [153]. Autoindexing finds all of the diffraction spots, indexes the found spots, and calculates the crystal lattice and unit cell dimensions. Following this, there is refinement of the cell parameters and crystal orientation, which is achieved by integrating small segments of data (3-5°). Further, an estimate of mosaic spread in the crystal is generated (caused by crystals being comprised of domains that are imperfectly aligned). Mosflm and XDS then integrate raw pixel intensities of the dataset in order to determine the intensities of the reflections (I) and the background values (σ/I). A resolution cut off is generally determined by looking at the signal:noise ratio: I/σ(I), with a cut off value >1 and a $CC_{1/2}$ value >0.5.

The data measured by the integration programs are not on a common scale. This is

due to variations in diffraction power, changes in the intensity of the X-ray beam, and the length of the X-ray path in different crystal orientations. Therefore, reflections of the same index are compared between frames, and the intensities re-scaled accordingly. This process, known as scaling, is accomplished by the program Scala [154] for data indexed by Mosflm, and by XSCALE [155] for data indexed by XDS. After scaling, the data is merged so that multiple observations of the same reflection are merged into a single, average intensity. Thus scaling and merging may be thought of as an attempt to model the changes occurring during the diffraction experiment. Because of this role of scaling, the most useful statistics on the self-consistency of the data are only obtained after scaling and merging and not directly from integration. The quality of the merged data is given by R_{merge}, the linear merging residual index (Equation 2.1), where N is the number of redundant observations for reflection *hkl*, and $\bar{I}$ is the average intensity of each reflection. However, R_{merge} always results in higher R-factors when the multiplicity is higher. Therefore, often modified versions of R_{merge} are used: R_{meas} and R_{pim} (Equations 2.2 and 2.3). In which, R_{meas} takes into account the multiplicity, and R_{pim} takes into account the multiplicity and the increasing precision as more and more observations are merged.

$$R_{merge} = \frac{\sum_{hkl} \sum_{i=1}^{n} |I_i(hkl) - \bar{I}(hkl)|}{\sum_{hkl} \sum_{i=1}^{n} I_i(hkl)} \tag{2.1}$$

$$R_{meas} = \frac{\sum_{hkl} \sqrt{\frac{n}{n-1}} \sum_{i=1}^{n} |I_i(hkl) - \bar{I}(hkl)|}{\sum_{hkl} \sum_{i=1}^{n} I_i(hkl)} \tag{2.2}$$

$$R_{pim} = \frac{\sum_{hkl} \sqrt{\frac{1}{n-1}} \sum_{i=1}^{n} |I_i(hkl) - \bar{I}(hkl)|}{\sum_{hkl} \sum_{i=1}^{n} I_i(hkl)} \tag{2.3}$$

2.7.6 Solving the Phase Problem Using Molecular Replacement

The output from the scaling and merging is a complete hkl dataset with intensities, but with no phase information. The calculation of the electron density requires the sum of the structure factors and phases of all of the diffraction points in reciprocal space (Equation 2.4).

$$p(xyz) = \frac{1}{V} \sum_{hkl} |F(hkl)| e^{-2\pi i[hx+ky+lz-\phi(hkl)]} \tag{2.4}$$

Therefore, in order to work out the electron density, the phase information must be obtained through other means. The phase data from similar solved structures can be used as an estimate starting point; this is used as the basis of molecular replacement (MR). Molecular replacement was carried out using Phaser-MR. The structures of homologous proteins that shared at least 60% sequence identity were used as starting models with loops removed. Sometimes an all alanine model was created to remove the phase information from the differing side-chains in the model to unknown structure. The space group, and the solvent content, estimated using the Matthews Coefficient [156] were also inputs. Phaser-MR then used maximal likelihood methods in translation and rotational functions (combined with statistical analyses) to work out the likelihood of the data, given the input model, compared to a model of random atom. Multiple solutions were produced, which were examined by their log likelihood gain (LLG) score. The Z-scores were also calculated by comparing LLG values calculated with a random set of rotation and translation functions, judged as the number of standard deviations the solution LLG was above or below the mean value of the random set. The solution with the highest LLG score (above 120) and the highest Z-score (above 8) were taken on for model rebuilding

and refinement.

2.7.7 Model Rebuilding and Refinement

The solution from MR contains extra electron density than is predicted by the molecular replacement model, and vice versa. These density maps can be visualised using Coot [157]. The $2F_{obs}$-F_{calc} map shows a good representation of all of the observed electron density from the data. The F_{obs}-F_{calc} map shows the difference between the electron density from the data and the electron density calculated from the model. Any region that has electron density seen in the model that is not seen in the map (negative) will typically appear in red and should be removed. Whereas any electron density that is not represented by the model but is present in the data (positive) will typically appear in green and can be built into manually in Coot.

Rounds of rebuilding of the model to more closely fit the electron density from the experimental data, were inter-dispersed with rounds of refinement in BUSTER [158]. BUSTER uses maximal likelihood methods to refine the atomic model, minimising the difference between the experimental data and the calculated data from the model. The Buster program outputs the refined model and the electron density maps (observed and calculated in light of the refined model). BUSTER contains commands that allow modelling of missing model structures such as loops, which results in positive output density that can be built into.

The quality of each refinement step was assessed using the R factors, R_{free} and R_{work}. The R factors assess the difference between the observed and calculated structure factors (Equation 2.5). Whereas R_{work} is calculated from the working set of reflections, the R_{free} uses a small subset of reflections (around 5%) that are not

used in refinement.

$$R = \frac{\sum_{hkl} |F_{hkl}^{obs} - F_{hkl}^{calc}|}{\sum_{hkl} F_{hkl}^{obs}} \tag{2.5}$$

Successive refinement and rebuilding should allow the construction of a good order that explains the R-factor. As more of the structure is rebuilt, the difference between the observed and calculated structure factors should become closer together and therefore the R-factors smaller. The final R_{work} should be roughly between 0.20 and 0.30 at a resolution of 3Å to signify that the built structure fits the observed data well. The final R_{free} value should be higher than the R_{work} but at the most 0.05 above the R_{work} to show that data was not over-interpreted.

2.8 Electron Cryo-Microscopy Methods

In electron cryo-microscopy (Cryo-EM) experiments, an electron beam is passed through a protein sample that has been frozen in a near native condition in a thin layer of amorphous ice. The electron microscope operates at high vacuum and below liquid nitrogen temperatures, to keep the ice in an amorphous state and reduce radiation damage. The electrons that are transmitted through the sample are then detected using a direct electron detector (DED). Cryo-EM images of proteins display very high levels of background noise due to the very low electron dose that is needed to avoid radiation damage. Therefore, in order to improve the signal to noise ratio, hundreds of thousands of single protein particle images must be aligned according to the same molecule orientation and averaged. This is a very intensive process and uses a lot of computational power. Recent improvements in image processing methods as well as the ever increasing abilities of computational microprocessors, have made this process less time consuming and easier to handle. These computational improvements

have thus contributed greatly to the "resolution revolution", in which smaller proteins (<50kDa) to higher resolutions (< 3) are being solved by cryo-EM. Another main reason for the "resolution revolution" in cryo-EM has been due to the advances made in electron detectors. The DEDs have a very thin sensing layer that reduces lateral charge spread of the electrons detected and so results in higher resolution than the previously used charge coupled device (CCD) sensors. Moreover, the electrons are detected directly (as the name suggests), which means image distortions are minimised without having to transfer the signal through fiber-optic or optical lenses. Another advantage of DEDs over CCD sensors is that they have a higher frame rate, with no dead time between frames. This allows intrinsic dose fractionation during image acquisition, which can be used for motion correction, damage compensation, and other image processing techniques.

2.8.1 Grid Preparation

2.8.1.1 Glow Discharge

Protein samples in solution were (usually) applied onto holey or continuous carbon films that cover a special EM grid. Since carbon is generally a hydrophobic substance, they were treated by glow discharge to make the grids more hydrophilic, allowing the aqueous sample to be applied and spread evenly across the grid. When glow discharging grids, a high voltage was applied across the partially evacuated chamber. This electron potential ionised the residual gas within the chamber, creating negatively charged ions that are then deposited on the carbon surface of the grid. These negatively charged ions made the carbon surface of the grid more hydrophilic. The glow discharge machine used was the Q150T Plus Coater (Quorum) at 15 mA for 60 seconds.

2.8.1.2 Grid Vitrification

In the process of vitrification, the sample is cooled so rapidly that the water does not have time to crystallise, instead it forms an amorpheus solid which does little/no damage to the sample structure [159]. This is usually achieved by plunging the sample-coated grid very quickly into liquid ethane so that it is to cooled to -180°C in $< 10^{-4}$ seconds.

A Vitrobot^{TM} (ThermoFischer Scientific) was used in the vitrification of grids. Generally, grids were set up at 22°C and 100% humidity. The grids were loaded into the vitrobot chamber using Vitrobot Tweezers (FEI,Ted Pella 47000-500)and 3 μL of sample was applied to the carbon coated side of the grid. The general freezing settings included a 30 second wait time after the sample was applied to the grid, a 3 second blot time to remove excess solution and a 2 second drain time before being plunged into the liquid ethane. The blot force was varied in order to optimise ice thickness and spreading. Once the grid had been plunged into the liquid ethane (a cryogen), the grids were carefully released from the vitrobot tweezers and transferred to a grid box for storage, all under liquid nitrogen.

2.8.1.3 Preferred Orientation

When samples are vitrified on a holey cryoEM grid, there are often issues with the sample accumulating at the air-water interface. The hydrophobic properties of the surface of the protein determines how the protein orients itself at the air-water interface with hydrophobic patches being repelled from the solution. Therefore, when proteins accumulate at the air-water interface, this often leads to the proteins being in very similar orientations (preferred orientations). As a result, rather than the different views of the protein tumbling in solution being captured when the sample is vitrified and seen, one or two views are seen and the other views are not.

2.8.1.3.1 Making Graphene Oxide Grids One way to overcome this problem of preferred orientation is by having a continuous very thin carbon support film (<5μm) across the holes. The sample often has an affinity for the carbon film, thus drawing the particles away from the air-water interface. Grids with continuous support films can either be made in house or bought. In this thesis, both bought Quantifoil® 2nm carbon grids, and homemade graphene oxide grids were used.

Graphene-based support layers have the advantage over the continuous carbon that unlike carbon, graphene is almost electron transparent and so contrast should not be reduced as much. Graphene itself is extremely hydrophobic but can be made hydrophilic by chemical modification or by hydrogen plasma treatment. An alternative is to use the naturally hydrophilic graphene oxide. Graphene oxide grids were made according to the protocol in [160], with some changes. In short, the Graphene Oxide Solution (Graphenea) was diluted 20x in MQ H_2O to 0.2 mg/ml and spun at 1600rpm for 2 minutes to remove any aggregates. Meanwhile, the grids (usually Quantifoil® Cu 200 1.2/1.3) were glow discharged using a (Quorum Q150T Plus Coater) at 50mA for 60 seconds. Each grid was held in anti-capillary tweezers (Dumont N5AC) and 3 μl of GO was applied to the grid for 1 minute and then blotted on Whatman Filter Paper. The grid was then moved to a 20μL droplet of ddH_2O and picked this droplet up before being blotted, this was repeated one more time on the carbon side of the grid before the third droplet of ddH_2O was picked up using the back side of the grids. The grids were dried for 5 minutes before being transferred to a gridbox for storage. The grids were vitrified within 24 hours of preparation, without further need to glow discharge.

2.8.1.4 Transfer of Cryo-EM grids to the Microscope

2.8.1.4.1 Grid Clipping In order to autoload multiple grids onto the Talos Arctica 200kV and the Titan Krios 300kV (ThermoFisher Scientific), they had to be

clipped into autogrid cartridges. Tweezers were used to load C clips into the C-clip insertion tool (ThermoFisher Scientific) and the C-clip position at the rim of the tool by gently pressing the insertion tool against a flat surface. The autogrid assembly station (ThermoFisher Scientific) was cooled using liquid N_2. Once cooled, the C-clip insertion tool loaded with C-clips, the autogrid tweezers were placed in the assembly station to cool them down to liquid N_2 temperature. The C-clip ring was transferred to the loading position on the assembly station, with the flat side positioned against the base of the station. The grid was transferred from the gridbox onto the top of the C-clip ring using the cooled autogrid tweezers, under liquid N_2. The station was moved to the "closed" position and the C-clip insertion tool inserted vertically over the top of the grid. The button on the top of the tool was pressed to release the C-clip onto the C-clip ring. The autogrid was inspected to ensure that the grid was securely clipped. The autogrids were transferred to the autogrid container, with the C-clip facing towards the notch on the grid box.

2.8.1.4.2 Grid Loading The loading station and NanoCab (ThermoFisher Scientific) were both cooled in liquid N_2. The autogrids were transferred to the station and positioned in the grid cassette using the autogrid tweezers with the C-clip facing the top of the cassette. The pre-cooled nanocab was docked against the loading station an the cartridge arm used to transfer the cassette quickly into the NanoCab. The NanoCab could then be used to dock the cassette onto the microscope.

2.8.2 Grid Screening

Grid screening was generally carried out on the Talos Arctica operating at 200kV and equipped with a Falcon 3 detector (FEI). At low magnification the grid was assessed for "good" ice distribution and thickness across the gridsquare (Atlas). Good gridsquares were noted for future potential data collection. Once a gridsquare

had been chosen, then holes within that gridsquare were then looked at, at a higher magnification using the FEI-EPU, to assess the distribution and clarity of the particles within the holes.

2.8.3 Data Acquisition

Once a grid had been screened and observed to contain a good distribution of particles within the holes as well as even and thin ice across the grid, data could be collected. Data collection was normally carried out on the Titan Krios operating at 300kV, connected to a Falcon 3 detector (FEI). Movies were collected in counting mode with a sampling of 0.86Å/pixel, 12 s exposure, and total dose of 50 electrons.

2.8.4 Data Processing

The first part of processing was to correct for blurring of the images caused by electron-beam induced motion [161]. This was carried out using Simple-unblur [162].

In cryo-EM specimens, the contrast in the micrographs between the background and the protein particles is very low, since electrons interact weakly with the atoms in the protein. However, by defocussing the microscope by 1-3μm, phase contrast can be generated in the micrograph images. The contrast transfer function (CTF) is an oscillating function that describes how contrast generated by the microscope varies as a function of resolution in the micrographs. As resolution increases, reversals in phase contrast occur more often, and the overall signal decreases. Therefore, to computationally correct for these effects you must CTF-correct your images, which was carried out using CTFFIND4 [163].

On the motion-corrected and CTF-corrected micrographs, particles were picked

using EMAN2 [164]. The boxes could be picked either completely manually or semi-automatically using the input of manually picked particles as templates. The boxes were extracted using a box 240Å x 240Å and subject to multiple rounds of reference-free 2D classification in Relion2.0 resulting in the production of multiple 2D classes [165].

The way in which the 2D classes were used to make a 3D model is based on "Projection Theorem" [166, 167], which states that the "fourier transform of a 2D projection equals the central section through its 3D Fourier transform perpendicular to the direction of the projection". An initial 3D model was made using the angular reconstruction method, which allowed the angular distribution of the 2D projections to be related to the 3D object. In short, this method is based on the fact that two different 2D projections of a 3D object always have a common one dimensional (1D) line in their projections, where amplitude and phase are similar in two images. If this common line can be found, a 3D object can be reconstructed from 2D projections by aligning the 2D projections along the common lines. The common line is found by producing a graph called a sinogram, which plots the phase and amplitude of the fourier transform of 2D classes. The phase and amplitude that match the best between the two classes can then be used to identify the common line and relative orientation of the 2D projections. Once multiple relative orientations of 2D reprojections have been established, the angular restrictions were used to place the various class averages FTs in a 3D FT space. Once all 2D FTs (2D classes) were inserted in a 3D space, inverse FT of this 3D FT will produce a 3D structure. This process was used to create an initial model, and was carried out in Relion [165].

The initial model was then used as a starting point for 3D classification. During classification, different 3D classes can be produced that represent sample heterogeneity, for example caused by different protein conformations and/or different compositions of complexes. The particles that made up the predominant 3D class, were then used

for movie processing in the 3D refinement step. The objective of this refinement step is to determine the position of all of the particles in relation to the 3D reference molecule. This is an iterative process that aligned individual images (N.B. not class averages) on the projections created from the initial model. (These individual images, although similar to the overall 2D class were not identical). After many iterative rounds, the refinement ends with a higher resolution structure than the initial model with more stringent angular constrain. The 3D reconstruction is iterated until resolution stopped to improve [168].

2.9 Immunogen Design

2.9.1 Protein Repair One Stop Shop (PROSS)

PROSS is an algorithm that has been developed to improve the expression and stability of proteins, whilst maintaining their function [169]. The general protocol is described in detail in **Section 4.4.4.3.1**. Three changes were made to the general protocol:

1. **Sequence Alignment Used (Section 4.4.4.3.2)**

   An automatic sequence alignment generated by the PROSS program searching BLASTP included many sequences that had <31% sequence identity, which would likely not retain the functional epitope on Subdomain 3.

2. **Functional Residues were "Fixed" (Section 4.4.4.3.3)**

   The functional residues of the DB9 epitope on subdomain 3 were fixed and therefore not allowed to be mutated by the program. Theses residues were: 86, 87, 90, 93, 98, 100, 101, 102, 103, 106, 109, 111, 113, 116, 117.

3. **Introducing Mutations at "Interface" Region (Section 4.4.4.3.5)**

   The hydrophobic residues that made contact between subdomain 2 and subdo-

main 3 were highly conserved and not mutated by PROSS. Therefore, these now solvent exposed residues had to be mutation to make them more hydrophilic.

Chapter 3

Targeting Cytoadherence in Severe Malaria by Blocking EPCR binding *Pf*-EMP1s

3.1 Introduction to *Plasmodium falciparum* Erythrocyte Membrane Proteins

During the erythrocytic cycle, the parasite within the RBC modulates the erythrocyte membrane by extending the parasitophorous membrane into a network to the RBC surface [170]. Parasitic proteins are shuttled to the erythocytic cell surface membrane such as *P. falciparum* histidine-rich protein (*Pf*HRP), which is involved in the formation of knobs (symmetric membrane arrangements) on the iRBC surface [171]. More proteins, like the erythrocytic membrane proteins (*Pf*EMP1) occupy these knobs during the intra-erythrocyte ring and trophozoite stages [172]. *Pf*EMP1 proteins play major roles in the sequestration of the iRBC by cytoadherence to the vascular endothelium through interactions with various endothelial receptors [70, 173, 174]. The sequestration of the iRBCs that results from these interactions not only allows the iRBC to avoid removal through splenic clearance [175] but also contributes to many of the pathologies of severe disease [176]. Specifically, the

interaction between *Pf*EMP1s expressed on iRBCs and brain endothelial receptors has been strongly implicated in the onset of the very fatal form of cerebral malaria [70].

3.1.1 Antigenic Variation

*Pf*EMP1 proteins are encoded by a large family of genes (called the *var* genes) located at the telomeres and centromeres of the chromosomes. Individual parasites have approximately 60 *var* gene variants, and switching between single transcribed *var* genes leads to changes in cytoadhesive phenotype, as well as contributing greatly to clonal antigenic variation and immune escape. The *var* genes can be separated into three groups: A, B, and C based on location within the chromosome, sequence and promoter sequence [177]. Group A *Pf*EMP1s have been related to severe malaria and appear to be less diverse [178]. Another layer of variation is added due to the fast or slow coordinated switching of gene variants expressed between one generation and the next [179]. This switching of *var* gene expression between parental and daughter cells is a very complex, highly evolved process which also allows the parasite to evade adaptive immune system. In addition, there is evidence that the *var* genes can undergo mitotic recombination, further increasing diversity [180].

3.1.2 Architecture

Despite the great sequence diversity between *Pf*EMP1 proteins, they all share a conserved architecture. *Pf*EMP1s are anchored to the infected erythrocyte membrane by a single transmembrane helix. They have a small cytosolic domain and a large extracellular multidomain region required for cytoadherence **(Figure 3.1)** [174]. The small cytosolic domain, known as the acidic terminal segment (ATS) due to its net negative charge [181], is highly conserved in sequence and structure [182–184].

The ATS is largely flexible with a small structured core that provides minimal conformational restraint for a central flexible segment. The disordered nature of this region is believed to create large intermolecular interfaces that provide highly specific but reversible interactions [185], both with parasite proteins released into the cytosol (such as PFI1780w [183]) and components of the cytoskeleton of the host cell (like spectrin [186]). Thereby, these cytosolic interactions contribute to the localisation of *Pf*EMP1s in the knob structures.

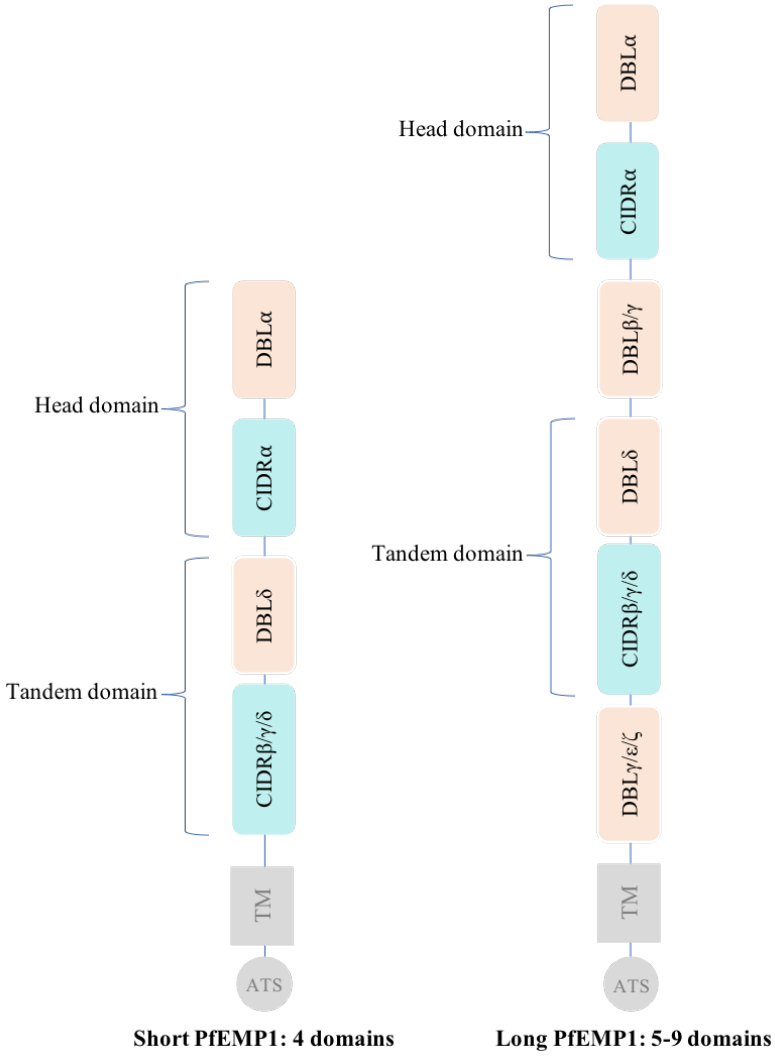


Figure 3.1: Architecture of *Pf*EMP1 Proteins The overall organisations of DBL and CIDR domains in short and long *Pf*EMP1s, with the single transmembrane helix labelled as TM, and the cytosolic domain labelled as ATS. (Based on figure in [187]).

The extracellular region of *Pf*EMP1s consists of two main types of domain: CIDR

(cysteine-rich interdomain region) domains and DBL (Duffy binding like) domains. The DBL domains show homology to the *P. falciparum* EBA-175 (erythrocyte binding antigen) as well as *P. vivax* and *P. knowlesi* Duffy binding proteins, all involved in erythrocyte invasion [188, 189]. The CIDR domains were first discovered in *Pf*EMP1 proteins [181] and the C-terminal half of the domain shares some similarity in fold, particularly in the placement of the cysteines, with the DBL domains [190].

The size of *Pf*EMP1 proteins varies between ∼160kDa and 300kDa, and the number and organisation of the CIDR and DBL domains within a *Pf*EMP1 protein varies [191]. According to sequence, the CIDR domains can be classified into three different types (α, β, γ) and DBL domains into five different types $(\alpha, \beta, \gamma, \delta, \epsilon)$ [192]. The most common ectodomain architecture of the *Pf*EMP1 consists of a head structure of a DBLα-CIDRα tandem, followed by another DBLδ-CIDRβ/γ tandem [181, 193]. In larger *Pf*EMP1s, additional DBLβ/γ domains are inserted between the head and the tandem domains, and/or between the tandem domain and the transmembrane domain **(Figure 3.1)**.

Furthermore, adding another layer of complexity to *Pf*EMP1s, different arrangements of DBL and CIDR domains belonging to particular subclasses are often found in certain combinations, called domain cassettes (DCs) [194]. Some DCs have been associated with the development of severe disease, due to the conserved presence of ligand binding domains, for example DC8 and DC13 **(Figure 3.2)** [195].

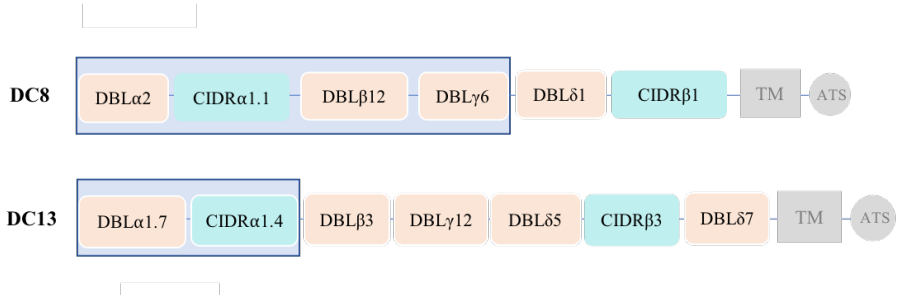


Figure 3.2: Domain Cassettes Associated with Severe Malaria
Two examples of DCs associated with severe malaria. The DCs for DC8 and DC13 are highlighted by the blue box

3.1.3 Endothelial Binding Partners

*Pf*EMP1 proteins have multiple binding partners on the endothelium, including ICAM-1 [196], CD36 [197], EPCR [198], as well as the carbohydrate chondroitin sulphate A (CSA) [199]. Most of these endothelial receptors bind to single CIDR or DBL domains within the *Pf*EMP1, however the binding of CSA requires multiple domains of the *Pf*EMP, Var2CSA **(Figure 3.3)**. Some *Pf*EMP1s with different combinations of DBL and CIDR domains can bind to a variety of these ligands, allowing one *Pf*EMP1 to contact multiple endothelial ligands at once. For example, a single *Pf*EMP1 has been shown to bind to both EPCR and ICAM-1 [200]. It should be noted that **Figure 3.3 A** provides a simplistic view and so for example, although DBLβ domain has been linked to ICAM-1 binding [201], not all DBLβ-containing *Pf*EMP1s will bind to ICAM-1 [202]. The overall organisation within the *Pf*EMP1, and from which *var* gene group the *Pf*EMP1 is expressed, add another layer of complexity in predicting their binding to endothelial receptors.

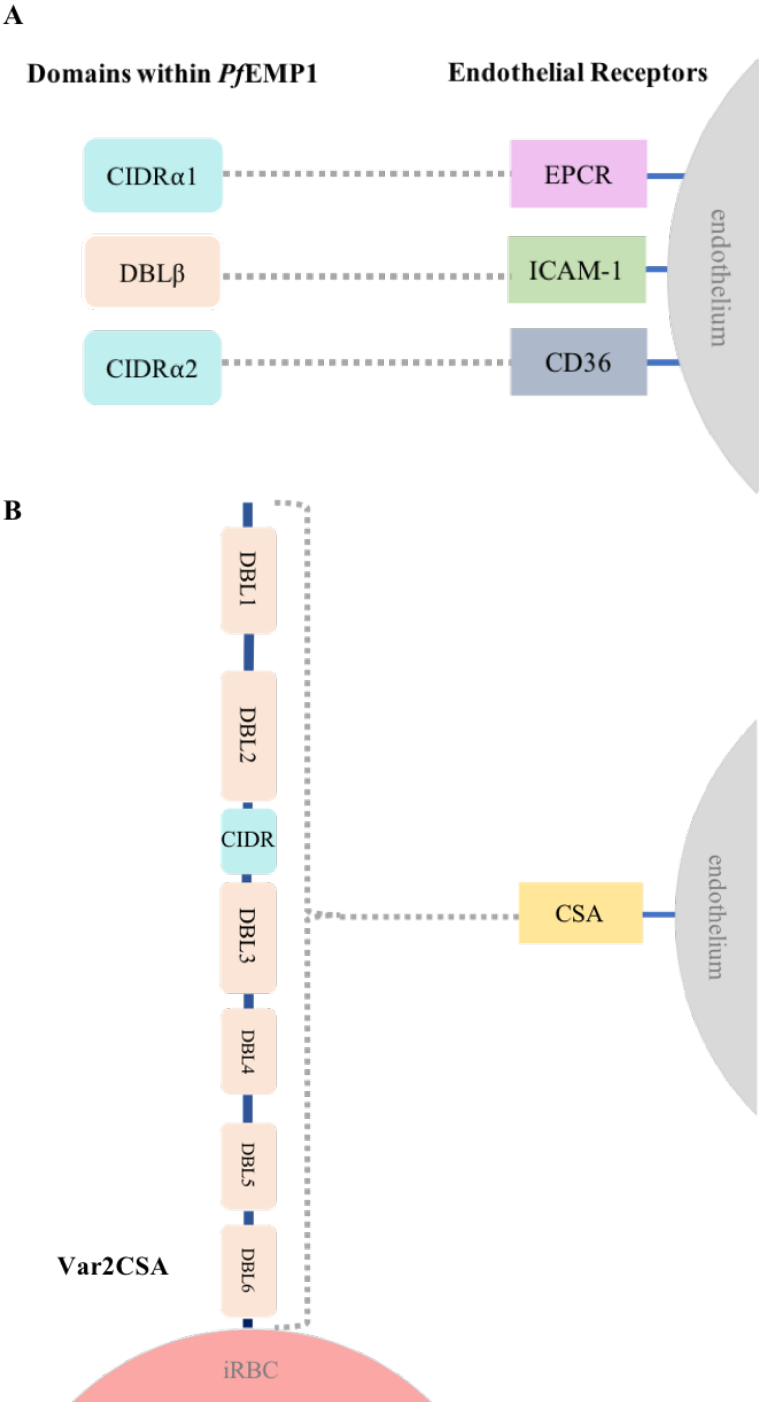


Figure 3.3: *Pf*EMP1 Endothelial Binding Partners A) Individual domains of *Pf*EMP1 proteins have been shown to be the minimal binding units to a variety of endothelial receptors. B) For the binding of CSA to the unusual *Pf*EMP1, Var2CSA, the whole ectodomain is required for binding.

3.1.4 The Interaction of *Pf*EMP1 with EPCR

Prior to the discovery of EPCR as a *Pf*EMP1 binding partner, the expression of two DCs, DC8 and DC13, were linked to the development of severe malaria [195]. The interaction between *Pf*EMP1 DC8 and DC13, and EPCR was then identified by the screening of human plasma proteins against the *Pf*EMP1 IT4var20. This identified EPCR as the sole receptor for this *Pf*EMP1 [198]. The link between EPCR and severe malaria was then further corroborated through studies which showed that gene variants of EPCR correlated with protection against the severe disease [203]. In addition, abnormal expression of EPCR in the brain endothelium was shown to be associated with severe cerebral disease [204].

The minimal binding region for this interaction was mapped to the CIDRα1 domains within the head region of the *Pf*EMP1, which binds with similar affinity to EPCR compared to the whole *Pf*EMP1 [198]. Sequence analysis of over 800 CIDRα1 domains allowed this class of domain to be further divided into different subclasses (1.1-1.8) and nearly all CIDRα1s bind to EPCR (all except 1.2 and 1.3) **(Figure 3.4)** [205]. This sequence analysis also revealed that despite being able to bind EPCR, the EPCR-binding CIDRα1s share very little sequence identity, with only 6% of residues absolutely conserved.

3.1.4.1 The Natural Role of EPCR

EPCR is an ubiquitous endothelial receptor that belongs to the major histocompatibility complex (MHC) class superfamily. However, in its central cavity it has a lipid bound rather than a peptide as seen in other MHC class proteins [206, 207]. The natural role of EPCR is to bind protein C (PC) in the bloodstream, which is generated by the thrombin-thrombomodulin complex responding to tissue damage [208]. Once protein C has bound to EPCR, it becomes activated (activated protein

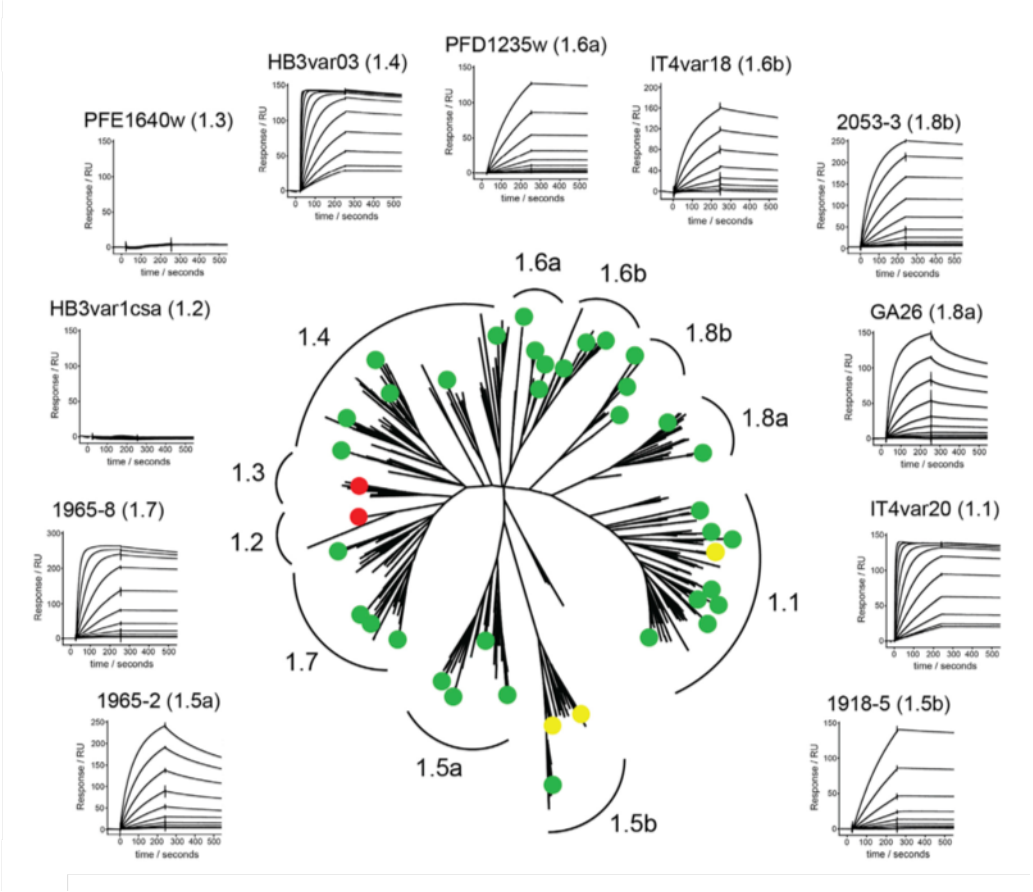


Figure 3.4: Phylogenetic Tree Showing the Different Subclasses of CIDRα1 Domains The circles represent the degree of EPCR binding by ELISA with positive (green), weakly positive (yellow), and negative (red). SPR traces of the different CIDRα domains of each subclass (in brackets) showing ability to bind to EPCR [205].

C, APC) and can bind to protease activated receptor 1 (PAR-1) *in cis* (i.e. on the same cell) [209]. PAR-1 transduces the signal into the cell via β-arrestin which recruits dishevelled-2 (Dvl-2) [210]. Dvl-2 is responsible for binding to downstream effectors that mediate anti-inflammatory and cyto-protective effects [211] **(Figure 3.5 A)**. Further binding [198] and structural studies [205] highlight that the binding of CIDRα1 domains from *Pf*EMP1s to EPCR prevents the binding of protein C, since these two proteins share the same binding surface. It has been hypothesised that it is advantageous for the malaria parasite to use an important existing binding site, making it difficult for the host to mutate EPCR to prevent binding to parasite protein, without adversely affecting interactions with protein C.

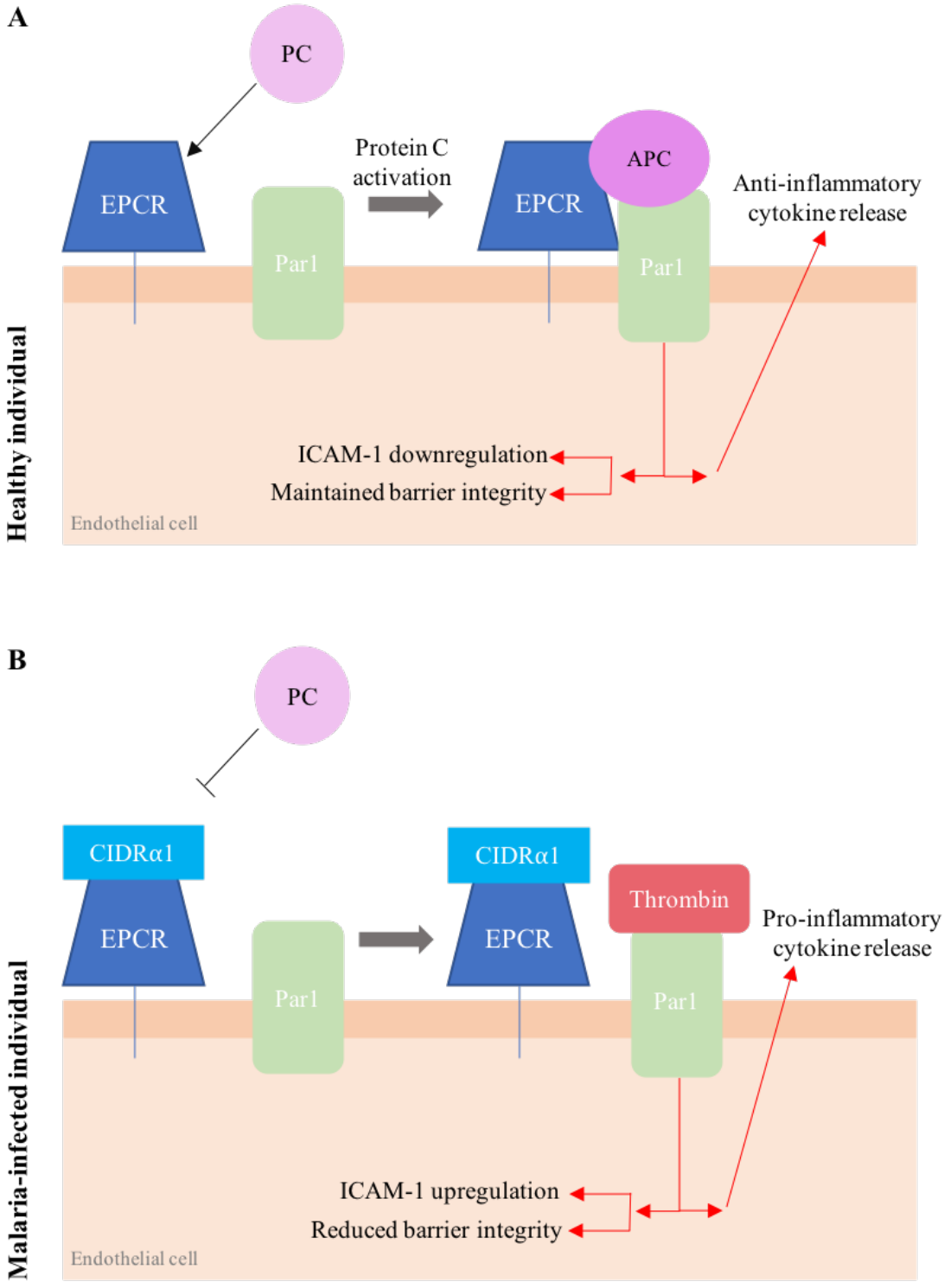


Figure 3.5: Proposed Model of EPCR-binding *Pf*EMP1 proteins Modulating the Immune Response The PAR-1 signal pathway in the absence (top panel) and presence (bottom panel) of the EPCR-binding *Pf*EMP1s.

This has led to the proposal of an immuno-modulatory function of *Pf*EMP1 binding. In this model, in infected individuals, protein C is blocked by CIDRα1 from binding to EPCR. Therefore, PAR-1 (protease activated receptor 1) is free to bind to thrombin, and then transmits intracellular signals that lead to the production of pro-inflammatory cytokines and increased permeability of the endothelial barrier. In this way CIDRα1 binding blocks activated protein C-mediated cytoprotective signalling, leading to pro-inflammatory, pro-apoptotic responses and loss of the endothelial barrier **(Figure 3.5 B)**. In addition, when protein C binding to EPCR is blocked by CIDRα, this causes increased expression of ICAM1, another *Pf*EMP1-endothelial receptor, which may lead to enhanced sequestration of infected erythrocytes [198].

3.1.4.2 Structural Analysis of the CIDRα1:EPCR Interaction

Two structures of the CIDRα1.4:EPCR complex revealed a binding patch, on a helix-kink-helix motif in the CIDRα1.4 domain. Seven out of the nine directly interacting residues are located on this helix-kink-helix motif. Moreover, a phenylalanine (F656) at the centre of this patch is required for high-affinity binding **(Figure 3.6 A and B)** [205]. Indeed, mutation of this F656 to alanine results in a 100-fold increase in the dissociation-rate of the complex. However, even within this EPCR-binding region of the CIDRα1 domains, the sequence is not highly conserved **(Figure 3.6 C)** [205]. Structural studies and sequence analysis indicate that although the residues themselves are not conserved, their chemical characteristics are, with a hydrophilic surface surrounding the hydrophobic protusion at the kink of the helix-kink-helix, enabling these sequence diverse CIDRα1s to bind EPCR.

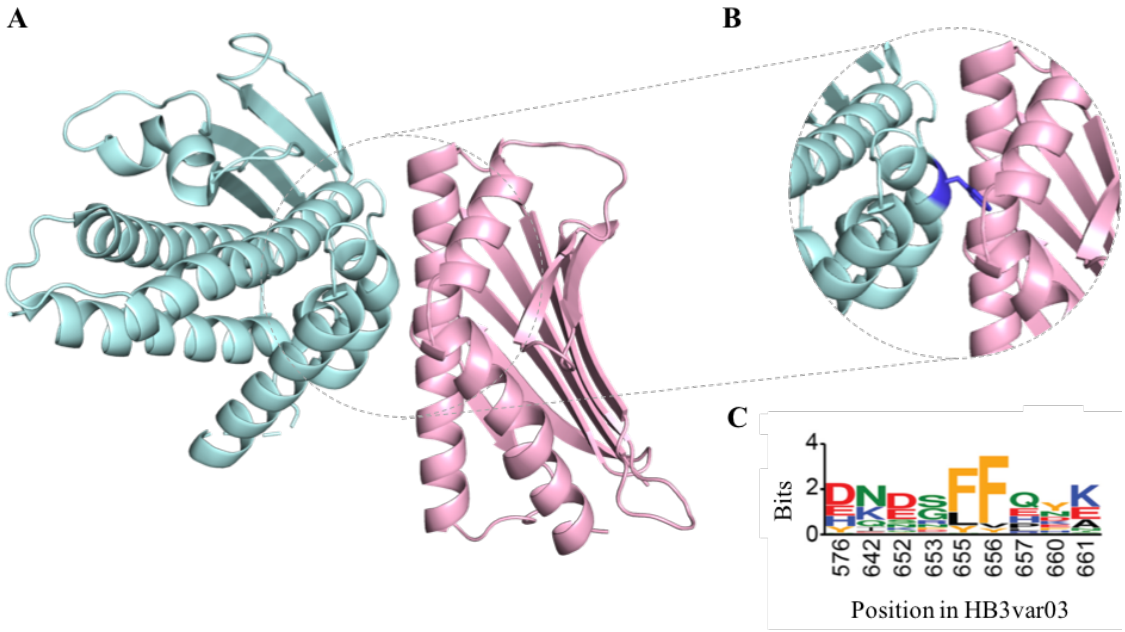


Figure 3.6: CIDRα1:EPCR Structure A) The structure of CIDRα1.4 (HB3var03) (cyan) in complex with EPCR (pale pink) (PDB:4V3D). B) A close-up of the binding site, showing the phenylalanine (F656) (blue) crucial for the interaction. C) Sequence logo from all EPCR-binding CIDRα1s showing the variation in the residues that directly interact with EPCR. The sequence logo was made using WebLogo [212].

3.1.4.3 *Pf*EMP1s as Vaccine Candidates

Due to the huge diversity in *Pf*EMP1s that can be expressed in one parasite, as well as the ability of the *var* genes to undergo mitotic recombination, the pool of *Pf*EMP1s that need to be targeted is immense [180, 213]. As a result, generating immunity to malaria via *Pf*EMP1s not only depends on the acquisition of antibodies that target a crucial interaction, but also those that are protective against a range of *Pf*EMP1 strains (strain-transcending). Despite seeming like an impossible challenge, as children grow up in malarial endemic areas they gain naturally acquired immunity with antibodies that learn to recognise *Pf*EMP1 antigens expressed on their iRBCs [214]. It is thought that although the sequence diversity of *Pf*EMP1s is great, their function of binding to endothelial receptors constrains their antigenic diversity in these regions. Therefore, the parasite is in a trade-off between antigenic novelty for

immune evasion and conservation to ensure effective cytoadhesion [214–216].

The discovery that the interaction between EPCR and the CIDRα1 domain of *Pf*EMP1s is crucial for cerebral malaria, suggests that the CIDRα1 domains show a good potential as a vaccine candidate [198, 203, 204]. Moreover, naturally acquired human antibodies that bind to CIDRα1.1 (IT4var20) and CIDRα1.4 (HB3var03) can block EPCR-binding by recombinant CIDRα1 domains, as well as block endothelial cell binding by parasite-infected erythrocytes [205]. Moreover, these inhibitory antibodies can individually reduce the binding of the EPCR to both CIDRα1.1 and CIDRα1.4. These provisional experiments suggest that although different subclasses of CIDRα1 domains are very sequence diverse, the architecture of the helix-kink-helix EPCR-binding site is conserved, allowing mAbs to inhibit EPCR binding to different CIDRα1 domains.

3.1.5 Design of an Immunogen to Block *Pf*EMP1:EPCR Interaction

In recent years there have been multiple advances in structure-guided immunogen design. Most examples start with a structure of an inhibitory antibody in complex with the pathogen surface protein, which is followed by grafting of these antibody epitopes onto smaller scaffolds [138, 217]. In some cases, these design strategies have led to the production of smaller immunogens that specifically elicit the production of inhibitory antibodies [135, 137, 218]. Since there is currently no structural information about the epitopes of antibodies that inhibit CIDRα1 binding to EPCR, the usual immunogen design method described above could not be used. However, having identified the conserved chemistry and shape of the EPCR-binding site, an immunogen was designed based on this helix-kink-helix region. It was hoped that by grafting the EPCR-binding site onto a smaller protein, a molecule could be designed

which could be used as a tool to assess whether inhibitory antibodies targeted the EPCR binding site directly. Moreover, it was hoped that if used as an immunogen, the immune response could be focused to this region of the CIDRα1 and therefore a greater proportion of inhibitory antibodies could be produced than when immunising with the whole CIDRα1 domain. This conserved binding site might also be able to produce cross inhibitory antibodies that block multiple subsets of CIDRα1 binding to EPCR.

The immunogen design process was based on the structure of the CIDRα1 domains bound to EPCR. These structures showed that there were nine residues of the CIDRα that directly interacted with EPCR, and seven of these residues were found on the helix-kink-helix **(Figure 3.7)**.

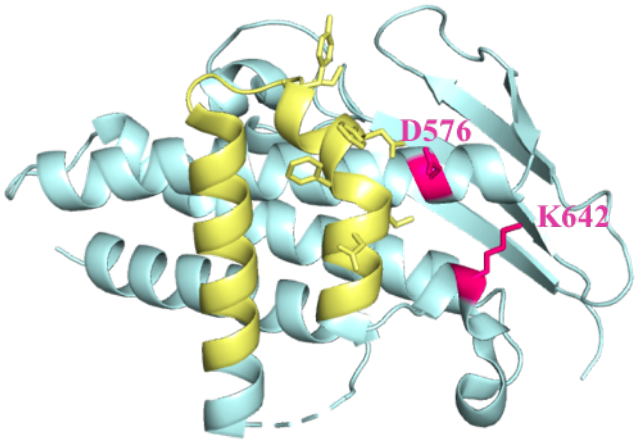


Figure 3.7: CIDRα1.4 EPCR-binding Residues The side chains of the EPCR-interacting residues are shown as sticks. The helix-kink-helix motif is coloured in yellow, and the two interacting residues that lie outside of this region are coloured in pink and labelled.

In order to confirm that the absence of the two binding residues found outside the helix-kink-helix region did not ablate EPCR-binding, a CIDRα1 double mutant (D576A and K642A) was produced. The affinity for EPCR was assessed by SPR and although there was a 500 hundred-fold reduction in affinity, from 0.4nM to 220nM, it still bound with high nanomolar affinity similar to that of other CIDRα1s **(Table**

3.1) [205]. As a result, the helix-kink-helix was grafted onto a helical scaffold in order to produce a protein that contained the EPCR-binding site of the CIDRα1, but not other non-functional regions of the CIDRα1.

	$\mathbf{k}_{on}$ $(\mathbf{M}^{-1}s^{-1})$	$\mathbf{k}_{off}$ $(\mathbf{s}^{-1})$	$\mathbf{K}_d$ **(nM)**
CIDRα1.4	53000	2.00x10^{-4}	0.37
CIDRα1.4 D576A K642A	9100	2.00x10^{-3}	220

Table 3.1: Comparison of EPCR-binding by CIDRα1.4 and CIDRα1.4 Double Mutant

The method used to achieve this grafting was Fold-from-loops [137]. The first step was to find a scaffold onto which the EPCR-binding site could be grafted. A three-helical scaffold (PDB code: 3LHP, chain S) that had previously been used for immunogen design [218] was chosen and deemed suitable for grafting because the two longer helices of the helix-kink-helix motif could be matched to the path of the scaffold within 2.5Å.

The initial design contained the sequence from Asn648 to Lys678 of the CIDRα1.4 (HB3var03) domain inserted in between the Gly70 and Asp103 residues of the three-helical bundle **(Figure 3.8)**. This hybrid sequence was compiled and using a Rosetta program computationally allowed to refold around the binding site (Asn72-Trp93), which was held in a fixed position. C_α-C_α restraints for the helical bundle portion were also applied to drive the refolding towards the helical correct solution. The models produced were then allowed to "relax" by removing constraints to a local energy minima. The resulting models were scored according Rosetta score (a measure of the intramolecular interaction strength) and root square mean deviation (RMSD) to the starting models.

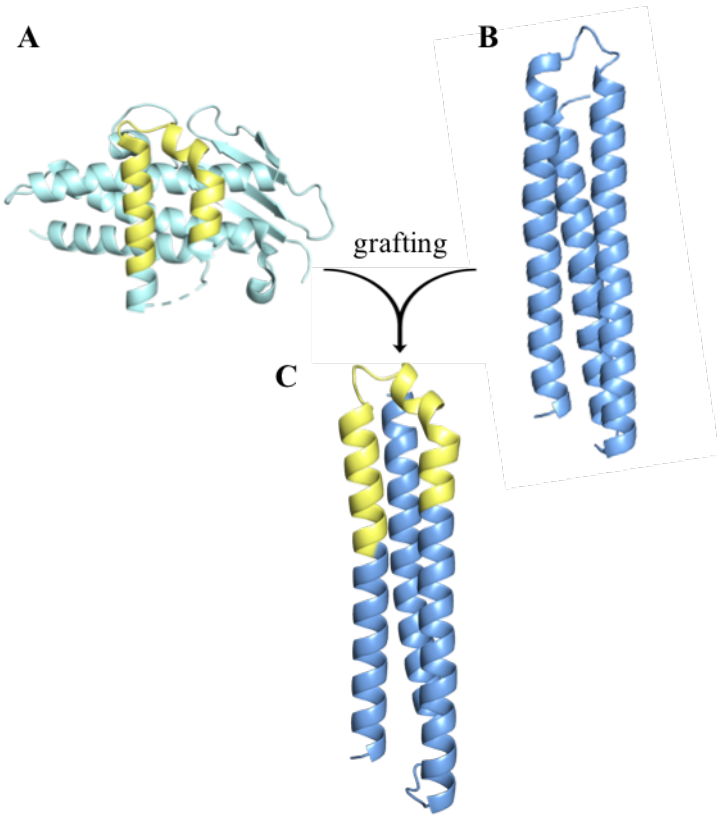


Figure 3.8: Grafting of Epitope on Helical Scaffold The grafting of the EPCR-binding site from CIDRα1.4 (A) onto a three-helical scaffold (B) to produce a three-helical bundle that contains the desired binding site (C).

The model that had the lowest RMSD was chosen for further design in the Fold-from-loops method. All residues in the model, except the seven directly involved in binding and three residues thought to be crucial for the maintenance of the kinked-helix, were allowed to mutate. Multiple rounds of sequence design involved mutation followed by structural relaxation. The output models were ranked by the Rosetta score and the RMSD to the helix-kinked-helix and the models with the lowest RMSD were taken forward **(Figure 3.9 B)**. All of the output models had relatively large RMSD scores (>3Å) compared to the binding epitope. Looking at the overlay of the predicted structure of the output models with the CIDRα1 structure, even the best scoring outputs of the round 1 models deviated from the kinked-helix of the original structure **(Figure 3.9 A)**. In particular, the loop in the epitope, between the kinked helix

and the following helix did not pack properly, affecting the conformation of the kinked helix side chains. Therefore, in order to improve the stability, an extension was added to the N-terminal helix of the scaffold to provide extra support for the EPCR-binding site. The use of the longer scaffold generated round 2 models, whose epitopes were much closer to that of the CIDRα1.4 EPCR-binding site **(Figure 3.9 A)**.

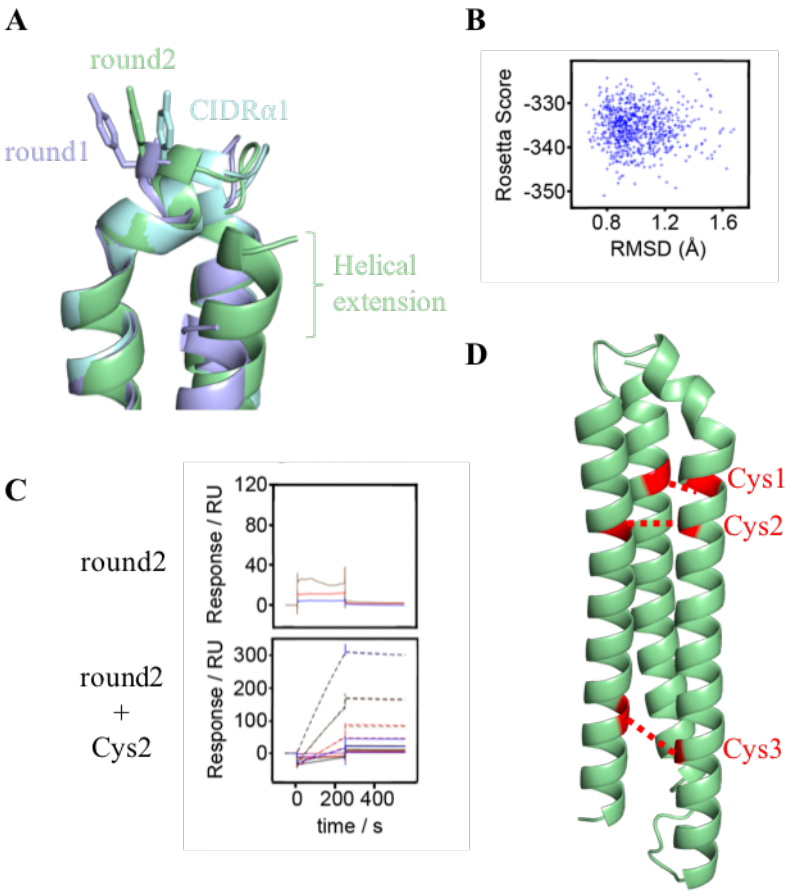


Figure 3.9: Rounds of Bundle Design A) An overlay of the helix-kinked-helix region of the round1 model (lilac), round 2 model (pale green) and the original CIDRα1.4 (pale blue). The helical extension to the scaffold added between round1 and round 2 of the design is also highlighted. B) An example graph showing the output models scored by Rosetta score and RMSD to the kinked helix motif. C) Preliminary SPR curves to show the binding of the round 2 model, and the round 2 model plus Cys2 to EPCR. D) The positions of the pairs of cysteine residues introduced to stabilise the fold of the round 2 design by the addition of disulphide bonds.

Upon small scale expression and purification of the best round 2 models, it was discovered that none of these models bound to EPCR **(Figure 3.9 C)**. To gain clues why these construct did not bind, the AbInitioRelax program in Rosetta was used to predict the conformation of the Fold-from-loop output models. These results highlighted that the constructs had not folded into the conformation output by the Fold-from-loops protocol. Therefore, in an attempt to stabilise the fold of the kinked-helix and to reinstate the EPCR-binding, three pairs of residues were mutated to cysteine both in combination and separately in order to introduce stabilising disulphide bonds **(Figure 3.9 D)**. Initial trials on these cysteine mutants showed that introducing the Cys2 pair of mutations reinstated the ability to bind to EPCR in initial SPR trials **(Figure 3.9 C)**. Henceforth, this construct will be referred to as Bundle1.4.

3.2 Production, Structural and Biophysical Characterisation of Immunogens against the CIDRα1:EPCR interaction

3.2.1 Project Aims

1. Characterise the fold and the function of the design of immunogen, via CD, SPR, and X-ray crystallography
2. Design a panel of immunogens to represent the sequence diversity of the EPCR-binding CIDRα1s, and characterise their binding and immunogenicity
3. Obtain structural information of inhibitory mouse antibodies bound to EPCR to gain insight into the position of inhibitory immunogenic epitopes and use for further immunogen design

3.2.2 Declaration

A former DPhil student, Clinton Lau, designed the original Bundle1.4 construct, using the Fold-from-loops design process outlined in **Section 3.1.5**. Immunisations and subsequent analysis were carried out by collaborators, Thomas Lavstsen, University of Copenhagen (outlined in **Section 3.2.10**).

3.2.3 Designing Panel of Immunogen Bundles to Represent Sequence Diversity

From previous studies exploring the CIDRα1:EPCR interactions, it has been shown that there are six EPCR-binding subclasses of CIDRα1. Namely, CIDRα1.1 and

CIDRα1.4-1.8 **(Figure 3.4)**. Although it was hoped that immunisations with one class of immunogen (i.e. the previously designed Bundle 1.4) would be able to raise antibodies that block all classes of EPCR-binding CIDRα1s, due to the conserved chemistry of the binding site, it was not clear whether this would be the case. Therefore, a panel of bundles were designed that represented the sequence diversity of the different classes of EPCR-binding CIDRα1s that could be tested in parallel.

Following the successful design of Bundle1.4, which bound to EPCR, the bundle scaffold was kept the same and only those residues at the surface which directly contact EPCR were mutated to represent the different sequences of the subclasses. For each different CIDRα1 subclass, an alignment of all known sequences was made and the variant which most closely matched the consensus was chosen as a representative for that subtype. For example, for the subtype CIDRα1.1, the IT4var20 sequence had the most similarity to the consensus sequence made from all known CIDRα1.1 sequences. The names of the CIDRα1 domains that were chosen to represent their subtype are shown in **Table 3.2** and the protein sequences are shown in **Appendix A.1.1**.

Bundle Subclass	Based on CIDRα1 Domain
Bundle1.1	IT4var20
Bundle1.4	HB3var03
Bundle1.5	1965_2
Bundle1.6	PFD1235w
Bundle1.7	2083_1
Bundle1.8	1702_3

Table 3.2: CIDRα1 domains on which each subclass of Bundle was Designed

GeneStrings of all of the bundles were ordered and they were cloned into the pEt15bcon vector backbone, using the primers specified in **Appendix A.1.2 and A.1.3**. The vector contains the DNA sequence for an N-terminal His-tag for pu-

rification of the recombinant protein and a C-terminal SpyTag was included in the GeneStrings for immunisation purposes. The Bundle protein sequences can be seen in **Figure 3.10**, with the mutations in the EPCR-binding site compared to Bundle1.4 highlighted in green.

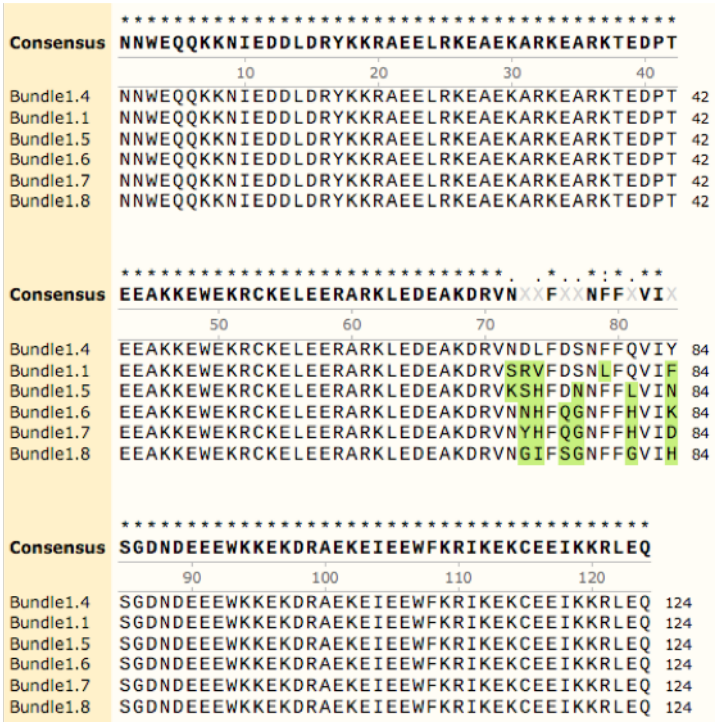


Figure 3.10: Bundle Sequence Alignment An alignment of the bundle sequences with the mutations in the EPCR-binding site from Bundle1.4 highlighted in green for each construct.

3.2.4 Bundle Expression and Purification

The bundles were grown up in BL21(DE3) *E. coli* cells in 2xYT media broth until OD_{600}=0.8-1.0, before inducing expression with 1 mM IPTG at 25°C overnight. The bundles were expressed almost exclusively in the insoluble fraction of the cell pellet. Subsequently, the inclusion bodies were completely denatured in 6 M guanidine hydrochloride, before being refolded whilst bound to nickel resin via their His-tag in a decreasing gradient of guanidine hydrochloride **(Materials and Methods 2.4.4)**. The eluted fractions were analysed by SDS-PAGE to assess protein content, and

for crystallography the flexible N-terminal purification tags were removed using TEV protease cleavage **(Materials and Methods 2.4.5)**. These preparations were then gel filtered on S75 16/60 (GE Healthcare) to ensure purity of homogeneous monomeric protein **(Figures 3.11 & 3.12)**. The presence of the multiple bands for each of the purified bundles was always seen and was thought to be a gel artifact due to partial reduction of the disulphide bond in each of the constructs, causing two species to migrate slightly different on the gel. Additionally, proteolysis of the sample could be causing the presence of multiple bands on a reduced gel.

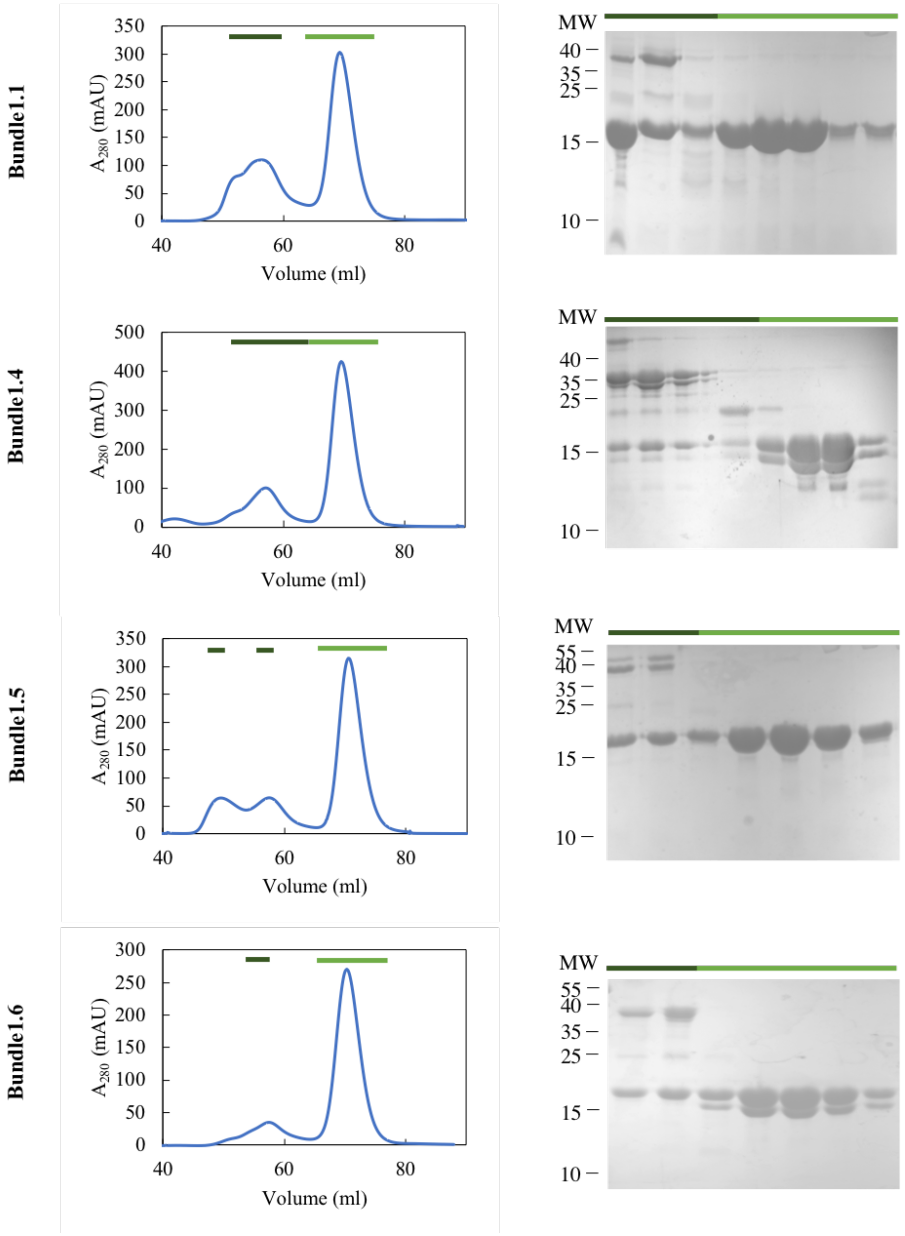


Figure 3.11: Purification of Bundle constructs SEC trace (left) and corresponding reduced SDS-PAGE gel (right) showing the elution of these proteins as a monodispersed peak. The light green bar represents the fractions that were pooled for further studies.

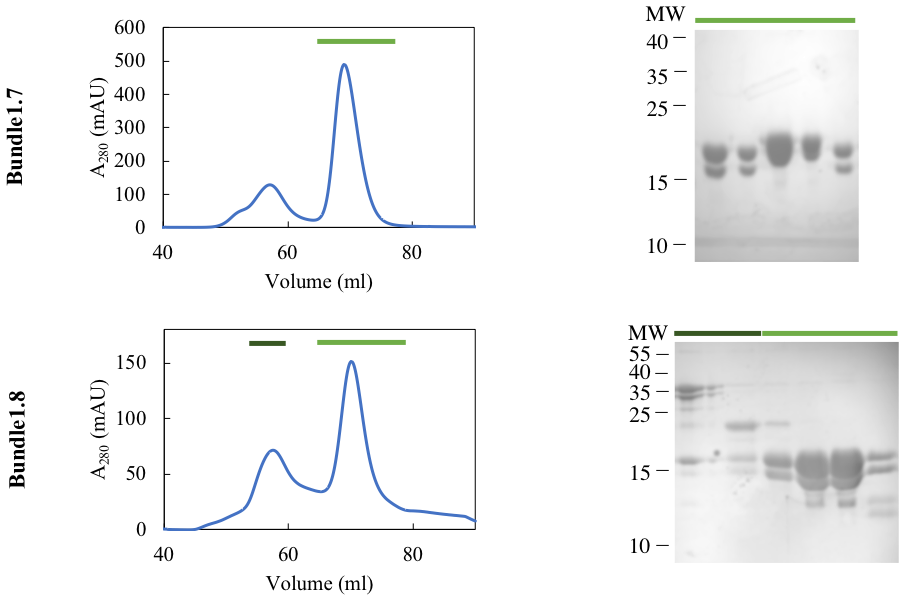


Figure 3.12: Purification of Bundle Constructs Continued SEC trace (left) and corresponding reduced SDS-PAGE gel (right) showing the elution of these proteins as a monodispersed peak. The light green bar represents the fractions that were pooled for further studies.

3.2.5 EPCR Expression and Purification

An EPCR construct had already been cloned into the pExpreS^2-1 vector and purification protocol optimised in the laboratory by Clinton Lau [205]. The pExpreS^2-1 vector **(Figure 2.2)** had four N-terminal elements: a BiP leader sequence for protein secretion; a biotin acceptor (BAP) tag; a hexa-histidine tag and a TEV protease cleavage site. These tags allowed for secretion and purification of the protein as well as being useful for biophysical experiments. A stable S2 cell line transfected with this EPCR construct was thawed and expanded to four litres **(Materials and Methods 2.2.2)**. The EPCR-containing supernatant was buffer exchanged and concentrated, before being applied to a Ni-NTA column **(Materials and Methods 2.2.2.3)**. Size exclusion gel chromatography column (S75 16/60, GE Healthcare) was performed with the EPCR eluting as a monodisperse peak.

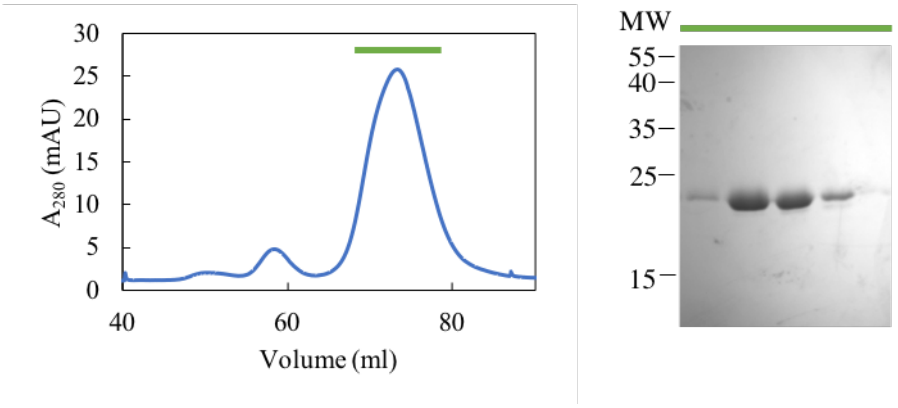


Figure 3.13: Purification of EPCR SEC trace (left) and corresponding reduced SDS-PAGE gel (right) showing the elution of these proteins as a monodispersed peak. The light green bar represents the fractions that were pooled for further studies.

EPCR from S2 cells was incubated overnight with TEV protease, again removing uncleaved protein and TEV protease using NiNTA **(Materials and Methods 2.4.5)**. The EPCR construct also had four predicted glycosylation sites, which could be glycosylated by the S2 expression system. These glycans were removed using a combination of endoglycosidases EndoHf and EndoF3 **(Materials and Methods**

2.4.7), with overnight reactions showing glycan removal by a shift on SDS-PAGE gel **(Figure 3.14)**.

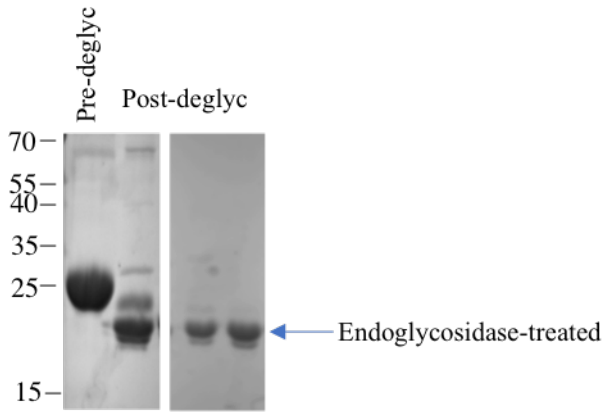


Figure 3.14: Deglycosylation of EPCR A reduced SDS-PAGE gel showing the pre-deglycosylation (pre-deglyc) EPCR sample and the fully deglycosylated EPCR samples (post-deglyc).

3.2.6 Assessing the Fold of the Bundle Immunogens

Since the bundle constructs were expressed and refolded from inclusion bodies, it was important to check that the purified protein isolated from the gel filtration contained secondary structure. To check this, circular dichroism (CD) spectra were taken of the six expressed bundles. A CD spectrum was taken between 190nm and 250nm at room temperature for each bundle, correcting for the buffer contribution. The spectra showed that all of the bundle constructs contained a high alpha helical content as expected, characterised by CD minima at 222nm and 208nm **(Figure 3.15)**.

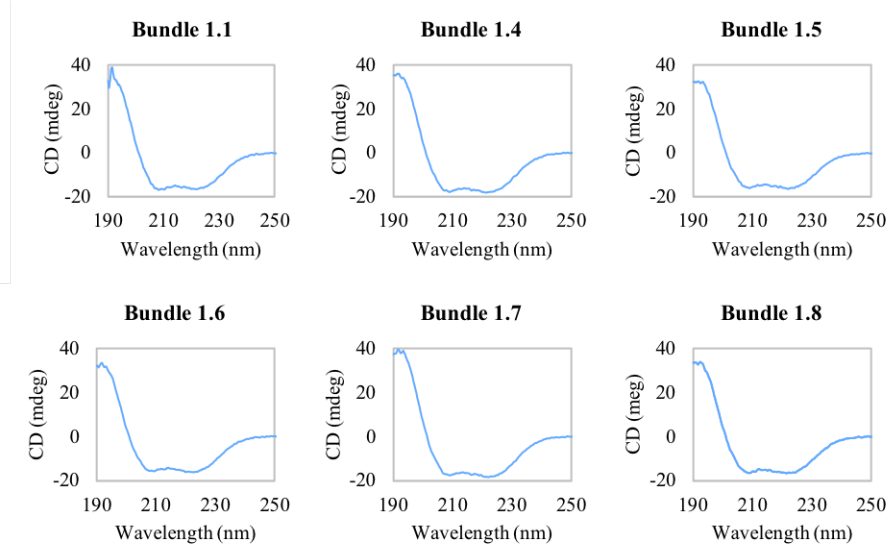


Figure 3.15: CD Spectra of Bundles

Furthermore, in order to assess the stability of these bundles (an important feature of potential vaccines), a CD thermal melt experiment was carried out **(Materials and Methods 2.6.1)** between 20°C and 90°C **(Figure 3.16)**. When proteins unfold they lose these highly ordered structures and the CD bands change, therefore this experiment can be used to determine melting temperatures. All of the bundles had similar melting temperatures above 60°C **(Table 3.3)**, which indicates that the proteins are stable as potential immunogens in the hot climates where malaria is prevalent.

Construct	Melting Temperature °C
Bundle1.1	61
Bundle1.4	60
Bundle1.5	67
Bundle1.6	63
Bundle1.7	65
Bundle1.8	61

Table 3.3: Melting Temperatures of the Different Bundles

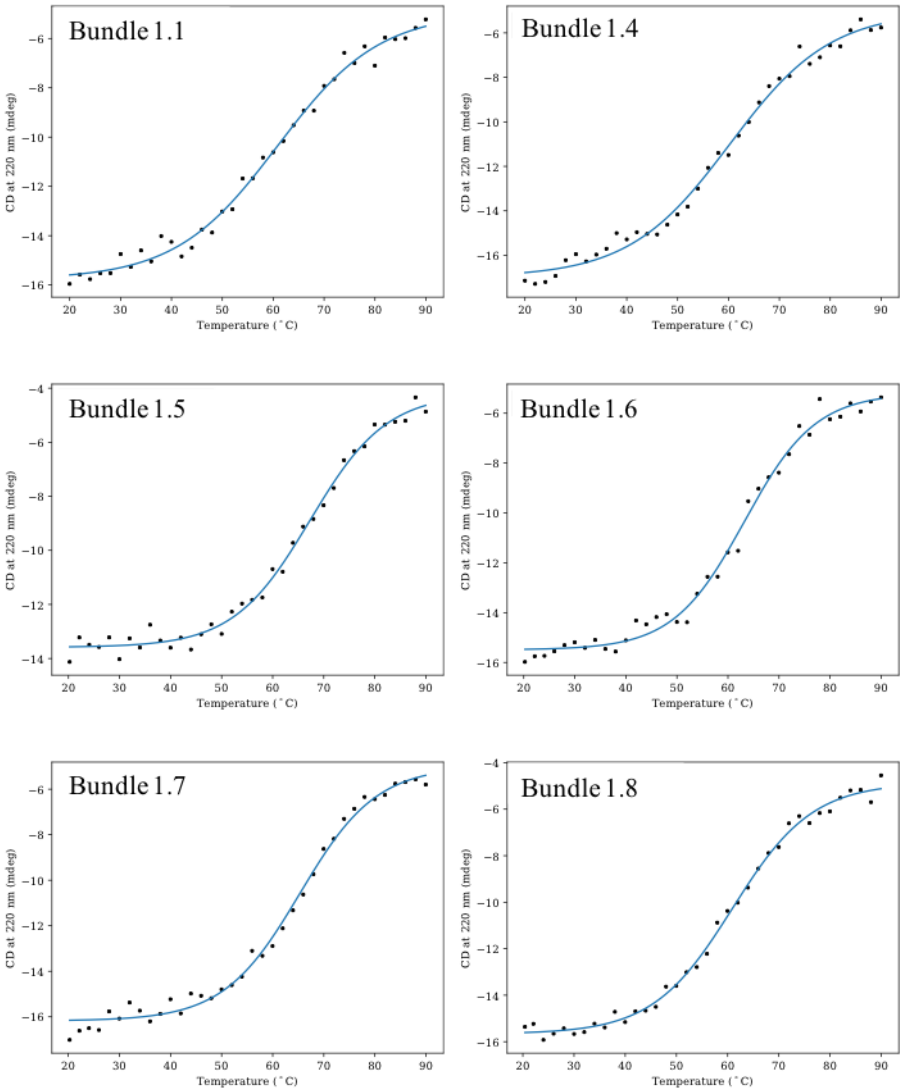


Figure 3.16: CD melting curves of Bundles at 220nm

3.2.7 Binding of Bundles to Recombinant EPCR

In order to check the function of the recombinant bundles, their binding to EPCR by surface plasmon resonance (SPR) was used to show that they had been correctly refolded and successfully mimicked the EPCR-binding sites of their CIDRα1 counterparts.

EPCR was coupled to the capture chip by biotinylating the biotin acceptor peptide (BAP), included in the S2 EPCR construct **(Materials and Methods 2.6.2.2.2)**.

This system allowed the presentation of a surface of EPCR in the same orientation as on the cell, as the BAP tag was designed on the N-terminus of the EPCR construct, which is proximal to the membrane helix *in vivo*. The bundles were then flowed over the EPCR-coated chip. Bundle1.1, Bundle1.4, Bundle1.5, and Bundle1.6 all bound to EPCR **(Figure 3.17)**.

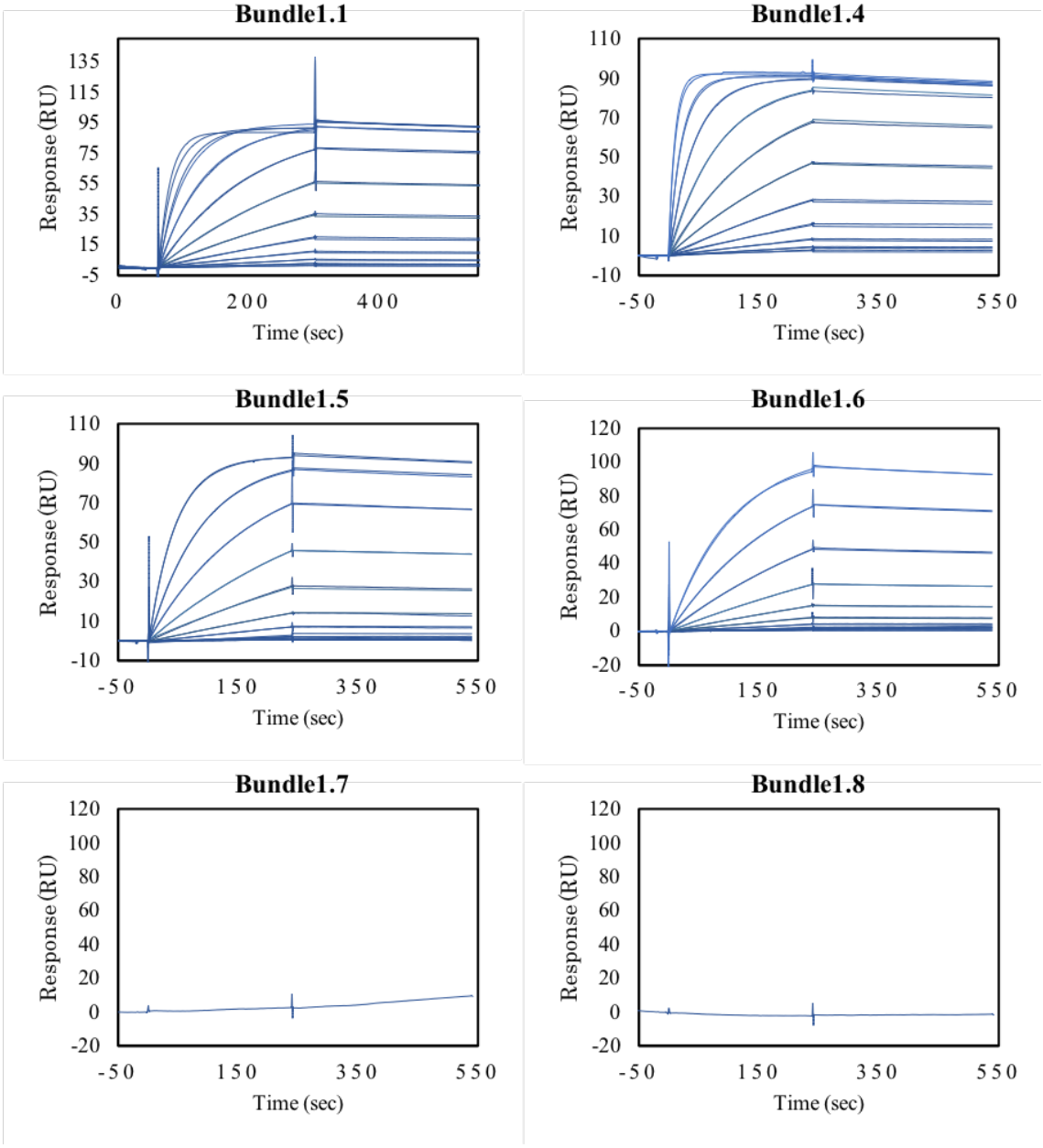


Figure 3.17: SPR curves showing binding of Bundles to EPCR
A 2-fold dilution series of each Bundle was analysed, with a starting concentration of 16 μM, and twelve curves in total for each Bundle.

And so the kinetics of these interactions could be assessed by flowing over a two-fold dilution series of each bundle in turn, starting at 16μM **(Table 3.4)**. All of the bundles bound to the EPCR with nanomolar affinity and a slow off-rate (x10^{-4} s^{-1}) characteristic of CIDRα1 binding to EPCR.

	$\mathbf{k}_{on}$ ($\mathbf{M}^{-1}s^{-1}$)	$\mathbf{k}_{off}$ ($\mathbf{s}^{-1}$)	$\mathbf{K}_d$ (nM)
Bundle1.1	3680	1.31x10^{-4}	36.0
Bundle1.4	6090	1.57x10^{-4}	26.0
Bundle1.5	1400	1.36x10^{-4}	97.0
Bundle1.6	635	1.62x10^{-4}	250

Table 3.4: Kinetic Analysis of Bundles Binding to EPCR

In order to assess whether the EPCR-binding site grafted onto the bundle scaffold matched the original CIDRα1.4, the binding kinetics of the Bundle1.4 was compared with those of the original CIDRα1.4. However, as described previously, since only seven out of the nine interacting residues were grafted onto the Bundle1.4, a better comparison is with the CIDRα1.4 mutant in which the two residues outside of the grafted region are mutated to nullify their effect on the binding **(Section 3.1.5)**. The mutant bound with K_d=220nM affinity, and the bundle with K_d=26.0nM affinity **(Table 3.5)**, showing that the Bundle1.4 successfully mimics the EPCR-binding of the CIDRα1 domain.

