Structural Studies of the Human $GABA_B$ Receptor using Cryo-Electron Microscopy



Zhaoyang Sun

Doctoral Training Centre - Medical Science Division
Division of Structural Biology
St Catherine's College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
Trinity 2018

Abstract

The metabotropic gamma-aminobutyric acid type-B receptor ($GABA_BR$) mediates slow and prolonged inhibitory neurotransmission in the mammalian nervous system. $GABA_BR$ belongs to the Class C GPCR family and is the first GPCR reported to function as an obligate heterodimer. Structurally, $GABA_BR$ displays distinctive features for its lack of the cysteine-rich domain found conserved among other Class C receptors. $GABA_BR$'s extracellular Venus flytrap (VFT) domain is directly connected to the transmembrane domain (TMD), followed by a small cytoplasmic domain. To date, structures of an isolated extracellular VFT domain and an intracellular coiled-coil domain have been solved using X-ray crystallography. No structure of the transmembrane domain (TMD) or the full-length receptor has been reported.

This project focused on the recombinant production and structural analysis of the full-length human $GABA_BR$ using single-particle cryo-Electron Microcopy (cryo-EM). It was determined that the HEK293 $GnTI^-$ cells were a suitable host cell system for production of the integral human $GABA_BR$. Following extensive optimisation of constructs and expression conditions, stable cell lines were generated which yielded sufficient material for structural work. $GABA_BR$ was isolated from the membranes using a combination of detergents that maintained the receptor's ligand binding activity. The high purity heterodimeric $GABA_BR$ samples obtained were used to screen various reconstitution systems compatible with cryo-EM, including amphipols, salipro particles and MSP nanodiscs. The MSP2N2 lipid nanodisc was found to be optimal for reconstituting $GABA_BR$, yielding intact receptor capable of binding to G protein. Thus far, medium resolution (10 ~ 15Å) models have been reconstructed for the full-length

$GABA_BR$. These models provided new insights regarding $GABA_BR$ domain organisations and receptor dynamics. In addition, a ~4Å reconstruction was obtained for the extracellular VFT domain alone and showed density for regions missing in the published crystal structures. The work presented in this thesis provides the foundation for solving higher-resolution $GABA_BR$ structures and for their detailed mechanistic investigation.

Declaration of work

I, Zhaoyang Sun, confirm that this thesis is my own work and it has not been submitted anywhere for any award. Where information has been derived from other sources, it has been indicated in the thesis.

Printed: ZHAOYANG SUN

Date: July 2019

Signature：

Acknowledgements

I would like to express my sincere gratitude to my supervisor Prof. Radu Aricescu for offering me the opportunity to work on this stimulating project and for providing help and advice over the course of my study. In addition, the enthusiasm and joy he has for research has bas been a great inspiration. I would also like to thank supervisor Prof. Juha Huiskonen, who introduced me to the world of Cryo-EM, for having me work on the interesting virus projects and his ongoing support for the $GABA_BR$ project.

It has been an unique experience to have spent time in both STRUBI and LMB and I am very grateful for the people I have met. The Aricescu group has been a wonderful source of professional advice and personal friendships. In particular, I very much appreciate the useful comments from Dr. Lina Malinauskaite, Dr Duncan Laverty and Ms. Christina Heroven for this thesis. I would also like to thank Dr. Paul Miller, Dr Veronica Chang and Dr. Simonas Masiulis for their advice in the lab. It was great to do this Ph.D together with Ms. Suzanne Scott, who has been a great lab mate and cake provider. I would like to thank Dr. Sai Li and Dr. Chris Tate for their helpful advices on my projects. Many thanks to the people who helped to keep the facilities running smoothly, in particular Ms. Margaret Jones, Ms. Weixian Lu and Dr. Giuseppe Cannone.

I wish to thank the Wellcome Trust for their Financial Support.

I have been incredibly lucky to have had the support and friendship from a fantastic group of people in Oxford and Cambridge. To Bo, Qing, Claire, Kun, Hongjia, Yuewen, Chen, Xiaoting, Weimin, Tao, and Ting, for your care and encouragements. Thank you all for sharing this journey with me, with all the love and joy and food you have provided.

Lastly, I wish to thank my parents for their love and their readiness to support me wherever and whenever needed.

Holly 昭阳

October 2018

Table of Contents

Abstract ***1***
Declaration of work ***3***
Acknowledgement ***4***
Table of Contents ***5***
List of figures ***8***
List of Tables ***10***
Abbreviations ***11***

Chapter 1. INTRODUCTION ***12***

1.1 Introduction to G protein-coupled receptors (GPCRs) **12**
1.1.1 GPCRs function and classification 12
1.1.2 Structural determinations of GPCRs 14

1.2 The metabotropic $GABA_BR$ as a GPCR **15**
1.2.1 Discovery of the $GABA_BR$ 15
1.2.2 Neuronal inhibition by $GABA_BR$ 16

1.3 $GABA_BR$ functions as an obligate heterodimer **19**
1.3.1 $GABA_BR$ subunits and isoforms 19
1.3.2 Functions of $GABA_BR$ subunits 21

1.4 Structural organisation of heterodimeric $GABA_BR$ **22**
1.4.1 $GABA_BR$'s distinctive features from other Class C GPCR 22
1.4.2 $GABA_BR$ extracellular domain arrangements 24
1.4.3 $GABA_BR$ Transmembrane domain organisation 29
1.4.4 $GABA_BR$ Cytoplasmic domain organisation 29

1.5 Pharmacology of $GABA_BR$ **31**
1.5.1 $GABA_BR$ agonists 31
1.5.2 $GABA_BR$ antagonists 32
1.5.3 $GABA_BR$ allosteric modulators 33

1.6 $GABA_BR$ ligand recognition mechanisms **35**
1.6.1 Structural plasticity of the $GABA_BR$ VFT 35
1.6.2 $GABA_BR$ antagonist recognition 36
1.6.3 $GABA_BR$ agonist recognition 38

1.7 Aims **39**

Chapter 2. METHODS ***41***

2.1 Molecular Biology **41**
2.1.1 Construct design 41
2.1.2 Polymerase Chain Reactions (PCR) 41
2.1.3 Cloning and transformation 43

2.2 Mammalian Cell culture **44**
2.2.1 Growing and passaging of adherent cells 44
2.2.2 Growing and passaging of suspension cells 44
2.2.3 DNA plasmids preparation for transient transfection 45
2.2.4 Transient transfection and expression in adherent cells 45
2.2.5 Transient transfection and expression in suspension cells 46
2.2.6 Lentiviral cell lines generation and protein expression 46
2.2.7 Stable cell line expression in WAVE bioreactors 48

2.3 Bacterial protein expression and purification **48**
2.3.1 Expression and purification of MSP proteins 48

2.3.2 Expression and purification of saposin proteins 49

2.4 $GABA_BR$ solubilisation and purification 51
2.4.1 Detergent purification from small-scale adherent cells 51
2.4.2 Detergent purification from suspension cells 52
2.4.3 Non-Detergent Membrane solubilisation 53
2.4.4 Size Exclusion Chromatography (SEC) 54
2.4.5 Generation of 1D4-CNBr affinity beads 55

2.5 Scintillation proximity assay 56

2.6 $GABA_BR$ reconstitution 57
2.6.1 On-beads Reconstitution into amphipols/nanodisc 57
2.6.2 Reconstitution of purified proteins 58

2.7 G protein complex formation 58

2.8 Chemical crosslinking 58
2.8.1 Chemical crosslinking using BS^3 59
2.8.2 Chemical crosslinking using the CovalX kit 59

2.9 EM sample preparation 59
2.9.1 Negative stain EM grid preparation 59
2.9.2 Cryo-EM grid preparation 60

2.10 Cryo-EM Data processing 61
2.10.1 Single particle analysis using RELION 61
2.10.2 Single particle analysis using cryoSPARC 62

Chapter 3. EXPRESSION OF HUMAN $GABA_BR$ IN MAMMALIAN CELLS 64

3.1 Introduction 64

3.2 Results 66
3.2.1 Design of human $GABA_BR$ synthetic constructs 66
3.2.2 $GABA_BR$ expression and surface trafficking in HEK293 $GnTI^-$ cells 67
3.2.3 Protein expression level optimisation by protein engineering 70
3.2.4 Improving expression in transient transfected suspension cells 73
3.2.5 Generation of polyclonal stable cell lines 77

3.3 Discussion 80

3.4 Conclusion 84

Chapter 4. SAMPLE PREPARATION FOR CRYO-EM STUDIES 85

4.1 Introduction 85

4.2 Results 87
4.2.1 Detergent screening 87
4.2.2 Attempts in direct solubilisation 93
4.2.3 Optimisations on affinity purification 98
4.2.4 Effect of different gel filtration columns 102
4.2.5 Optimisation of purification buffer condition 103
4.2.6 Reconstitution of detergent solubilised receptors 106

4.3 Discussion 110
4.3.1 Establishment of a multi-step workflow for Cryo-EM sample preparation 110
4.3.2 Monitoring receptor stability and function 113
4.3.3 $GABA_BR$ reconstitution for cryo-EM studies 115

4.4 Conclusion 117

Chapter 5. CRYO-EM ANALYSIS OF $GABA_BR$ RECONSTITUTED IN AMPHIPOLS. 118

5.1 Introduction 118

5.2 Results 119

5.2.1 $GABA_BR$ reconstitution into amphipols 119

5.2.2 Data processing and analysis of $GABA_{B1b}$-bRIL/$GABA_{B2}$-amphipols particles 123

5.2.3 Data processing and analysis of $GABA_{B1b}$-bRIL/$_{B2}$-Cys-amphipols particles 128

5.3 Discussion 134

5.3.1 Amphipols reconstitution for cryo-EM studies 134

5.3.2 Model reconstruction of $GABA_{B1b}$-bRIL/$_{B2}$ particles 135

5.3.3 Model reconstruction of $GABA_{B1b}$-bRIL/$_{B2}$-cys particles 137

5.4 Conclusion 138

Chapter 6. CRYO-EM DATA ANALYSIS OF $GABA_BR$ RECONSTITUTED IN LIPID NANODISCS 139

6.1 Introduction 139

6.2 Results 141

6.2.1 Reconstitution into salipro nanodiscs 141

6.2.2 Reconstitution into MSP1D1 nanodiscs 144

6.2.3 Data processing and analysis of the MSP1D1 nanodiscs particles 147

6.2.4 Reconstitution into MSP2N2 nanodiscs 153

6.2.5 Data processing and analysis of the $GABA_BR$-WT-2N2 particles 155

6.2.6 Forming the GPCR-G protein complex 162

6.2.7 Data processing and analysis of the WT-$GABA_BR$-MSP2N2-GFPminiGo particles 166

6.3 Discussion 170

6.3.1 Choosing the suitable nanodisc reconstitution system 170

6.3.2 $GABA_BR$ exhibited large degree of inter-domain flexibility 172

6.3.3 Stabilisation of the receptor TMD by G-protein 174

6.4 Conclusion 175

Chapter 7. Concluding remarks and future work 176

References 178

Appendix A 191

Appendix B 221

List of Figures

Figure 1.1 Phylogenetic relationships between human GPCRs. 13
Figure 1.2 Schematic model for G protein activation and deactivation cycle. 17
Figure 1.3 Locations and functions of $GABA_BR$. 18
Figure 1.4 Schematic model of a functional $GABA_BR$ heterodimer. 22
Figure 1.5 Differences in domain organisations between mGluR and $GABA_BR$. 23
Figure 1.6 Domain organisations of $GABA_BR$. 24
Figure 1.7 VFT domain of $GABA_{B1b}$ and $GABA_{B2}$ displays high structural similarities. 25
Figure 1.8 Heterodimeric VFT upper lobe interface. 26
Figure 1.9 Heterodimeric VFT lower lobe interface. 28
Figure 1.10 Crystal structure of a $GABA_{B1/B2}$ intracellular coiled-coil. 30
Figure 1.11 Chemical structures of representative $GABA_BR$ agonists, antagonists and allosteric modulators. 34
Figure 1.12 $GABA_BR$ VFT oscillates between open and closed conformations. 36
Figure 1.13 Antagonist recognition mechanism. 37
Figure 1.14 Agonist recognition mechanism. 39
Figure 2.1 Reproducibility of Shimazdu FSEC analysis. . 54
Figure 2.2 Schematic illustration of the scintillation proximity assay. 56
Figure 2.3 Schematic Representation of RELION-based data processing work flow. 62
Figure 2.4 Schematic Representation of CryoSPARC data processing work flow. 63
Figure 3.1 List of initial constructs generated. 66
Figure 3.2 Western blot analysis of the recombinantly expressed protein from HEK293 $GnTI^-$ cells. 68
Figure 3.3 Confirmation of $GABA_{B1/B2}$ heterodimer formation. 69
Figure 3.4 Localisation of recombinantly expressed $GABA_BRs$ in adherent cells. 70
Figure 3.5 Effects of protein engineering on expression levels. 73
Figure 3.6 Effects on variations of DNA ratios used for transfection while keeping total amount of DNA for each transfection the same. 75
Figure 3.7 Comparison of protein expression levels in suspension culture under different conditions. 76
Figure 3.8 $GABA_BR$ expression levels in constitutively active cell line. 77
Figure 3.9 Comparison of $GABA_BR$ expression levels in constitutive active cell lines and inducible cell lines. 79
Figure 3.10 Comparison of expression levels of cells grown in suspension bottles and WAVEbag. 80
Figure 4.1 Analysis of solubilisation efficiency of various detergents. 89
Figure 4.2 SEC analysis of protein purified using detergent mixture. 91
Figure 4.3 Binding activities of detergent solubilised $GABA_BR$. 92
Figure 4.4 SEC-based thermostability analysis. 93
Figure 4.5 Schematic illustration of membrane solubilisation using SMA. 94
Figure 4.6 Analysis of SMA extraction products. 95
Figure 4.7 Analysis of saposin D extraction products. 97
Figure 4.8 Analysis of purification specificity of each tag. 99
Figure 4.9 Purification efficiency analysis of SBP-tag and 1D4-tag. 100
Figure 4.10 Optimisation of the 1D4-tag purification step. 101
Figure 4.11 Comparison of 1D4-tag purification products from large scale 2 L purification and 500 ml medium scale purification. 102
Figure 4.12 Representative SEC results from double-purified $GABA_BR$. 103
Figure 4.13 SEC profiles of $GABA_BR$ purified in different buffer conditions. 105
Figure 4.14 Effects of small molecule supplements on receptor thermostability. 106
Figure 4.15 Representative micrograph of detergent-purified $GABA_BR$ particles. 107
Figure 4.16 SEC profiles of $GABA_BR$ reconstituted into different systems. 109
Figure 4.17 Solubilisition systems tested in this chapter. 117
Figure 5.1 Schematic illustration of the amphipols exchange process. 119
Figure 5.2 Biochemical characterisation of amphipol-reconstituted $GABA_BR$. 121
Figure 5.3 Amphipols-reconstituted $GABA_BR$ Cryo-EM sample. 122
Figure 5.4 Model of the recombinant $GABA_{B1b}$-bRIL/$GABA_{B2}$ receptor. 123

Figure 5.5 2D classification results of $GABA_{B1b}$-bRIL/ $GABA_{B2}$-amphipols particles. _____ *124*
Figure 5.6 Representative 2D class averages. _____ *125*
Figure 5.7 3D reconstruction model of $GABA_{B1}$-Bril/$_{B2}$ reconstituted in amphipols. _____ *127*
Figure 5.8 Localised reconstruction of ECD region with VFT crystal structure fitted. _____ *128*
Figure 5.9 Location of point directed mutations in the receptor. _____ *129*
Figure 5.10 Characterisation of $GABA_{B1b}$-bRIL/$_{B2}$-cys mutant receptor. _____ *130*
Figure 5.11 Representative 2D classification averages of $GABA_{B1b}$-bRIL/$_{B2}$-cys particles. _____ *132*
Figure 5.12 3D reconstruction of $GABA_{B1b}$-bRIL/$_{B2}$-cys particles. _____ *133*
Figure 6.1 Analysis of $GABA_BR$-salipro particle reconstitution. _____ *142*
Figure 6.2 Analysis of $GABA_BR$ Salipro particle Cryo-EM sample. _____ *143*
Figure 6.3 $GABA_BR$ reconstitution into MSP1D1 nanodiscs. _____ *145*
Figure 6.4 Analysis of MSP1D1 reconstituted $GABA_BR$ particles. _____ *146*
Figure 6.5 Particle picking for MSP1D1 reconstituted $GABA_BR$. _____ *148*
Figure 6.6 2D averages of MSP1D1 reconstituted particles. _____ *149*
Figure 6.7 Representative 2D class averages of half-sized particles. _____ *150*
Figure 6.8 Data processing of $GABA_BR$ ECD particles. _____ *151*
Figure 6.9 Heteromeric ECD VFT EM model. _____ *152*
Figure 6.10 Analysis of MSP2N2 reconstituted particles. _____ *155*
Figure 6.11 Micrograph of MSP2N2-reconstituted particles. _____ *156*
Figure 6.12 Confirmation of intact $GABA_BR$ reconstituted in MSP2N2 nanodiscs. _____ *157*
Figure 6.13 2D classification results of MSP2N2-reconstituted particles. _____ *158*
Figure 6.14. $GABA_BR$-MSP2N2 nanodisc particles reconstruction. _____ *161*
Figure 6.15 $GABA_BR$-miniG_{ao} complex formation under low Na^+ environment. _____ *164*
Figure 6.16 Effects of chemical cross-linking treatments on the $GABA_BR$-miniG_{ao} incubated sample. _ *165*
Figure 6.17 Representative 2D classification results from non-crosslinked $GABA_BR$/miniG_{ao} sample. _ *167*
Figure 6.18 3D reconstruction model of MSP2N2 reconstituted particles. _____ *169*