	$\mathbf{k}_{on}$ ($\mathbf{M}^{-1}s^{-1}$)	$\mathbf{k}_{off}$ ($\mathbf{s}^{-1}$)	$\mathbf{K}_d$ (nM)
Bundle1.4	3680	1.31x10^{-4}	26.0
CIDRα1.4	53000	2.00x10^{-4}	0.37
CIDRα1.4 Mutant	9100	2.00x10^{-3}	220

Table 3.5: Comparison of Bundle1.4 and CIDRα1.4 Binding to EPCR

Despite the success of grafting the EPCR-binding sites of four of the different subtypes onto the three helical bundles, the bundles designed from the CIDRα1.7

and CIDRα1.8 sequences (Bundle1.7 and Bundle1.8) showed no binding to EPCR even at 16 μM **(Figure 3.17)**. This was further confirmed by attempting small scale complex formation on a SEC column **(Figure 3.18)**. As can be seen neither the Bundle1.7, nor the Bundle1.8, co-eluted with EPCR. The Bundle1.7 and Bundle1.8 eluted later than the EPCR, highlighting that these two proteins do not form a complex. This suggested that the EPCR-binding mechanism may be slightly different for these subclasses of CIDRα1, potentially involving residues outside of the grafted helix-kinked-helix motif.

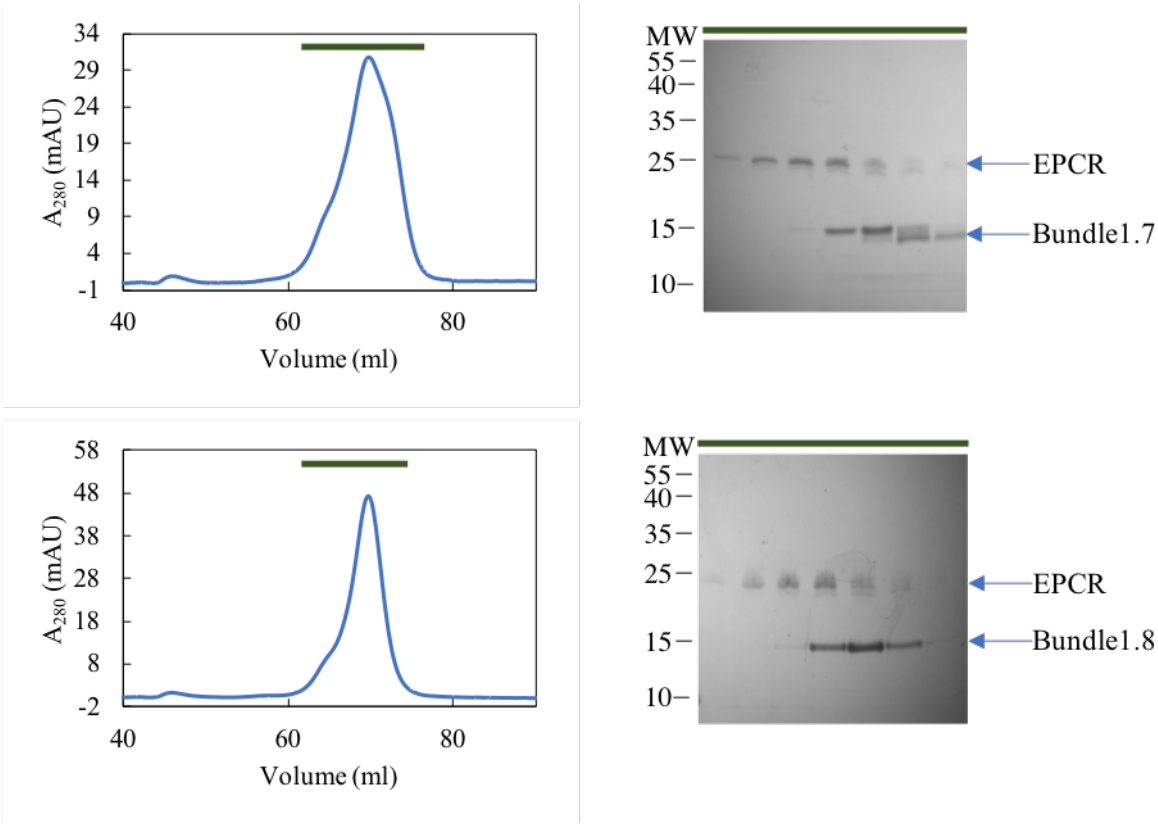


Figure 3.18: Bundle1.7:EPCR and Bundle1.8:EPCR Attempted Complex Formation SEC trace (left) and corresponding reduced SDS-PAGE gel (right) showing that these proteins do not co-elute. The dark green bar represents the corresponding regions of the SEC trace and SDS-PAGE gel.

3.2.8 Structural Basis of the Bundle:EPCR Interaction

3.2.8.1 Complex Formation

For the four Bundles that did bind to EPCR, structural studies were carried out to confirm that the Bundles bound to EPCR in the same way as their parental CIDRα1. For complex formation, the purified Tev-cleaved Bundle and the Tev-cleaved and deglycosylated EPCR were incubated together for thirty minutes at room temperature. SEC (GE Healthcare, S200 16/60) was then carried out to ensure that both components eluted as one monodisperse peak showing that a complex had been formed **(Figure 3.19)**.

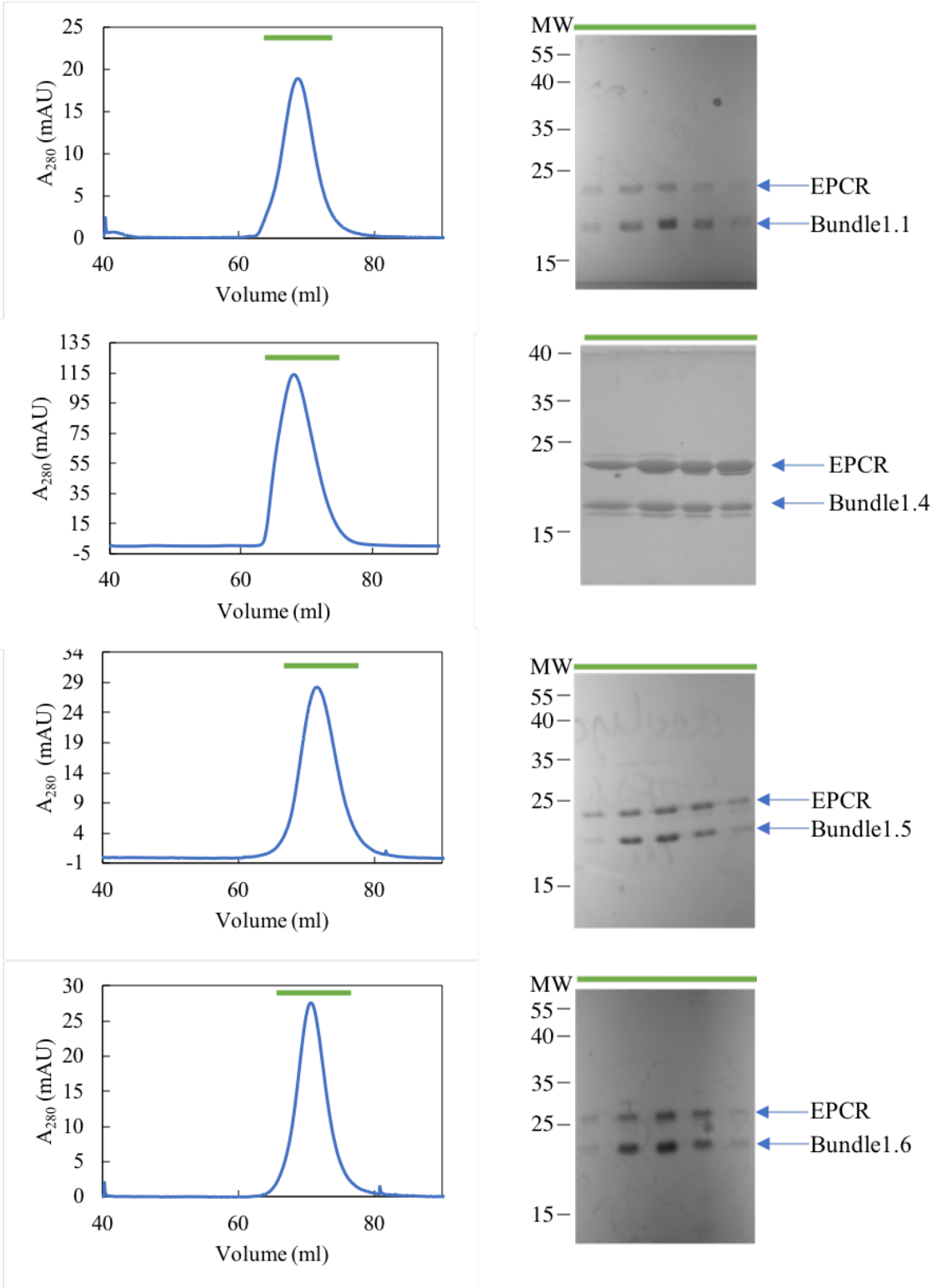


Figure 3.19: Bundle:EPCR Complex Formation SEC trace (left) and corresponding reduced SDS-PAGE gel (right) showing the elution of these proteins as a monodispersed peak. The light green bar represents the fractions that were pooled for further studies.

3.2.8.2 Bundle1.4:EPCR Crystallisation

The Bundle1.4:EPCR construct was crystallised first. The peak which corresponded to the Bundle1.4:EPCR complex **(Figure 3.19)** was pooled and concentrated to 10.7 mg/ml. Likewise, the Bundle1.4 alone peak **(Figure 3.11)** was pooled and concentrated to 12.5 mg/ml. The following screens were set up in 96-well sitting drop plates, using the commercial screens Morpheus, Midas, JCSG+ and Proplex (all Molecular Dimensions) at both 4 °C and 18 °C **(Materials and Methods 2.7.2.1)**.

The bundle alone did not crystallise with 80% of the drops being clear of precipitation. The bundle was then methylated **(Materials and Methods 2.4.9)** in an attempt to decrease its solubility and promote crystallisation **(Figure 3.20)**. The small upwards shift on the SDS-PAGE gel suggested that methylation was successful. The methylated bundle was also concentrated more highly to 25 mg/ml to try reduce solubility. Although this did increase the amount of precipitation with 60% drops clear, there was still no crystallisation observed for the bundle alone.

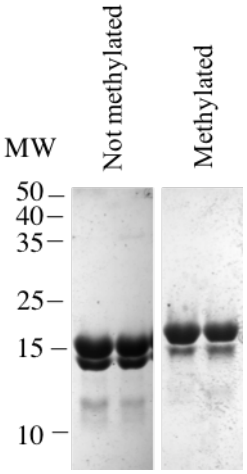


Figure 3.20: Methylation of Bundle1.4 A reduced SDS-PAGE gel of the bundle before and after methylation

However, crystals of the Bundle1.4:EPCR complex grew in two different forms in many ProPlex conditions at 4°C. The biggest and most clearly defined crystals grew

as hexagonal bi-pyramids in 0.1 M HEPES pH 7, 1 M Sodium Citrate (condition G12). Other large crystals grew as irregular cuboids in 0.1 M HEPES pH 7, 18% w/v PEG 12000 (condition F4). SDS-PAGE analysis followed by silver stain of the gel was used to confirm that the two different crystal forms contained both Bundle1.4 and EPCR **(Figure 3.21)**.

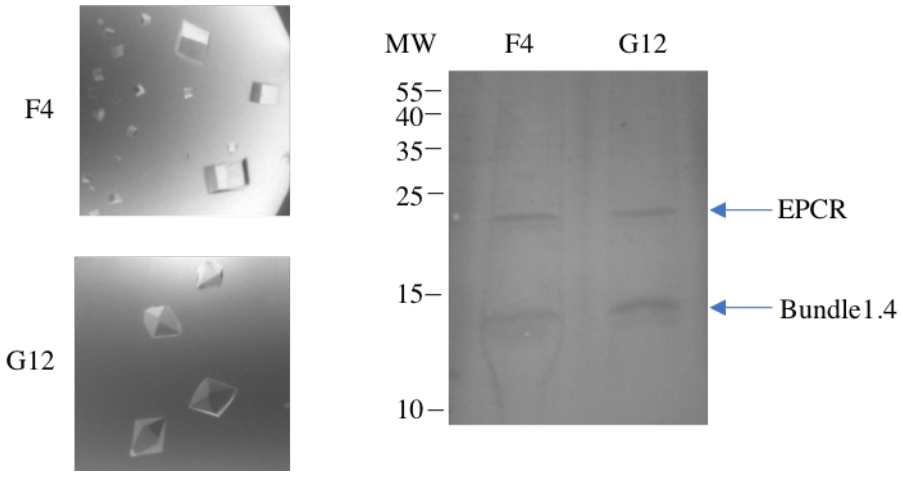


Figure 3.21: Silver stain of Bundle1.4:EPCR crystals Left: the different forms of crystals, F4 as irregular cuboids, and G12, as hexagonal bi-pyramids. Right: the silver stained reduced SDS-PAGE gel of both F4 and G12 crystals.

Optimisation screens of both types of crystal were set up using the Silver Bullets additive screen, both with and without crystal seeds **(Materials and Methods 2.7.2.2)**. Crystals formed abundantly and the best were picked and harvested after roughly one week of growth. The crystals were looped into drops of mother liquor containing 25% glycerol and cryo-cooled in liquid nitrogen for storage and data collection.

3.2.8.3 Data Collection and Solution of the Structures of Bundle1.4:EPCR

Data were collected for the Bundle1.4:EPCR complex at the Diamond Light Source on beamline I-03. Initially a series of cryocooled crystals were mounted on the

gonionmeter and shot at three angles (0°, 45°, and 90°) to check for diffraction quality and resolution limits. This identified a crystal from the G12 optimisation conditions with 2 M sodium citrate, 0.1 M HEPES pH7 25% glycerol, which diffracted to 3.11Å. A full dataset was collected from this crystal, using a collection strategy suggested in the outputs of EDNA and Mosflm. This dataset was processed automatically by Xia2, which suggested a space group of $P3_221$ and the data collection statistics can be seen in **Table 3.6**.

Space Group	$P3_221$
Cell Parameters (Å)	a=b=112.73, c= 168.49
(°)	α=β=90, γ=120
Resolution (Å)	31.78-3.11 (3.22-3.11)
R_{meas}	0.079 (1.35)
R_{pim}	0.031 (0.246)
I/σ (I)	15.9 (3.89)
$CC_{1/2}$	1.00 (0.80)
Completeness (%)	99.6 (99.9)
Multiplicity	4.80 (4.90)

Table 3.6: Data collection statistics for the Bundle1.4:EPCR complex (Values in parentheses are the highest-resolution shell).

Phase information was obtained by using EPCR (PDB code: 4V3D) as a search model for molecular replacement. The top result for this search showed two EPCR molecules per asymmetric unit, which was consistent with the Matthew's coefficient estimation. Additional electron density (for the Bundle1.4) next to one of the EPCR molecules was seen. By building short helices into this density in Coot, and refining the dataset in Buster with the missing atoms demand, additional electron density appeared. Successive rounds of refinement and rebuilding allowed this three-helical bundle to be built, and using the predicted structure for the bundle from Rosetta as a guide, the bundle sequence could be assigned to the main chain. Most of the

bundle could be resolved except the loop of residues 35-50, which could not be seen in the electron density. Rather unusually, it appeared that in each asymmetric unit there was one Bundle:EPCR complex and one free EPCR molecule **(Figure 3.22 A)**. Looking at the crystal packing in **Figure 3.22 B** there was a bundle from one asymmetric unit (grey) and an EPCR from another asymmetric unit (green) that packed directly against the unbound EPCR (lilac), making it impossible for an additional bundle to bind.

The refinement statistics can be seen in **Table 3.7**. The R_{work} value of 0.208 suggested a good fit of model to data at the current resolution (3.11Å), while the slightly higher R_{free} value of 0.247 showed that independent, unprocessed reflections could be explained by the model and suggested that the model did not overinterpret the data. The Ramachandran plot (made using the RAMPAGE server [219]) showed that a high number of residues are in the favoured region (96.1%), with nearly all the rest in the allowed region (3.4%) **(Figure 3.23)**. There are two outliers (making up the remaining 0.5%) which are two threonines, one from each copy of EPCR (T88). This particular residue is strained in all solved EPCR structures, suggesting that it is modelled correctly.

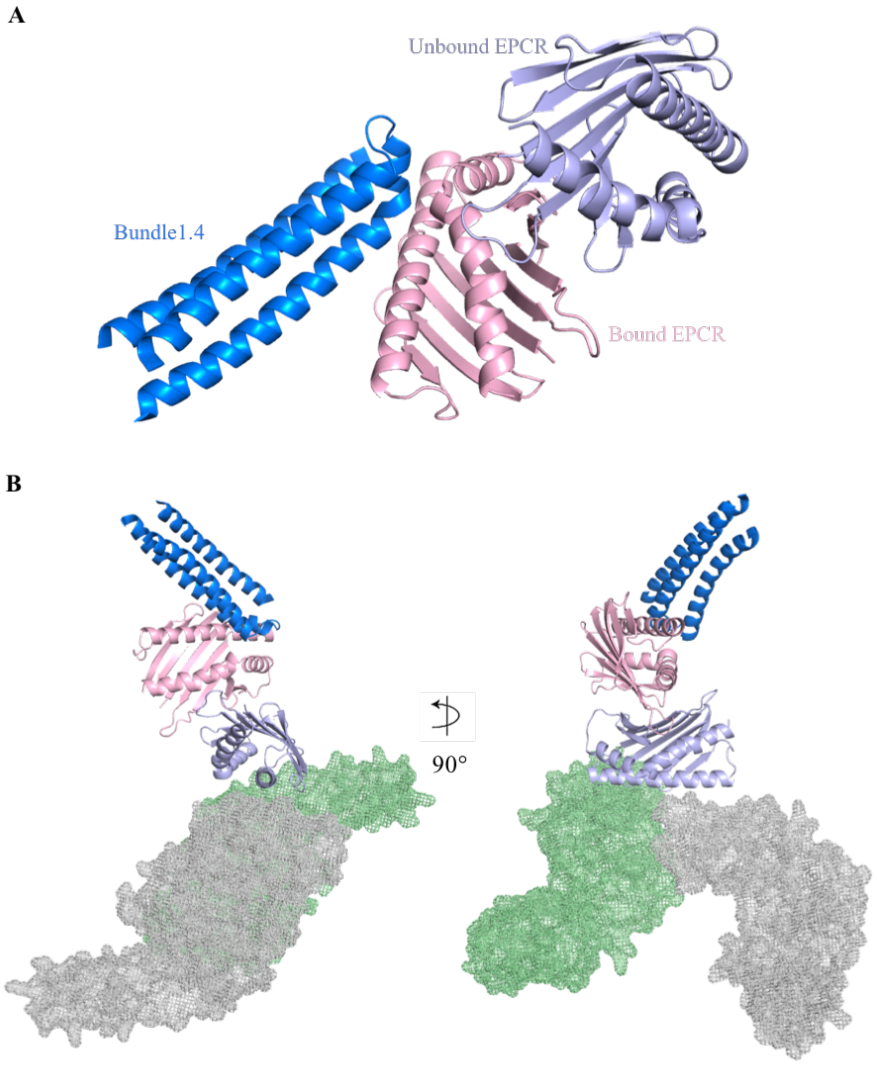


Figure 3.22: Asymmetric Unit and Crystal Packing in Bundle1.4:EPCR Crystal A) The Bundle1.4 (blue) bound to one copy of EPCR (pink) with another free unbound copy of EPCR (lilac) in each asymmetric unit of the crystal. B) The crystal packing shown with three asymmetric units, one displayed as a ribbon (same colours as A), and the other two displayed as surface mesh one in green and one in grey.

Resolution (Å)	31.78-3.11
R_{work}	0.208
R_{free}	0.247
rmsd bond lengths (Å)	0.014
rmsd bond angles (°)	1.88
Ramachandran plot	
Favoured region (%)	96.1
Allowed region (%)	3.4
Disallowed region (%)	0.5

Table 3.7: Refinement statistics of the Bundle1.4:EPCR complex

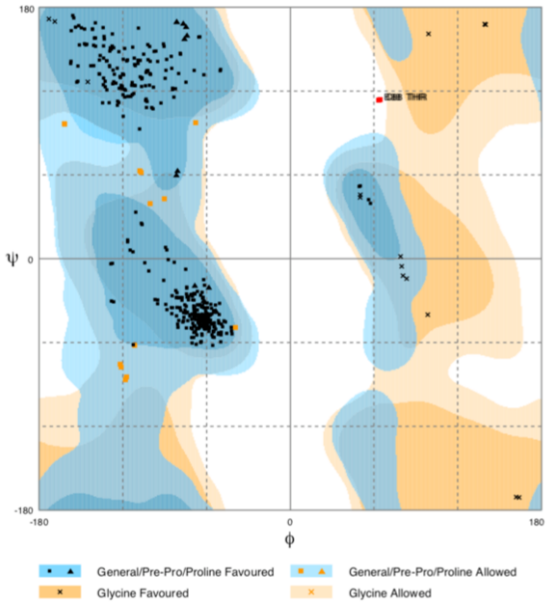


Figure 3.23: Ramachandran Plot of Bundle1.4:EPCR Structure

3.2.8.4 Determining the Structures of the Other Bundle:EPCR Complexes

Having solved the Bundle1.4:EPCR structure, it was a much easier process to determine the other Bundle:EPCR structures. As the other bundles have >95% sequence identity to the Bundle1.4 and were predicted to bind to EPCR via the same mode, it was thought that they would likely crystallise in the same conditions. Therefore, only optimisation plates of the condition in which Bundle1.4:EPCR crystals grew and diffracted needed to be set up. As a result Silver Bullet additive screens were set up in 0.1 M HEPES pH 7, 1 M Sodium Citrate (ProPlex, G12) **(Materials and Methods 2.7.2.2)**. For the Bundle1.6:EPCR complex, hexagonal bipyramidal crystals grew to a size large enough for data collection. However, crystals originally did not grow for Bundle1.1:EPCR nor Bundle1.5:EPCR. In order to promote crystal growth, Bundle1.4:EPCR crystals were used as cross-seeds to nucleate the crystal growth of the other Bundle:ECPR complexes **(Materials and Methods 2.7.2.2)**. This was successful and large hexagonal bi-pyramidal crystals grew for both the Bundle1.1:EPCR and Bundle1.5:EPCR complexes. These crystals were harvested, stored and diffraction data collected in the same way as the Bundle1.4:EPCR crystals. The data collection statistics can be seen in **Table 3.8**.

The unit cell sizes for the different Bundle:EPCR complexes were similar to that of the Bundle1.4:EPCR complex. Due to this and the fact that the protein composition was almost identical, it was assumed that the crystal packing was identical to Bundle1.4:EPCR. Therefore, an alanine model (in which all residues were mutated to alanine), with all loops and the binding site removed, based on the Bundle1.4:EPCR structure solution (with two copies of EPCR and one copy of Bundle) was used as a search model in Molecular Replacement (Phaser). After refinement, electron density could be seen for the loop regions and binding site in the model, with more electron density seen after successive rounds of rebuilding and refinement. In addition,

	Bundle1.1:EPCR	Bundle1.5:EPCR	Bundle1.6:EPCR
Space Group	$P3_221$	$P3_221$	$P3_221$
Cell Parameters (Å)	a=b=113.13, c=168.00	a=b=115.76, c=171.99	a=b=115.14, c=171.85
(°)	α=β=90, γ=120	α=β=90, γ=120	α=β=90, γ=120
Resolution (Å)	84.69-3.43 (3.55-3.43)	86.61-2.67 (2.77-2.67)	99.77-2.93 (3.04-2.93)
R_{meas}	0.119 (1.54)	0.131 (1.53)	0.125 (3.30)
R_{pim}	0.0384 (0.482)	0.0421 (0.551)	0.0408 (1.07)
I/σ (I)	10.31 (1.52)	11.11 (1.19)	9.71 (0.94)
$CC_{1/2}$	0.999 (0.843)	0.996 (0.686)	0.998 (0.616)
Completeness (%)	98.8 (96.7)	99.3 (96.4)	98.6 (96.0)
Multiplicity	9.70 (10.1)	9.49 (7.40)	9.50 (9.40)

Table 3.8: Data Collection Statistics for the Other Bundle:EPCR Complexes Data collection statistics for Bundle1.1:EPCR, Bundle1.5:EPCR and Bundle1.6:EPCR. (Values in parentheses are the highest resolution shell).

since an alanine model had been used, electron density for the side chains could be used to assign the expected sequence and the side chains that differed between the bundles confirmed that the correct solution had been found. All of the Bundle:EPCR structures, like that of Bundle1.4:EPCR had one bound and one unbound EPCR in their unit cells. The solutions reached again were verified by analysing their low R factors and that the R_{work} value was below the R_{free} **(Table 3.9)**. Moreover, looking at the Ramachandran plots **(Figure 3.24)**, at least 90% of the residues in all of the structures have chi and psi angles that place them in the favoured region of the plot.

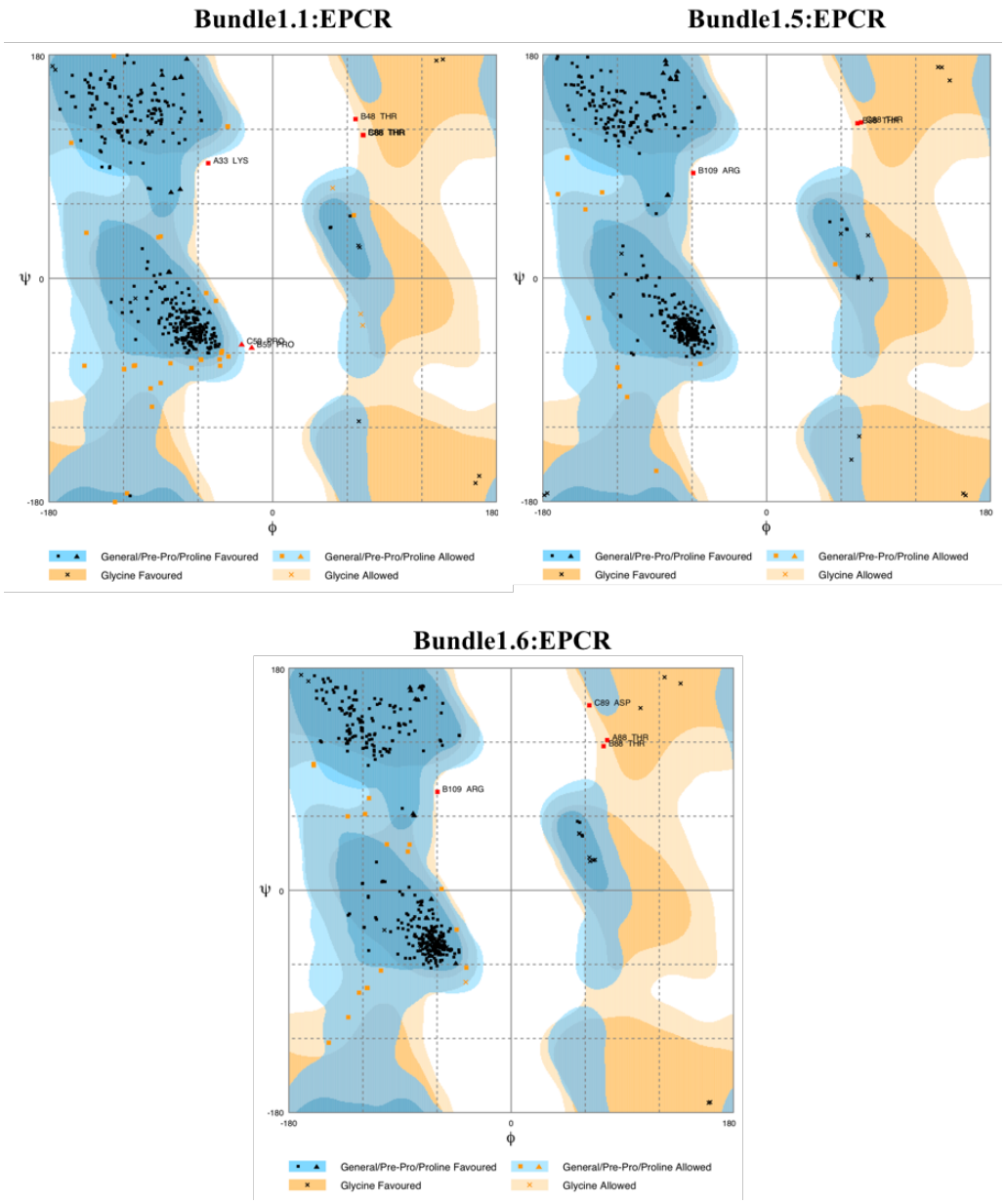


Figure 3.24: Ramachandran Plots for Bundle:EPCR Structures

	Bundle1.1:EPCR	Bundle1.5:EPCR	Bundle1.6:EPCR
Resolution (Å)	3.43	2.67	2.93
R_{work}	0.210	0.234	0.237
R_{free}	0.263	0.263	0.264
rmsd bond lengths (Å)	0.02	0.02	0.02
rmsd bond angles (°)	2.00	1.92	1.95
Ramachandran Plot			
Favoured Region (%)	90.3	95.7	94.5
Allowed Region (%)	8.6	3.6	4.4
Disallowed Region (%)	1.1	0.7	1.1

Table 3.9: Data Refinement Statistics for the Bundle:EPCR Complexes

3.2.9 Analysis of the Crystal Structures

3.2.9.1 Comparison of the Rosetta Model and the Structure of Bundle1.4

The structure of Bundle1.4 was compared with the predicted structure of the model produced in the Rosetta design process **(Figure 3.25A)**. By aligning the two structures, it was clear that the EPCR-binding site adopted a very similar conformation to that predicted by the model, with the Asn72-Trp93 region of the helix-kink-helix motif overlaying with an RMSD of 1.2Å. Elsewhere, the path of the helical bundle diverged from the design with an overall RMSD of 2.0Å, with the most difference lying in residues at the beginning and end of the first helix (Asn1-Lys8, and Lys30-Glu34). The variation in RMSD between the two structures is shown clearly in **Figure 3.25B**. The EPCR-binding site is mainly dark blue, highlighting it's similarity to the model. Whereas, the both ends of the N-terminal helix of Bundle1.4 deviate

from the predicted model, illustrated by the white and red colouring.

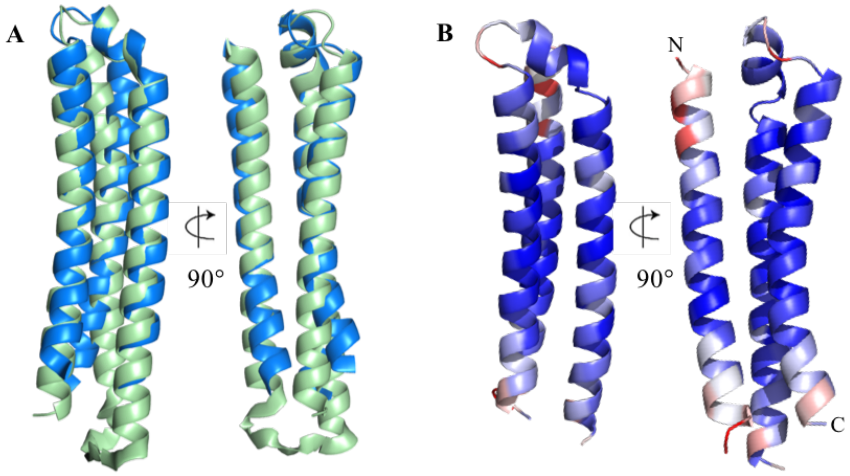


Figure 3.25: Comparison of Bundle1.4 to Rosetta Model A) An overlay of the Rosetta Model predicted structure (green) and Bundle1.4 structure (light blue). B) The Bundle1.4 structure coloured by RMSD to the Rosetta Model, with blue being the areas of lowest RMSD and red being the highest areas of RMSD. The spectrum range is from 0.35Å to 3.78Å.

3.2.9.2 Comparison of the Bundle1.4:EPCR Complex to the CIDRα1.4:EPCR Complex

The Bundle1.4:EPCR structure was also compared directly to the CIDRα1.4:EPCR structure from which the EPCR-binding site was grafted. Overlaying the two structures showed that the Bundle1.4 binds to EPCR by the same mechanism as the CIDRα1.4 **(Figure 3.26)**. An alignment of these two structures showed that the EPCR-binding region (Asn72-Trp93) overlays with an RMSD of 0.77Å, highlighting that the position of the amino acids match very closely. As expected and designed, the EPCR-binding site of the Bundle1.4, is very similar to that of the equivalent CIDRα1 bound to the EPCR. The central phenylalanine (Phe80) protrudes into a hydrophobic cleft in the EPCR (consisting of hydrophobic portions of 8 amino acid side chains plus a small portion of the bound lipid). This protruding Phe80

is positioned at the tip of a kinked helix, which is stabilised by a phenylalanine (Phe75) and a valine (Val82), that pack together on the opposite side of the kink. This demonstrates that the designed synthetic protein contains the core of the EPCR-binding site of the CIDRα1 domains, and binds in the predicted manner to EPCR.

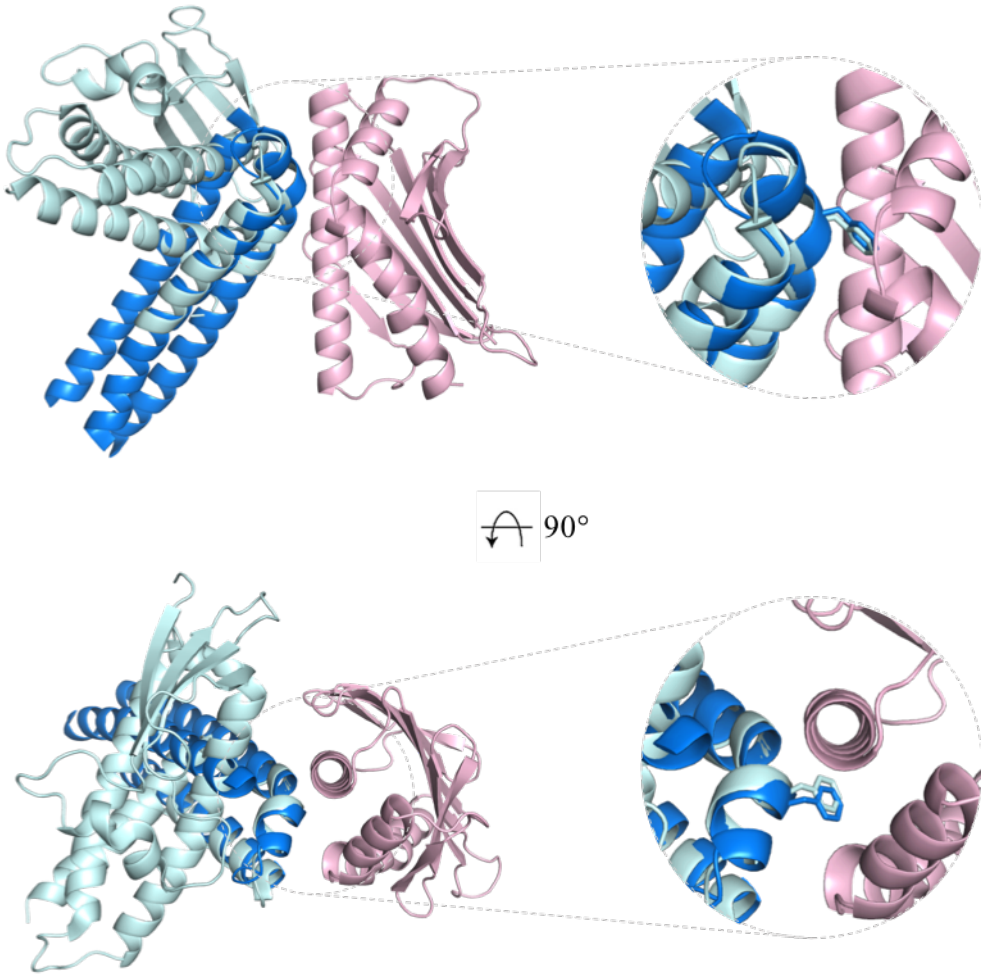


Figure 3.26: Comparison of Bundle1.4:EPCR and CIDRα1.4:EPCR Structures An overlay of the Bundle1.4 (dark blue) and the CIDRα1.4 (teal) bound to EPCR (pink), with a close up view of the binding site, showing the position of the phenylalanine crucial for binding (F80 in the Bundle1.4, and F656 in the CIDRα1.4.

3.2.9.3 Comparison of the Bundles Binding to EPCR

Overlaying the structures of the four different Bundle:EPCR complexes, showed that they all adopted the helix-kink-helix EPCR-binding site **(Figure 3.27 A)**. The positions of the sidechains within the helix-kink-helix binding site are very similar **Figure 3.27 B)**. Moreover, the conserved residues, shown in brick red, adopt very similar conformations in the different Bundles especially the protruding phenylalanine that is crucial for EPCR-binding.

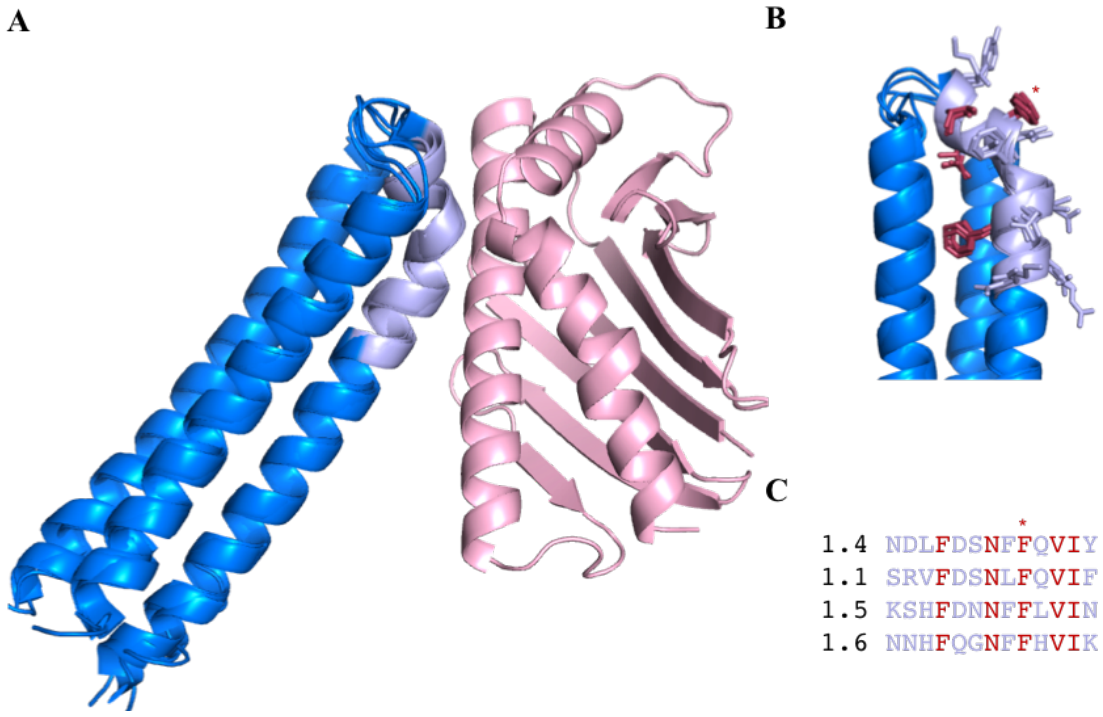


Figure 3.27: Comparison of Different Bundles A) An overlay of the different subclasses of bundles 1.1, 1.4. 1.5 and 1.6 (blue) bound to EPCR (light pink). B) An overlay of the bundles showing the grafted area of the EPCR binding site (pale blue) and side chains. The side chains of the residues that are conserved across all the bundles are shown in brick red, with asterisk masking the crucial central phenylalanine. C) The sequences for the different bundles between residues 72-84 as colour coded in B).

3.2.10 Bundles as Immunogens

3.2.11 Declaration

The following immunisations and accompannying experiments were carried out by members of the Lavstsen group at the University of Copenhagen.

3.2.11.1 Binding of the Bundle1.4 to Inhibitory Antibodies raised by Immunisations with CIDRα1.4

Having successfully designed and characterised the bundles, it was tested whether these potential immunogens bound to inhibitory antibodies either from immunisation of rats with CIDRα1 domains, or from sera of humans in malaria endemic areas. This would show whether antibodies that bind to the EPCR-binding helix-kink-helix motif were present and also whether these antibodies were inhibitory.

Antibodies from immunisations of rats with the CIDRα1.4 domain, HB3var03, were affinity purified from sera on columns coupled either to HB3var03 or the equivalent Bundle1.4. When affinity purified on the CIDRα1.4 (HB3var03) column, the IgG in the run-through (RT) only contained $\sim$20% of the inhibition of CIDRα1 binding to EPCR compared to the original sera **(Figure 3.28)**. This suggested that the majority of the inhibitory antibodies bound to the CIDRα1.4 column. However, in contrast, when affinity purified on a Bundle1.4 column, the RT had the same capacity as the original sera, suggesting that the majority of the inhibitory antibodies did not bind to Bundle1.4. Despite this, the antibodies that were subsequently eluted from the Bundle1.4 column did have some inhibitory activity, with a 50% dilution of the eluted antibodies causing a 40% decrease in CIDRα1 binding to EPCR. This can be compared to the antibodies eluted from the HB3var03 column (diluted between 50% and 3.1%), which all showed nearly 100% inhibition of CIDRα1 binding to

EPCR **(Figure 3.28)**. As a result, despite the Bundle1.4 adopting the correct EPCR-binding site conformation, it was not recognised by the majority of CIDRα1 reactive antibodies that inhibit EPCR binding.

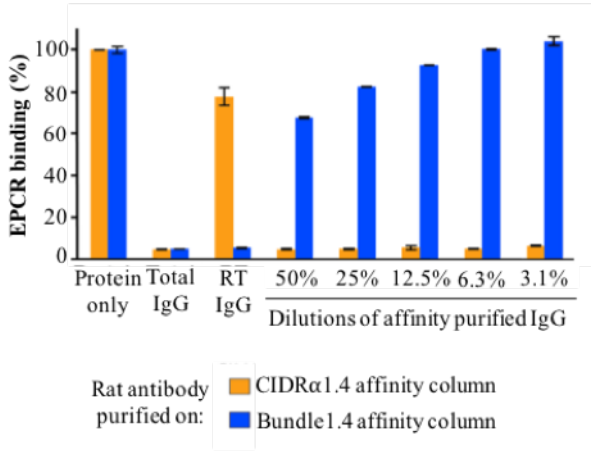


Figure 3.28: Binding of Bundle1.4 to Antibodies from Immunised Rats Antibodies produced from rats immunised with CIDRα1.4 domains were either affinity purified on a CIDRα1.4 column or a Bundle1.4 column. The different purification fractions were then assessed for the ability to prevent the CIDRα1.4 binding to EPCR by ELISA. Protein only indicates the level of EPCR binding in the presence of no antibody, which was set as a reference at 100% binding. Total IgG shows the binding of CIDRα1.4 to the EPCR in the presence of 0.55mg/ml total IgG. RT (run through) IgG shows the binding of CIDRα1.4 to EPCR in the presence of the antibodies that did not bind to the affinity column. The remaining columns show the binding of CIDRα1.4 to EPCR in the presence of the eluted and affinity purified antibodies at different dilutions.

In order to quantify this result, the total antibody concentration from the affinity purified sera was set to 25 μg/ml and the inhibitory capacity elucidated. For all IgG that was present in the sera, there was a 70% reduction in CIDRα1 binding to the EPCR **(Figure 3.29)**. In comparison, for the antibodies purified on the CIDRα1.4 affinity column, at 25 μg/ml, there was a 90% reduction in EPCR-binding by CIDRα1. However, the antibodies purified on the Bundle1.4 affinity column only resulted in a 40% decrease in CIDRα1 binding to EPCR **(Figure 3.29)**. This indicated that most of the inhibitory antibodies raised by the immunisation of rats with the whole CIDRα1 domain recognised epitopes outside the helix-kinked-helix

motif. Moreover, the antibodies that bound to these epitopes, although not directly overlapping with the EPCR-binding site, contributed to the inhibition of EPCR binding.

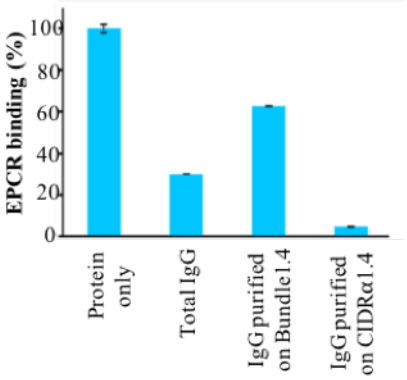


Figure 3.29: Quantification of Binding by Bundle1.4 to Antibodies from Immunised Rats Quantification of the experiment shown in **Figure 3.28**, in which each fraction of antibodies was quantified at 25 μg/ml. Again in the presence of no antibodies (protein alone) the CIDRα1.4 binding to EPCR was set to 100% and used as a reference.

A similar experiment was also carried out using antibodies purified from the sera of semi-immune humans from a malaria endemic area in Tanzania. These antibodies represented those that appear from natural acquired immunity rather than those produced in response to immunisations. The human IgG was purified using a protein G sepharose column and antibodies selected by the ability to inhibit HB3var03 binding to EPCR by an enzyme-linked immunosorbent assay (ELISA). This purified antibody pool was then tested for its ability to prevent the binding of five different subclasses of CIDRα1 domain to EPCR. The pool of antibodies had varying ability to inhibit the different CIDRα1 subclasses binding to EPCR **(Table 3.10** and **Figure 3.30)**.

CIDRα1 Subclass	CIDRα1 Name	Inhibition of EPCR Binding by Total IgG (%)
CIDRα1.1	IT4var20	50
CIDRα1.4	HB3var03	>90
CIDRα1.5	1965_2	0
CIDRα1.7	IT4var22	50
CIDRα1.8	1702_3	20

Table 3.10: Ability of Natural Antibody Pool to Inhibit EPCR Binding to Different CIDRα1 Domains

In a second experiment, antibodies from this human serum pool were purified on the Bundle1.4 affinity column and again tested against the different CIDRα1 subclasses for EPCR binding. The antibodies eluted from the Bundle1.4 column retained around 40% inhibition of CIDRα1.1 binding to EPCR compared to the original antibody mixture **(Figure 3.30)**. However, for the other CIDRα1 subclasses the vast majority of the inhibitory antibodies were found in the column run through. Therefore, again it was seen that these naturally acquired inhibitory antibodies, like those produced upon immunisation, did not solely recognise the helix-kink-helix motif of the EPCR-binding site. Nevertheless, antibodies that did bind to this motif in the Bundle1.4 appeared to block EPCR binding as expected. This indicated there are additional inhibitory epitopes that block EPCR binding but do not bind directly to the central region of the EPCR-binding site.

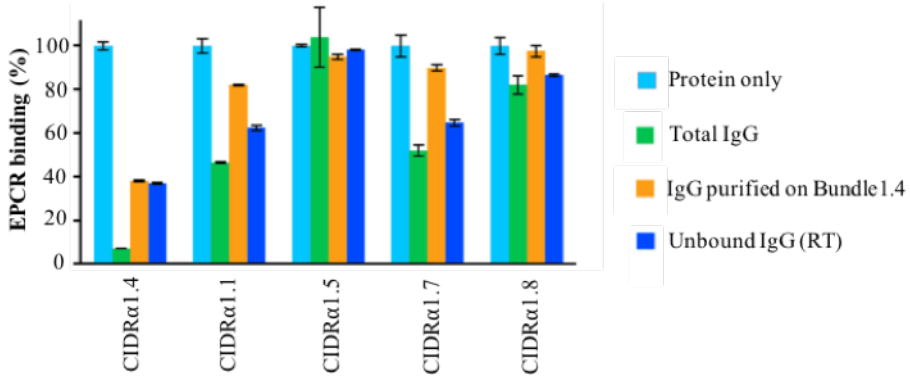


Figure 3.30: Binding by Bundle1.4 to Antibodies from Humans in Malaria Endemic Area Sera from individuals in a malaria endemic region were tested for the ability to inhibit the binding of different CIDRα1.4 subclasses to EPCR. The total antibody pool, antibodies purified on the Bundle1.4 and those that did not bind to the Bundle1.4 column (RT) were all tested and represented as a fraction of the binding in the presence of no antibody (protein alone).

3.2.11.2 Analysis of Sera from Immunisations with Bundles

In order to assess the ability of the Bundles to raise antibodies that inhibited the binding of EPCR by CIDRα1 domains, rats were immunised with the Bundle1.4.

Looking at the total pool of antibodies produced, these reduced the binding of CIDRα1.4 (HB3var03) to EPCR by 40% compared to the binding seen when no antibodies were present (protein only) **(Figure 3.31 A)**. Again the antibodies were purified on both the CIDRα1.4 column and the Bundle1.4 column. The antibodies bound to both columns (which both contained the helix-kink-helix motif) with very little inhibitory activity in the run through **(Figure 3.31 A)**. This was quantified by setting the total IgG from the Bundle1.4 immunisations to 25 μg/ml, and compared to 25 μg/ml total IgG from the CIDRα1.4 immunisations, there was reduced inhibition of EPCR binding by $\sim$50% **(Figure 3.31 B)**. This highlights that the proportion of EPCR-blocking antibodies in the total IgG produced from immunisation with the Bundle1.4 was lower than the proportion produced from immunisation with the CIDRα1.4.

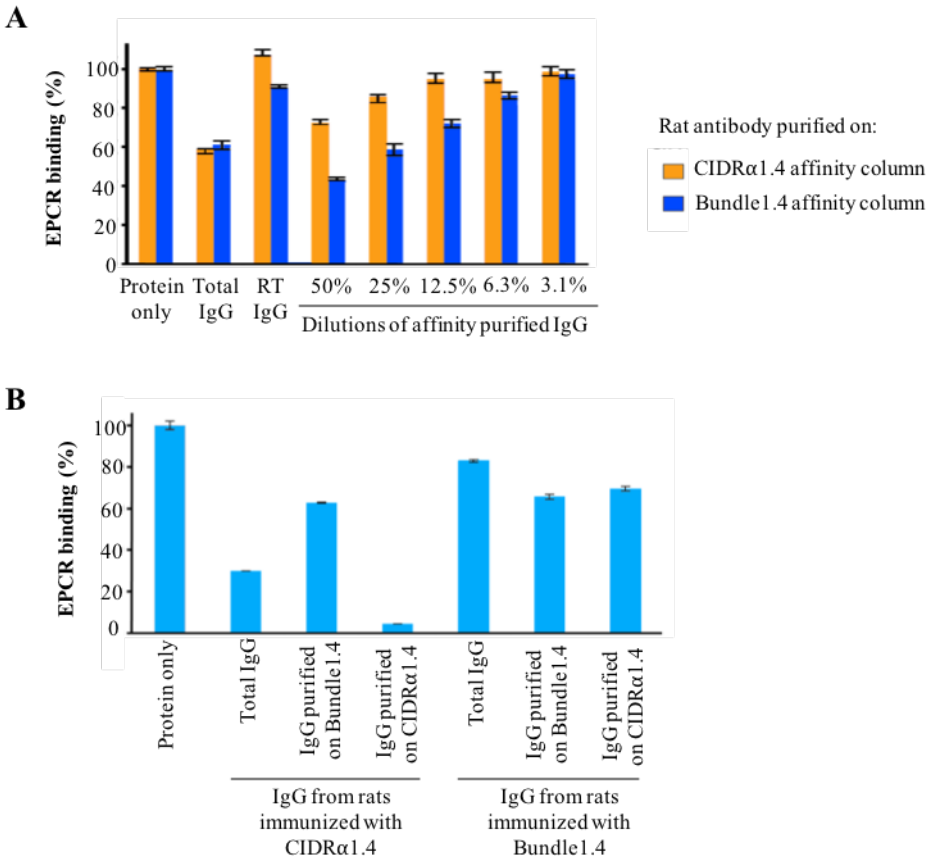


Figure 3.31: Immunogenicity of Bundle1.4 A) Antibodies from rats immunised with Bundle1.4 were affinity purified using either a CIDRα1.4 (HB3var03) column or a Bundle1.4 column. The purified antibodies were assessed for their ability to prevent CIDRα1.4 binding to EPCR. Protein alone shows the level of binding in the absence of antibody and was set as a reference at 100%. The total IgG showed the percentage of binding in the presence of 0.4 mg/ml total IgG. RT IgG shows the EPCR binding in the presence of the antibodies that did not bind to the affinity column. The remaining columns show EPCR binding in the presence of affinity purified antibodies at different dilutions. B) Prevention of CIDRα1.4 binding to EPCR was quantified at 25 μg/ml. Antibodies were from rats immunised either with CIDRα1.4 or Bundle1.4. The EPCR binding in the presence of total IgG or of IgG affinity purified on the Bundle1.4 or the CIDRα1.4 was expressed as a fraction of the binding in the absence of antibody (protein alone).

One of the aims of making the Bundles was to produce cross-reactive antibodies that blocked EPCR-binding to multiple CIDRα1 subclasses. It was thought that the bundles, by only containing the highly conserved helix-kink-helix motif would result in the production antibodies with a higher cross-reactivity than immunisations

with the whole CIDRα1 domain. This was assessed by testing the ability of the antibodies produced from immunisations with Bundle1.4 to bind to multiple different CIDRα1 domains in a luminex assay. As demonstrated in **Figure 3.4**, there is great sequence diversity of EPCR-binding CIDRα1 domains, with only 6% of residues completely conserved. In order to represent this sequence variation, between four and eight constructs for each subclass of CIDRα1 (that were the most phylogenetically distinct within their subclass) were chosen for the CIDRα1 panel **(Figure 3.32)**. The antibodies, produced from Bundle1.4 immunisations, did recognise the cognate CIDRα1.4 domain, HB3var03 (from which the Bundle1.4 was designed) as well as four other CIDRα1.4 domains **(Figure 3.32 top panel)**. However, the antibodies did not bind to any other subclasses of CIDRα domains. The CIDRα1.4 domains, to which the IgG did bind, showed EPCR binding inhibition of <40% **(Figure 3.32 bottom panel)**. As a result, since Bundle1.4 only produced sera that recognised the CIDRα1.4 subclass, it did not appear that upon immunisation Bundles produced more cross-reactive antibodies than the CIDRα1 domains.

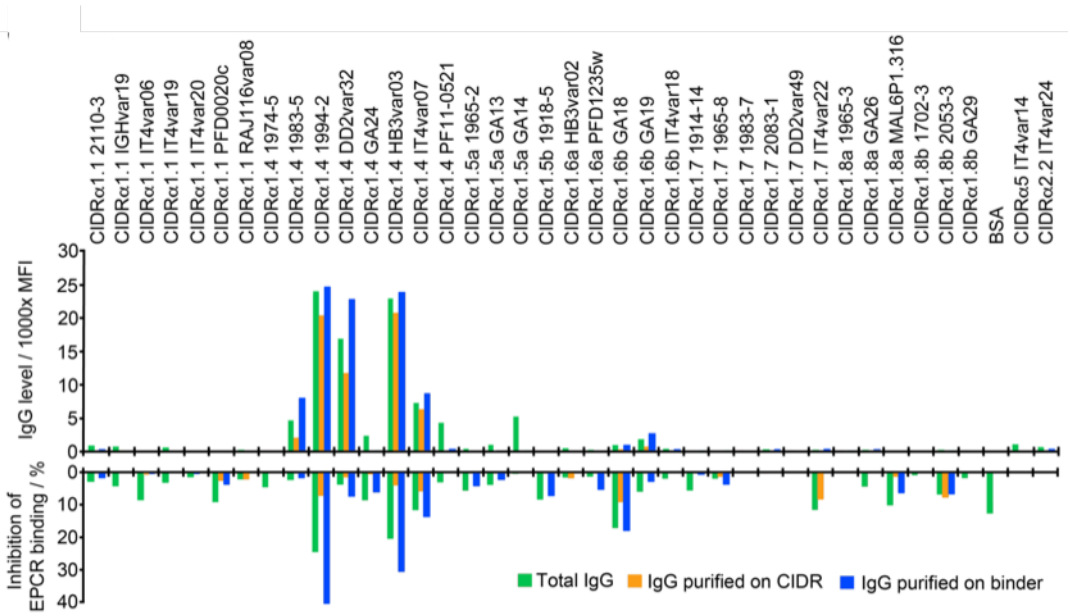


Figure 3.32: Cross-reactivity of Antibodies Generated from Immunisation with Bundle1.4 In this luminex assay, Antibodies from rat immunisations with Bundle1.4 was tested against a panel of CIDRαs, either as total IgG, or after affinity purification on the Bundle1.4 or CIDRα1.4 (HB3var03) column. The CIDRα2/5 domains are controls not expected to bind EPCR. The upper panel shows IgG binding levels (in mean fluorescence intensity, MFI). The lower panel shows the inhibition of the binding of these CIDRα domains to EPCR.

3.2.11.3 The Use of VLPs in Immunisations

In other projects in the laboratory as well as in the literature [220], it has been seen that immunisation of animals with small immunogens often results in a lower overall titre of antibodies produced. Since the Bundles were 18kDa compared to the 26kDa of CIDRα1 domains, it was thought that this might explain the decrease in immunogenicity. Attaching multiple immunogens to larger particles like viral-like particles (VLPs) have previously been used to increase immunogenicity [221–223]. Therefore, in an effort to boost the overall immune response, it was decided to couple the bundles to a VLP. Furthermore, since the individual Bundle1.4 did not appear to be cross-reactive, multiple different bundle subclasses could be coupled to a single VLP in order to improve the cross-reactivity of the antibodies produced.

3.2.11.3.1 SpyCatcher:SpyTag Coupling Immunogens can be coupled to VLPs by using a SpyCatcher:SpyTag coupling system [224]. The SpyTag is a 13 residue fragment that can form an isopeptide bond to covalently attach itself to the SpyCatcher complementary domain **(Figure 3.33 A,B, and C)**. Therefore if an immunogen of choice is coupled to a SpyTag and the VLP is expressed with exposed SpyCatcher domains, the SpyCatcher-VLPs and SpyTag-immunogen interaction allows VLPs to be arrayed with an immunogen of choice **(Figure 3.33 D)**.

In order to make use of this coupling system, all of the Bundles were expressed with a SpyTag inserted at their C-terminus to allow subsequent interaction with an Acinetobacter bacteriophage VLP which displayed one N-terminal SpyCatcher per capsid subunit [225]. Assembled VLPs were quality assessed by Dynamic Light Scattering (DLS).

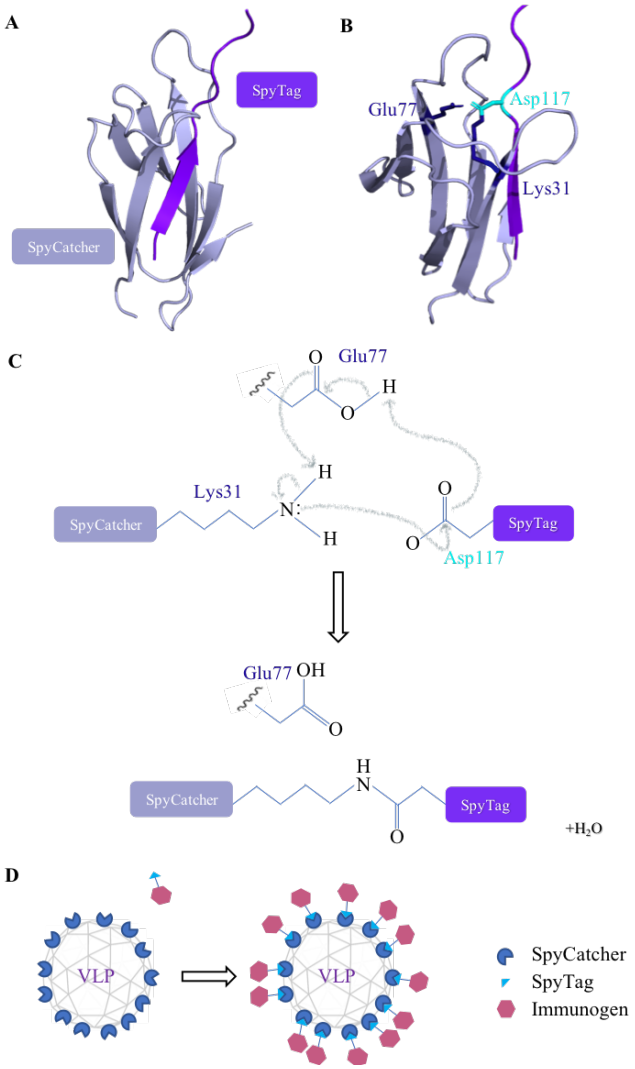


Figure 3.33: SpyCatcher:SpyTag Coupling A) The reformation of the β-sandwich by the SpyTag (dark purple) and the SpyCatcher (light purple). B) The residues involved in the covalent linking of the SpyTag and SpyCatcher are highlighted. C) Reaction meachanism of the peptide bond foramtion. D) VLPs made up of particles containing SpyCatcher domains can be used to covalently attach multiple immunogens with SpyTags to one particle.

3.2.11.3.2 Immunisations with VLP-Bundles The Bundle1.4 was coupled to SpyCatcher-VLPs via its SpyTag, and the ability of the sera from the Bundle1.4-VLP immunisations to bind to the panel of CIDRα1 domains, as well as their ability to inhibit binding of these CIDRα1s to EPCR was tested. This could be directly compared to the binding and inhibition demonstrated by sera from the Bundle1.4 not coupled to VLPs **(Figure 3.34)**. As with the "no VLP" construct, immunisations

with the Bundle1.4-VLP resulted in the production of antibodies that recognised the cognate CIDRα1.4 and inhibited its binding to EPCR. However, the sera from the Bundle1.4-VLP did not convincingly bind to other CIDRα1.4 domains, nor to any of the other subclasses of CIDRα1 **(Figure 3.34)**. Therefore, immunisations with the Bundle1.4 attached to a VLP did not improve the cross-reactivity of the sera produced.

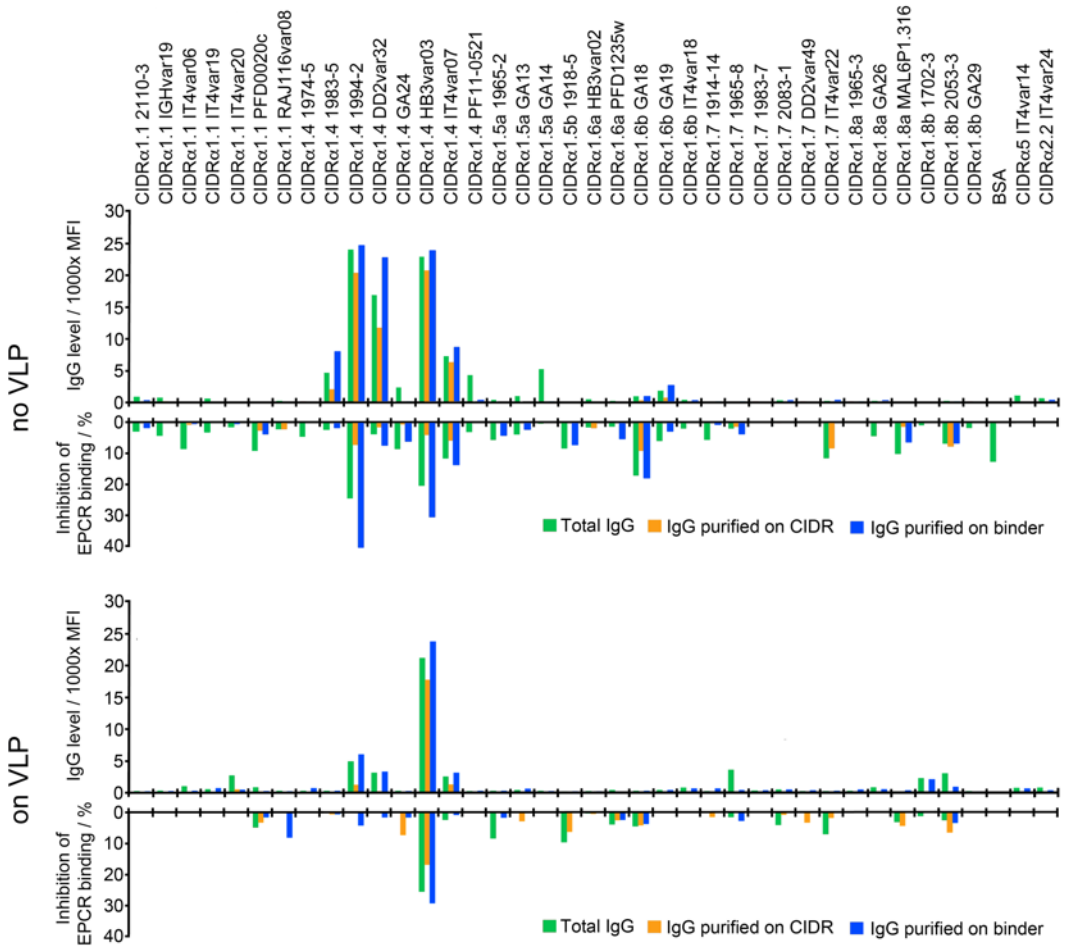


Figure 3.34: Immunogenicity Cross-reactivity of Antibodies Generated from Immunisation with Bundle1.4-VLP In this luminex assay, antibodies from rat immunisations with Bundle1.4 both off (labelled "no VLP") and on VLP (labelled "on VLP") was tested against a panel of CIDRαs, either as total IgG, or after affinity purification on the Bundle1.4 or CIDRα1.4 (HB3var03) column. The CIDRα2/5 domains are controls not expected to bind EPCR. The upper panel shows IgG binding levels (in mean fluorescence intensity, MFI). The lower panel shows the inhibition of the binding of these CIDRα domains to EPCR.

In an attempt to increase the CIDRα1 binding response and cross-reactivity of the sera produced, all the Bundles were coupled to VLP particles by the SpyTag-SpyCatcher system and then mixed together. Four rats were immunised with this mixture of Bundle-VLPs, the sera collected and IgG purified on protein G sepharose. The total pool of IgG blocked the binding of CIDRα1.4 (HB3var03) to EPCR by 40% compared to the binding seen when no antibodies were present **(Figure 3.35)**. Again antibodies were purified on both the CIDRα1.4 column and the Bundle1.4 column. The antibodies bound to both columns (which both contained the helix-kink-helix motif) with very little inhibitory activity in the run through **(Figure 3.35)**. For immunisations with the mixture of the Bundles on VLPs, the sera produced roughly the same amount of inhibition as the sera produced from immunisations with just Bundle1.4 ($\sim$40% inhibition of EPCR binding), however the amount of IgG produced was not quantified.

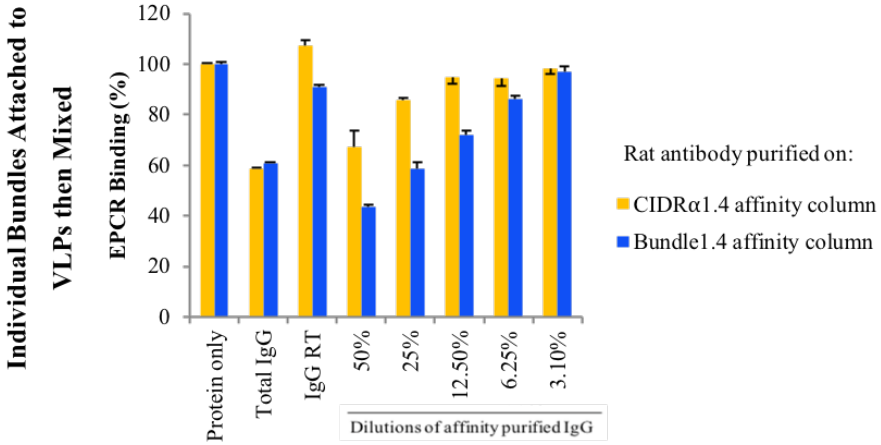


Figure 3.35: Immunogenicity of a Mixture of Bundles Attached to VLPs IgG from rats immunised with a mixture of Bundles attached to VLPs were affinity purified using either a CIDRα1.4 (HB3var03) column or a Bundle1.4 column. The purified antibodies were assessed for their ability to prevent CIDRα1.4 binding to EPCR. Protein alone shows the level of binding in the absence of antibody and was set as a reference at 100%. The total IgG shows the percentage of binding in the presence of 0.4 mg/ml total IgG. RT IgG shows the EPCR binding in the presence of the antibodies that did not bind to the affinity column. The remaining columns show EPCR binding in the presence of affinity purified antibodies at different dilutions.

As a final experiment, the cross-reactivity of the IgG produced from immunisations with a mixture of all of the Bundles both on and off VLPs was assessed **(Figure 3.36)**. The purified antibodies from the sera were tested for binding to a variety of different subclasses of CIDRα1 domains. Again as seen with the immunisations with the individual Bundle1.4 off the VLP **(Figure 3.34)**, binding to the cognate HB3var03 CIDRα1.4 (as well as two other CIDRα1.4 domains) was observed **(Figure 3.36 A)**. Unlike immunisations with Bundle1.4-VLP, the coupling of the mixture of Bundles to VLPs raised sera that interacted with three additional CIDRα1.4 domains compared with the sera from immunisations off VLPs **(Figure 3.36 A)**. In this experiment, there was binding to a greater range of CIDRα1.4 domains "on-VLP" compared with "off-VLP", and the overall IgG level of binding was higher, indicating more binding antibodies were present. However, for both mixtures of Bundles, on and off VLP, again very little cross-reactivity was seen to the other subclasses of CIDRα1s, despite the multiple bundle subclasses being used in the immunisations. This suggests that the epitope presented by the Bundles (designed on a CIDRα1.4 binding site) may be more relevant to CIDRα1.4 protein than the other subclasses. Therefore, despite the introduction of mutations to represent the different CIDRα1 subclass sequences at the EPCR-binding site, the bundle only promoted the production of CIDRα1.4 reactive antibodies.

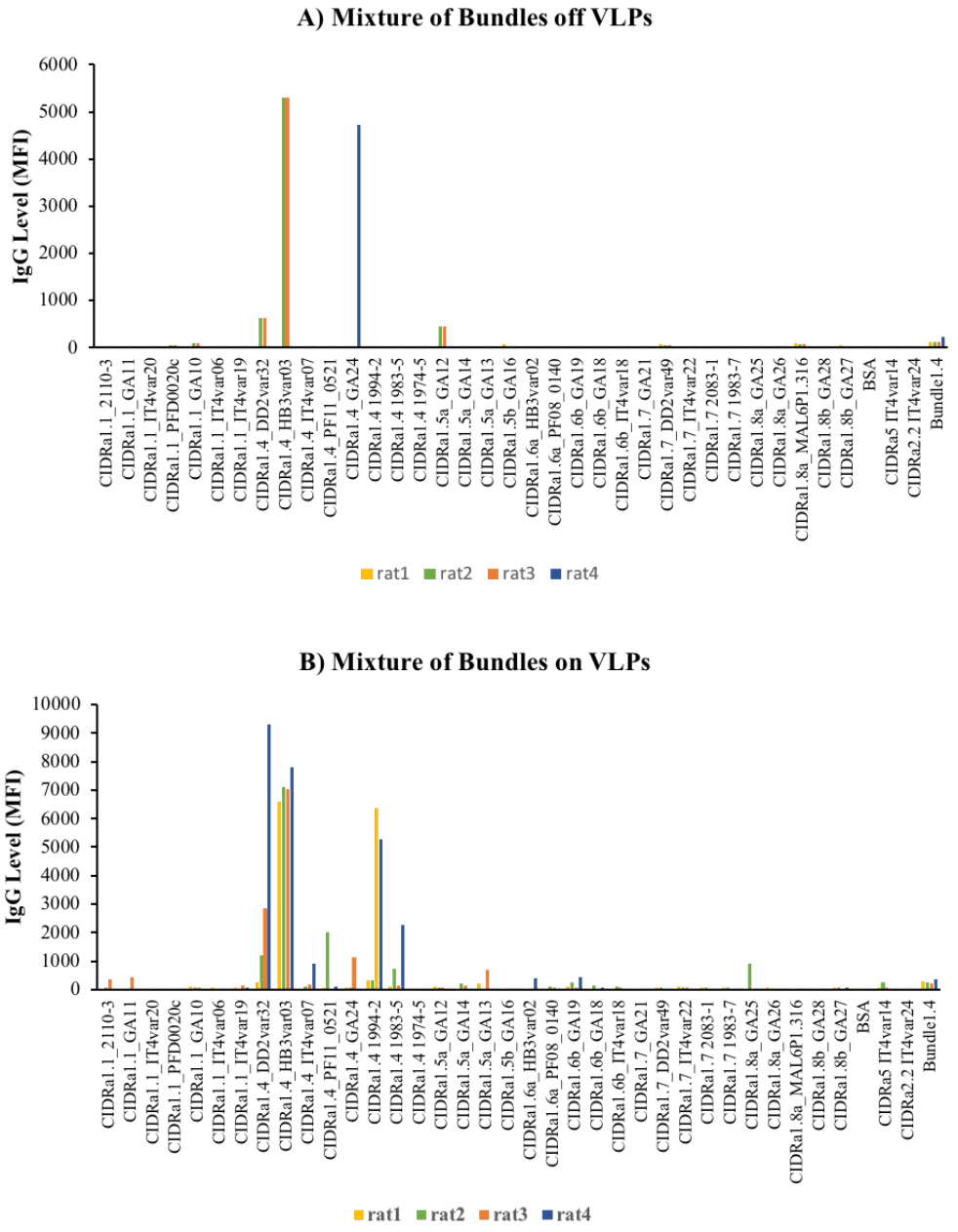


Figure 3.36: Cross-Reactivity from Immunisations with all Bundles attached to VLP Total IgG from rat immunisations with all Bundles attached to a VLP tested for binding against a panel of CIDRαs is shown. The CIDRα2/5 domains are controls not expected to bind EPCR. The graph shows IgG binding levels (in mean fluorescence intensity, MFI), with the bars representing the four rat sera tested against each CIDRα domain.

3.2.12 Discussion

In recent years, new vaccine design protocols have used structural methods to understand how protective, inhibitory antibodies bind and to then design synthetic proteins that mimic the structural feature of these protective epitopes [138, 217]. This has been successfully used to design novel immunogens, which upon immunisation produce broadly inhibitory antibodies [135, 137, 218]. In the absence of structural information about epitopes of inhibitory antibodies against the CIDRα1:EPCR interaction, an adaptation of this method was used in which the EPCR binding site was mimicked. Structural and sequence analysis of the EPCR-binding site indicated that it was chemically conserved across the different subclasses of EPCR-binding CIDRα1s. This feature, as well as being a site of functional importance, made this an attractive site to target. The structural helix-kink-helix motif of the EPCR-binding site was grafted onto a three-helical scaffold, which was remodelled using a Rosetta-based approach to produce a small Bundle that mimicked the EPCR binding site. In an attempt to increase the likelihood of production of cross-reactive antibodies, a panel of the six Bundles was created to represent the sequence-diverse subclasses of CIDRα1s that bound to EPCR.

The design and production of the Bundles were successful. The CD data highlighted that the insoluble Bundles were refolded correctly and were stable at RT. The SPR data showed that four of these Bundles bound to EPCR with a high nanomolar affinity and slow off-rate, characteristic of the CIDRα1:EPCR interaction [205]. Moreover, the crystal structures of these Bundles bound to EPCR showed that the conformation was similar to that of the Rosetta-predicted design and in addition very closely matched the binding by the cognate CIDRα1.4 domain [226]. Therefore, for these four Bundles, synthetic proteins had been created that successfully mimicked the core EPCR-binding site of an EPCR-binding *Pf*EMP1.

The Bundle1.7 and Bundle1.8 did not bind to EPCR either by SPR or SEC and so in these cases an EPCR-binding site was not successfully recapitulated on the synthetic protein. It was unclear why these two Bundle mutations did not work. The original Bundle1.4 was designed on the CIDRα1.4 structure and sequence, whereas the other Bundles just had five or six point mutations introduced in the EPCR-binding motif to represent the sequence diversity. Despite only having structures of the CIDRα1.4 subclass bound to EPCR, due to the conserved chemistry of the residues in the binding site, it was assumed that the mechanism of binding to EPCR would be the same for the other CIDRα1s. This indeed seemed to be the case for Bundle1.1, Bundle1.5, and Bundle1.6. However, perhaps for the Bundle1.7 and Bundle1.8 other residues outside this motif were involved in the binding to EPCR. Looking at the mutations introduced into Bundle1.7 and Bundle1.8, six residues had been mutated in both of these constructs compared with Bundle1.4 **(Figure 3.37)**. For Bundle1.8, only one of these mutations was shared with another Bundle that bound to EPCR, and so the introduction of the other five residues may have disrupted the direct interactions. Specifically, the introduction of three glycines and one serine in the place of residues with longer side chains would make these hydrogen bonding groups too far away from the EPCR residues involved in Bundle1.4 binding.

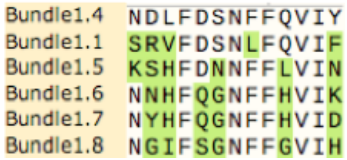


Figure 3.37: EPCR-Binding Site Sequence for the Different Bundles The sequences for the bundle residues 72-84 are shown and the mutations made compared to the Bundle1.4 sequence are highlighted in green.

For Bundle1.7, four out of six of these point mutations were shared with at least one other Bundle that did show binding to EPCR **(Figure 3.37)**. The other two mutations, D73Y and Y84D, may have disrupted crucial interactions with residues in EPCR. For example, the introduction of the Tyr at position Asp73 may cause steric

clashes with the EPCR Ser71, and the also Y84D mutation, may have prevented the hydrogen bonding to the Gln149 and Arg87 residues **(Figure 3.38 A)**. However, in Bundle1.5, the introduction of Asn84 made a different hydrogen bond to the EPCR Leu82, which could also be adopted by the similar sidechain of Asp84 introduced in Bundle1.7 **(Figure 3.38 B)**.

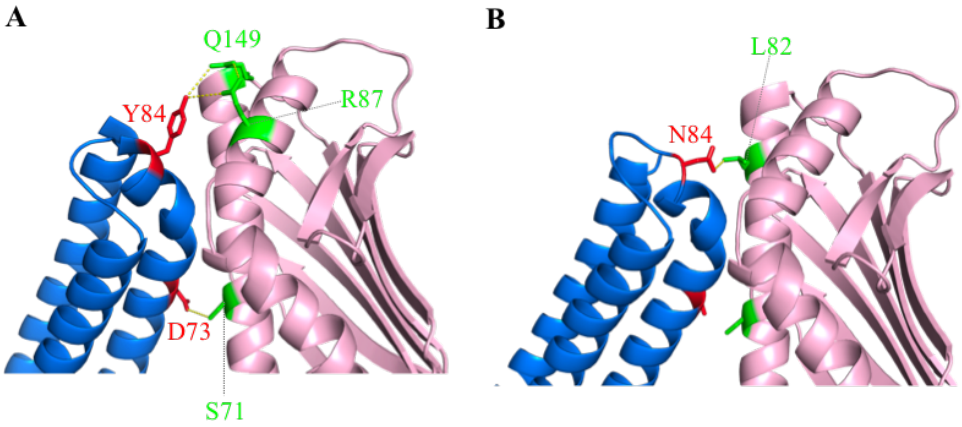


Figure 3.38: Interactions with EPCR of Residues Mutated in Bundle1.7 A) Directly contacting residues in Bundle1.4:EPCR structure, with the Bundle residues labelled in red and the EPCR residues labelled in green. B) The hydrogen bond formed between N84 in Bundle1.5 (highlighted in red) and L82 of EPCR (highlighted in green) is shown.

Although looking at the structures of the other Bundles bound to EPCR can give us some clues as to why Bundle1.7 and Bundle1.8 did not bind to EPCR, it is unclear why these two Bundle mutations did not work. In order to answer this question fully and gain greater understanding of CIDRα1 binding to EPCR, structures of some of the other CIDRα1 subclasses bound to EPCR are needed. These structures would answer the question as to whether the CIDRα1 domains all bind to EPCR via the same conserved helix-kinked-helix interface or whether other regions of the domain are involved in some of the subclasses. To probe this, initial crystallisation screens were set up for CIDRα1.7 and CIDRα1.8 in complex with EPCR, both methylated and non-methylated, however all attempts to produce crystals as of yet have been unsuccessful.

Interaction studies of Bundles from the sera produced from immunisations with the cognate CIDRα1 domains, showed that the Bundles did bind to some of the antibodies, and that some of these antibodies were inhibitory to EPCR binding. However, these antibodies were less effective in both quantity and quality than the antibodies that bind to the CIDRα1 domain. Moreover, comparison of antibodies produced by immunisation of rats with the CIDRα1.4 domain and the Bundle1.4 showed that although the Bundle did generate antibodies that bound to its cognate CIDRα1 domain and inhibited EPCR, these antibodies were again less effective and less abundant than the equivalent antibodies raised through immunisation with the CIDRα1 domain. Furthermore, the antibodies produced by immunisation with the Bundle showed no more cross-reactivity than those raised by the CIDRα1 domains. In addition, coupling the Bundles, alone or in combination, to VLPs, made no improvement to either abundance of antibodies produced nor in improving the cross-reactivity of the antibodies produced. Indeed, immunisations with the mixture of Bundles, both on and off VLPs, only resulted in production of CIDRα1.4 reactive antibodies. This could be because either the CIDRα1.4 reactive antibodies were easier to raise or because the design process was based on the EPCR-binding site of CIDRα1.4 and despite mutations being introduced to represent sequence diversity of the other CIDRα1 subclasses, the Bundle binding site was more relevant to the CIDRα1.4 than the other subclasses. This observation also suggests that the other subclasses of CIDRα1 may bind to EPCR via a slightly different mode than CIDRα1.4, and the structures of different CIDRα1:EPCR complexes would be helpful.

Overall, there is no evidence that focusing the immune response to the helix-kinked-helix functional binding site generates more cross-reactive and cross-inhibitory antibodies [226]. These findings indicate that many of the epitopes for the antibodies that inhibit EPCR-binding to CIDRα1 domains are not solely located at the EPCR-

binding helix-kink-helix site. Therefore, targeting the functional binding site of the protein in vaccine design may not be the best approach. In contrast, previously it has been shown that knowledge of the structure of the epitope of an antibody known to be produced through immunisation is a valid starting point. In this method the whole epitope is recapitulated and it is also possible to elicit such an antibody from the human germline. Therefore, for future vaccine design against the *Pf*EMP1:EPCR interaction, the molecular and structural features of an epitope of an inhibitory monoclonal antibody is needed. This could then be used as a starting point for future structure-guided design efforts.

3.3 Structural Studies of Mouse Monoclonal Antibodies Raised Against CIDRα1s To Guide Future Immunogen Design

3.3.1 Project Aim

- Structurally characterise the epitope of an inhibitory mouse monoclonal antibody to a CIDRα1 domain by X-ray crystallography

3.3.2 Declaration

The immunisations of mice with CIDRα1 domains was carried out by our collaborators, the Lavstsen Group at the University of Copenhagen. They also provided us with the genomic CIDRα1 DNA and furthermore characterised and isolated the inhibitory mouse monoclonal antibodies.

Two of the CIDRα1 constructs had already been cloned by a former DPhil student in the lab, Clinton Lau. Clinton also had previously optimised expression and purification protocols for CIDRα1 domains. An undergraduate student, Mateusz Kotowski, under my supervision designed, cloned and expressed the CIDR$_{-\beta,-loop}$ constructs.

3.3.3 Monoclonal Antibody Fab Fragment Generation

Monoclonal mouse antibodies were obtained from our collaborators, who used full-length CIDRα1 domains of different subclasses to separately immunise mice to raise antibodies. The sera from these mice were collected, IgG purified and assessed using ELISA/Luminex assays for its ability to bind to the CIDRα1 domain used for immunisation, as well as inhibition to EPCR binding. B cell sorting was then used to identify the mAb that inhibited the binding of each of the five subclasses of CIDRα1 to EPCR **(Table 3.11)**. Although an inhibitory mAb was not isolated from the mice immunised with CIDRα1.5 (GA12), it was discovered that the B4 mAb raised by CIDRα1.7, also bound to the GA12 CIDRα1.5.