List of Tables

Table 1.1 GPCR classification based on phylogenetic relationships._____ 13
Table 1.2 Number of GPCR structures solved._____ 14
Table 1.3 Binding affinities of representative $GABA_BR$ ligands._____ 35
Table 1.4 Existing crystal structures of $GABA_B$ VFT. _____ 35
Table 3.1 Constructs designed for expression level tests. _____ 71
Table 4.1 Binding activities of $GABA_BR$ solubilised in different detergents. _____ 90
Table 5.1 Data collection statistics for $GABA_{B1b}$-bRIL/ $GABA_{B2}$-amphipols _____ 123
Table 5.2 Data collection statistics for B1b-bRIL/B2-Cys data set_____ 131
Table 6.1 Data collection statistics for $GABA_BR$_WT_MSP1D1 _____ 147
Table 6.2 Data collection statistics for WT $GABA_BR$-MSP2N2-GFPminiGo._____ 166

Abbreviations

bRIL	apocytochrome b562RIL
BSA	bovine serum albumin
CHS	cholesteryl hemisuccinate
Cryo-EM	Cryo-Electron Micrsocopy
DDM	n-dodecyl β-D-maltopyranoside
DDTM	n-Dodecyl-β-D-thiomaltopyranoside
DM	n-decyl β-D-maltopyranoside
DMEM	Dulbecco's Modified Eagle's Medium
DMNG	decyl maltose neopentyl glycol
DMSO	dimethyl sulphoxide
ECD	extracellulae domain
FCS	fetal calf serum
FSEC	fluorescence-detection size-exclusion chromatography
GABA	γ-aminobutyric acid
$GABA_BR$	γ-aminobutyric acid type-B receptor
GFP	green fluorescent protein
$GNTI^-$	N-acetylglucosaminyltransferase I deficient
GPCR	G protein-coupled receptor
HPLC	high performance liquid chromatography
ICD	intracellular domain
mGluR	metabotropic glutamate receptor
MSP	membrane scaffolding protein
NEAA	non-essential amino acid
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	protein data bank
PEI	polyethylenimine
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
VPA	valproic acid

Chapter 1. INTRODUCTION

1.1 Introduction to G protein-coupled receptors (GPCRs)

1.1.1 GPCRs function and classification

G protein-coupled receptors (GPCRs) are a family of membrane-bound receptors that transduce signals across the cell membrane. The receptors elicit their actions through the heterotrimeric G proteins inside the cell [1]. GPCRs are involved in a wide range of essential physiological activities, in response to a broad range of stimuli such as light, ions, neurotransmitters and hormones [2]. GPCRs represent the largest family of receptors in the human body, with more than 800 genes encoded in the genome [3]. Receptors in the family can be classified into 5 distinctive classes (Table 1.1), based on their sequence identity within the transmembrane region (Figure 1.1) [4, 5]. The Glutamate class (also known as class C) GPCRs are mainly involved in synaptic transmissions, including the metabotropic glutamate receptors (mGluRs), calcium-sensing receptor and gamma-aminobutyric acid type B receptor ($GABA_BR$), as well as 3 taste receptors [5]. The Rhodopsin-like class (or Class A) is the largest class, containing receptors for a wide range of functions. Representative receptors from this class include the adrenoceptors, chemokine receptors and dopamine receptors [5]. The Secretin class is a small class, consisted of receptors responding to polypeptide hormones [5]. The Adhesion class receptors have close phylogenetically relations to the secretin class [5]. Most receptors in adhesion class are still orphans with no endogenous ligand assigned [6]. The smallest class Frizzled includes 10 frizzled receptors of the WNT signaling pathway and one smoothened receptor of the Hedgehog pathway [5, 7].

Table 1.1 GPCR classification based on phylogenetic relationships.

Receptors are grouped into 5 classes following the GRAFS nomenclature system [5].Table summarised number of receptors in each class using data collected from the IUPHAR/BPS Guide on 19/09/2018.

Class	Glutamate	Rhodopsin	Adhesion	Frizzled & Smoothened	Secretin
Receptors with known ligands	12	197	0	11	15
Orphans	8	87	26	0	-
Sensory	3	435	-	-	-
Total	22	719	33	11	15

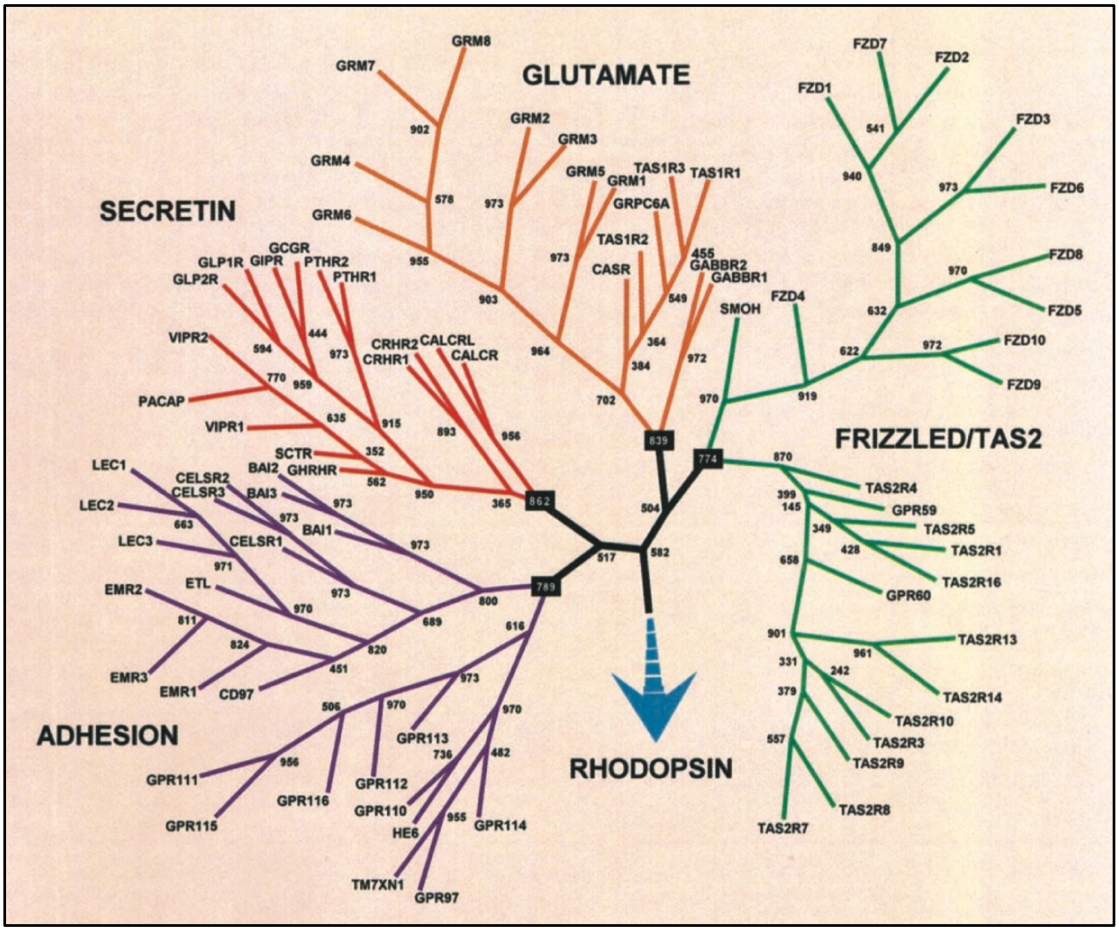


Figure 1.1 Phylogenetic relationships between human GPCRs.

The phylogenetic tree had receptors clustered into 5 classes based on the GRAFS classification system [5]. Individual receptors in class A was not shown due to the large size of the class. Figure was adapted from [5].

1.1.2 Structural determinations of GPCRs

High-resolution protein structures can provide useful information to better understand receptor functions and aid future drug development. In 2000, a 2.8 Å bovine rhodopsin crystal structure was obtained [8] and marked the first high-resolution GPCR structure reported. Since then, structures of another 51 receptors have been solved (Table 1.2) [9]. Most of the receptors solved belong to the Class A family, which mainly consists of a 7-helices transmembrane domain (TMD) and a short extracellular N-terminus. The underrepresentation of structures from other classes could be due to a combination of reasons, including but not limited to, low protein expression level, instability in detergent and high conformational flexibility.

Table 1.2 Number of GPCR structures solved.

Class	Total	Class A (Rhodopsin)	Secretin	Adhesion	Class C (Glutamate)	Class F (Frizzled/ Smoothened)
Unique Receptors**	52	45	4	0	2	1
Receptor - ligand complexes	364	310	25	0	13	16
Receptor-G protein complexes	4	2	2	0	0	0
Active-state receptors*	18	16	2	0	0	0

*Active state is defined as agonist-bound and opened intracellular TM bundle. **Receptor-orthologues are only counted once. Table was modified based on data from the GPCRdb database [9] on 20180922.

Traditionally, studies of GPCR structures mainly relied on X-ray crystallography [10]. In order to obtain well-diffracting crystals, receptors often had to be heavily engineered to improve stability and crystal packing. In addition, many of the structures solved were in the inactivated state [10], which is the more stable form of

a GPCR. Recent advances in the cryo-Electron Microscopy (cryo-EM) technology had made high-resolution structural determination of smaller sized proteins (<300kDa range) possible, using a combination of better imaging equipment and processing software. Since the cryo-EM "resolution revolution", more of the recently reported structures were activated state receptors in complex with a G-protein. Representative receptor structures in activated state include the Class A 5-HT_{1b} receptor [11] and class B GLP-1 receptor [12]. So far no full-length structure from the Glutamate class (Class C) and the Adhesion class family has been described.

1.2 The metabotropic $GABA_BR$ as a GPCR

1.2.1 Discovery of the $GABA_BR$

Gamma-aminobutyric acid (GABA) is the chief inhibitory neurotransmitter in the mammalian nervous system [13]. In 1979, $GABA_BR$ was discovered by Bowery and colleagues as a novel receptor that responded to GABA but not bicuculline, which is a competitive antagonist to the only known GABA receptor at the time (now known as $GABA_AR$) [14, 15].

Our current understanding of the GABAergic pathways suggest that GABA exerts its effect through three types of receptors [16]: type A ($GABA_AR$), type B ($GABA_BR$), and Type C ($GABA_\rho R$). $GABA_AR$ and $GABA_\rho R$ are both ionotropic receptors that differ in their assembling subunits [17, 18]. Upon release of GABA, they provide neuronal responses with fast onset and short duration (within ms) [19]. $GABA_BR$, on the other hand, modulates synaptic activities over a slower timescale through secondary messenger networks [19]. $GABA_BR$ uses G-proteins to affect ion-channel

activities and adenylyl cyclase functions [2–4], which results, for example, in the hyperpolarisation of neurons or the reduced release of other neurotransmitters.

1.2.2 Neuronal inhibition by $GABA_BR$

Upon activation by an extracellular ligand, $GABA_BR$ acts on the intracellular heterotrimeric G-proteins, which further transduce signals to various downstream effectors inside the cell. The heteromeric G-protein is consisted of G_α, G_β and G_γ subunits, where G_β and G_γ form a stable dimeric functional complex (Figure 1.2) [20]. G_α exhibits intrinsic GTPase activity and its signalling activity is dependent on the type of molecule bound - with GTP as an activating factor and its hydrolysed product GDP as an inactivating factor [21]. GPCR activation promotes the release of GDP on the G_α subunit in exchange for GTP (Figure 1.2) [22]. Binding of GTP leads to dissociation of G protein trimer from the receptor and separation of G_α from $G_{\beta\gamma}$ subunits (Figure 1.2) [23]. Depending on their local environment, the separated G_α and $G_{\beta\gamma}$ subunits move on to mediate various downstream signalling pathways. Signalling is terminated when G_α and $G_{\beta\gamma}$ re-associate together due to hydrolysis of GTP on G_α (Figure 1.2) [24].

As an important class of signal transducers, heteromeric G proteins have evolved to a large, diverse group of proteins. There are currently 18 different G_α, 5 G_β and 12 G_γ proteins encoded in the human genome [25]. Heteromeric G proteins can be categorised into 4 classes based on phylogenetic relationships of their G_α subunit (namely $G_{\alpha s}$, $G_{\alpha i}$, $G_{\alpha q}$ and $G_{\alpha 12}$) [25]. $GABA_BR$ predominantly couples to the $G_{\alpha i}$ class of G-proteins [26-31], where the “I” stands for inhibition. This is in contrast to the $G_{\alpha s}$ class, where the “s” stands for stimulation. Both classes regulate activity of the membrane enzyme adenylyl cyclase, with $G_{\alpha i}$ acting as an inhibitory signal and $G_{\alpha s}$ as

a stimulatory signal [32]. Modulation of adenylyl cyclase changes the level of an important intracellular secondary messenger cyclic AMP and a cascade of downstream pathways dependent on it. In addition to G_{α}, $G_{\beta\gamma}$ released from $GABA_BR$ activation also serve to signal to downstream effectors. The main effector proteins include presynaptic calcium ion channels and postsynaptic potassium ion channels [28].

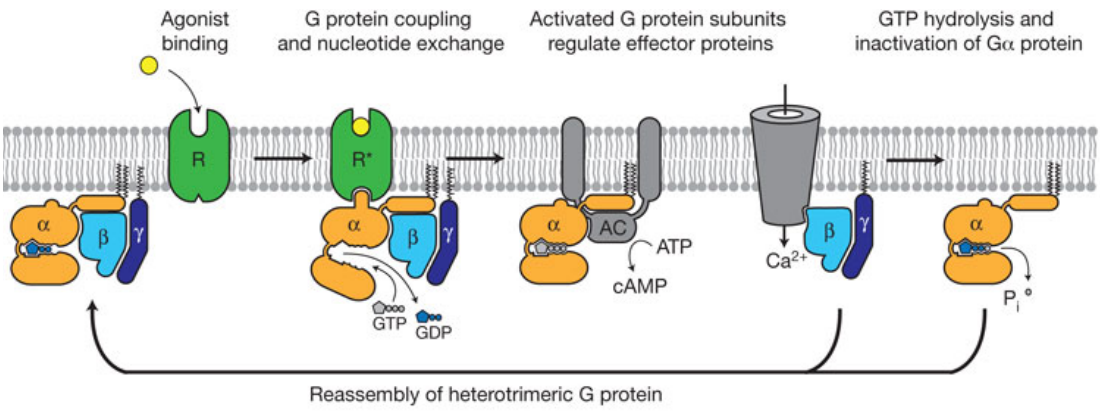


Figure 1.2 Schematic model for G protein activation and deactivation cycle.

Agonist binding induces conformational changes on the intracellular site of the GPCR (marked as "R", green) to allow association of the heterotrimeric G-protein (yellow, light blue and dark blue for G_{α}, G_{β} and G_{γ} subunit). Upon formation of a GPCR-G protein complex, GDP on the G_{α} subunit is exchanged to a GTP. GTP-bound G-protein dissociates from the receptor and separate into G_{α} and $G_{\beta\gamma}$ for downstream signalling pathways. Hydrolysis of GTP deactivates G_{α} and leads to its re-association with $G_{\beta\gamma}$. Signalling of $G_{\beta\gamma}$ stops once they re-assemble with the GDP-bound G_{α}. Figure was adapted from [22].