Mouse Immunised With		**mAb Name**
CIDRα1.1	IT4var20	Jan
CIDRα1.4	HB3var03	Anker
CIDRα1.6	PF08	Sixten
CIDRα1.7	GAO20	B4
CIDRα1.8	MAL6P	Otto

Table 3.11: Name of Inhibitory Mouse mAb from Immunisations with Different CIDRα1s

Once the full-length mAbs had been received, their isotypes were identified as IgG1 **(Materials and Methods 2.5.1)**. As a result, in order to obtain the Fab fragments for structural studies, they were cleaved with Ficin protease overnight at 37°C **(Materials and Methods 2.5.2)**. Since Ficin resulted in the digestion of the Fc region of the mAb, no further purification step was needed to separate Fc/uncleaved mAb from the Fab fragments **(Figure 3.39)**.

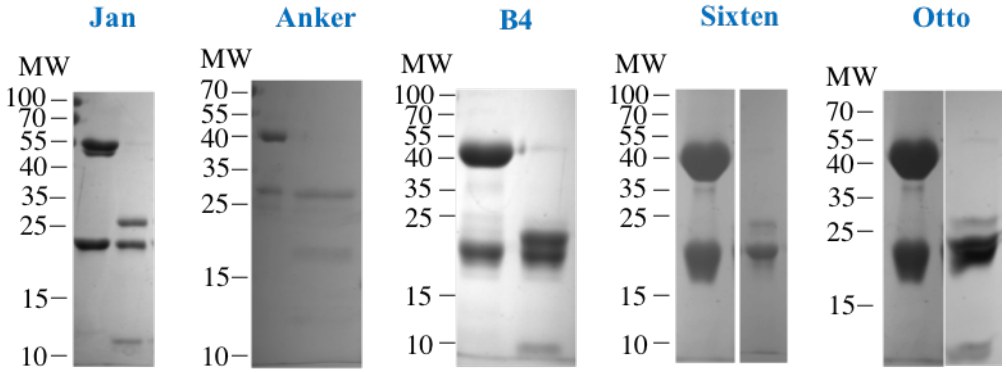


Figure 3.39: Ficin Cleavage of Mouse mAbs Reduced SDS-PAGE gels showing the ficin cleavage of the mouse mAbs. The left lane for each construct is the pre-cleaved mAb and the right lane for each construct, the post-ficin Fab fragment.

3.3.4 CIDRα1 Cloning, Expression and Purification

The six CIDRα1 domains that were used for the immunsiations **(Table 3.12)** were cloned from the genomic DNA (provided by the Lavstsen group at the University of Copenhagen) into pET15bcon vector **(Materials and Methods Figure 2.1)**. The sequences and the primers are specified in **(Appendix A.2.1 and A.2.2)**. As previously mentioned this vector contained an N-terminal His-tag for purification that could be removed by TEV protease cleavage.

CIDRα1.1	IT4var20
CIDRα1.4	HB3var03
CIDRα1.5	GA12
CIDRα1.6	PF08
CIDRα1.7	GAO20
CIDRα1.8	MAL6P

Table 3.12: Name of construct chosen to represent each CIDRα1 Subtype

CIDRα1 constructs for HB3var03 (CIDRα1.4) and IT4var20 (CIDRα1.1) had already been cloned by a previous DPhil student, Clinton Lau, into pET15bcon. Since the

HB3var03 CIDRα1 construct had successfully crystallised in complex with EPCR [205], the other CIDRα1 constructs were designed with similar predicted domain boundaries. The alignment of the shorter sequences for the CIDRα1 constructs can be seen in **Figure 3.40**, with the secondary structures of solved CIDRα domains also shown.

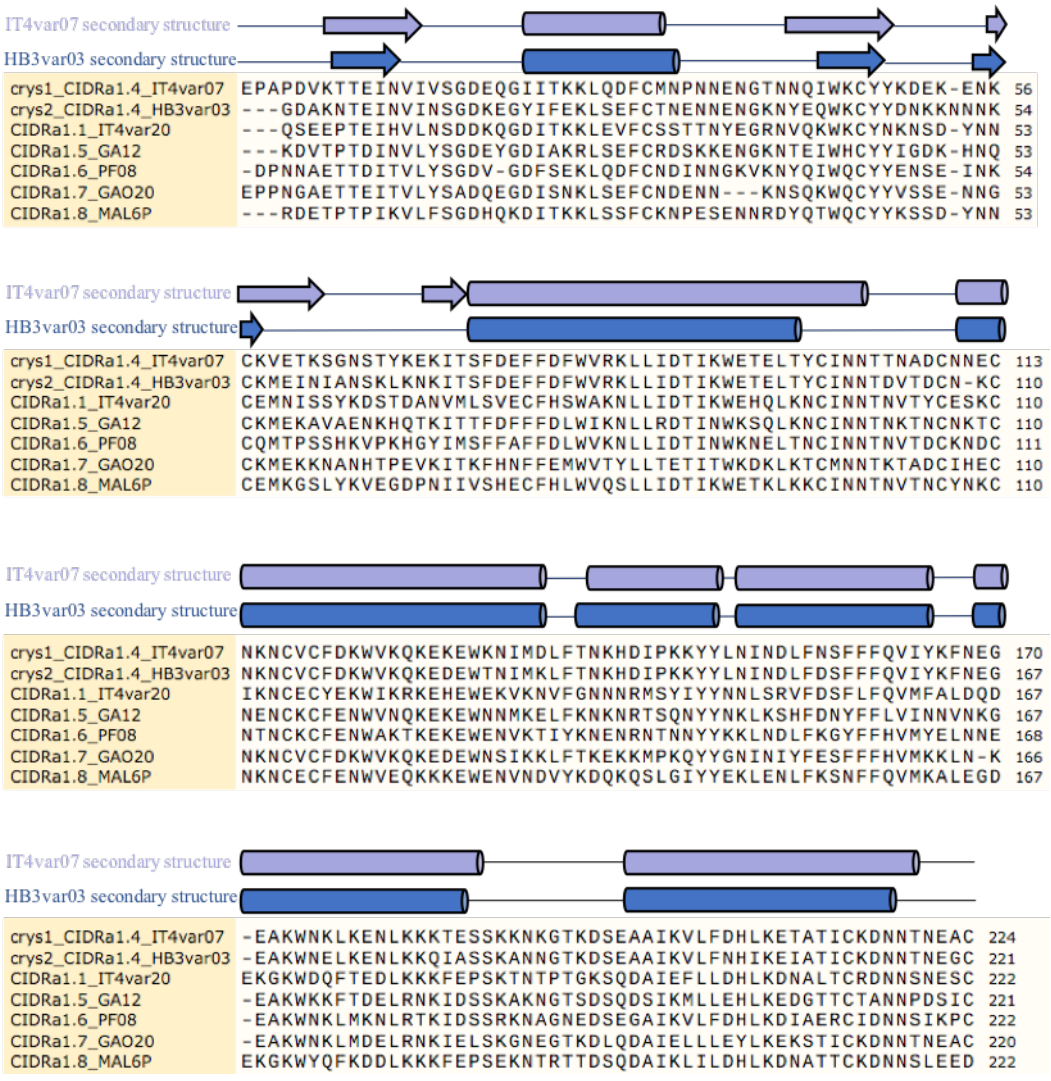


Figure 3.40: An Alignment of the CIDRα1 Sequences Compared to already Crystallised Constructs

The expression and purification of multiple CIDRα1 domains had already been optimised by Clinton Lau and so the same general protocol was followed as outlined below.

The constructs were grown up in 2xYT media until an OD_{600}=0.8 was reached and expression was induced by the addition of 1 mM IPTG at 25°C overnight. As expected, all the CIDRα1 domains expressed solely in the insoluble inclusion bodies of the cell pellet. The inclusion bodies were solubilised using guanidine hydrochloride and refolded in a step-wise decreasing gradient of guanidine hydrochloride whilst attached via their His-tag to a Ni column **(Materials and Methods 2.4.4)**. The eluted refolded protein was buffer exchanged into PBS and TEV-cleaved overnight **(Materials and Methods 2.4.5)**. To further purify and remove any aggregation the TEV-cleaved CIDRα1 was loaded onto a size exclusion column (GE Healthcare, S75 10/30, S75 16/60, or S200 10/30) **(Figure 3.41)**. The overall yield of CIDRα1 domains produced varied from 0.05 mg/L to 0.20 mg/L *E.coli* cells.

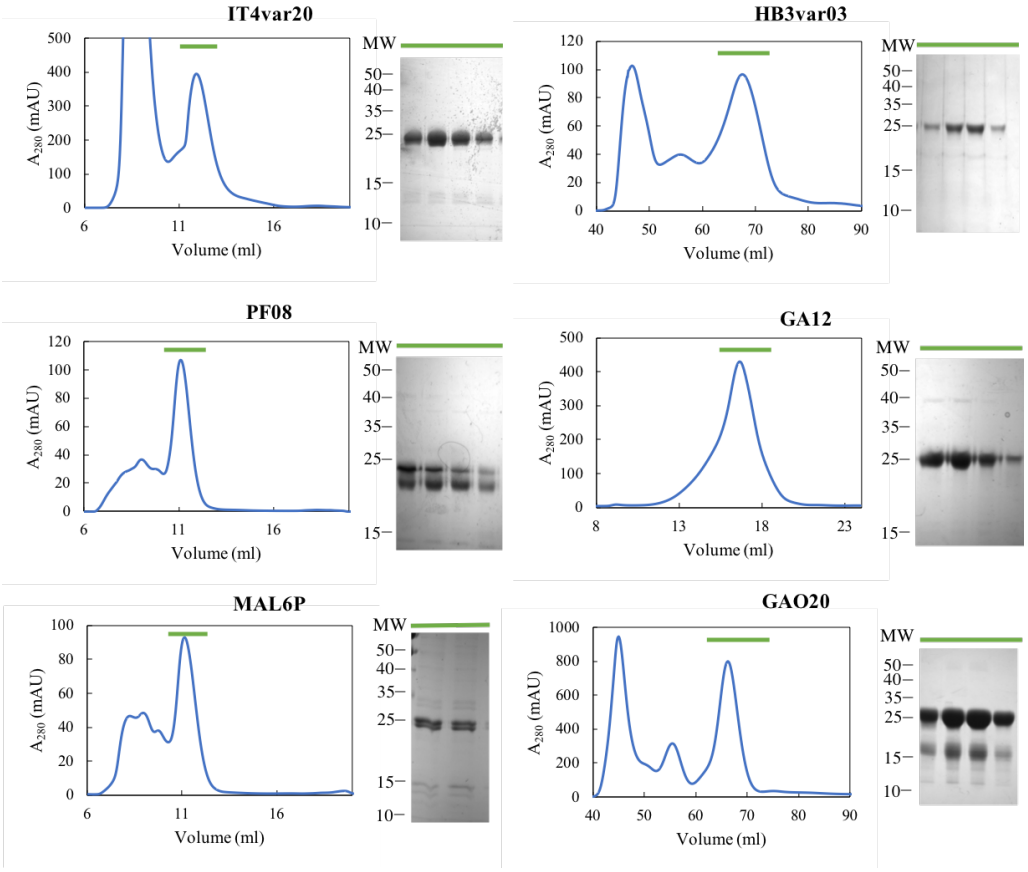


Figure 3.41: CIDRα1 Purification SEC trace (left) and corresponding non-reduced SDS-PAGE gel (right) showing the elution of the complex as one peak. The light green bar represents the corresponding regions of the SEC trace and SDS-PAGE gel, that were pooled for future experiments. IT4var20, PF08, and MAL6P were from an S75 10/30 SEC, HB3var03 and GAO20 were from an S75 16/60 SEC, and GA12 was from an S200 10/30 SEC.

3.3.5 Assessing the Fold of the CIDRα1 Domains

Since the CIDRα1 domains were all refolded from the insoluble inclusion bodies of the cell pellet, it was important to check that they were correctly refolded. This was done using CD, and all of the spectra of the CIDRα1 domains showed the characteristic α-helical trace with minima at 208nm and 222nm **(Figure 3.42)**.

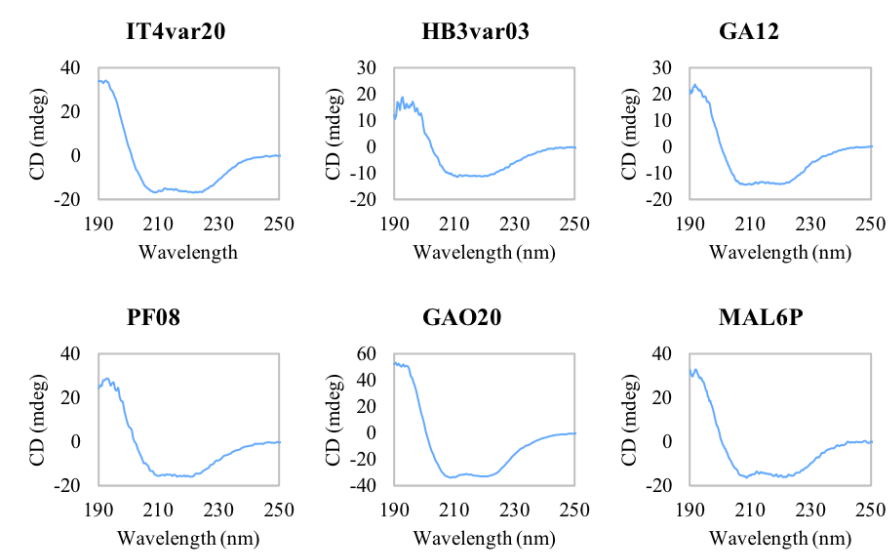


Figure 3.42: CD Spectra of CIDRα1 Domains

3.3.6 SPR of CIDRα1 Binding to EPCR

In order to confirm that the CIDRα1 domains had completely folded properly with a functional EPCR-binding site, the binding of the recombinant CIDRα1 domains to EPCR was characterised by SPR **(Figure 3.43)**. In these experiments 150 RU biotinylated EPCR was coupled to a CAPture chip and a two-fold dilution series starting at 1 μM of each CIDRα1 domain was analysed for binding **(Materials and Methods 2.6.2)**.

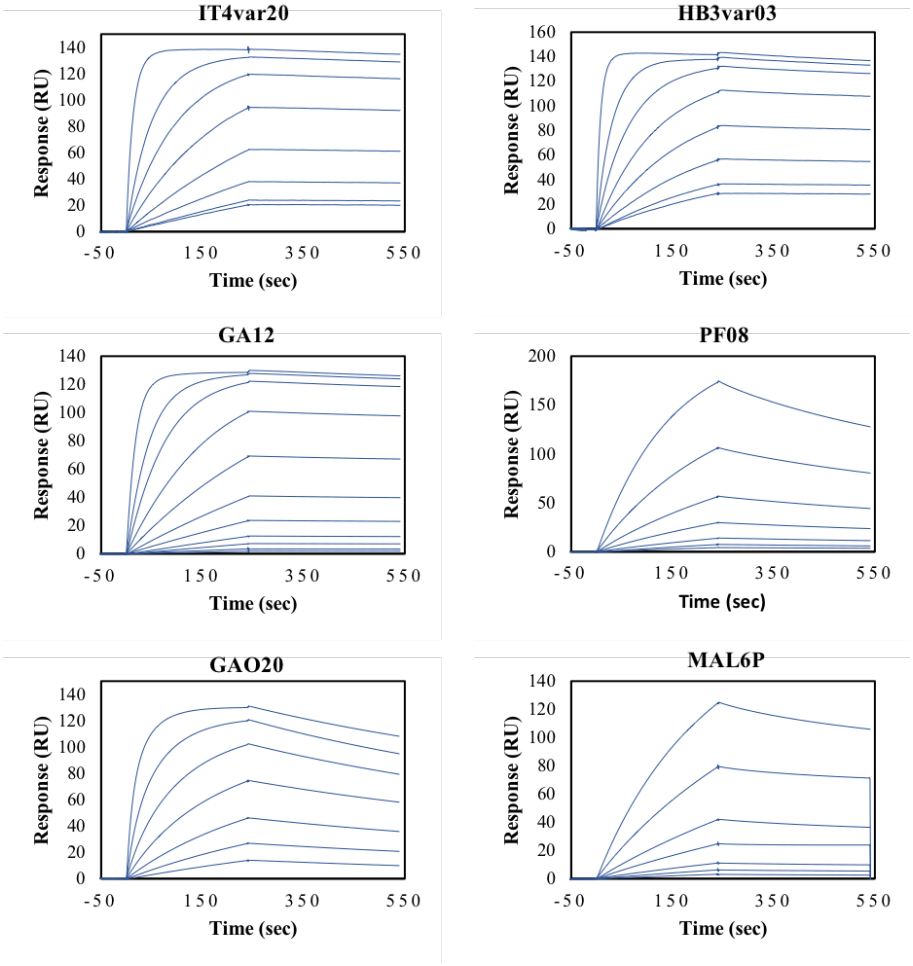


Figure 3.43: Binding of CIDRα1 Domains to EPCR Two-folded dilution SPR curves of the CIDRα1 domains starting at 1μM binding to immobilised EPCR.

All of the interaction data could be globally fit with a 1:1 interaction model using the BIAevaluation software (GE Healthcare). The CIDRα1 domains bound to EPCR with nanomolar affinity and a slow off-rate **(Table 3.13)**. The PFO8 and MAL6P had much slower on-rates (K_{on}) (up to 95x less) than the other CIDRα1 constructs, which explains the lower affinity seen for these CIDRα1 domains. The refolding of these purifications may not have been as successful in recapitulating the EPCR-binding site and so explains this slightly decreased affinity. Nevertheless, these constructs showed nanomolar affinity binding and so were proceeded with for future experiments.

	$\mathbf{k}_{on}$ ($\mathbf{M}^{-1}s^{-1}$)	$\mathbf{k}_{off}$ ($\mathbf{s}^{-1}$)	$\mathbf{K}_d$ (nM)
IT4var20	332000	1.23x10^{-4}	0.37
HB3var03	532000	1.97x10^{-4}	0.37
GA12	51600	1.16x10^{-4}	2.25
PF08	5600	1.16x10^{-3}	207
GAO20	313000	7.59x10^{-4}	24.1
MAL6P	7750	5.78x10^{-4}	74.6

Table 3.13: Kinetic Binding Analysis of CIDRα1 Domains Binding to EPCR

3.3.7 Binding of the CIDRα1 to mAbs

Having assessed the fold and function of the CIDRα1 domains, the binding of the CIDRα1 domains to their mAbs was checked **(Figure 3.44)**. In these experiments, protein A/G was irreversibly amine-coupled to a CM5 chip **(Materials and Methods 2.6.2)**. Then for each run, 10 μg/ml of the antibody was immobilised on the chip before a two-fold dilution series of the CIDRα1 domain starting at a concentration of 1 μM was flowed over. Unfortunately, there were issues in these experiments with the antibody uncoupling from the chip as the experiment progressed. As a result only curves for four out of the six CIDRα1 domains against their antibody could be obtained. Although these curves did indeed confirm binding of the CIDRα domain, as expected, no analysis of the kinetics of the interactions could be obtained as the BIAevaultion software could not fit the kinetic curves. Since binding to the antibodies could also be confirmed by co-elution in size exclusion chromatography **(Section 3.3.9)**, all of the constructs were continued with, despite this lack of SPR data.

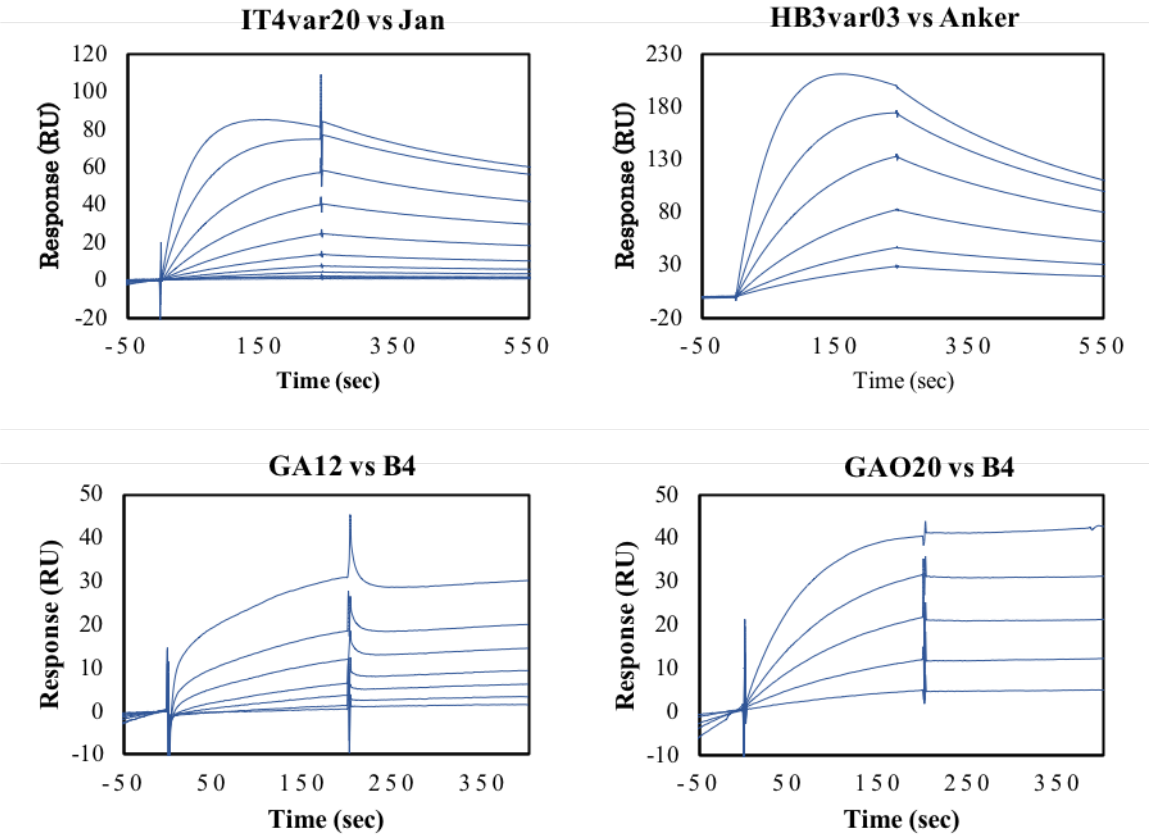


Figure 3.44: Binding of the CIDRα1 domains to Antibodies
Two-folded dilution SPR curves of the CIDRα1 domains starting at 1μM binding to immobilised mAbs.

3.3.8 Competition SPR of CIDRα1 binding to EPCR in presence of inhibitory mAb

Finally, in order to confirm that the antibodies were indeed inhibitory to EPCR-binding, an SPR competition experiment was conducted. Biotinylated EPCR was coupled to the CAPture chip, and in separate experiments the same volume of following were flowed over:

- 1 μM CIDRα1
- 1 μM antibody
- 0.5 μM CIDRα1 + 0.5 μM antibody mixture

The results of these experiments can be seen in **Figure 3.45**. It was clear that Jan and Anker were completely inhibitory of EPCR-binding to IT4var20 and HB3var03, respectively, since when in complex with their CIDRα1 domains there was no longer any binding to EPCR. The presence of equimolar B4 reduced the binding of GAO20 to EPCR by around 80%. As did Otto for MAL6P, however EPCR-binding was only reduced by $\sim$38%. Antibody B4 which was raised against CIDRα1.7 GAO20 but had been shown to bind CIDRα1.5 GA12, did not result in any inhibition of GA12 binding to EPCR. Despite this, the GA12:B4 complex was continued with since information on how B4 could be inhibitory to one CIDRα1 subclass and not another would be interesting and useful for future understanding. Unfortunately, the competition data for Sixten on PF08 could not be obtained due to lack of antibody.

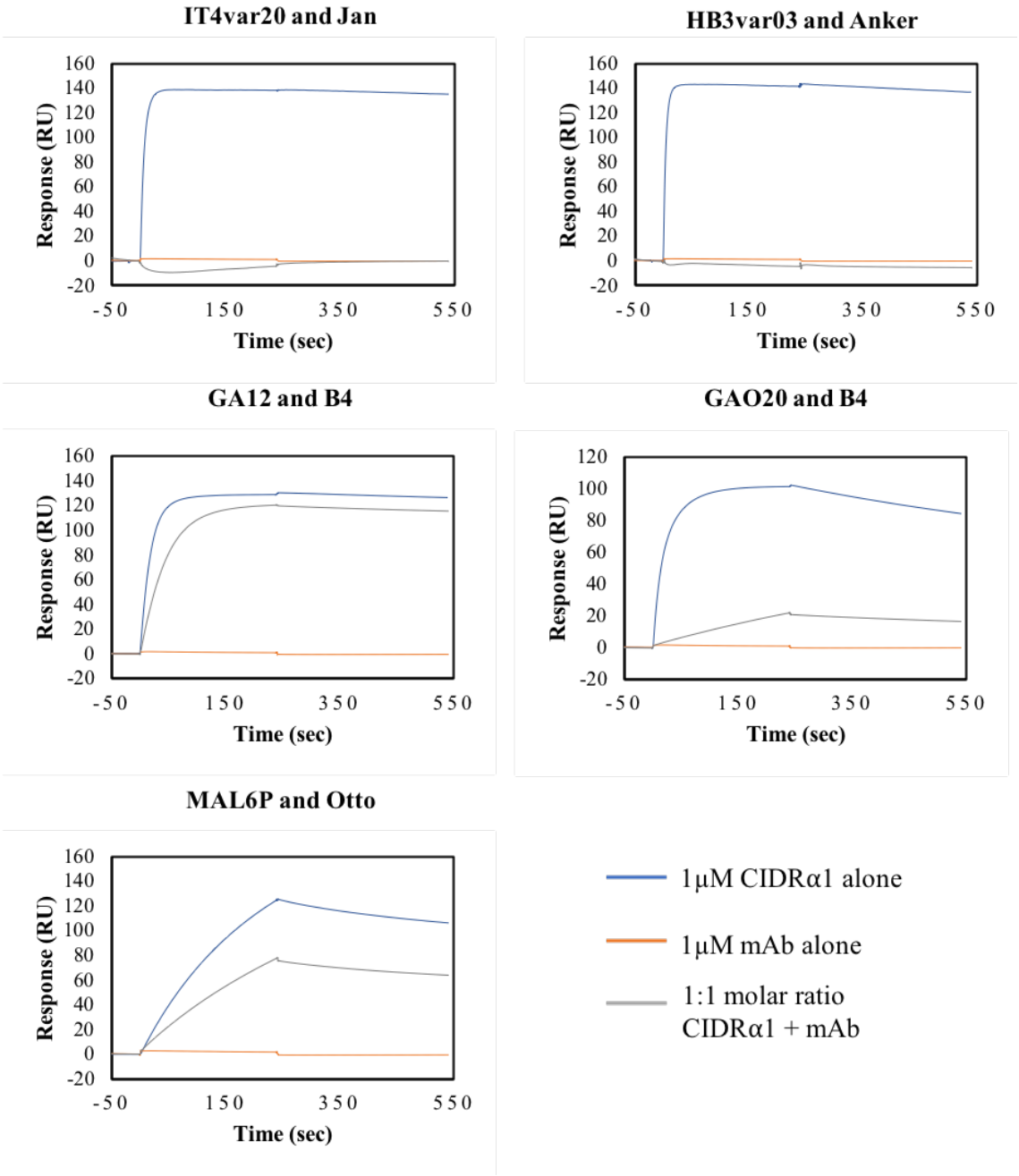


Figure 3.45: EPCR Binding Competition Between CIDRα1 domains both in Presence and Absence of mAbs

3.3.9 Attempting to Crystallise the CIDRα1:Fab complexes

The purified CIDRα1 domains and the Fab fragments were mixed in a 1.1:1 ratio (CIDRα1:Fab) for half an hour at room temperature. The complexes were then loaded on size exclusion columns, S200 3.2/30, 10/30, or 16/60 (GE Healthcare) depending on how much protein was available. The proteins eluted as one peak showing that complexes had been formed, which was confirmed by SDS-PAGE gel analysis **(Figure 3.46)**.

For all of the complexes formed, except PF08:Sixten, for which there was not enough protein, crystal plates were set up in an attempt to obtain the crystal structure of a CIDRα1 domain bound to an inhibitory Fab. The following broad crystallisation screens were set up: PACT Premier, JCSG plus, ProPlex and Morpheus, at both 4°C and 18°C **(Materials and Methods 2.7.2.1)**. Also, broad crystallisation screens were set up with complexes that had been more highly concentrated to ~20 mg/ml or methylated **(Materials and Methods 2.4.9)**, in the hope that this would decrease the solubility of the protein and aid crystallisation. In total over 100 different crystallisation screens were set up. A summary of the crystal plates set up of the different CIDRα1:Fab complexes can be seen in **Table 3.14**.

Complex	~10 mg/ml	~20 mg/ml	methylated
IT4var20:Jan	✓	✓	✓
HB3var03:Anker	✓	✓	✓
GA12:B4	✓		✓
GAO20:B4	✓		✓
MAL6P:Otto	✓	✓	✓

Table 3.14: Summary of CIDRα1:Fab Crystal Plates Set Up

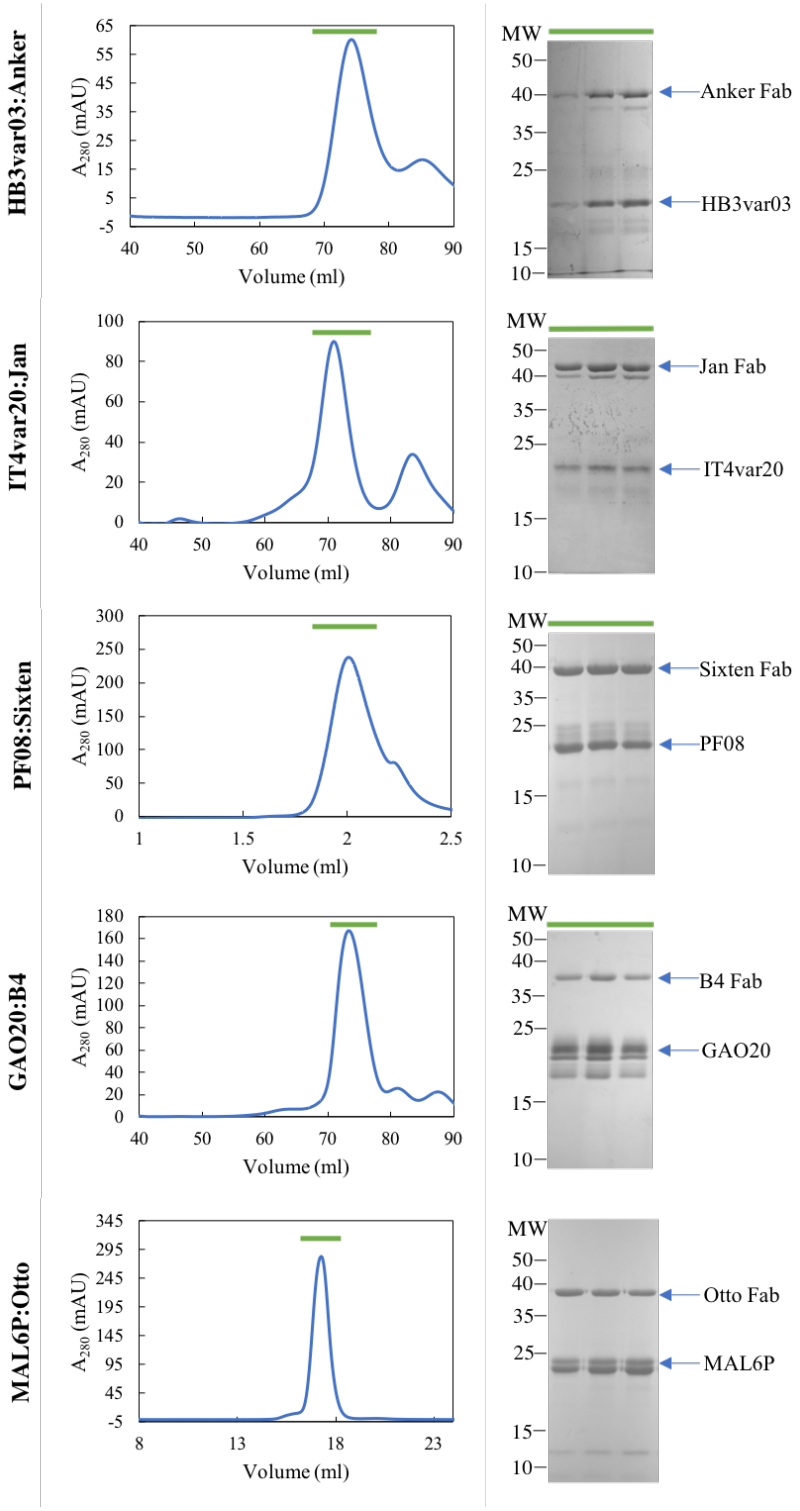


Figure 3.46: CIDRα1:Fab Complex Purification SEC trace (left) and corresponding non-reduced SDS-PAGE gel (right) showing the elution of the complex as one peak. The light green bar represents the corresponding regions of the SEC trace and SDS-PAGE gel, that were pooled for future experiments. (N.B. Different sizes of S200 size exclusion columns were used depending on how much protein was available).

3.3.9.1 Optimisation of crystal plates

The initial broad screens led to very few promising looking conditions, however in two methylated HB3var03:Anker droplets there were overnucleated microcrystals **(Figure 3.47)**. In an attempt to grow larger crystals, optimisation silver bullets screens of the condition 25% w/v polyethylene glycol 1500, 10 w/v malate-MES-tris pH6, were set up both with and without seeds **(Materials and Methods 2.7.2.2)**. Unfortunately, these optimisation screens did not lead to the growth of larger crystals, from which diffraction data could be collected.

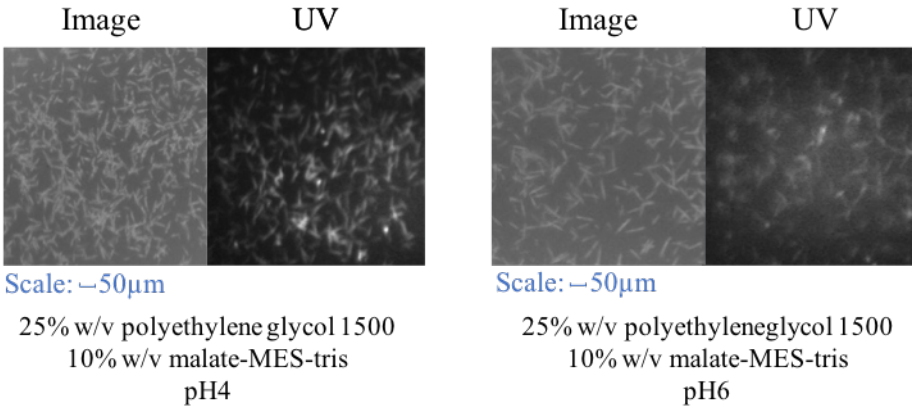


Figure 3.47: HB3var03:Anker Microcrystals The image and corresponding UV image of the microneedles that had formed in two droplets, with the buffer conditions that were in these droplets.

Also, after a few days some small crystalline objects appeared in the IT4var20:Jan crystal condition: 0.01 M Zinc Chloride, 0.1 M MES, pH6, 20% PEG 6000. Again silver bullet optimisation plates were set up in this condition both with and without seeds **(Materials and Methods 2.7.2.2)**. After around 25 days, the seeded optimisation screen lead to the growth of a couple of larger crystals **(Figure 3.48)**. The crystals were transferred into a drop of well solution containing 25% glycerol and were cryo-cooled in liquid nitrogen for storage and data collection.

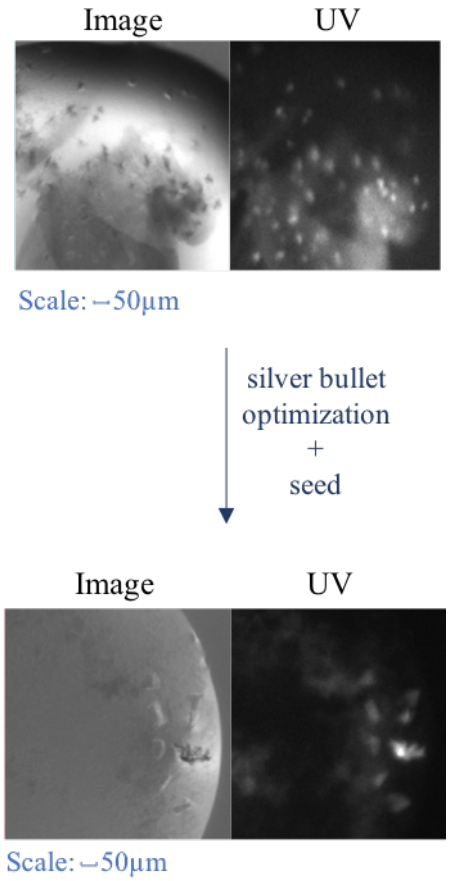


Figure 3.48: IT4var20:Jan Initial Crystals and Optimisation The image and corresponding UV image of the initial small crystals formed and the larger crystals formed after optimisation.

3.3.9.2 Data Collection for IT4var20:Jan Fab Crystal

Data were collected for the IT4var20:Jan Fab complex at the Diamond Light Source on beamline I-03. For the crystal that diffracted to the highest resolution, a dataset was collected using a collection strategy suggested in the outputs of EDNA and Mosflm. Looking at the raw diffraction images it was clear that there were diffraction spots missing or blurred in some regions of the image. It was therefore unsurprising that most of the programmes failed to process the data automatically, however DIALS2 in Xia2 was successful. This programme suggested a space group of $P2_12_12_1$ and the data collection statistics can be seen in **Table 3.15**.

Space Group	$P2_12_12_1$
Cell Parameters (Å)	a= 76.48, b=118.99, c= 122.18
(°)	α=β=γ=90
Resolution (Å)	85.25-3.13
R_{meas}	1.12 (2.96)
I/σ (I)	4.5 (1.2)
$CC_{1/2}$	0.814 (0.601)
Completeness (%)	99.6 (98.9)
Multiplicity	11.3 (10.9)

Table 3.15: Data collection statistics for IT4var20:Jan Fab

Looking at the Matthew's Coefficient for this dataset, it was suggested that there would be one copy of the Jan:IT4var20 complex per asymmetric unit of the unit cell. Therefore, molecular replacement was attempted in Phaser on this dataset using an all-alanine version of the solved structure of HB3var03 (PDB code: 4V3D Chain A) with loops removed as one search model and an all-alanine version of a Fab fragment (PDB code: 4U0R Chains B and C) as another search model. However, no molecular replacement solutions with a signal-to-noise Z-score above 8 were found (all TFZ and RFZ were below 5).

Xtriage was used to analyse the quality of the dataset **(Figure 3.49)**. This analysis showed that there were ice rings present which masks any diffraction spots at this resolution. Moreover, there was anisotropy present, which could be seen by the absence of spots in certain regions of the diffraction pattern **(Figure 3.49 A)**. This affect occurred since there was more disorder in the crystal in some directions than in others. As a result information needed to solve the structure was missing in certain resolution shells. From **Figure 3.49 B** it can be seen that the mean intensity of spots between 3.49-4.88Å was much lower than expected for a protein. As a result of the data that was missing at these resolutions, the molecular replacement was not

successful in finding a solution.

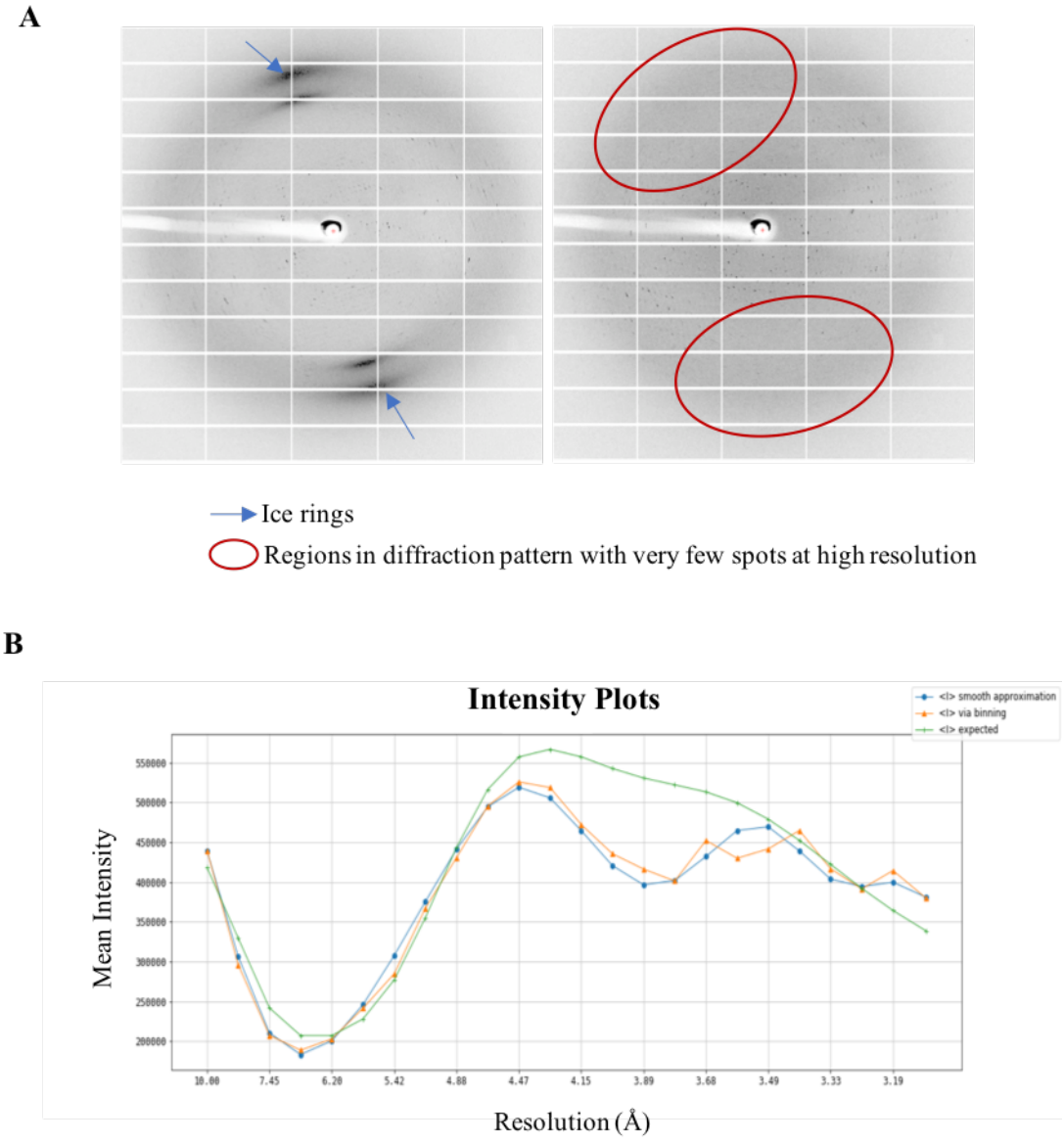


Figure 3.49: Anisotropy in IT4var20:Jan Dataset

Unfortunately, no more of these crystals were observed in the time frame of the project and so further optimisation of data collection and crystal quality could not be assessed.

3.3.10 Design of Potentially More Crystallisable CIDRα1 Constructs

3.3.10.1 Design Rational

Following the failure of any of the mouse mAbs to crystallise with EPCR, it was decided to optimise the design of the CIDRα1 domains in order to make them potentially more crystallisable.

The first design step was to remove the β-sheet from the CIDRα1 domain. The removal of this β-sheet in the past has been used to aid crystallography of CIDR domains [227] **(Figure 3.50 A and B)**. The second design step was to remove the flexible loop at the base of the two helices that form the EPCR binding site and replace it with a shorter linker **(Figure 3.50 C)**. This region in the electron density of the CIDRα1.4:EPCR structure was not visible, due to being unstructured. It was therefore thought that the removal of this flexible loop may aid crystallography of the complex by decreasing flexible surface regions.

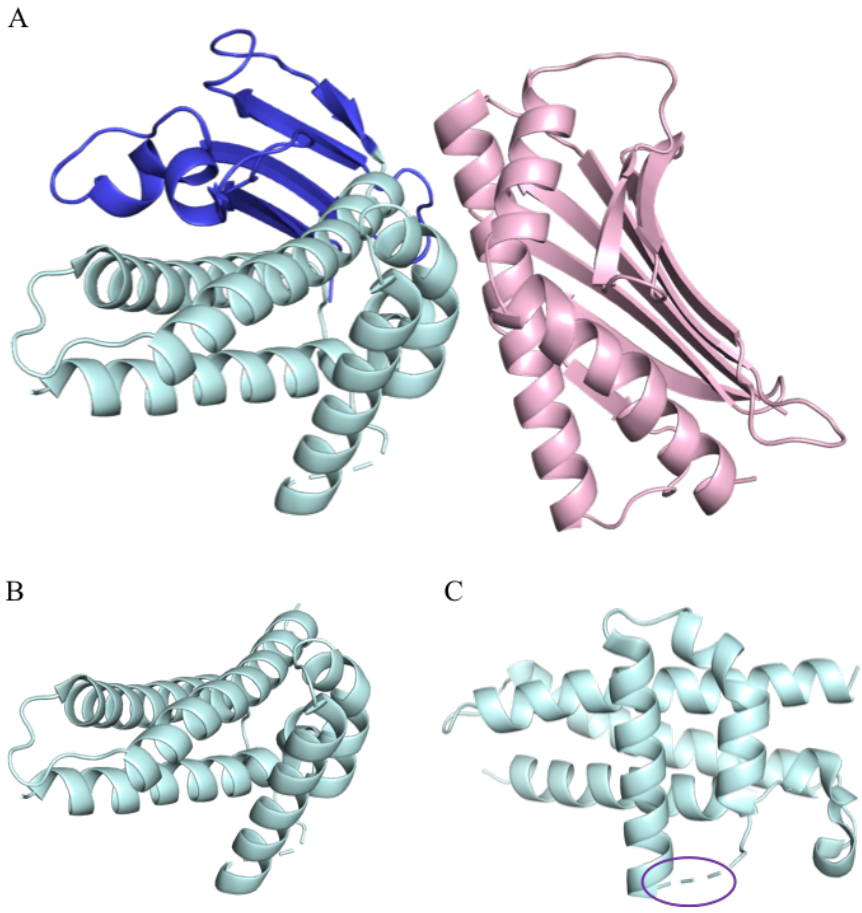


Figure 3.50: Overview of the minus β and minus loop constructs A) CIDRα1.4 (cyan) bound to EPCR (pink) with the β sheet highlighted in dark blue, not involved in the interaction. B) CIDRα1.4 minus β construct. C) CIDRα1.4 minus β construct with the loop to be shortened indicated by a purple circle.

The solved structure of the HB3var03 was used to determine the length of linker that should be introduced instead of the flexible loop **(Figure 3.51)**. The distance between the last structured residues of each helix, which the loop connected was measured as 13.1Å. Since one amino acid in an extended chain measures ~2.5Å, six amino acids (GSSGSS) were introduced to bridge this gap with a shorter loop. This loop would likely be less flexible, but not too short that there would be strain on the connecting helices. The third construct that was made was the double deletion of both the β-sheet and the loop, this construct was called minus beta-loop.

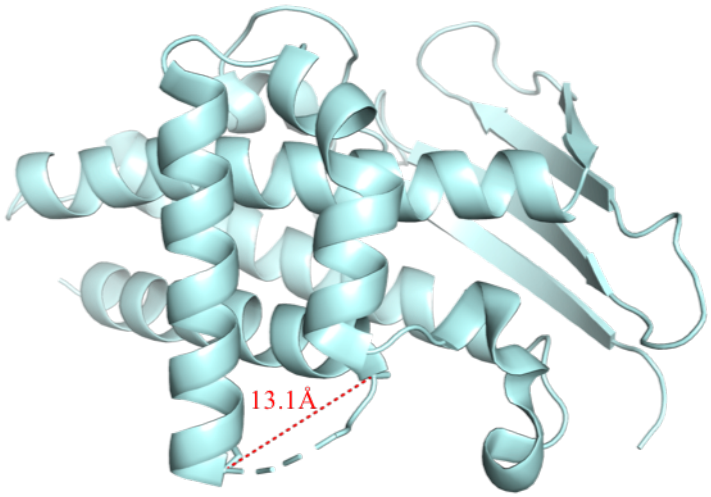


Figure 3.51: Linker length in Minus Loop Design

3.3.10.2 Construct Design

Only mutants for the IT4var20 (CIDRα1.1) and HB3var03 (CIDRα1.4) were designed and made initially due to time constraints. These two domains were chosen since the Jan and Anker antibodies showed the greatest degree of inhibition to EPCR binding **(Figure 3.45)** and were also the most available, which would allow more experiments to be carried out without delay. Therefore, the minus loop, minus beta, and the minus beta-loop constructs for both HB3var03 and IT4var20 were designed **(Figure 3.52)**.

3.3.10.3 Cloning, Expression and Purification of Designed CIDRα1 Constructs

The IT4var20 and HB3var03 minus loop, minus beta, and minus beta-loop constructs were cloned into pET15bcon **(Materials and Methods Figure 2.1)**, so that they could be expressed with a TEV-cleavable N-terminal His tag. The primers used for this mutagenesis cloning can be found in **Appendix A.2.3**.

Like their CIDRα1 domain counterparts, the new constructs were grown up in

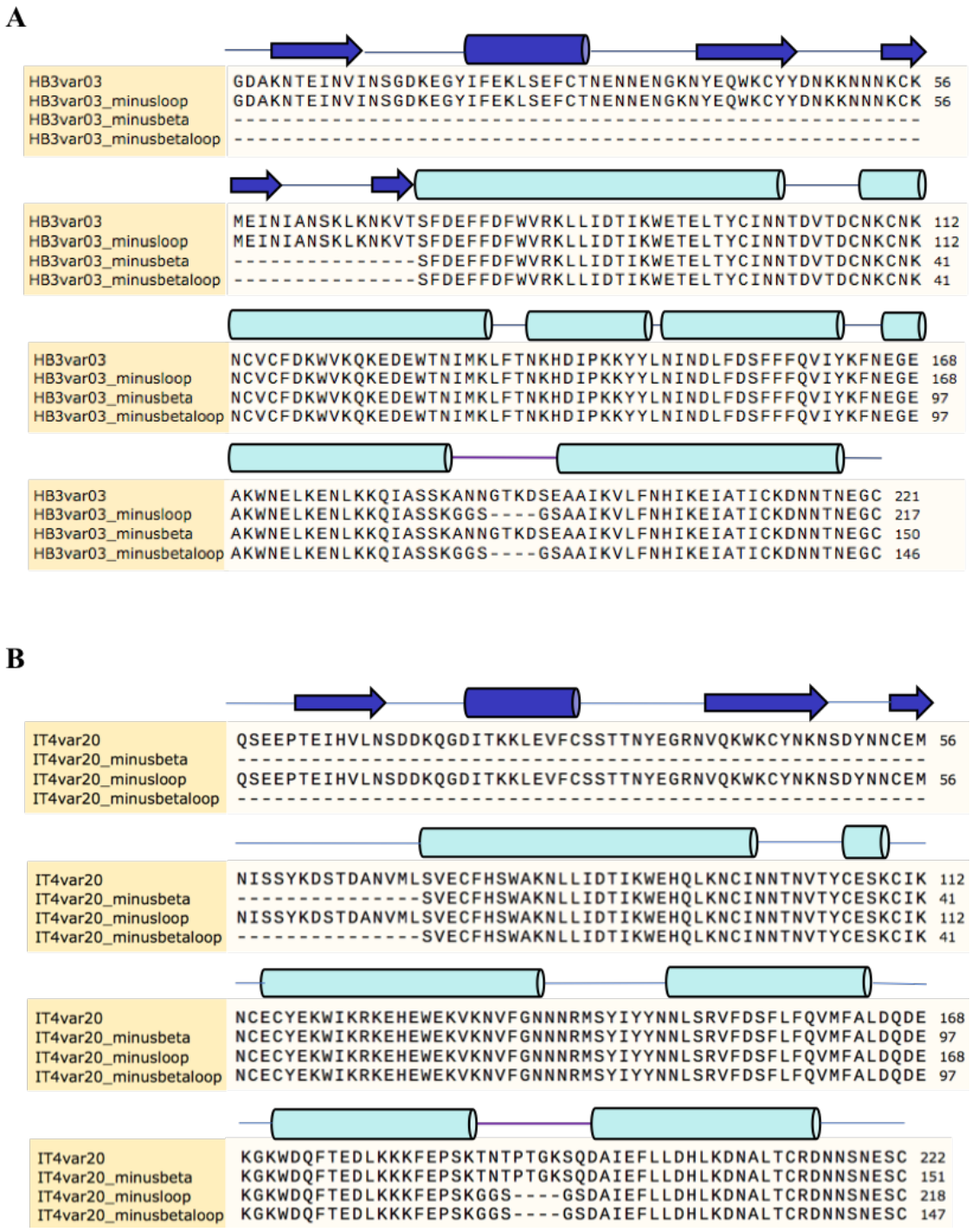


Figure 3.52: Summary of HB3var03 and IT4var20 Constructs A) An alignment of the designed HB3var03 constructs with the secondary structure from the solved HB3var03 structure, shown above the sequence. The β-strand region shown in dark blue, the loop to be removed shown in purple, and the rest of the protein shown in cyan. B) An alignment of the designed IT4var20 constructs with the predicted secondary structure from the Phyre2 server, shown above the sequence. The β-strand region shown in dark blue, the loop to be removed shown in purple, and the rest of the protein shown in cyan.

BL21(DE3) cells until an OD_{600}=0.8 was reached and were expressed overnight at 25°C by the addition of 1 mM IPTG. As expected the minus loop and beta-loop constructs were expressed in the insoluble inclusion body of the BL21(DE3) cells pellets. The cell pellets were solubilised in guanidine hydrochloride before being refolded whilst bound to the nickel resin via their His-tag in a decreasing gradient of guanidine hydrochloride **(Materials and Methods 2.4.4)**. The His-tag was removed via TEV cleavage **(Materials and Methods 2.4.5)** before being purified on a size exclusion column (S75 10/30 GE Healthcare) **(Figure 3.53)**. Unfortunately, the minus-beta constructs did not seem to express.

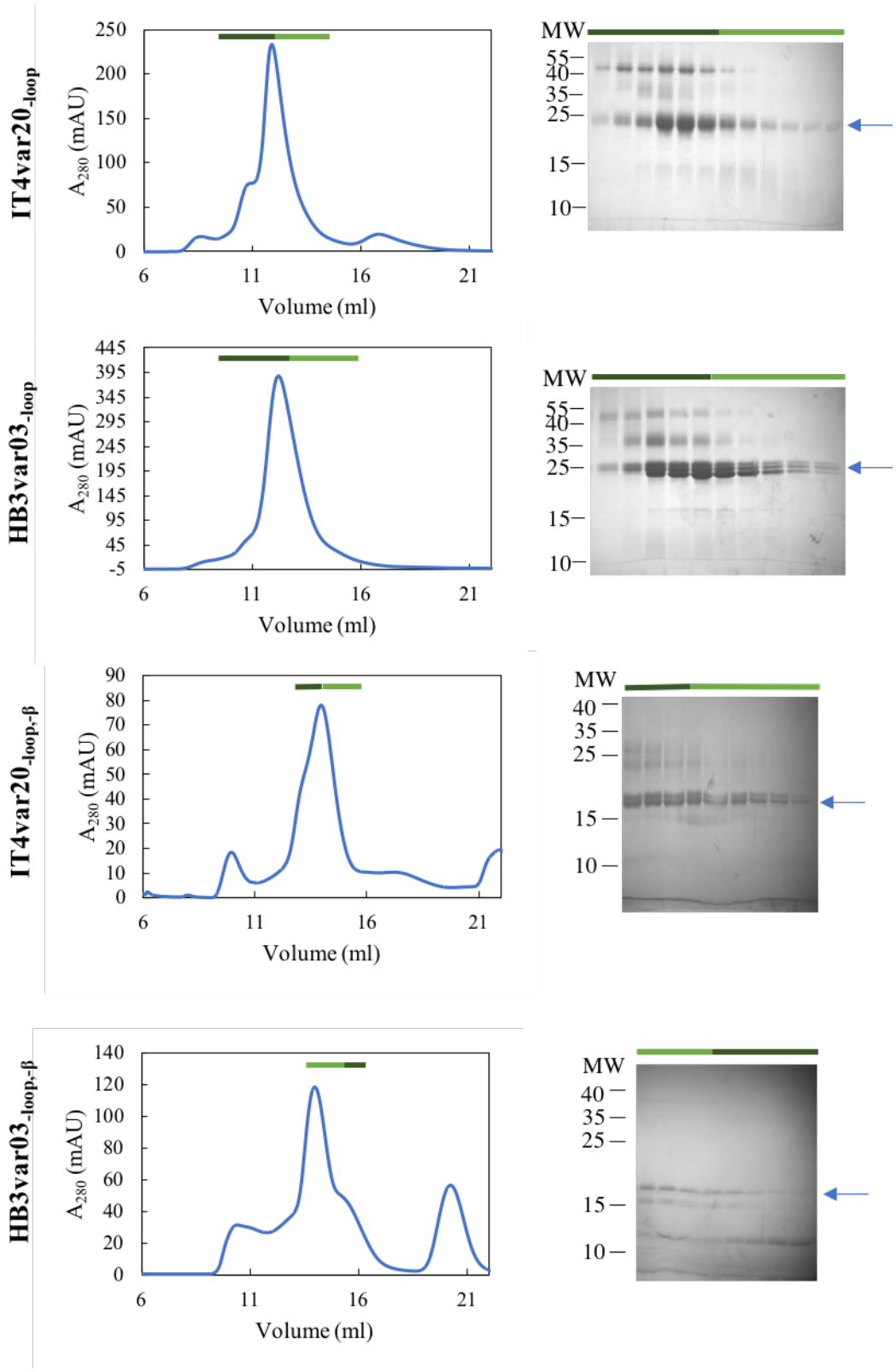


Figure 3.53: Purification of IT4var20$_{-loop}$, HB3var03$_{-loop}$, IT4var20$_{-loop,\beta}$ and HB3var03$_{-loop,\beta}$ The green bars represent the corresponding fractions of the S75 10/30 run in the size exclusion trace (left) and the non-reduced SDS-PAGE gel (right). The light green bar represents the fractions that were pooled for further studies.

3.3.10.4 Assessing the Fold of the CIDRα1 Constructs

To assess the secondary structure of the newly designed constructs, CD spectra were taken between 200 and 250 nm **(Materials and Methods 2.6.1)**. The shape of these spectra with minima at 208 nm and 222 nm were characteristic of proteins that had alpha-helical secondary structure, which was expected for these constructs **(Figure 3.54)**.

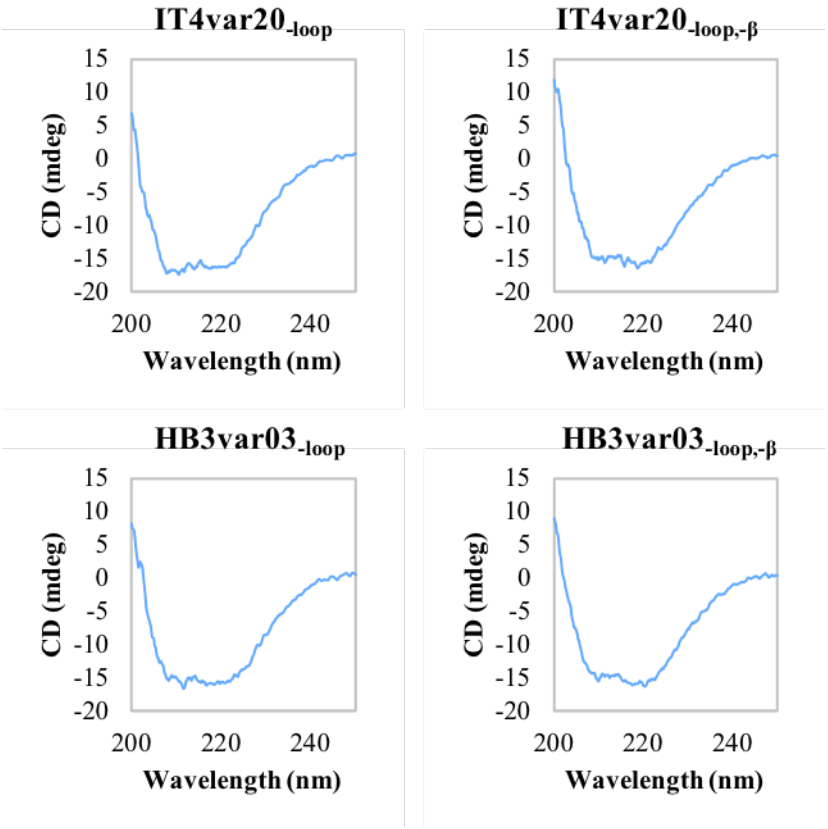


Figure 3.54: CD of Minus Loop and/or Beta Constructs

3.3.10.5 Binding of the Designed CIDRα1 Constructs to EPCR

In order to assess whether the designed constructs were properly folded and retained their function despite the deletions introduced, their binding to EPCR was checked via SPR. Biotinylated EPCR was immobilised on a CAPture chip and two-fold dilution series of each construct starting at 1 μM were flowed over and binding

assessed **(Materials and Methods 2.6.2)**. The binding curves can be seen in **Figure 3.55**.

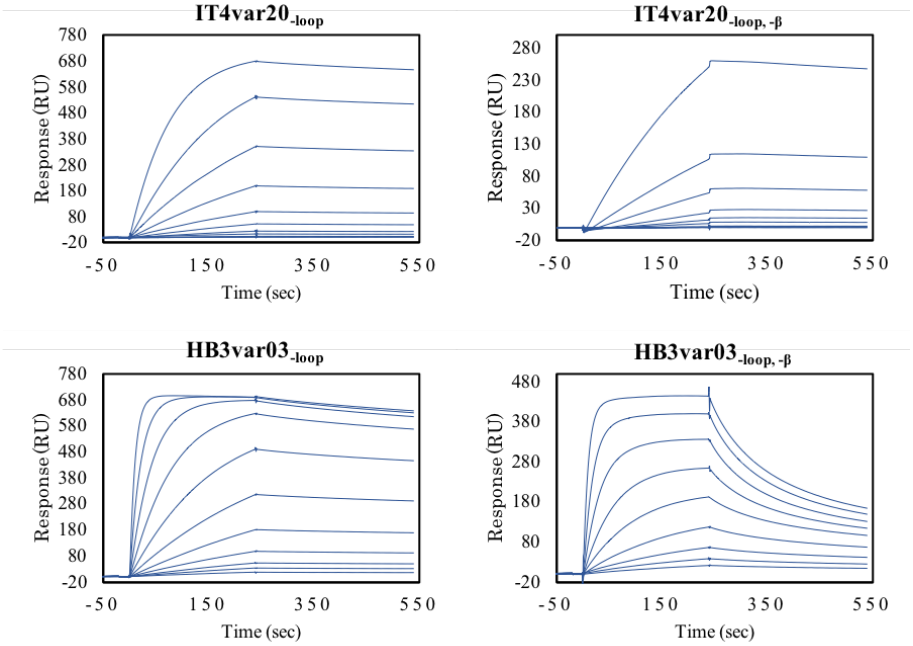


Figure 3.55: SPR curves showing binding of CIDRα1 constructs binding to EPCR

The kinetic curves were fit using the BIAevaluation software, assuming a 1:1 interaction, to determine the kinetic constants and the affinity of the interactions **(Table 3.16)**. All of the constructs still bound to EPCR. Both of the minus-loop constructs retained very fast association rates, slow dissociation rates, and low nanomolar affinity characteristic of their cognate CIDRα1 domains binding to EPCR. The HB3var03$_{-loop,-\beta}$ construct also mimicked the binding of the cognate HB3var03 to EPCR, although it had a 16 times faster off-rate. In contrast, the IT4var20$_{-loop,-\beta}$ binding to EPCR had a an extremely slow association rate to IT4var20 but an identical off-rate, which resulted in a $\sim$3000 decrease in affinity. The deletion of the β-sheet and the shortening and mutation of the loop both lead to an overall change in the architecture of the domain, which may result in slight changes in conformation of the EPCR-binding site. These small changes could explain the reduced affinities seen for these designed constructs. Despite some changes in the binding profiles

and affinity of the constructs to EPCR compared to their cognate CIDRα domains, nearly all of the constructs retained the ability to bind EPCR with nanomolar affinity showing that they were mostly folded correctly.

	$\mathbf{k}_{on}$ ($\mathbf{M}^{-1}s^{-1}$)	$\mathbf{k}_{off}$ ($\mathbf{s}^{-1}$)	$\mathbf{K}_d$ (nM)
IT4var20	332000	1.23x10^{-4}	0.37
IT4var20$_{-loop}$	14800	1.75x10^{-4}	11.8
IT4var20$_{-loop,-\beta}$	137	1.23x10^{-4}	1010
HB3var03	532000	1.97x10^{-4}	0.37
HB3var03$_{-loop}$	92200	3.25x10^{-4}	3.52
HB3var03$_{-loop,-\beta}$	150200	3.24x10^{-3}	21.6

Table 3.16: Kinetic Binding Analysis of CIDRα1 Domains Binding to EPCR

3.3.10.6 Binding of the Designed CIDRα1 Constructs to Inhibitory Mouse mAbs

Having demonstrated that the constructs were all folded and maintained their binding to EPCR, the binding of these potentially more crystallisable constructs were tested against the mouse mAbs by SPR. In each case protein A/G was coupled to a CM5 chip **(Materials and Methods 2.6.2)** before the mouse mAb was bound to the protein A/G. Once the chip had been prepared different concentrations of the CIDRα1 constructs were flowed over to assess the binding kinetics to the mouse mAbs. The HB3var03 constructs were tested on an Anker-coupled chip and the IT4var20 constructs were tested on a Jan-coupled chip **(Figure 3.56)**.

The constructs all bound to their mAb with similar looking binding profiles to their cognate CIDRα1 domains. This showed that the deletions introduced did not affect the ability of these constructs to bind to the mAbs, and their epitopes were not

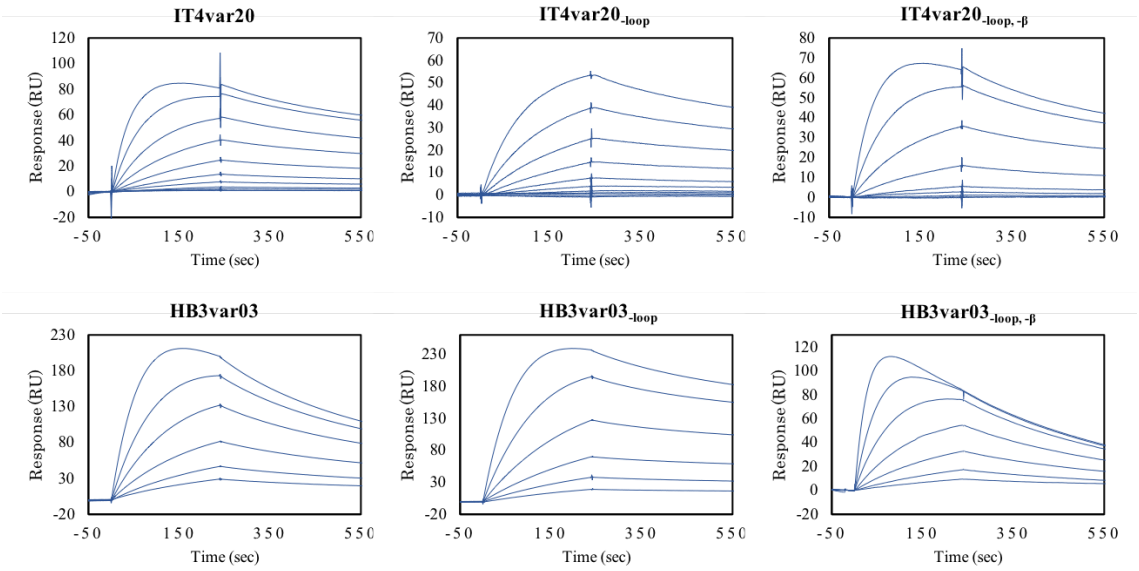


Figure 3.56: SPR curves showing binding of CIDRα1 constructs binding to their mAb The binding of the top row constructs (all IT4var20) were tested against the Jan mAb, and the binding of the bottom row constructs (all HB3var03) were tested against the Anker mAb.

disrupted by these deletions. Unfortunately, as with the CM5 chip system before, there was a problem with the antibodies uncoupling from the chip as the experiment was run. As a result, the kinetic software in the BIAevaluation programme was not able to confidently fit the curves and so no kinetic analysis could be obtained. However, these SPR experiments showed that the constructs stably bound to the mAbs and so there could be confidence that stable complexes could be formed for crystallography.

3.3.10.7 Attempted Crystallisation of the Designed CIDRα1 Constructs with their Inhibitory Mouse Fabs

Due to the limited amount of protein obtained in each preparation and time constraints, it was decided to only set up crystal plates with the double mutant constructs (minus beta-loop). Since this mutant removed both the loop and beta-sheet, it was thought that there would be more chance of this crystallising than the construct

with just the removed loop. Unfortunately, the formation of complex and setting up crystal plates was only achieved for the IT4var20$_{-loop,-\beta}$ and Jan Fab.

The purified IT4var20$_{-loop,-\beta}$ construct and the Jan Fab were mixed together in a 1:1 molar ratio for 30 min at RT. The complex was then concentrated and applied on a size exclusion chromatography column (S200 10/30, GE Healthcare).

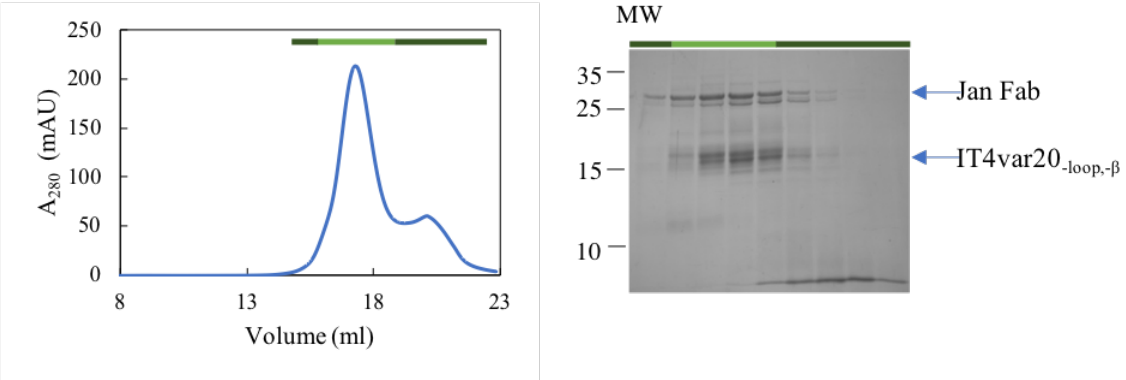


Figure 3.57: Purification of IT4var20$_{-loop,\beta}$:JanFab Complex. The size exclusion chromatography trace (left) and corresponding reduced SDS-PAGE gel (right) of the complex formation. The coloured bars demonstrate which part of the peak represents which region of the gel. With the light green bar showing the pooled fractions for crystallography.

The IT4var20$_{-loop,\beta}$:JanFab Complex was concentrated to 11.3 mg/ml and the following broad crystallisation screens were set up: PACT Premier, JCSG plus, ProPlex and Morpheus, at 18°C **(Materials and Methods 2.7.2.1)**. Regrettably, again no crystals were observed in these plates.

3.3.11 Discussion

In the previous section, it was shown that designing a vaccine that solely targeted the EPCR-binding site of the *Pf*EMP1 did not produce more effective or cross-reactive antibodies than using the whole CIDRα1 domain. As a result, in order to gain a greater understanding of the epitopes of inhibitory antibodies against this interaction, it was attempted to get the crystal structure of a CIDRα1 domain in complex with

an inhibitory mouse antibody. Using this structural information, it was hoped that novel vaccines could be designed, which upon immunisation produced inhibitory antibodies, as has previously been done successfully [137, 218].

Our collaborators immunised mice with individual CIDRα1 domains that represented a spread of the different EPCR-binding CIDRα1 subclasses. Antibodies were identified that were inhibitory against their immunised CIDRα1 domain binding to EPCR. Shorter, more crystallisable CIDRα1 constructs, based on the domain boundaries of the successfully crystallised HB3var03 werre cloned, expressed and purified. CD spectra and SPR experiments confirmed that these CIDRα1 domains were correctly refolded from inclusion bodies and retained their functional EPCR-binding site. SPR curves also showed that these constructs bound to the antibodies, and that the mAbs were inhibitory to the binding of EPCR to the CIDRα domain used for immunisation. One non-inhibitory mAb interaction was also characterised.

Many crystallisation plates were set up of the CIDRα1 domains in complex with the Fab fragment of their binding mAb. However, unfortunately no diffracting crystals from which a structure could be solved were produced. It was unclear why none of these complexes produced diffracting crystals. The process of why and when some proteins will crystallise is little understood and cannot be predicted. As a result, many screens which encompass hundreds of different combinations of buffers, precipitants, pHs etc. are set up in the hope that crystals will form in one of them. It could just be a case that the right condition was not found. One difficulty in setting up lots of plates of this complex was that the yield of pure CIDRα1 domains obtained after purification was very small ($\sim$0.2 mg/L *E.coli*) and so with each preparation only a few plates could be set up. Moreover, since the CIDRα1 domains were purified from inclusion bodies, there was variation in how successful the refolding protocol was in each preparation. This may have contributed to some purifications being of better quality than others, as well as possibly introducing heterogeneity into the

sample.

In an attempt to increase the chances of crystallisation of the CIDRα1 domains, constructs were designed in which a flexible loop and/or the β-sheet were removed. The removal of this β-sheet has proved successful in aiding crystallisation in the past [227]. SPR binding experiments confirmed that the constructs with these deletions still bound to both EPCR and their inhibitory mAb. However, the affinity of the constructs for EPCR was reduced compared to the full-length CIDRα1 domains. This was likely due to the large deletions of the β-sheet and shortening of the loop, slightly affected the fold and packing of the EPCR-binding site, which resulted in EPCR binding with lower affinities. Since the inhibitory mAbs bound to all of the constructs, it can be concluded that the epitope of the mAb does not lie in the loop region near the EPCR-binding site (nor in the β-sheet). Unfortunately, due to time constraints and low yield only one set of crystallisation plates could be set up for the IT4var20$_{-loop,\beta}$:JanFab complex to further probe the epitope of this inhibitory mAb, which did not produce crystals.

In the future, crystallisation of the other minus-loop, minus beta-loop constructs could be set up in complex with their Fab fragments to complete these screens. If crystallisation continues to prove problematic, information about the positions of the inhibitory mAbs' epitopes could be obtained through hydrogen-deuterium exchange experiments (HDX). Although this method would not give information about the individual residues involved in directly binding to the mAb, it will give an idea of the surface of the CIDRα1 which contacts the mAb. This information would be useful in narrowing down the areas of the CIDRα1 which form the epitopes of inhibitory mAbs. These regions could then be the focus for future vaccine design. In addition, optimisation of soluble CIDRα1 expression in an alternative expression system may reduce variability between preparations and increase yield, which might contribute to a greater number of crystal plates being set up with more uniform

protein, potentially aiding crystallography.

Overall, so far attempts to obtain structural information on the binding of inhibitory mAbs to CIDRα1 domains has proved unsuccessful. Continued research into the structural features of an epitope of an inhibitory antibody is crucial for future vaccine design in targeting EPCR-binding *Pf*EMP1s.

Chapter 4

Targeting Invasion in *Plasmodium vivax* Malaria via *Pv*DBP

4.1 Introduction to Reticulocyte Invasion by *Plasmodium vivax*

As with *P. falciparum*, the invasion of merozoites into RBCs is an essential event in the parasite life cycle, and therefore a good stage to target with a potential vaccine. Moreover, the events of the erythrocytic cycle are correlated with clinical symptoms of malaria, and therefore therapeutic prevention of parasite entry into RBCs could alleviate disease.

The invasion of RBCs by merozoites is a highly complicated and multi-step process that is dependent on a cascade of specific molecular interactions between host cell receptors and parasite ligands [228]. These interactions initiate an intracellular cascade of molecular events that progress from initial attachment, reorientation, commitment, followed by penetration into the RBCs [25]. Not much is known about the *P. vivax* invasion pathway, mostly due to an inability to produce a stable *P. vivax* culture *in vitro*. In comparison, continuous *P. falciparum in vitro* cultures have been around since the 1970s [229, 230] and have been successfully used to identify most of the proteins involved in the multiple redundant pathways of *P. falciparum*

erythrocyte invasion **(Figure 4.1)** [23, 91].

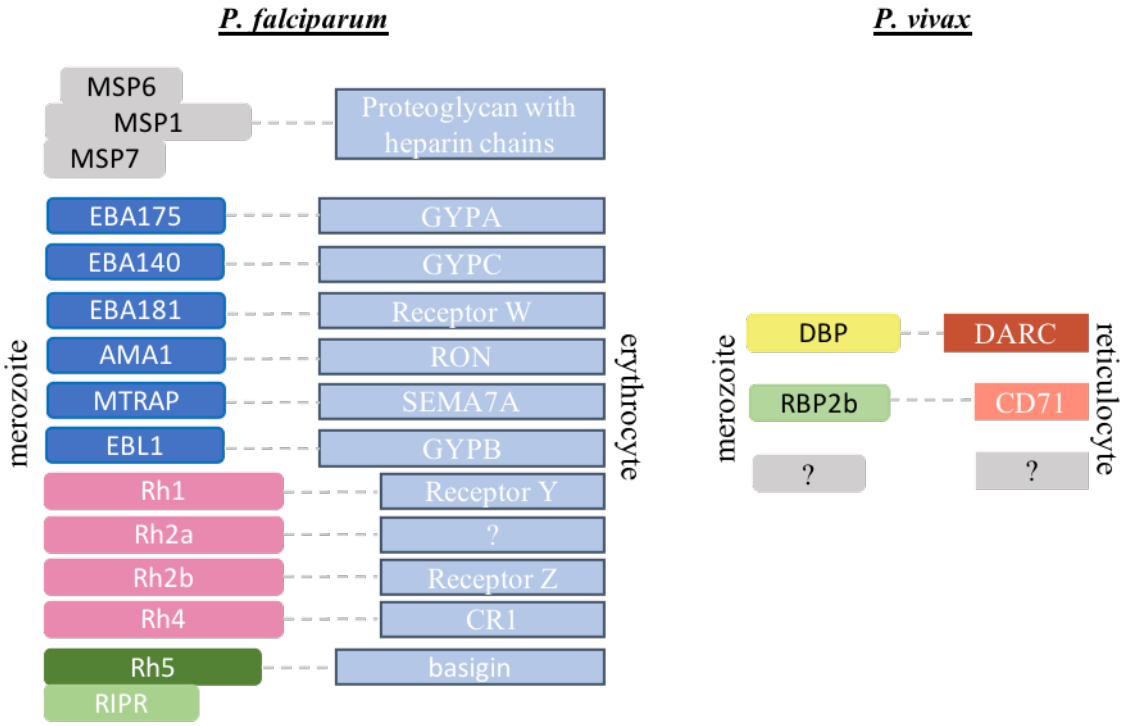


Figure 4.1: Comparison of Red Blood Cell Invasion Pathways in *Plasmodium falciparum* and *Plasmodium vivax* Merozoite receptors are shown in colours that represent groups that have redundancy. The left panel shows that there are many redundant receptors involved in invasion in the *P. falciparum* erythrocyte invasion pathway, whereas the right panel shows that *P. vivax*invasion is almost completely dependent on *Pv*DBP:DARC interaction.

Despite the problems with culturing *P. vivax*, one blood cell receptor has been known to be crucial in *P. vivax* invasion into RBCs since the 1970s: Duffy antigen receptor for chemokines (DARC). In 1975, Miller *et al.*, identified that individuals in Western and Central Africa, who lacked DARC on their RBCs, were resistant to *P. vivax* invasion [231]. Later it was shown that it was the interaction of the parasite Duffy binding protein (DBP) with DARC that was crucial for the formation of the tight junction during invasion **(Figure 4.1)** [232–235].

As described in **Section 1.1.5.3**, *P. vivax* only invades reticulocytes. Since DARC is present on both reticulocytes and mature erythrocytes, it cannot be this interaction

alone that governs the selective entry of *P. vivax* [236]. Recently, the interaction between CD71 (a reticulocyte receptor) and a *P. vivax* reticulocyte binding protein (RBP2b) was discovered [237]. It was shown that when this interaction was blocked by mAbs *in vitro*, invasion of *P. vivax* into reticulocytes was also inhibited. Therefore, it is likely that this CD71:*Pv*RBP2b interaction may be critical for the recognition of reticulocytes during *P. vivax* invasion and might provide the specificity for CD71-reticulocytes **(Figure 4.1)**.

Over the past decade there has been evidence that in a small number of cases, *P. vivax* can infect and cause disease in DARC-negative individuals (people that do not express DARC on their reticulocytes) [238–240]. This may suggest the presence of an alternative, DARC-independent invasion route [241]. An alternative pathway has yet to be discovered, which could be due to the current inability to probe the mechanisms of invasion because of a lack of *in vitro* continuous culture. Another option is that this is a relatively new adaptation, developed by *P. vivax* as a survival strategy in individuals with a Duffy negative phenotype. Some evidence suggests that a *Pv*DBP gene duplication, which is most prevalent in areas with a high-frequency of Duffy negative individuals, may be evolutionarily selected in order to overcome the constraints imposed by the lack of the DARC on the reticulocyte [240]. The genes for other RBPs have also recently been discovered, however, their roles are currently not understood. Our understanding of *P. vivax* invasion into reticulocytes is far from complete, and more research is required to identify other invasion ligands for both the the DARC-dependent and alternative pathways.

Despite the growing evidence that in some individuals there may be a Duffy-independent invasion pathway, there is a plethora of evidence that the *Pv*DBP:DARC interaction is the only invasion route in most *P. vivax* infections. Therefore, unlike in *P. falciparum* where multiple redundant invasion patwhays would need to be targeted to block invasion, P. vivax invasion can be stopped by targeting just the

*Pv*DBP protein and hence is an ideal vaccine target.

4.1.1 *Plasmodium vivax* Duffy Binding Protein

During invasion, *Pv*DBP is expressed and secreted from the micronemes in the merozoite and binds to its host ligand DARC [233, 242, 243]. *Pv*DBP has six extracellular regions followed by a type I transmembrane domain and a short cytoplasmic tail **(Figure 4.2 A)**. There are two cysteine-rich domains (regions II and VI) in the ectodomain, with *Pv*DBP region II (*Pv*DBP-RII) alone shown to be sufficient for DARC binding [188, 189]. *Pv*DBP-RII belongs to the Duffy binding-like (DBL) domain family and contains roughly 330 amino acids. Like other DBL domains, *Pv*DBP-RII has three distinct subdomains with multiple (in this case twelve) conserved cysteine residues that form intramolecular disulphide bonds **(Figure 4.2 B)**. Other erythrocyte binding proteins (EBPs) of the DBL family, which include the *P. falciparum* erythrocyte binding antigen *Pf*EBA-175, have been shown to mediate binding via similar conserved domains [243, 244].

As expected from the sequence analysis, the first crystal structure of *Pv*DBP-RII, revealed that it had a very similar structure to other known erythrocyte binding DBL domains, with twelve helices and three subdomains, each stabilised by disulphide bonds **(Figure 4.3)** [245, 246].

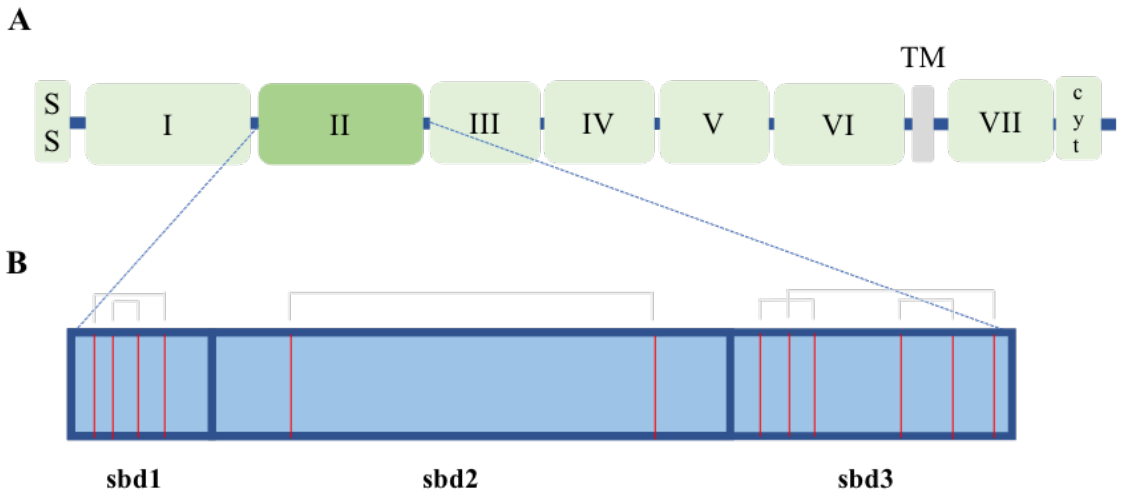


Figure 4.2: *Pv*DBP Domain Architecture A) The overall domain architecture of the *Pv*DBP protein, with the seven main domains numbered I-VII. The domain II is highlighted in a darker green. SS = signal sequence. cyt =cytoplasmic domain. B) The close up of *Pv*DBP-RII shows the subdomain regions (sbd1=subdomain1, sbd2=subdomain2, and sbd3=subdomain3) and the conserved cysteines in red and their pairings (grey brackets) within the subdomains.