$GABA_BR$ is broadly expressed in the nervous system in both inhibitory and excitatory synapses [13, 33]. Presynaptic $GABA_BR$ (Figure 1.3) have been found to function as both auto-receptors, modulating levels of GABA, and hetero-receptors, modulating levels of other neurotransmitters, such as glutamate, dopamine, serotonin and acetylcholine [13, 34]. Presynaptic $GABA_BR$ achieves these effects by inhibiting voltage-gated calcium channels via $G_{\beta\gamma}$ subunits [27, 29]. Calcium channels at presynaptic terminals are involved in triggering neurotransmitter release [35]. Inhibition of the influx of calcium ions at presynaptic terminals thus inhibits the

release of neurotransmitters [36]. Postsynaptic $GABA_BR$ (Figure 1.3) is responsible for inducing a slow inhibitory effect upon the release of GABA [28]. Again, the $G_{\beta\gamma}$ subunits act as receptor transducers to activate potassium channels [28]. The opening of postsynaptic potassium channels allows the efflux of potassium ions, resulting in cell hyperpolarisation [37].

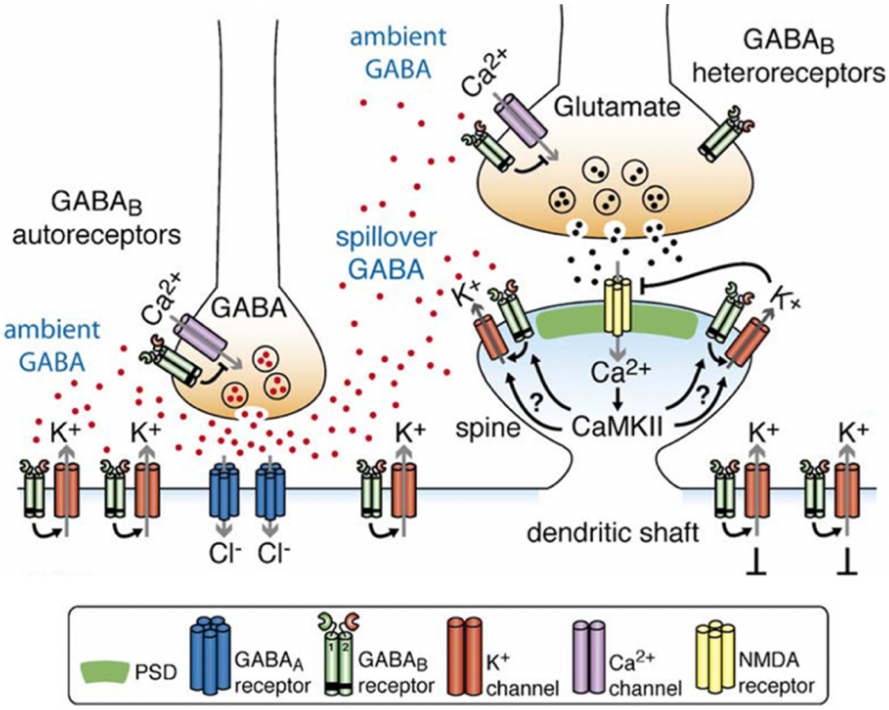


Figure 1.3 Locations and functions of $GABA_BR$.

Presynaptic $GABA_BR$ prevents neurotransmitter release by inhibiting the activity of voltage gated calcium channels. Presynaptic $GABA_BR$ can function as either auto-receptors to regulate release of GABA or as hetero-receptors to regulate release of other neurotransmitters. Postsynaptic $GABA_BR$ causes hyperpolarisation of the cell by activating potassium channels. Figure was modified from [38].

$GABA_BR$ has also been found to modulate forskolin-stimulated adenylyl cyclase activity [13]. Depending on the downstream pathways, $GABA_BR$ mediates through either the G_α or $G_{\beta\gamma}$ subunits. The G_α subunit inhibits the I, III, V and VI types of adenylyl cyclase, and the $G_{\beta\gamma}$ subunits stimulate some other types (II, IV and VII) [13]. Their effects on neuronal activity through the cyclic AMP pathway are expected to last over

a longer time scale compared to when they are coupled to ion-channels [13]. It has been also suggested that $GABA_BR$ can influence synaptic plasticity through the cyclic-AMP pathway [13].

Recent proteomic studies of native $GABA_BR$ have identified a diverse set of proteins that are potentially associated with $GABA_BR$-mediated cellular signaling events [39]. For example the K^+ channel tetramerisation-domain containing (KCTD) proteins, the potassium/sodium hyperpolarization-activated cyclic nucleotide-gated channel 2 and the amyloid beta A4 protein [39].

1.3 $GABA_BR$ functions as an obligate heterodimer

1.3.1 $GABA_BR$ subunits and isoforms

The $GABA_BR$ cDNA was isolated by expression cloning upon the development of a high affinity $GABA_BR$ radioligand ^{125}I-CGP64213 [40]. The first cDNA clone isolated was named $GABA_{B1a}$ and encoded a protein made of 960 amino acids. Later, a shorter clone corresponding to 884 amino acids was identified by cDNA library screening and was named $GABA_{B1b.}$ Several other variants were identified in the screen, but isoforms $GABA_{B1a}$ and $GABA_{B1b}$ remained the most predominant ones [40, 41]. Sequence alignment of $GABA_{B1a}$ and $GABA_{B1b}$ showed that the two proteins shared the same 797 residues counting from the C-terminus (sequence alignment in appendix B Figure B.1). $GABA_{B1a}$, the bigger isoform, has two additional sushi domains at the receptor N-terminus [40]. The sushi domains, also known as complement control modules or short consensus repeats, are common protein interactions motifs. The sushi domains on $GABA_{B1a}$ function as an axonal targeting signal to selectively traffic this isoform into the axons [42], while both $GABA_{B1a}$ and $GABA_{B1b}$ can be found on the dendritic side of

neurons [43, 44]. It was proposed that the default destination for $GABA_BR$ is the somatodendritic compartment, unless the sushi domain is present to bind to axonal sorting proteins [42]. Prevention of the correct folding of sushi domains was found to impair $GABA_{B1a}$ axonal trafficking without affecting receptor signal transduction [42]. Whether the different isoforms exhibit different functionalities is still controversial. While most studies suggest that they are not functionally and pharmacologically distinct, there is also evidence that they can regulate stress-related diseases differently [45, 46].

Recombinantly expressed $GABA_BRs$ using the clones mentioned above were found to bind to $GABA_BR$ radioligand but failed to couple to G proteins [40]. Further screening experiments using $GABA_{B1a}$ linked to sequence tag revealed a novel protein made of 941 amino acids [47]. The newly identified protein had 34% sequence identity with $GABA_{B1a}$ (sequence alignment in appendix B Figure B.1). Binding studies showed the novel protein was not responding to GABA nor $GABA_{B1}$ agonist baclofen [47]. However, when co-expressed together with $GABA_{B1a}$, the co-expressed product was able to bind to GABA and initiate downstream signaling events [48]. At the time, GPCRs were commonly thought to act as homomers and $GABA_BR$ was the first one appeared to function as a heteromeric receptor. The hypothesis that $GABA_BR$ heterodimerisation is required for function was soon confirmed by studies from three individual groups [49-51]. (Details regarding functions of each $GABA_BR$ subunit is described in section 1.3.2). The name of the novel protein was thus changed to $GABA_{B2}$ [48]. To date, there is no other confirmed $GABA_{B2}$ splice variant.

1.3.2 Functions of $GABA_BR$ subunits

$GABA_BR$ homodimers have been tested in multiple heterogeneous systems and have been shown to be functionally inactive [48]. It is now universally accepted that assembly of the heterodimer is required for forming functional $GABA_BR$. Each subunit of the $GABA_BR$ is specialised for different purposes (Figure 1.4): the $GABA_{B1}$ subunit for binding to orthosteric ligands; the $GABA_{B2}$ subunit for receptor cell surface trafficking, activating G-proteins and allosterically modulating ligand affinity [48, 52, 53]. The exact mechanism of how the signal is transduced from the $GABA_{B1}$ subunit to the $GABA_{B2}$ subunit is not known yet due to the lack of structural information.

To date, all $GABA_BR$ agonists and antagonists identified bind to an orthosteric site located at the cleft in the $GABA_{B1}$ VFT (Figure 1.4) [54]. No essential ligand-binding site has been reported at the $GABA_{B2}$ VFT and sequence alignment with $GABA_{B1}$ showed that the binding residues are not conserved in $GABA_{B2}$ [54]. Where key residues involved in ligand recognition in the $GABA_{B1}$ subunit are highly conserved across various species [55]. However, $GABA_{B1}$ will not be trafficked to the cell surface until its association with $GABA_{B2}$ [56]. In addition, $GABA_{B2}$ is the only subunit needed for and capable of G-protein coupling [57]. Chimera receptor dimers with two $GABA_{B2}$ transmembrane regions are functional, whereas mutated receptors with two $GABA_{B1}$ transmembrane regions are not [58]. Removal of any intracellular loops in B2 results in nonfunctional receptors. Mutagenesis studies have identified $GABA_{B2}$ residues Lys^{586}, Met^{587} and Lys^{590} in the second and third intracellular loops are important for G-protein activation [57]. The $GABA_{B2}$ subunit cytoplasmic terminus has been found

to play an important, but not essential, role in G-protein coupling [58]. Complete deletion of B2 C-termini results in significant attenuation to the receptor function [58].

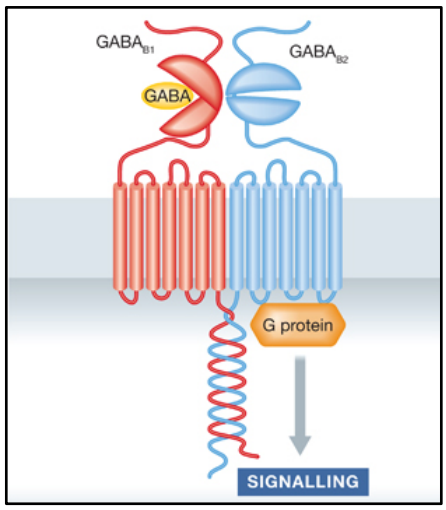


Figure 1.4 Schematic model of a functional GABA$_B$R heterodimer.

An orthosteric ligand binding site is found in the GABA$_{B1}$ subunit extracellular domain. Current hypothesis is that ligand-binding induces conformational changes in the receptor, allowing the GABA$_{B2}$ subunit to activate G proteins inside the cell. Figure was modified from [59]

1.4 Structural organisation of heterodimeric GABA$_B$R

1.4.1 GABA$_B$R's distinctive features from other Class C GPCR

GABA$_B$R belongs to the class C GPCR family which is found to always form dimeric receptors [2]. Receptors can form either as homomers, such as the Ca^{2+} sensing receptors [60], as heterodimers, such as the GABA$_B$R and taste receptors [61, 62], or both, such as the mGluRs [63-65]. The assembly of one GABA$_{B1}$ and one GABA$_{B2}$ subunit forms the minimal functional unit of a GABA$_B$R [51]. Tetramers and higher-order oligomers may also exist *in vivo* through GABA$_{B1}$-GABA$_{B1}$ interactions between heterodimers [66-68].

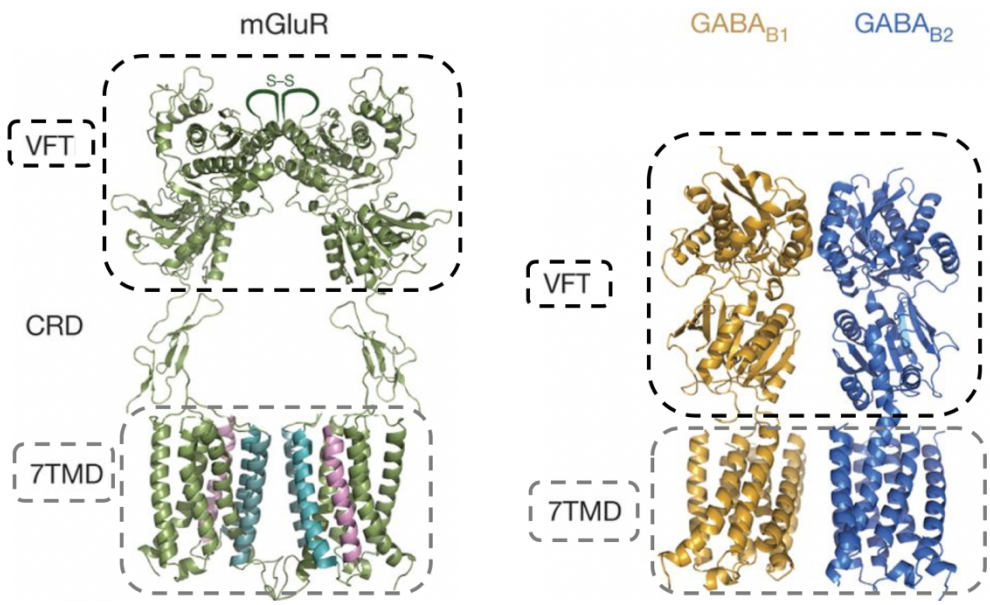


Figure 1.5 Differences in domain organisations between mGluR (left) and $GABA_BR$ (right).

The mGluR contains all key features conserved in a Class C GPCR, with the extracellular VFT connected to a CRD followed by the 7-helical TMD. In contrast, $GABA_BR$ has the VFT directly linked to the TMD. The exact way of how the TMD is connected to the extracellular domain remains unknown in both receptors. Models shown here represent the receptors in the resting conformation. mGluR homodimer model is based on the $mGlu_3$ VFT crystal structure (PDB 3SM9), $mGlu_3$ CRD crystal structure (PDB 2E4U) and $mGlu_5$ TMD crystal structure (PDB 4OO9). $GABA_BR$ heterodimer model is based on the heteromeric VFT crystal structure (PDB 4MR7) and $mGlu_5$ TMD crystal structure. Figure was modified from [65].

Structurally, all class C receptors possess a large N-terminal extracellular region where a highly conserved Venus Fly-Trap (VFT) domain is located [60, 69, 70]. It is named as the VFT because the movements in this domain resemble the plant Venus Fly-Trap when touched by a fly. Another domain found in all other class C receptors, but not in $GABA_BR$, is an extracellular cysteine rich domain (CRD) located between the VFT and the TMD (Figure 1.5) [40, 71]. TMD of class C receptors adapt the univeral GPCR 7-helices arrangement, and is follwed by a relatively short cytoplasmic domain [65]. Compared to other domains, the intracellular C-terminus region is more variable across different receptors in this class, possibly for selectivity of coupling to signaling protein [60].

The overall architectures of the $GABA_{B1}$ and $GABA_{B2}$ subunits are very similar (Figure 1.5right, Figure 1.6a). Each subunit consists of a large extracellular domain, followed by a 7-helices TMD and a cytoplasmic tail [4,8][72]. Detailed structural arrangements in each domain are described in section 1.4.2-1.4.4.

1.4.2 $GABA_BR$ extracellular domain arrangements

The $GABA_BR$ extracellular domain is primarily composed of the VFT, with the only exception being that $GABA_{B1a}$ subunit having sushi domains at the N-terminal to the VFT (Figure 1.6a). Studies so far suggest that only the second sushi domain, i.e. the one closer to the VFT, is compactly folded [73]. Nuclear Magnetic Resonance experiments have shown that this sushi domain appears to have a barrel-like structure stabilised by disulfide bonds (Figure 1.6b) [73].

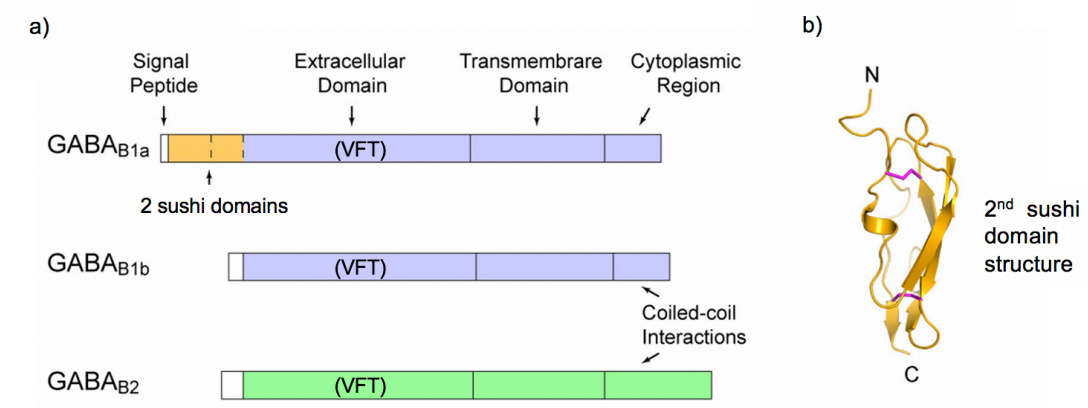


Figure 1.6 Domain organisations of $GABA_BR$.

(a) Schematic representation of the $GABA_{B1a}$, $GABA_{B1b}$ and $GABA_{B2}$. Overall domain organisation is very similar in different subunits, except that $GABA_{B1a}$ has two sushi domains (yellow) at the N-terminus. (b) NMR structure (PDB ID 1SRZ) of the second sushi domain in $GABA_{B1a}$. Disulfide bonds are labelled in magenta.