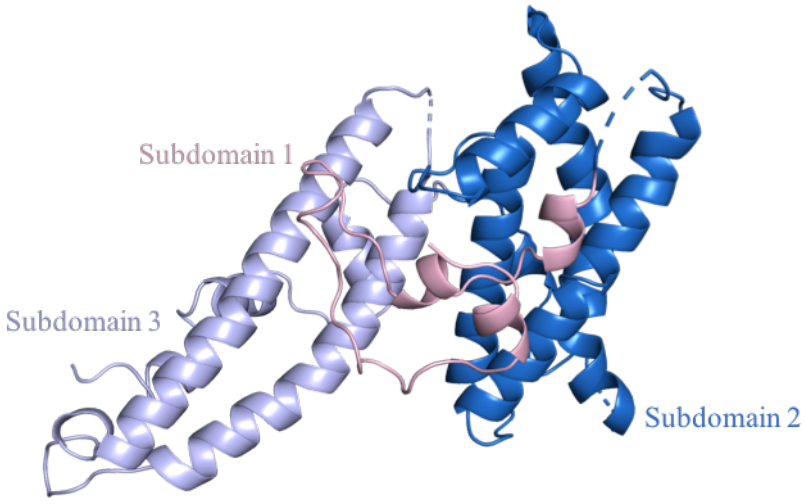


Figure 4.3: *Pv*DBP-RII Structure The structure of *Pv*DBP-RII (PDB code: 3RRC) [245]. The three distinct subdomains shown in different colours: subdomain1 (pale pink), subdomain2 (royal blue), and subdomain3 (lilac).

4.1.2 DARC

DARC (officially known as ACKR1, atypical chemokine receptor 1) is a transmembrane G-protein coupled receptor (GPCR) that lacks the signaling motif that allows the intracellular domain to bind to G proteins [232]. Like all GPCRs, DARC is predicted to have seven transmembrane helices, with a short N-terminal domain and C-terminal sequence **(Figure 4.4)**. It has been shown that the extracellular N-terminal 60 amino acid sequence blocks *Pv*DBP binding to reticulocytes [247]. It is therefore, thought that this N-terminal region, may form most of the binding site for *Pv*DBP-RII required for *P. vivax* invasion.

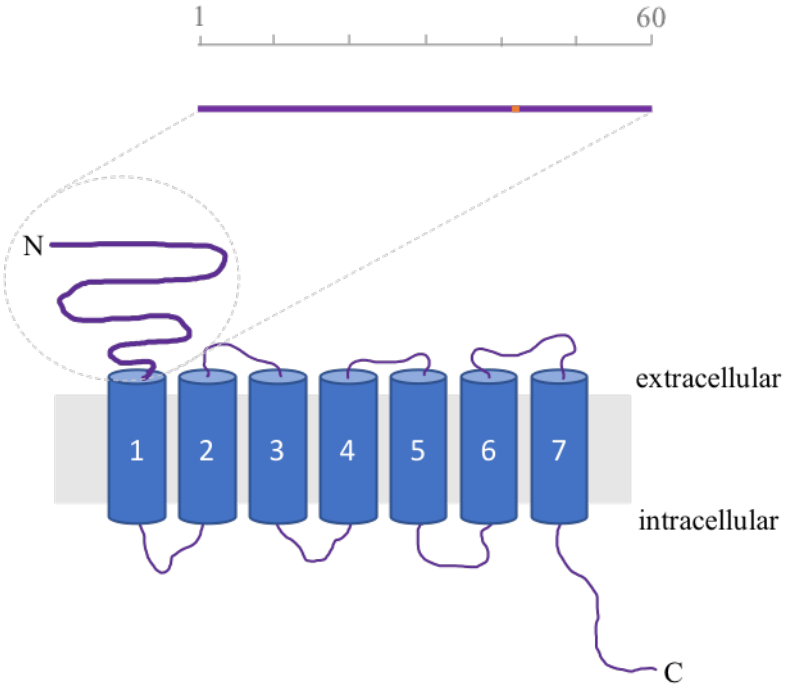


Figure 4.4: Schematic of DARC Architecture The seven helices, typical to GPCRs, are shown in blue, with the N and C termini marked. A close-up of the N-terminal 60 residue fragment is shown with the residue mutated from Gly to Asp in the FY∗A and FY∗B alleles highlighted in orange.

The host role of DARC on RBCs is not completely understood. When discovered, DARC was identified as a promiscuous chemokine receptor, binding to both CXC

and CC chemokines [232]. It has since been suggested that DARC acts as chemokine scavenger, lowering the concentration of chemokines in circulation, which can be toxic in certain conditions [248]. In addition, DARC is not only expressed on erythrocytes, but also in endothelial cells, even in Duffy negative individuals [249]. On endothelial cells, there is evidence that DARC mediates chemokine-induced leukocyte migration across the endothelial barrier leading to inflammation and other pro-inflammatory effects [250]. Since the Duffy negative phenotype only affects expression on erythrocytes and not endothelial cells, this may explain why the DARC gene and protein evolutionarily remain, despite not being necessary on erythrocytes for survival.

4.1.3 DARC:*Pv*DBP Interaction

Although the *Pv*DBP:DARC interaction has been shown to be crucial for the formation of a tight junction, it is not completely understood how the binding of *Pv*DBP to DARC leads to subsequent invasion of the merozoite into the reticulocyte. This lack of knowledge is partially due to the full-length structures of *Pv*DBP and particularly DARC being uncharacterised. As previously mentioned, a fragment containing the N-terminal 60 residues of DARC has been shown to bind to *Pv*DBP-RII [247]. Structural studies using this fragment and *Pv*DBP-RII have suggested a step-wise binding of one $DARC_{1-60}$ fragment to a *Pv*DBP-RII, that then induces the dimerisation of *Pv*DBP-RII, finally allowing a heterotetramer of 2:2 *Pv*DBP-RII:$DARC_{1-60}$ to be formed **(Figure 4.5)** [251].

From this model it has been proposed that the dimerisation of *Pv*DBP-RII stabilised by the binding to the DARC homodimer (GPCRs often dimerise), may activate signal transduction in the *P. vivax* to initiate the erythrocyte invasion pathway. This hypothesis is supported by the further structural and epitope mapping studies, which indicate that *Pv*DBP may exist as a dimer in solution [245]. The structure showed

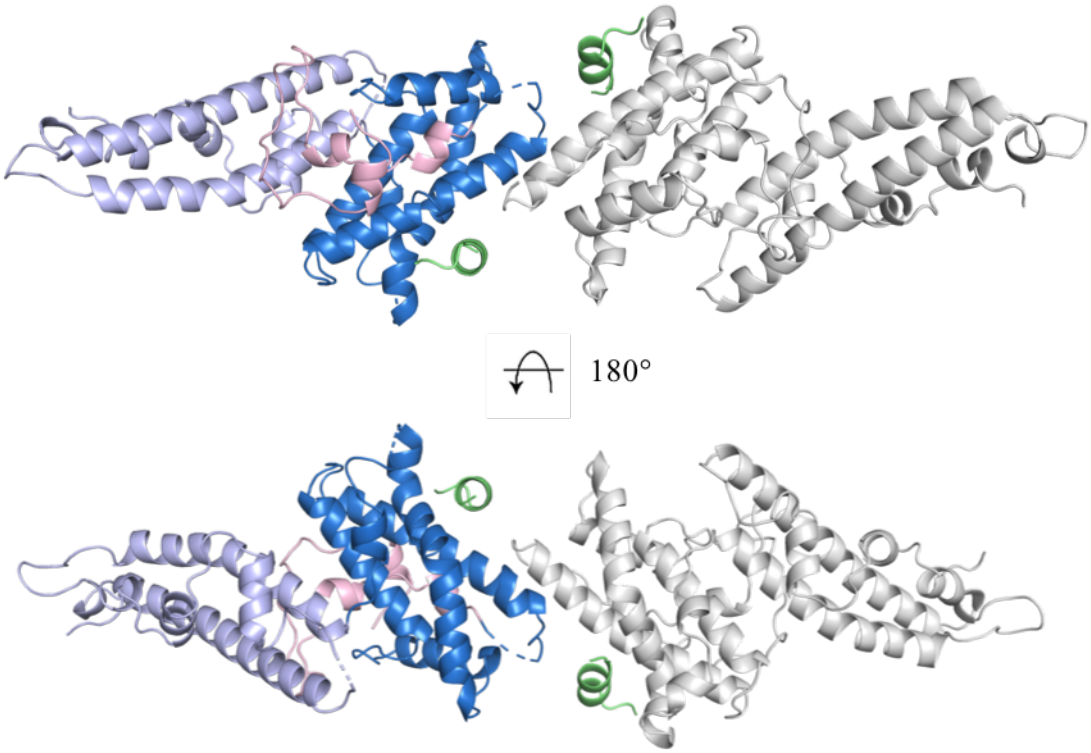


Figure 4.5: Heterotetramer of *Pv*DBP-RII:DARC$_{19-30}$ One copy of *Pv*DBP-RII is shown in grey. The other copy of *Pv*DBP-RII is coloured according to it's subdomains: subdomain 1 in pink, subdomain 2 in blue, and subdomain 3 in violet. The two DARC$_{19-30}$ fragments are shown in green. (Structure from [251]).

that this DARC fragment bound to subdomain 2 of *Pv*DBP-RII, at the predicted dimer interface. Despite *Pv*DBP-RII being polymorphic, this DARC binding site is highly conserved amongst *P. vivax* strains [246].

Even with this proposed model, some questions still remain regarding the *Pv*DBP-RII:DARC$_{1-60}$ interaction. Firstly, in the solved structure, only eleven out of the sixty residues of the DARC$_{1\text{-}60}$ had enough density to be built with any certainty. This region of DARC contained residues 19 to 30 (DARC$_{19\text{-}30}$) and formed a short length of helix **(Figure 4.6)**.

The rest of the DARC fragment could not be seen, indicating that it was not stabilised in the crystal and there was a degree of flexibility. Therefore, the position of the other residues of the DARC extracellular fragment is unclear. Specifically, the stabilised

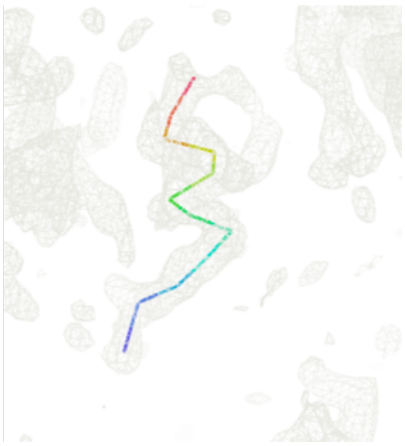


Figure 4.6: Electron Density Map for the DARC Fragment The backbone of $DARC_{19-30}$ is shown coloured as a rainbow with the N-terminus being blue and the C-terminus red. The electron density map is shown as a grey mesh and is at σ level 1. (Electron Density and $DARC_{19-30}$ from [251]).

DARC region did not contain the Tyr41 residue, which when sulphated *in vivo* has been shown to be crucial for binding to *Pv*DBP [252]. Moreover, another tyrosine has also been shown to be sulphated (Tyr30) in native DARC. Since the $DARC_{1\text{-}60}$ construct used to determine this structure was expressed in *E. coli* bacterial cells, the tyrosines would not have been post-translationally sulphated in this prokaryotic system. As a result of this lack of sulphation and electron density, especially of the crucial Tyr41, it is unlikely that this predicted model is the complete story of DARC binding to *Pv*DBP-RII. This has been further confirmed through mutagenesis studies, which identified an additional patch of residues on subdomain 2 thought to be involved in the binding of *Pv*DBP-RII to DARC [234, 253]. This suggested that other DARC residues may be involved in the binding to *Pv*DBP-RII. Greater understanding of the interaction between DARC and *Pv*DBP-RII may give us more insight into designing a new vaccine that targets this binding event and subsequent RBC invasion.

4.1.4 *Pv*DBP-RII as a Vaccine Candidate

Vaccines that target invasion into reticulocytes aim to slow down and neutralise this asexual growth stage of the parasite. There is great need for a blood stage vaccine that acts quickly and kills the *P. vivax* before gametocyte production and subsequent transmission. *Pv*DBP-RII is an ideal vaccine candidate to target RBC invasion, since in most *P. vivax* infections this invasion is thought to be solely dependent on the *Pv*DBP-RII:DARC interaction and there are currently no known redundant pathways.

Further evidence that targeting RBC invasion with this interaction comes from the protection against *P. vivax* invasion conferred in Duffy-negative individuals [95]. The DARC gene is highly polymorphic and has two major alleles FY*A and FY*B, which differ by the substitution of a single G to A nucleotide producing a Gly44Asp substitution in the DARC polypeptide. The Fy(a-b-) phenotype corresponds to the Duffy-negative phenotype in which DARC is no longer expressed on RBCs [254], hence providing protection against *P. vivax* invasion [231]. The molecular basis for this was attributed to a single mutation in the upstream sequence of the silent FY*B allele, which has a single T to C substitution at nucleotide 46 in the promoter region of DARC. This mutation disrupts the binding of a GATA1 erythroid transcription factor, stopping the expression of DARC in RBCs [255]. However, this mutation does not appear to affect expression of DARC in non-erythroid tissues, which explains the absence of any abnormality in the regulation of inflammation or chemokine levels [256, 257].

Moreover, it has been shown that people in areas of high transmission of *P. vivax* develop a high-titre of antibodies that block multiple alleles of *Pv*DBP binding to DARC, and this confers protection against *P. vivax* infection [258]. This acquired immunity gives protection to adults who have been repeatedly infected but not to

young children, which is the age group most at risk. In addition, antibodies acquired naturally from repeated exposure to malaria block the binding of recombinant *Pv*DBP-RII *in vitro* [259]. All of these reasons provide a good rationale to base a vaccine on the *Pv*DBP-RII domain.

4.1.5 Pre-clinical and Clinical Trials Using *Pv*DBP-RII

Since *Pv*DBP-RII has been implicated as the key domain required for *P. vivax* invasion into erythrocytes, pre-clinical trials in mice and rabbits using *Pv*DBP-RII-based immunogens were performed. These induced inhibitory antibodies that block binding of recombinant *Pv*DBP-RII to DARC [260–262]. Furthermore, more recent Phase I human trials have shown that using *Pv*DBP-RII as a vaccine leads to strain-transcending inhibitory antibodies, both with *Pv*DBP-RII injected as an adenovirus [129] and directly [53].

4.1.6 Studies of Inhibitory Monoclonal Antibodies Against *Pv*DBP-RII

Gaining information about the structural features of the epitopes of inhibitory antibodies, suggests which regions of the protein should be used in immunogens to induce the production of these desirable antibodies. Inhibitory sera acquired from people in endemic *P. vivax* areas of Papua New Guinea have been screened against a linear peptide array [263] and the inhibitory epitopes have been shown to bind to a similar region to where the $DARC_{19-30}$ peptide binds (subdomain 2) [251]. This suggests that antibodies that directly block the binding of DARC and the dimerisation of *Pv*DBP-RII are necessary for inhibition. Inhibitory mouse mAbs have also been obtained from pre-clinical trials that block DARC binding *in vitro* have had their binding sites mapped. Both structural studies and mass-spectrometry based

approaches suggest that these epitopes lie on subdomains 1 and 3 of *Pv*DBP-RII, which is distant to the DARC binding site on subdomain 2 **(Figure 4.7)** [264]. The way in which these mouse mAbs prevent the binding of *Pv*DBP-RII:DARC and subsequent invasion is not well understood.

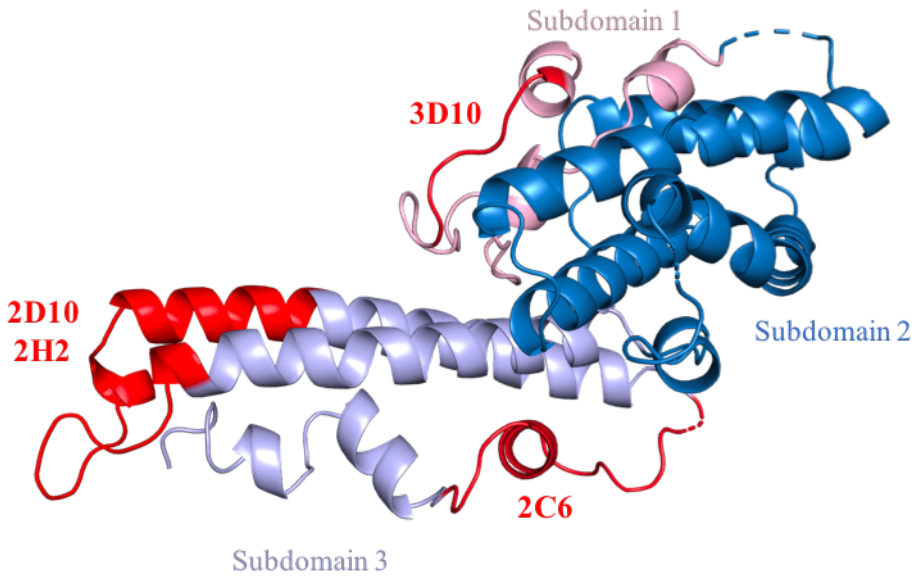


Figure 4.7: Epitopes of Inhibitory Mouse mAbs The epitopes of four inhibitory mouse mAbs mapped by structural studies and mass spectrometry approaches are shown in red. The mouse mAbs 2D10 and 2H2 have overlapping epitopes on the tip of subdomain3 (lilac), whereas the 2C6 epitope is located on the other side of subdomain 3. The 3D10 epitope is located on subdomain1 (pale pink). Epitopes from [264].

There is also currently no available information about how human mAbs bind to *Pv*DBP-RII and inhibit reticulocyte invasion. It is also unknown whether the mouse mAbs and peptide array studies provide a good model for predicting the epitopes of inhibitory human mAbs. Recently, there has been an increase in structure-guided vaccine design, in which the structure of a human inhibitory mAb in complex with its target protein, has been used to identify the epitopes that induce the production of these antibodies. This is then followed by grafting these epitopes onto smaller molecules to focus the immune response to the regions of the protein that lead to the production of the most inhibitory antibodies. This technique has been developed and shown to be successful in targeting HIV-1 [265–267]. In order to use this technique to

design a vaccine against the *Pv*DBP:DARC interaction, the epitope of a human mAb that inhibits this binding and subsequent invasion must be structurally characterised in complex with *Pv*DBP-RII.

4.1.7 Identification of DB9 as an Inhibitory Human mAb

From human Phase I safety trials with a *Pv*DBP-RII viral vector, collaborators in the Jenner Institute, obtained human mAbs. mAbs that inhibited the binding of recombinant *Pv*DBP-RII to recombinant $DARC_{1-60}$ were identified using an ELISA (enzyme-linked immmunosorbent assay). Although, this is a good high throughput method of identifying potentially inhibitory mAbs, it often produces false positives [268]. Therefore, a more stringent identification method is to test the ability of the antibodies to inhibit erythrocyte invasion, using a growth inhibition assay (GIA). Since *P. vivax* cannot be cultured, a *P. knowlesi* system that has had the *Pk*DBP genes replaced with the *Pv*DBP gene, was used [269]. Out of the panel of *Pv*DBP-RII mAbs, this identified three highly neutralising antibodies, namely DB1, DB10, and DB9 **(Figure 4.8A and B)**. To assess for broad reactivity, clinical isolates of *P. vivax* from Thai patients in endemic regions were used to test the ability of these antibodies to inhibit different strains of *P. vivax*. Only one of the three highly neutralising antibodies, DB9, showed high strain-transcendence in ability to potently inhibit (∼65-90% inhibition) reticulocyte invasion in 10/11 of the isolates tested **(Figure 4.8C)** [268].

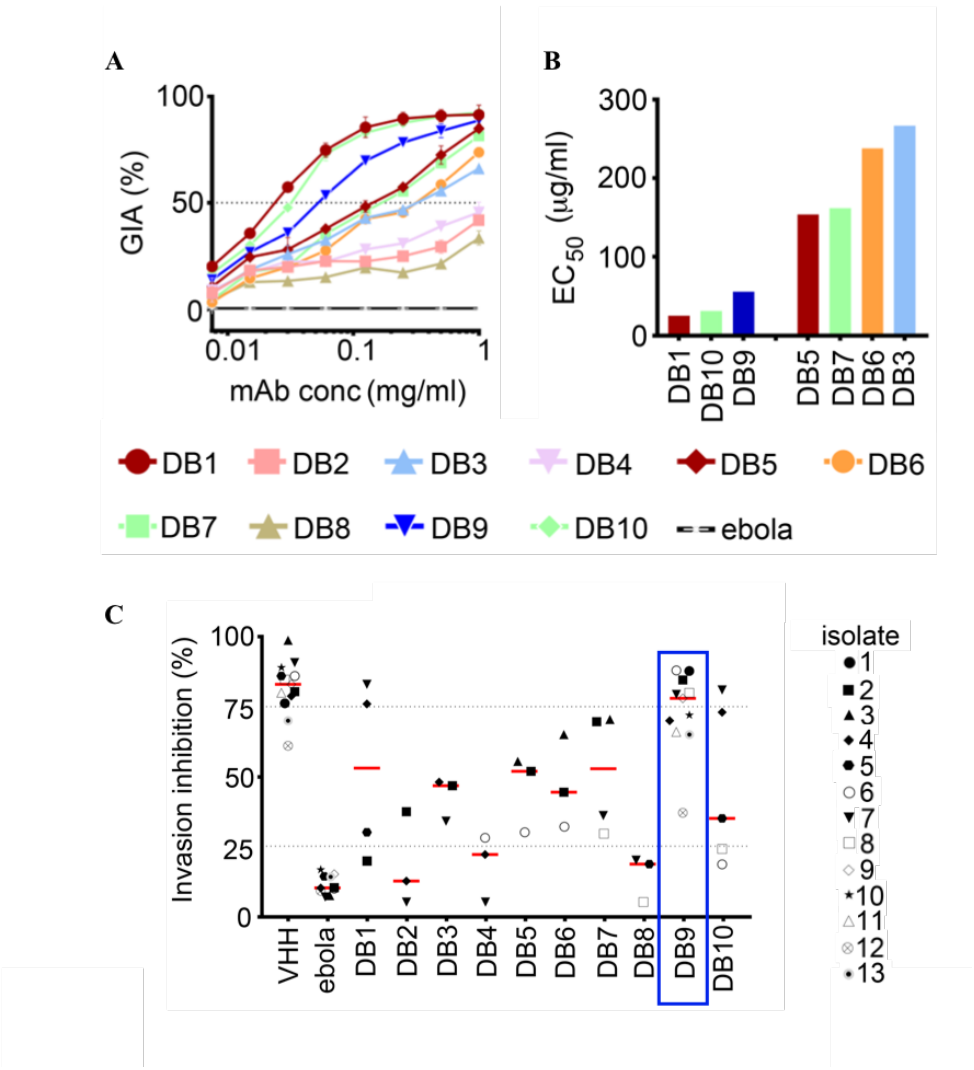


Figure 4.8: Inhibition Analysis of the Panel of Human mAbs A) Assays of growth inhibitory activity (GIA) by the panel of mAbs. Inhibition was tested in a two-fold dilution series starting at 1 mg/ml. Data points represent the mean of three technical replicates, while the error bars represent the standard deviation. B) The EC_{50} values (interpolated by non-linear regression) are shown in ascending rank order for the the seven mAbs which reached >50% GIA at the maximum concentration (1 mg/ml). C) *Plasmodium vivax ex vivo* invasion assays performed on different strains of *P. vivax*. Each data point represents the % inhibition for each antibody. A recombinant human IgG1 anti-*Ebolavirus* mAb (ebola) was used as a negative control at 1 mg/ml. (Adapted from [268]).

4.2 Structural Characterisation of *Pv*DBP-RII Bound to a Broadly Inhibitory Human mAb

4.2.1 Project Aims

- To express and purify the *Pv*DBP-RII:DB9 Fab complex
- To kinetically characterise the binding of DB9 to *Pv*DBP-RII
- To structurally characterise the binding epitope of DB9 Fab on the *Pv*DBP-RII
- To suggest a model which explains how DB9 inhibits the *Pv*DBP:DARC interaction and subsequent invasion of the parasite into erythrocytes.

4.2.2 Declaration

Tom Rawlinson, Draper Group, Jenner Institute, University of Oxford provided us with the most broadly inhibitory human mAb, DB9, from Phase I safety trials. The recombinant *Pv*DBP-RII was cloned and the original purification protocol was optimised by Samuel Gerrard, Higgins Group.

4.2.3 Expression and Purification of *Pv*DBP-RII

The *Pv*DBP-RII gene was previously cloned in the laboratory into the pET15bcon vector backbone **(Materials and Methods Figure 2.1 and Appendix A.3.1)**. This vector contained the DNA sequence for an N-terminal histidine tag plus a TEV-cleavage site for the purification of recombinant protein. *Pv*DBP-RII was expressed in the *E.coli* strain BL-21, almost entirely in the insoluble inclusion body. The

original purification protocol of *Pv*DBP-RII required multiple rounds of optimisation to develop the final protocol. Some methods that were altered in an attempt to improve the yield of the *Pv*DBP-RII purification included:

- Solubilisation of the cell pellet directly into guanidine vs inclusion body purification protocol
- Buffer exchange out of arginine refolding buffer by TFF vs dialysis
- Refolding by flash dilution vs on-column gradient refold
- Nickel purification in guanidine before refolding vs nickel purification after refolding

The finalised *Pv*DBP-RII purification protocol is outlined in detail in **Materials and Methods 2.4.4.1**. In short, after the lysis of the *E. coli* cell pellet, the insoluble fraction was thoroughly washed to remove any membranes and cellular components, before being solubilised in 6 M guanidine hydrochloride. The purified inclusion body was flash diluted into a buffer containing 400 mM arginine, 20 mM EDTA, 3 mM glutathione reduced and 0.3 mM glutathione oxidised and left to refold for 20 hours at 4°C. The folded *Pv*DBP-RII was buffered exchanged by TFF and incubated on Ni-NTA beads, with soluble His-tagged *Pv*DBP-RII eluted in high imidazole buffer. The elution fractions were then concentrated and buffer exchanged by dialysis into PBS, before overnight cleavage of the His tag with TEV protease. The cleaved *Pv*DBP-RII was separated from the His-tagged TEV and any uncleaved *Pv*DBP-RII by reapplying to Ni-NTA. The flow-through was then applied on a S200 10/30 (GE Healthcare) size exclusion column **(Figure 4.9)**.

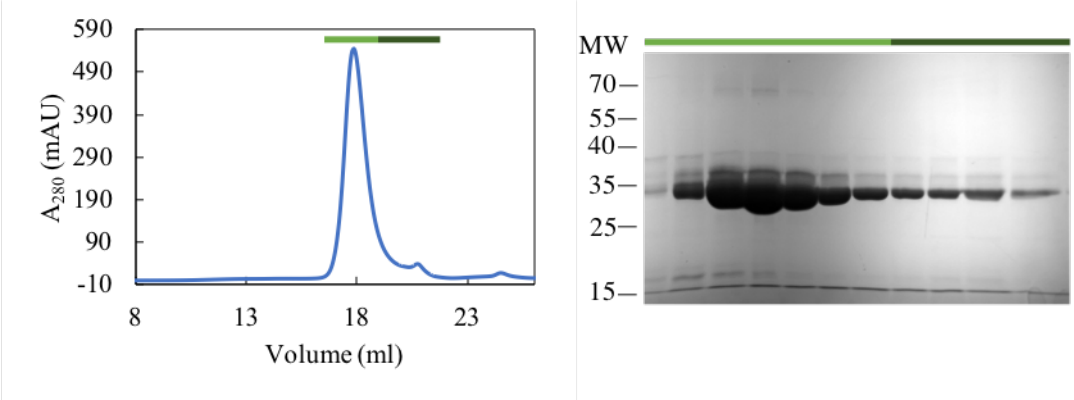


Figure 4.9: *Pv*DBP-RII Purification The green bars represent the corresponding fractions of the S200 10/30 run in the size exclusion trace (left) and the SDS-PAGE gel (right). The light green bar represents the fractions that were pooled for further studies.

4.2.4 Preparation of DB9 Fab Fragment

The human mAb, DB9, raised against *Pv*DBP-RII, was given to us by our collaborators at the Jenner institute. The DB9 mAb belonged to isotype class IgG1. For crystallisation, only the Fab fragment of the antibody was required, since this is the part of the antibody that binds to the antigen. Also Fab fragments have much reduced flexibility compared to total antibodies, and so are more likely to crystallise. Papain protease was chosen to cleave the mAb into Fab fragments, as it is routinely used for the preparation of Fab fragments from IgG1. The DB9 was concentrated to 2 ml (at ∼2 mg/ml) and incubated with papain protease at 37°C overnight, and the mAb was cleaved into Fab and Fc regions of the antibody **(Materials and Methods 2.5.2)**. The Fab and Fc portions were then separated by incubating the mixture with protein A agarose beads that specifically bind to the Fc region of antibodies. SDS-PAGE analysis confirmed that the Fab portion of the mixture was present in the flow-through and wash, whilst the Fc and any uncleaved mAb bound to the protein A and was present in the elution fraction **(Figure 4.10 A)**. The protein A flow-through and wash was then concentrated to 1 ml, and applied to SEC column S200 10/30 (GE Healthcare) to remove any aggregates **(Figure 4.10**

B and C).

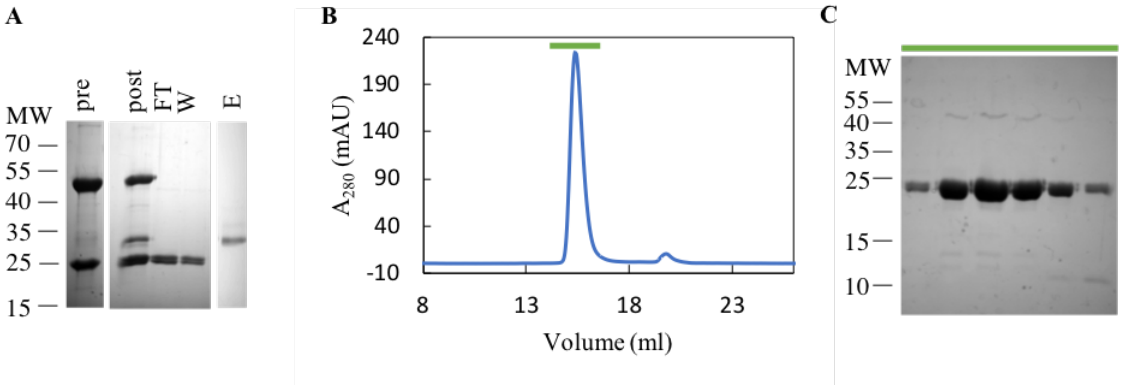


Figure 4.10: DB9Fab Purification A) Papain cleavage and subsequent protein A of DB9. Labels are explained as follows: pre=pre-papain cleavage, post = post papain cleavage, FT = protein A flow-through, W = protein A wash, W end = end of protein A wash, E = protein A elution. B) SEC trace of the cleaved DB9Fab from the FT and W fractions of A). C) Reduced SDS-PAGE gel of the DB9Fab fractions from the SEC. The green bar shows the corresponding fractions on the trace and the gel and the whole peak was pooled.

4.2.5 Assessing the Binding of *Pv*DBP-RII to DB9

In order to check that the refolded *Pv*DBP-RII was functional, its binding to the DB9 mAb was assessed by SPR. The DB9 mAb was captured on a protein A-coupled CM5 chip (**Materials and Methods 2.6.2**), before a two-fold dilution series of *Pv*DBP-RII (starting at 1 μM) was tested for binding (**Figure 4.11**). The kinetic curves were fit using the BIAevaluation software, assuming a 1:1 interaction, to determine the kinetic constants and the affinity of the interaction **(Table 4.1)**. This showed that *Pv*DBP-RII bound DB9 with a 2.63 nM affinity.

	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})	K_d (nM)
***Pv*DBP-RII**	$3.71x10^5$	$9.76x10^{-4}$	2.63

Table 4.1: Kinetic Analysis of *Pv*DBP-RII Binding to DB9

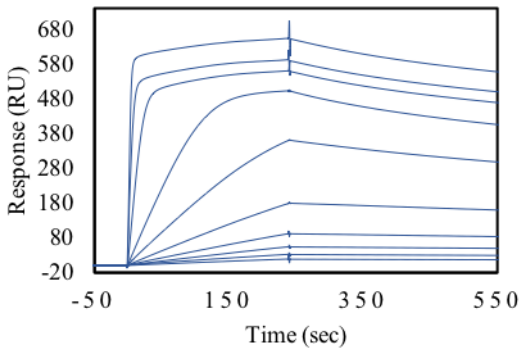


Figure 4.11: Binding of *Pv*DBP-RII to DB9 by SPR DB9 was bound to a protein A-coupled CM5 SPR chip. A two-fold dilution series of *Pv*DBP-RII from 1 μM was flowed over and binding to DB9 assessed.

4.2.6 Crystallisation of the *Pv*DBP-RII:DB9Fab Complex

The purified *Pv*DBP-RII and DB9Fab were mixed in a 1.2:1 molar ratio and incubated at RT for 1 hour. The complex was purified using SEC S200 16/60 (GE Healthcare) in 20 mM Hepes pH7.5, 150 mM NaCl **(Figure 4.12)**. The complex peak was pooled and concentrated to 12 mg/ml.

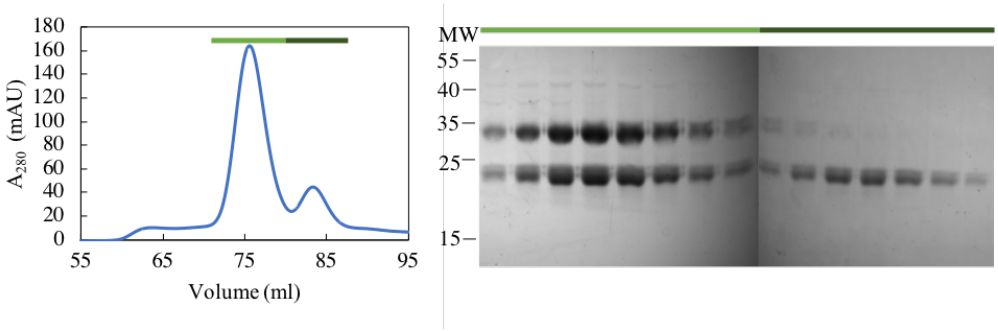


Figure 4.12: *Pv*DBP-RII:DB9Fab Complex Formation The green bars represent the corresponding fractions of the S200 10/30 run in the size exclusion trace (left) and the reduced SDS-PAGE gel (right). The light green bar represents the fractions that were pooled for further studies.

Broad crystallisation trials were set up for the *Pv*DBP-RII:DB9Fab complex to maximise the chance that crystallisation would occur **(Materials and Methods**

2.7.2.1). Crystals grew as needles in a reservoir solution of 45% v/v polypropylene glycol 400, 10% v/v ethanol. The crystals were washed in a drop of well solution and were then cryo-cooled in liquid nitrogen for storage and data collection. (No glycerol needed to be added as a cryoprotectant, since this condition already contained the cryoprotectant polypropylene glycol 400).

4.2.7 Data Collection and Solution of the *Pv*DBP-RII:DB9-Fab Complex

Data were collected for the *Pv*DBP-RII:DB9Fab complex at the Diamond Light Source on beamline I-03. A full dataset was collected, using a collection strategy suggested in the outputs of EDNA and Mosflm from the three initial diffraction images. The data were processed automatically by FastDP and Xia2, which both suggested a space group of $P6_222$. This allowed collection of a complete dataset to 3.00Å, with a signal-to-noise ratio in the outer shell: $I/\sigma I = 1.20$. The R_{pim} value in the outer shell was 0.431, within the generally accepted limits for data quality **(Table 4.2)**.

Xia2 was used to automatically process the data, with Mosflm used for indexing and Scala for scaling the data **(Materials and Methods 5.2.5)**. The Matthews coefficient estimation predicted that there was one copy of the *Pv*DBP-RII:DB9Fab complex per asymmetric unit. Molecular replacement in Phaser was used to get phase information for this dataset. The structures of the variable and constant domains of a human Fab (PDB: 3DIF) were used as separate search models. The top result for this search showed one Fab molecule with the variable and constant domains correctly put back together in the unit cell. With the Fab molecules fixed in position, a second molecular replacement search was carried out using the structure of *Pv*DBP-RII (PDB 4NUV) with all loops and helices removed except the core

Space Group	$P6_222$
Cell Parameters (Å)	a=b=173.56, c= 169.73
(°)	α=β=90, γ=120
Resolution (Å)	84.94-3.00 (3.09-3.00)
R_{meas}	0.177 (1.93)
R_{pim}	0.0399 (0.431)
I/σ (I)	12.8 (1.20)
$CC_{1/2}$	0.998 (0.696)
Completeness (%)	99.8 (99.9)
Multiplicity	19.6 (20.1)

Table 4.2: Data collection statistics for the *Pv*DBP-RII:DB9Fab Structure (Values in parentheses are the highest-resolution shell).

four helices of subdomain 2 (residues 265-291, 297-315, 324-336, and 342-363). One molecule of *Pv*DBP-RII was found, with extra density consistent with the loops and helices that had been removed in the search model. After multiple rounds of refinement and rebuilding, using Buster and Coot, respectively, all of the helices and the loops were resolved. Final solution parameters are shown in **(Table 4.3)**.

The higher value of R_{free} (0.221) compared to R_{work} (0.177) shows that independent, unprocessed reflections are explained by the model and suggests that the model did not over-interpret the data. Moreover, the Ramachandran plot showed 95% of residues in the favoured region, 5% in the allowed region and 0% in the disallowed region, again indicating that this is the correct structure **(Figure 4.13)**.

Resolution (Å)	84.94-3.00
R_{work}	0.177
R_{free}	0.221
rmsd bond lengths (Å)	0.01
rmsd bond angles (°)	1.24
Ramachandran plot	
Favoured region (%)	95.0
Allowed region (%)	5.0
Disallowed region (%)	0.0

Table 4.3: Refinement statistics of the *Pv*DBP:DB9 complex

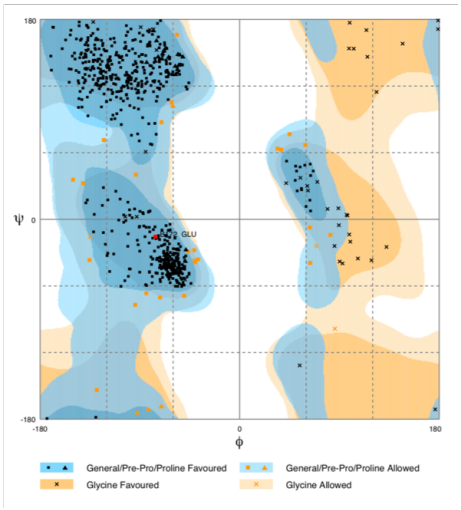


Figure 4.13: Ramachandran Plot of *Pv*DBP:DB9 Structure

4.2.8 Analysis of the PvDBP-RII:DB9Fab structure

As can be seen in **Figure 4.14**, DB9Fab interacts solely with subdomain 3 of *Pv*DBP-RII, which is distal to the dimerisation interface and DARC binding site on subdomain 2. Therefore, it is unlikely that this highly inhibitory mAb inhibits invasion by directly blocking DARC binding and subsequent dimerisation.

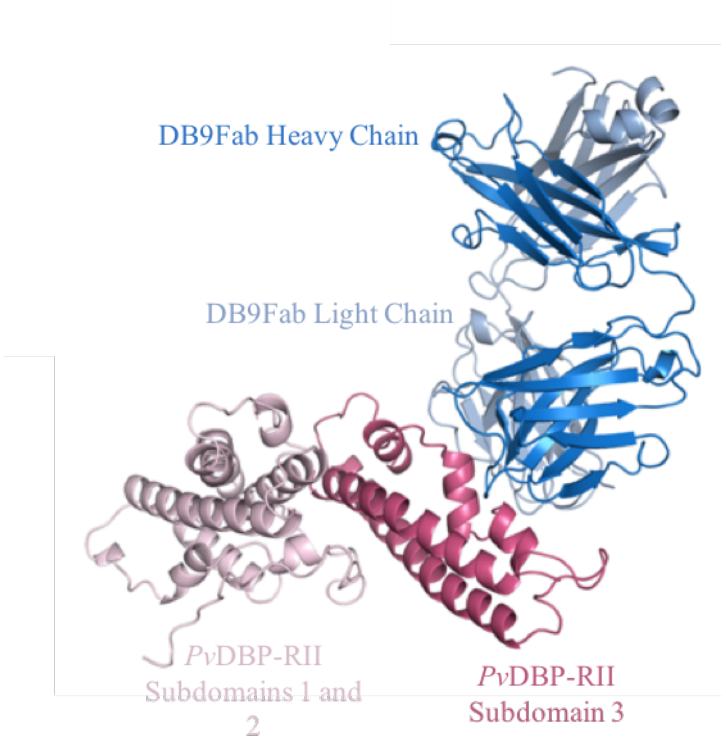


Figure 4.14: Structure of *Pv*DBP-RII bound to DB9 Fab The DB9Fab heavy chain is shown in dark blue, and the light chain in pale blue. The *Pv*DBP-RII is shown in pink, with the subdomain 3 highlighted in dark pink, and the rest of the domain shown in light pink.

There are eighteen direct contacts made between the *Pv*DBP-RII and DB9Fab **(Table 4.4)**. 12/18 of these are hydrogen bonds made between the heavy chain CDR (complementarity determining region) loops of the DB9Fab and the *Pv*DBP-RII. The remaining six contacts consist of three hydrogen bonds, two hydrophobic interactions, and one cation-pi interaction between the light chain CDR loops and the *Pv*DBP-RII. Also, 15/18 contacts on the *Pv*DBP-RII are made via the side chains as opposed to the main chain.

Looking at the DB9 epitope on *Pv*DBP-RII, it shows that this region does not have any polymorphisms and so is highly conserved across the different strains of *P. vivax* **(Figure 4.15)**. This explains the strain-transcendence of DB9 and also indicates that this region would be a good vaccine candidate as it seems resistant to selection pressure. From the structure it was determined that most of the contacts (15/18)

PvDBP-RII residue	Group	Fab Chain	Residue	Group	Interaction
K412	Side Chain	Heavy Chain	Y74	Side Chain	Hydrogen bond
D476	Side Chain	Light Chain	S75	Side Chain	Hydrogen bond
E484	Side Chain	Heavy Chain	E20	Main Chain	Hydrogen bond
E484	Side Chain	Heavy Chain	G45	Main Chain	Hydrogen bond
N486	Side Chain	Heavy Chain	R118	Side Chain	Hydrogen bond
E487	Side Chain	Light Chain	Y68	Side Chain	Hydrogen bond
E487	Side Chain	Light Chain	S75	Main Chain	Hydrogen bond
V488	Side Chain	Light Chain	Y68	Side Chain	Hydrophobic
V488	Side Chain	Light Chain	L74	Side Chain	Hydrophobic
N492	Side Chain	Heavy Chain	T121	Main Chain	Hydrogen bond
E493	Side Chain	Heavy Chain	T52	Side Chain	Hydrogen bond
R497	Side Chain	Light Chain	W51	Side Chain	Cation-pi
R497	Main Chain	Heavy Chain	T121	Side Chain	Hydrogen bond
R497	Side Chain	Heavy Chain	T121	Main Chain	Hydrogen bond
G499	Main Chain	Heavy Chain	T52	Main Chain	Hydrogen bond
A500	Main Chain	Heavy Chain	S51	Main Chain	Hydrogen bond
E503	Side Chain	Heavy Chain	Y54	Side Chain	Hydrogen bond
E503	Side Chain	Heavy Chain	Y79	Side Chain	Hydrogen bond

Table 4.4: Interaction Table for *Pv*DBP-RII binding to DB9

were made with the side chain rather than the main chain. The ability of DB9 to bind to many different strains is therefore consistent with the lack of polymorphism in this region of the sequence.

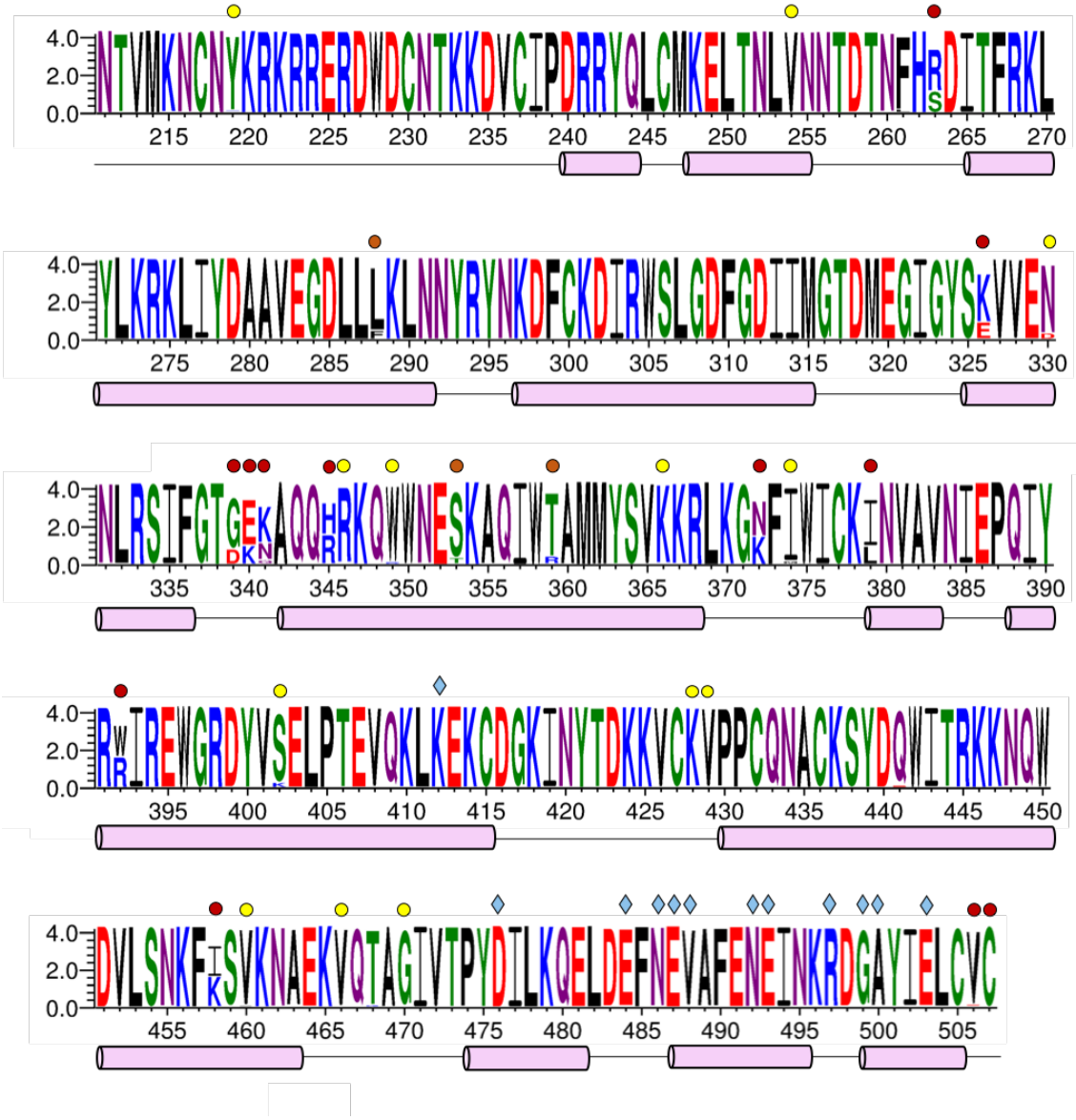


Figure 4.15: Sequence logo representing sequence conservation across *Pv*DBP-RII Sequence logo derived from 383 sequences of *Pv*DBP-RII from *Plasmodium vivax* isolates. Underneath the logo is the residue number from the SalI *Pv*DBP-RII variant. Cylinders represent the location of helices while lines represent loops. Above the sequence, blue kites indicate residues which directly contact DB9. Yellow, orange and red circles represent residues with sequence entropies of 0.15-0.3, 0.3-0.45 and >0.45 respectively. The sequence logo was made with WebLogo [212].

4.2.9 Discussion

Elucidating the epitopes of broadly inhibitory mAbs improves our understanding of the regions of the protein required to produce the most desired antibodies [130]. Here, the broadly inhibitory human mAb, DB9, is shown to bind distal to both the predicted DARC binding site and dimerisation interface, on the non-polymorphic subdomain 3. Docking the DB9Fab onto the structure of the dimer of *Pv*DBP-RII bound to the $DARC_{19-30}$ peptide, shows that the DB9 emerges from the same side of the *Pv*DBP-RII as the C-terminus of DARC **(Figure 4.16)**. This raises the possibility that DB9 may sterically prevent the *Pv*DBP from approaching a DARC-coated surface, thereby blocking receptor engagement and subsequent invasion.

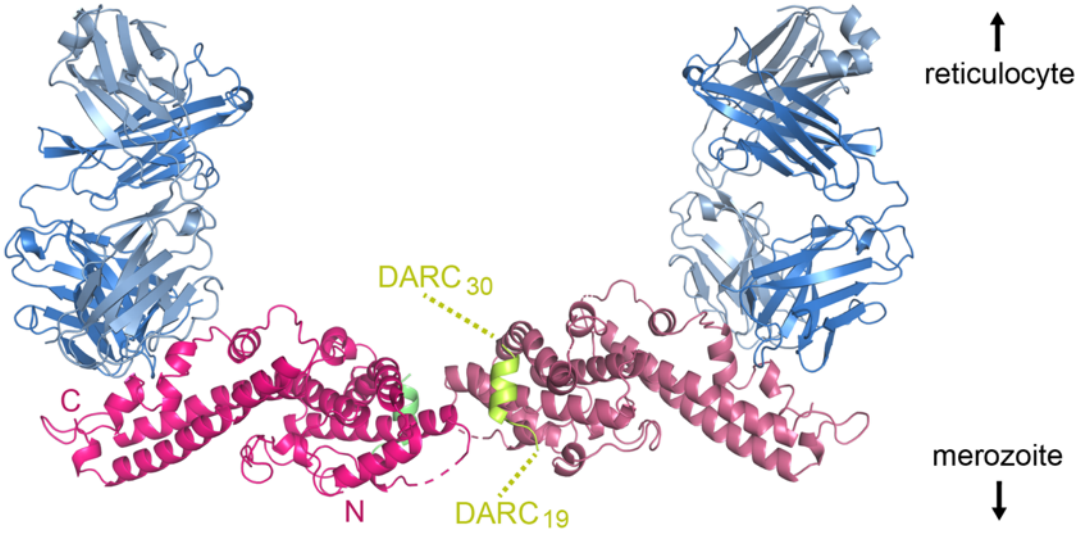


Figure 4.16: Overlay of *Pv*DBP-RII:DB9Fab Structure on the *Pv*DBP-RII:DARC Heterotetramer The *Pv*DBP-RII:DARC heterotetramer structure from [251] is shown in pinks and greens, respectively. The DB9 Fabs are positioned by overlaying the *Pv*DBP-RII domains with our structure.

This steric occlusion hypothesis by DB9 is supported by the structural and epitope mapping data of inhibitory mouse monoclonal mAbs [264]. The inhibitory potential of these mAbs was only based on a recombinant protein-based assay, which has been shown to only partially indicate the efficacy in parasite-based neutralisation assays

[268]. However, the epitope of one of the mAbs, 2C6, is adjacent to that of DB9 and may sterically prevent DARC receptor engagement via a similar mechanism to DB9 **(Figure 4.17)**. The 2D10 epitope appears on the opposite side of subdomain 3 to the DB9 but in the same plane and so could prevent reticulocyte invasion again by sterically occluding *Pv*DBP-RII:DARC interaction and/or the subsequent association of the two membranes **(Figure 4.17)**.

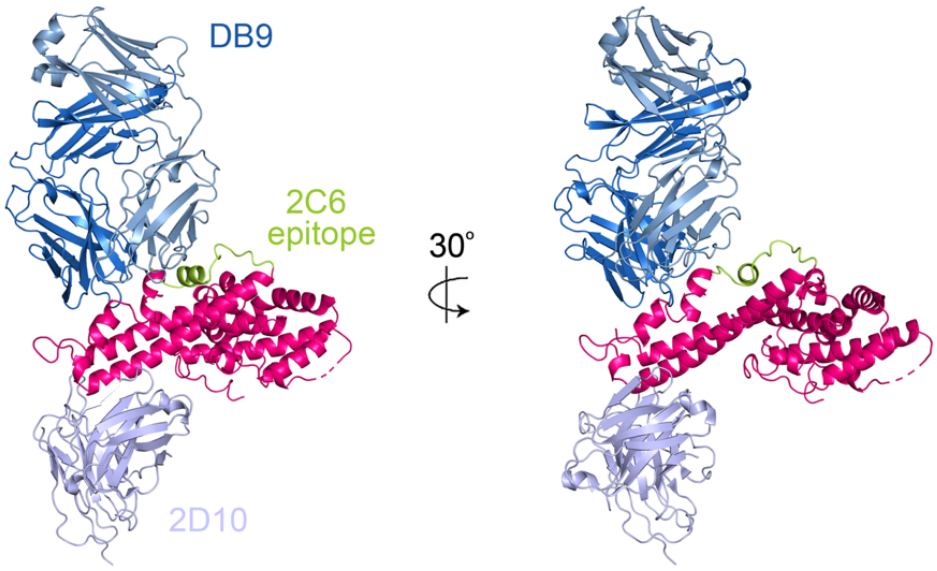


Figure 4.17: Comparison of DB9 Epitope with Inhibitory Mouse mAbs Two views of the DB9Fab (blue) bound to the *Pv*DBP-RII (pink) are shown with the epitope of the 2C6 mAb highlighted (green) and the Fv structure of the mAb 2D10 shown (lilac).

It has previously been thought that the most inhibitory antibodies against *Pv*DBP would directly block the binding of DARC or inhibit dimerisation [245, 270]. This has been recently supported by identification of inhibitory human mAbs that bind at the dimer interface [271], at a site which also lacks polymorphisms. Here, an alternative non-polymorphic region of *Pv*DBP-RII has been identified to target with future vaccine design. As a result of this research, an immunogen based on subdomain 3 may aid in the production of broadly inhibitory mAbs, whilst reducing the production of non-strain transcending mAbs, which are induced by other more polymorphic regions of *Pv*DBP-RII.

4.3 Triple Complex Formation of *Pv*DBP-RII with DB9Fab and DARC ectodomain

4.3.1 Rational

The only structural information available regarding DARC binding to *Pv*DBP-RII, is obtained from the structure of *Pv*DBP-RII bound to a $DARC_{1-60}$ ectodomain fragment. However, the sixty amino acid DARC fragment was highly flexible so the majority of it could not be seen in the electron density. Moreover, the DARC fragment used in this study was not sulphated, which has been shown to be crucial for binding to *Pv*DBP-RII. For these reasons, it is likely that the proposed model of DARC binding to *Pv*DBP-RII is not the full story. Since diffracting crystals of *Pv*DBP-RII:DB9Fab had been produced and the DB9Fab was distal to the proposed DARC binding site, it was thought that a triple complex could be formed using a DARC ectodomain fragment. The triple complex, with the presence of the DB9Fab may aid crystallisation, allowing us to elucidate further the mechanism of DARC binding to *Pv*DBP-RII.

4.3.2 Project Aims

- Express and purify sulphated $DARC_{1-60}$
- Form a triple complex of *Pv*DBP-RII:DB9Fab:$DARC_{1-60}$
- Use the DB9Fab as a crystallisation chaperone to obtain the structure of sulphated $DARC_{1-60}$ in complex with *Pv*DBP-RII

4.3.3 Declaration

The DARC construct had been previously cloned in the laboratory by a former student, Caroline Lenz, who also showed that the construct expressed in HEK cells was successfully sulphated at both tyrosines by native mass spectrometry. A summer student, Georgina Benn, expressed and purified the $DARC_{1\text{-}60}$ ectodomain under my supervision.

4.3.4 DARC Ectodomain Cloning, Expression and Purification

The sixty residue ectodomain **(Appendix A.3.2)** had previously been subcloned into the pHLSec vector, which encodes an N-terminal six-histidine tag separated from the recombinant protein via a TEV cleavage for purification **(Materials and Methods Figure 2.4)**. This DARC construct was transfected into Expi293S cells as described in **Materials and Methods 2.2.4**. The supernatant was harvested four days after transfection and the DARC ectodomain expressed very well, allowing ∼10 mg of protein to be purified from 1 L of culture.

The DARC ectodomain was purified on a Ni-NTA column via its histidine tag **(Materials and Methods 2.4.3)**. The protein eluted on the nickel column contained $DARC_{1-60}$ in different glycosylation states as indicated by the multiple bands on the SDS-PAGE gel **(Figure 4.18)**. In order to make the DARC protein homogeneous, it was buffer exchanged into PBS and treated with excess PNGaseF to remove any N-linked glycosylations present **(Materials and Methods 2.4.7)**. As can be seen in **Figure 4.18**, this resulted in the disappearance of the upper bands, with only fully deglycosylated $DARC_{1-60}$ present. (During future purifications a much smaller amount of PNGaseF (1:100 molar ratio to recombinant protein) was used

so that it could be separated from the deglycosylated DARC more easily by size exclusion).

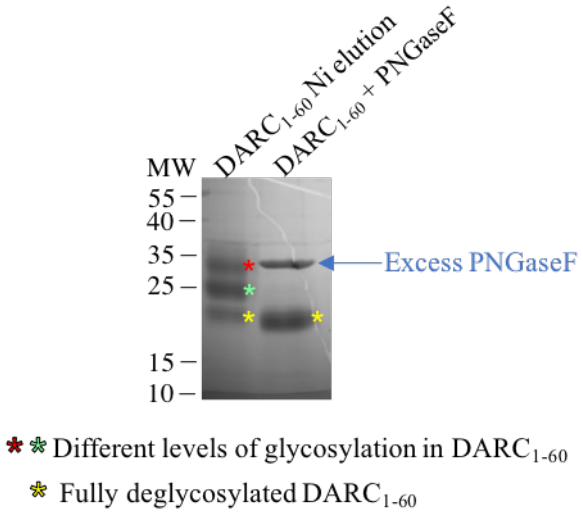


Figure 4.18: Deglycosylation of DARC$_{1-60}$ Ectodomain An SDS-PAGE gel to show the DARC$_{1-60}$ sample before and after treatment with excess PNGaseF. The bands representing different levels of glycosylation are highlighted by red and green asterisks, with the fully deglycosylated state represented by a yellow asterisk.

The DARC construct was then further purified via SEC (S75 16/60, GE Healthcare) to remove any aggregates present **(Figure 4.19)**.

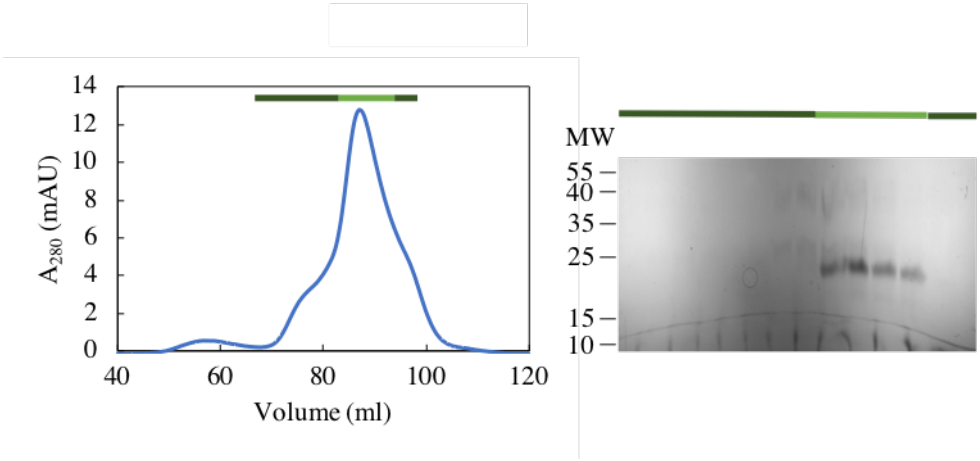


Figure 4.19: Purification of DARC$_{1-60}$ Ectodomain The size exclusion chromatography trace (left) and corresponding SDS-PAGE gel (right) of the triple complex formation. The coloured bars demonstrate which part of the peak represents which region of the gel. With the light green bar showing the fractions pooled for crystallography.

4.3.5 Crystallisation of DARC:PvDBP-RII:DB9Fab Triple Complex

Purified *Pv*DBP-RII, DB9Fab and DARC$_{1\text{-}60}$ were mixed in a 1:1:1 ratio and incubated at RT for 1 hour before being applied to size exclusion chromatography (S200 16/60, GE Healthcare) **(Figure 4.20)**.

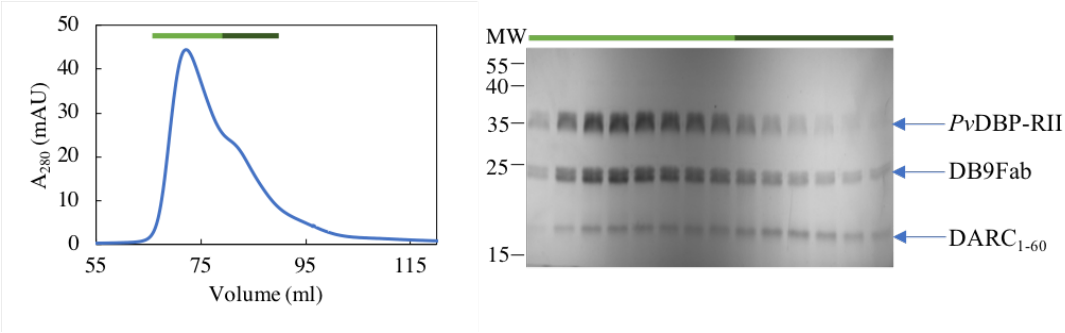


Figure 4.20: *Pv*DBP-RII:DB9Fab:DARC Complex Formation The size exclusion chromatography trace (left) and corresponding SDS-PAGE gel (right) of the triple complex formation. The coloured bars demonstrate which part of the peak represents which region of the gel. With the light green bar showing the fractions pooled for crystallography.

The peak corresponding to the triple complex was pooled and concentrated to 10

mg/ml. Broad crystallisation screens (ProPlex, PACT, MIDAS, Morpheus and JCSG+) were set up at 4°C as described in **Materials and Methods 2.7.2.1**. Cuboid crystals formed in two conditions: 0.2 M sodium acetate, 0.1 M sodium citrate, pH 5.5, 10% w/v PEG 4000; and 0.2 M ammonium acetate, 0.1 M sodium acetate, pH 4.0, 15% w/v PEG 4000. In order to check that these crystals contained all three components of the triple complex, a few crystals were harvested, washed, and run on an SDS-PAGE gel and the gel silver stained **(Materials and Methods 2.3.1)**. As can be seen in **Figure 4.21** all three components of the complex were present in the crystal. Therefore, silver bullet optimisation screens were set up with and without seeds as described in **Materials and Methods 2.7.2.2** to produce more crystals. The crystals were harvested and cryo-protected by transferring into a solution of their mother liquor containing 25% glycerol, before being cryo-cooled and stored in liquid nitrogen.

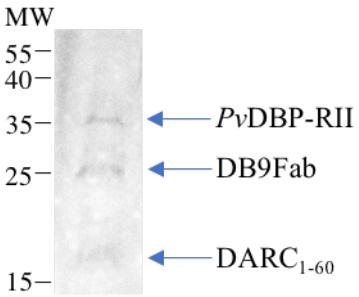


Figure 4.21: Crystals contain *Pv*DBP-RII:DB9Fab:DARC Complex An SDS-PAGE gel of the harvested and washed crystal, which had been silver stained.

Data was collected at Diamond Light Source, on Beamline I-03. Initially the cryocooled crystals were mounted on the goniometer and shot at three angles to check for diffraction quality and resolution limits. Unfortunately all of the crystals that were screened only showed diffraction spots to around 7Å, which was too low resolution for a large dataset collection, and subsequent structure solution **(Figure 4.22)**.

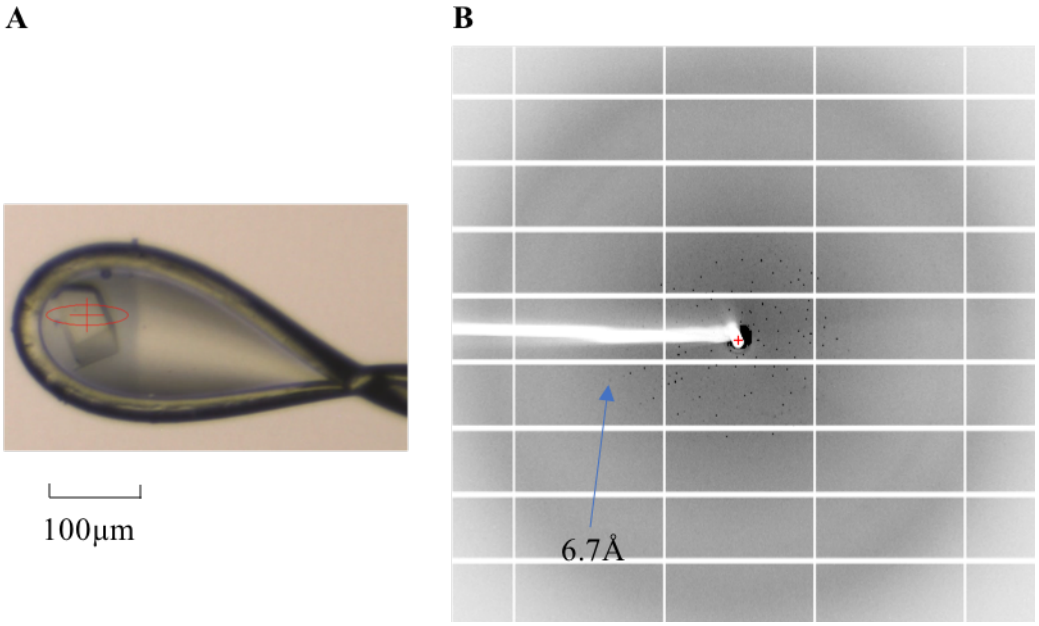


Figure 4.22: *Pv*DBP-RII:DB9Fab:DARC Crystal in Loop and Diffraction Image A) An example of a cuboid crystal mounted in the loop. B) An example of a diffraction image, with spots seen to 6.7Å.

Initial processing of these images by Xia2, suggested a space group of P1 with a large unit cell that could contain multiple copies of the triple complex **(Table 4.5)**. However, since $\alpha \approx \beta \approx 90$°and $\gamma \approx 120$°, which suggest a trigonal (P3) or hexagonal (P6) space group, the automatic indexing of this low resolution dataset was not convincing.

Space Group	P1
Cell Parameters (Å)	a=123.47 b=138.61, c=237.84
(°)	α=91.38 β=95.61, γ=116.31

Table 4.5: Space Group and Unit Cell Parameters for *Pv*DBP-RII:DB9Fab:DARC Complex

4.3.5.1 Further Optimisation of Crystallisation

In an attempt to improve the diffraction resolution of the triple complex, the crystal growth was optimised. One reason that the resolution might have been limited

was that not all of the complexes contained $DARC_{1-60}$. Due to the small size of the $DARC_{1-60}$ fragment, this heterogeneity may not have been seen on the SDS-PAGE gel or the SEC trace. The heterogeneity in the crystal might have led to the low resolution diffraction. Therefore, plates of the *Pv*DBP-RII:DB9Fab:$DARC_{1-60}$ complex were set up, with excess $DARC_{1-60}$ added *in situ* as the plate was made. In order to keep the concentration of the complex as high as possible to promote crystallisation, the drops were set up as follows: 100 nl of complex, 50 nl well condition and 50 nl DARC at 10 mg/ml in the well condition. In this way the complex would be completely saturated with $DARC_{1-60}$. In addition, new plates were set up where the triple complex had been mixed in a 2:1:1 molar ratio of $DARC_{1-60}$:*Pv*DBP-RII:DB9Fab and then further purified by SEC to remove the excess DARC. Neither of these attempts improved the crystal packing and so the resolution of the datasets collected remained similar.

Low resolution can also be caused by flexibility in the protein and so in order to remove any flexible loops or regions, *in situ* limited proteolysis was attempted. The theory of limited proteolysis is that the cleavage sites of any flexible sites will be exposed and so can be nicked, whereas the sites in the folded, stable part of the protein would be protected from cleavage. In this way, the flexible regions can be digested whilst leaving the overall core structure unchanged, potentially aiding crystallisation. Therefore, in the conditions in which crystals already had formed, crystal plates were set up with different ratios of different proteases **(Figure 4.23 A)**. Although needle shaped crystals formed in many drops of these screens, despite seeding and using different ratios of protein, the crystals could not be grown to a size suitable for data collection **(Figure 4.23 B)**.

Lastly, from the previously published structure, the only part of the $DARC_{1-60}$ ectodomain fragment that was ordered and not flexible was $DARC_{19-30}$. Using this information, and the evidence that Tyr41 is required to be sulphated for binding

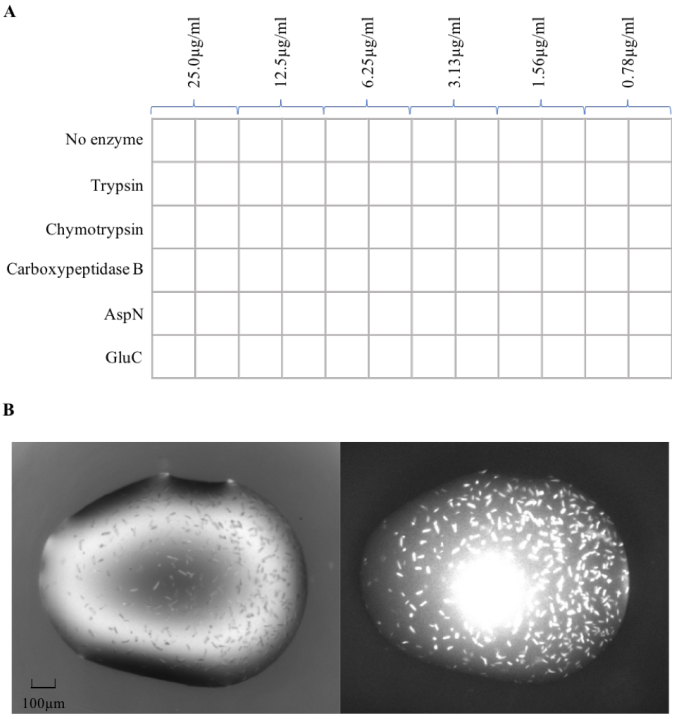


Figure 4.23: Limited Proteolysis Screen for Crystal Optimisation A) An overview of the 72-well screen set up with five different proteases at six concentrations in duplicate, with the "no enzyme" row as a negative control. B) An image of a drop containing needle microcrystals (left) and the UV image of the same drop (right).

to *Pv*DBP-RII, new constructs were designed that removed the potentially flexible N-terminal part of the DARC ectodomain; $DARC_{18-50}$, $DARC_{18-55}$, and $DARC_{18-60}$. These constructs were successfully cloned into pHLSec and transfected into Expi293 cells. However, the expression levels were very low and only a small amount of protein could be purified (between 50 μg and 200 μg per 200 mL of culture) **(Figure 4.24)**. There was not enough protein for a complex between *Pv*DBP-RII and these shortened DARC ectodomains to be formed. Further expression optimisation and/or larger scale purification in the future should be carried out to enable crystallisation screens with these constructs to be set up.

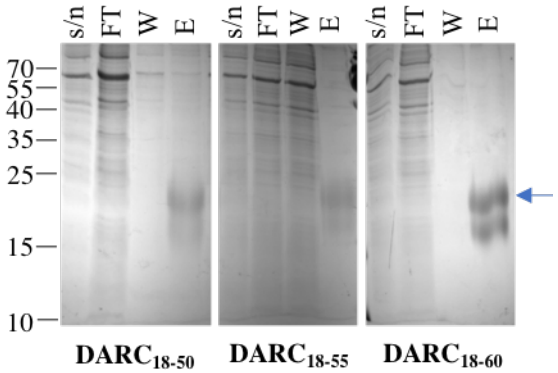


Figure 4.24: SDS-PAGE Gels of Ni-NTA Purification of New DARC Constructs The Ni-NTA purification of the three DARC constructs is shown with s/n=supernatant, FT=flow-through, W=wash, and E=elution. The DARC constructs visible in the elution fractions are highlighted by the blue arrow. The multiple bands present are likely due to different glycosylation states of the DARC constructs.

4.3.6 Discussion of the Triple Complex Attempt

DARC is thought to bind at the dimerisation interface of *Pv*DBP-RII, however, the structural data that supports this is not complete and may not reveal the whole story [251]. It was thought that using the DB9Fab, as a crystallisation chaperone, the triple complex of $DARC_{1-60}$:*Pv*DBP-RII:DB9Fab could be used to resolve more of the $DARC_{1-60}$ residues involved in the *Pv*DBP binding. Although crystals formed, they did not diffract past 7Å resolution. Saturation of the crystals with DARC, limited proteolysis and shorter DARC construct design were attempted in order to improve crystal packing and diffraction resolution, however, these proved unsuccessful. One reason this may have not worked could be due to the fact that the *Pv*DBP-RII:DB9Fab crystallised in a monomeric form, whereas DARC is thought to bind at the dimer interface. This would mean that despite only adding a small peptide, the overall crystal packing in the triple complex might have complete change from the form seen in the double *Pv*DBP-RII:DB9Fab complex.

Moreover, the ability to perform very broad crystallisation screens was limited by the amount of *Pv*DBP-RII that could be produced. Ideally, the limited proteolysis

would have been performed over broad crystallisation screens rather than over the already optimised screens. Furthermore, a greater variety of fine optimisation screens would have been formed. In addition, due to time constraints, it was not possible to optimise the expression of the shortened DARC constructs but this could be repeated in the future. Both double complexes with *Pv*DBP-RII and triple complexes with *Pv*DBP-RII:DB9 could be formed in order to try to gain insight into DARC binding and particularly the role of the sulphated Tyr41.

Although these shortened DARC ectodomain fragments may be useful in crystallography and understanding individual residues involved in binding, this will not provide us with a complete understanding of DARC binding to *Pv*DBP in a cellular context. In order to further elucidate how this interaction leads to reticulocyte invasion, structural information on how either the full length proteins interact, or at least larger portions, is needed. Due to the size of the overall dimeric complex, this could potentially be studied in the future by cryoEM to gain further insight.

4.4 Structure-Guided Immunogen Design

4.4.1 Rational

The structure of *Pv*DBP-RII:DB9Fab identified subdomain 3 as a potential new vaccine candidate, since all interacting residues between the *Pv*DBP-RII and the broadly inhibitory DB9 mAb were located in this non-polymorphic subdomain of *Pv*DBP-RII. Therefore, smaller immunogens based on subdomain 3 were designed in order to focus the immune response to produce inhibitory antibodies against this region.

4.4.2 Project Aims

- Design smaller immunogens that maintain the DB9 epitope from *Pv*DBP-RII
- Clone, express and purify the immunogens
- Characterise the fold and the binding of the immunogens to DB9
- Structurally characterise the immunogens binding to DB9
- Assess the ability of the immunogens to raise inhibitory antibodies in rabbits

4.4.3 Declaration

The expression and purification of the designed immunogens was done with the help of James Nugent, an undergraduate project student. Also, under my guidance as project supervisor, James Nugent undertook the biophysical and structural characterisation of these recombinant proteins. The rabbit immunisations, isolation of sera and testing of the inhibition of the monoclonal antibodies produced was carried out by the Draper Group, Jenner Institute (University of Oxford).

4.4.4 Design of Subdomain3-based Immunogens

It is hard to predict from the sequence alone whether a subdomain will fold in isolation, without the packing interactions against the other subdomains to stabilise the overall conformation. Therefore, multiple strategies of design were employed to maximise the chances that folded immunogen, with a DB9 epitope could be formed. An overview of these design strategies can be seen in **Figure 4.25**, and are explained in further detail in the following sections.

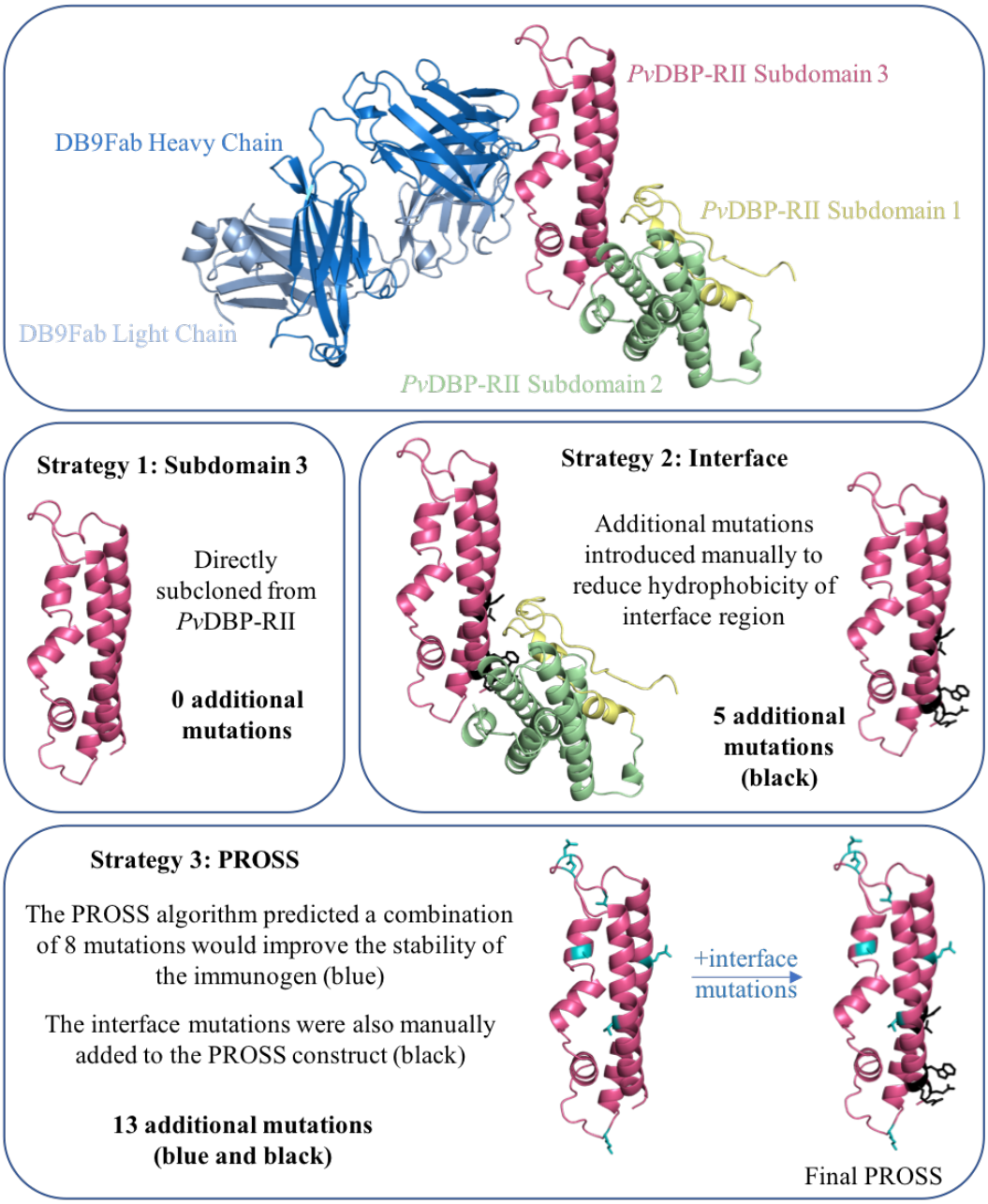


Figure 4.25: Overview of Immunogen Design Strategies Top panel: the *Pv*DBP-RII:DB9 structure with all of the subdomains and chains coloured and labelled. Remaining panels: outline the three strategies, with the mutations introduced highlighted on the structures and in bold text.

4.4.4.1 Strategy 1: Subdomain 3 (SBD3)

The first construct was produced by subcloning the subdomain 3 region (residues 211-508) from the *Pv*DBP-RII sequence **(Strategy 1 in Figure 4.25)**. Boundaries for this construct were modelled using the structure of *Pv*DBP-RII and primers for this subcloning into pET15bcon are detailed in **Appendix A.3.3**.

4.4.4.2 Strategy 2: Interface

Analysing the *Pv*DBP-RII structure, highlighted that the subdomain 3 made hydrophobic contacts and hydrogen bonds with the subdomains 1 and 2. Therefore, when subdomain 3 was produced in isolation, these previously buried hydrophobic residues would become solvent exposed, which would be energetically unfavourable thus potentially preventing correct folding. As a result, five hydrophobic and polar residues in this interfacial region were manually mutated to make them more charged and hydrophilic **(Figure 4.26)**. The resulting construct was called "Interface" **(Strategy 2 in Figure 4.25)**. A GeneArt DNA string (ThermoFisher) was ordered with restriction sites to allow subsequent cloning into pET15bcon **Appendix A.3.3**.

```
Subdomain3 PQIYRWIREWGRDYVSELPTEVQKLKEKCDGKINYTDKKVCKVPPCQNACKSYDQWITRKK 61
Interface  PDIYEKIREWGRDYVSELPTEVQKLKEKCDGKINYTDKKVCKVPPCQNACKSYDQWITRKK 61

Subdomain3 NQWDVLSNKFISVKNAEKVQTAGIVTPYDILKQELDEFNEVAFENEINKRDGAYIELCVC 121
Interface  NEWDELSNKFISVKNAEKVQTAGIVTPYDILKQELDEFNEVAFENEINKRDGAYIELCVC 121
```

Figure 4.26: Mutations Introduced at Subdomain 3:Subdomain 2 Interface An alignment of the Subdomain 3 and Interface sequences, with the mutations made in the Interface construct highlighted in blue.

4.4.4.3 Strategy 3: Protein Repair One-Stop Shop (PROSS)

In the laboratory, the Protein Repair One-Stop Shop (PROSS) algorithm has successfully been used in the thermal stabilisation of another malaria immunogen, based

on the *Pf*RH5 invasion protein [272]. It was therefore decided to use PROSS as an additional strategy to design a thermal stable immunogen in this project.

4.4.4.3.1 PROSS Background and Overview PROSS is an algorithm that has been developed to improve the expression and stability of proteins whilst keeping their function unchanged [169]. It does this by selecting all amino acid mutations that optimize the protein's computed energy, subject to constraints inferred from homologous sequences. It therefore generates designs that improve a range of molecular parameters that are associated with stability whilst preserving function of the protein. The PROSS Workflow is summarised in **Figure 4.27**.

The first step is to generate an alignment of sequences homologous to the original model sequence. The homologous sequences are collected using BLASTP [273] with hits excluded if they cover less than 60% of the query sequence or if their sequence identity to the query sequence is lower than 20%. An alignment is created with MUSCLE [274] and is used to identify mutations that are most likely to occur through the natural diversity of the protein family, whilst also omitting potentially deleterious changes that are rarely seen in nature. By decreasing the subset of potential mutations, the PROSS algorithm also minimises the risks of false positives being introduced.

The next step involves Rosetta computational mutation scanning [275], in which each mutation suggested from the previous step is singly modelled against the desired designed structure. The energy difference between the wild-type structure and single point mutation models is calculated and given in a value of $\Delta\Delta G_{calc}$. Any mutation that produces a model that has a $\Delta\Delta G_{calc} \leq$ -0.45 Rosetta energy units is predicted to individually stabilise the initial model.

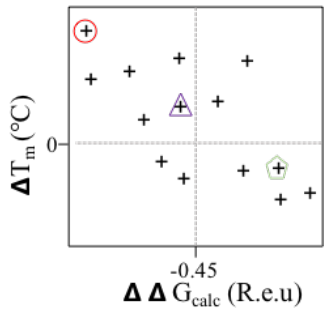
Finally, combinatorial sequence optimisation is used to design variants, which contain many of the mutations identified, and have an improved native-state energy compared



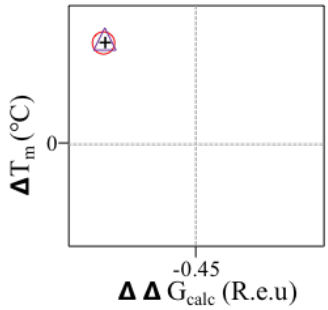
An alignment of the sequences of homologous proteins is generated

--GKQA---LPW---MAR-
--GKDA---LPW---MAR-
--GKQA---LTW---MAR-
--GKQA---LPW---MGR-

Individual point mutations are selected based on those which are seen in natural selection in the alignment



Each identified mutation is individually modelled against the WT Subdomain3 structure and scored according to RMSD to original model structure and Rosetta folding energy



Different individual stabilising point mutations identified in the previous step are tried in combination to see whether the multiple mutations result in an improved native-state energy compared to the original model protein

--GKDA---LPW---MGR-

A sequence with multiple stabilising mutations is identified

Figure 4.27: An Overview of the PROSS Workflow

to the initial model.