Crystal structures of the $GABA_{B1b/B2}$ heterodimer VFT domain have been reported [55]. The VFT in both subunits exhibits a bi-lobular structure, consisting of

two lobes of similar size (Figure 1.7a). Each lobe has a central beta sheet, flanked by alpha helices on the outer side [55, 74]. The two lobes are connected by three loops which act as the linker region of the fly-trap [55, 74]. VFT of the $GABA_{B1b}$ and $GABA_{B2}$ subunit display high structural similarities between each other, with a root mean square deviation of 1.3 Å for the upper lobes and 1.1 Å for the lower lobes (Figure 1.7b) [55].

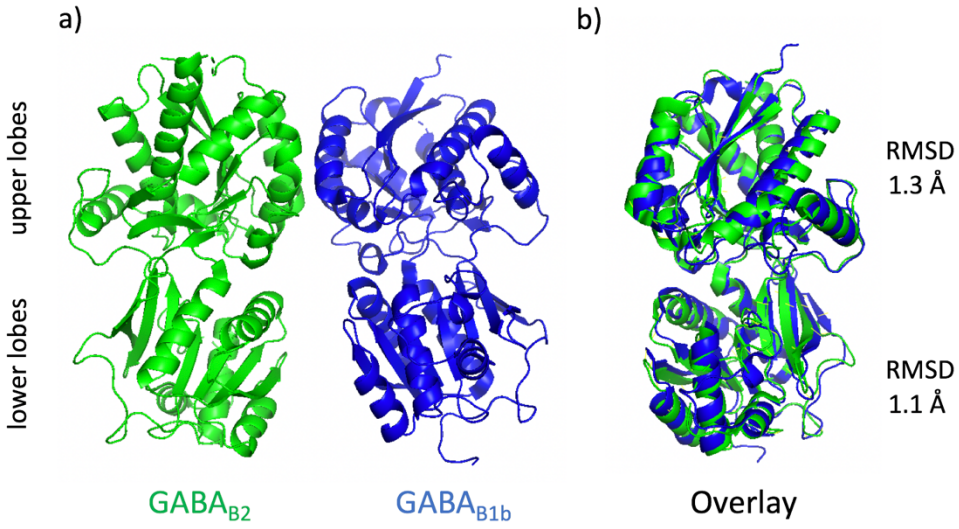


Figure 1.7 VFT domain of $GABA_{B1b}$ and $GABA_{B2}$ displays high structural similarities.

(a) Crystal structure of $GABA_{B1b}/_{B2}$ heterodimer. The VFT in each subunit is made of two lobes connected by three loops. Each lobe contains a central beta sheet flanked by alpha helices on two sides. $GABA_{B1b}$ and $GABA_{B2}$ subunit is coloured blue and green respectively. (b) Superimposition of $GABA_{B1b}$ and $GABA_{B2}$. Structures of the two subunits are highly similar, giving a root mean square deviation of 1.3 Å and 1.1 Å for the upper and lower lobes when superimposed. Figures and calculations were generated in PyMOL [75] using resting state receptor crystal structure (PDB 4MQE).

When forming the heterodimer, VFT domain of the two subunits bind side-by-side to each other while facing in opposite directions [55]. There is an interface between the upper lobes of the two subunits in all conformations (Figure 1.8a), which is believed to be the driving force for extracellular domain dimerisation [55]. The interface is mediated across three regions by helices from both subunits (Figure 1.8b).

One of the regions at the centre (region I, Figure 1.8b) of the interface is critical for dimeric formation and involves mainly hydrophobic interactions [55]. Contacts in region I are highly conserved in all reported $GABA_{B1b/B2}$ structures. The other two regions located at the sides (region II and III, Figure 1.8b) are more variable, involving mainly hydrogen bonding and a few salt bridges [55].

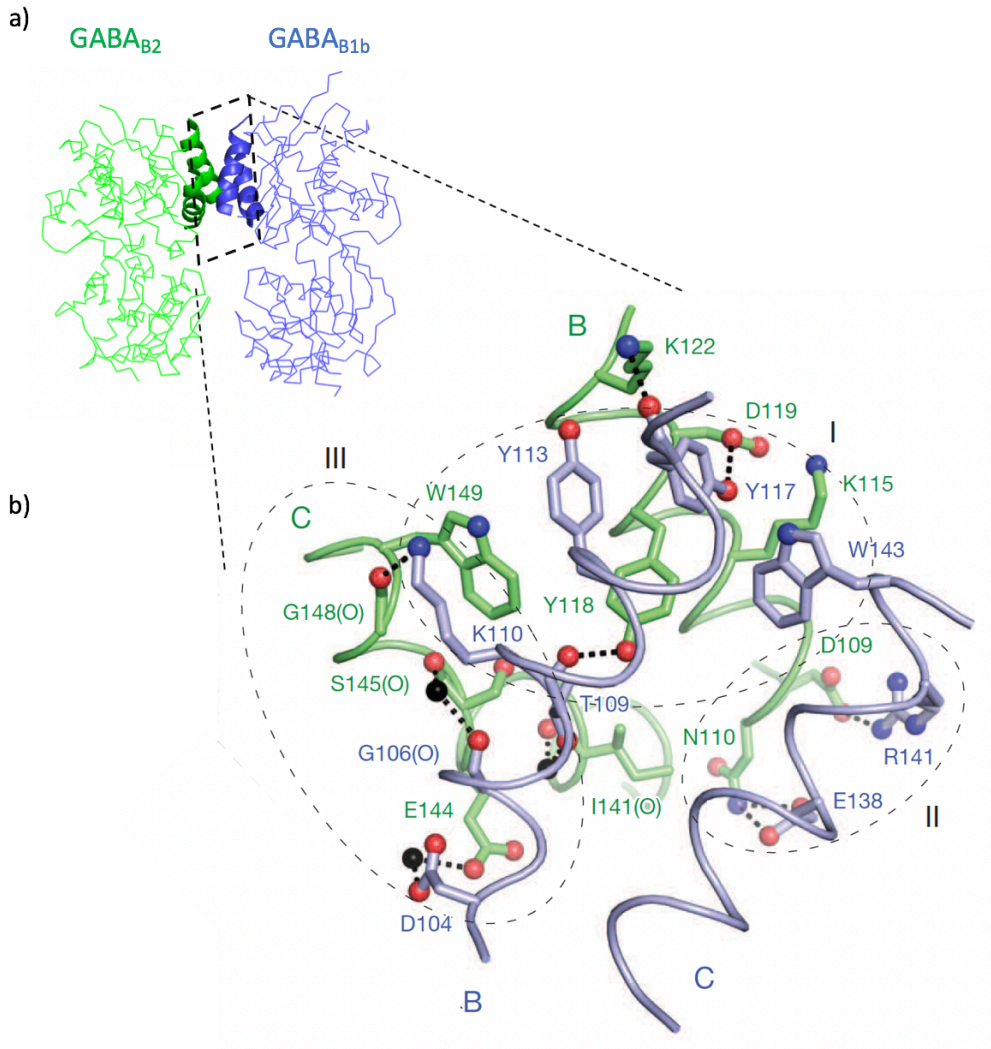


Figure 1.8 Heterodimeric VFT upper lobe interface.

(a) Structure of the $GABA_{B1b}/_{B2}$ heterodimer VFT in resting state. $GABA_BR$ in the resting state forms one inter-subunit interface at the upper lobes (marked by dashed lines). Interfacial interaction is mediated by helices B and C (highlighted as cartoon) of both subunits. (b) Three contact regions at the interface. Region I is located at the centre of the interface. Region I involves mainly aromatic stacking contacts, using residues $GABA_{B1b}$-Tyr^{113}, $GABA_{B1b}$-Tyr^{117} and $GABA_{B2}$-Tyr^{118}, and some hydrogen bonds around the hydrophobic core. Region II and III are found on the sides of Region I. Region II contains several hydrogen bonds and one salt-bridge, and region III is consisted of some hydrogen bonds. Location of each region is marked by dashed line. Hydrogen bonds were indicated by thick dashed line. Figure (a) was generated in pyMOL using crystal structure 4MQE. Figure (b) was adapted from [74].

According to the crystal structures, the lower two lobes only come into close contact when there is an agonist bound to the $GABA_{B1}$VFT (Figure 1.9a) [55]. These observations are in agreement with fluorescence resonance energy transfer assays, where they demonstrated that the $GABA_{B1}$ and $GABA_{B2}$ ectodomains come into close proximity in the activated state [76]. Disruption of the lower lobe contacts using large N-linked glycan prevents agonist-induced receptor activation [77], therefore formation of the lower lobe interface is necessary for $GABA_BR$ activation [55].