4.4.4.3.2 Alignment for *Pv*DBP-RII Subdomain 3 The first step in the PROSS workflow was to generate an alignment of sequences homologous to those of *Pv*DBP-RII subdomain 3. However, when an automatic alignment was created by searching BLASTP for homologous sequences, it included many protein sequences of DBL (duffy binding like) domains, mostly from the *Pf*EMP1 family. Although, *Pv*DBP-RII belongs to the DBL domain family, the structure of the subdomain 3 region of DBP appears to differ greatly from the corresponding region in the majority of *Pf*EMP1 DBL domains **(Figure 4.28)**. Sequence alignment analysis shows that there is great sequence diversity, with as little sequence identity as 22% between *Pv*DBP-RII subdomain 3 and the same region of *Pf*EMP1 DBL domains. Therefore, in order to conserve subdomain 3's function (i.e. as an epitope of DB9), an alignment was made manually that only contained DBP sequences (>67% sequence identity), rather than all DBL domains. (The sequence IDs for those used in the alignment can be seen in **Appendix A.4**).

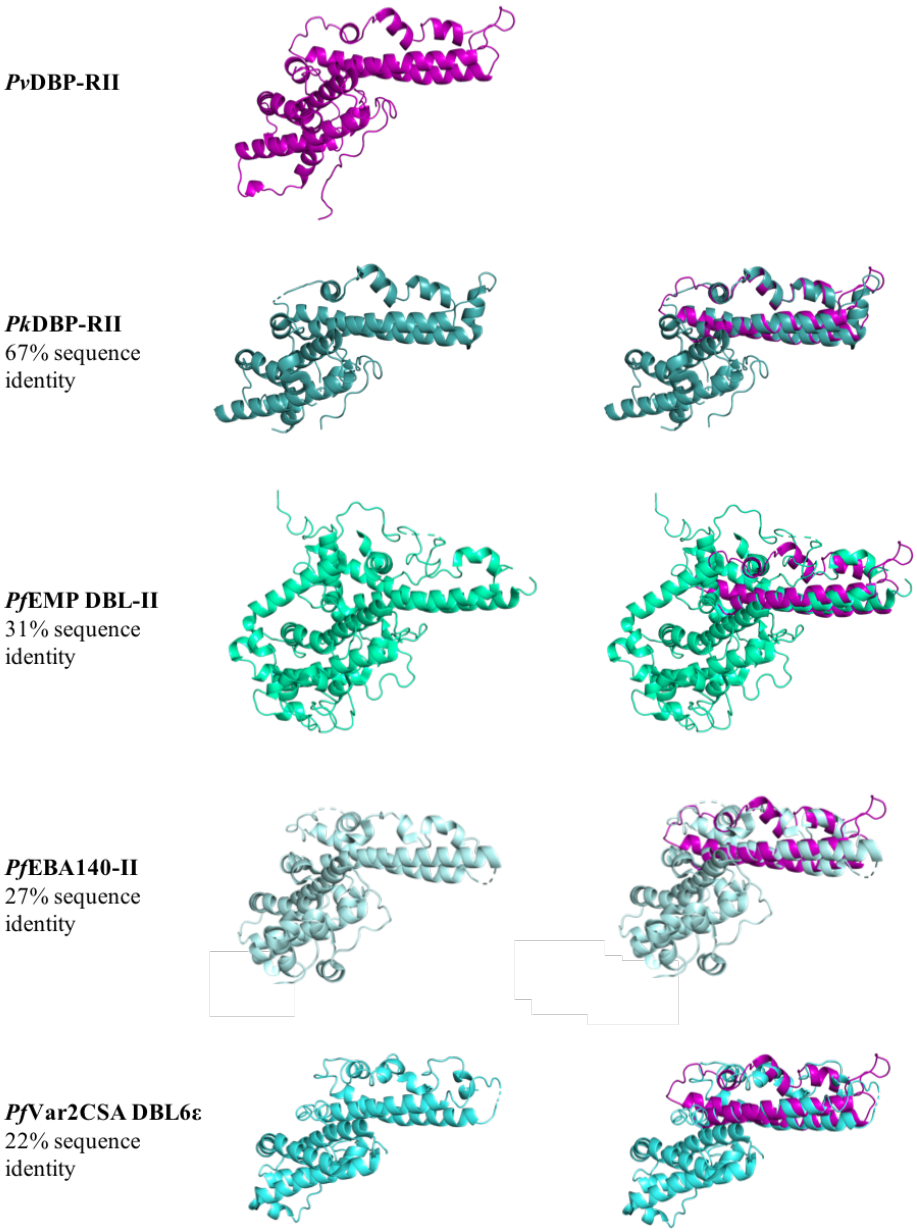
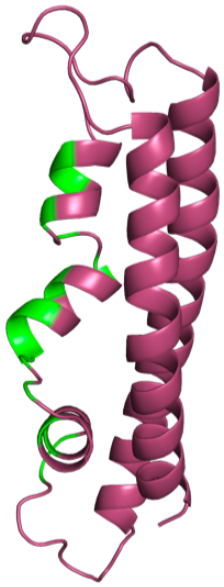


Figure 4.28: A Comparison of the Structure of *Pv*DBP-RII and Other Known DBL domains The structures of *Pv*DBP-RII and other DBP and DBL domains, with the overlay of the subdomain 3 region in pink to higlight difference in this region. The sequence identity between the subdomain 3 regions of these proteins compared to *Pv*DBP-RII is also shown.

4.4.4.3.3 Fixing Functional Residues In addition, the fifteen residues that directly bound to DB9 were "fixed" and so could not be mutated by the algorithm, which further helped to preserve the epitope (highlighted in green in **Figure 4.29**).



PQIYRWIREWGRDYVSELPTEVQKLKEKCDGKIAYTDKKVCKVPPCQNACKSYDQWITRKK
NQWDVLSNKFISVKNAEKVQTAGIVTPYDILKQELDEFNEVAFENEINKRDGAYIELCVC

Figure 4.29: "Fixed" residues not allowed to mutate in PROSS algorithm The fixed residues were all those directly interacting with DB9 and are highlighted in both the structure and sequence in green.

4.4.4.3.4 PROSS Stabilising Mutations PROSS produced seven stabilised variants with various numbers of mutations compared to the input Subdomain 3 sequence. The first design (Design1) contained the fewest number of mutations (two point mutations), whereas the last design (Design7) contained eight point mutations in combination **(Figure 4.30)**.

WT Subdomain 3
PQIYRWIREWGRDYVSELPTEVQKLKEKCDGKINYTDKKVCKVPPCQNACKSYDQWITRKK
NQWDVLSNKFISVKNAEKVQTAGIVTPYDILKQELDEFNEVAFENEINKRDGAYIELCVC

Design 1
PQIYRWIREWGRDYVSELPTEVQKLKEKCDGKINYTDKKVCKVPPCQNACKSYDQWITRKK
NQWDVLSNKFISVKNAEKVQTAGIVTPYDILKQELDEFNEVAFENEINKRDGAYIELCVC

Design 2
PQIYRWIREWGRDYVSELPTEVQKLKEKCDGKINYTDKKVCKVPPCQNACKSYDQWITRKK
NQWDVLSNKFISVKNAEKVQTAGIVTPYDILKQELDEFNEVAFENEINKRDGAYIELCVC

Design 3
PQIYRWIREWGRDYVSELPTEVQKLKEKCDGKINYTDKKVCKVPPCQNACKSYDQWITRKK
NQWDVLSNKFISVKNAEKVQTAGIVTPYDILKQELDEFNEVAFENEINKRDGAYIELCVC

Design 4
PQIYRWIREWGRDYVSELPTEVQKLKEKCDGKINYTDKKVCKVPPCQNACKSYDQWITRKK
NQWDVLSNKFISVKNAEKVQTAGIVTPYDILKQELDEFNEVAFENEINKRDGAYIELCVC

Design 5
PQIYRWIREWGRDYVSELPTEVQKLKEKCDGKINYTDKKVCKVPPCQNACKSYDQWITRKK
NQWDVLSNKFISVKNAEKVQTAGIVTPYDILKQELDEFNEVAFENEINKRDGAYIELCVC

Design 6
PQIYRWIREWGRDYVSELPTEVQKLKEKCDGKINYTDKKVCKVPPCQNACKSYDQWITRKK
NQWDVLSNKFISVKNAEKVQTAGIVTPYDILKQELDEFNEVAFENEINKRDGAYIELCVC

Design 7
PQIYRWIREWGRDYVSELPTEVQKLKEKCDGKINYTDKKVCKVPPCQNACKSYDQWITRKK
NQWDVLSNKFISVKNAEKVQTAGIVTPYDILKQELDEFNEVAFENEINKRDGAYIELCVC

Figure 4.30: Positions of PROSS-predicted stabilising mutations An overview of the different designs that PROSS produced and the mutations introduced in each design highlighted in yellow.

From previous experience of using PROSS in the laboratory, it was found to be conservative in the number of mutations it suggested. Therefore as a first attempt the Design7 was taken forward. The eight point mutations, in relation to the fixed residues of the DB9 epitope, can be seen on the sequence logo in **Figure 4.31**.

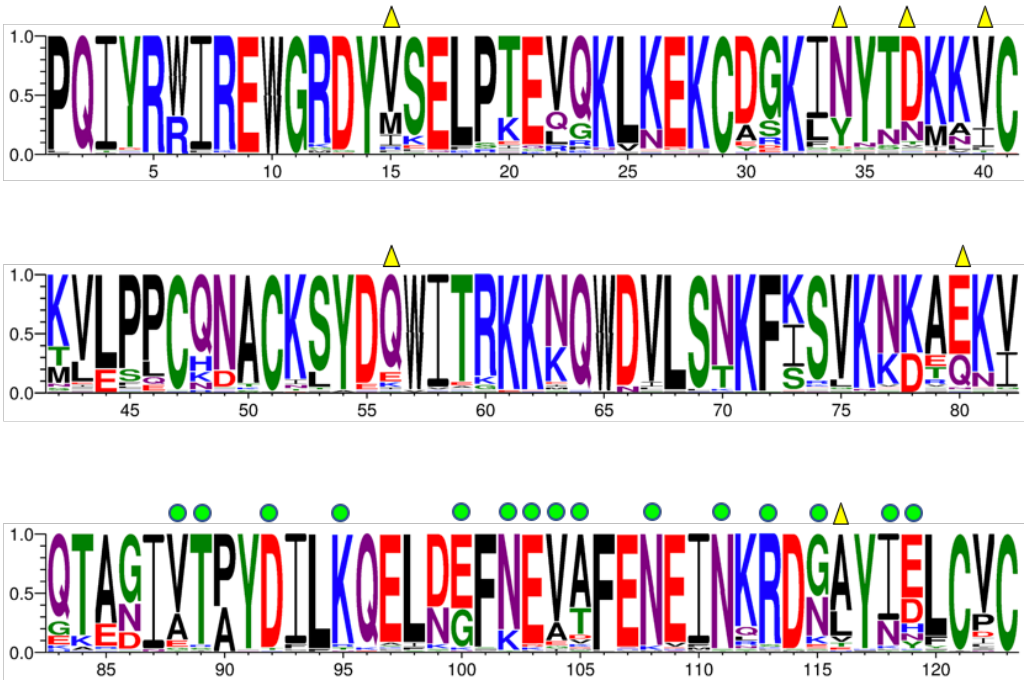


Figure 4.31: Positions of PROSS-predicted stabilising mutations A sequence logo showing the polymorphisms in the subdomain 3 sequences with the PROSS introduced mutations shown by the yellow triangles, and the DB9 contacting residues to be fixed highlighted by green circles. The sequence logo was made with WebLogo [212].

4.4.4.3.5 Mutations Added at Subdomain 2:Subdomain 3 Interface

Another factor that was taken into account when using the PROSS algorithm was that the residues that make up the interface between subdomain 2 and subdomain 3 were conserved within all DBP-RII domains. Therefore, these residues would not have been mutated by PROSS, since there was no natural variation in the homologous sequences. But looking at the structure, this normally buried surface would be solvent exposed in the designed immunogen. As a result, the five residues in the "Interface" construct were also manually changed in this PROSS design. Therefore, the final "PROSS" design contained eight PROSS-stabilising mutations plus the five interface manually-introduced mutations **(Strategy 3 in Figure 4.25)**. A GeneArt DNA string (ThermoFischer) was ordered for this construct and cloned into pET15bcon (**Appendix A.3.3)**

4.4.4.4 Sequences of the Proposed Immunogens

The designed immunogens taken forward were Subdomain 3, Interface, and PROSS **(Figure 4.32)**.

```
Subdomain3 PQIYRWIREWGRDYVSELPTEVQKLKEKCDGKINYTDKKVCKVPPCQNACKSYDQWITRKK 61
Interface  PDIYEKIREWGRDYVSELPTEVQKLKEKCDGKINYTDKKVCKVPPCQNACKSYDQWITRKK 61
PROSS      PDIYEKIREWGRDYMSELPTEVQKLKEKCDGKIYYTGKKTCKVPPCQNACKSYDEWITRKK 61

Subdomain3 NQWDVLSNKFISVKNAEKVQTAGIVTPYDILKQELDEFNEVAFENEINKRDGAYIELCVC 121
Interface  NEWDELSNKFISVKNAEKVQTAGIVTPYDILKQELDEFNEVAFENEINKRDGAYIELCVC 121
PROSS      KEWDELSNKFISVKNAENVQTAGIVTPYDILKQELDEFNEVAFENEINKRDGTYIELCVC 121
```

Figure 4.32: Sequence Alignment of Proposed Immunogen Constructs Highlighted in blue are the residues that have been changed compared to the original subdomain 3 sequence.

4.4.5 Cloning, Expression and Purification of Designed Immunogens

All of the immunogen constructs were cloned into the pET15bcon plasmid, which contains a cleavable N-terminal histidine tag for purification **(Materials and Methods Figure 2.1)**. The immunogens were expressed in *E. coli* BL21-DE3 cells.

Like *Pv*DBP-RII, Subdomain 3 expressed solely in the insoluble fraction of the cell pellet, with no band visible in the post-lysis supernatant sample of the SDS-PAGE gel **(Figure 4.33)**. The Subdomain 3 therefore had to be solubilised in guanidine hydrochloride before being refolded on a nickel column in a 10% step-wise decreasing gradient of guanidine **(Materials and Methods 2.4.4)**. Not only did the refolding protocol prove much easier than for *Pv*DBP-RII, which does not refold on a nickel column, but also Subdomain 3 expressed at greater levels than *Pv*DBP-RII at 0.96 mg/L cell pellet, compared to 0.3 mg/L.

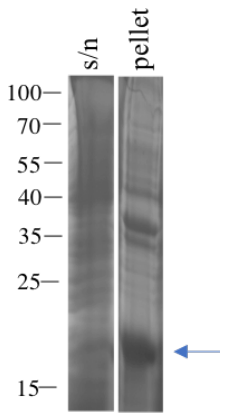


Figure 4.33: Insoluble Expression of Subdomain 3 SDS-PAGE gels showing the s/n (supernatant) and the resuspended pellet from the spin post-lysis. The s/n contains only the soluble proteins, and the pellet contains the insoluble proteins. The band representing Subdomain 3 is indicated by the blue arrows.

The Interface and PROSS immunogens both expressed solubly in the cells **(Figure 4.34)**. Post-lysis, the soluble supernatant was flowed over Ni-NTA beads and the PROSS/Interface proteins eluted in 200 mM imidazole **(Materials and Methods 2.4.3)**. As well as expressing solubly, the PROSS and Interface protein constructs also gave a much higher yields at 5.20 mg/L cell pellet and 2.20 mg/L cell pellet, respectively.

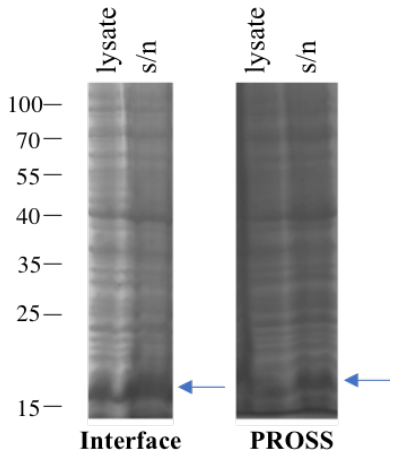


Figure 4.34: Soluble Expression of Interface and PROSS SDS-PAGE gels showing the lysate, which is the whole cell lysate and so contains all proteins both soluble and insoluble. s/n=supernatant from the spin post-lysis and so contains only the soluble proteins. The bands representing interface and PROSS are indicated by the blue arrows.

Post Ni-NTA purification, all constructs were dialysed into PBS and the His-tags removed by TEV cleavage overnight at RT **(Materials and Methods 2.4.5)**. The TEV-cleaved proteins were further purified to remove any other contaminants or aggregates using size exclusion chromatography (S75, GE Healthcare). The single monodisperse peaks highlight that the folded, monomeric protein was isolated for each construct **(Figure 4.35)**.

In all of these constructs, the monomeric protein appeared as a double band on the SDS-PAGE gel **(Figure 4.35)**. In order to check that there was no contamination/degradation of the proteins, for the Subdomain 3 construct each band was cut from the gel and sent for in-gel trypsin mass spectrometry. The analysis of both bands confirmed that each contained the whole sequence of Subdomain 3. As a result, it was thought that the appearance as a double band was a gel artifact caused by partial reduction of the three disulphide bonds present in each construct.

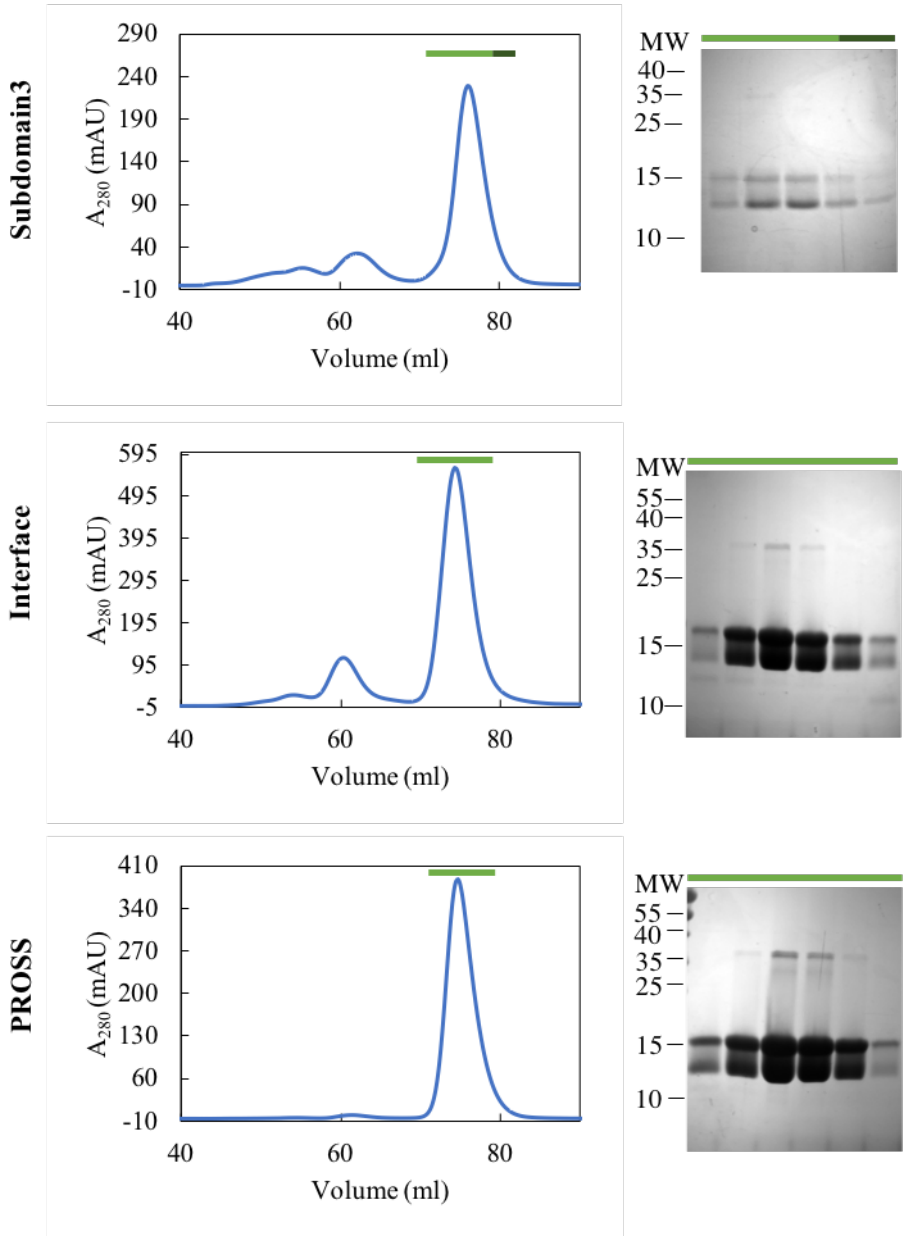


Figure 4.35: Purification of Interface and PROSS The size exclusion trace (left) and corresponding reduced SDS-PAGE gel (right) with the equivalent fractions shown by the green bar. The light green section represents the fractions which were pooled for future studies.

4.4.6 Assessing the Fold of the Immunogens

It was important to check that the immunogens were properly folded (especially for the immunogens expressed in the insoluble fraction). This was done using CD. From CD spectra with minima at 208nm and 220 nm, it was clear that PROSS, Interface

and SBD3 all contained predominantly alpha-helical secondary structure, as expected (left panel of **Figure 4.36**). Moreover, from the thermal melts of the immunogens from 25°C to 90°C it was clear that all of the constructs retained their secondary structure until 60°C (right panel of **Figure 4.36**). Indeed PROSS did not fully melt even at 90°C and so is more thermally stable than Interface and Subdomain3 **(Table 4.6)**.

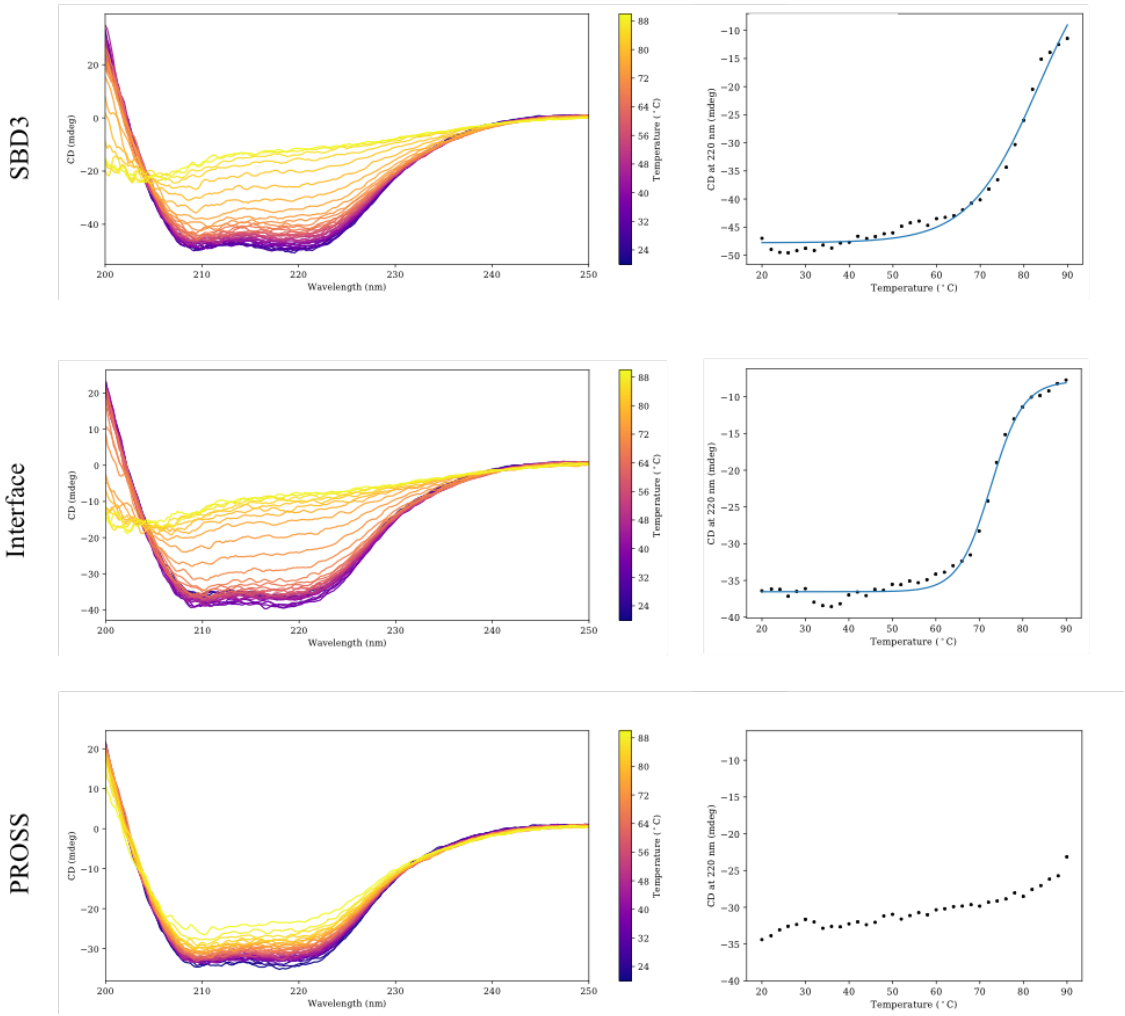


Figure 4.36: CD Spectra of Immunogens Left: the thermal melt curves between wavelengths 200 and 250 nm, at temperatures ranging from 20°C to 90°C. Right: The thermal denaturation of the alpha-helical protein measured at 220nm to gain the melting temperature.

Construct	Tm °C
Subdomain 3	83.2
Interface	72.7
PROSS	n/a

Table 4.6: Melting Temperatures of Immunogen Constructs

4.4.7 Binding of the Immunogens to DB9

In order to confirm that the immunogens successfully recapitulated the DB9 epitope, the binding of the immunogens to DB9 was assessed using SPR. In each case, protein A/G was coupled to a CM5 SPR chip. First, 10μg/ml of DB9 mAb was coupled to the protein A/G and then the immunogen was flown over the chip in a two-fold dilution series from 1μM **(Figure 4.37)**. The Subdomain 3 construct bound to the DB9 with similar affinity to the recombinant DBP-RII, and a comparably slow off-rate **(Table 4.7)**. The Interface construct had a 10-30 times faster off-rate than the Subdomain 3 and *Pv*DBP-RII constructs. The mutations introduced at the "interface" region could have caused slight changes in the fold of the DB9 epitope, explaining this increased off-rate.

PROSS, however, bound to the DB9 with a similar on-rate but had a very fast off-rate **(Table 4.7)**. Upon re-analysis of the designed PROSS sequence and the interacting site in the *Pv*DBP-RII:DB9Fab structure, it was noticed that the A114 in the original sequence had been changed to T114 as one of the mutations introduced by the PROSS algorithm **(Figure 4.31)**. Although this side chain is not directly involved in binding to the DB9Fab, the main chain carbonyl of A114 does make a hydrogen bond with the heavy chain of the DB9Fab **(Table 4.4)**. By increasing the size of the side chain, from alanine to threonine, this interaction and other nearby interactions may have been disrupted. Therefore, another construct was produced in which the T114 was mutated back to the A114, this construct was referred to as

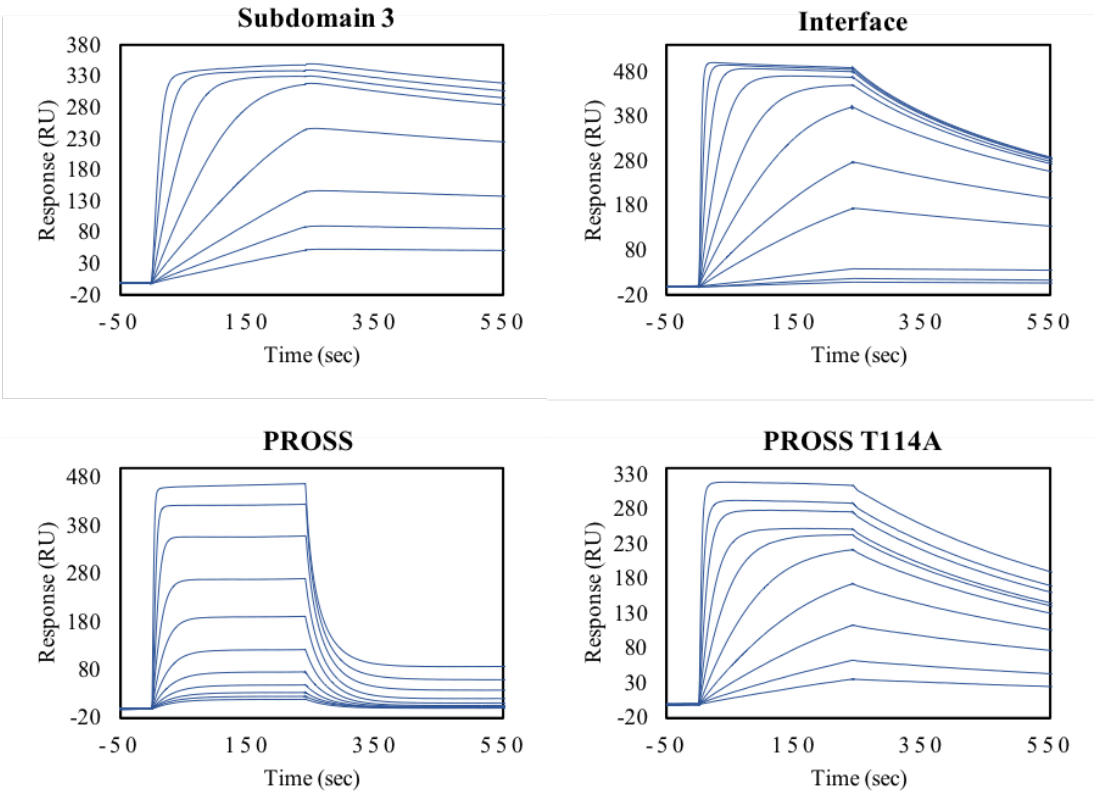


Figure 4.37: Immunogen Binding to DB9 SPR curves showing the binding of the immunogens to DB9. Each set of curves represents a two-fold dilution series from 1μM.

PROSS-T114A **(Appendix A.3.3)**. Through this point mutation, the slow off-rate of the interaction with DB9 was recovered **(Figure 4.37 and Table 4.7)**, and so all further experiments were continued with the PROSS-T114A construct.

	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})	K_d (nM)
***Pv*DBP-RII**	6.94x10^5	9.76x10^{-4}	2.63
Subdomain3	8.26x10^5	3.28x10^{-4}	0.397
Interface	7.07x10^5	2.71x10^{-3}	3.82
PROSS-T114A	6.94x10^5	1.54x10^{-3}	2.00

Table 4.7: Kinetic Analysis of Immunogens Binding to DB9

4.4.8 Structural Validation of the Designed Immunogens

4.4.8.1 Crystallisation of Subdomain3:DB9Fab Complex

The purified Subdomain 3 and the DB9Fab were mixed in a 1.1:1 ratio (Subdomain3:DB9Fab) for half an hour at room temperature before being loaded onto a size exclusion column S200 10/30 (GE Healthcare). The complex eluted as one peak **(Figure 4.38 A)**, however, the peak eluted much later than expected. Despite this, complex formation was confirmed via SDS-PAGE **(Figure 4.38 B)** and the peak was pooled and concentrated to 11 mg/ml.

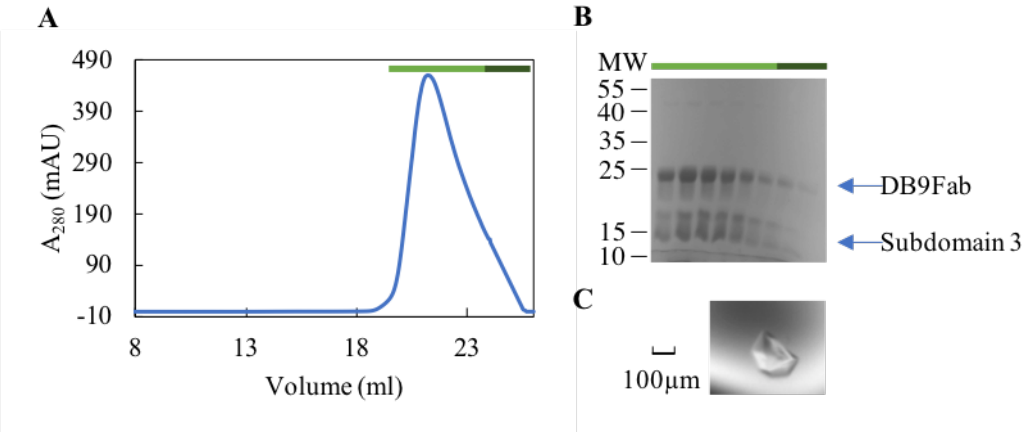


Figure 4.38: Subdomain3:DB9Fab Complex Formation and Crystallisation The size exclusion trace (A) and corresponding reduced SDS-PAGE gel (B) with the equivalent fractions shown by the green bar. The light green section represents the fractions which were pooled for future studies. C) A to scale image of the Subdomain3:DB9Fab crystal.

The following broad crystallisation commercial screens were set up: MIDAS, JCSG plus, ProPlex, and Morpheus, at 4°C (as described in **Materials and Methods 2.7.2.1**). A crystal grew in ProPlex B7 with a reservoir solution of 0.2 M ammonium acetate, 0.1 M sodium acetate pH4.0, 15% w/v PEG 4000 **(Figure 4.38 C)**. The crystal was transferred into a drop of well solution and was then cryo-cooled in liquid nitrogen for storage and data collection.

4.4.8.2 Data Collection and Solution of Subdomain3:DB9Fab Complex Structure

For the Subdomain3:DB9Fab complex, data was collected at the Diamond Light Source on beamline I03. The initial data showed that the Subdomain:DB9Fab crystal diffracted to 1.58Å. A full dataset was collected, using a collection strategy suggested by EDNA. This dataset was processed automatically by Xia2, which suggested a space group of $P2_12_12_1$, and the data collection statistics can be seen in **Table 4.8**.

Space Group	$P2_12_12_1$
Cell Parameters (Å)	a=76.43, b=126.19, c= 129.86
	α=β=γ=90°
Resolution (Å)	90.50-1.55 (1.55-1.58)
R_{meas} (%)	6.08 (78.5)
R_{pim} (%)	2.36 (30.8)
I/σ (I)	14.77 (1.96)
$CC_{1/2}$	1 (0.6)
Completeness (%)	98.3 (95.7)
Multiplicity	6.5 (6.3)

Table 4.8: Data collection statistics for the Subdomain3:DB9Fab Structure (Values in parentheses are the highest-resolution shell).

The Matthews coefficient estimation was used to predict (with a probability of 99%) that two Subdomain3:DB9Fab complexes were in each asymmetric unit of the crystal. Therefore, the structure was solved by molecular replacement, using two polyalanine models of the structure of the DB9Fab **(Section 4.2.7)**, with all loops removed as search models, in Phaser. Once the two Fabs had correctly been positioned in the unit cell and electron density representing the Subdomain 3 seen in the space adjacent to the predicted positions of the Fab CDR loops, a second molecular replacement

search was performed. In this second search, two models of the Subdomain 3 with all loops and DB9 interacting helices removed was used to find the positions of the two Subdomain 3 moelcules. Two SBD3:DB9Fab complexes were found in each unit cell, as predicted by the Matthews coefficient, with extra density consistent with the loops and helices that had been removed in the search models. After multiple rounds of refinement and rebuilding, using Buster and Coot respectively, all of the helices and loops were resolved. The final solution parameters are shown in **(Table 4.9)**.

Resolution (Å)	90.50-1.55
R_{work}	0.227
R_{free}	0.243
rmsd bond lengths (Å)	0.01
rmsd bond angles (°)	1.15
Ramachandran plot	
Favoured region (%)	98.0
Allowed region (%)	2.00
Disallowed region (%)	0.00

Table 4.9: Refinement statistics of the Subdomain3:DB9Fab Complex

The R_{work} value of 0.227 suggested a good fit of the model to the data at 1.55Å, while the slightly higher R_{free} value of 0.243 showed that independent, unprocessed reflections could be explained by the model and confirmed that the model did not overinterpret the data. The Ramachandran plot **(Figure 4.39)** showed that 98% of the residues are in the allowed region, further suggesting that this was modelled correctly.

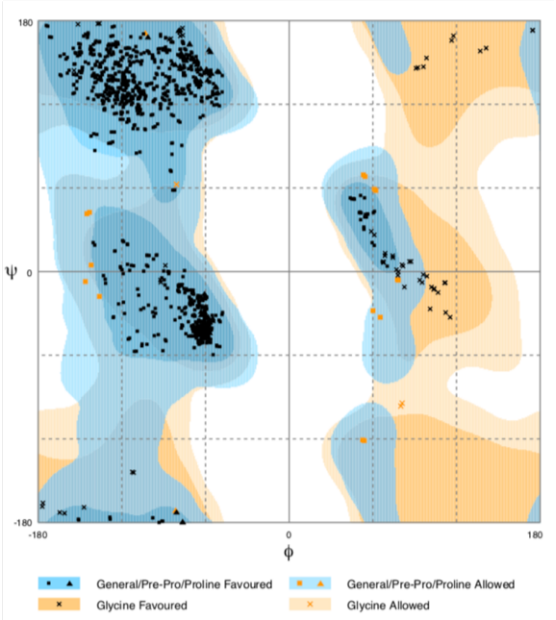


Figure 4.39: Ramachandran Plot for the Subdomain3:DB9Fab Structure

4.4.8.3 Analysis of the Subdomain3:DB9Fab Complex Structure

As expected, the DB9Fab binding site on the subdomain 3 is the same as that using the whole *Pv*DBP-RII domain, with the interacting residues identical **(Figure 4.40)**. Overlaying the Subdomain3:DB9Fab structure with that of *Pv*DBP-RII:DB9Fab shows an overall RMSD of 0.43Å, with the DB9 epitope region having an RMSD of 0.34Å.

This showed that the designed immunogen successfully mimics the DB9 epitope and can be used as a potential immunogen.

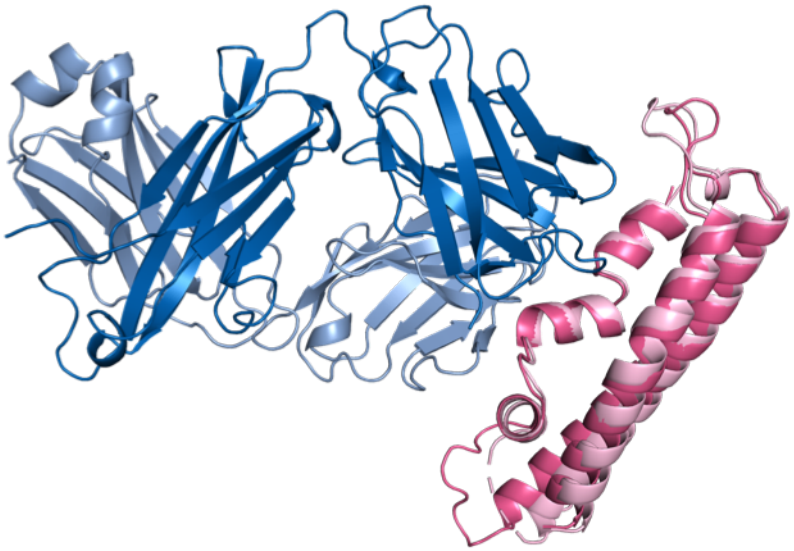


Figure 4.40: Overlay of the Subdomain3:DB9Fab Structure with the Subdomain3 from the previous *Pv*DBP-RII:DB9Fab Structure The two structures were aligned by the heavy chain of the DB9Fab fragment, with the dark and light blue representing the heavy and light chains of DB9Fab, respectively. The light pink is the structure of the subcloned Subdomain3, whereas the dark pink is the subdomain 3 region from the *Pv*DBP-RII structure.

4.4.8.4 Attempted Crystallisation of Interface:DB9Fab

Since Interface and PROSS-T114A were soluble and expressed highly, they were seen as more desirable potential vaccines due to the relative ease and cost of production. Due to limited amount of DB9 provided, and the delay in obtaining the PROSS-T114A construct, an attempt was made to obtain the crystal complex of Interface bound to DB9Fab.

In a similar way to Subdomain3:DB9Fab complex, Interface:DB9Fab complex was formed by the mixing of purified Interface and DB9Fab in a 1.1:1 (Interface:DB9Fab) ratio for half an hour at RT **(Figure 4.41)**.

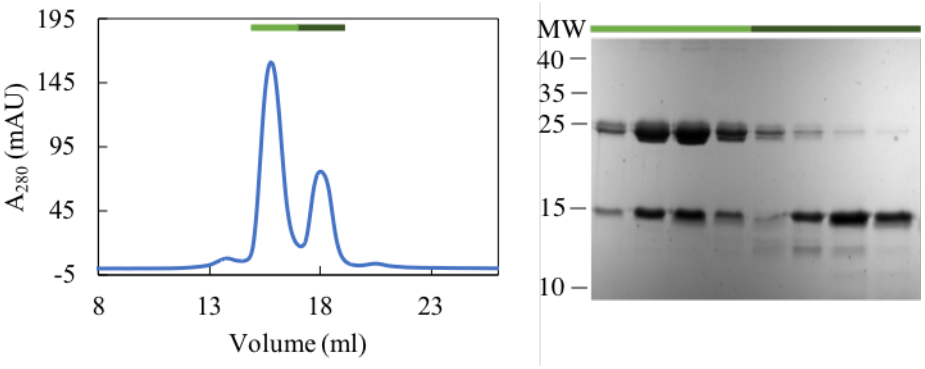


Figure 4.41: Interface:DB9Fab Complex Formation The size exclusion trace (left) and corresponding reduced SDS-PAGE gel (right) with the equivalent fractions shown by the green bar. The light green section represents the fractions which were pooled for future studies

The peak that represented the purified complex was pooled and concentrated to 12 mg/ml and broad crystallisation screens set up in the same sitting drop vapour diffusion as Subdomain3:DB9Fab. Overnucleated crystals grew in a couple of conditions: ProPlex D12 (0.1 M Tris pH8.5 20% w/v PEG 6000) and also in ProPlex C8 (0.2 M NaCl, 0.1 M Tris pH8.0, 20% w/v PEG 4000). To check whether these overnucleated crystals contained the complex, some were harvested and washed in mother liquor and an SDS-PAGE gel was developed using silver stain **(Figure 4.42)**. Silver bullet optimisation plates of both ProPlex D12 and ProPlex C8 were set up using the overnucleated crystals as seeds **(Methods 2.7.2.2)**. However, no larger crystals grew from which data could be collected.

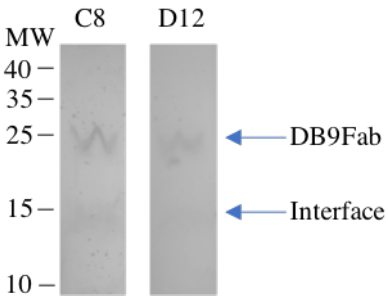


Figure 4.42: Interface:DB9Fab Crystal Contains Both Components A reduced SDS-PAGE gel of the harvested and washed Interface:DB9Fab crystals, visualised using silver stain.

Looking at the packing of the Subdomain3:DB9Fab structure it can be seen that three of the five residues changed between the Interface and SBD3 were directly involved in crystal contacts in the Subdomain3:DB9Fab structure **Figure 4.43)**. It can be speculated that these mutations therefore would disrupt the packing in this particular crystal structure and prevent this crystal from forming.

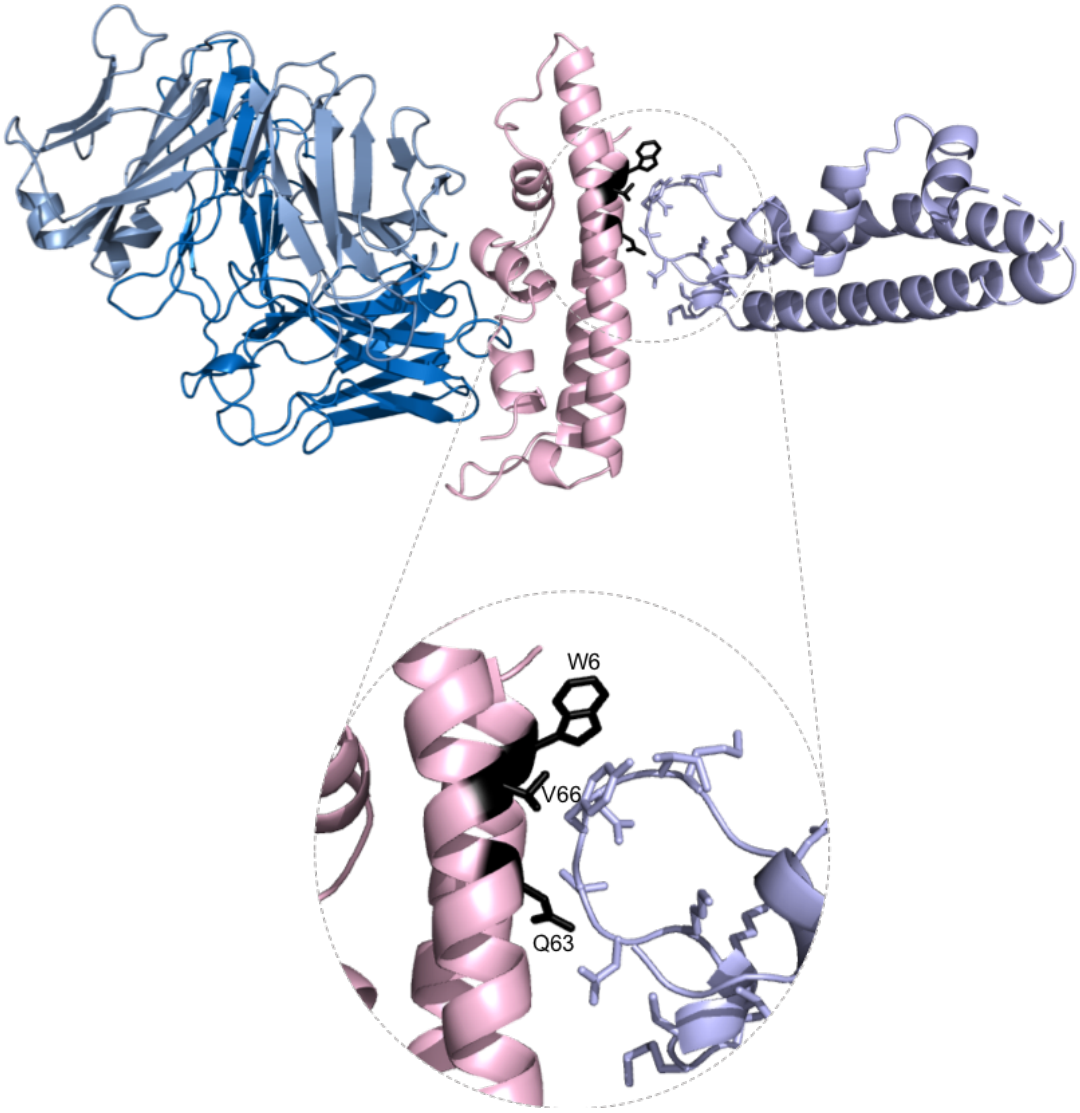


Figure 4.43: Subdomain3:DB9Fab Crystal Contacts in Interface Region Subdomain 3 in light pink and the Fab Heavy in dark blue and the Light Chain in light blue as usual. The residues mutated in the Interface construct that make crystal contacts with another copy of Subdomdain 3 (mauve) are highlighted in black.

4.4.9 Immunisations of Rabbits with Designed Immunogens

Although the designed immunogens had been shown to be folded and functional as they bound to DB9 with a similar affinity to *Pv*DBP-RII, this did not necessarily mean that they would be good vaccine candidates and upon immunisation raise antibodies that prevent reticulocyte invasion. To test their immunogenicity, rabbits were immunised with each of the immunogens and the sera analysed for inhibitory antibodies that block both the binding of recombinant *Pv*DBP-RII to recombinant DARC, and that inhibit reticulocyte invasion.

4.4.9.1 Endotoxin Removal

Since all the immunogens were expressed in *E. coli* cells, they had to be treated to remove the endotoxins that compose the outer cell membrane of all gram-negative bacteria. This was achieved using Pierce High-Capacity Endotoxin Removal Resin (ThermoScientific).

4.4.9.2 Coupling of Immunogens to VLPs

As described previously in **Section 3.2.11.3**, it has been seen that by attaching small immunogens to larger particles, like VLPs, the overall immunogenicity and quantity of antibodies produced can be increased [221–223]. Since the Subdomain 3-based constructs are 15kDa compared to the 35.5kDa of *Pv*DBP-RII, it was thought that the immunogenicity could be increased by attaching the immunogen to a VLP. One way to attach immunogens to VLPs is by using a SpyCatcher:SpyTag coupling system [224] (explained in **Section 3.2.11.3.1**).

4.4.9.3 Design of Interface Mutant For VLP Display

Despite the PROSS-T114A construct being potentially the most promising vaccine candidate because of its higher yield and thermal stability, due to the issues with the original PROSS construct binding to DB9, and the delay in obtaining the PROSS-T114A construct, the Interface construct was coupled to VLPs.

In order to couple the Interface immunogen to VLPs, a SpyTag had to be introduced. One consideration to take in the coupling of small immunogens to VLPs, is where to place the SpyTag so that the desired DB9 epitope is likely to be presented and accessible on the surface of the VLP. Thereby leading to the production of the desired DB9-like antibodies. If the tag had been introduced at the N or C-terminus, it was postulated that the DB9 epitope may be more occluded and therefore less likely to be recognised by the immune system **(Figure4.44 A and B)**. Therefore, a construct was designed in which the SpyCatcher tag would allow the immunogen to be displayed perpendicularly on the VLP, with the DB9 epitope putatatively most exposed **(Figure4.44 C)**.

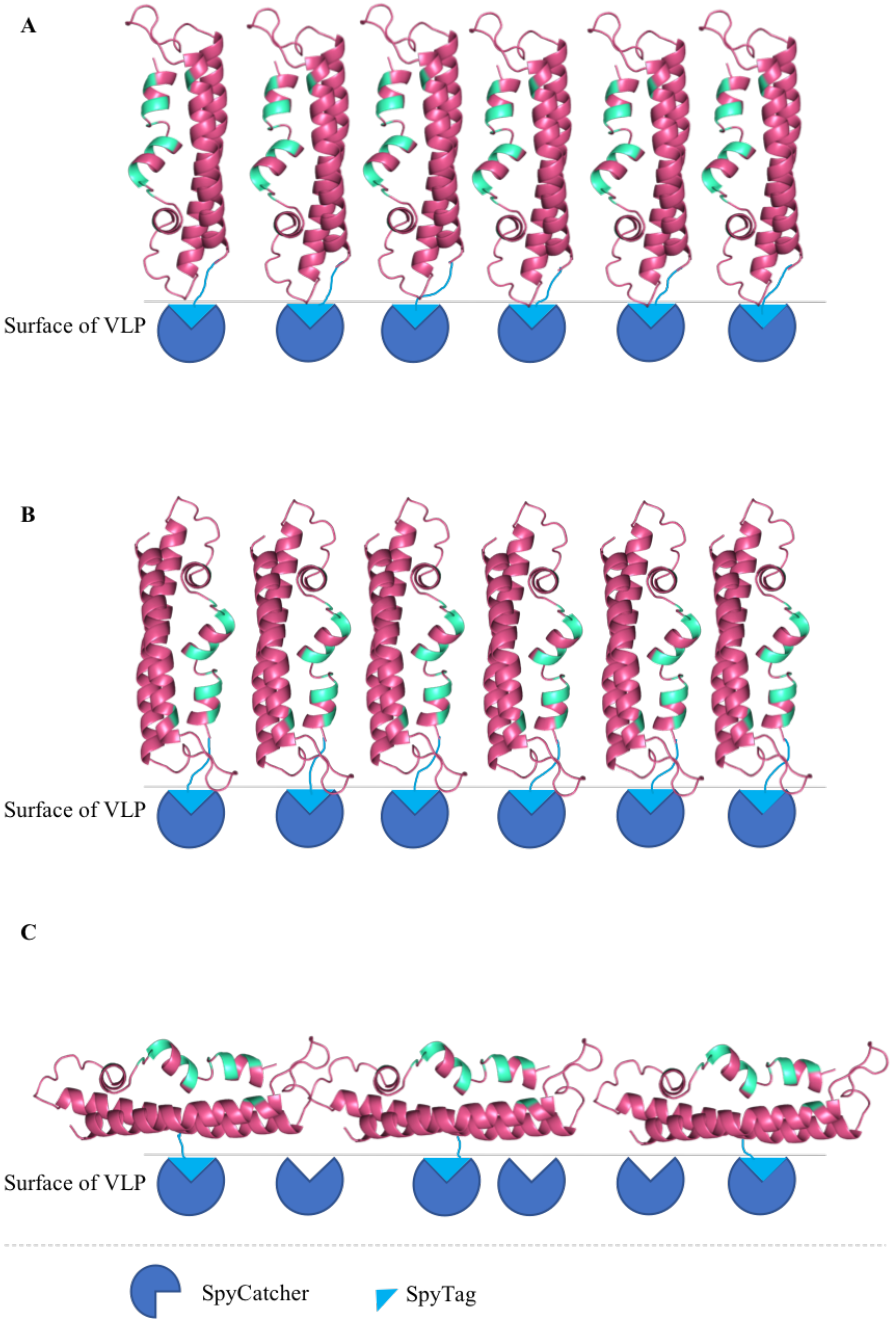


Figure 4.44: Where to put the SpyTag? Display of Immunogens on VLP The orientation of Interface on VLP surface if the SpyTag is attached to the N-terminus (A), C-terminus (B) or to a sidechain in the middle of the sequence on the opposite side of the immunogen to the DB9 epitope (C). The DB9 epitope in each case is shown in green cyan.

A synthetic peptide, which had a SpyTag coupled to a linker that contained a maleimide group was made (H_2N-SpyTag-MalPEG2-amide). The maleimide group forms covalent bonds with thiols, for example on free cysteines, allowing coupling of the SpyTag to specific positions on the protein **(Figure 4.45 A)**. Therefore a free cysteine was introduced by mutation of Arg59, R59C. The position of the mutated Arg59 allowed the SpyTag to be coupled to the backside of the helix on the opposite side of the protein to the DB9 epitope, which would likely be exposed after SpyTag:SpyCatcher coupling **(Figure 4.45 B)**.

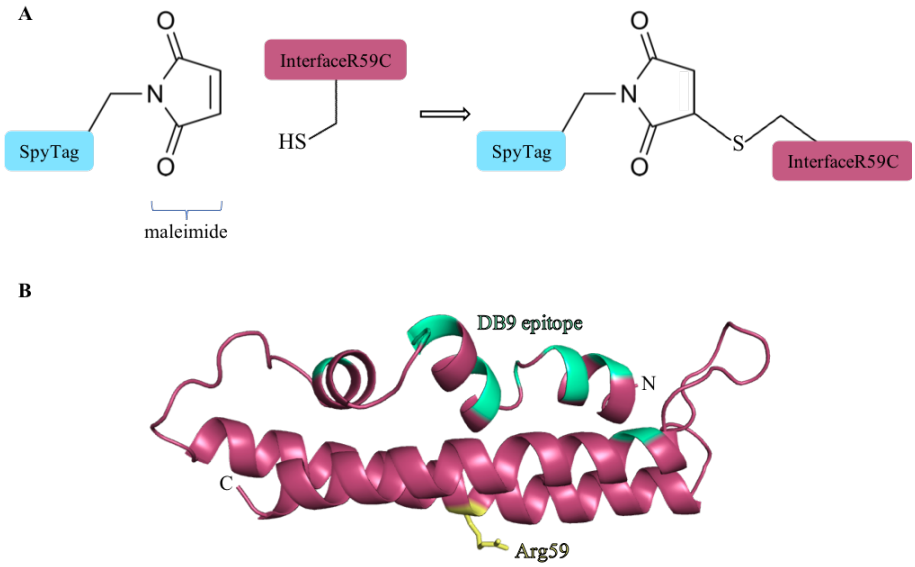


Figure 4.45: SpyTag-maleimide interaction with InterfaceR59C
A) Reaction mechanism of the maleimide group interacting with the free thiol. B) The model structure of Interface, highlighting the position of the Arg59 residue (yellow) compared to the DB9 epitope (green-cyan).

4.4.9.4 Purification of InterfaceR59C

An R59C mutant of the Interface construct was cloned by site directed mutagenesis **(Materials and Methods 2.1.5)**, with the primers for this reaction seen in the **Appendix A.3.3**. The InterfaceR59C construct was expressed and purified in an identical way to the Interface construct. A higher proportion of the protein folded into dimers than with the Interface construct. This was likely due to the free cysteine

in the Interface-R59C forming intermolecular disulphide bonds. However, despite this the monomer peak could clearly be separated by size exclusion **(Figure 4.46)**. Again, as with the other immunogen constructs, multiple bands could be seen on the SDS-PAGE gel, thought to be due to partial reduction of the disulphide bonds causing the species with different numbers of disulphide bonds to migrate differently.

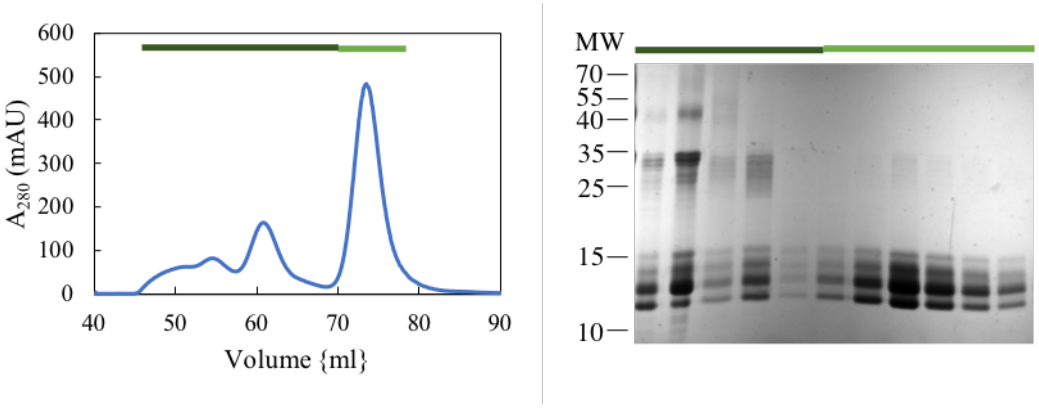


Figure 4.46: Purification of Interface R59C The size exclusion trace (left) and corresponding reduced SDS-PAGE gel (right) with the equivalent fractions shown by the green bar. The light green section represents the fractions which were pooled for future studies.

4.4.9.5 Coupling of the InterfaceR59C to the VLP

The Interface-R59C was firstly coupled to the SpyTag-linker by incubating them together overnight at RT in a 2:1 molar ratio (Interface-R59C:SpyTag-linker) to ensure that no free linker would be present **(Figure 4.47 A)**. The upwards shift on the SDS-PAGE gel represented the successful coupling of the Interface R59C construct to the SpyTag-maleimide linker molecule.

Once formed, the InterfaceR59C-SpyTag construct was incubated with the HBsAg-Spy Catcher-VLPs (gifted from the Biswas Group, Jenner Institute). The HBsAg-SpyCatcher-VLP consists of the hepatitis B surface antigen (HBsAg), a protein able to assemble into VLPs that each contain ∼60 HBsAg subunits. Each HBsAg subunit has a SpyCatcher domain genetically fused into it, allowing coupling to

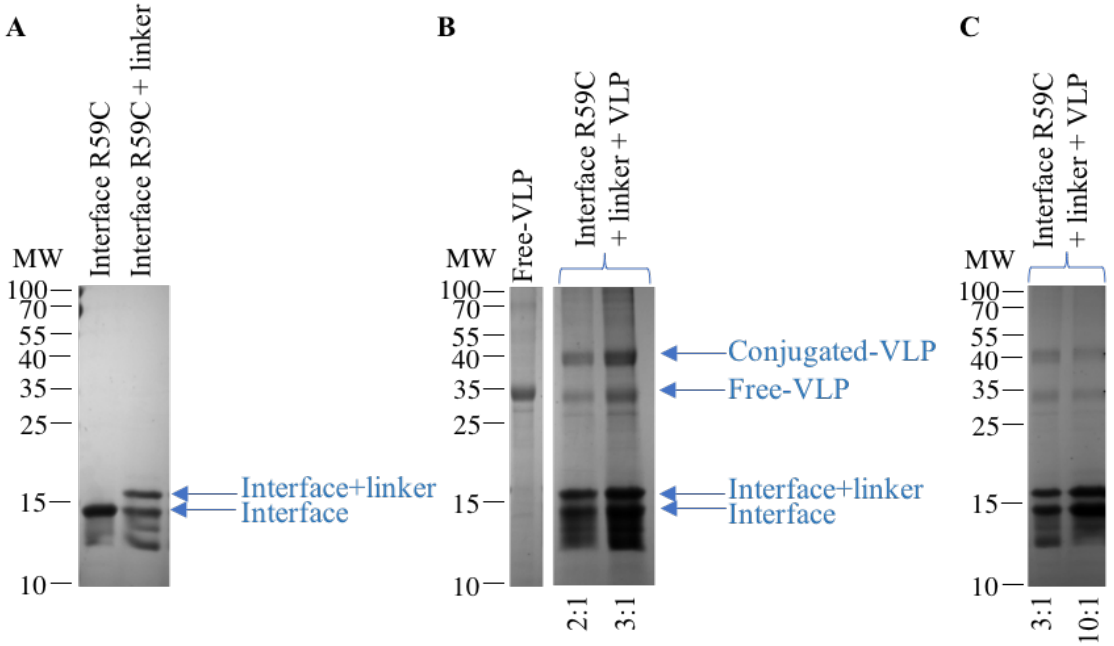


Figure 4.47: Interface R59C conjugation with Malemide Spytag Linker before with Spy-catcher VLP A) SDS-PAGE gel showing purified Interface R59C before and after coupled to the SpyTag linker. B) SDS-PAGE gel showing the coupling of the InterfaceR59C-SpyTag to the VLP-SpyCatcher at two different ratios, 2:1 and 3:1. C) SDS-PAGE gel showing the coupling of the InterfaceR59C-SpyTag to the VLP-SpyCatcher at ratios 3:1 and 10:1.

SpyTag-tagged proteins. The InterfaceR59C-SpyTag construct was mixed with the HBsAg-SpyCatcher-VLPs in 2:1 and 3:1 ratios, respectively, for 3 hours. The 3:1 ratio produced a greater proportion of coupled VLP subunits compared to free VLP subunits **(Figure 4.47 B)**, however there was still a large amount of uncoupled VLP present. In an attempt to increase the proportion of coupled-VLP to free-VLP, the ratio of InterfaceR59C-Spytag:HBsAg-SpyCatcher-VLPs was increased to 10:1. However, this increase did not appear to have any affect on the proportion of coupled-VLP to free-VLP, compared to the 3:1 incubation **(Figure 4.47 C)**.

From the gels in **(Figure 4.47 B and C)** it was clear that there was still a large proportion of VLP that did not have Interface coupled to it (Free-VLP). This may be due to steric occlusion not allowing all HBsAg subunits in a single VLP to be coupled to InterfaceR59C-SpyTag domains. Therefore, it was decided to continue

with this partially coupled sample. One disadvantage of using this sample is that for immunisations the concentration of conjugated-VLP can only be estimated. The excess of InterfaceR59C-maleimide-Spytag (and InterfaceR59C) was removed from the VLP mix by size exclusion chromatography (Superose 6 10/30, GE Healthcare) **Figure 4.48)**.

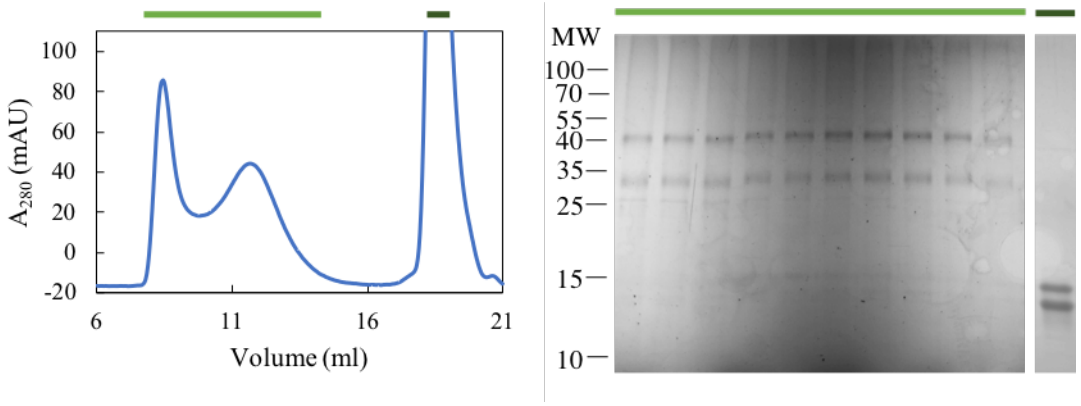


Figure 4.48: Purification of the Interface-Conjugated VLP The size exclusion trace (left) and corresponding SDS-PAGE gel (right) with the equivalent fractions shown by the green bar. The light green section represents the fractions which were pooled for future studies.

4.4.9.6 Analysis of Coupling

In order to check that the immunogens were covalently attached with their DB9 epitopes exposed, a crude SPR binding experiment was performed, in which, DB9 was bound to a protein A-coupled CM5 chip, and then 1μM of the purified coupled-VLP was flowed over. As a negative control test, completely uncoupled VLP was flowed over to show that any binding that occurred was due to the Interface specifically binding to the DB9 **(Figure 4.49)**.

This was a crude binding test as it only showed that there was binding in the Interface-coupled VLP and there was no binding in the uncoupled VLP. Due to the presentation of multiple Interface immunogens on one VLP particle, it was impossible to determine any binding kinetics. As can be seen in **Figure 4.49** the

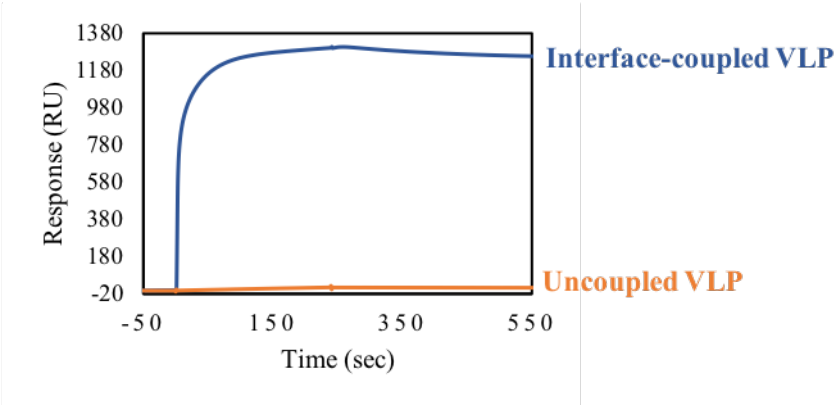


Figure 4.49: SPR binding curve of Interface-coupled VLP and uncoupled VLP to DB9

Interface-coupled VLP bound to the DB9 chip with a very slow off rate, which could be explained due to the avidity affect of having multiple Interface constructs on one VLP particle.

4.4.9.7 Immunisation Outcomes

Pairs of rabbits were each immunised with Subdomain 3, Interface, Interface-coupled VLP, and PROSS-T114A. As a negative control for the Interface-coupled VLP, two rabbits were also immunised with uncoupled-VLP. And as a positive control, two rabbits were immunised with recombinant *Pv*DBP-RII. An ELISA binding assay showed that all the immunisations, except the one with uncoupled-VLP, resulted in sera which could bind to recombinant *Pv*DBP-RII **(Figure 4.50 A)**. This was promising as it showed that antibodies that recognised *Pv*DBP-RII had been produced, and therefore that the immunisations themselves had likely been successful.

Despite this, only IgG from Subdomain 3 and Interface rabbits inhibited parasite growth **(Figure 4.50 B)**. Although it is positive that the immunisations with Subdomain 3 and Interface did result in inhibition of parasite growth, it is worrying that the IgG from the positive control rabbits immunised with *Pv*DBP-RII did not

result in any growth inhibition. It is also hard to explain why the PROSS and VLP-interface constructs also resulted in no growth inhibition.

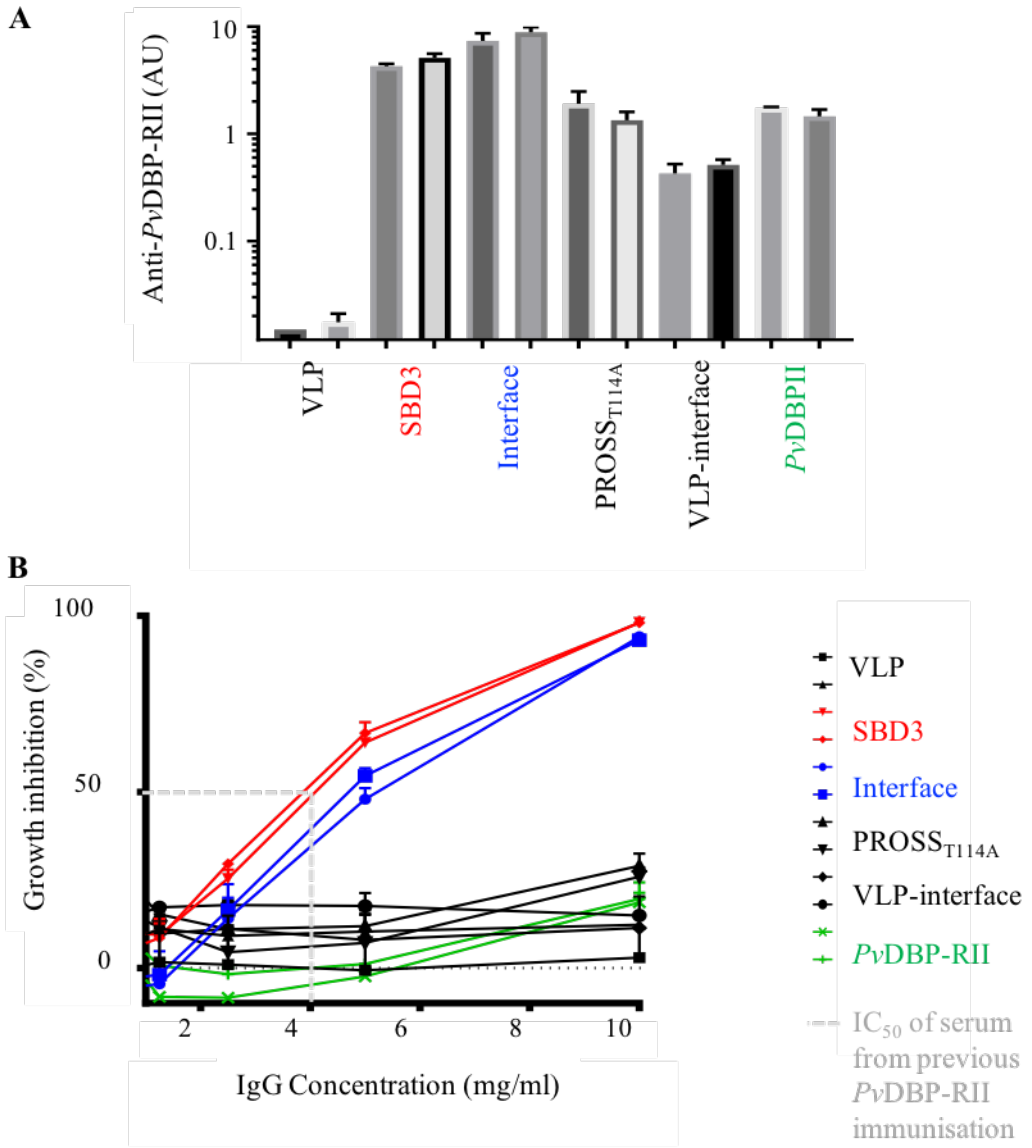


Figure 4.50: Immunisation Results Using Designed Immunogens A) An ELISA assay of immunised rabbit serum binding to recombinant *Pv*DBP-RII. The x-axis shows the sera came from rabbits immunised with the different immunogens. The y-axis shows binding of the sera to *Pv*DBP-RII. B) A growth inhibition assay in which purified IgG from each immunised rabbit was tested to see whether the presence of the IgG inhibited parasite growth. The dashed line represents the IC_{50} of the serum from a previous *Pv*DBP-RII rabbit immunisation.

From a previous attempt of rabbit immunisations, with the same concentration

of *Pv*DBP-RII and adjuvant injected, the sera in a GIA assay had an IC_{50} value of 4 mg/ml (grey dotted line on **Figure 4.50**). Using this as a benchmark it is clear that immunisations with Subdomain 3, and to a lesser extent Interface, led to the production of antibodies that are nearly as inhibitory as those produced by immunisations with *Pv*DBP-RII. In the future, to confirm and quantify this finding, as well as check the ability of the other constructs to induce the production of inhibitory antibodies, this experiment must be repeated.

4.4.10 Discussion of Immunogen Design

The main goal in malaria vaccine design is to make an immunogen that produces broadly inhibitory mAbs, which can stop growth of multiple different strains of malarial parasite. Elucidating the structure of a human broadly inhibitory mAb, DB9, in complex with *Pv*DBP-RII identified the Subdomain 3 region as an epitope that could elicit the production of these types of desirable antibodies to target *P. vivax* infection.

Immunisation with the whole *Pv*DBP-RII domain generates broadly inhibitory DB9-like mAbs, as well as mAbs against other surfaces of *Pv*DBP-RII, which may not be inhibitory. In order to "focus" the immune response into producing a greater proportion of the broadly inhibitory mAbs, immunogens based solely on the Subdomain 3 region were designed.

Mutations introduced at the hydrophobic interface with the other subdomains and the use of the PROSS strategy were used to increase the solubility and stability of the designed immunogens. These methods were successful since the resulting constructs, Interface and PROSS-T114A, were both solubly expressed, unlike the subcloned Subdomain 3 or the *Pv*DBP-RII constructs. All of the designed immunogens successfully mimicked the DB9 epitope, as demonstrated by the binding of these constructs

to DB9 with similar affinities to the *Pv*DBP-RII, as well as the Subdomain3:DB9 structure. Moreover, the melting temperature of PROSS was greatly increased compared to that of Subdomain 3, making this construct potentially better, more stable candidate for vaccine delivery.

In order to increase the size and immunogenicity of these small immunogens, the Interface construct was coupled to a VLP via a SpyTag:SpyCatcher system as has previously been done [220, 276]. The Interface immunogen was coupled in such a way that the desirable DB9 epitope would be most exposed on the VLP surface. Immunisations in rabbits were inconclusive since the sera taken from rabbits immunised with the positive control *Pv*DBP-RII did not show any growth inhibition. However, the Subdomain 3 and Interface sera did give a growth inhibitory response with similar IC_{50} values to sera from a previous *Pv*DBP-RII immunisation. This is promising but this experiment will have to be repeated to gain conclusive evidence to quantify this result. Furthermore, with IgG purified from these future immunisations the proportion of inhibitory antibodies compared to those generated by *Pv*DBP-RII immunisations could also be evaluated.

Overall, these studies and the literature highlight that the invasion of merozoites into reticulocytes via the *Pv*DBP:DARC interaction remains a good target for potential immunogens. This pathway is solely responsible for the majority of RBC invasion in *P. vivax* infection, and so targeting it will eliminate invasion, as demonstrated by the lack of *P .vivax* disease in Duffy negative individuals [231]. Here, the structural characterisation of the epitope of a strain-transcending inhibitory human mAb bound to *PV*DBP-RII enabled a set of new immunogens to be designed. These immunogens successfully encapsilated the epitope of the inhibitory antibody, shown by binding and structural studies, and they themselves produced sera that could inhibit *Pv*DBP:DARC-RII binding. Further immunisations and experiments on the antibodies produced, are needed to properly characterise both the quantity and

inhibitory "quality" of the antibodies produced. Structure-guided design based on the epitopes of inhibitory antibodies aims to produce a greater proportion of strain-transcending antibodies, whilst reducing the amount of non-inhibitory epitopes presented. For further understanding of this interaction and future vaccine design, more structural information about the DARC binding *Pv*DBP is required, ideally of the full-length constructs.

Chapter 5

Structural Studies of Var2CSA

5.1 Introduction to PAM and Var2CSA

5.1.1 Pregnancy-associated Malaria

Pregnancy-associated malaria (PAM) as the name suggests affects pregnant women and can cause severe clinical symptoms in the mother as well as disrupting fetal development [277]. Anaemia and hypertension in first-time pregnant women causes an increase in maternal morbidity and mortality [81]. Moreover, the fetus suffers from growth restriction, which can lead to low birth rate [278], infant anaemia [279] and increased risk of neonatal morbidity and mortality [280]. More than 50 million pregnancies each year are at risk to this *P. falciparum* malaria infection [281], and in 2015, 400,000 cases of maternal anaemia and 15% of maternal deaths globally were attributed to PAM [282]. Like other *P. falciparum* infections one of the major contributors to the severe symptoms of PAM is the cytoadherence of iRBCs to endothelial tissues. The complications arising from PAM have been attributed to the accumulation of iRBCs in the placenta via the binding to the carbohydrate chondroitin sulphate A (CSA) in the intravillous spaces. [82].

5.1.2 Var2CSA, an unusual *Pf*EMP1, Required for Placental Cytoadhesion

The protein associated with the binding to CSA has been attributed to an unusual *Pf*EMP1, called Var2CSA. Multiple lines of evidence support that Var2CSA is the PAM adhesion protein, including:

1. Var2CSA is the only *var* gene transcribed in CSA-binding laboratory isolates and iRBCs [283–285]
2. Knock-outs of the Var2CSA encoding *var* gene results in iRBCs unable to bind CSA [286], and the iRBCs are no longer sequestrated in the placental cryosections in these knockouts *in vitro* [287]
3. Antibodies against Var2CSA, gained either from multigravid women in endemic areas or through immunisations with recombinant protein provide protection against PAM [284, 288, 289]

Like other *Pf*EMP1s, the ectodomain of Var2CSA comprises of CIDR domains and DBL domains, however, it does not have the same domain organisation as other *Pf*EMP1s. Instead of a head region and tandem region of DBL and CIDR domain repeats, as described in **Section 3.1**, Var2CSA has six DBL domains and one CIDR-like domain **(Figure 5.1)** [283]. Between the DBL1 and DBL2, and DBL2 and DBL3 domains there are interdomain (ID) regions, and before the identification of the CIDR-like domain, the sequence between DBL2 and DBL3 was split into two regions called ID2a and ID2b **(Figure 5.1)**.

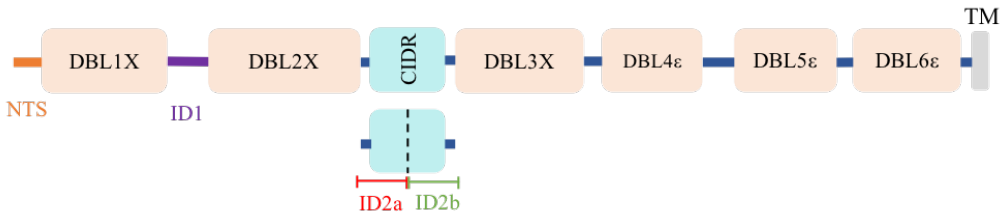


Figure 5.1: Architecture of Var2CSA The overall organisation of the DBL and CIDR-like domain in Var2CSA. The subclass of DBL is also labelled with X=unknown and ϵ=epsilon. NTS=N-terminal sequence, ID1=interdomain1, ID2a=interdomain2a, and ID2b=interdomain2b. The width of the box representing each domain and interdomain is proportional to the number of amino acids within each region.

5.1.3 Var2CSA Binding to the Carbohydrate CSA

Unlike other *Pf*EMP1s, Var2CSA binds a carbohydrate CSA rather than a protein endothelial receptor. CSA is present in the placenta as glycosaminoglycan (GAG) side chains covalently linked to a core protein to form a chondroitin sulphate proteoglycan (CSPG). The binding of infected erythrocytes (expressing the full-length Var2CSA ectodomain) to the placenta endothelium has been shown to be highly specific to CSA and is not inhibited by other GAGs [290]. In an attempt to determine the CSA-binding site of Var2CSA that provides this specificity, the binding of CSA to individual domains has been characterised. The individual domains DBL2X, DBL3X, DBL5ϵ, and DBL6ϵ, have been shown to bind to CSA [291–293]. However, this binding has been shown to be non-specific, with the individual domains binding other highly charged and sulphated GAGs sometimes with higher affinity than they bind to CSA [294]. As a result it is hard to predict how much information can be gathered from these *in vitro* binding tests in relation to binding *in vivo*, with many false positives being observed. Moreover, the affinity with which all the individual domains (except DBL2) bind to CSA, is >1000-fold lower than the affinity of the full-length protein **(Table 5.1)**. This together with small angle X-ray scattering (SAXS) data, which predicts the ectodomain to form a highly compact structure [295, 296], points

Recombinant Construct	K_D (μM)
Full-length ectodomain (DBL1-DBL6)	0.127
DBL2	5.3
DBL3	344
DBL5	150
DBL6	92
DBL1-DBL2	0.77
DBL1-CIDR	1.58
DBL1-DBL3	0.58
DBL1-DBL5	0.115

Table 5.1: Affinity of Recombinant DBL Constructs for CSPG
All of these affinity measurements were determined by SPR experiments in the following publications [295, 298, 299]

to the requirement of higher domain organisation involving the whole ectodomain to provide specific and high affinity nanomolar CSA binding [199, 295, 297]. The low micromolar affinity for CSA shown by VarCSA, as well as further CSA-binding studies, have been used to narrow down the binding site region to DBL1-DBL2 , with the major interaction site comprising of DBL2X and potentially $CIDR_{PAM}$ [199, 298]. These fragments bind *in vitro* with nanomolar affinity comparable to the full-length Var2CSA ectodomain, and so may represent the core binding site **(Table 5.1)**.

5.1.4 Structural Investigations of Var2CSA Domains

The full-length structure of the ∼350kDa Var2CSA ectodomain has not been elucidated. The only information available about its overall higher-order structure is from low resolution SAXS data, which shows a compact domain organisation [295, 296, 300]. The crystal structures of the DBL3X [301, 302] and DBL6ϵ [297] individual domains from Var2CSA have been solved, and both of these domains have

the same architecture as previously solved DBL domains, as expected. In addition to the individual domains, the structure of the double DBL3X-DBL4ϵ construct has been solved [303]. The conserved packing interface between theses two domains provides clues to how the compact organisation of the overall Var2CSA may be achieved with the DBL domains packing against one another **(Figure 5.2)**.

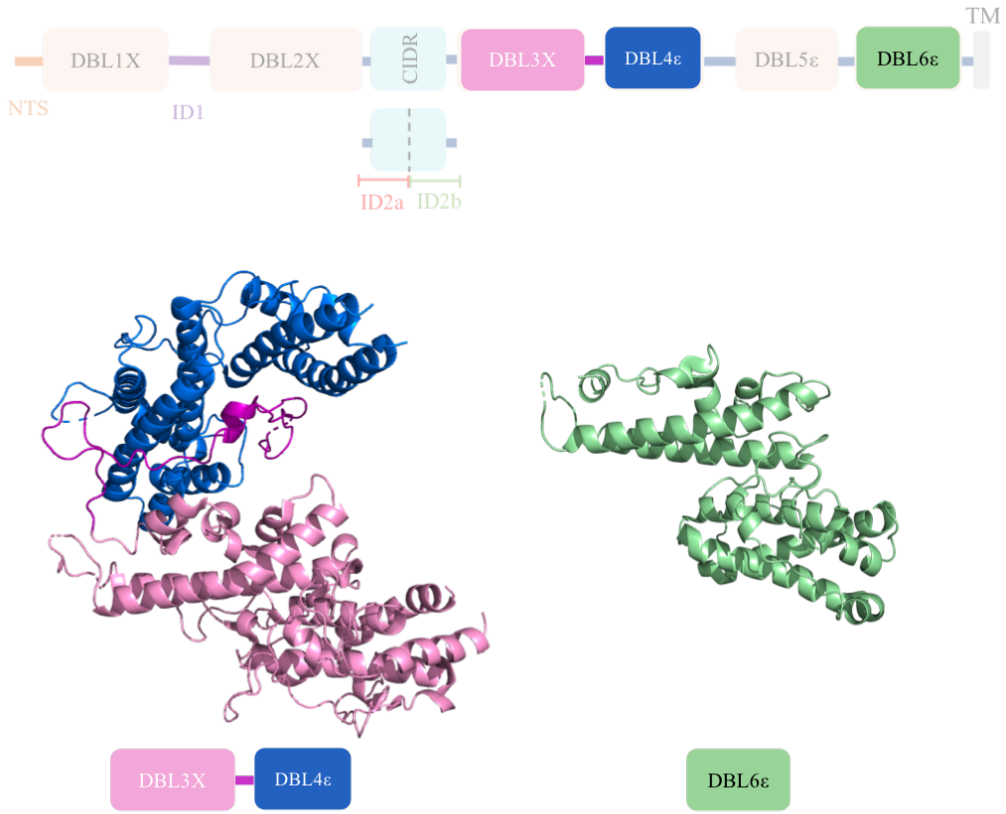


Figure 5.2: Solved Structures of Var2CSA DBL Domains Cartoon representation of the solved crystal structures of DBL3X-DBL4ϵ (PDB code: 4P1T) and DBL6ϵ (PDB code: 2WAU) constructs.

5.1.5 Sequence Conservation of Var2CSA

Like other *Pf*EMP1s, Var2CSA has been shown in multiple studies to have great sequence and antigenic polymorphism [283, 304–306]. However, an extensive analysis of over 100 Var2CSA sequences identified that these individual polymorphisms are often linked with other polymorphisms and can form different correlated segments

[307]. As a result, when looking at the correlated segments, the diversity between strains is much reduced. Even the monkey infective *P. reichenowi*, which is thought to have diverged from *P. falciparum* 6-10 million years ago, shows the same correlated segment polymorphisms. This suggests that rather than individual polymorphisms arising, there is slow mutation of the different segment types, which creates the combinatorial diversity seen in Var2CSA [307].

5.1.6 Var2CSA as a Vaccine

This reduced diversity caused by overlapping polymorphisms between geographically diverse parasite isolates [307], as well as the presence of only one Var2CSA gene found in all isolates [286, 308] gives hope that a strain-transcending vaccine can be designed and produced. Furthermore, naturally acquired immunity after the first or second pregnancy to PAM has been associated with the production of cross-reactive adhesion-blocking antibodies that target Var2CSA [288, 289, 309, 310]. This feature also makes Var2CSA a promising vaccine candidate. Only eight human mAbs, that interact with Var2CSA-expressing iRBCs, have been isolated from multigravid women who have developed natural immunity to PAM [304]. Further epitope mapping showed that four of the mAbs bind to DBL5ϵ, and three to DBL3X. The last mAb, PAM1.4, although it binds to full-length Var2CSA, it does not bind to single domains and is believed to recognise a conformational epitope **(Figure 5.3)**.

These studies showed that these isolated mAbs exhibit a degree of strain-specificity since they do not recognise Var2CSA on iRBCs from multiple different strains [304]. Moreover, only three mAbs (PAM1.4, PAM3.10, and PAM8.1) inhibit binding of Var2CSA-expressing iRBCs to CSA, however this level of inhibition is relatively low at ∼15-25% [311]. Combining all of the mAbs together leads to an increase in inhibition to ∼40%, although this is less than half the inhibition achieved by soluble CSA [311]. These investigations highlight that these mAbs (that bind to several

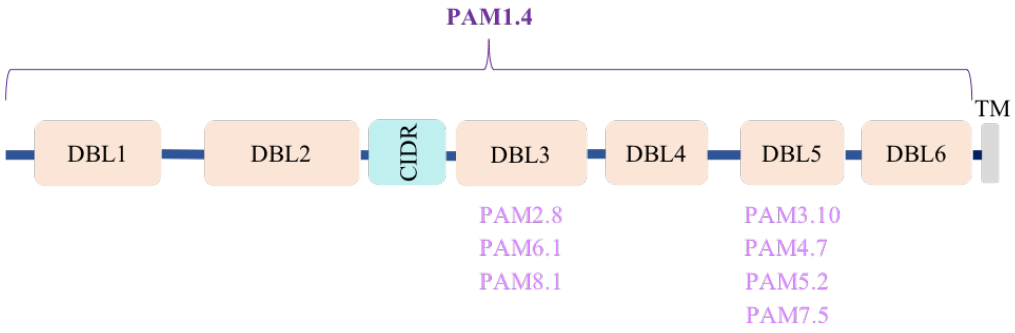


Figure 5.3: mAbs against Var2CSA The human mAbs are labelled in purple and the domains of Var2CSA which they bind are indicated by the brackets. The precise binding site of PAM1.4 is not known but it is known to bind a multi-domain conformational epitope.

epitopes) can together inhibit the binding of Var2CSA to CSA. However, a single highly cross-reactive and inhibitory mAb from naturally immune individuals has not yet been isolated.

The difficulties of purifying the full-length ectodomain of Var2CSA have thus far made designing a vaccine, that would be able to produce broadly inhibitory mAbs, a hard task. As a result individual domains have been previously used as immunogens, and although they produce polyclonal antibodies that interact with CSA-binding isolates, most do not block adhesion [312–315]. Small multidomain fragments have been shown to induce the production of antibodies that interact with homologous parasite surfaces more highly than the individual domains [316]. This is consistent with the *in vitro* binding data, which predicts that the individual domains do not exhibit the full CSA-binding phenotype. The N-terminal region of Var2CSA has consistently been used as an immunogen to promote the production of anti-adhesion antibodies [296, 316–318]. In particular, the NTS-DBL2 region has been shown to produce antibodies that block the binding of multiple isolates to CSA by >50% and so is of particular interest [317].

The current leading vaccine candidate, PAMVAC, is based on a ID1-ID2a construct. This construct, upon immunisation, induces the production of strain-transcending

inhibitory antibodies that block adhesion of Var2CSA-expressing cells to CSA [319]. However, it is expected that once a better understanding of the overall Var2CSA ectodomain structure and its binding to CSA is gained, a vaccine that targets adhesion more directly and with greater efficacy could be produced.

5.1.7 Multimeric IgM Binding to Var2CSA

Natural non-specific IgM binds to some strains of parasite iRBCs via *Pf*EMP1s, including Var2CSA, expressed on the cell surface [320, 321]. This IgM:*Pf*EMP1 interaction is mainly associated with the rosetting of multiple uninfected RBCs around an iRBC, which leads to impeded blood flow [322]. However, rosetting has rarely been observed in CSA-binding PAM isolates, suggesting that IgM may play a different role when binding to Var2CSA-expressing iRBCs [323, 324]. It is thought that IgM, via Var2CSA, bound to the iRBCs could cross-bridge to host IgM-binding receptors expressed on placental cells, thus leading to increased sequestration of the parasitised cells [325]. Furthermore, incubation of Var2CSA-expressing iRBCs with IgM prevents the subsequent labelling by some mAbs that target DBL3 and DBL5 [326]. This suggests that IgM has evolved to bind to immunogenic epitopes that are not critical for CSA-binding, blocking IgG binding and hence reducing IgG induced phagocytosis. However, not all mAbs tested are blocked by IgM binding, namely PAM1.4, which binds a conformational epitope [326]. If the epitope of mAbs like this could be determined, then vaccines could be designed to target inhibitory epitopes that are available despite non-specific IgM binding. Various studies have attempted to identify the regions in Var2CSA responsible for IgM binding and have been reported to involve DBL2X, DBL6ϵ, and particularly DBL5ϵ [321], or only DBL6ϵ [327]. It is therefore unclear where IgM binds to Var2CSA. Further knowledge of the epitopes bound by non-specifc IgM (and inhibitory mAbs) will be crucial for the development of a vaccine that will lead to the production of effective inhibitory

mAbs *in vivo*.

5.1.8 Initial Structural Studies of Full-length Var2CSA

In an attempt to gain more information about the full-length Var2CSA structure, previously in the Higgins laboratory an NTS-DBL6 construct was cloned into the pAcGP67A vector and expressed in a lepidopteran cell/baculovirus expression system **(Materials and Methods 2.2.3)**. Initial structural studies of this full-length construct included setting up negative stain electron microscopy (EM) grids to gain an insight of the overall architecture **(Figure 5.4)**.

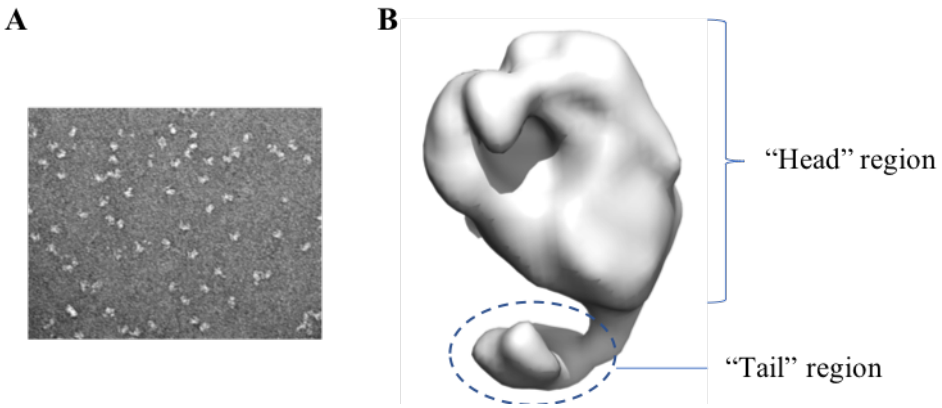


Figure 5.4: Negative Stain of NTS-DBL6 A) Representative negative stain micrograph and B) reconstruction of NTS-DBL6 construct

The negative stain reconstruction showed a compact main structure ("head") with a smaller globular ("tail") region protruding from it, which was consistent with the overall volume of the SAXS data already obtained [295]. From this and other data that indicated that immunisation with DBL5 could elicit the production of cross-reactive mAbs [328], and DBL6 was not required for the overall binding to CSA [329], it was hypothesised that this smaller globular "tail" region may be the DBL6 domain. As a result, in an attempt to produce a more globular structure, which may be more likely to crystallise, an NTS-DBL5 construct was also cloned and expressed in this system.

5.2 Structural Studies of Full Length Var2CSA

5.2.1 Rational

Multiple sources of evidence suggested the need for the full-length Var2CSA ectodomain both for CSA-binding and for the production of adhesion-blocking antibodies upon immunisation. Structural information about the overall higher-order organisation of this large protein would greatly help further the understanding of both CSA-binding and reveal which regions of the protein the adhesion-blocking antibodies target. This knowledge could then be used in the design of a vaccine to block this interaction which is crucial for the sequestration of iRBCs in the placenta in PAM.

In initial studies in the laboratory carried out by Professor Matthew Higgins, the expression and purification protocol for both the NTS-DBL5 and NTS-DBL6 constructs was optimised. Also, initial negative stain EM had been carried out on NTS-DBL6, and several crystal plates were set up for both constructs but no crystals were produced.

5.2.2 Project Aims

- Set up broad crystallisation trials for both the NTS-DBL6 and NTS-DBL5 constructs
- Optimise cryoEM grid preparation and data collection for the NTS-DBL6 and NTS-DBL5 constructs
- Identify and design new constructs that may be more likely to crystallise

5.2.3 Declaration

Professor Matthew Higgins had previously cloned the Var2CSA NTS-DBL6 and NTS-DBL5 constructs into the pAcGP67A vector, which contained an N-terminal His-tag for affinity purification.

5.2.4 Var2CSA construct expression and purification

The NTS-DBL6 and NTS-DBL5 constructs were expressed in a Lepidopteran cell/baculovirus expression system as described in **Materials and Methods 2.2.3**. The NTS-DBL5 domain expressed at much higher levels than that of NTS-DBL6 (4 mg/L compared to 1 mg/L).

The supernatant of the High-Five cells was collected four days after inoculation, and was concentrated and buffer-exchanged by TFF. The concentrated supernatant was applied onto a Ni-NTA column and the Var2CSA constructs purified via their His-tag **(Materials and Methods 2.4.3)**. The eluted fractions were further purified by size exclusion to remove any aggregates (Superose 6 10 30, GE Healthcare). In order to gain a single monodisperse peak, the monomer peak had to be pooled and re-applied to the SEC column **(Figure 5.5)**.

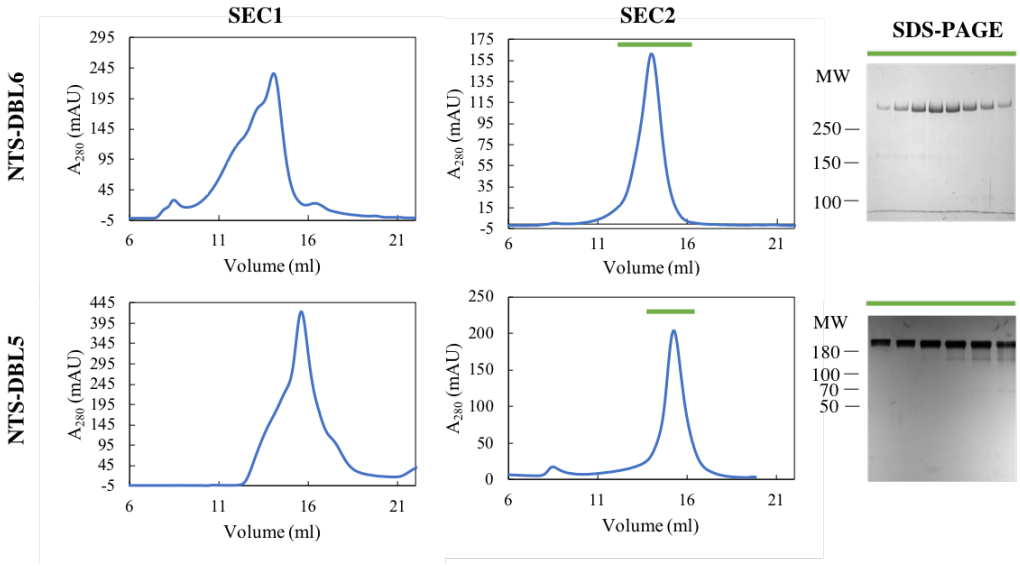


Figure 5.5: Purification of NTS-DBL6 and NTS-DBL5. The top panel shows the purification of the NTS-DBL6 construct, and the bottom panel of the NTS-DBL5 construct. The first column shows the first size exclusion chromatography traces (SEC1), and the second column shows the size exclusion traces of the initial monomer peak pooled and run over the SEC column again (SEC2). The third column shows the SDS-PAGE gels of the fractions from the corresponding SEC2 (SDS-PAGE). The coloured bars demonstrate which part of the peak represents which region of the gel, and the fractions that were pooled fractions for future experiments.

5.2.5 Attempts to Crystallise Var2CSA

5.2.5.1 Initial Crystallisation Screens

Purified NTS-DBL5 and NTS-DBL6 were concentrated to between 9 mg/ml and 15 mg/ml, deglycosylated **(Materials and Methods 2.4.7)**, and broad crystallisation screens were set up as described in **Materials and Methods 2.7.2.1**. Since the Var2CSA ectodomain is very lysine-rich with lysine representing 12.5% of the overall sequence, methylation of these long, flexible side chains could greatly reduce this flexibility and so aid crystallisation. Therefore, the constructs were also methylated in order to promote crystal formation **(Materials and Methods 2.4.9)**. In total, ∼20 96-well screens were set up for the methylated and/or deglycosylated NTS-DBL5

construct at either 4°C or 18°C, equating to nearly 2000 different conditions. Only five methylated and deglycosylated plates at 4°C could be set up for the NTS-DBL6 construct due to the lower expression levels. However, despite multiple different plates for both constructs being set up, no crystals were observed.

5.2.5.2 Making a Complex with CSA

In an attempt to promote crystallisation, a complex of NTS-DBL5 was made with its carbohydrate binding partner, CSA. Since it has been shown that multiple domains are required for CSA binding, it was thought that the CSA-binding might stabilise the overall fold of the protein, making it less flexible and therefore more prone to crystallise.

CSA consists of a repeating dimer unit (β-glucuronic acid-(1-3)-N-acetyl-β-galactosamine-4-sulfate) **(Figure 5.6)**, which was purchased as a polysaccharide from Sigma Aldrich. Previous experiments have indicated that the size of the CSA fragment that binds to Var2CSA was ∼12 dimer units [330]. The excess chain of polysaccharide that did not directly bind to the Var2CSA would be highly flexible and reduce the chance of crystallisation. Therefore, this excess CSA chain was trimmed away by incubation with Chondroitinase ABC (Sigma Aldrich) **(Materials and Methods 2.4.10)**.

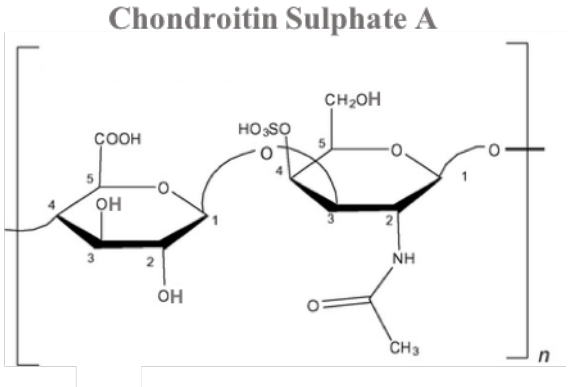


Figure 5.6: Chondroitin Sulphate A Architecture CSA exists as an alternating copolymer of β-glucuronic acid-(1-3)-N-acetyl-β-galactosamine-4-sulfate.

After mixing the CSA polysaccharide with the Var2CSA construct, it was incubated with chondroitinase ABC for 30 minutes at 37°C before being applied to SEC. The CSA, as a carbohydrate could not be seen on an SDS-PAGE gel, nor could a shift be seen on the chondroitinase-treated sample SEC trace, since the molecular weight of the cleaved CSA was very small compared to the Var2CSA protein. However, if the Var2CSA in complex with the CSA polymer was applied to the SEC column then a shift to left was seen, this was most likely due to multiple Var2CSA constructs binding to one CSA polymer leading to much larger multimers. This hypothesis was further supported by the fact that on addition of chondroitinase ABC to the Var2CSA:CSA polymer mix, the peak shifted back to the position of the monomer, since the linker regions not protected by the binding to Var2CSA could be cleaved **(Figure 5.7)**.

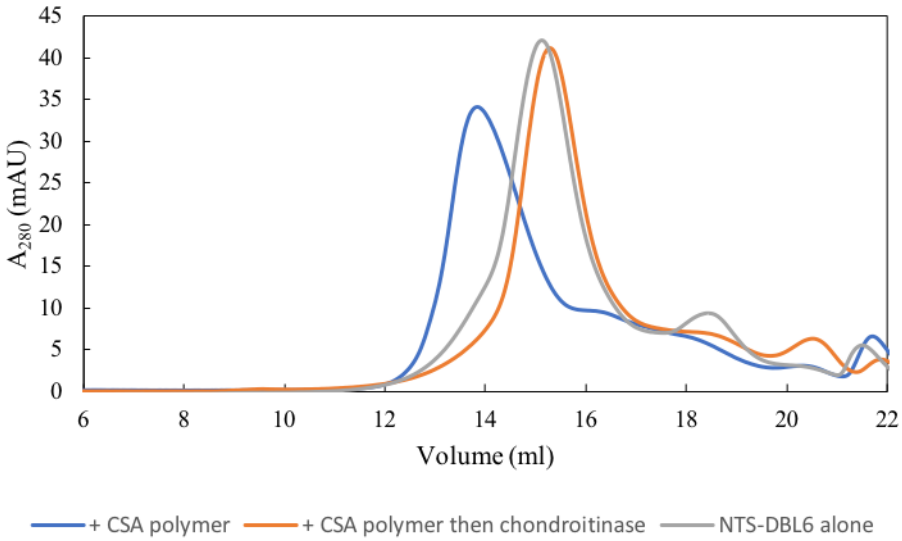


Figure 5.7: Complex formation of Var2CSA with CSA Before and After Chondroitinase Treatment The SEC traces for NTS-DBL6 alone (grey), NTS-DBL6 plus CSA polymer (blue) and the NTS-DBL6 plus CSA polymer and chondroitinase incubation (orange).

A complex of NTS-DBL5:CSA was produced in this way, concentrated and the following broad crystallisation screens set up at 4°C: MIDAS, Morpheus, JCSG+, PACT Premier, and ProPlex. However, no crystals were observed in the time-frame

of this project.

5.2.5.3 Making a Complex with PAM1.4

Another way to promote crystallisation can be to make a complex with a Fab fragment as a crystallisation chaperone. Similar to a binding partner, it is thought that the Fab fragment may stabilise the fold and reduce flexibility of the protein, or allow new crystal contacts to form. One previously characterised mAb, PAM1.4, which has been shown to bind to a conformational epitope comprising of multiple domains of Var2CSA, was kindly provided by Lea Barford, University of Copenhagen [304]. Fab fragments were generated by papain cleavage of the PAM1.4 mAb, followed by the separation of the Fab and any uncleaved mAb or Fc fragments using protein A resin **(Materials and Methods 2.5.2)**. The purified Fab fragments were present in the Protein A flow-through and wash **(Figure 5.8)**.

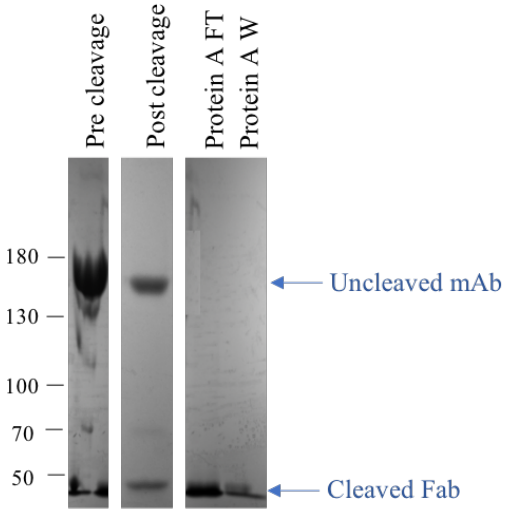


Figure 5.8: Formation of PAM1.4 Fab Non-reduced SDS PAGE gels with the first two lanes showing PAM1.4 mAb before (Pre cleavage) and after (Post cleavage) papain cleavage. The post cleavage fraction was then mixed with Protein A resin and the flow through (FT) and wash (W) fractions shown in the last two lanes.

The NTS-DBL6 and NTS-DBL5 constructs were mixed with the purified PAM1.4 Fab fragment (in a 1:1 ratio), and incubated for 30 minutes at RT, before being

applied onto the Superose 6 10/30 (GE Healthcare) **(Figure 5.9)**. The complex peak was pooled and concentrated to ∼10 mg/ml and crystal plates were set up. A full initial screen of the five customary plates were set up for the NTS-DBL5:PAM1.4Fab complex **(Materials and Methods 2.7.2.1)**. Once more not enough protein was obtained to set up a full screen of the NTS-DBL6:PAM1.4Fab complex and only a couple of plates could be made. Unfortunately again no crystals were produced.

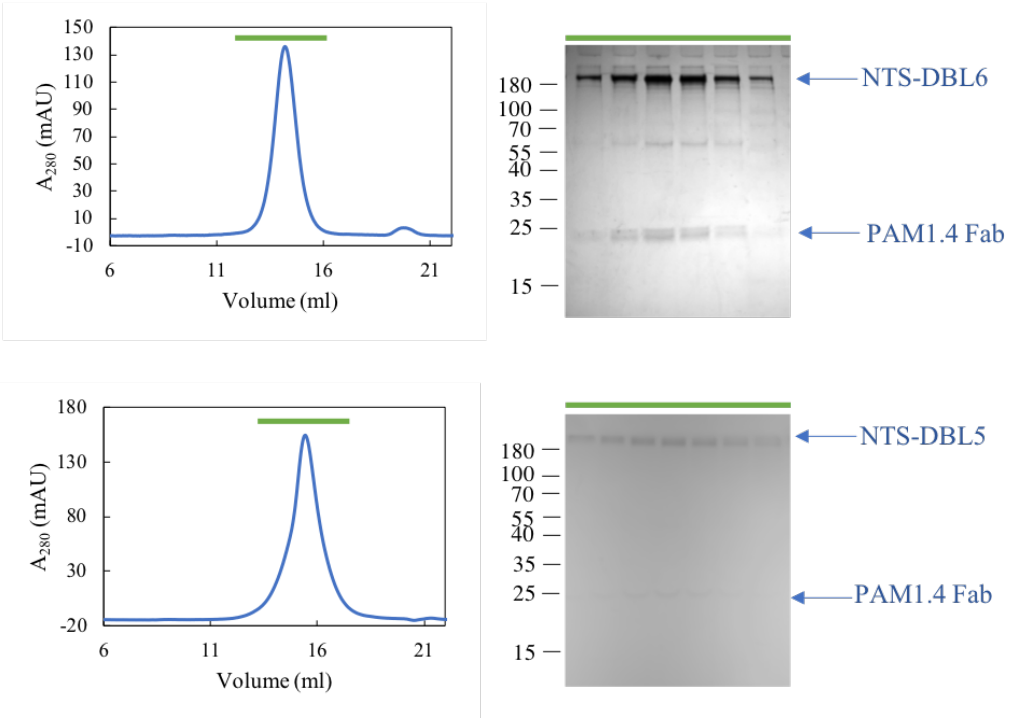


Figure 5.9: Formation of NTS-DBL6:PAM1.4 Fab and NTS-DBL5:PAM1.4 Fab Complexes The SEC trace (left) and corresponding reduced SDS-PAGE gel (right) for the NTS-DBL6:PAM1.4Fab complex (top) and the NTS-DBL5:PAM1.4Fab complex (bottom). The green bar represents the corresponding areas between the SEC trace and the gel, and also the fractions which were pooled for further studies.

5.2.6 Attempting Full-length Var2CSA Structure Determination by CryoEM

In recent years, during the electron cryo-microscopy (cryoEM) revolution [331], more and more structures of proteins <500kDa have been solved to high resolution (≤3Å) using this technique [332–334]. Therefore, it was hoped that structural information about the overall architecture of the NTS-DBL6 (∼310kDa) construct and/or the NTS-DBL5 (∼265kDa) construct could be obtained through this method.

The purified NTS-DBL6 and NTS-DBL5 constructs were applied onto grids and vitrified in liquid ethane as described in **Materials and Methods 2.8.1.2**, before the grids were clipped and loaded onto either the Talos Arctica 200kV and the Titan Krios 300kV microscopes **(Materials and Methods 2.8.1.4)**. Multiple different construct concentrations, construct manipulations (for example addition of CSA and/or PAM1.4Fab as described in **Section 5.2.5**) and different grid types were screened to try to optimise the ice thickness and distribution of particles on the grids.

5.2.6.1 Initial Screens and Optimisation on Holey Grids

Initial screens were carried out on Holey Quantifoil R 1.2/1.3 Cu 200 grids. For both the NTS-DBL6 and NTS-DBL5 samples, the ice thickness was optimised by using the grid preparation and freezing conditions seen in **Table 5.2**.

The first experiment compared methylated NTS-DBL6 **(Materials and Methods 2.4.9)** with non-methylated NTS-DBL6, since altering the surface properties of a protein can greatly affect the distribution of particles within the holes. For example, more hydrophobic patches of the protein interact more favourably with the air-water interface than polar regions. Visual inspection of the micrographs showed that on the

Grid	Quantifoil R 1.2/1.3 Cu 200
Glow Discharge	60s at 15mV

Vitrobot Settings	
Humidity	100%
Temperature	20°C
Wait Time	30s
Blot Time	3s

Table 5.2: Optimised Grid Preparation and Vitrobot Freezing Parameters

non-methylated grids the particles aggregated into patches or "rafts" in certain areas of the hole, leaving other parts of the hole completely empty of particles **(Figure 5.10 A)**. However, when the NTS-DBL6 was methylated the protein was more uniformally distributed across the holes of the grid, with discrete particles discernible for picking **(Figure 5.10 B)**. As a result, all future experiments used methylated protein.

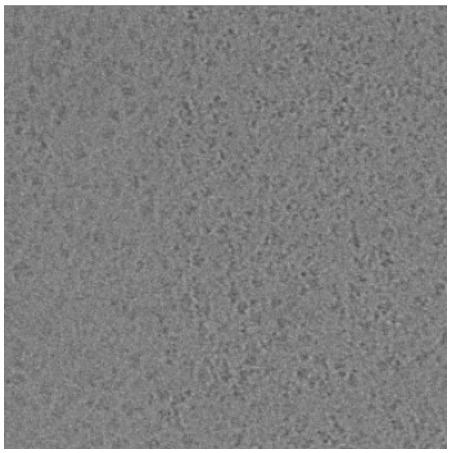


Figure 5.10: Representative Micrographs Comparing Non-methylated vs Methylated NTS-DBL6 Samples Examples of "rafts" of aggregated particles are highlighted with a red arrow. Empty spaces with no particles are highlighted with a yellow arrow. Examples of discrete individual particles which can be picked are demonstrated by green circles.

Both the NTS-DBL6 and NTS-DBL5 methylated constructs were applied to the grids, alone and in combination with binding partners. The sample concentration of 0.3 mg/ml for both NTS-DBL6 and NTS-DBL5 constructs allowed for evenly but densely distributed particles across the grid.

Larger datasets were collected for both NTS-DBL6 and NTS-DBL5, to allow comparison of the 2D classes and 3D reconstructions generated by each construct **(Materials and Methods 2.8.3)**. Due to the relatively small size of the constructs (NTS-DBL6: 310kDa, NTS-DBL5:265kDa), the PAM1.4Fab was added to increase the overall size of the particles in the hope that this would not only make particle picking easier, but also aid 2D classification. A uniform distribution of discrete particles was seen in the holes of the grids for both constructs plus PAM1.4Fab **(Figure 5.11)**.

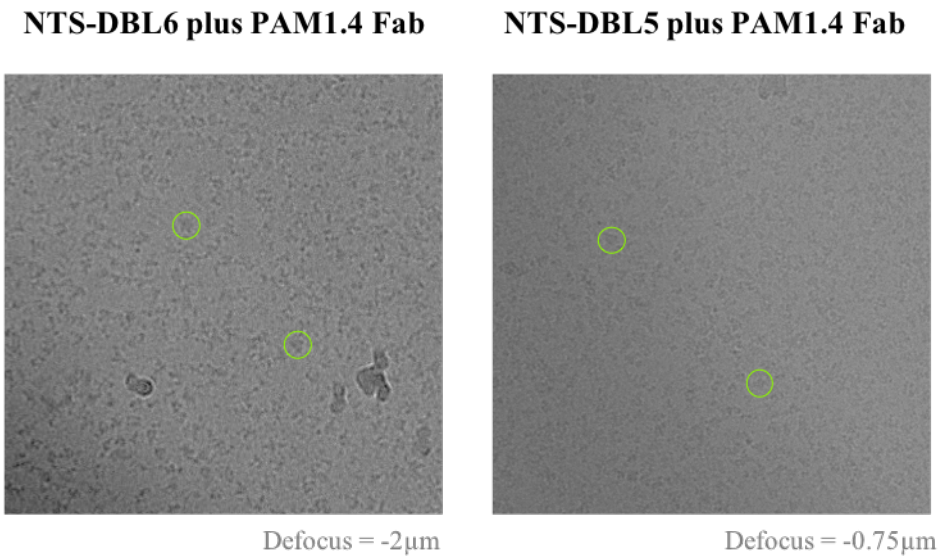


Figure 5.11: Representative Micrographs for NTS-DBL6 plus PAM1.4Fab vs NTS-DBL5 plus PAM1.4Fab Examples of individual particles that could be picked are shown by the green circles. The NTS-DBL6 plus PAM1.4Fab micrograph was taken at a defocus of -2μm compared to the NTS-DBL6 plus PAM1.4Fab micrograph at -0.75μm, explaining the lower contrast seen the latter.

The micrographs were motion and CTF-corrected using the SIMPLE preprocessing command **(Materials and Methods 2.8.4)**. The particles were initially manually picked from 200 processed migrographs using EMAN box picker tool. The particles

from the remaining micrographs were automatically picked using the manually picked particles as references by a program in SIMPLE. Once all the particles had been picked, they were extracted and classified by RELION into 2D classes representing particles vitrified in different orientations in the ice. These 2D classes therefore represented different projections of the overall 3D shape. The top fifteen 2D classes, ranked by the number of particles that make up that class, for both the NTS-DBL6 plus PAM1.4Fab and NTS-DBL5 plus PAM1.4Fab can be seen in **Figure 5.12**. The main represented view, examples of which are highlighted by the red boxes, looked similar for both samples. This 2D class also had the most fine detail shown by the presence of lines within the density representing helices.

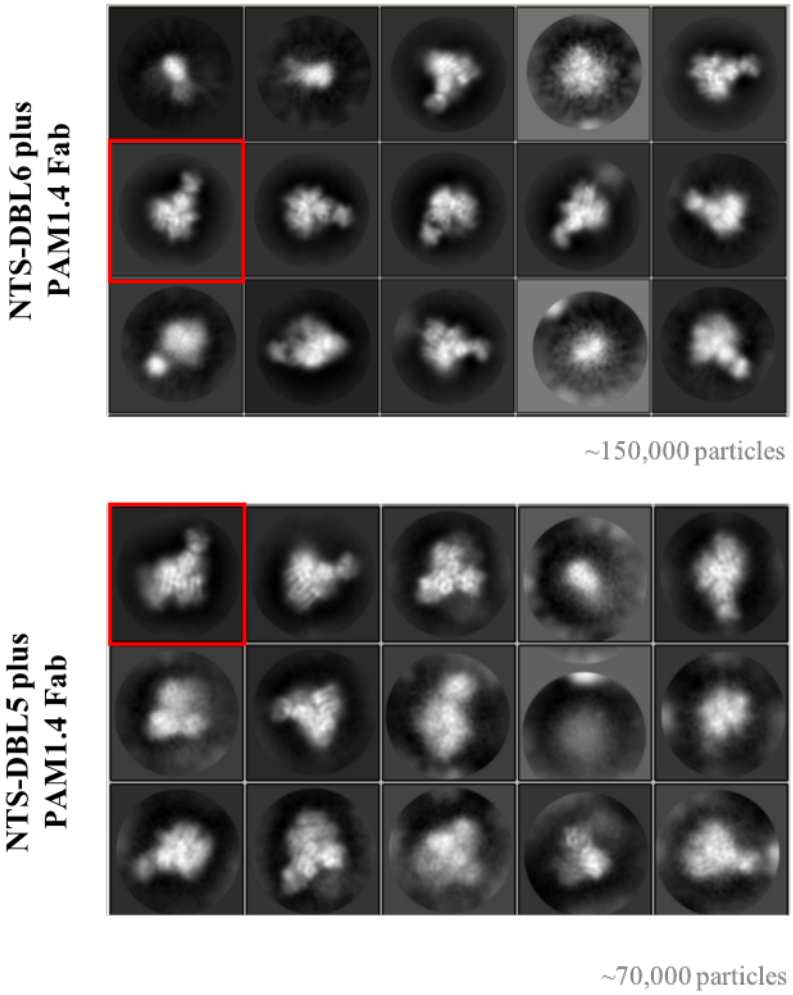


Figure 5.12: 2D Reconstructions for NTS-DBL6 vs NTS-DBL5
The 2D classes were sorted by number of particles contributing to each class with the classes with the most particles found in the top left corner and reading across the row. The red boxes highlight an example of the most represented, predominant view. Only the top fifteen classes are shown.

The 2D class averages were then aligned to produce a 3D reconstruction in RELION **(Materials and Methods 2.8.4)**. The 3D reconstructions showed a "front" view, which looked similar to the highlighted predominant 2D class, had a definite shape, with the discrete domains discernible **(Figure 5.13)**. Indeed, the crystal structures of DBL domains could be fitted into the NTS-DBL6 reconstruction, showing that the density was approximately the right size **(Figure 5.14)**.

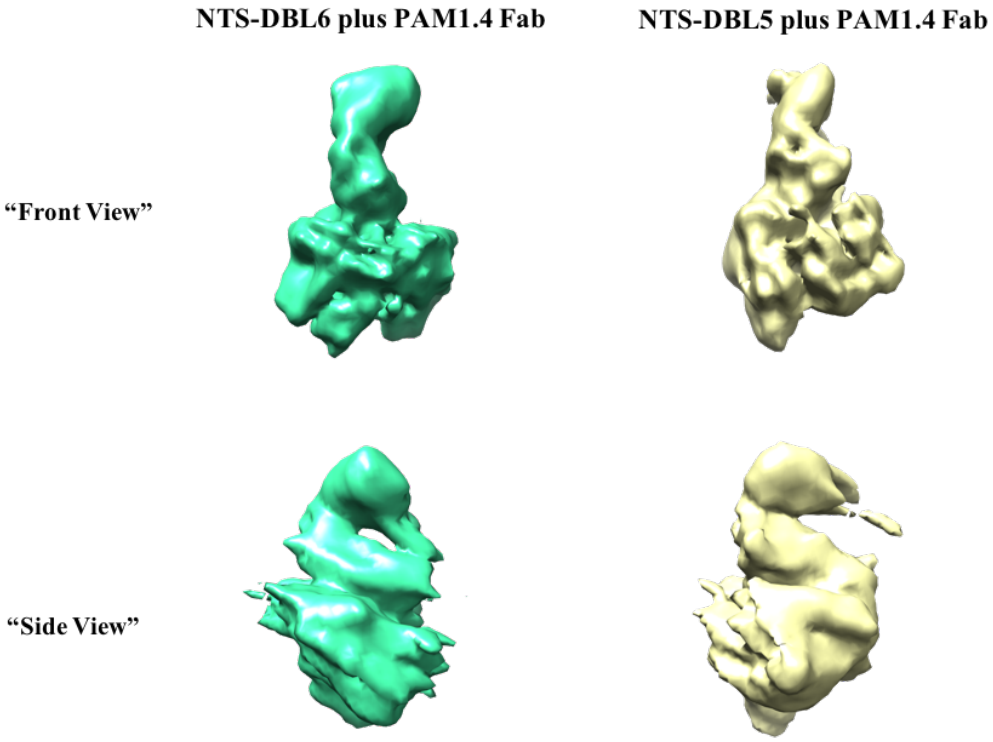


Figure 5.13: 3D Reconstructions for NTS-DBL6 vs NTS-DBL5 The electron density maps for the 3D reconstructions of NTS-DBL6 (green) and NTS-DBL5 (yellow) plus PAM1.4Fab at two different views rotated by 90°around the y-axis.

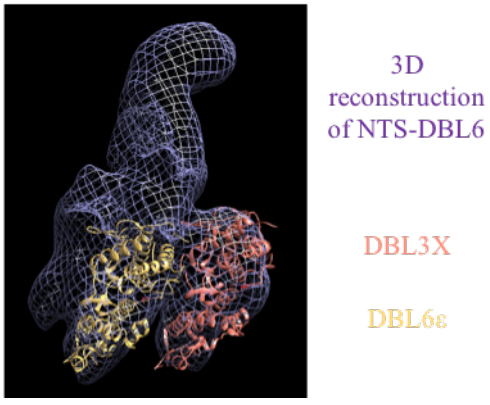


Figure 5.14: DBL domain size in reference to Var2CSA Cry-oEM Density The 3D reconstruction of NTS-DBL6 shown as a purple mesh. The solved crystal structures of DBL3X (PDB code: and DBL6ϵ (PDB code: 2WAU) constructs.

However, when the 3D reconstruction was rotated 90° around its y-axis, the "side view" appeared stretched and had very few details **(Figure 5.13)**. This was due to missing information due to a lack of 2D views from this direction. This lack of

certain views was likely due to the particles in the vitrified ice not being randomly oriented and not representing the sample as it would be in solution. The particles oriented themselves preferentially in the ice, potentially due to surface properties such as charge or overall shape, meaning that some 2D views like the "front views" were represented a lot in the sample and other views like the "side views" were barely represented at all, as was clear in the 2D classes (**Figure 5.12)**.

One observation that could be made by comparing the 3D reconstructions of the NTS-DBL6 sample compared to the NTS-DBL5 sample was that they looked remarkably similar **(Figure 5.13)**. The small globular "tail" protruding from the main "body" of the density was present in both the NTS-DBL5 and NTS-DBL6 reconstructions, showing that this "tail" region was not DBL6. This may explain why no improvement in the crystallisability of the NTS-DBL5 construct compared to the NTS-DBL6 construct has been seen.

In order to check that the addition of PAM1.4 did not contribute to the preferred orientation of the particles, smaller datasets (∼20,000 particles) were collected for NTS-DBL6 with and without the addition of PAM1.4Fab. Looking at the 3D reconstructions, the "side" views again were very stretched and had little detail compared with the "front" and "back" views for both constructs **(Figure 5.15)**. Therefore, the sample without PAM1.4 also had preferred orientation on the grid. In the plus PAM1.4Fab sample, there was no obvious extra density in the "front" or "back" views compared with the minus PAM1.4Fab sample. The stretched and streaky nature of the "side" views made it hard to interpret whether the protruding density was "real" or not, especially since this was not seen for the larger datasets **(Figure 5.13)**.

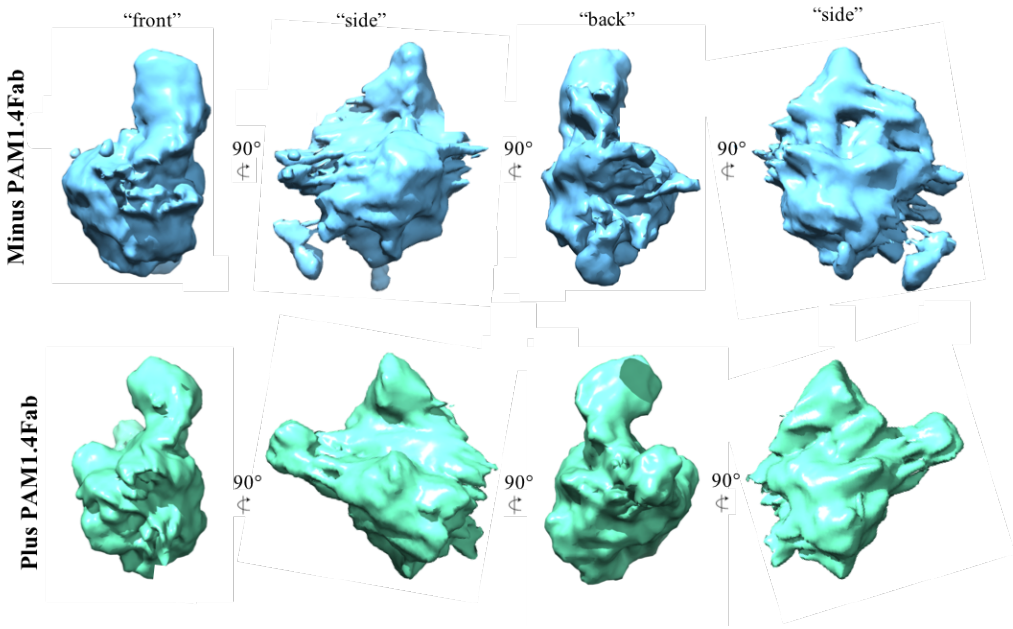


Figure 5.15: 3D Reconstructions for NTS-DBL6 Plus and Minus the Addition of PAM1.4Fab The 3D reconstructions seen from different views rotated 90°around the y-axis

5.2.6.2 The Preferred Orientation Problem

Particles having preferred orientation when vitrified on grids is a fairly common problem in cryoEM [335, 336]. In order to increase the signal to noise ratio of the particles to the background in cryoEM, it is a requirement to have very thin films of vitrified ice. These thin layers have very large surface:volume ratios (especially in holey grids) and so a protein diffusing freely within this layer will collide with the air-water interface >1000 times per second [337]. When grid freezing, even if the vitrobot is set to have a blot time of 1 second, then there will be a few seconds between sample application and vitrification in liquid ethane, meaning each particle will collide with the air-water interface at least a few thousand times. Hydrophobic patches on the surface of the protein will favourably interact with the air-water interface, leading to certain views being represented more frequently than others in the micrograph images. Moreover, there is evidence that upon interaction with the air-water interface, partial or complete denaturation of the protein may occur [338, 339]. Therefore, to

overcome the problem of preferred orientation, the interaction of the proteins with the air-water interface in the thin layer of ice must be reduced. There are multiple different ways that this has successfully been achieved [340], including:

1. **Changing the surface properties of the protein**

   This can be achieved by changing the pH and other buffer components, which may affect protein stability and surface properties [341]. Moreover, methylation and the addition of binding partners can change the overall surface properties of the protein. In addition, recently, the use of PEGylated samples has had a striking effect on both the distribution and the angular spread of proteins on cryoEM grids [342].

2. **Use of surfactants**

   Adding surfactants, like the detergent dodecylmaltoside DDM and lauryl maltose neopentyl glycol (LMNG), can prevent the denaturation of soluble proteins at the air-water interface by providing a protective monolayer that acts as a barrier between the particles and the hydrophobic side of air-water interface [335, 336]. In addition, fluorinated detergents, such as fluorinated Fos-choline-8, have been added to grid samples with great effect [343–345]. This addition results in decreasing non-specific aggregation, improving particle distribution on cryoEM grids, as well as leading to better vitrification for cryoEM data collection.

3. **Use of grids with continuous support films**

   Hydrophilic graphene, graphene oxide and thin-carbon continuous support films lying across the holes, can be used to move the particles away from the air-water interface [160, 346]. Quite often the particles will interact with the support rather than the air-water interface, leading to reorientation. However, due to the presence of a support the signal-to-noise ratio is lower, which makes it imperative that the support is as thin as possible when collecting high

resolution datasets. Moreover, there is a risk that the sample will interact with the again leading to preferred orientation of the particles. However, if these views are different to those displayed in holey grids, then multiple datasets from the different grids can be merged to obtain complete angular reconstitution [347].

4. **Collecting tilted datasets**

   By tilting the grid and collecting tilted datasets of preferentially orientated samples may allow the multiple views needed for complete 3D reconstruction to be obtained [348]. However, tilting the grids leads to greater ice thickness and a defocus gradient is introduced across each micrograph, making data processing more challenging. In addition, since holey grids are used in this method collisions with the air-water interface could still lead to sample denaturation which may limit the resolution achieved.

5.2.6.2.1 Attempts to Solve the Preferred Orientation Problem As described previously, the presence of the PAM1.4Fab did not help prevent preferred orientation of the particles on the grid **(Figure 5.15)**. The Var2CSA constructs were also purified in complex with CSA to see whether this binding changed the overall conformation of the protein resulting in a more even distribution of orientations in the ice. However, a comparison of the 2D reconstructions, from grids with and without CSA, showed multiple classes with the main predominant view associated with preferred orientation **(Figure 5.16)**. The preferred orientation was confirmed by 3D reconstruction, which showed streaky "side" views as seen previously.

Neither did the addition of detergents to the buffer during grid preparation. Grids were set up with buffers with different concentrations of DDM in relation to its critical micelle concentration (cmc). Since Var2CSA ectodomain is a soluble protein, micelles do not need to form, just enough detergent has to be added to change the

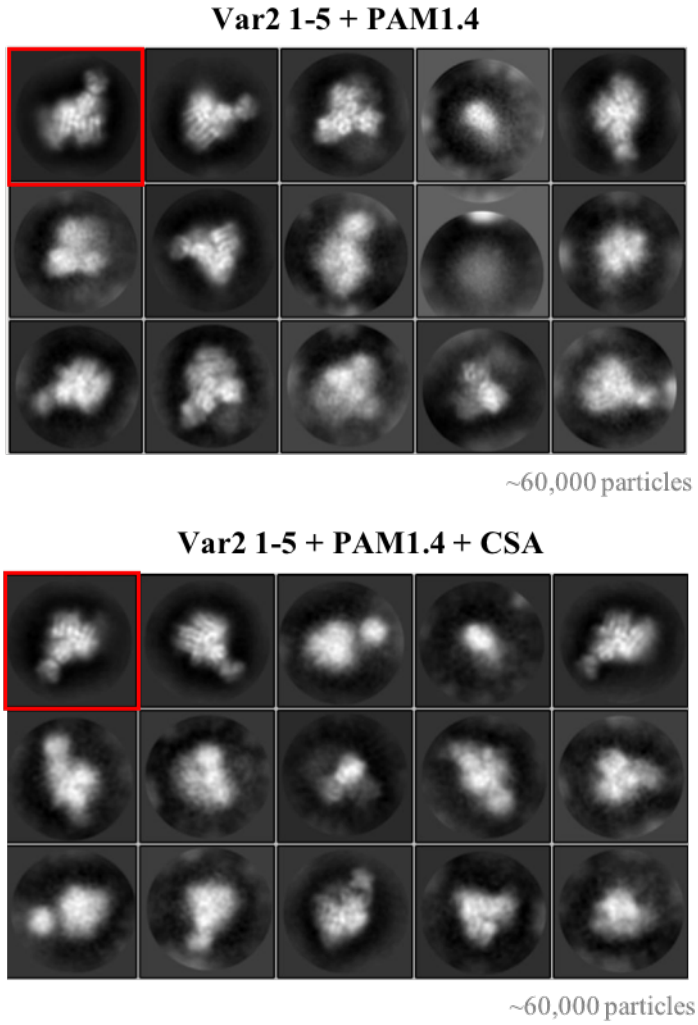


Figure 5.16: 2D Classification for NTS-DBL5 with PAM1.4Fab Plus and Minus CSA The 2D classes were sorted by number of particles contributing to each class with the classes with the most particles found in the top left corner and reading across the row. The red boxes highlight an example of the most represented, predominant view. Only the top fifteen classes are shown.

surface properties of the drop. Therefore, concentrations ranging from 0.125cmc to 1cmc were attempted. Unfortunately, the addition of the DDM detergent at any concentration led to very thin ice and a non-uniform ditribution of particles, with the appearance of aggregated rafts of particles and parts of the hole with no particles at all **(Figure 5.17)**.

Different detergents, due to distinctions in their hydrophobic tail region, have varying

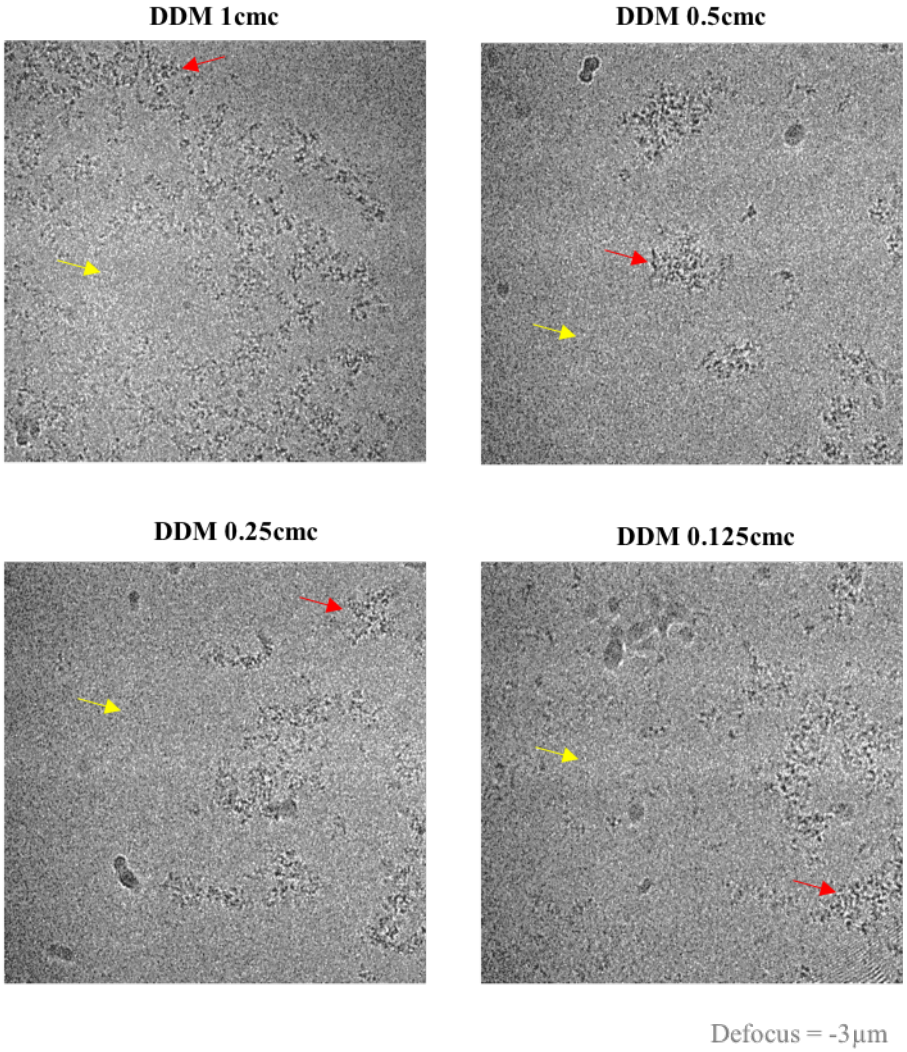


Figure 5.17: Representative Micrographs of Grids Prepared with DDM Detergent in the Sample Buffer Multiple different concentrations of DDM were added to the sample buffer in relation to its critical micelle concentration (cmc), and representative micrographs can be seen. Examples of rafts of aggregated particles are highlighted with a red arrow. Empty spaces with no particles are highlighted with a yellow arrow.

effects on the surface properties of a the drop of sample applied to the grid and so other detergents may not lead to the aggregation seen when using DDM in the buffer. As a result grids were set up with the following detergents added to the sample buffer at 0.5cmc: lauryl maltose neopentyl glycol (LMNG), lauryldimethylamine N-oxide (LDAO), Undecyl β-D-maltoside (UDM), and Fos-choline-8 (fluorinated). Addition of LDAO and UDM led to the aggregation of particles into rafts like

DDM, however on the addition of LDAO the rafts appeared to be smaller in overall size **(Figure 5.18)**. The LMNG and Fos-choline-8 (fluorinated) samples did not appear to lead to the formation of rafts of particles as much as the other detergents, however the thinning of the ice by the addition of the detergents led to particles of more "dotty" appearance which may be due to denaturation of the particles at the air-water interface **(Figure 5.18)**. No data could be collected from these grids due to the non-uniform distribution of particles across the holes with no discrete particles of good quality for picking. Further optimisation of these grids to improve the particle distribution is needed in the future. A wider range of concentrations of all the detergents tested could be tried, as well as other detergents with different properties.

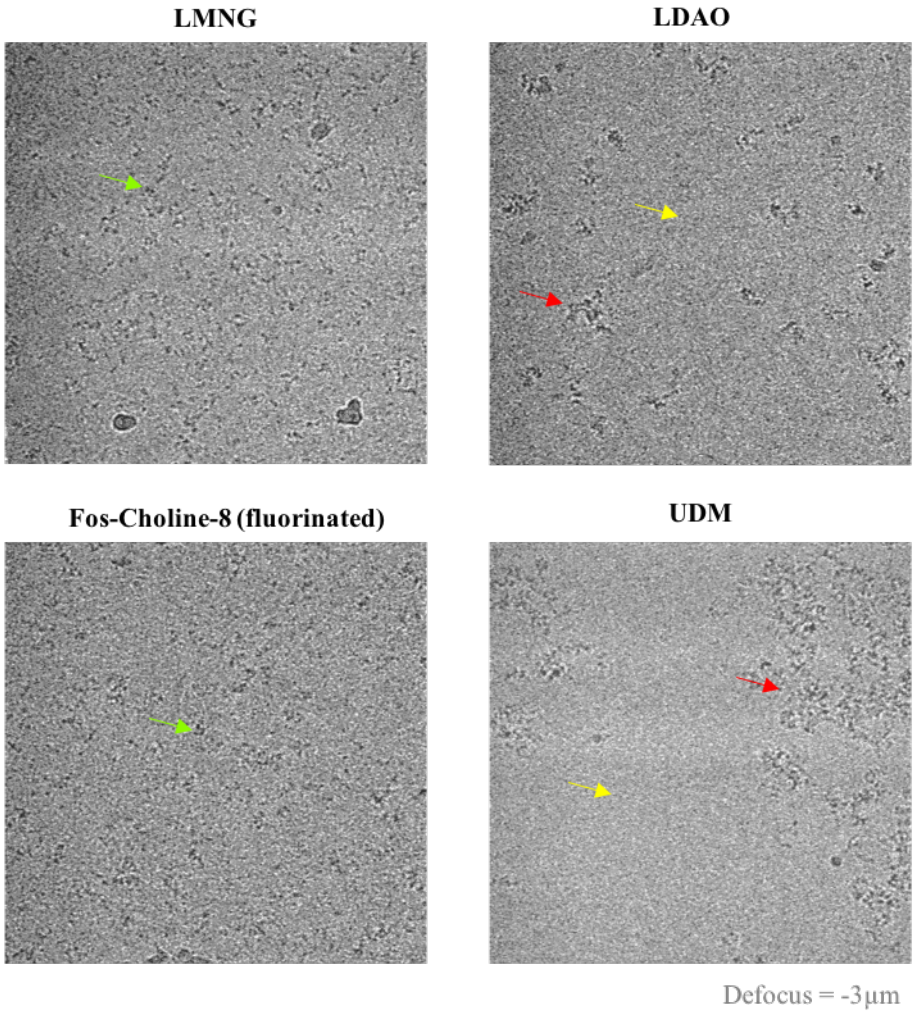


Figure 5.18: Representative Micrographs of Grids Prepared with Different Detergents in the Sample Buffer Examples of rafts of aggregated particles are highlighted with a red arrow. Empty spaces with no particles are highlighted with a yellow arrow. Examples of "dotty" particles thought to be denatured are shown by the green arrow.

Using grids with continuous supports was also attempted. Graphene oxide grids were produced **(Materials and Methods 2.8.1.3.1)**, whereas thin-layer carbon grids were manufactured (Quantifoil R 1.2/1.3 with Ultrathin Carbon). Both the graphene oxide grids and the thin layer carbon grids had a good distribution of particles but the relative contrast of the particles to the background was better for the graphene oxide grids and therefore a dataset was collected from these grids **(Figure 5.19)**.

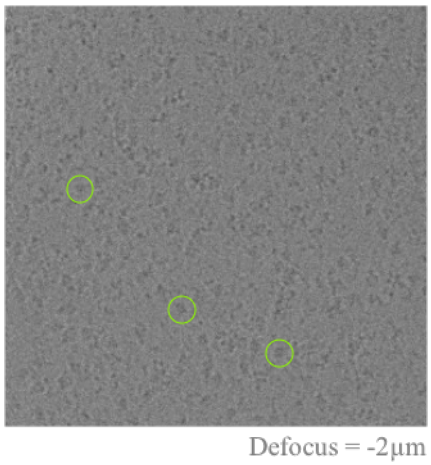


Figure 5.19: Representative Micrograph for NTS-DBL6 plus PAM1.4 on Graphene Oxide Grids Examples of individual particles that could be picked are shown by the green circles.

In the 2D classes from these graphene oxide grids, a few different views could be seen (highlighted in green boxes) in addition to those previously observed (shown in the red box) **(Figure 5.20)**.

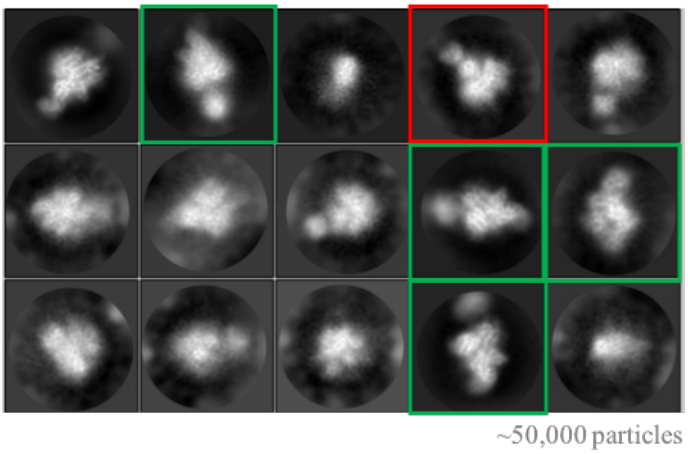


Figure 5.20: 2D Classification for NTS-DBL6 plus PAM1.4 on Graphene Oxide Grids The 2D classes are ordered according to the number of particles representing each class. The new views seen in the 2D reconstructions highlighted by green boxes, with the preferred orientation view highlighted by green boxes.

When making a 3D reconstruction, these different views did appear to reduce the stretched and streaky nature of the "side views" but the overall resolution and level of detail seen in the "front" view was visibly lower than the resolution seen from

the holey grids **(Figure 5.21)**. This lower resolution was probably a result of the overall signal to noise ratio being lower on these graphene oxide grids compared to the holey grids. In the future, the ice thickness must be optimised to be as thin as possible without causing the particles to aggregate, to increase the signal-to-noise ratio and overall resolution.

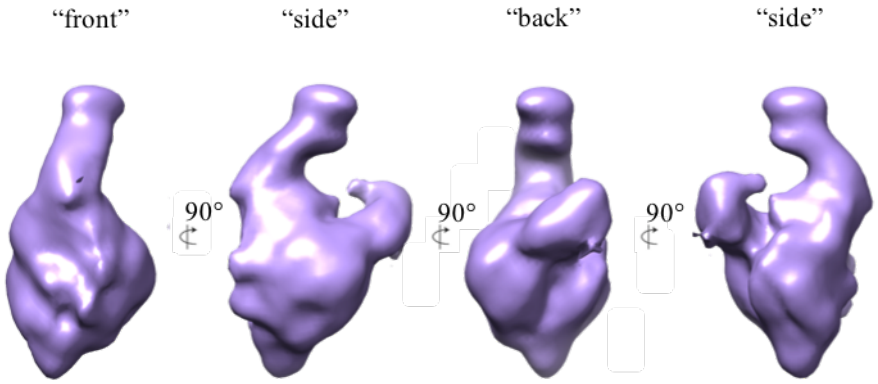


Figure 5.21: 3D Reconstructions for NTS-DBL6 plus PAM1.4 on Graphene Oxide Grids The 3D reconstruction seen from different views rotated 90°around the y-axis

5.2.6.2.2 Future Directions for Solving the Preferred Orientation Problem

In the future, grid preparation and freezing of the graphene oxide grids (and/or the thin layer carbon grids) must be optimised in order to improve the resolution achieved. This optimisation could include the addition of a detergent to the buffer to promote thinner ice and/or additional binding partners. Additionally, collecting a tilted dataset on grids with a gold support (to reduce beam-induced motion) could be attempted to overcome the problem of preferred orientation [348].

5.2.6.3 Conclusions So Far from cryoEM Experiments

1. Comparison of the NTS-DBL5 and NTS-DBL6 particles and reconstructions showed that they both had a very similar overall shape with a large "head" domain and smaller "tail" domain. Therefore, the "tail" observed in the

negative stain EM of NTS-DBL6 is not DBL6.

2. Both the NTS-DBL5 and NTS-DBL6 samples exhibited preferred orientation on holey grids, without a uniform distribution of angular projection views of the individual particles, and so the overall 3D shape could not be determined.

3. Preparing the sample on a continuous support film, such as on graphene oxide grids, appeared to reduce preferred orientation. However, the overall resolution also decreased.

5.2.7 Ongoing and Future Directions

5.2.7.1 Limited Proteolysis to Find Smaller Stable Fragments of Var2CSA

In order to identify smaller and more stable fragments of Var2CSA, limited proteolysis on the NTS-DBL5 and the NTS-DBL6 domains was performed. The proteases preferentially have access to the cleavage sites that are exposed at the surface of the overall fold, and so limited proteolysis cuts in these disordered regions and loops, which might be hindering crystallisation [349]. A time course of proteolytic digestion was carried out using: trypsin, chymotrypsin, carboxypeptidase B, endoproteinase GluC and endoproteinase AspN, firstly on the NTS-DBL5 construct **(Figure 5.22)**.

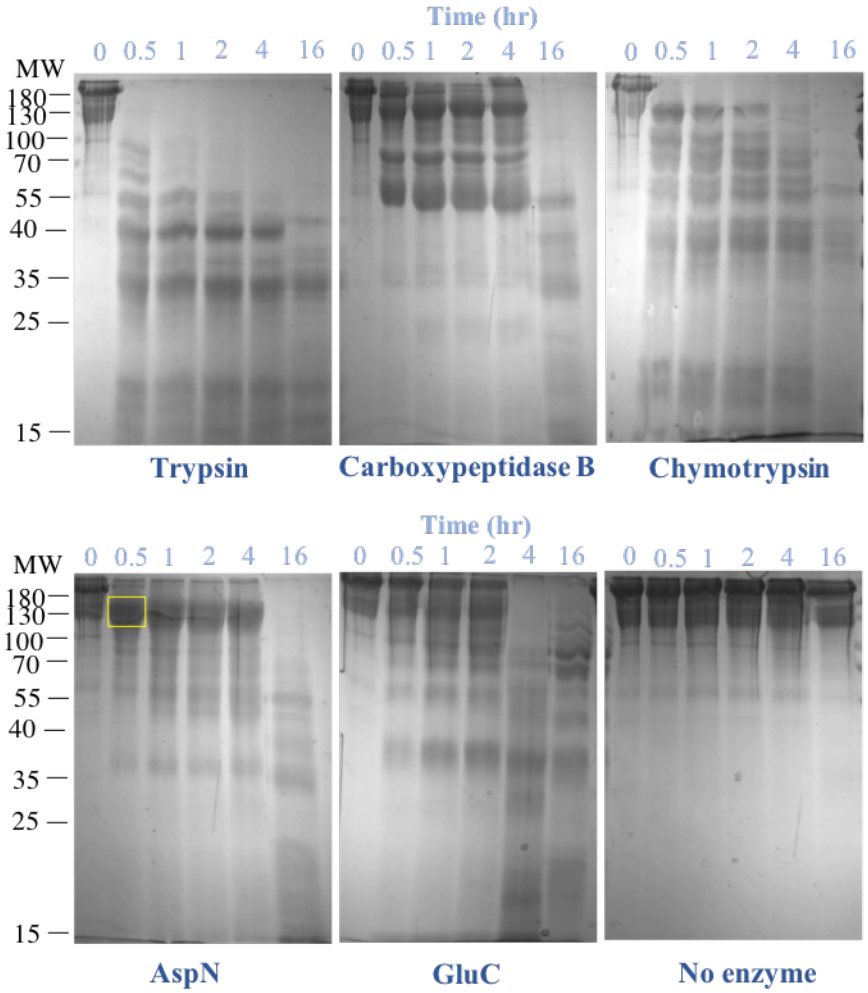


Figure 5.22: Limited Proteolysis of NTS-DBL5 Reduced SDS-PAGE gels of the limited proteolysis time course for NTS-DBL5 with five different enzymes (trypsin, CB=carboxypeptidase B, chymotrypsin, AspN, GluC) and no enzyme as a negative control. The band highlighted by the yellow box represents a stable fragment that we aimed to purify.

Only digestion with AspN produced a stable fragment from NTS-DBL5, highlighted by a yellow box **(Figure 5.22)**. It was hypothesised that this was the same fragment that was also seen in the GluC digests and the no enzyme negative control. However, the AspN digestion also resulted in the presence of the multiple smaller degradation products as well as the stable fragment, therefore the digestion would need to be optimised before structural studies could be carried out. For further analysis, the band was excised and sent for in-gel trypsin mass spectrometry analysis [350]. An excised band for the full NTS-DBL5 construct was also analysed for direct comparison

(Figure 5.23).

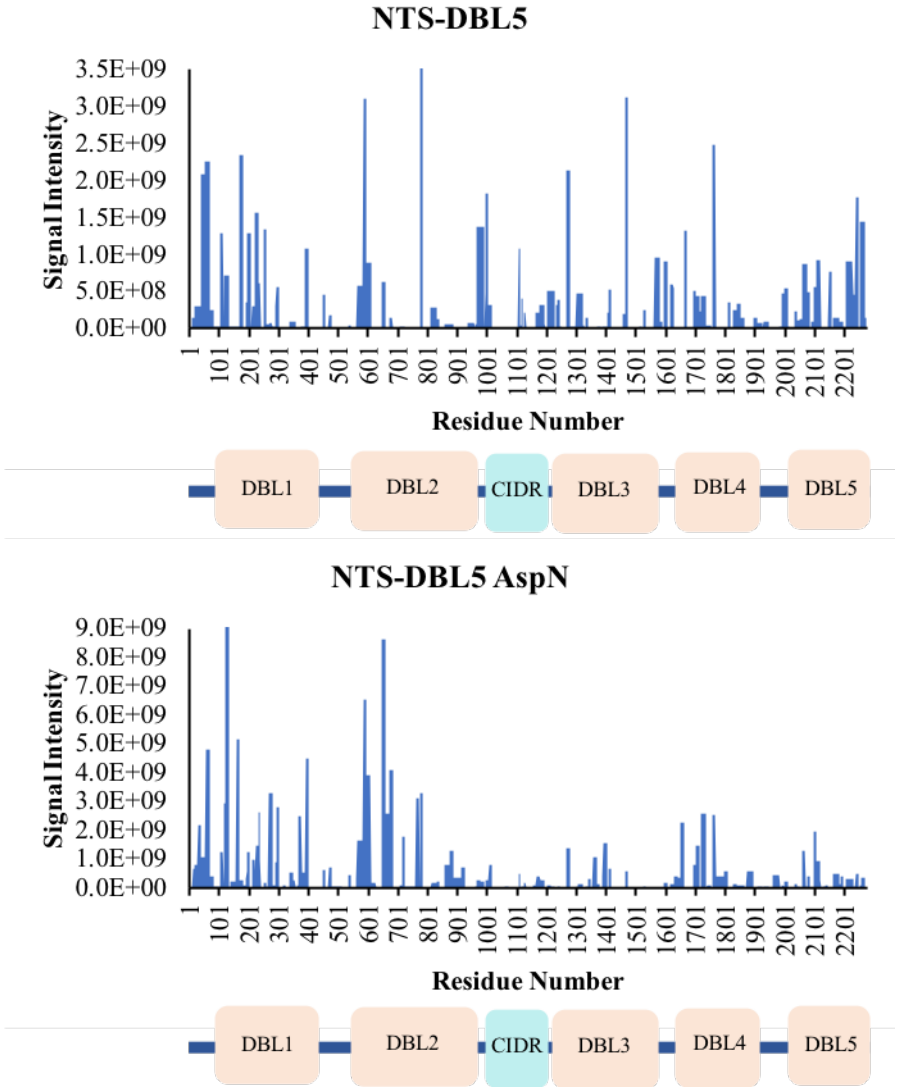


Figure 5.23: In-gel Trypsin Digest Mass Spectrometry of NTS-DBL5 and AspN Fragment A plot of the residue number of the fragments detected and the size of the signal each fragment produced for the full length NTS-DBL5 (top), and the band after AspN digest (bottom). Underneath the plots is a schematic of the domain architecture showing how the residue number relates to the domain boundaries in the NTS-DBL5 construct.

Whereas the NTS-DBL5 construct had fragments of relatively high signal intensity across the whole construct, the AspN-treated protein had fragments of higher signal intensity at the N-terminus in comparison to the rest of the construct, potentially representing an NTS-DBL2 fragment. Since there were still some fragments detected

in the remaining parts of the ApsN treated construct, the AspN-digested fragment was purified by SEC and sent for analysis by native mass spectrometry **(Figure 5.24)** [351].

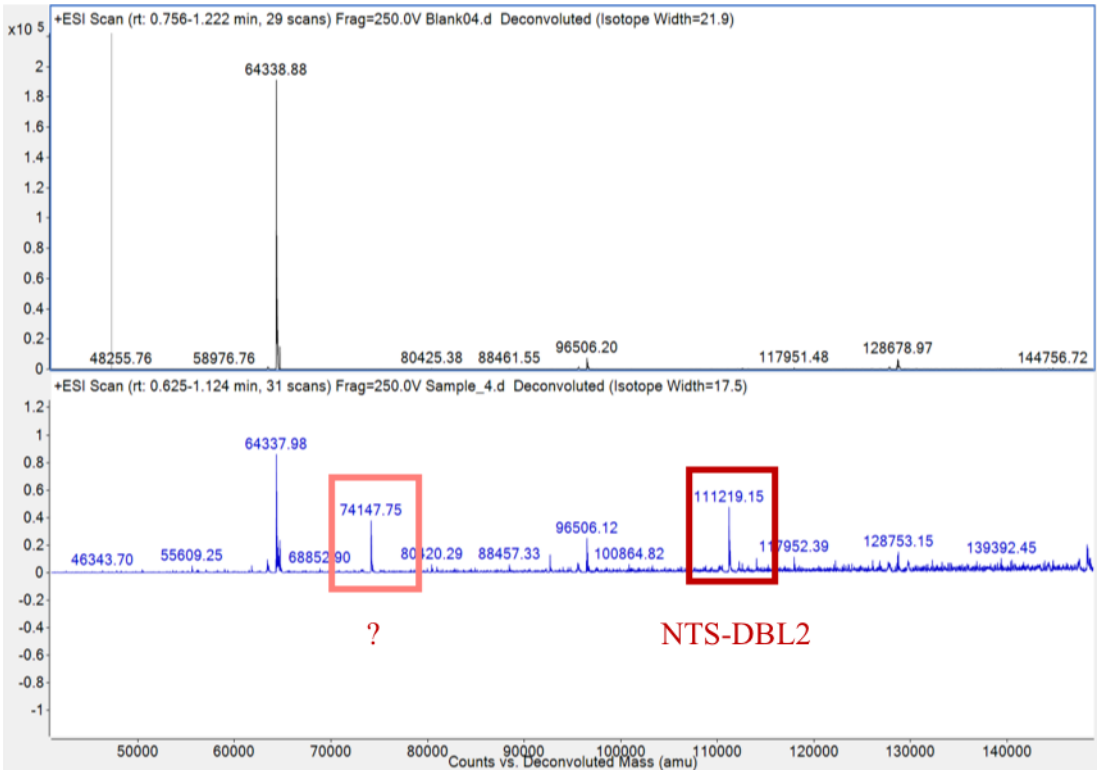


Figure 5.24: Native Mass Spectrometry of NTS-DBL5 Digested with AspN Spectra from native mass spectrometry experiment with the blank buffer sample in the top panel and the NTS-DBL5 AspN digested sample in the bottom panel. The main peaks present in the NTS-DBL5 AspN panel but not in the blank spectrum are highlighted by the red boxes. The dark red box is putatively NTS-DBL2, whereas the species represented by the light red box is unknown.

In the native mass spectrometry spectrum there were only two main peaks that were not seen in the blank control (marked by boxes in **Figure 5.24**). The peak at molecular weight of 111.2kDa (highlighted by the red box) had a very similar mass to that predicted by the NTS-DBL2 sequence (110.8kDa). Taking this with the in-gel trypsin digest mass spectrometry, it was suggested that the AspN digestion product was NTS-DBL2. It is unclear what species the other peak at 74.1 kDa represented. Unfortunately, data was not collected for species that were greater than 145kDa.

To confirm that the AspN digestion of NTS-DBL5 resulted in the production of an N-terminal fragment, potentially NTS-DBL2, with the C-terminal section completely digested, a non-reduced SDS-PAGE gel was run. The presence of one band in an identical place in both the reduced and non-reduced gels showed that rather than small flexible loops being digested away from the whole protein, there was a large digestion and only one species thought to be NTS-DBL2 remained **(Figure 5.25)**.

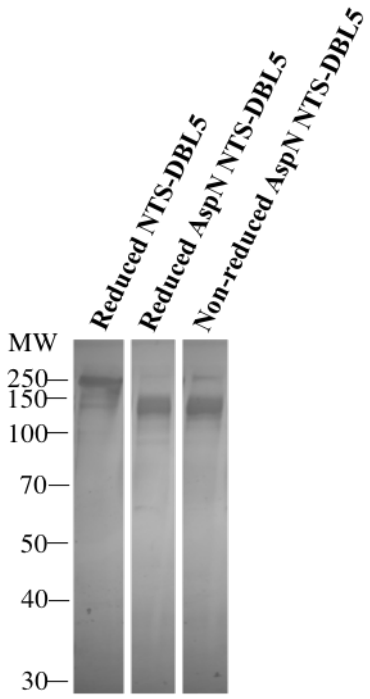


Figure 5.25: AspN Digestion of NTS-DBL5 Reduced and non-reduced samples of the same AspN digestion of NTS-DBL5 run an SDS-PAGE gel. A full-length uncleaved NTS-DBL5 sample is also shown as a reference to where full-length protein runs on the SDS-PAGE gel.

The digestion of NTS-DBL6 with AspN at 37°C resulted in the production of two stable fragments **(Figure 5.26 A)**. The lower band of the NTS-DBL6 AspN digest ran on the SDS-PAGE gel in a similar place to the NTS-DBL5 AspN digestion product. In order to check whether NTS-DBL6 was truly digested into two products or whether it was just nicked with the two halves remaining as a single complex, the post AspN sample was run on a non-reduced gel **(Figure 5.26 B)**. This resulted

in the appearance of just one band that ran at the same position as the full-length NTS-DBL6 construct, suggesting that AspN cleaved a loop/loops without disrupting the overall fold.

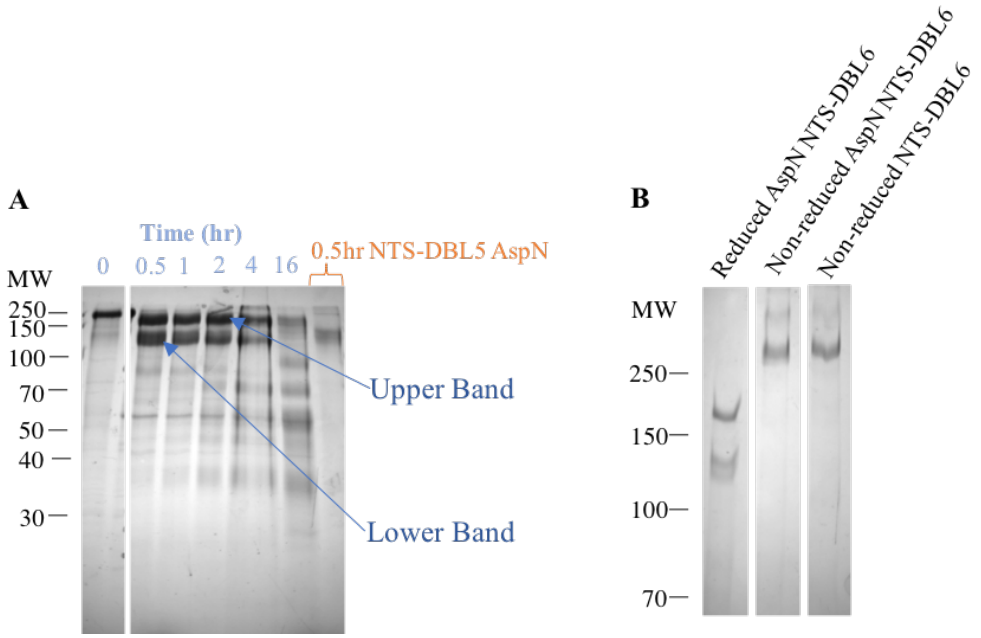


Figure 5.26: AspN Digestion of NTS-DBL6 A) A reduced SDS-PAGE gel of the time course of NTS-DBL6 digestion by AspN, with the completed digestion of NTS-DBL5 by AspN also run as a reference. B) A reduced SDS-PAGE sample of NTS-DBL6 treated with AspN compared to a non-reduced sample of the same cleavage. The non-reduced sample of NTS-DBL6 was used as a reference to show where the full length protein ran on the gel.

In-gel trypsin mass spectrometry data showed that the lower band contained the NTS-DBL2 fragment (as expected from the band on the SDS-PAGE gel running in the same place as the fragment generated from NTS-DBL5 AspN cleavage) **(Figure 5.27)**. The upper band from the digest contained a digestion product that spanned DBL3-DBL6 **(Figure 5.27)**.

There were no fragments detected in the CIDR region in either the upper band or lower band samples, whereas there were fragments detected in the full length NTS-DBL6 sample **(Figure 5.27)**. This suggested that the $CIDR_{PAM}$ domain was removed either partially or completely by AspN cleavage, highlighting that there is a

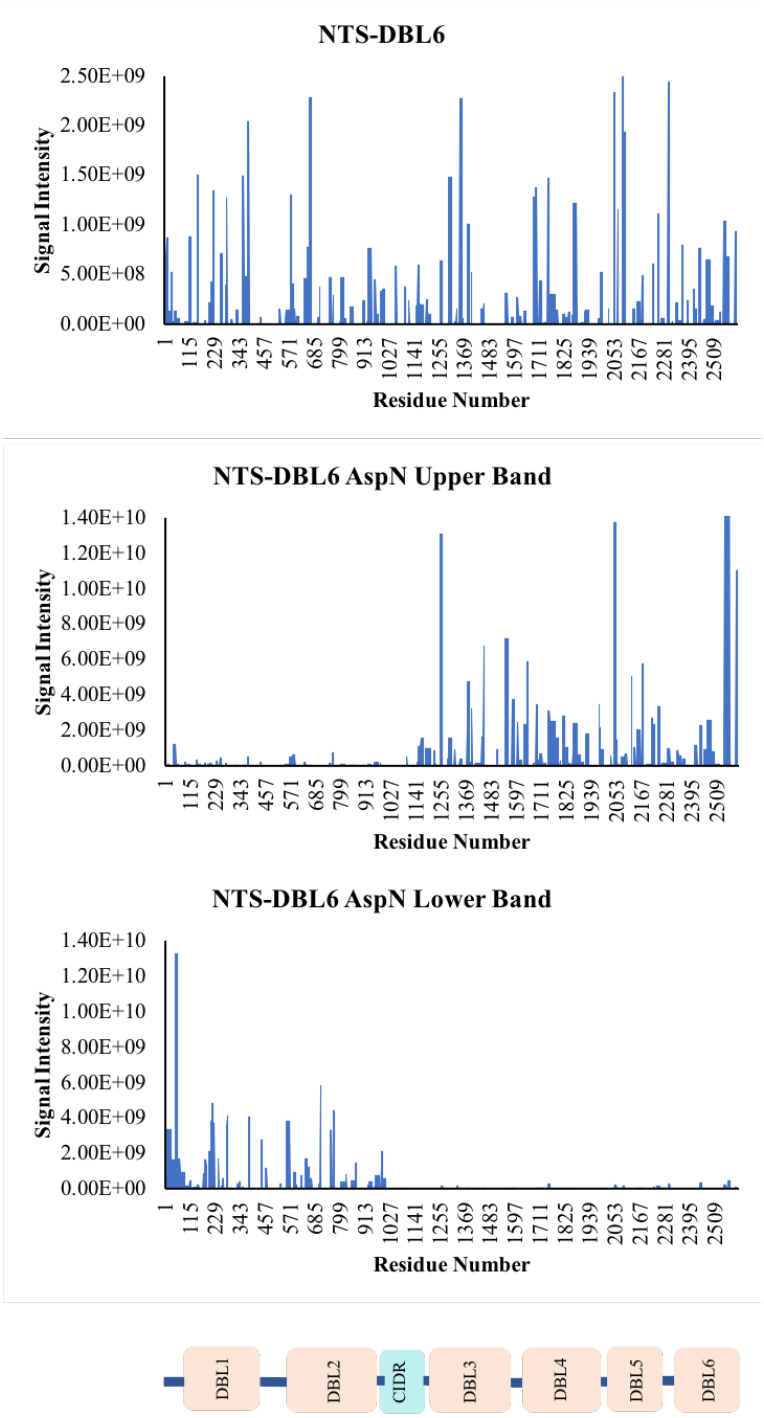


Figure 5.27: In-gel Trypsin Digest Mass Spectrometry of NTS-DBL6 and AspN Fragments A plot of the residue number of the fragments detected and the size of the signal each fragment produced for the full length NTS-DBL6 (top), the upper band after AspN digest (middle) and the lower band after AspN digest (bottom). Underneath the plots is a schematic of the domain architecture showing how the residue number relates to the domain boundaries in the NTS-DBL6 construct.

flexible region between DBL2 and DBL3 which could be cleaved or nicked by AspN creating two fragments, NTS-DBL2/CIDR and CIDR/DBL3-DBL6. These remained together as a complex in non-reducing conditions, suggesting that the quaternary fold of the NTS-DBL6 involved an inter-domain disulphide bond between these two fragments. The digestion of NTS-DBL5 with AspN resulted only in the production of one stable digestion product, thought to be NTS-DBL2. The presence of only one digestion product compared to two for NTS-DBL6 suggested that the DBL6 masked additional AspN cleavage sites that then became available in the NTS-DBL5 construct. The stabilising affect of DBL6 could also be due to its involvement in an interdomain disulphide bond. This hypothesis was supported by the presence of a free cysteine, which in the solved DBL6 structure was solvent exposed and therefore could be available to bond with another free cysteine on another domain **(Figure 5.28)**.

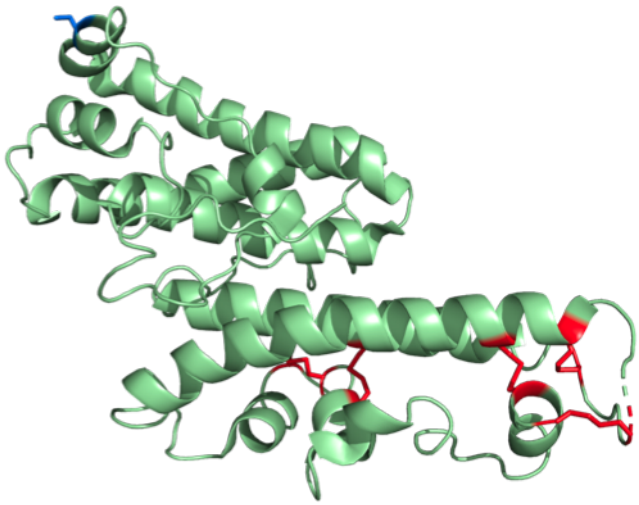


Figure 5.28: Free Cysteine in DBL6ϵ Domain The solved structure of DBL6 [297] (green) with all cysteines shown as side chains. The paired cysteines forming intra-domain disulphide bonds are shown in red, with the one unpaired, solvent exposed cysteine shown in blue).

Initial broad crystallisation screens were set up using purified AspN-digested NTS-DBL5 **(Materials and Methods 2.7.2.1)**, however no crystals were produced. In

the future, when more NTS-DBL6 protein is obtained, broad crystallisation screens of AspN-digested NTS-DBL6 will be conducted and this sample will be used for cryoEM studies. The removal of a flexible region might help with 2D classification of the particles, as well as altering the surface properties which may prevent preferred orientation.

5.2.7.2 New Var2CSA Constructs Design and Expression Trials

Since no crystallography nor cryoEM had yet been successful, it was decided to design some new shorter constructs that may crystallise more easily **(Figure 5.29 A)**, alone and in combination with antibody Fab fragments or nanobodies. The domain boundaries for these constructs were determined by using secondary structure prediction results from the Phyre2 server [352].

The results from the limited proteolysis experiments, described in **Section 5.2.7.1**, were used to decide which constructs should be made. The digestion of NTS-DBL6 with AspN indicated that a flexible region between DBL2 and DBL3 was removed without disrupting the rest of the structure. Since the cleavage sites could not be determined accurately, the following constructs were designed: NTS-DBL2, NTS-CIDR$_{PAM}$, and NTS-DBL3. The other fragment produced in the digestion of NTS-DBL6 (thought to be stabilised by the presence of DBL6) was represented by the constructs CIDR$_{PAM}$-DBL6 and DBL3-DBL6. In addition to these constructs, since a different expression system was being used, an S2 NTS-DBL6 construct was designed in the hope that this construct would express better than the lepidoptoeran/baculovirus counterpart so future experiments could be carried out more quickly. In addition, since multiple constructs were already being cloned and transfected, an NTS-DBL4 construct was also designed as this had not yet been tried and the CSA-binding site is thought to be in the N-terminal region **(Figure 5.29 A)**.

Furthermore, limited proteolysis of NTS-DBL6 with AspN suggested the possibility that $CIDR_{PAM}$ is connected to the rest of the protein by flexible linkers. Therefore, as well as the shorter constructs, NTS-DBL6 constructs in which the $CIDR_{PAM}$ domain had been deleted and replaced with linkers were also designed **(Figure 5.29 B)**. Since the distance between the C-terminus of DBL2 and the N-terminus if DBL3 was unknown linkers of different lengths were designed. The distance between the N- and C-termini of the solved structures of CIDR domains was used as a reference and measured as 51.2Å, which equates to 15 residues in an extended chain. However, the residues at the termini in the solved structure were not seen in the electron density and therefore this was likely to be an underestimate. To compensate for these long flexible linkers, composed mainly of glycine and serine residues, additional aspartate residues were added so that this flexible region could be removed by AspN cleavage for future structural studies **(Figure 5.29 B)**.

These constructs were cloned from the NTS-DBL6 construct in the pAcGP67A vector, into the S2 cell vector pExpresS2.1, using the primers shown in **Appendix A.5**. It was decided to move from the lepidopteran/baculovirus system to the insect cell S2 system, as this would be less time consuming to produce multiple new stable cell lines. Small expression trials were carried out for the successfully transfected S2 cell lines, and expression was characterised by western blot **(Figure 5.30)**. The NTS-DBL2, NTS-DBL4, NTS-DBL6, and ΔCIDR-40aalinker, all expressed in a secreted form. However, NTS-DBL3, NTS-CIDR, DBL3-DBL6, and CIDR-DBL6 constructs did not express. Transfections for these constructs will be repeated. For all the constructs that did express, the S2 stable cell lines will be expanded and large-scale expression and purification will take place for crystallisation.

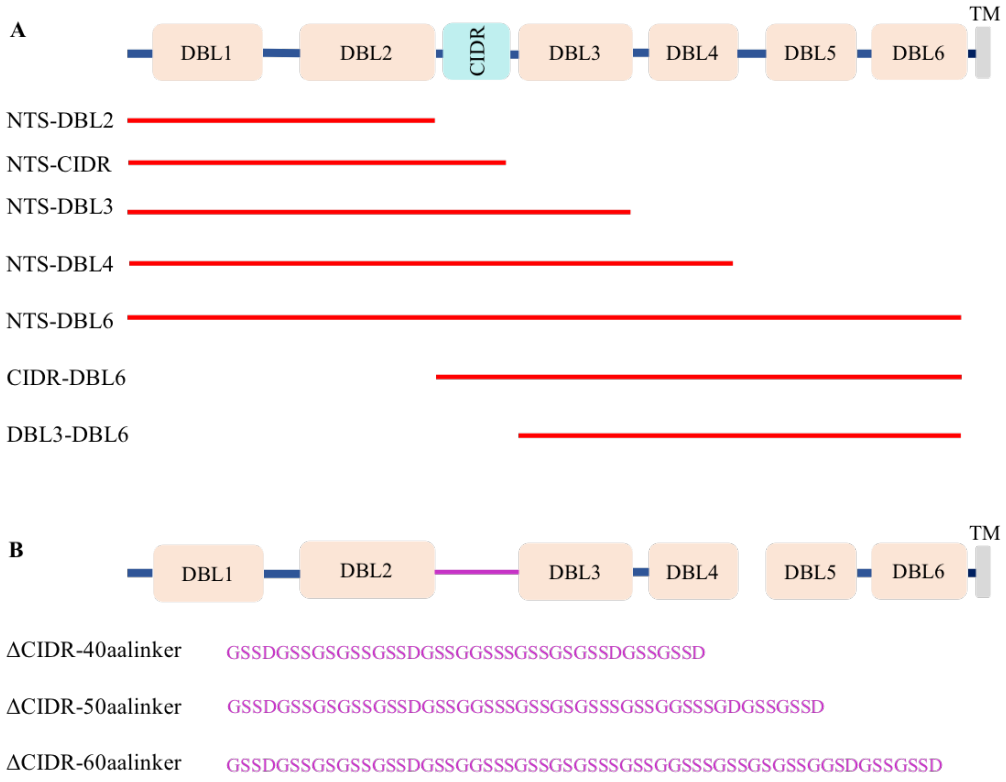


Figure 5.29: Boundaries of New Constructs A) An overview of the shorter constructs designed with the domain boundaries shown in relation to the overall sequence by the red lines. B) The ΔCIDR constructs with the CIDR domain replaced with the linker sequences shown in purple.

5.2.7.3 Nanobodies

Collaborators from the university of Brussels (Yann Sterckx) had previously given us sixteen nanobodies, which had been raised against and bound to the DBL1-ID2a region of Var2CSA, which was predicted to contain the CSA binding site **(Figure 5.31)**.

In addition, at the beginning of the project when it was thought that the DBL6 was the flexible tail region of the protein and that the NTS-DBL5 construct would be more globular and more likely to crystallise, the NTS-DBL5 construct was also used to immunise llamas to generate nanobodies **(Figure 5.31)**. It was hoped that the nanobodies produced might stabilise the quaternary structure of the NTS-DBL5 by binding in crevices between the domains. Biotinylated NTS-DBL5 was

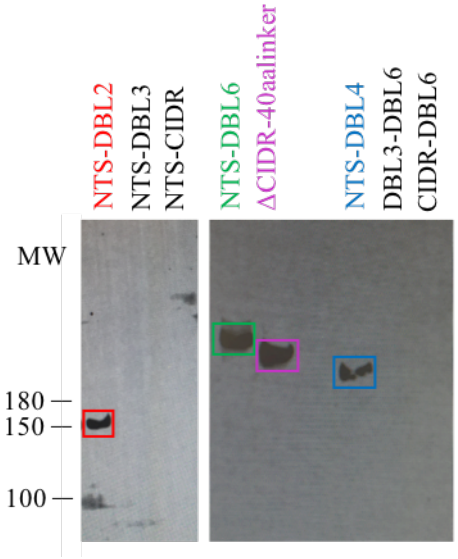


Figure 5.30: Expression Trials for New Var2CSA Constructs Western blot of the S2 cell supernatant for the stable transfections of multiple different constructs. The western blot was probed with an anti-His conjugated antibody, which targeted the His-tag of these constructs, showing up as a band on the blot.

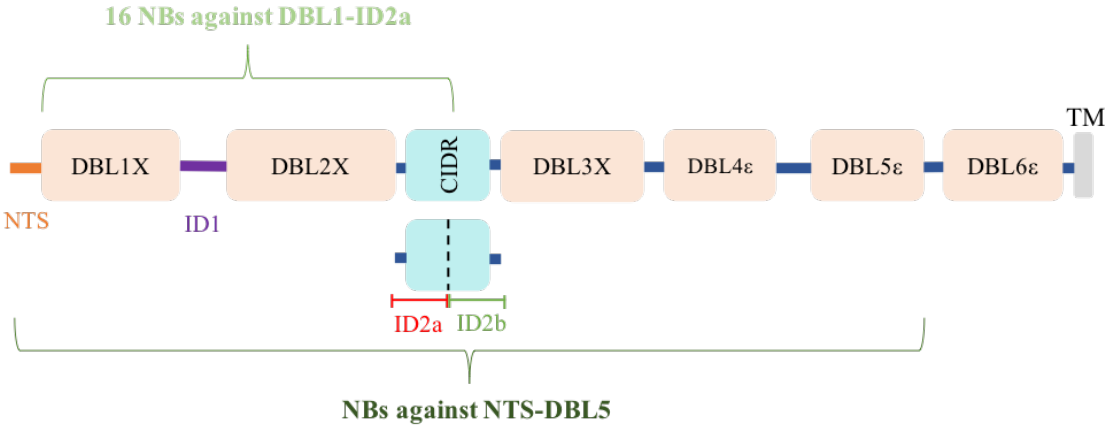


Figure 5.31: Nanobodies against Var2CSA

produced, for screening and this identified nanobodies that interacted with NTS-DBL5 **(Materials and Methods 2.4.8.2)**. There were 30 successful nanobodies identified, which in the future could be screened for use with the original DBL1-ID2a nanobodies for crystallography. Moreover, multiple nanobodies could be bound to the Var2CSA constructs simultaneously, which might help prevent preferred orientation in cryoEM.

To characterise which nanobody might have the most stabilising effect on the overall fold, potentially by simultaneously contacting different domains, the individual domains of Var2CSA were expressed and purified. The rational was that a "stabilising" nanobody would bind to the NTS-DBL5/6 construct but would not bind or would bind more weakly to the individual domains since its epitope comprised of multiple domains. In preparation for this, expression trials in *E.coli* SHuffle cells were carried out on the DBL domains and three of them expressed to a high enough level for purification **(Figure 5.32)**. The remaining DBL domains and the expression of the CIDR domain will need to be optimised in the future, before this characterisation of the nanobodies can take place.

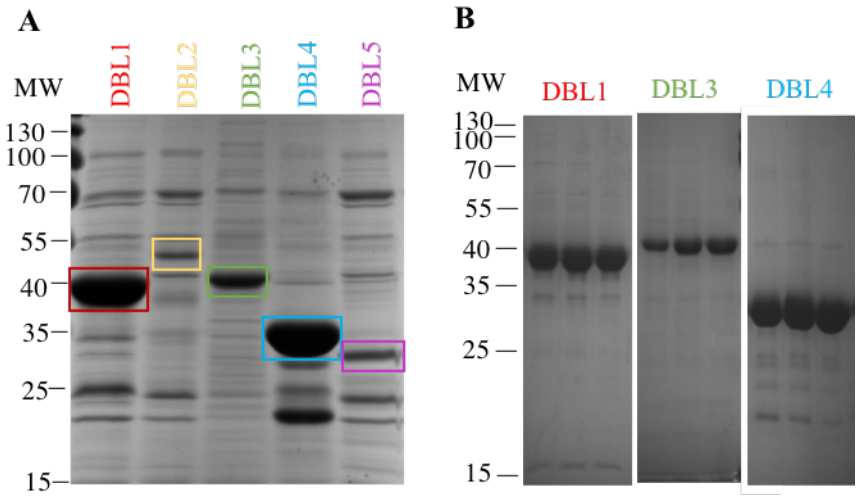


Figure 5.32: Expression and Purification of Individual DBL Domains A) SDS-PAGE gel of bacterial expression of DBL domains, with the band representing the DBL expression highlighted. B) For those that expressed highly, the SDS-PAGE of the SEC fractions of the purified DBL domains.

5.3 Discussion

The full length structure of Var2CSA, needed for specific binding to CSA, will greatly aid the design of a vaccine that targets this interaction, which is crucial

for placental cytoadhesion in PAM [199, 295, 298]. During this project, steps have been taken forward towards determining this structure. Clues from low resolution cryoEM 3D reconstructions give us an insight into the overall shape of Var2CSA ectodomain. Consistent with the previous negative stain reconstruction, the overall shape consists of a large globular part ("head") and a much smaller globular part ("tail") connected together. The NTS-DBL5 reconstruction had the same overall shape as the NTS-DBL6 construct, and therefore this "tail" region is not solely the DBL6 domain as previously hypothesised.

Limited proteolysis data showed that upon digestion of NTS-DBL6 with AspN, two stable fragments of around 130kDa and 180kDa were produced. Mass spectrometry suggested these to comprise of regions NTS-DBL2 and DBL3-DBL6, respectively, with the precise cleavage sites unknown. Interestingly, the limited proteolysis of NTS-DBL5 with AspN only resulted in the formation of one stable fragment, identified as NTS-DBL2. This suggested that DBL6 protected an otherwise exposed AspN cleavage site, and thereby stabilised the C-terminal fragment of this cleavage. One clue as to how the DBL6 may have stabilised the CIDR/DBL3-DBL6 fragment was the presence of a free cysteine on the surface of DBL6, which could form a stabilising interdomain disulphide bond. Based on these limited proteolysis experiments, smaller multidomain constructs have also been designed and cloned. Future structural studies of these multidomain fragments, as well as the full-length protein treated with AspN may prove more successful (both by X-ray crystallography and cryoEM) than those of the whole uncleaved full-length construct.

There is also still much work to be done to obtain the structure of full-length Var2CSA by cryoEM. Further optimisation of the graphene-oxide grids, as well as potentially obtaining tilted datasets, may overcome preferred orientation. Moreover, the additional antibodies and nanobodies can be added to produce a larger complex reorientating particles by changing the overall shape and exposed surface properties of

Var2CSA. Even if a high resolution structure cannot be obtained, a lower resolution structure could still be very helpful, especially when combined with X-ray crystal structures of multiple domain fragments. In addition, given the number of nanobodies obtained, gold-labelled nanobodies can provide information about the positions of different domains within even low resolution density [353]. Furthermore, gold-labelled CSA particles could be used to show the position of the binding site. There are still many things to attempt in this project over the coming years in order to gain further understanding of the full-length structure of Var2CSA and its binding to CSA.

Chapter 6

Concluding Remarks

6.1 Summary of Main Findings from this Thesis

The main aim of this thesis was to gain insights into how knowledge of the structures of potential malarial vaccine candidates, especially in complex with a binding partner or an inhibitory antibody, can aid future immunogen design.

6.1.1 Characterisation of Immunogens Against EPCR-binding CIDRα1 Domains

In this project, the vaccine targets were *Pf*EMP1s that bind to EPCR, leading to cytoadherence and sequestration of iRBCs in the brain, a pathophysiology crucial in the development of severe malaria. Previously, the structure of the EPCR-binding domain of a *Pf*EMP1, CIDRα1.4, in complex with EPCR was used to characterise the structural features of the binding site, namely a helix-kink-helix motif. Sequence analysis revealed that the chemistry of the binding site region was conserved in different variants of *Pf*EMP1. Therefore, the helix-kink-helix motif was then grafted onto a small three-helical scaffold to produce an immunogen, Bundle1.4.

Subsequently other immunogens were produced by introducing mutations into the

binding site motif of Bundle1.4, to recapitulate the six different subclasses of the EPCR-binding CIDRα1s, representing sequence diversity. The six designed immunogen bundles were cloned, expressed in *E. coli* inclusion bodies, refolded and purified. Assessment of their fold by circular dichroism showed that they all contained predominantly α-helical structure. Their ability to bind to EPCR was tested by surface plasmon resonance, which demonstrated that four out of the six bundle immunogens bound to EPCR. It was speculated that for the CIDRα1 subclasses represented by the two bundle immunogens that did not bind to EPCR, regions outside of the helix-kink-helix motif may be involved in binding. Also, the assumption that the same EPCR-binding mode, based on the conserved chemistry of the helix-kink-helix, for the different CIDRα1 subclasses was questioned. Structures, solved by X-ray crystallography, of the four binding bundles in complex with EPCR, showed that the EPCR-binding site had been successfully introduced into the immunogens, and that the mechanism of binding was the same to that of the CIDRα1:EPCR interaction.

The Bundle1.4 was then used to pull-down antibodies from sera of rats immunised with CIDRα1.4. This demonstrated that although Bundle1.4 bound to some inhibitory antibodies, the vast majority of inhibitory antibodies did not bind. Therefore, it was concluded that most inhibitory epitopes either lie completely outside of the functional EPCR-binding site or contain other regions that are not in the helix-kink-helix motif. Similarly, as an immunogen the Bundle1.4 generated antibodies that could recognise the cognate CIDRα1.4 domain and inhibit EPCR binding, however, the antibodies produced were less effective and abundant than the equivalent antibodies raised through immunisation with the CIDRα1.4 domain. In addition, immunisations with the Bundle1.4, both when coupled to a VLP and when uncoupled, did not increase the cross-reactivity of the pool of antibodies to other CIDRα1 subclasses compared to immunisations with the CIDRα1.4. Indeed, even when a mixture of bundles was

used for immunisations, only antibodies that recognised members of the CIDRα1.4 subclass were produced. This suggested that the EPCR-binding site grafted onto the bundle was more representative of the CIDRα1.4 antigen than the other subclasses. This makes sense, since the original Bundle1.4 was designed from the structure of CIDRα1.4:EPCR complex. This finding, together with the fact that two of the other Bundles did not bind to EPCR, suggested that the mode of EPCR binding by the different CIDRα1 subclasses may not be the same, as previously thought. In the future in order to understand this, the structures of different CIDRα1 domains bound to EPCR must be elucidated.

In addition, these studies caution against an approach in which a binding surface is grafted onto a smaller scaffold for immunisation. Although targeting a functional site directly may seem intuitive, these regions may not be the most immunogenic and therefore few inhibitory antibodies are raised. In contrast, knowledge of the structure of the epitope of an effective antibody that is known to be elicited through human immunisation, is a valid starting point for such a study. This allows certainty that a complete epitope is recapitulated and also that it is possible to elicit such an antibody from the human germline. Therefore, future vaccine design efforts targeting an EPCR-binding *Pf*EMP1 protein require the molecular characterisation of human monoclonal antibodies, which can be used to generate an epitope-focused vaccine.

6.1.2 The Molecular Characterisation of the Epitope of an Inhibitory Human mAb Against *P. vivax* Invasion and Subsequent Epitope-Focused Vaccine Design

In the second project, the potential vaccine target was *Pv*DBP, which binds to the DARC receptor and is essential for *P. vivax* invasion into reticulocytes. A strain-

transcending human monoclonal antibody, elicited by immunisation with *Pv*DBP-RII, that blocked the invasion of *P. vivax* into erythrocytes, was isolated. The structure of the Fab fragment of this antibody, DB9, was solved in complex with the parasite *Pv*DBP-RII protein, by X-ray crystallography. This revealed that the epitope was solely on subdomain 3 of the *Pv*DBP-RII, distal to its DARC receptor binding site and dimerisation interface on subdomain 2.

Immunogens, based on the subdomain 3 were designed, cloned, expressed and purified. Their fold was assessed by circular dichroism, and binding studies by surface plasmon resonance showed that the DB9-epitope had been successfully incorporated. This was further clarified by obtaining the crystal structure of the subdomain 3 immunogen in complex with DB9 Fab, which highlighted that the binding mode to DB9 was the same as the *Pv*DBP-RII domain. Upon immunisation of rabbits with the subdomain 3 immunogens, encouragingly, the sera contained antibodies that did block the binding of *Pv*DBP-RII to DARC. However, the quantity and inhibitory-quality of the antibodies produced, compared to those generated upon immunisation with *Pv*DBP-RII, could not be elucidated. These results provide further promise for epitope-focused vaccine design methods. In the future, further immunisation experiments with the subdomain 3 immunogens is needed to assess the quality and quantity of the antibodies produced in comparison to the *Pv*DBP-RII immunisations.

6.1.3 Towards the Structure of Var2CSA

In the final project, the vaccine target was an unusual *Pf*EMP1, called Var2CSA, which is involved in the cytoadherence of infected red blood cells to the placenta. Var2CSA mediates this sequestration by binding to the carbohydrate CSA. Unlike the previous projects in which single domains of the parasite protein bind to receptors, the binding to CSA requires multiple domains and potentially the overall quaternary

fold of ectodomain for the interaction. Therefore, the main aim of this project was to gain insight into the structure of the Var2CSA ectodomain by both X-ray crystallography and cryoEM. Low resolution cryoEM structures highlighted that the ectodomain forms a compact shape with one large globular region and a smaller flexibly-linked region. The graphene-oxide grids provided promise of overcoming the preferred orientation problem seen with the particles on holey grids. Further optimisation of the sample and the ice thickness of these grids may enable datasets of increased resolution to be obtained. Moreover, limited proteolysis experiments identified a flexible region between DBL2 and DBL3 domains that could be removed to promote crystallisation. In addition, new shorter multidomain constructs have been designed and cloned, which may provide easier targets to crystallise in the future. Insights into the full-length structure of the Var2CSA ectodomain will not only enable us to gain further understanding of how CSA binds, but also enable the epitopes of inhibitory antibodies to be elucidated for future vaccine design against this target.

6.2 Towards a Vaccine Against Malaria

Malaria remains a major global killer, leading to nearly 500,000 deaths each year, primarily of African children under the age of five [6]. An effective vaccine, which is still lacking, therefore is an urgent medical imperative. During its complicated life cycle both in mosquitoes and the human host, the *Plasmodium* parasite undergoes changes in morphology to enable it to invade different cell types, including red-blood cells. Once inside the red blood cell, the parasite expresses its own proteins on the red blood cell surface that interact with endothelial receptors, causing sequestration of the infected red blood cells to the tissue, a major cause of the symptoms of the disease. Targeting the blood stage of the life cycle with a vaccine is attractive since

at both invasion and during cyoadherence, the parasite exposes cell surface proteins to the host immune system. These parasite specific proteins therefore provide good targets for potential vaccines.

However, not only are there different species of human-infective malaria, but within each species, there are multiple different strains that express antigenically distinct variants of the same protein, for example *Pf*EMP1. This makes producing a vaccine that will block multiple different parasite strains a difficult challenge. However, the isolation of cross-reactive inhibitory antibodies from immunisations and from people in malarial-endemic areas provide hope that these antibodies can be produced. This combined with epitope-focused vaccine design methods, in which the structural epitopes, defined by protective antibodies, are isolated and incorporated into immunogens, show great promise for malarial vaccine development [137]. In the future, structural-based vaccine design efforts should focus on the the molecular characterisation of human monoclonal antibodies in complex with the parasite protein, which can then be used to generate an epitope-focused vaccine. However, even if immunogen design is successful in generating cross-reactive antibodies for one blood stage target, due to complexity of the parasite's life cycle, targeting one protein alone may not work (especially in *P. falciparum* disease, where there are many redundant pathways). Therefore, multicomponent vaccines that contain the inhibitory antibody epitopes of many malarial targets at different life cycle stages, may hold even greater promise.

Appendix A
Appendix

A.1 CIDRα1 and Bundle Sequences and Primers

A.1.1 Sequences of CIDRα1s on which Bundles were Designed

CIDRα1 Class	**CIDRα1 Name**	**Protein Sequence**
CIDRα1.1	IT4var20	PDCVVQCKGGKCTEDKKNDKCRSKIIKKIL QSEEPTEIHVLNSDDKQGDITKKLEVFCSSTT NYEGRNVQKWKCYNKNSDYNNCEMNISSY KDSTDANVMLSVECFHSWAKNLLIDTIKWE HQLKNCINNTNVTYCESKCIKNCECYEKWIK RKEHEWEKVKNVFGNNNRMSYIYYNNLSRV FDSFLFQVMFALDQDEKGKWDQFTEDLKK KFEPSKTNTPTGKSQDAIEFLLDHLKDNALT CRDNNSNESC
CIDRα1.4	HB3var03	PDCGVECKNETCTPKTVIYPDCGKNEKYEP PGDAKNTEINVINSGDKEGYIFEKLSEFCTN ENNENGKNYEQWKCYYDNKKNNNKCKME INIANSKLKNKITSFDEFFDFWVRKLLIDTIK WETELTYCINNTDVTDCNKCNKNCVCFDK WVKQKEDEWTNIMKLFTNKHDIPKKYYLN INDLFDSFFFQVIYKFNEGEAKWNELKENLK KQIASSKANNGTKDSEAAIKVLFNHIKEIATI CKDNNTNEGC

CIDRα1 Class	**CIDRα1 Name**	**Protein Sequence**
CIDRα1.5	1965-2	PDCGVECKNGRCKKKTDADGNCGKTQIYDI PKDVTPTDINVLYSGDEYGDIAKRLSEFCRD SKKENGKNTEIWHCYYIGDKHNQCKMEKA VAENKHQTKITTFDFFFDLWIKNLLRDTINW KSQLKNCINNTNKTNCNKTCNENCKCFENW VNQKEKEWNNMKELFKNKNRTSQNYYNKL KSHFDNYFFLVINNVNKGEAKWKKFTDELR NKIDSSKAKNGTSDSQDSIKMLLEHLKEDGT TCTANNPDSICNTPGTDARSLEPASKIEPQN KDTRTNPC
CIDRα1.6	PFD1235w	PECGVQCSGTTCTPKKVIHPNCKDKETYEP GDAKTTDITVLYSGDEEGDIAQKLQDFCND KNKENDENYEKWQCYYKSSEINKCQMTPSS HKVPKHGYIMSFYAFFDLWVKNLLIDSINW KNDLTNCINNTNVTDCKNDCNTNCKCFEN WAKTKENEWKKVKTIYKNENGNTNNYYKK LNNHFQGYFFHVMKELNKEEKWYKLMEDL KEKIDSSNLKNGTKDSEGAIKVLFDHLKDIA ERCIDNNSKDSC
CIDRα1.7	2083-1	PDCGVDCSSGTCEVKNKPDGNCGKNVKYE FPKGKPTTEITVLYSADQEGDISNKLSEFCK NPNDYDGKNYEKWQCYYKSSKDIKCHMTN LKQKVPKHLKVMIFDEFFDMWVTYLLTETI TWKDKLKKCMNNIKTADCIDECNKNCVCF DKWVKQKGKEWSSIKKLLRKKYNIPQQYY HKINYHFQGYFFHVMDKLNQDEAKWKELT EELRKKIDSSKEKSGTKDLQDAIELLLDHLK DNATICKDNNTNEACDPTGNPTQNPC
CIDRα1.8	1702-3	PDCIVVCDYKGCKENKNDENCKSKRTYSLP PDVNSTEIEVLFSGDNQEDIVEKLSSFCKNTN NENGENVEKWECYYESEHNNKCKMTSPKH KDEKRPTVMIFDEFFDFWVTHLIKDTIKWE SDINDCINNTNVTDCDSACNENCKCFDEWV EKKEVEWGNVKKVLGNRSENLNIYYNKLNG IFSGFFFGVMHELKKKEAKQGVKAEEAQEA EETQEAQEEEAKWNKLTAKLQEIIDSSKGKA DTANSQDAIKPLLEYIKETATTCIDNNSLAVE NCPKTKINPC

A.1.2 GeneString Sequences for Bundles

Construct Name	Sequence
Bundle1.1	CCATGGCTCATATTGTTATGGTGGATGCATACAA ACCGACCAAAGGTGGTAGCAATAATTGGGAACA GCAGAAAAAAAACATCGAGGATGATCTGGACCG TTACAAAAAACGTGCAGAAGAACTGCGTAAAGAA GCCGAAAAAGCACGCAAAGAAGCACGTAAAACC GAAGATCCGACCGAAGAGGCCAAAAAAGAATGG GAAAAACGTTGTAAAGAACTGGAAGAACGTGCC CGTAAACTGGAAGATGAAGCAAAAGATCGTGTT AGCCGTGTTTTTGATAGCAACCTGTTTCAGGTGA TTTTCAGCGGTGATAACGATGAAGAAGAGTGGA AAAAAGAAAAAGACCGTGCCGAGAAAGAAATCG AAGAATGGTTTAAACGCATCAAAGAAAAATGCGA AGAAATCAAAAAACGCCTGGAACAAGAGAACCTG TATTTTCAGGGTCATCATCATCACCATCACTAAG GATCC
Bundle1.4	CCATGGCTCATATTGTTATGGTGGATGCATACAA ACCGACCAAAGGTGGTAGCAATAATTGGGAACA GCAGAAAAAAAACATCGAGGATGATCTGGACCG TTACAAAAAACGTGCAGAAGAACTGCGTAAAGAA GCCGAAAAAGCACGCAAAGAAGCACGTAAAACC GAAGATCCGACCGAAGAGGCCAAAAAAGAATGG GAAAAACGTTGTAAAGAACTGGAAGAACGTGCC CGTAAACTGGAAGATGAAGCAAAAGATCGTGTG AACGACCTGTTTGATAGCAACTTTTTTCAGGTGA TCTATAGCGGTGATAACGATGAAGAAGAGTGGA AAAAAGAAAAAGACCGTGCCGAGAAAGAAATCG AAGAATGGTTTAAACGCATCAAAGAAAAATGCGA AGAAATCAAAAAACGCCTGGAACAAGAGAACCTG TATTTTCAGGGTCATCATCATCACCATCACTAAG GATCC

Construct Name	Sequence
Bundle1.5	CCATGGCTCATATTGTTATGGTGGATGCATACAA ACCGACCAAAGGTGGTAGCAATAATTGGGAACA GCAGAAAAAAAACATCGAGGATGATCTGGACCG TTACAAAAAACGTGCAGAAGAACTGCGTAAAGAA GCCGAAAAAGCACGCAAAGAAGCACGTAAAACC GAAGATCCGACCGAAGAGGCCAAAAAAGAATGG GAAAAACGTTGTAAAGAACTGGAAGAACGTGCC CGTAAACTGGAAGATGAAGCAAAAGATCGTGTG AAAAGCCACTTCGACAACAACTTTTTTCTGGTGA TTAACAGCGGTGATAACGATGAAGAAGAGTGGA AAAAAGAAAAAGACCGTGCCGAGAAAGAAATCG AAGAATGGTTTAAACGCATCAAAGAAAAATGCGA AGAAATCAAAAAACGCCTGGAACAAGAGAACCTG TATTTTCAGGGTCATCATCATCACCATCACTAAG GATCC
Bundle1.6	CCATGGCTCACATCGTGATGGTGGACGCTTACAA GCCCACCAAGGGTGGATCCAACAACTGGGAACA GCAGAAAAAAAACATCGAAGACGACCTGGACCG TTACAAAAAACGTGCGGAAGAACTGCGTAAAGAA GCGGAAAAAGCGCGTAAAGAAGCGCGTAAAACC GAAGACCCGACCGAAGAAGCGAAAAAAGAATGG GAAAAACGTTGCAAAGAACTGGAAGAACGTGCG CGTAAACTGGAAGACGAAGCGAAAGACCGTGTT AACAACCACTTCCAGGGTAACTTCTTCCACGTTA TCAAATCTGGTGACAACGACGAAGAAGAATGGA AAAAAGAAAAAGACCGTGCGGAAAAAGAAATCG AAGAATGGTTCAAACGTATCAAAGAAAAATGCG AAGAAATCAAAAAACGTCTGGAACAGGAGAACCT GTACTTCCAGGGCCACCACCACCACCACCACTAA GGATCC

Construct Name	Sequence
Bundle1.7	CCATGGCTCATATTGTTATGGTGGATGCATACAA ACCGACCAAAGGTGGTAGCAATAATTGGGAACA GCAGAAAAAAAACATCGAGGATGATCTGGACCG TTACAAAAAACGTGCAGAAGAACTGCGTAAAGAA GCCGAAAAAGCACGCAAAGAAGCACGTAAAACC GAAGATCCGACCGAAGAGGCCAAAAAAGAATGG GAAAAACGTTGTAAAGAACTGGAAGAACGTGCC CGTAAACTGGAAGATGAAGCAAAAGATCGTGTG AACTATCATTTCCAGGGCAACTTTTTTCATGTGA TCGATAGCGGTGATAACGATGAAGAAGAGTGGA AAAAAGAAAAAGACCGTGCCGAGAAAGAAATCG AAGAATGGTTTAAACGCATCAAAGAAAAATGCGA AGAAATCAAAAAACGCCTGGAACAAGAGAACCTG TATTTTCAGGGTCATCATCATCACCATCACTAAG GATCC
Bundle1.8	CCATGGCTCATATTGTTATGGTGGATGCATACAA ACCGACCAAAGGTGGTAGCAATAATTGGGAACA GCAGAAAAAAAACATCGAGGATGATCTGGACCG TTACAAAAAACGTGCAGAAGAACTGCGTAAAGAA GCCGAAAAAGCACGCAAAGAAGCACGTAAAACC GAAGATCCGACCGAAGAGGCCAAAAAAGAATGG GAAAAACGTTGTAAAGAACTGGAAGAACGTGCC CGTAAACTGGAAGATGAAGCAAAAGATCGTGTG AACGGTATTTTTAGCGGCAACTTTTTTGGTGTGA TTCACAGCGGTGATAACGATGAAGAAGAGTGGA AAAAAGAAAAAGACCGTGCCGAGAAAGAAATCG AAGAATGGTTTAAACGCATCAAAGAAAAATGCGA AGAAATCAAAAAACGCCTGGAACAAGAGAACCTG TATTTTCAGGGTCATCATCATCACCATCACTAAG GATCC

A.1.3 Primers for Cloning Bundles into pET15bcon

Primer	Sequence
Forward	GGATCTGGATCCAATAATTGGGAACAGCAGAAA AAAAACATCGAGGATGATCTGG
Reverse	GGATCCGCTAGCTTATTTGGTCGGTTTGTATGCA TCCACCATAAC

A.1.4 Protein Sequences for Bundles

Bundle	Sequence
Bundle1.1	NNWEQQKKNIEDDLDRYKKRAEELRKEAEKARKE ARKTEDPTEEAKKEWEKRCKELEERARKLEDEAK DRVSRVFDSNLFQVIFSGDNDEEEWKKEKDRAEKE IEEWFKRIKEKCEEIKKRLEQGGSAHIVMVDAYKPT K
Bundle1.4	NNWEQQKKNIEDDLDRYKKRAEELRKEAEKARKE ARKTEDPTEEAKKEWEKRCKELEERARKLEDEAK DRVNDLFDSNFFQVIYSGDNDEEEWKKEKDRAEKE IEEWFKRIKEKCEEIKKRLEQGGSAHIVMVDAYKPT K
Bundle1.5	NNWEQQKKNIEDDLDRYKKRAEELRKEAEKARKE ARKTEDPTEEAKKEWEKRCKELEERARKLEDEAK DRVKSHFDNNFFLVINSGDNDEEEWKKEKDRAEKE IEEWFKRIKEKCEEIKKRLEQGGSAHIVMVDAYKPT K
Bundle1.6	NNWEQQKKNIEDDLDRYKKRAEELRKEAEKARKE ARKTEDPTEEAKKEWEKRCKELEERARKLEDEAK DRVNNHFQGNFFHVIKSGDNDEEEWKKEKDRAEK EIEEWFKRIKEKCEEIKKRLEQGGSAHIVMVDAYKP TK
Bundle1.7	NNWEQQKKNIEDDLDRYKKRAEELRKEAEKARKE ARKTEDPTEEAKKEWEKRCKELEERARKLEDEAK DRVNYHFQGNFFHVIDSGDNDEEEWKKEKDRAEK EIEEWFKRIKEKCEEIKKRLEQGGSAHIVMVDAYKP TK
Bundle1.8	NNWEQQKKNIEDDLDRYKKRAEELRKEAEKARKE ARKTEDPTEEAKKEWEKRCKELEERARKLEDEAK DRVNGIFSGNFFGVIHSGDNDEEEWKKEKDRAEKE IEEWFKRIKEKCEEIKKRLEQGGSAHIVMVDAYKPT K

A.2 Sequences and Primers for CIDRα1 Constructs

A.2.1 Sequences for CIDRα1 Constructs

CIDRα1 Class	CIDRα1 Name	DNA Sequence
CIDRα1.1	IT4var20	CCCGATTGTGTAGTTCAATGTAAGGGTGG AAAATGTACAGAAGATAAGAAAAATGACA AATGCAGGAGCAAAATAATCAAAAAAATT CTACAAAGTGAGGAACCTACTGAAATTCA TGTCCTCAATAGTGATGATAAACAAGGTG ATATTACAAAAAAATTAGAAGTTTTTTGT TCGAGCACAACTAATTACGAAGGTCGAAA TGTTCAAAAATGGAAATGCTATAATAAAA ATAGTGATTATAATAATTGTGAAATGAAC ATTTCATCATACAAAGATTCAACAGATGC TAACGTTATGTTATCTGTTGAATGTTTTC ATTCGTGGGCTAAAAATTTATTAATAGAC ACCATAAAGTGGGAACATCAACTTAAGAA TTGCATAAATAATACTAATGTCACGTATT GTGAAAGTAAATGTATCAAAAATTGTGAA TGTTATGAAAAATGGATTAAACGAAAAGA ACATGAATGGGAAAAAGTGAAGAATGTAT TCGGAAATAATAATAGAATGTCGTACATT TATTATAATAACCTTAGTCGTGTTTTTGA TAGTTTTTTATTTCAAGTTATGTTTGCGC TTGACCAAGACGAAAAAGGAAAATGGGAT CAATTTACGGAAGATTTAAAAAAAAAATT TGAACCTTCCAAAACAAACACACCTACTG GAAAGTCACAAGATGCAATAGAATTCTTA TTAGATCACTTAAAAGATAATGCCTTAAC ATGCAGAGACAATAATTCAAACGAATCAT GT

CIDRα1 Class	CIDRα1 Name	DNA Sequence
CIDRα1.4	HB3var03	CCAGACTGCGGAGTCGAATGTAAGAATGA AACATGCACACCAAAAACAGTGATATATC CGGATTGTGGAAAAAATGAAAAATATGAA CCTCCTGGTGATGCAAAAAATACTGAAAT TAATGTAATTAATAGTGGTGATAAAGAAG GTTATATTTTTGAAAAATTAAGTGAATTT TGTACTAACGAAAATAATGAAAATGGTAA AAACTATGAACAATGGAAATGCTATTACG ATAATAAAAAAAAATAATAATAAATGTAAA ATGGAAATAAATATTGCAAACAGTAAGCT TAAAAATAAAATAACGTCATTTGATGAAT TTTTTGATTTTTGGGTTAGAAAATTATTA ATAGACACTATAAAGTGGGAAACCGAACT TACGTATTGTATAAATAATACTGATGTCA CGGATTGTAATAAATGTAACAAAAATTGC GTATGTTTTGACAAATGGGTTAAACAAAA AGAAGACGAATGGACAAATATAATGAAAC TATTCACAAACAAACACGATATACCGAAA AAATATTATCTTAATATTAATGATCTTTT TGATAGTTTTTTTTTCCAAGTTATATATA AGTTTAACGAAGGAGAAGCAAAATGGAAT GAACTTAAAGAAAATTTAAAAAAGCAAAT TGCGTCTTCCAAAGCAAATAACGGAACCA AAGATTCAGAAGCTGCAATAAAAGTGTTG TTTAATCACATAAAAGAAATTGCAACAAT ATGCAAAGATAATAATACAAACGAAGGAT GT

CIDRα1 Class	CIDRα1 Name	DNA Sequence
CIDRα1.5	GA12	AAAGATGTGACCCCGACCGATATTAACGT GCTGTATAGCGGCGATGAATATGGCGATA TTGCGAAACGCCTGAGCGAATTTTGCCGC GATAGCAAAAAAGAAAACGGCAAAAACAC CGAAATTTGGCATTGCTATTATATTGGCG ATAAACATAACCAGTGCAAAATGGAAAAA GCGGTGGCGGAAAACAAACATCAGACCAA AATTACCACCTTTGATTTTTTTTTTGATC TGTGGATTAAAAACCTGCTGCGCGATACC ATTAACTGGAAAAGCCAGCTGAAAAACTG CATTAACAACACCAACAAAACCAACTGCA ACAAAACCTGCAACGAAAACTGCAAATGC TTTGAAAACTGGGTGAACCAGAAAGAAAA AGAATGGAACAACATGAAAGAACTGTTTA AAAACAAAAACCGCACCAGCCAGAACTAT TATAACAAACTGAAAAGCCATTTTGATAA CTATTTTTTTCTGGTGATTAACAACGTGA ACAAAGGCGAAGCGAAATGGAAAAAATTT ACCGATGAACTGCGCAACAAAATTGATAG CAGCAAAGCGAAAAACGGCACCAGCGATA GCCAGGATAGCATTAAAATGCTGCTGGAA CATCTGAAAGAAGATGGCACCACCTGCAC CGCGAACAACCCGGATAGCATTTGCAACA CCCCGGGCACCGATGCGCGCAGCCTGGAA CCGGCGAGCAAAATTGAACCGCAGAACAA AGATACCCGCACCAACCCGTGC

CIDRα1 Class	CIDRα1 Name	DNA Sequence
CIDRα1.6	PF08	CCGGACTGCGGGGTTGAATATAAGAATG GAAGATATACAGCTAAAGATCAAAAATAT CCGGATTGTAGAAATGAAAAATATGATCC TAATAATGCAGAAACCACTGACATTACTG TCCTCTATAGTGGTGATGTTGGTGATTTT TCTGAAAAATTACAAGATTTTTGTAATGA CATAAATAATGGGAAAGTTAAAAACTATC AAATATGGCAATGCTATTATGAAAATAGT GAGATTAATAAATGTCAAATGACACCATC ATCACACAAAGTTCCAAAACATGGTTACA TTATGTCATTTTTTGCTTTTTTTGATTTG TGGGTTAAGAATCTATTAATAGACACTAT AAATTGGAAAAACGAACTTACAAATTGTA TAAATAATACTAATGTCACGGATTGTAAA AATGATTGTAACACAAATTGTAAATGTTT TGAAAATTGGGCTAAAACAAAAGAAAAAG AATGGGAAAATGTGAAGACGATATACAAA AATGAAAACAGAAACACGAACAATTATTA TAAAAAACTTAATGATCTTTTTAAAGGCT ATTTTTTTCATGTTATGTATGAGCTTAAC AATGAAGAAGCAAAATGGAATAAACTTAT GAAAAATTTAAGAACAAAAATTGATTCTT CCAGAAAAAATGCAGGTAACGAAGATTCA GAAGGCGCAATAAAAGTGTTGTTCGATCA CTTAAAAGATATAGCTGAAAGATGCATAG ACAATAATTCAATCAAACCATGT

CIDRα1 Class	CIDRα1 Name	DNA Sequence
CIDRα1.7	GAO20	CAACGGTGCTTACGAGCCCCCCAACGGTG
		CTGAGACTACCGAGATCACCGTGCTGTAC
		TCCGCTGACCAAGAGGGCGACATCTCCAA
		CAAGCTGTCCGAGTTCTGCAACGACGAGA
		ACAACAAGAACTCCCAAAAGTGGCAGTGC
		TACTACGTGTCCTCCGAGAACAACGGTTG
		CAAGATGGAAAAGAAGAACGCCAACCACA
		CCCCCGAAGTGAAGATCACCAAGTTCCAC
		AACTTCTTCGAGATGTGGGTCACCTACCT
		GCTGACCGAGACTATCACCTGGAAGGACA
		AGCTCAAGACTTGCATGAACAACACCAAG
		ACCGCTGACTGCATCCACGAGTGTAACAA
		GAACTGCGTGTGCTTCGATAAGTGGGTCA
		AGCAAAAGGAAGATGAGTGGAACTCCATC
		AAGAAGCTGTTCACCAAGGAAAAGAAGAT
		GCCCAAGCAGTACTACGGCAACATCAACA
		TCTACTTCGAGTCGTTCTTCTTCCACGTG
		ATGAAGAAGCTGAACAAGGAAGCCAAGT
		GGAACAAGCTGATGGACGAGCTGCGTAA
		CAAGATCGAGCTGTCCAAGGGCAACGAG
		GGCACCAAGGACCTGCAGGACGCTATCGA
		GCTGCTGCTCGAGTACCTGAAGGAAAAGT
		CCACCATCTGCAAGGACAACAACACTAAC
		GAGGCTTGCGACCCCACC

CIDRα1 Class	CIDRα1 Name	DNA Sequence
CIDRα1.8	MAL6P	CCCGCTTGTGTAGTTGAATGTGACGGTG GAAAATGTGAAGAGAAAAAAAACAGTGA TGGTACTTGTATTGAAGCACAGATATATA CAGTTGTACGAGATGAGACGCCTACTCCA ATTAAAGTTCTCTTTAGCGGAGACCACCA GAAAGATATAACAAAAAAATTAAGTTCAT TCTGTAAGAACCCAGAAAGTGAAAATAAT AGGGACTATCAAACATGGCAATGCTATTA TAAAAGTAGTGATTATAATAATTGTGAAA TGAAAGGTTCATTATACAAGGTTGAAGGA GATCCTAACATTATTGTATCTCATGAATG TTTTCACTTGTGGGTGCAAAGTTTATTAA TAGACACCATAAAGTGGGAAACCAAACTT AAGAAGTGCATAAATAATACTAACGTCAC GAATTGTTATAATAAATGTAACAAAAATT GTGAATGTTTTGAAAATTGGGTTGAACAA AAAAAAAAAGAATGGGAAAATGTAAACGA TGTATACAAAGATCAGAAACAAAGTTTGG GCATTTATTATGAAAAACTCGAAAATCTT TTTAAAAGTAATTTTTTTCAAGTTATGAA AGCTCTTGAGGGAGATGAAAAAGGAAAAT GGTATCAATTTAAGGACGATTTAAAAAAA AAATTTGAACCTTCCGAAAAAAACACACG TACTACAGATTCGCAAGATGCAATAAAAC TCATATTAGATCACTTAAAAGATAATGCC ACAACATGCAAAGACAATAATTCATTAGA AGAAGATGAAAATTGC

A.2.2 Primers for CIDRα1 Constructs

Construct	Primer Direction	Sequence
GA12	Forward	CGCAAATAATCCGGATAGT ATTTGTTAAGCTAGCTCGA GGTACCATATGC
GA12	Reverse	GCATATGGTACCTCGAGCT AGCTTAACAAATACTATCC GGATTATTTGCG
PF08	Forward	CCTGACGGGATCCGATCCG AATAATG
PF08	Reverse	GGCGTAGCTAGCTTAGCAC GGTTTAATG
GAO20	Forward	GCGGATCCGAGCCCCCCAA CGGTGCTGAGAC
GAO20	Reverse	GCGCTAGCTTAGCAAGCCT CGTTAGTGTTGTTGTCCTT GC
MAL6P	Forward	CCTGACGGGATCCAGAGAT GAAACCCCG
MAL6P	Reverse	GGCGTAGCTAGCTTAGTCC TCTTCCAGG

A.2.3 Primers for CIDRα1 Minus Beta, Minus Loop Constructs

Construct	Primer Direction	Sequence
HB3var03$_{-loop}$ Loop Deletion	Forward	CAAATTGCGTCTTCCAAAGCTGCAATAAAAGTG
HB3var03$_{-loop}$ Loop Deletion	Reverse	CACTTTTATTGCAGCTTTGGAAGACGCAATTTG
HB3var03$_{-loop}$ Linker Insertion	Forward	GCGTCTTCCAAAGGCGGCAGCGGCGGCAGCGCTGCAATAAAAG
HB3var03$_{-loop}$ Linker Insertion	Reverse	CTTTTATTGCAGCGCTGCCGCCGCTGCCGCCTTTGGAAGACGC
HB3var03$_{-\beta}$	Forward	CGTACGGATCCTCATTTGATGAATTTTTTGATTTTTGG
HB3var03$_{-\beta}$	Reverse	CGCATCATATGTTAACATCCTTCGTTTGTATTATTATC
IT4var20$_{-loop}$ Loop Deletion	Forward	TTTGAACCTTCCAAAGATGCAATAGAATTC
IT4var20$_{-loop}$ Loop Deletion	Reverse	GAATTCTATTGCATCTTTGGAAGGTTCAAA
IT4var20$_{-loop}$ Linker Insertion	Forward	CCTTCCAAAGGCGGCAGCGGCGGCAGCGATGCAATAG
IT4var20$_{-loop}$ Linker Insertion	Reverse	CTATTGCATCGCTGCCGCCGCTGCCGCCTTTGGAAGG
IT4var20$_{-beta}$	Forward	CGTACGGATCCTCTGTTGAATGTTTTCATTCGTGGG
IT4var20$_{-beta}$	Reverse	CGCATCATATGTTAACATGATTCGTTTGAATTATTGTCTCT

A.3 Sequences and Primers for *Pv*DBP-RII Project

A.3.1 DBP Sequence

	Protein Sequence
*Pv*DBP-RII	MGHHHHHHENLYFQGGSGNTVMKNCNYKRKRRER DWDCNTKKDVCIPDRRYQLCMKELTNLVNNTDTN FHRDITFRKLYLKRKLIYDAAVEGDLLLKLNNYRYN KDFCKDIRWSLGDFGDIIMGTDMEGIGYSKVVENNL RSIFGTDEKAQQRRKQWWNESKAQIWTAMMYSVK KRLKGNFIWICKLNVAVNIEPQIYRWIREWGRDYVS ELPTEVQKLKEKCDGKINYTDKKVCKVPPCQNACK SYDQWITRKKNQWDVLSNKFISVKNAEKVQTAGIV TPYDILKQELDEFNEVAFENEINKRDGAYIELCVCS

A.3.2 DARC Construct

	Protein Sequence
DARC	MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGTM GNALHRAELSPSTENSSQLDFEDVWNSSYGVNDSFP DGDYDANLEAAAPAHSANLLDDSLVPRGSGLNDIFE AQKIEWHEGSHHHHHH

A.3.3 Immunogen Primers

Construct	Primer Direction	Sequence
Subdomain3	Forward	CCTGACGGGATCCCCCCAGATCTA CCGGTGGATCAGAGAG
Subdomain3	Reverse	GGCGTAGCTAGCCTAGCTGCACAC GCACAGCTCGATG
PROSS T114A	Forward	CAATAAACGTGATGGCGCCTATAT TGAACTGTGC
PROSS T114A	Reverse	GCACAGTTCAATATAGGCGCCATC ACGTTTATTG
Interface R59C	Forward	GATCAGTGGATTACCTGCAAAAAA AACGAATGG
Interface R59C	Reverse	CCATTCGTTTTTTTTGCAGGTAAT CCACTGATC

A.4 Gene Identifiers for DBP Sequences in PROSS Alignment

CCP37649.1 BAM17299.1 C2C6JA BAH96577.1 AAG53619.1 AAC47175.1 ACD76811.1 AHV83764.1 AAC47188.1 ACD76812.1 AHB23289.1 AAC47190.1 AHB23287.1 AAC47189.1 AGH67770.1 AAC47187.1 AAC47182.1 AKS26862.1 ACJ54192.1 ACJ54187.1 AAG53621.1 ABR08404.1 AHV83760.1 AFD18603.1 AAG30849.1 ACD76813.1 AAG53617.1 AAF25492.1 ACD76803.1 ACD76805.1 AAG53626.1 AAA61633.1 ABR14008.1 AAA61768.1 ACD76809.1 AAA61766.1 ABA39301.1 AAL79058.1 AAL79120.1 AAL79070.1 AAL79053.1 AAL79071.1 ACY92023.1 ACY92009.1 ACY92055.1 AKU77039.1 ACY92049.1 AAF25485.2 AKU77041.1 AKU77040.1 ACY92057.1 AFD18605.1 ABR14007.1 ACB42425.1 ACY92040.1 AAF25494.2 AAG53623.1 AFM44716.1 AFM44717.1 AMQ67631.1 AMQ67621.1 AMQ67627.1 AMQ67633.1 XP019917116.1 AMQ67625.1 AGL76504.1 AGL76478.1 AJH76907.1 AGL76451.1 AJH76927.1 AKC01008.1 SBO28752.1 AGL76487.1 AGL76500.1 SBO22280.1 CAQ42950.1 AGL76455.1 XP019913776.1 AGL76456.1 BAM17299.1 GAW80084.1 GAW84276.1 BAM17300.1 SBT83545.1 SBS98843.1 SCA59679.1

A.5 Var2CSA New Constructs Primers

Construct	Primer Direction	Sequence
NTS-	Forward	GCCGAATTCATGGACAGCACCAGCACAATCGCCAACAAGATC
-DBL2	Reverse	GCGGCGGCCGCTTACTTGCACTGGCAGGCGTACTTGTACCGGTAG
-CIDR	Reverse	CGCGCGGCCGCTTAATTGGTGTCGGTGCTCTCGTTCTCTTTGCACTTCTTC
-DBL3	Reverse	GCGGCGGCCGCTTAGCAGCTGCAGATGCTGCTGTAATCGCCGTAG
-DBL4	Reverse	GCGGCGGCCGCTTACTTGTCCATCTTGTTCAGCACCTCGTTCATCAGGTCC
-DBL6	Reverse	GCGGCGGCCGCTTAAGGGCAGTCGCACTTGCTCTTGAAGGTATCTTCC
CIDR-	Forward	GCCTCTAGAATGAAGAGAGGGTATAAGAACGATAACTACGAACTGTGCAAGTACAACG
DBL3-	Forward	GCCTCTAGAAGCGAGACATCCTGCGACCTGAACGCC
ΔCIDR-linker	Forward	GGATCCTCCGGCAGCTCCGACATCCCCCCTCGCACCCAGAACCTG
ΔCIDR-linker	Reverse	GGAGCTGCCATCGGAGCTGCCGGGGATCTTGCACTGGCAGGC

Publications

T. A. Rawlinson, N. M. Barber, F. Mohring, J. S. Cho, V. Kosaisavee, S. F.Gérard, D. G. W. Alanine, G. M. Labbé, S. C. Elias, S. E. Silk, D. Quinkert, J. Jin, J. M. Marshall, R. O. Payne, A. M. Minassian, B. Russell, L. Rénia,F. H. Nosten, R. W. Moon, M. K. Higgins, and S. J. Draper, **"Structural basis for inhibition of Plasmodium vivax invasion by a broadly neutralizing vaccine-induced human antibody"** Nature Microbiology, 2019.

N. M. Barber, C. K. Y. Lau, L. Turner, G. Watson, S. Thrane, J. P. A. Lusingu, T. Lavstsen, and M. K. Higgins, **"Structure-guided design of a synthetic mimic of an EPCR-binding PfEMP1 protein"** bioRxiv, p. 749432, 1 201

Bibliography

[1] R. L. Miller, S. Ikram, G. J. Armelagos, R. Walker, W. B. Harer, C. J. Shiff, D. Baggett, M. Carrigan, and S. M. Maret, "Diagnosis of Plasmodium falciparum infections in mummies using the rapid manual ParaSight™-F test," *Transactions of The Royal Society of Tropical Medicine and Hygiene*, vol. 88, pp. 31–32, 1 1994.

[2] A. Laveran, *Nature parasitaire des accidents de l'impaludisme : description d'un nouveau parasite trouve dans le sang des malades atteints de fievre palustre*. 1881.

[3] R. Ross, "THE RÔLE OF THE MOSQUITO IN THE EVOLUTION OF THE MALARIAL PARASITE.," *The Lancet*, vol. 152, pp. 488–490, 8 1898.

[4] B. Grassi, "Studi di uno zoologo sulla malaria / Battista Grassi.," *Studi di uno zoologo sulla malaria / Battista Grassi.*, vol. 3, 1900.

[5] R. Carter and K. N. Mendis, "Evolutionary and historical aspects of the burden of malaria," *Clin Microbiol Rev*, vol. 15, 2002.

[6] WHO, *World Malaria Report 2018*. 2018.

[7] J. W. W. Stephens, "A New Malaria Parasite of Man," *Annals of Tropical Medicine & Parasitology*, vol. 16, pp. 383–388, 12 1922.

[8] P. C. C. Garnham, "The Primate Malarias," *Science*, vol. 177, pp. 50 LP – 50, 7 1972.

[9] N. J. White, "Plasmodium knowlesi: The Fifth Human Malaria Parasite," *Clinical Infectious Diseases*, vol. 46, pp. 172–173, 1 2008.

[10] K. Mendis, B. J. Sina, P. Marchesini, and R. Carter, "The neglected burden of Plasmodium vivax malaria," *Am J Trop Med Hyg*, vol. 64, 2001.

[11] R. Rappuoli and A. Aderem, "A 2020 vision for vaccines against HIV, tuberculosis and malaria," *Nature*, vol. 473, p. 463, 5 2011.

[12] R. Rosenberg, "Plasmodium vivax in Africa: hidden in plain sight?," *Trends Parasitol*, vol. 23, 2007.

[13] N. M. Anstey, N. M. Douglas, J. R. Poespoprodjo, and R. N. Price, "Chapter Three - Plasmodium vivax: Clinical Spectrum, Risk Factors and Pathogenesis," in *Advances in Parasitology* (S. I. Hay, R. Price, and J. K. B. T. A. i. P. Baird, eds.), vol. 80, pp. 151–201, Academic Press, 2012.

[14] R. N. Price, E. Tjitra, C. A. Guerra, S. Yeung, N. J. White, and N. M. Anstey, "Vivax malaria: neglected and not benign," *The American journal of tropical medicine and hygiene*, vol. 77, pp. 79–87, 12 2007.

[15] K. E. Battle, C. A. Guerra, N. Golding, K. A. Duda, E. Cameron, R. E. Howes, I. R. F. Elyazar, J. K. Baird, R. C. Reiner Jr., P. W. Gething, D. L. Smith, and S. I. Hay, "Global database of matched Plasmodium falciparum and P.

vivax incidence and prevalence records from 1985–2013," *Scientific Data*, vol. 2, p. 150012, 8 2015.

[16] J. P. Vanderberg and U. Frevert, "Intravital microscopy demonstrating antibody-mediated immobilisation of Plasmodium berghei sporozoites injected into skin by mosquitoes," *International Journal for Parasitology*, vol. 34, no. 9, pp. 991–996, 2004.

[17] R. Amino, S. Thiberge, B. Martin, S. Celli, S. Shorte, F. Frischknecht, and R. Ménard, "Quantitative imaging of Plasmodium transmission from mosquito to mammal," *Nature Medicine*, vol. 12, no. 2, pp. 220–224, 2006.

[18] G. Pradel, S. Garapaty, and U. Frevert, "Proteoglycans mediate malaria sporozoite targeting to the liver," *Molecular Microbiology*, vol. 45, no. 3, pp. 637–651, 2002.

[19] M. M. Mota, G. Pradel, J. P. Vanderberg, U. Frevert, R. S. Nussenzweig, V. Nussenzweig, and A. Rodríguez, "Migration of Plasmodium Sporozoites through Cells before Infection," *Science*, vol. 291, no. 5501, pp. 141–144, 2001.

[20] M. M. Mota, J. C. R. Hafalla, and A. Rodriguez, "Migration through host cells activates Plasmodium sporozoites for infection," *Nature Medicine*, vol. 8, no. 11, pp. 1318–1322, 2002.

[21] A. Sturm, R. Amino, C. Van De Sand, T. Regen, S. Retzlaff, A. Rennenberg, A. Krueger, J.-M. Pollok, R. Menard, and V. T. Heussler, "Manipulation of host hepatocytes by the malaria parasite for delivery into liver sinusoids," *Science*, vol. 313, no. 5791, pp. 1287–1290, 2006.

[22] K. Baer, C. Klotz, S. H. I. Kappe, T. Schnieder, and U. Frevert, "Release of Hepatic Plasmodium yoelii Merozoites into the Pulmonary Microvasculature (Plasmodium Merozoite Release in the Lung)," *PLoS Pathogens*, vol. 3, no. 11, p. e171, 2007.

[23] G. E. Weiss, P. R. Gilson, T. Taechalertpaisarn, W.-H. Tham, N. W. M. de Jong, K. L. Harvey, F. J. I. Fowkes, P. N. Barlow, J. C. Rayner, G. J. Wright, A. F. Cowman, and B. S. Crabb, "Revealing the Sequence and Resulting Cellular Morphology of Receptor-Ligand Interactions during Plasmodium falciparum Invasion of Erythrocytes," *PLOS Pathogens*, vol. 11, p. e1004670, 2 2015.

[24] D. G. W. Alanine, D. Quinkert, R. Kumarasingha, S. Mehmood, F. R. Donnellan, N. K. Minkah, B. Dadonaite, A. Diouf, F. Galaway, S. E. Silk, A. Jamwal, J. M. Marshall, K. Miura, L. Foquet, S. C. Elias, G. M. Labbé, A. D. Douglas, J. Jin, R. O. Payne, J. J. Illingworth, D. J. Pattinson, D. Pulido, B. G. Williams, W. A. de Jongh, G. J. Wright, S. H. I. Kappe, C. V. Robinson, C. A. Long, B. S. Crabb, P. R. Gilson, M. K. Higgins, and S. J. Draper, "Human Antibodies that Slow Erythrocyte Invasion Potentiate Malaria-Neutralizing Antibodies," *Cell*, vol. 178, no. 1, pp. 216–228, 2019.

[25] A. F. Cowman and B. S. Crabb, "Invasion of Red Blood Cells by Malaria Parasites," *Cell*, vol. 124, pp. 755–766, 2 2006.

[26] H. Sakamoto, S. Takeo, A. G. Maier, J. Sattabongkot, A. F. Cowman, and T. Tsuboi, "Antibodies against a Plasmodium falciparum antigen PfMSPDBL1 inhibit merozoite invasion into human erythrocytes," *Vaccine*, vol. 30, no. 11, pp. 1972–1980, 2012.

[27] V. K. Goel, X. Li, H. Chen, S.-C. Liu, A. H. Chishti, and S. S. Oh, "Band 3 is a host receptor binding merozoite surface protein 1 during the Plasmodium falciparum invasion of erythrocytes," *Proceedings of the National Academy of Sciences*, vol. 100, pp. 5164 LP – 5169, 4 2003.

[28] S. Lopaticki, A. G. Maier, J. Thompson, D. W. Wilson, W.-H. Tham, T. Triglia, A. Gout, T. P. Speed, J. G. Beeson, J. Healer, and A. F. Cowman, "Reticulocyte and erythrocyte binding-like proteins function cooperatively in invasion of human erythrocytes by malaria parasites," *Infection and immunity*, vol. 79, pp. 1107–1117, 3 2011.

[29] S. Besteiro, J.-F. Dubremetz, and M. Lebrun, "The moving junction of apicomplexan parasites: a key structure for invasion," *Cellular Microbiology*, vol. 13, pp. 797–805, 6 2011.

[30] G. E. Ward, L. H. Miller, and J. A. Dvorak, "The origin of parasitophorous vacuole membrane lipids in malaria-infected erythrocytes," *Journal of Cell Science*, vol. 106, pp. 237 LP – 248, 9 1993.

[31] D. T. Riglar, D. Richard, D. W. Wilson, M. J. Boyle, C. Dekiwadia, L. Turnbull, F. Angrisano, D. S. Marapana, K. L. Rogers, C. B. Whitchurch, J. G. Beeson, A. F. Cowman, S. A. Ralph, and J. Baum, "Super-Resolution Dissection of Coordinated Events during Malaria Parasite Invasion of the Human Erythrocyte," *Cell Host & Microbe*, vol. 9, no. 1, pp. 9–20, 2011.

[32] M. Aikawa, L. H. Miller, J. Johnson, and J. Rabbege, "Erythrocyte entry by malarial parasites. A moving junction between erythrocyte and parasite," *The Journal of cell biology*, vol. 77, pp. 72–82, 4 1978.

[33] M. J. Blackman and V. B. Carruthers, "Recent insights into apicomplexan parasite egress provide new views to a kill," *Current Opinion in Microbiology*, vol. 16, no. 4, pp. 459–464, 2013.

[34] A. M. Talman, O. Domarle, F. E. McKenzie, F. Ariey, and V. Robert, "Gametocytogenesis : the puberty of Plasmodium falciparum," *Malaria Journal*, vol. 3, no. 1, p. 24, 2004.

[35] M. E. Smalley, J. Brown, and N. M. Bassett, "The rate of production of Plasmodium falciparum gametocytes during natural infections," *Transactions of The Royal Society of Tropical Medicine and Hygiene*, vol. 75, pp. 318–319, 1 1981.

[36] M. C. Bruce, P. Alano, S. Duthie, and R. Carter, "Commitment of the malaria parasite Plasmodium falciparum to sexual and asexual development," *Parasitology*, vol. 100, no. 2, pp. 191–200, 1990.

[37] A. S. I. Aly, A. M. Vaughan, and S. H. I. Kappe, "Malaria Parasite Development in the Mosquito and Infection of the Mammalian Host," *Annual Review of Microbiology*, vol. 63, pp. 195–221, 9 2009.

[38] T. Kariu, T. Ishino, K. Yano, Y. Chinzei, and M. Yuda, "CelTOS, a novel malarial protein that mediates transmission to mosquito and vertebrate hosts," *Molecular Microbiology*, vol. 59, pp. 1369–1379, 3 2006.

[39] A. S. I. Aly and K. Matuschewski, "A malarial cysteine protease is necessary for Plasmodium sporozoite egress from oocysts," *The Journal of Experimental Medicine*, vol. 202, pp. 225 LP – 230, 7 2005.

[40] M. G., *The Epidemiology and Control of Malaria.* Amen House, Warwick Square, London E.C.4.: Oxford University Press, 1957.

[41] P. W. Gething, I. R. F. Elyazar, C. L. Moyes, D. L. Smith, K. E. Battle, C. A. Guerra, A. P. Patil, A. J. Tatem, R. E. Howes, M. F. Myers, D. B. George, P. Horby, H. F. L. Wertheim, R. N. Price, I. Müeller, J. K. Baird, and S. I. Hay, "A long neglected world malaria map: Plasmodium vivax endemicity in 2010," *PLoS neglected tropical diseases*, vol. 6, pp. e1814–e1814, 9 2012.

[42] I. Mueller, M. R. Galinski, J. K. Baird, J. M. Carlton, D. K. Kochar, P. L. Alonso, and H. A. del Portillo, "Key gaps in the knowledge of Plasmodium vivax, a neglected human malaria parasite," *The Lancet Infectious Diseases*, vol. 9, pp. 555–566, 9 2009.

[43] M. R. Galinski, C. C. Medina, P. Ingravallo, and J. W. Barnwell, "A reticulocyte-binding protein complex of plasmodium vivax merozoites," *Cell*, vol. 69, pp. 1213–1226, 6 1992.

[44] J. Li and E.-T. Han, "Dissection of the Plasmodium vivax reticulocyte binding-like proteins (PvRBPs)," *Biochemical and Biophysical Research Communications*, vol. 426, pp. 1–6, 9 2012.

[45] R. J. Wilson, G. Pasvol, and D. J. Weatherall, "Invasion and growth of Plasmodium falciparum in different types of human erythrocyte," *Bulletin of the World Health Organization*, vol. 55, no. 2-3, pp. 179–186, 1977.

[46] G. Pasvol, D. J. Weatherall, and R. J. M. Wilson, "The Increased Susceptibility of Young Red Cells to Invasion by the Malarial Parasite Plasmodium falciparum," *British Journal of Haematology*, vol. 45, pp. 285–295, 6 1980.

[47] R. Suwanarusk, B. M. Cooke, A. M. Dondorp, K. Silamut, J. Sattabongkot, N. J. White, and R. Udomsangpetch, "The Deformability of Red Blood Cells Parasitized by Plasmodium falciparum and P. vivax," *The Journal of Infectious Diseases*, vol. 189, pp. 190–194, 1 2004.

[48] R. HEGNER, "RELATIVE FREQUENCY OF RING-STAGE PLASMODIA IN RETICULOCYTES AND MATURE ERYTHROCYTES IN MAN AND MONKEY," *American Journal of Epidemiology*, vol. 27, pp. 690–718, 5 1938.

[49] L. J. Chwatt, "Infection of Reticulocytes by Plasmodium Falciparum and Plasmodium Malariae in Hyperendemic Indigenous Malaria," *Annals of Tropical Medicine & Parasitology*, vol. 42, pp. 101–112, 4 1948.

[50] S. Handayani, D. T. Chiu, E. Tjitra, J. S. Kuo, D. Lampah, E. Kenangalem, L. Renia, G. Snounou, R. N. Price, N. M. Anstey, and B. Russell, "High Deformability of Plasmodium vivax-Infected Red Blood Cells under Microfluidic Conditions," *The Journal of Infectious Diseases*, vol. 199, pp. 445–450, 2 2009.

[51] S. F. Kitchen and M. F. Boyd, "On the Infectiousness of Patients Infected with Plasmodium Vivax and Plasmodium Falciparum 1," *The American Journal of Tropical Medicine and Hygiene*, vol. s1-17, pp. 253–262, 3 1937.

[52] T. Bousema and C. Drakeley, "Epidemiology and Infectivity of Plasmodium falciparium and Plasmodium vivax gametocytes in relation to malaria control and elimination," *Clinical Microbiology Reviews*, vol. 24, pp. 377 LP – 410, 4 2011.

[53] K. Singh, P. Mukherjee, A. R. Shakri, A. Singh, G. Pandey, M. Bakshi, G. Uppal, R. Jena, A. Rawat, P. Kumar, R. Bhardwaj, S. S. Yazdani, D. Hans, S. Mehta, A. Srinivasan, K. Anil, R. L. Madhusudhan, J. Patel, A. Singh, R. Rao, S. Gangireddy, R. Patil, S. Kaviraj, S. Singh, D. Carter, S. Reed, D. C. Kaslow, A. Birkett, V. S. Chauhan, and C. E. Chitnis, "Malaria vaccine candidate based on Duffy-binding protein elicits strain transcending functional antibodies in a Phase I trial," *NPJ vaccines*, vol. 3, p. 48, 9 2018.

[54] T. Harinasuta and D. Bunnang, "Malaria, principles and practice of malariology," *First edittion. Edinburg: Churchill Livingstone*, 1988.

[55] H. M. Gilles and D. A. Warrell, *Bruce-Chwatt's essential malariology.* No. Ed. 3, Edward Arnold (Publisher) Ltd., 1996.

[56] C. L. Mackintosh, J. G. Beeson, and K. Marsh, "Clinical features and pathogenesis of severe malaria," *Trends in Parasitology*, vol. 20, no. 12, pp. 597–603, 2004.

[57] L. H. Miller, H. C. Ackerman, X.-z. Su, and T. E. Wellems, "Malaria biology and disease pathogenesis: insights for new treatments," *Nature Medicine*, vol. 19, p. 156, 2 2013.

[58] N. J. White and M. Ho, "The pathophysiology of malaria," in *Advances in parasitology*, vol. 31, pp. 83–173, Elsevier, 1992.

[59] A. M. Dondorp, P. A. Kager, J. Vreeken, and N. J. White, "Abnormal Blood Flow and Red Blood Cell Deformability in Severe Malaria," *Parasitology Today*, vol. 16, no. 6, pp. 228–232, 2000.

[60] L. H. Miller, S. Usami, and S. Chien, "Alteration in the rheologic properties of plasmodium knowlesi-infected red cells. A possible mechanism for capillary obstruction," *The Journal of Clinical Investigation*, vol. 50, pp. 1451–1455, 7 1971.

[61] A. Bachmann, C. Esser, M. Petter, S. Predehl, V. von Kalckreuth, S. Schmiedel, I. Bruchhaus, and E. Tannich, "Absence of Erythrocyte Sequestration and Lack of Multicopy Gene Family Expression in Plasmodium falciparum from a Splenectomized Malaria Patient," *PLOS ONE*, vol. 4, p. e7459, 10 2009.

[62] P. H. David, M. Hommel, L. H. Miller, I. J. Udeinya, and L. D. Oligino, "Parasite sequestration in Plasmodium falciparum malaria: spleen and antibody modulation of cytoadherence of infected erythrocytes," *Proceedings of the National Academy of Sciences*, vol. 80, pp. 5075 LP – 5079, 8 1983.

[63] S. Pichyangkul, P. Saengkrai, and H. K. Webster, "Plasmodium falciparum pigment induces monocytes to release high levels of tumor necrosis factor-α and interleukin-1β," *The American journal of tropical medicine and hygiene*, vol. 51, no. 4, pp. 430–435, 1994.

[64] D. Ghigo, R. Todde, H. Ginsburg, C. Costamagna, P. Gautret, F. Bussolino, D. Ulliers, G. Giribaldi, E. Deharo, and G. Gabrielli, "Erythrocyte stages of Plasmodium falciparum exhibit a high nitric oxide synthase (NOS) activity and release an NOS-inducing soluble factor.," *Journal of Experimental Medicine*, vol. 182, no. 3, pp. 677–688, 1995.

[65] L. Schofield and F. Hackett, "Signal transduction in host cells by a glycosylphosphatidylinositol toxin of malaria parasites.," *The Journal of Experimental Medicine*, vol. 177, pp. 145 LP – 153, 1 1993.

[66] M. E. Molyneux, "Impact of malaria on the brain and its prevention," *The Lancet*, vol. 355, no. 9205, pp. 671–672, 2000.

[67] J. Crawley, S. Smith, P. Muthinji, K. Marsh, and F. Kirkham, "Electroencephalographic and clinical features of cerebral malaria," *Archives of Disease in Childhood*, vol. 84, pp. 247 LP – 253, 3 2001.

[68] H. J. Shikani, B. D. Freeman, M. P. Lisanti, L. M. Weiss, H. B. Tanowitz, and M. S. Desruisseaux, "Cerebral Malaria: We Have Come a Long Way," *The American Journal of Pathology*, vol. 181, no. 5, pp. 1484–1492, 2012.

[69] G. L. Birbeck, M. E. Molyneux, P. W. Kaplan, K. B. Seydel, Y. F. Chimalizeni, K. Kawaza, and T. E. Taylor, "Blantyre Malaria Project Epilepsy Study (BMPES) of neurological outcomes in retinopathy-positive paediatric cerebral malaria survivors: a prospective cohort study," *The Lancet Neurology*, vol. 9, no. 12, pp. 1173–1181, 2010.

[70] M. Bernabeu, S. A. Danziger, M. Avril, M. Vaz, P. H. Babar, A. J. Brazier, T. Herricks, J. N. Maki, L. Pereira, A. Mascarenhas, E. Gomes, L. Chery, J. D. Aitchison, P. K. Rathod, and J. D. Smith, "Severe adult malaria is associated with specific PfEMP1 adhesion types and high parasite biomass," *Proceedings of the National Academy of Sciences*, vol. 113, pp. E3270 LP – E3279, 6 2016.

[71] H. Brown, G. Turner, S. Rogerson, M. Tembo, J. Mwenechanya, M. Molyneux, and T. Taylor, "Cytokine Expression in the Brain in Human Cerebral Malaria," *The Journal of Infectious Diseases*, vol. 180, pp. 1742–1746, 11 1999.

[72] K. Dorovini-Zis, K. Schmidt, H. Huynh, W. Fu, R. O. Whitten, D. Milner, S. Kamiza, M. Molyneux, and T. E. Taylor, "The Neuropathology of Fatal Cerebral Malaria in Malawian Children," *The American Journal of Pathology*, vol. 178, no. 5, pp. 2146–2158, 2011.

[73] I. A. Clark and K. A. Rockett, "Points in question Sequestration, cytokines, and malaria pathology," *International Journal for Parasitology*, vol. 24, no. 2, pp. 165–166, 1994.

[74] I. A. Clark and K. A. Rockett, "The cytokine theory of human cerebral malaria," *Parasitology Today*, vol. 10, no. 10, pp. 410–412, 1994.

[75] I. A. Clark and L. M. Alleva, "Is human malarial coma caused, or merely deepened, by sequestration?," *Trends in Parasitology*, vol. 25, no. 7, pp. 314–318, 2009.

[76] J. Dunst, F. Kamena, and K. Matuschewsk, "Cytokines and Chemokines in Cerebral Malaria Pathogenesis," *Front. Cell. Infect. Microbiol.*, vol. 7, 2017.

[77] V. A. White, S. Lewallen, N. Beare, K. Kayira, R. A. Carr, and T. E. Taylor, "Correlation of retinal haemorrhages with brain haemorrhages in children dying of cerebral malaria in Malawi," *Transactions of The Royal Society of Tropical Medicine and Hygiene*, vol. 95, pp. 618–621, 11 2001.

[78] R. J. Maude, N. A. V. Beare, A. A. Sayeed, C. C. Chang, P. Charunwatthana, M. A. Faiz, A. Hossain, E. B. Yunus, M. G. Hoque, M. U. Hasan, N. J. White, N. P. J. Day, and A. M. Dondorp, "The spectrum of retinopathy in adults with Plasmodium falciparum malaria," *Transactions of The Royal Society of Tropical Medicine and Hygiene*, vol. 103, pp. 665–671, 7 2009.

[79] N. A. V. Beare, S. Lewallen, T. E. Taylor, and M. E. Molyneux, "Redefining cerebral malaria by including malaria retinopathy," *Future Microbiology*, vol. 6, pp. 349–355, 3 2011.

[80] M. J. Boivin, M. J. Gladstone, M. Vokhiwa, G. L. Birbeck, J. G. Magen, C. Page, M. Semrud-Clikeman, F. Kauye, and T. E. Taylor, "Developmental outcomes in Malawian children with retinopathy-positive cerebral malaria," *Tropical Medicine & International Health*, vol. 16, pp. 263–271, 3 2011.

[81] B. J. Brabin, Z. Premji, and F. Verhoeff, "An Analysis of Anemia and Child Mortality," *The Journal of Nutrition*, vol. 131, pp. 636S–648S, 2 2001.

[82] M. Fried and P. E. Duffy, "Adherence of Plasmodium falciparum to Chondroitin Sulfate A in the Human Placenta," *Science*, vol. 272, pp. 1502 LP – 1504, 6 1996.

[83] S. COHEN, I. A. McGREGOR, and S. CARRINGTON, "Gamma-Globulin and Acquired Immunity to Human Malaria," *Nature*, vol. 192, no. 4804, pp. 733–737, 1961.

[84] E. Sergent and L. Parrot, "L'immunité, la prémunition et la résistance innée," *Archives de l'institut Pasteur d'Algérie*, vol. 13, pp. 279–319, 1935.

[85] R. Koch, "Dritter Bericht über die Thätigkeit der Malaria-Expedition," *DMW-Deutsche Medizinische Wochenschrift*, vol. 26, no. 17, pp. 281–284, 1900.

[86] S. Gupta, R. W. Snow, C. A. Donnelly, K. Marsh, and C. Newbold, "Immunity to non-cerebral severe malaria is acquired after one or two infections," *Nature Medicine*, vol. 5, no. 3, pp. 340–343, 1999.

[87] A. F. Cowman, J. Healer, D. Marapana, and K. Marsh, "Malaria: Biology and Disease," *Cell*, vol. 167, pp. 610–624, 10 2016.

[88] K. MARSH and S. KINYANJUI, "Immune effector mechanisms in malaria," *Parasite Immunology*, vol. 28, pp. 51–60, 1 2006.

[89] L. S. Mueller and Ivo, "Clinical Immunity to Malaria," 2006.

[90] A. Scherf, J. J. Lopez-Rubio, and L. Riviere, "Antigenic Variation in Plasmodium falciparum," *Annual Review of Microbiology*, vol. 62, pp. 445–470, 9 2008.

[91] W.-H. Tham, J. Healer, and A. F. Cowman, "Erythrocyte and reticulocyte binding-like proteins of Plasmodium falciparum," *Trends in Parasitology*, vol. 28, pp. 23–30, 1 2012.

[92] A. C. Allison, "Protection Afforded by Sickle-cell Trait Against Subtertian Malarial Infection," *British Medical Journal*, vol. 1, pp. 290 LP – 294, 2 1954.

[93] R. Cholera, N. J. Brittain, M. R. Gillrie, T. M. Lopera-Mesa, S. A. S. Diakité, T. Arie, M. A. Krause, A. Guindo, A. Tubman, H. Fujioka, D. A. Diallo, O. K. Doumbo, M. Ho, T. E. Wellems, and R. M. Fairhurst, "Impaired cytoadherence of Plasmodium falciparum-infected erythrocytes containing sickle hemoglobin," *Proceedings of the National Academy of Sciences*, vol. 105, pp. 991 LP – 996, 1 2008.

[94] J. Vega-Rodríguez, B. Franke-Fayard, R. R. Dinglasan, C. J. Janse, R. Pastrana-Mena, A. P. Waters, I. Coppens, J. F. Rodríguez-Orengo, P. Srinivasan, M. Jacobs-Lorena, and A. E. Serrano, "The glutathione biosynthetic pathway of Plasmodium is essential for mosquito transmission," *PLoS pathogens*, vol. 5, pp. e1000302–e1000302, 2 2009.

[95] L. H. Miller, S. J. Mason, D. F. Clyde, and M. H. McGinniss, "The Resistance Factor to Plasmodium vivax in Blacks," *New England Journal of Medicine*, vol. 295, pp. 302–304, 8 1976.

[96] G. Harrison, "Mosquitoes, malaria and man: a history of the hostilities since 1880.," *Mosquitoes, malaria and man: a history of the hostilities since 1880.*, 1978.

[97] U. D'Alessandro, B. O. Olaleye, W. McGuire, M. C. Thomson, P. Langerock, S. Bennett, and B. M. Greenwood, "A comparison of the efficacy of insecticide-treated and untreated bed nets in preventing malaria in Gambian children," *Transactions of The Royal Society of Tropical Medicine and Hygiene*, vol. 89, pp. 596–598, 11 1995.

[98] K. Hargreaves, L. L. Koekemoer, B. D. Brooke, R. H. Hunt, J. Mthembu, and M. Coetzee, "Anopheles funestus resistant to pyrethroid insecticides in South Africa," *Medical and Veterinary Entomology*, vol. 14, pp. 181–189, 6 2000.

[99] D. Martinez-Torres, F. Chandre, M. S. Williamson, F. Darriet, J. B. Bergé, A. L. Devonshire, P. Guillet, N. Pasteur, and D. Pauron, "Molecular characterization of pyrethroid knockdown resistance (kdr) in the major malaria vector Anopheles gambiae s.s.," *Insect Molecular Biology*, vol. 7, pp. 179–184, 5 1998.

[100] C. J. McMeniman, R. V. Lane, B. N. Cass, A. W. C. Fong, M. Sidhu, Y.-F. Wang, and S. L. O'Neill, "Stable Introduction of a Life-Shortening Wolbachia Infection into the Mosquito Aedes aegypti," *Science*, vol. 323, no. 5910, pp. 141–144, 2009.

[101] J. Carlson, E. Suchman, and L. Buchatsky, "Densoviruses for control and genetic manipulation of mosquitoes," *Advances in virus research*, vol. 68, pp. 361–392, 2006.

[102] M. Q. Benedict and A. S. Robinson, "The first releases of transgenic mosquitoes: an argument for the sterile insect technique," *Trends in parasitology*, vol. 19, no. 8, pp. 349–355, 2003.

[103] C. D. Fitch, R. Chevli, H. S. Banyal, G. Phillips, M. A. Pfaller, and D. J. Krogstad, "Lysis of Plasmodium falciparum by ferriprotoporphyrin IX and a chloroquine-ferriprotoporphyrin IX complex.," *Antimicrobial Agents and Chemotherapy*, vol. 21, pp. 819 LP – 822, 5 1982.

[104] R. E. Martin, R. V. Marchetti, A. I. Cowan, S. M. Howitt, S. Bröer, and K. Kirk, "Chloroquine Transport via the Malaria Parasite's Chloroquine Resistance Transporter," *Science*, vol. 325, pp. 1680 LP – 1682, 9 2009.

[105] D. A. Fidock, T. Nomura, A. K. Talley, R. A. Cooper, S. M. Dzekunov, M. T. Ferdig, L. M. B. Ursos, A. bir Singh Sidhu, B. Naudé, K. W. Deitsch, X.-z. Su, J. C. Wootton, P. D. Roepe, and T. E. Wellems, "Mutations in the P. falciparum Digestive Vacuole Transmembrane Protein PfCRT and Evidence for Their Role in Chloroquine Resistance," *Molecular Cell*, vol. 6, pp. 861–871, 10 2000.

[106] J. Wang, C.-J. Zhang, W. N. Chia, C. C. Y. Loh, Z. Li, Y. M. Lee, Y. He, L.-X. Yuan, T. K. Lim, M. Liu, C. X. Liew, Y. Q. Lee, J. Zhang, N. Lu, C. T. Lim, Z.-C. Hua, B. Liu, H.-M. Shen, K. S. W. Tan, and Q. Lin, "Haem-activated promiscuous targeting of artemisinin in Plasmodium falciparum," *Nature Communications*, vol. 6, p. 10111, 12 2015.

[107] N. Klonis, M. P. Crespo-Ortiz, I. Bottova, N. Abu-Bakar, S. Kenny, P. J. Rosenthal, and L. Tilley, "Artemisinin activity against Plasmodium falciparum requires hemoglobin uptake and digestion," *Proceedings of the National Academy of Sciences*, vol. 108, pp. 11405 LP – 11410, 7 2011.

[108] C. L. Hartwig, A. S. Rosenthal, J. D'Angelo, C. E. Griffin, G. H. Posner, and R. A. Cooper, "Accumulation of artemisinin trioxane derivatives within neutral

lipids of Plasmodium falciparum malaria parasites is endoperoxide-dependent," *Biochemical Pharmacology*, vol. 77, no. 3, pp. 322–336, 2009.

[109] H. M. Ismail, V. Barton, M. Phanchana, S. Charoensutthivarakul, M. H. L. Wong, J. Hemingway, G. A. Biagini, P. M. O'Neill, and S. A. Ward, "Artemisinin activity-based probes identify multiple molecular targets within the asexual stage of the malaria parasites Plasmodium falciparum 3D7," *Proceedings of the National Academy of Sciences*, vol. 113, pp. 2080 LP – 2085, 2 2016.

[110] K. Stover, T. King, and J. Robinson, "Artemether-Lumefantrine: An Option for Malaria," *The Annals of pharmacotherapy*, vol. 46, pp. 567–577, 4 2012.

[111] A. M. Dondorp, F. Nosten, P. Yi, D. Das, A. P. Phyo, J. Tarning, K. M. Lwin, F. Ariey, W. Hanpithakpong, S. J. Lee, P. Ringwald, K. Silamut, M. Imwong, K. Chotivanich, P. Lim, T. Herdman, S. S. An, S. Yeung, P. Singhasivanon, N. P. J. Day, N. Lindegardh, D. Socheat, and N. J. White, "Artemisinin Resistance in Plasmodium falciparum Malaria," *New England Journal of Medicine*, vol. 361, pp. 455–467, 7 2009.

[112] S. Picot, P. Olliaro, F. de Monbrison, A.-L. Bienvenu, R. N. Price, and P. Ringwald, "A systematic review and meta-analysis of evidence for correlation between molecular markers of parasite resistance and treatment outcome in falciparum malaria," *Malaria Journal*, vol. 8, no. 1, p. 89, 2009.

[113] Y. Yuthavong, S. Kamchonwongpaisan, U. Leartsakulpanich, and P. Chitnumsub, "Folate metabolism as a source of molecular targets for antimalarials," *Future Microbiology*, vol. 1, pp. 113–125, 6 2006.

[114] J. Yuvaniyama, P. Chitnumsub, S. Kamchonwongpaisan, J. Vanichtanankul, W. Sirawaraporn, P. Taylor, M. D. Walkinshaw, and Y. Yuthavong, "Insights into antifolate resistance from malarial DHFR-TS structures," *Nature Structural & Molecular Biology*, vol. 10, no. 5, pp. 357–365, 2003.

[115] T. Triglia, J. G. T. Menting, C. Wilson, and A. F. Cowman, "Mutations in dihydropteroate synthase are responsible for sulfone and sulfonamide resistance in Plasmodium falciparum," *Proceedings of the National Academy of Sciences*, vol. 94, pp. 13944 LP – 13949, 12 1997.

[116] T. N. C. Wells, R. H. van Huijsduijnen, and W. C. Van Voorhis, "Malaria medicines: a glass half full?," *Nature Reviews Drug Discovery*, vol. 14, p. 424, 5 2015.

[117] C. Arama and M. Troye-Blomberg, "The path of malaria vaccine development: challenges and perspectives," *Journal of Internal Medicine*, vol. 275, pp. 456–466, 5 2014.

[118] C. H. Coelho, J. Y. A. Doritchamou, I. Zaidi, and P. E. Duffy, "Advances in malaria vaccine development: report from the 2017 malaria vaccine symposium," *NPJ vaccines*, vol. 2, p. 34, 11 2017.

[119] R. CARTER and D. H. CHEN, "Malaria transmission blocked by immunisation

with gametes of the malaria parasite," *Nature*, vol. 263, no. 5572, pp. 57–60, 1976.

[120] R. W. Gwadz, "Successful immunization against the sexual stages of Plasmodium gallinaceum," *Science*, vol. 193, pp. 1150 LP – 1151, 9 1976.

[121] K. R. Talaat, R. D. Ellis, J. Hurd, A. Hentrich, E. Gabriel, N. A. Hynes, K. M. Rausch, D. Zhu, O. Muratova, R. Herrera, C. Anderson, D. Jones, J. Aebig, S. Brockley, N. J. MacDonald, X. Wang, M. P. Fay, S. A. Healy, A. P. Durbin, D. L. Narum, Y. Wu, and P. E. Duffy, "Safety and Immunogenicity of Pfs25-EPA/Alhydrogel®, a Transmission Blocking Vaccine against Plasmodium falciparum: An Open Label Study in Malaria Naïve Adults," *PLOS ONE*, vol. 11, p. e0163144, 10 2016.

[122] N. J. MacDonald, V. Nguyen, R. Shimp, K. Reiter, R. Herrera, M. Burkhardt, O. Muratova, K. Kumar, J. Aebig, K. Rausch, L. Lambert, N. Dawson, J. Sattabongkot, X. Ambroggio, P. E. Duffy, Y. Wu, and D. L. Narum, "Structural and Immunological Characterization of Recombinant 6-Cysteine Domains of the Plasmodium falciparum Sexual Stage Protein Pfs230," *Journal of Biological Chemistry*, vol. 291, pp. 19913–19922, 9 2016.

[123] D. F. Clyde, H. Most, V. C. Mccarthy, and J. P. Vanderberg, "Immunization of man against sporozite-induced falciparum malaria," *The American journal of the medical sciences*, vol. 266, no. 3, p. 169, 1973.

[124] "Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: final results of a phase 3, individually randomised, controlled trial," *The Lancet*, vol. 386, no. 9988, pp. 31–45, 2015.

[125] R. O. Payne, K. H. Milne, S. C. Elias, N. J. Edwards, A. D. Douglas, R. E. Brown, S. E. Silk, S. Biswas, K. Miura, R. Roberts, T. W. Rampling, N. Venkatraman, S. H. Hodgson, G. M. Labbé, F. D. Halstead, I. D. Poulton, F. L. Nugent, H. de Graaf, P. Sukhtankar, N. C. Williams, C. F. Ockenhouse, A. K. Kathcart, A. N. Qabar, N. C. Waters, L. A. Soisson, A. J. Birkett, G. S. Cooke, S. N. Faust, C. Woods, K. Ivinson, J. S. McCarthy, C. L. Diggs, J. Vekemans, C. A. Long, A. V. S. Hill, A. M. Lawrie, S. Dutta, and S. J. Draper, "Demonstration of the Blood-Stage Plasmodium falciparum Controlled Human Malaria Infection Model to Assess Efficacy of the P. falciparum Apical Membrane Antigen 1 Vaccine, FMP2.1/AS01," *The Journal of Infectious Diseases*, vol. 213, pp. 1743–1751, 2 2016.

[126] P. Srinivasan, G. C. Baldeviano, K. Miura, A. Diouf, J. A. Ventocilla, K. P. Leiva, L. Lugo-Roman, C. Lucas, S. Orr-Gonzalez, D. Zhu, E. Villasante, L. Soisson, D. L. Narum, S. K. Pierce, C. A. Long, C. Diggs, P. E. Duffy, A. G. Lescano, and L. H. Miller, "A malaria vaccine protects Aotus monkeys against virulent Plasmodium falciparum infection," *npj Vaccines*, vol. 2, no. 1, p. 14, 2017.

[127] A. D. Douglas, A. R. Williams, J. J. Illingworth, G. Kamuyu, S. Biswas, A. L. Goodman, D. H. Wyllie, C. Crosnier, K. Miura, G. J. Wright, C. A.

Long, F. H. Osier, K. Marsh, A. V. Turner, A. V. S. Hill, and S. J. Draper, "The blood-stage malaria antigen PfRH5 is susceptible to vaccine-inducible cross-strain neutralizing antibody," *Nature Communications*, vol. 2, p. 601, 12 2011.

[128] A. D. Douglas, G. C. Baldeviano, C. M. Lucas, L. A. Lugo-Roman, C. Crosnier, S. J. Bartholdson, A. Diouf, K. Miura, L. E. Lambert, J. A. Ventocilla, K. P. Leiva, K. H. Milne, J. J. Illingworth, A. J. Spencer, K. A. Hjerrild, D. G. Alanine, A. V. Turner, J. T. Moorhead, K. A. Edgel, Y. Wu, C. A. Long, G. J. Wright, A. G. Lescano, and S. J. Draper, "A PfRH5-Based Vaccine Is Efficacious against Heterologous Strain Blood-Stage Plasmodium falciparum Infection in Aotus Monkeys," *Cell Host & Microbe*, vol. 17, pp. 130–139, 1 2015.

[129] R. O. Payne, S. E. Silk, S. C. Elias, K. H. Milne, T. A. Rawlinson, D. Llewellyn, A. R. Shakri, J. Jin, G. M. Labbé, N. J. Edwards, I. D. Poulton, R. Roberts, R. Farid, T. Jørgensen, D. G. Alanine, S. C. de Cassan, M. K. Higgins, T. D. Otto, J. S. McCarthy, W. A. de Jongh, A. Nicosia, S. Moyle, A. V. Hill, E. Berrie, C. E. Chitnis, A. M. Lawrie, and S. J. Draper, "Human vaccination against Plasmodium vivax Duffy-binding protein induces strain-transcending antibodies," *JCI insight*, vol. 2, p. e93683, 6 2017.

[130] D. R. Burton, "Antibodies, viruses and vaccines," *Nature Reviews Immunology*, vol. 2, no. 9, pp. 706–713, 2002.

[131] D. R. Burton, P. Poignard, R. L. Stanfield, and I. A. Wilson, "Broadly Neutralizing Antibodies Present New Prospects to Counter Highly Antigenically Diverse Viruses," *Science*, vol. 337, no. 6091, pp. 183–186, 2012.

[132] M. L. Azoitei, B. E. Correia, Y.-E. A. Ban, C. Carrico, O. Kalyuzhniy, L. Chen, A. Schroeter, P.-S. Huang, J. S. Mclellan, P. D. Kwong, D. Baker, R. K. Strong, and W. R. Schief, "Computation-guided backbone grafting of a discontinuous motif onto a protein scaffold," *Science (New York, N.Y.)*, vol. 334, no. 6054, p. 373, 2011.

[133] M. L. Azoitei, Y.-E. A. Ban, J.-P. Julien, S. Bryson, A. Schroeter, O. Kalyuzhniy, J. R. Porter, Y. Adachi, D. Baker, E. F. Pai, and W. R. Schief, "Computational Design of High-Affinity Epitope Scaffolds by Backbone Grafting of a Linear Epitope," *Journal of Molecular Biology*, vol. 415, no. 1, pp. 175–192, 2012.

[134] B. E. Correia, Y.-E. A. Ban, M. A. Holmes, H. Xu, K. Ellingson, Z. Kraft, C. Carrico, E. Boni, D. N. Sather, C. Zenobia, K. Y. Burke, T. Bradley-Hewitt, J. F. Bruhn-Johannsen, O. Kalyuzhniy, D. Baker, R. K. Strong, L. Stamatatos, and W. R. Schief, "Computational Design of Epitope-Scaffolds Allows Induction of Antibodies Specific for a Poorly Immunogenic HIV Vaccine Epitope," *Structure*, vol. 18, no. 9, pp. 1116–1126, 2010.

[135] G. Ofek, F. Guenaga, W. Schief, J. Skinner, D. Baker, R. Wyatt, and P. D.

Kwong, "Elicitation of structure-specific antibodies by epitope scaffolds," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 42, pp. 17880–17887, 2010.

[136] J. S. Mclellan, B. E. Correia, M. Chen, Y. Yang, B. S. Graham, W. R. Schief, and P. D. Kwong, "Design and characterization of epitope-scaffold immunogens that present the motavizumab epitope from respiratory syncytial virus," *Journal of molecular biology*, vol. 409, no. 5, p. 853, 2011.

[137] B. E. Correia, J. T. Bates, R. J. Loomis, G. Baneyx, C. Carrico, J. G. Jardine, P. Rupert, C. Correnti, O. Kalyuzhniy, V. Vittal, M. J. Connell, E. Stevens, A. Schroeter, M. Chen, S. MacPherson, A. M. Serra, Y. Adachi, M. A. Holmes, Y. Li, R. E. Klevit, B. S. Graham, R. T. Wyatt, D. Baker, R. K. Strong, J. E. Crowe, P. R. Johnson, and W. R. Schief, "Proof of principle for epitope-focused vaccine design," *Nature*, vol. 507, p. 201, 2 2014.

[138] D. R. Burton, "What Are the Most Powerful Immunogen Design Vaccine Strategies? Reverse Vaccinology 2.0 Shows Great Promise," *Cold Spring Harbor perspectives in biology*, vol. 9, no. 11, 2017.

[139] D. R. Burton and L. Hangartner, "Broadly Neutralizing Antibodies to HIV and Their Role in Vaccine Design," *Annual Review of Immunology*, vol. 34, pp. 635–659, 5 2016.

[140] J.-P. Julien and H. Wardemann, "Antibodies against Plasmodium falciparum malaria at the molecular level," *Nature Reviews Immunology*, 2019.

[141] S. J. Draper, B. K. Sack, C. R. King, C. M. Nielsen, J. C. Rayner, M. K. Higgins, C. A. Long, and R. A. Seder, "Malaria Vaccines: Recent Advances and New Horizons," *Cell host & microbe*, vol. 24, pp. 43–56, 7 2018.

[142] C. Crosnier, L. Y. Bustamante, S. J. Bartholdson, A. K. Bei, M. Theron, M. Uchikawa, S. Mboup, O. Ndir, D. P. Kwiatkowski, M. T. Duraisingh, J. C. Rayner, and G. J. Wright, "Basigin is a receptor essential for erythrocyte invasion by Plasmodium falciparum," *Nature*, vol. 480, p. 534, 11 2011.

[143] J. Baum, L. Chen, J. Healer, S. Lopaticki, M. Boyle, T. Triglia, F. Ehlgen, S. A. Ralph, J. G. Beeson, and A. F. Cowman, "Reticulocyte-binding protein homologue 5 – An essential adhesin involved in invasion of human erythrocytes by Plasmodium falciparum," *International Journal for Parasitology*, vol. 39, no. 3, pp. 371–380, 2009.

[144] J. Volz, A. Yap, X. Sisquella, J. Thompson, N. Lim, L. Whitehead, L. Chen, M. Lampe, W.-H. Tham, D. Wilson, T. Nebl, D. Marapana, T. Triglia, W. Wong, K. Rogers, and A. Cowman, "Essential Role of the PfRh5/PfRipr/CyRPA Complex during Plasmodium falciparum Invasion of Erythrocytes," *Cell Host & Microbe*, vol. 20, no. 1, pp. 60–71, 2016.

[145] R. O. Payne, S. E. Silk, S. C. Elias, K. Miura, A. Diouf, F. Galaway, H. de Graaf, N. J. Brendish, I. D. Poulton, O. J. Griffiths, N. J. Edwards, J. Jin, G. M. Labbé, D. G. Alanine, L. Siani, S. Di Marco, R. Roberts, N. Green, E. Berrie, A. S. Ishizuka, C. M. Nielsen, M. Bardelli, F. D. Partey, M. F. Ofori, L. Barfod,

J. Wambua, L. M. Murungi, F. H. Osier, S. Biswas, J. S. McCarthy, A. M. Minassian, R. Ashfield, N. K. Viebig, F. L. Nugent, A. D. Douglas, J. Vekemans, G. J. Wright, S. N. Faust, A. V. Hill, C. A. Long, A. M. Lawrie, and S. J. Draper, "Human vaccination against RH5 induces neutralizing antimalarial antibodies that inhibit RH5 invasion complex interactions," *JCI insight*, vol. 2, p. e96381, 11 2017.

[146] R. W. B. T. M. i. E. Woody, "[4] Circular dichroism," in *Biochemical Spectroscopy*, vol. 246, pp. 34–71, Academic Press, 1995.

[147] W. C. Johnson Jr., "Protein secondary structure and circular dichroism: A practical guide," *Proteins: Structure, Function, and Bioinformatics*, vol. 7, pp. 205–214, 1 1990.

[148] R. P. H. Kooyman, R. M. Corn, A. Wark, H. J. Lee, E. Gedig, G. Engbers, L. Walstrom, N. J. de Mol, D. R. Hall, P. Yager, T. Chinowsky, E. Fu, K. Nelson, A. McWhirter, M. J. E. Fischer, A. M. C. Lokate, J. B. Beusink, G. J. M. Pruijn, W. Knoll, A. Kasry, J. Liu, T. Neumann, L. Niu, H. Park, H. Paulsen, R. Robelek, F. Yu, P. Schuck, R. B. M. Schasfoort, and A. J. Tudos, *Handbook of Surface Plasmon Resonance*. The Royal Society of Chemistry, 2008.

[149] G. Winter, "xia2: an expert system for macromolecular crystallography data reduction," *Journal of Applied Crystallography*, vol. 43, pp. 186–190, 2 2010.

[150] M.-F. Incardona, G. P. Bourenkov, K. Levik, R. A. Pieritz, A. N. Popov, and O. Svensson, "EDNA: a framework for plugin-based applications applied to X-ray experiment online data analysis," *Journal of Synchrotron Radiation*, vol. 16, pp. 872–879, 11 2009.

[151] A. G. W. Leslie and H. R. Powell, "Processing diffraction data with MOSFLM," in *Evolving methods for macromolecular crystallography*, vol. 41, pp. 41–51, 7 2007.

[152] A. Leslie, "The integration of macromolecular diffraction data," *Acta Crystallographica Section D*, vol. 62, pp. 48–57, 1 2006.

[153] W. Kabsch, "XDS," *Acta crystallographica. Section D, Biological crystallography*, vol. 66, pp. 125–132, 2 2010.

[154] P. Evans, "Scaling and assessment of data quality," *Acta Crystallographica Section D*, vol. 62, pp. 72–82, 1 2006.

[155] W. Kabsch, "Integration, scaling, space-group assignment and post-refinement," *Acta Crystallographica Section D*, vol. 66, pp. 133–144, 2 2010.

[156] B. W. Matthews, "Solvent content of protein crystals," *Journal of Molecular Biology*, vol. 33, no. 2, pp. 491–497, 1968.

[157] P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan, "Features and development of Coot," *Acta Crystallographica Section D*, vol. 66, pp. 486–501, 4 2010.

[158] G. Bricogne, E. Blanc, M. Brandl, C. Flensburg, P. Keller, W. Paciorek, P. Roversi, A. Sharff, O. S. Smart, C. Vonrhein, and T. O. Womack, "BUSTER version X.Y.Z.," *Cambridge, United Kingdom: Global Phasing Ltd*, 2016.

[159] M. Adrian, J. Dubochet, J. Lepault, and A. W. McDowall, "Cryo-electron microscopy of viruses," *Nature*, vol. 308, no. 5954, pp. 32–36, 1984.

[160] R. S. Pantelic, J. C. Meyer, U. Kaiser, W. Baumeister, and J. M. Plitzko, "Graphene oxide: A substrate for optimizing preparations of frozen-hydrated samples," *Journal of Structural Biology*, vol. 170, pp. 152–156, 4 2010.

[161] A. F. Brilot, J. Z. Chen, A. Cheng, J. Pan, S. C. Harrison, C. S. Potter, B. Carragher, R. Henderson, and N. Grigorieff, "Beam-induced motion of vitrified specimen on holey carbon film," *Journal of Structural Biology*, vol. 177, no. 3, pp. 630–637, 2012.

[162] C. F. Reboul, M. Eager, D. Elmlund, and H. Elmlund, "Single-particle cryo-EM—Improved ab initio 3D reconstruction with SIMPLE/PRIME," *Protein Science*, vol. 27, pp. 51–61, 1 2018.

[163] A. Rohou and N. Grigorieff, "CTFFIND4: Fast and accurate defocus estimation from electron micrographs," *Journal of Structural Biology*, vol. 192, no. 2, pp. 216–221, 2015.

[164] G. Tang, L. Peng, P. R. Baldwin, D. S. Mann, W. Jiang, I. Rees, and S. J. Ludtke, "EMAN2: An extensible image processing suite for electron microscopy," *Journal of Structural Biology*, vol. 157, no. 1, pp. 38–46, 2007.

[165] D. Kimanius, B. Forsberg, and E. Lindahl, "Accelerated Cryo-EM Structure Determination with Parallelisation using GPUs in Relion-2," *Biophysical Journal*, vol. 112, no. 3, pp. 575a–575a, 2017.

[166] D. J. DE ROSIER and A. KLUG, "Reconstruction of Three Dimensional Structures from Electron Micrographs," *Nature*, vol. 217, no. 5124, pp. 130–134, 1968.

[167] R. G. Hart, "Electron Microscopy of Unstained Biological Material: The Polytropic Montage," *Science*, vol. 159, pp. 1464 LP – 1467, 3 1968.

[168] E. V. Orlova and H. R. Saibil, "Structural Analysis of Macromolecular Assemblies by Electron Microscopy," *Chemical Reviews*, vol. 111, pp. 7710–7748, 12 2011.

[169] A. Goldenzweig, M. Goldsmith, S. E. Hill, O. Gertman, P. Laurino, Y. Ashani, O. Dym, T. Unger, S. Albeck, J. Prilusky, R. L. Lieberman, A. Aharoni, I. Silman, J. L. Sussman, D. S. Tawfik, and S. J. Fleishman, "Automated Structure- and Sequence-Based Design of Proteins for High Bacterial Expression and Stability," *Molecular cell*, vol. 70, p. 380, 4 2018.

[170] M. Marti, R. T. Good, M. Rug, E. Knuepfer, and A. F. Cowman, "Targeting Malaria Virulence and Remodeling Proteins to the Host Erythrocyte," *Science*, vol. 306, pp. 1930 LP – 1933, 12 2004.

[171] D. W. Taylor, M. Parra, G. B. Chapman, M. E. Stearns, J. Rener, M. Aikawa, S. Uni, S. B. Aley, L. J. Panton, and R. J. Howard, "Localization of Plasmodium falciparum histidine-rich protein 1 in the erythrocyte skeleton under knobs," *Molecular and Biochemical Parasitology*, vol. 25, no. 2, pp. 165–174, 1987.

[172] S. A. Kyes, Z. Christodoulou, A. Raza, P. Horrocks, R. Pinches, J. A. Rowe, and C. I. Newbold, "A well-conserved Plasmodium falciparum var gene shows an unusual stage-specific transcript pattern," *Molecular Microbiology*, vol. 48, pp. 1339–1348, 6 2003.

[173] C. Magowan, W. Wollish, L. Anderson, and J. Leech, "Cytoadherence by Plasmodium falciparum-infected erythrocytes is correlated with the expression of a family of variable proteins on infected erythrocytes.," *The Journal of Experimental Medicine*, vol. 168, pp. 1307 LP – 1320, 10 1988.

[174] D. I. Baruch, B. L. Pasloske, H. B. Singh, X. Bi, X. C. Ma, M. Feldman, T. F. Taraschi, and R. J. Howard, "Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and adherence receptor on the surface of parasitized human erythrocytes.," *Cell*, vol. 82, pp. 77–87, 7 1995.

[175] S. M. Moghimi, "Mechanisms of splenic clearance of blood cells and particles: towards development of new splenotropic agents," *Advanced Drug Delivery Reviews*, vol. 17, no. 1, pp. 103–115, 1995.

[176] L. H. Miller, D. I. Baruch, K. Marsh, and O. K. Doumbo, "The pathogenic basis of malaria," *Nature*, vol. 415, no. 6872, p. 673, 2002.

[177] S. M. Kraemer and J. D. Smith, "Evidence for the importance of genetic structuring to the structural and functional specialization of the Plasmodium falciparum var gene family," *Molecular Microbiology*, vol. 50, pp. 1527–1538, 12 2003.

[178] A. T. R. Jensen, P. Magistrado, S. Sharp, L. Joergensen, T. Lavstsen, A. Chiucchiuini, A. Salanti, L. S. Vestergaard, J. P. Lusingu, R. Hermsen, R. Sauerwein, J. Christensen, M. A. Nielsen, L. Hviid, C. Sutherland, T. Staalsoe, and T. G. Theander, "Plasmodium falciparum associated with severe childhood malaria preferentially expresses PfEMP1 encoded by group A var genes," *The Journal of experimental medicine*, vol. 199, pp. 1179–1190, 5 2004.

[179] A. Bachmann, S. Predehl, J. May, S. Harder, G. D. Burchard, T.-W. Gilberger, E. Tannich, and I. Bruchhaus, "Highly co-ordinated var gene expression and switching in clinical Plasmodium falciparum isolates from non-immune malaria patients," *Cellular Microbiology*, vol. 13, pp. 1397–1409, 9 2011.

[180] A. Claessens, W. L. Hamilton, M. Kekre, T. D. Otto, A. Faizullabhoy, J. C. Rayner, and D. Kwiatkowski, "Generation of Antigenic Diversity in Plasmodium falciparum by Structured Rearrangement of Var Genes During Mitosis," *PLOS Genetics*, vol. 10, p. e1004812, 12 2014.

[181] X.-z. Su, V. M. Heatwole, S. P. Wertheimer, F. Guinet, J. A. Herrfeldt, D. S. Peterson, J. A. Ravetch, and T. E. Wellems, "The large diverse gene family

var encodes proteins involved in cytoadherence and antigenic variation of Plasmodium falciparum-infected erythrocytes," *Cell*, vol. 82, no. 1, pp. 89–100, 1995.

[182] S. M. Kraemer and J. D. Smith, "A family affair: var genes, PfEMP1 binding, and malaria disease," *Current Opinion in Microbiology*, vol. 9, no. 4, pp. 374–380, 2006.

[183] C. Mayer, L. Slater, M. C. Erat, R. Konrat, and I. Vakonakis, "Structural analysis of the Plasmodium falciparum erythrocyte membrane protein 1 (PfEMP1) intracellular domain reveals a conserved interaction epitope," *The Journal of biological chemistry*, vol. 287, pp. 7182–7189, 3 2012.

[184] T. Lavstsen, A. Salanti, A. Jensen, D. Arnot, and T. Theander, "Sub-grouping of Plasmodium falciparum 3D7 var genes based on sequence analysis of coding and non-coding regions," *Malar. J.*, vol. 2, 2003.

[185] A. K. Dunker, M. S. Cortese, P. Romero, L. M. Iakoucheva, and V. N. Uversky, "Flexible nets," *The FEBS Journal*, vol. 272, pp. 5129–5148, 10 2005.

[186] E. E. Cutts, N. Laasch, D. M. Reiter, R. Trenker, L. M. Slater, P. J. Stansfeld, and I. Vakonakis, "Structural analysis of P. falciparum KAHRP and PfEMP1 complexes with host erythrocyte spectrin suggests a model for cytoadherent knob protrusions," *PLoS Pathogens*, vol. 13, no. 8, p. e1006552, 2017.

[187] J. D. Smith, "The role of PfEMP1 adhesion domain classification in Plasmodium falciparum pathogenesis research," *Molecular and biochemical parasitology*, vol. 195, pp. 82–87, 7 2014.

[188] C. E. Chitnis and L. H. Miller, "Identification of the erythrocyte binding domains of Plasmodium vivax and Plasmodium knowlesi proteins involved in erythrocyte invasion," *The Journal of experimental medicine*, vol. 180, pp. 497–506, 8 1994.

[189] B. K. Sim, C. E. Chitnis, K. Wasniowska, T. J. Hadley, and L. H. Miller, "Receptor and ligand domains for invasion of erythrocytes by Plasmodium falciparum," *Science*, vol. 264, pp. 1941 LP – 1944, 6 1994.

[190] M. M. Klein, A. G. Gittis, H.-P. Su, M. O. Makobongo, J. M. Moore, S. Singh, L. H. Miller, and D. N. Garboczi, "The Cysteine-Rich Interdomain Region from the Highly Variable Plasmodium falciparum Erythrocyte Membrane Protein-1 Exhibits a Conserved Structure," *PLOS Pathogens*, vol. 4, p. e1000147, 9 2008.

[191] J. H. Leech, J. W. Barnwell, L. H. Miller, and R. J. Howard, "Identification of a strain-specific malarial antigen exposed on the surface of Plasmodium falciparum-infected erythrocytes," *The Journal of experimental medicine*, vol. 159, pp. 1567–1575, 6 1984.

[192] J. D. Smith, G. Subramanian, B. Gamain, D. I. Baruch, and L. H. Miller, "Classification of adhesive domains in the Plasmodium falciparum Erythrocyte

Membrane Protein 1 family," *Molecular and Biochemical Parasitology*, vol. 110, no. 2, pp. 293–310, 2000.

[193] S. M. Kraemer, S. A. Kyes, G. Aggarwal, A. L. Springer, S. O. Nelson, Z. Christodoulou, L. M. Smith, W. Wang, E. Levin, C. I. Newbold, P. J. Myler, and J. D. Smith, "Patterns of gene recombination shape var gene repertoires in Plasmodium falciparum: comparisons of geographically diverse isolates," *BMC genomics*, vol. 8, p. 45, 2 2007.

[194] T. S Rask, D. A Hansen, T. Theander, A. Pedersen, and T. Lavstsen, *Plasmodium falciparum Erythrocyte Membrane Protein 1 Diversity in Seven Genomes – Divide and Conquer*, vol. 6. 9 2010.

[195] T. Lavstsen, L. Turner, F. Saguti, P. Magistrado, T. S. Rask, J. S. Jespersen, C. W. Wang, S. S. Berger, V. Baraka, A. M. Marquard, A. Seguin-Orlando, E. Willerslev, M. T. P. Gilbert, J. Lusingu, and T. G. Theander, "Plasmodium falciparum erythrocyte membrane protein 1 domain cassettes 8 and 13 are associated with severe malaria in children," *Proceedings of the National Academy of Sciences*, vol. 109, pp. E1791 LP – E1800, 6 2012.

[196] D. P.-G. Howell, E. A. Levin, A. L. Springer, S. M. Kraemer, D. J. Phippard, W. R. Schief, and J. D. Smith, "Mapping a common interaction site used by Plasmodium falciparum Duffy binding-like domains to bind diverse host receptors," *Molecular Microbiology*, vol. 67, pp. 78–87, 1 2008.

[197] B. A. Robinson, T. L. Welch, and J. D. Smith, "Widespread functional specialization of Plasmodium falciparum erythrocyte membrane protein 1 family members to bind CD36 analysed across a parasite genome," *Molecular Microbiology*, vol. 47, pp. 1265–1278, 3 2003.

[198] L. Turner, T. Lavstsen, S. S. Berger, C. W. Wang, J. E. V. Petersen, M. Avril, A. J. Brazier, J. Freeth, J. S. Jespersen, M. A. Nielsen, P. Magistrado, J. Lusingu, J. D. Smith, M. K. Higgins, and T. G. Theander, "Severe malaria is associated with parasite binding to endothelial protein C receptor," *Nature*, vol. 498, p. 502, 6 2013.

[199] M. Dahlbäck, L. M. Jørgensen, M. A. Nielsen, T. M. Clausen, S. B. Ditlev, M. Resende, V. V. Pinto, D. E. Arnot, T. G. Theander, and A. Salanti, "The chondroitin sulfate A-binding site of the VAR2CSA protein involves multiple N-terminal domains," *The Journal of biological chemistry*, vol. 286, pp. 15908–15917, 5 2011.

[200] F. Lennartz, Y. Adams, A. Bengtsson, R. W. Olsen, L. Turner, N. T. Ndam, G. Ecklu-Mensah, A. Moussiliou, M. F. Ofori, B. Gamain, J. P. Lusingu, J. E. V. Petersen, C. W. Wang, S. Nunes-Silva, J. S. Jespersen, C. K. Y. Lau, T. G. Theander, T. Lavstsen, L. Hviid, M. K. Higgins, and A. T. R. Jensen, "Structure-Guided Identification of a Family of Dual Receptor-Binding PfEMP1 that Is Associated with Cerebral Malaria," *Cell host & microbe*, vol. 21, pp. 403–414, 3 2017.

[201] J. D. Smith, A. G. Craig, N. Kriek, D. Hudson-Taylor, S. Kyes, T. Fagan, R. Pinches, D. I. Baruch, C. I. Newbold, and L. H. Miller, "Identification of a Plasmodium falciparum intercellular adhesion molecule-1 binding domain: a parasite adhesion trait implicated in cerebral malaria," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, pp. 1766–1771, 2 2000.

[202] A. Bengtsson, L. Joergensen, T. S. Rask, R. W. Olsen, M. A. Andersen, L. Turner, T. G. Theander, L. Hviid, M. K. Higgins, A. Craig, A. Brown, and A. T. R. Jensen, "A novel domain cassette identifies Plasmodium falciparum PfEMP1 proteins binding ICAM-1 and is a target of cross-reactive, adhesion-inhibitory antibodies," *Journal of immunology (Baltimore, Md. : 1950)*, vol. 190, pp. 240–249, 1 2013.

[203] I. Naka, J. Patarapotikul, H. Hananantachai, H. Imai, and J. Ohashi, "Association of the endothelial protein C receptor (PROCR) rs867186-G allele with protection from severe malaria," *Malaria Journal*, vol. 13, no. 1, p. 105, 2014.

[204] C. A. Moxon, S. C. Wassmer, D. A. Milner, N. V. Chisala, T. E. Taylor, K. B. Seydel, M. E. Molyneux, B. Faragher, C. T. Esmon, C. Downey, C.-H. Toh, A. G. Craig, and R. S. Heyderman, "Loss of endothelial protein C receptors links coagulation and inflammation to parasite sequestration in cerebral malaria in African children," *Blood*, vol. 122, pp. 842 LP – 851, 8 2013.

[205] C. K. Y. Lau, L. Turner, J. S. Jespersen, E. D. Lowe, B. Petersen, C. W. Wang, J. E. V. Petersen, J. Lusingu, T. G. Theander, T. Lavstsen, and M. K. Higgins, "Structural conservation despite huge sequence diversity allows EPCR binding by the PfEMP1 family implicated in severe childhood malaria," *Cell host & microbe*, vol. 17, pp. 118–129, 1 2015.

[206] L. Zoltan, M. Alexander, T. F. B., F. Gary, and E. C. T., "Human Protein C Receptor Is Present Primarily on Endothelium of Large Blood Vessels ," *Circulation*, vol. 96, pp. 3633–3640, 11 1997.

[207] V. Oganesyan, N. Oganesyan, S. Terzyan, D. Qu, Z. Dauter, N. L. Esmon, and C. T. Esmon, "The Crystal Structure of the Endothelial Protein C Receptor and a Bound Phospholipid," *Journal of Biological Chemistry*, vol. 277, pp. 24851–24854, 7 2002.

[208] M. Riewald, R. J. Petrovan, A. Donner, B. M. Mueller, and W. Ruf, "Activation of Endothelial Cell Protease Activated Receptor 1 by the Protein C Pathway," *Science*, vol. 296, pp. 1880 LP – 1882, 6 2002.

[209] L. O. Mosnier, R. K. Sinha, L. Burnier, E. A. Bouwens, and J. H. Griffin, "Biased agonism of protease-activated receptor 1 by activated protein C caused by noncanonical cleavage at Arg46," *Blood*, vol. 120, pp. 5237 LP – 5246, 12 2012.

[210] U. J. K. Soh and J. Trejo, "Activated protein C promotes protease-activated receptor-1 cytoprotective signaling through β-arrestin and dishevelled-2 scaf-

folds," *Proceedings of the National Academy of Sciences*, vol. 108, no. 50, pp. E1372–E1380, 2011.

[211] J.-S. Bae and A. R. Rezaie, "Protease activated receptor 1 (PAR-1) activation by thrombin is protective in human pulmonary artery endothelial cells if endothelial protein C receptor is occupied by its natural ligand," *Thrombosis and haemostasis*, vol. 100, pp. 101–109, 7 2008.

[212] G. E. Crooks, G. Hon, J.-M. Chandonia, and S. E. Brenner, "WebLogo: a sequence logo generator," *Genome research*, vol. 14, pp. 1188–1190, 6 2004.

[213] A. E. Barry, A. Leliwa-Sytek, L. Tavul, H. Imrie, F. Migot-Nabias, S. M. Brown, G. A. V. McVean, and K. P. Day, "Population genomics of the immune evasion (var) genes of Plasmodium falciparum," *PLoS pathogens*, vol. 3, pp. e34–e34, 3 2007.

[214] P. C. Bull, B. S. Lowe, M. Kortok, C. S. Molyneux, C. I. Newbold, and K. Marsh, "Parasite antigens on the infected red cell surface are targets for naturally acquired immunity to malaria," *Nature medicine*, vol. 4, pp. 358–360, 3 1998.

[215] M. A. Nielsen, T. Staalsoe, J. A. L. Kurtzhals, B. Q. Goka, D. Dodoo, M. Alifrangis, T. G. Theander, B. D. Akanmori, and L. Hviid, "Plasmodium falciparum Variant Surface Antigen Expression Varies Between Isolates Causing Severe and Nonsevere Malaria and Is Modified by Acquired Immunity," *The Journal of Immunology*, vol. 168, pp. 3444 LP – 3450, 4 2002.

[216] S. Maite, K. Don, and H. Hans, "How selection forces dictate the variant surface antigens used by malaria parasites," *Journal of The Royal Society Interface*, vol. 9, pp. 246–260, 2 2012.

[217] F. Sesterhenn, J. Bonet, and B. E. Correia, "Structure-based immunogen design — leading the way to the new age of precision vaccines," *Current Opinion in Structural Biology*, vol. 51, pp. 163–169, 2018.

[218] B. E. Correia, Y.-E. A. Ban, D. J. Friend, K. Ellingson, H. Xu, E. Boni, T. Bradley-Hewitt, J. F. Bruhn-Johannsen, L. Stamatatos, R. K. Strong, and W. R. Schief, "Computational Protein Design Using Flexible Backbone Remodeling and Resurfacing: Case Studies in Structure-Based Antigen Design," *Journal of Molecular Biology*, vol. 405, no. 1, pp. 284–297, 2011.

[219] S. C. Lovell, I. W. Davis, W. B. Arendall III, P. I. W. de Bakker, J. M. Word, M. G. Prisant, J. S. Richardson, and D. C. Richardson, "Structure validation by Cα geometry: ,ψ and Cβ deviation," *Proteins: Structure, Function, and Bioinformatics*, vol. 50, pp. 437–450, 2 2003.

[220] M. F. Bachmann and G. T. Jennings, "Vaccine delivery: a matter of size, geometry, kinetics and molecular patterns," *Nature Reviews Immunology*, vol. 10, p. 787, 10 2010.

[221] B. R. Halbroth and S. J. Draper, "Chapter One - Recent Developments in Malaria Vaccinology," vol. 88, pp. 1–49, Academic Press, 2015.

[222] F. Zabel, T. M. Kündig, and M. F. Bachmann, "Virus-induced humoral immunity: on how B cell responses are initiated," *Current Opinion in Virology*, vol. 3, no. 3, pp. 357–362, 2013.

[223] A. Roldão, M. C. M. Mellado, L. R. Castilho, M. J. T. Carrondo, and P. M. Alves, "Virus-like particles in vaccine development," *Expert Review of Vaccines*, vol. 9, pp. 1149–1176, 10 2010.

[224] K. D. Brune, D. B. Leneghan, I. J. Brian, A. S. Ishizuka, M. F. Bachmann, S. J. Draper, S. Biswas, and M. Howarth, "Plug-and-Display: decoration of Virus-Like Particles via isopeptide bonds for modular immunization," *Scientific reports*, vol. 6, p. 19234, 1 2016.

[225] S. Thrane, C. M. Janitzek, S. Matondo, M. Resende, T. Gustavsson, W. A. de Jongh, S. Clemmensen, W. Roeffen, M. van de Vegte-Bolmer, G. J. van Gemert, R. Sauerwein, J. T. Schiller, M. A. Nielsen, T. G. Theander, A. Salanti, and A. F. Sander, "Bacterial superglue enables easy development of efficient virus-like particle based vaccines," *Journal of Nanobiotechnology*, vol. 14, no. 1, p. 30, 2016.

[226] N. M. Barber, C. K. Y. Lau, L. Turner, G. Watson, S. Thrane, J. P. A. Lusingu, T. Lavstsen, and M. K. Higgins, "Structure-guided design of a synthetic mimic of an EPCR-binding PfEMP1 protein," *bioRxiv*, p. 749432, 1 2019.

[227] F.-L. Hsieh, L. Turner, J. R. Bolla, C. V. Robinson, T. Lavstsen, and M. K. Higgins, "The structural basis for CD36 binding by the malaria parasite," *Nature Communications*, vol. 7, p. 12837, 9 2016.

[228] D. Gaur, D. C. G. Mayer, and L. H. Miller, "Parasite ligand–host receptor interactions during invasion of erythrocytes by Plasmodium merozoites," *International Journal for Parasitology*, vol. 34, no. 13, pp. 1413–1429, 2004.

[229] J. D. HAYNES, C. L. DIGGS, F. A. HINES, and R. E. DESJARDINS, "Culture of human malaria parasites Plasmodium falciparum," *Nature*, vol. 263, no. 5580, pp. 767–769, 1976.

[230] W. Trager and J. B. Jensen, "Human malaria parasites in continuous culture," *Science*, vol. 193, pp. 673 LP – 675, 8 1976.

[231] L. H. Miller, S. J. Mason, J. A. Dvorak, M. H. McGinniss, and I. K. Rothman, "Erythrocyte receptors for (Plasmodium knowlesi) malaria: Duffy blood group determinants," *Science*, vol. 189, 1975.

[232] R. Horuk, C. E. Chitnis, W. C. Darbonne, T. J. Colby, A. Rybicki, T. J. Hadley, and L. H. Miller, "A receptor for the malarial parasite Plasmodium vivax: the erythrocyte chemokine receptor," *Science*, vol. 261, pp. 1182 LP – 1184, 8 1993.

[233] S. P. Wertheimer and J. W. Barnwell, "Plasmodium vivax interaction with the human Duffy blood group glycoprotein: Identification of a parasite receptor-like protein," *Experimental Parasitology*, vol. 69, pp. 340–350, 10 1989.

[234] K. M. VanBuskirk, E. Sevova, and J. H. Adams, "Conserved residues in the Plasmodium vivax Duffy-binding protein ligand domain are critical for erythrocyte receptor recognition," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, pp. 15754 LP – 15759, 11 2004.

[235] A. P. Singh, H. Ozwara, C. H. M. Kocken, S. K. Puri, A. W. Thomas, and C. E. Chitnis, "Targeted deletion of Plasmodium knowlesi Duffy binding protein confirms its role in junction formation during invasion," *Molecular Microbiology*, vol. 55, pp. 1925–1934, 3 2005.

[236] B. Malleret, F. Xu, N. Mohandas, R. Suwanarusk, C. Chu, J. A. Leite, K. Low, C. Turner, K. Sriprawat, R. Zhang, O. Bertrand, Y. Colin, F. T. M. Costa, C. N. Ong, M. L. Ng, C. T. Lim, F. Nosten, L. Rénia, and B. Russell, "Significant Biochemical, Biophysical and Metabolic Diversity in Circulating Human Cord Blood Reticulocytes," *PLOS ONE*, vol. 8, p. e76062, 10 2013.

[237] J. Gruszczyk, U. Kanjee, L.-J. Chan, S. Menant, B. Malleret, N. T. Y. Lim, C. Q. Schmidt, Y.-F. Mok, K.-M. Lin, R. D. Pearson, G. Rangel, B. J. Smith, M. J. Call, M. P. Weekes, M. D. W. Griffin, J. M. Murphy, J. Abraham, K. Sriprawat, M. J. Menezes, M. U. Ferreira, B. Russell, L. Renia, M. T. Duraisingh, and W.-H. Tham, "Transferrin receptor 1 is a reticulocyte-specific receptor for Plasmodium vivax," *Science*, vol. 359, pp. 48 LP – 55, 1 2018.

[238] J. R. Ryan, J. A. Stoute, J. Amon, R. F. Dunton, R. Mtalib, J. Koros, B. Owour, S. Luckhart, R. A. Wirtz, J. W. Barnwell, and R. Rosenberg, "Evidence for transmission of Plasmodium vivax among a duffy antigen negative population in Western Kenya," *Am J Trop Med Hyg*, vol. 75, 2006.

[239] C. E. Cavasini, L. C. Mattos, A. A. Couto, C. R. Bonini-Domingos, S. H. Valencia, W. C. Neiras, R. T. Alves, A. R. Rossit, L. Castilho, and R. L. Machado, "Plasmodium vivax infection among Duffy antigen-negative individuals from the Brazilian Amazon region: an exception?," *Trans R Soc Trop Med Hyg*, vol. 101, 2007.

[240] D. Menard, E. R. Chan, C. Benedet, A. Ratsimbasoa, S. Kim, P. Chim, C. Do, B. Witkowski, R. Durand, M. Thellier, C. Severini, E. Legrand, L. Musset, B. Y. M. Nour, O. Mercereau-Puijalon, D. Serre, and P. A. Zimmerman, "Whole Genome Sequencing of Field Isolates Reveals a Common Duplication of the Duffy Binding Protein Gene in Malagasy Plasmodium vivax Strains," *PLOS Neglected Tropical Diseases*, vol. 7, p. e2489, 11 2013.

[241] K. Gunalan, E. Lo, J. B. Hostetler, D. Yewhalaw, J. Mu, D. E. Neafsey, G. Yan, and L. H. Miller, "Role of Plasmodium vivax Duffy-binding protein 1 in invasion of Duffy-null Africans," *Proceedings of the National Academy of Sciences*, vol. 113, pp. 6271 LP – 6276, 5 2016.

[242] J. H. Adams, D. E. Hudson, M. Torii, G. E. Ward, T. E. Wellems, M. Aikawa, and L. H. Miller, "The duffy receptor family of plasmodium knowlesi is located within the micronemes of invasive malaria merozoites," *Cell*, vol. 63, no. 1, pp. 141–153, 1990.

[243] J. H. Adams, B. K. Sim, S. A. Dolan, X. Fang, D. C. Kaslow, and L. H. Miller, "A family of erythrocyte binding proteins of malaria parasites," *Proceedings of the National Academy of Sciences*, vol. 89, pp. 7085 LP – 7089, 8 1992.

[244] N. H. Tolia, E. J. Enemark, B. K. L. Sim, and L. Joshua-Tor, "Structural Basis for the EBA-175 Erythrocyte Invasion Pathway of the Malaria Parasite Plasmodium falciparum," *Cell*, vol. 122, pp. 183–193, 7 2005.

[245] J. D. Batchelor, J. A. Zahm, and N. H. Tolia, "Dimerization of Plasmodium vivax DBP is induced upon receptor binding and drives recognition of DARC," *Nature Structural &Amp; Molecular Biology*, vol. 18, p. 908, 7 2011.

[246] S. K. Singh, R. Hora, H. Belrhali, C. E. Chitnis, and A. Sharma, "Structural basis for Duffy recognition by the malaria parasite Duffy-binding-like domain," *Nature*, vol. 439, no. 7077, pp. 741–744, 2006.

[247] C. E. Chitnis, A. Chaudhuri, R. Horuk, A. O. Pogo, and L. H. Miller, "The domain on the Duffy blood group antigen for binding Plasmodium vivax and P. knowlesi malarial parasites to erythrocytes.," *The Journal of Experimental Medicine*, vol. 184, pp. 1531 LP – 1536, 10 1996.

[248] C. Tournamille, A. Filipe, K. Wasniowska, P. Gane, E. Lisowska, J.-P. Cartron, Y. Colin, and C. Le Van Kim, "Structure–function analysis of the extracellular domains of the Duffy antigen/receptor for chemokines: characterization of antibody and chemokine binding sites," *British Journal of Haematology*, vol. 122, pp. 1014–1023, 9 2003.

[249] S. C. Peiper, Z. X. Wang, K. Neote, A. W. Martin, H. J. Showell, M. J. Conklyn, K. Ogborne, T. J. Hadley, Z. H. Lu, J. Hesselgesser, and R. Horuk, "The Duffy antigen/receptor for chemokines (DARC) is expressed in endothelial cells of Duffy negative individuals who lack the erythrocyte receptor.," *The Journal of Experimental Medicine*, vol. 181, pp. 1311 LP – 1317, 4 1995.

[250] M. Pruenster, L. Mudde, P. Bombosi, S. Dimitrova, M. Zsak, J. Middleton, A. Richmond, G. J. Graham, S. Segerer, R. J. B. Nibbs, and A. Rot, "The Duffy antigen receptor for chemokines transports chemokines and supports their promigratory activity," *Nature immunology*, vol. 10, pp. 101–108, 1 2009.

[251] J. D. Batchelor, B. M. Malpede, N. S. Omattage, G. T. DeKoster, K. A. Henzler-Wildman, and N. H. Tolia, "Red blood cell invasion by Plasmodium vivax: structural basis for DBP engagement of DARC," *PLoS pathogens*, vol. 10, pp. e1003869–e1003869, 1 2014.

[252] H. Choe, M. J. Moore, C. M. Owens, P. L. Wright, N. Vasilieva, W. Li, A. P. Singh, R. Shakri, C. E. Chitnis, and M. Farzan, "Sulphated tyrosines mediate association of chemokines and Plasmodium vivax Duffy binding protein with

the Duffy antigen/receptor for chemokines (DARC)," *Molecular Microbiology*, vol. 55, pp. 1413–1422, 3 2005.

[253] D. Hans, P. Pattnaik, A. Bhattacharyya, A. R. Shakri, S. S. Yazdani, M. Sharma, H. Choe, M. Farzan, and C. E. Chitnis, "Mapping binding residues in the Plasmodium vivax domain that binds Duffy antigen during red cell invasion," *Molecular Microbiology*, vol. 55, pp. 1423–1434, 3 2005.

[254] R. R. Race and R. Sanger, "Blood Groups in Man, 6TH ED," *Immunological Communications*, vol. 9, pp. 419–420, 1 1980.

[255] C. Tournamille, Y. Colin, J. P. Cartron, and C. Le Van Kim, "Disruption of a GATA motif in the Duffy gene promoter abolishes erythroid gene expression in Duffy-negative individuals," *Nat Genet*, vol. 10, 1995.

[256] K. Wasniowska and T. J. Hadley, "The Duffy Blood Group Antigen: An Update," *Transfusion Medicine Reviews*, vol. 8, no. 4, pp. 281–288, 1994.

[257] A. Chaudhuri, J. Polyakova, V. Zbrzezna, and A. O. Pogo, "The coding sequence of Duffy blood group gene in humans and simians: restriction fragment length polymorphism, antibody and malarial parasite specificities, and expression in nonerythroid tissues in Duffy- negative individuals," *Blood*, vol. 85, pp. 615 LP – 621, 2 1995.

[258] C. L. King, P. Michon, A. R. Shakri, A. Marcotty, D. Stanisic, P. A. Zimmerman, J. L. Cole-Tobian, I. Mueller, and C. E. Chitnis, "Naturally acquired Duffy-binding protein-specific binding inhibitory antibodies confer protection from blood-stage Plasmodium vivax infection," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, pp. 8363–8368, 6 2008.

[259] P. Michon, T. Fraser, and J. H. Adams, "Naturally acquired and vaccine-elicited antibodies block erythrocyte cytoadherence of the Plasmodium vivax Duffy binding protein," *Infection and immunity*, vol. 68, pp. 3164–3171, 6 2000.

[260] B. T. Grimberg, R. Udomsangpetch, J. Xainli, A. McHenry, T. Panichakul, J. Sattabongkot, L. Cui, M. Bockarie, C. Chitnis, J. Adams, P. A. Zimmerman, and C. L. King, "Plasmodium vivax Invasion of Human Erythrocytes Inhibited by Antibodies Directed against the Duffy Binding Protein," *PLOS Medicine*, vol. 4, p. e337, 12 2007.

[261] S. C. de Cassan, A. R. Shakri, D. Llewellyn, S. C. Elias, J. S. Cho, A. L. Goodman, J. Jin, A. D. Douglas, R. Suwanarusk, F. H. Nosten, L. Rénia, B. Russell, C. E. Chitnis, and S. J. Draper, "Preclinical Assessment of Viral Vectored and Protein Vaccines Targeting the Duffy-Binding Protein Region II of Plasmodium Vivax," *Frontiers in immunology*, vol. 6, p. 348, 7 2015.

[262] A. Moreno, I. Caro-Aguilar, S. S. Yazdani, A. R. Shakri, S. Lapp, E. Strobert, H. McClure, C. E. Chitnis, and M. R. Galinski, "Preclinical assessment of the receptor-binding domain of Plasmodium vivax Duffy-binding protein as a vaccine candidate in rhesus macaques," *Vaccine*, vol. 26, pp. 4338–4344, 8 2008.

[263] P. Chootong, F. B. Ntumngia, K. M. VanBuskirk, J. Xainli, J. L. Cole-Tobian, C. O. Campbell, T. S. Fraser, C. L. King, and J. H. Adams, "Mapping epitopes of the Plasmodium vivax Duffy binding protein with naturally acquired inhibitory antibodies," *Infection and immunity*, vol. 78, pp. 1089–1095, 3 2010.

[264] E. Chen, N. D. Salinas, Y. Huang, F. Ntumngia, M. D. Plasencia, M. L. Gross, J. H. Adams, and N. H. Tolia, "Broadly neutralizing epitopes in the Plasmodium vivax vaccine candidate Duffy Binding Protein," *Proceedings of the National Academy of Sciences*, vol. 113, pp. 6277–6282, 5 2016.

[265] L. M. Walker and D. R. Burton, "Rational antibody-based HIV-1 vaccine design: current approaches and future directions," *Current opinion in immunology*, vol. 22, pp. 358–366, 6 2010.

[266] B. F. Haynes, G. Kelsoe, S. C. Harrison, and T. B. Kepler, "B-cell-lineage immunogen design in vaccine development with HIV-1 as a case study," *Nature biotechnology*, vol. 30, pp. 423–433, 5 2012.

[267] F. Klein, H. Mouquet, P. Dosenovic, J. F. Scheid, L. Scharf, and M. C. Nussenzweig, "Antibodies in HIV-1 vaccine development and therapy," *Science (New York, N.Y.)*, vol. 341, pp. 1199–1204, 9 2013.

[268] T. A. Rawlinson, N. M. Barber, F. Mohring, J. S. Cho, V. Kosaisavee, S. F. Gérard, D. G. W. Alanine, G. M. Labbé, S. C. Elias, S. E. Silk, D. Quinkert, J. Jin, J. M. Marshall, R. O. Payne, A. M. Minassian, B. Russell, L. Rénia, F. H. Nosten, R. W. Moon, M. K. Higgins, and S. J. Draper, "Structural basis for inhibition of Plasmodium vivax invasion by a broadly neutralizing vaccine-induced human antibody," *Nature Microbiology*, 2019.

[269] F. Mohring, M. N. Hart, T. A. Rawlinson, R. Henrici, J. A. Charleston, E. Diez Benavente, A. Patel, J. Hall, N. Almond, S. Campino, T. G. Clark, C. J. Sutherland, D. A. Baker, S. J. Draper, and R. W. Moon, "Rapid and iterative genome editing in the malaria parasite Plasmodium knowlesi provides new tools for P. vivax research," *eLife*, vol. 8, p. e45829, 6 2019.

[270] C. E. Chitnis and A. Sharma, "Targeting the Plasmodium vivax Duffy-binding protein," *Trends in Parasitology*, vol. 24, pp. 29–34, 1 2008.

[271] D. Urusova, L. Carias, Y. Huang, V. C. Nicolete, J. Popovici, C. Roesch, N. D. Salinas, B. Witkowski, M. U. Ferreira, J. H. Adams, M. L. Gross, C. L. King, and N. H. Tolia, "Structural basis for neutralization of Plasmodium vivax by naturally acquired human antibodies that target DBP," *Nature Microbiology*, 2019.

[272] I. Campeotto, A. Goldenzweig, J. Davey, L. Barfod, J. M. Marshall, S. E. Silk, K. E. Wright, S. J. Draper, M. K. Higgins, and S. J. Fleishman, "One-step design of a stable variant of the malaria invasion protein RH5 for use as a vaccine immunogen," *Proceedings of the National Academy of Sciences*, vol. 114, pp. 998 LP – 1002, 1 2017.

[273] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic

local alignment search tool," *Journal of Molecular Biology*, vol. 215, no. 3, pp. 403–410, 1990.

[274] R. C. Edgar, "MUSCLE: multiple sequence alignment with high accuracy and high throughput," *Nucleic Acids Research*, vol. 32, pp. 1792–1797, 3 2004.

[275] T. A. Whitehead, A. Chevalier, Y. Song, C. Dreyfus, S. J. Fleishman, C. De Mattos, C. A. Myers, H. Kamisetty, P. Blair, I. A. Wilson, and D. Baker, "Optimization of affinity, specificity and function of designed influenza inhibitors using deep sequencing," *Nature Biotechnology*, vol. 30, p. 543, 5 2012.

[276] A. C. Gomes, M. Mohsen, and M. F. Bachmann, "Harnessing Nanoparticles for Immunomodulation and Vaccines," *Vaccines*, vol. 5, p. 6, 2 2017.

[277] B. J. Brabin, C. Romagosa, S. Abdelgalil, C. Menendez, F. H. Verhoeff, R. McGready, K. A. Fletcher, S. Owens, U. d'Alessandro, and F. Nosten, "The sick placenta—the role of malaria," *Placenta*, vol. 25, no. 5, pp. 359–378, 2004.

[278] B. BRABIN, "An Assessment of Low Birthweight Risk in Primiparae as an Indicator of Malaria Control in Pregnancy," *International Journal of Epidemiology*, vol. 20, pp. 276–283, 3 1991.

[279] S. le Cessie, F. H. Verhoeff, G. Mengistie, P. Kazembe, R. Broadhead, and B. J. Brabin, "Changes in haemoglobin levels in infants in Malawi: effect of low birth weight and fetal anaemia," *Archives of Disease in Childhood - Fetal and Neonatal Edition*, vol. 86, pp. F182 LP – F187, 5 2002.

[280] D. Radeva-Petrova, K. Kayentao, F. O. ter Kuile, D. Sinclair, and P. Garner, "Drugs for preventing malaria in pregnant women in endemic areas: any drug regimen versus placebo or no treatment," *Cochrane Database of Systematic Reviews*, no. 10, 2014.

[281] M. Parise, B. Nahlen, C. Menendez, and R. Steketee, "The burden of malaria in pregnancy in malaria-endemic areas," *The American Journal of Tropical Medicine and Hygiene*, vol. 64, pp. 28–35, 1 2001.

[282] WHO, *Control and elimination of Plasmodium vivax malaria – A technical brief*. 2015.

[283] A. Salanti, T. Staalsoe, T. Lavstsen, A. T. R. Jensen, M. P. K. Sowa, D. E. Arnot, L. Hviid, and T. G. Theander, "Selective upregulation of a single distinctly structured var gene in chondroitin sulphate A-adhering Plasmodium falciparum involved in pregnancy-associated malaria," *Molecular Microbiology*, vol. 49, pp. 179–191, 7 2003.

[284] A. Salanti, M. Dahlbäck, L. Turner, M. A. Nielsen, L. Barfod, P. Magistrado, A. T. R. Jensen, T. Lavstsen, M. F. Ofori, K. Marsh, L. Hviid, and T. G. Theander, "Evidence for the Involvement of VAR2CSA in Pregnancy-associated Malaria," *The Journal of Experimental Medicine*, vol. 200, pp. 1197 LP – 1203, 11 2004.

[285] N. G. Tuikue Ndam, A. Salanti, G. Bertin, M. Dahlbäck, N. Fievet, L. Turner, A. Gaye, T. Theander, and P. Deloron, "High Level of var2csa Transcription by Plasmodium falciparum Isolated from the Placenta," *The Journal of Infectious Diseases*, vol. 192, pp. 331–335, 7 2005.

[286] N. K. Viebig, B. Gamain, C. Scheidig, C. Lépolard, J. Przyborski, M. Lanzer, J. Gysin, and A. Scherf, "A single member of the Plasmodium falciparum var multigene family determines cytoadhesion to the placental receptor chondroitin sulphate A," *EMBO reports*, vol. 6, pp. 775–781, 8 2005.

[287] N. K. Viebig, E. Levin, S. Dechavanne, S. J. Rogerson, J. Gysin, J. D. Smith, A. Scherf, and B. Gamain, "Disruption of Var2csa Gene Impairs Placental Malaria Associated Adhesion Phenotype," *PLOS ONE*, vol. 2, p. e910, 9 2007.

[288] T. Staalsoe, C. E. Shulman, J. N. Bulmer, K. Kawuondo, K. Marsh, and L. Hviid, "Variant surface antigen-specific IgG and protection against clinical consequences of pregnancy-associated Plasmodium falciparum malaria," *The Lancet*, vol. 363, pp. 283–289, 1 2004.

[289] P. E. Duffy, "Maternal immunization and malaria in pregnancy," *Vaccine*, vol. 21, no. 24, pp. 3358–3361, 2003.

[290] A. Alkhalil, R. N. Achur, M. Valiyaveettil, C. F. Ockenhouse, and D. C. Gowda, "GLYCOBIOLOGY AND EXTRACELLULAR MATRICES-Structural requirements for the adherence of Plasmodium falciparum-infected erythrocytes to chondroitin sulfate proteoglycans of human placenta.," *Journal of Biological Chemistry*, vol. 275, no. 51, pp. 40357–40364, 2000.

[291] N. Bir, S. S. Yazdani, M. Avril, C. Layez, J. Gysin, and C. E. Chitnis, "Immunogenicity of Duffy Binding-Like Domains That Bind Chondroitin Sulfate A and Protection against Pregnancy-Associated Malaria," *Infection and Immunity*, vol. 74, pp. 5955 LP – 5963, 10 2006.

[292] M. Avril, B. Gamain, C. Lépolard, N. Viaud, A. Scherf, and J. Gysin, "Characterization of anti-var2CSA-PfEMP1 cytoadhesion inhibitory mouse monoclonal antibodies," *Microbes and Infection*, vol. 8, no. 14, pp. 2863–2871, 2006.

[293] B. Gamain, A. R. Trimnell, C. Scheidig, A. Scherf, L. H. Miller, and J. D. Smith, "Identification of Multiple Chondroitin Sulfate A (CSA)–Binding Domains in the var2CSA Gene Transcribed in CSA-Binding Parasites," *The Journal of Infectious Diseases*, vol. 191, pp. 1010–1013, 3 2005.

[294] M. Resende, S. B. Ditlev, M. A. Nielsen, S. Bodevin, S. Bruun, V. V. Pinto, H. Clausen, L. Turner, T. G. Theander, A. Salanti, and M. Dahlbäck, "Chondroitin sulphate A (CSA)-binding of single recombinant Duffy-binding-like domains is not restricted to Plasmodium falciparum Erythrocyte Membrane Protein 1 expressed by CSA-binding parasites," *International Journal for Parasitology*, vol. 39, no. 11, pp. 1195–1204, 2009.

[295] A. Srivastava, S. Gangnard, A. Round, S. Dechavanne, A. Juillerat, B. Raynal, G. Faure, B. Baron, S. Ramboarina, S. K. Singh, H. Belrhali, P. England,

A. Lewit-Bentley, A. Scherf, G. A. Bentley, and B. Gamain, "Full-length extracellular region of the var2CSA variant of PfEMP1 is required for specific, high-affinity binding to CSA," *Proceedings of the National Academy of Sciences*, vol. 107, pp. 4884 LP – 4889, 3 2010.

[296] T. M. Clausen, S. Christoffersen, M. Dahlbäck, A. E. Langkilde, K. E. Jensen, M. Resende, M. Agerbæk, D. Andersen, B. Berisha, S. B. Ditlev, V. V. Pinto, M. A. Nielsen, T. G. Theander, S. Larsen, and A. Salanti, "Structural and functional insight into how the Plasmodium falciparum VAR2CSA protein mediates binding to chondroitin sulfate A in placental malaria," *The Journal of biological chemistry*, vol. 287, pp. 23332–23345, 7 2012.

[297] P. Khunrae, M. Dahlbäck, M. A. Nielsen, G. Andersen, S. B. Ditlev, M. Resende, V. V. Pinto, T. G. Theander, M. K. Higgins, and A. Salanti, "Full-Length Recombinant Plasmodium falciparum VAR2CSA Binds Specifically to CSPG and Induces Potent Parasite Adhesion-Blocking Antibodies," *Journal of Molecular Biology*, vol. 397, no. 3, pp. 826–834, 2010.

[298] A. Srivastava, S. Gangnard, S. Dechavanne, F. Amirat, A. Lewit Bentley, G. A. Bentley, and B. Gamain, "Var2CSA Minimal CSA Binding Region Is Located within the N-Terminal Region," *PLOS ONE*, vol. 6, p. e20270, 5 2011.

[299] S. Gangnard, N. G. Tuikue Ndam, S. Gnidehou, M. Quiviger, A. Juillerat, G. Faure, B. Baron, F. Viwami, P. Deloron, and G. A. Bentley, "Functional and immunological characterization of the var2CSA-DBL5 domain of a placental Plasmodium falciparum isolate," *Molecular and Biochemical Parasitology*, vol. 173, no. 2, pp. 115–122, 2010.

[300] R. N. Achur, M. Valiyaveettil, A. Alkhalil, C. F. Ockenhouse, and D. C. Gowda, "Characterization of proteoglycans of human placenta and identification of unique chondroitin sulfate proteoglycans of the intervillous spaces that mediate the adherence of Plasmodium falciparum-infected erythrocytes to the placenta.," *The Journal of biological chemistry*, vol. 275, pp. 40344–56, 12 2000.

[301] M. K Higgins, "The Structure of a Chondroitin Sulfate-binding Domain Important in Placental Malaria*S," *The Journal of biological chemistry*, vol. 283, pp. 21842–21846, 7 2008.

[302] K. Singh, A. G. Gittis, P. Nguyen, D. C. Gowda, L. H. Miller, and D. N. Garboczi, "Structure of the DBL3x domain of pregnancy-associated malaria protein VAR2CSA complexed with chondroitin sulfate A," *Nature Structural &Amp; Molecular Biology*, vol. 15, p. 932, 8 2008.

[303] S. Gangnard, A. Lewit-Bentley, S. Dechavanne, A. Srivastava, F. Amirat, G. A. Bentley, and B. Gamain, "Structure of the DBL3X-DBL4ϵ region of the VAR2CSA placental malaria vaccine candidate: insight into DBL domain interactions," *Scientific Reports*, vol. 5, p. 14868, 10 2015.

[304] L. Barfod, N. L. Bernasconi, M. Dahlbäck, D. Jarrossay, P. H. Andersen, A. Salanti, M. F. Ofori, L. Turner, M. Resende, M. A. Nielsen, T. G.

Theander, F. Sallusto, A. Lanzavecchia, and L. Hviid, "Human pregnancy-associated malaria-specific B cells target polymorphic, conformational epitopes in VAR2CSA: This article became available OnlineOpen after it was first published online on 14 December 2006 [14 February 2007].," *Molecular Microbiology*, vol. 63, no. 2, pp. 335–347, 2007.

[305] M. Dahlbäck, T. S. Rask, P. H. Andersen, M. A. Nielsen, N. T. Ndam, M. Resende, L. Turner, P. Deloron, L. Hviid, O. Lund, A. G. Pedersen, T. G. Theander, and A. Salanti, "Epitope Mapping and Topographic Analysis of VAR2CSA DBL3X Involved in P. falciparum Placental Sequestration," *PLOS Pathogens*, vol. 2, p. e124, 11 2006.

[306] A. R. Trimnell, S. M. Kraemer, S. Mukherjee, D. J. Phippard, J. H. Janes, E. Flamoe, X.-z. Su, P. Awadalla, and J. D. Smith, "Global genetic diversity and evolution of var genes associated with placental and severe childhood malaria," *Molecular and Biochemical Parasitology*, vol. 148, no. 2, pp. 169–180, 2006.

[307] J. Bockhorst, F. Lu, J. H. Janes, J. Keebler, B. Gamain, P. Awadalla, X.-z. Su, R. Samudrala, N. Jojic, and J. D. Smith, "Structural polymorphism and diversifying selection on the pregnancy malaria vaccine candidate VAR2CSA," *Molecular and Biochemical Parasitology*, vol. 155, no. 2, pp. 103–112, 2007.

[308] M. F. Duffy, A. G. Maier, T. J. Byrne, A. J. Marty, S. R. Elliott, M. T. O'Neill, P. D. Payne, S. J. Rogerson, A. F. Cowman, B. S. Crabb, and G. V. Brown, "VAR2CSA is the principal ligand for chondroitin sulfate A in two allogeneic isolates of Plasmodium falciparum," *Molecular and Biochemical Parasitology*, vol. 148, no. 2, pp. 117–124, 2006.

[309] P. Fried, M & Duffy, "Maternal malaria and parasite adhesion," *J Mol Med*, vol. 76, no. 3-4, pp. 162–171, 1998.

[310] C. H. Ricke, T. Staalsoe, K. Koram, B. D. Akanmori, E. M. Riley, T. G. Theander, and L. Hviid, "Plasma Antibodies from Malaria-Exposed Pregnant Women Recognize Variant Surface Antigens on Plasmodium falciparum-Infected Erythrocytes in a Parity-Dependent Manner and Block Parasite Adhesion to Chondroitin Sulfate A," *The Journal of Immunology*, vol. 165, pp. 3309 LP – 3316, 9 2000.

[311] L. Barfod, T. Dobrilovic, P. Magistrado, P. Khunrae, F. Viwami, J. Bruun, M. Dahlbäck, N. L. Bernasconi, M. Fried, D. John, P. E. Duffy, A. Salanti, A. Lanzavecchia, C. T. Lim, N. T. Ndam, M. K. Higgins, and L. Hviid, "Chondroitin Sulfate A-Adhering Plasmodium falciparum-Infected Erythrocytes Express Functionally Important Antibody Epitopes Shared by Multiple Variants," *The Journal of Immunology*, vol. 185, pp. 7553 LP – 7561, 12 2010.

[312] M. Avril, B. R. Kulasekara, S. O. Gose, C. Rowe, M. Dahlbäck, P. E. Duffy, M. Fried, A. Salanti, L. Misher, D. L. Narum, and J. D. Smith, "Evidence for Globally Shared, Cross-Reacting Polymorphic Epitopes in the Pregnancy-

Associated Malaria Vaccine Candidate VAR2CSA," *Infection and Immunity*, vol. 76, pp. 1791 LP – 1800, 4 2008.

[313] P. Fernandez, N. K. Viebig, S. Dechavanne, C. Lépolard, J. Gysin, A. Scherf, and B. Gamain, "Var2CSA DBL6-epsilon domain expressed in HEK293 induces limited cross-reactive and blocking antibodies to CSA binding parasites," *Malaria Journal*, vol. 7, no. 1, p. 170, 2008.

[314] A. Salanti, M. Resende, S. B. Ditlev, V. V. Pinto, M. Dahlbäck, G. Andersen, T. Manczak, T. G. Theander, and M. A. Nielsen, "Several domains from VAR2CSA can induce Plasmodium falciparum adhesion-blocking antibodies," *Malaria Journal*, vol. 9, no. 1, p. 11, 2010.

[315] A. V. Oleinikov, S. E. Francis, J. R. Dorfman, E. Rossnagle, S. Balcaitis, T. Getz, M. Avril, S. Gose, J. D. Smith, M. Fried, and P. E. Duffy, "VAR2CSA Domains Expressed in Escherichia coli Induce Cross-Reactive Antibodies to Native Protein," *The Journal of Infectious Diseases*, vol. 197, pp. 1119–1123, 4 2008.

[316] M. A. Nielsen, V. V. Pinto, M. Resende, M. Dahlbäck, S. B. Ditlev, T. G. Theander, and A. Salanti, "Induction of Adhesion-Inhibitory Antibodies against Placental Plasmodium falciparum Parasites by Using Single Domains of VAR2CSA," *Infection and Immunity*, vol. 77, pp. 2482 LP – 2487, 6 2009.

[317] J. Doritchamou, P. Bigey, M. A. Nielsen, S. Gnidehou, S. Ezinmegnon, A. Burgain, A. Massougbodji, P. Deloron, A. Salanti, and N. T. Ndam, "Differential adhesion-inhibitory patterns of antibodies raised against two major variants of the NTS-DBL2X region of VAR2CSA," *Vaccine*, vol. 31, no. 41, pp. 4516–4522, 2013.

[318] P. Bigey, S. Gnidehou, J. Doritchamou, M. Quiviger, F. Viwami, A. Couturier, A. Salanti, M. A. Nielsen, D. Scherman, P. Deloron, and N. Tuikue Ndam, "The NTS-DBL2X region of VAR2CSA induces cross-reactive antibodies that inhibit adhesion of several Plasmodium falciparum isolates to chondroitin sulfate A," *The Journal of infectious diseases*, vol. 204, no. 7, p. 1125, 2011.

[319] M. A. Nielsen, M. Resende, W. A. de Jongh, S. B. Ditlev, B. Mordmüller, S. Houard, N. T. Ndam, M. Agerbæk, M. Hamborg, A. Massougbodji, S. Issifou, A. Strøbæk, L. Poulsen, O. Leroy, P. G. Kremsner, J.-P. Chippaux, A. J. F. Luty, P. Deloron, T. G. Theander, C. Dyring, and A. Salanti, "The Influence of Sub-Unit Composition and Expression System on the Functional Antibody Response in the Development of a VAR2CSA Based Plasmodium falciparum Placental Malaria Vaccine," *PLOS ONE*, vol. 10, p. e0135406, 9 2015.

[320] A. M. Creasey, T. Staalsoe, A. Raza, D. E. Arnot, and J. A. Rowe, "Nonspecific Immunoglobulin M Binding and Chondroitin Sulfate A Binding Are Linked Phenotypes of Plasmodium falciparum Isolates Implicated in Malaria during Pregnancy," *Infection and Immunity*, vol. 71, pp. 4767 LP – 4771, 8 2003.

[321] N. Rasti, F. Namusoke, A. Chêne, Q. Chen, T. Staalsoe, M.-Q. Klinkert,

F. Mirembe, F. Kironde, and M. Wahlgren, "Nonimmune immunoglobulin binding and multiple adhesion characterize Plasmodium falciparum-infected erythrocytes of placental origin," *Proceedings of the National Academy of Sciences*, vol. 103, pp. 13795 LP – 13800, 9 2006.

[322] J. Alexandra Rowe, J. Shafi, O. Kai, K. Marsh, and A. Raza, "Nonimmune IgM, but not IgG binds to the surface of Plasmodium falciparum-infected erythrocytes and correlates with rosetting and severe malaria," *The American journal of tropical medicine and hygiene*, vol. 66, pp. 692–699, 6 2002.

[323] B. Maubert, N. Fievet, G. Tami, C. Boudin, and P. Deloron, "Plasmodium falciparum-isolates from Cameroonian pregnant women do not rosette," *Parasite*, vol. 5, pp. 281–283, 9 1998.

[324] S. J. Rogerson, J. G. Beeson, C. G. Mhango, F. K. Dzinjalamala, and M. E. Molyneux, "Plasmodium falciparum Rosette Formation Is Uncommon in Isolates from Pregnant Women," *Infection and Immunity*, vol. 68, pp. 391 LP – 393, 1 2000.

[325] D. M. Czajkowsky, A. Salanti, S. B. Ditlev, Z. Shao, A. Ghumra, J. A. Rowe, and R. J. Pleass, "IgM, Fc mu Rs, and malarial immune evasion," *Journal of immunology (Baltimore, Md. : 1950)*, vol. 184, pp. 4597–4603, 5 2010.

[326] L. Barfod, M. B. Dalgaard, S. T. Pleman, M. F. Ofori, R. J. Pleass, and L. Hviid, "Evasion of immunity to Plasmodium falciparum malaria by IgM masking of protective IgG epitopes in infected erythrocyte surface-exposed PfEMP1," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, pp. 12485–12490, 7 2011.

[327] J.-P. Semblat, A. Raza, S. A. Kyes, and J. A. Rowe, "Identification of Plasmodium falciparum var1CSA and var2CSA domains that bind IgM natural antibodies," *Molecular and biochemical parasitology*, vol. 146, pp. 192–197, 4 2006.

[328] M. Avril, M. M. Cartwright, M. J. Hathaway, M. Hommel, S. R. Elliott, K. Williamson, D. L. Narum, P. E. Duffy, M. Fried, J. G. Beeson, and J. D. Smith, "Immunization with VAR2CSA-DBL5 Recombinant Protein Elicits Broadly Cross-Reactive Antibodies to Placental Plasmodium falciparum-Infected Erythrocytes," *Infection and Immunity*, vol. 78, pp. 2248 LP – 2256, 5 2010.

[329] T. Saveria, A. V. Oleinikov, K. Wiliamson, R. Chaturvedi, J. Lograsso, G. J. Keitany, M. Fried, and P. Duffy, "Antibodies to Escherichia coli-expressed C-terminal domains of plasmodium falciparum variant surface antigen 2-chondroitin sulfate a (VAR2CSA) inhibit binding of CSA-adherent parasites to placental tissue," *Infection and Immunity*, vol. 81, pp. 1031 LP – 1039, 4 2013.

[330] A. Alkhalil, R. N. Achur, M. Valiyaveettil, C. F. Ockenhouse, and D. C. Gowda, "Structural Requirements for the Adherence ofPlasmodium falciparum-

infected Erythrocytes to Chondroitin Sulfate Proteoglycans of Human Placenta," *Journal of Biological Chemistry*, vol. 275, pp. 40357–40364, 12 2000.

[331] W. Kühlbrandt, "The Resolution Revolution," *Science*, vol. 343, pp. 1443 LP – 1444, 3 2014.

[332] A. Merk, A. Bartesaghi, S. Banerjee, V. Falconieri, P. Rao, M. I. Davis, R. Pragani, M. B. Boxer, L. A. Earl, J. L. S. Milne, and S. Subramaniam, "Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery," *Cell*, vol. 165, no. 7, pp. 1698–1707, 2016.

[333] A. Bartesaghi, A. Merk, S. Banerjee, D. Matthies, X. Wu, J. L. S. Milne, and S. Subramaniam, "2.2 Å resolution cryo-EM structure of β-galactosidase in complex with a cell-permeant inhibitor," *Science (New York, N.Y.)*, vol. 348, no. 6239, p. 1147, 2015.

[334] M. A. H. Jr, M. Wu, and G. C. Lander, "Achieving better-than-3-Å resolution by single-particle cryo-EM at 200 keV," *Nature Methods*, vol. 14, no. 11, p. 1075, 2017.

[335] R. M. Glaeser, "How good can cryo-EM become?," *Nature methods*, vol. 13, no. 1, p. 28, 2015.

[336] R. M. Glaeser and B.-G. Han, "Opinion: hazards faced by macromolecules when confined to thin aqueous films," *Biophysics Reports*, vol. 3, no. 1, pp. 1–7, 2017.

[337] K. A. Taylor and R. M. Glaeser, "Retrospective on the early development of cryoelectron microscopy of macromolecules and a prospective on opportunities for the future," *Journal of Structural Biology*, vol. 163, no. 3, pp. 214–223, 2008.

[338] H. Maity, M. Maity, M. M. G. Krishna, L. Mayne, and S. W. Englander, "Protein folding: The stepwise assembly of foldon units," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, pp. 4741 LP – 4746, 3 2005.

[339] G. Raffaini and F. Ganazzoli, "Protein Adsorption on a Hydrophobic Surface: A Molecular Dynamics Study of Lysozyme on Graphite," *Langmuir*, vol. 26, pp. 5679–5689, 4 2010.

[340] E. D'Imprima, D. Floris, M. Joppe, R. Sánchez, M. Grininger, and W. Kühlbrandt, "The deadly touch: protein denaturation at the water-air interface and how to prevent it," *bioRxiv*, p. 400432, 1 2018.

[341] A. Chari, D. Haselbach, J.-M. Kirves, J. Ohmer, E. Paknia, N. Fischer, O. Ganichkin, V. Möller, J. J. Frye, G. Petzold, M. Jarvis, M. Tietzel, C. Grimm, J.-M. Peters, B. A. Schulman, K. Tittmann, J. Markl, U. Fischer, and H. Stark, "ProteoPlex: stability optimization of macromolecular complexes by sparse-matrix screening of chemical space," *Nature Methods*, vol. 12, p. 859, 8 2015.

[342] J. N. Blaza, K. R. Vinothkumar, and J. Hirst, "Structure of the Deactive State of Mammalian Respiratory Complex I," *Structure (London, England : 1993)*, vol. 26, pp. 312–319, 2 2018.

[343] Z. L. Johnson and J. Chen, "Structural Basis of Substrate Recognition by the Multidrug Resistance Protein MRP1," *Cell*, vol. 168, no. 6, pp. 1075–1085, 2017.

[344] J. F. Fay, L. A. Aleksandrov, T. J. Jensen, L. L. Cui, J. N. Kousouros, L. He, A. A. Aleksandrov, D. S. Gingerich, J. R. Riordan, and J. Z. Chen, "Cryo-EM Visualization of an Active High Open Probability CFTR Anion Channel," *Biochemistry*, vol. 57, pp. 6234–6246, 10 2018.

[345] A. Blees, D. Januliene, T. Hofmann, N. Koller, C. Schmidt, S. Trowitzsch, A. Moeller, and R. Tampé, "Structure of the human MHC-I peptide-loading complex," *Nature*, vol. 551, p. 525, 11 2017.

[346] C. J. Russo and L. A. Passmore, "Controlling protein adsorption on graphene for cryo-EM using low-energy hydrogen plasmas," *Nature methods*, vol. 11, pp. 649–652, 6 2014.

[347] Y. Cheng, E. Wolf, M. Larvie, O. Zak, P. Aisen, N. Grigorieff, S. C. Harrison, and T. Walz, "Single Particle Reconstructions of the Transferrin–Transferrin Receptor Complex Obtained with Different Specimen Preparation Techniques," *Journal of Molecular Biology*, vol. 355, no. 5, pp. 1048–1065, 2006.

[348] Y. Z. Tan, P. R. Baldwin, J. H. Davis, J. R. Williamson, C. S. Potter, B. Carragher, and D. Lyumkis, "Addressing preferred specimen orientation in single-particle cryo-EM through tilting," *Nature Methods*, vol. 14, p. 793, 7 2017.

[349] S. J. Hubbard, "The structural aspects of limited proteolysis of native proteins," *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, vol. 1382, no. 2, pp. 191–206, 1998.

[350] M.-L. Huynh, P. Russell, and B. Walsh, "Tryptic Digestion of In-Gel Proteins for Mass Spectrometry Analysis BT - Two-Dimensional Electrophoresis Protocols," pp. 507–513, Totowa, NJ: Humana Press, 2009.

[351] A. C. Leney and A. J. R. Heck, "Native Mass Spectrometry: What is in the Name?," *Journal of the American Society for Mass Spectrometry*, vol. 28, pp. 5–13, 1 2017.

[352] L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, and M. J. E. Sternberg, "The Phyre2 web portal for protein modeling, prediction and analysis," *Nature Protocols*, vol. 10, p. 845, 5 2015.

[353] C. J. Ackerson, R. D. Powell, and J. F. Hainfeld, "Chapter Nine - Site-Specific Biomolecule Labeling with Gold Clusters," in *Cryo-EM Part A Sample Preparation and Data Collection* (G. J. B. T. M. i. E. Jensen, ed.), vol. 481, pp. 195–230, Academic Press, 2010.