Compared to the resting state structures, the $GABA_{B1}$ lower lobe rotates and moves towards the $GABA_{B2}$ upon agonist binding (Figure 1.9a) [55]. While the $GABA_{B1}$ VFT changes conformation in response to agonist binding, the $GABA_{B2}$ shows very little conformational change [55]. The relative movement of the two subunits, driven by $GABA_{B1}$, forms a second inter-subunit interface at the lower lobe in the activated state (Figure 1.9a) [55]. This interface is mediated predominantly by hydrophilic interactions, primarily hydrogen bonds [55]. Again, the interface can be divided to three regions of contacts (Region VI, I and VI, Figure 1.9b). Region IV and V both contains a large cluster of hydrogen bonds, while contacts in region VI are more scattered. It has been suggested that using hydrophilic contacts for the active state conformation will allow this interface to readily dissociate and return the receptor to its resting state [55]. Mutagenesis screenings suggest that certain residues engaged in the interface are critical for receptor activation and initiating G-protein coupling [55]. Key residues involved at the interface include the $GABA_{B1}$ subunit Thr^{198}, Glu^{201} and Ser^{225}, and $GABA_{B2}$ subunit Asp^{204}, Gln^{206}, Asn^{213} and Ser^{223} [74].

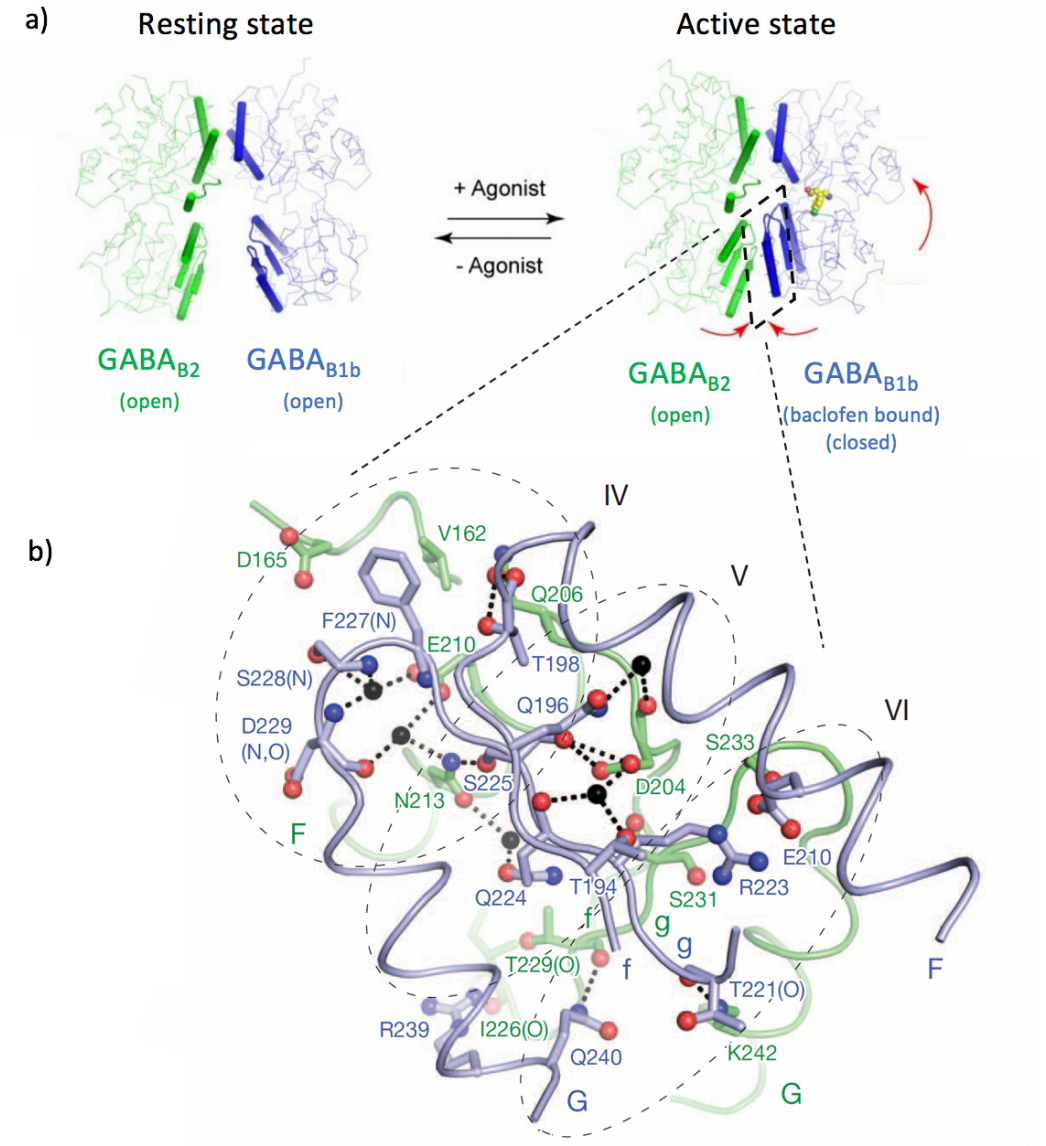


Figure 1.9 Heterodimeric VFT lower lobe interface.

(a) Comparison between resting state and active state. In the resting state crystal structure, only one interface is found between the upper lobes. In the agonist-bound (highlighted as spheres) active state structure, a second interface is formed across the lower lobes while maintaining the upper lobe interface. As compared to the resting state, $GABA_{B1}$ lower lobe displays an upward rotation, coupled by an inward movement towards the $GABA_{B2}$ in the active state. The $GABA_{B2}$ lower lobe displays only a small twist towards $GABA_{B1}$ direction. As a result, the lower lobe interface is created by relative movements between the two subunits, mainly driven by $GABA_{B1}$. The apo crystal structure (PDB 4MQE) and baclofen-bound crystal structure (PDB 4MS4) were used to show the resting state and the agonist-bound active state. Regions involved at the inter-subunit interface(s) are highlighted as cylinders. (b) Interactions at the lower lobe interface. This interface is mediated by two strand-loop-helix motifs from both subunits. The strands are at the centre of the interface, flanked by the helices. Strands and helices involved at the interface are strands f and g and helices H and G from each subunit. The interface can be divided into three regions (marked by dashed line). Region IV contains a cluster of hydrogen bonds mediated mainly through residues $GABA_{B1b}$ Thr^{198}, Ser^{225}, Ser^{228}, Asp^{229} and $GABA_{B2}$ Glu^{210}, Gln^{206}, Asn^{213}. Region V forms another cluster of hydrogen bonds involving $GABA_{B1b}$ Gln^{224}, Thr^{194}, Gln^{196} and $GABA_{B2}$ Asp^{204} and some overlapped residues ($GABA_{B1}$-Ser^{225}, $GABA_{B2}$-Asn^{213}) with Region IV. Region VI has isolated hydrogen bonds between $GABA_{B1}$-Gln^{240} and $GABA_{B2}$-Thr^{229}, and $GABA_{B1}$-Lys^{242} and $GABA_{B2}$-Thr^{221}. Hydrogen bonds were indicated by thick dashed line. Figure was modified from [55].

1.4.3 $GABA_BR$ Transmembrane domain organisation

No structure for the $GABA_BR$ TMD has been reported yet. This region is expected to follow the typical GPCR hepta-helical arrangement, giving a total of 14 transmembrane helices in the heterodimeric receptors. The transmembrane domain is expected to exist as pre-formed heterodimers, as both the extracellular and the intracellular regions of the receptor form stable heterodimers [55, 78, 79]. It has been proposed that closure in the VFT domain induces conformational changes in the receptor's TMD and leads on to activation of the $GABA_{B2}$ for downstream signaling [80]. Structures including the TMD are thus needed to understand the exact mechanisms regarding receptor activation and G-protein interactions.

1.4.4 $GABA_BR$ Cytoplasmic domain organisation

As described before, hetero-dimerisation is required for successful cell surface expression of the functional receptor [58, 78]. This is because the $GABA_{B1}$ cytoplasmic region contains a retention signal ($R^{806}SRR^{809}$) and a di-leucine internalisation motif ($E^{767}KSRLL^{774}$) controlling its trafficking [56, 81, 82]. If expressed alone, $GABA_{B1}$ will be trapped in the endoplasmic reticulum, in order to prevent cell surface display of non-functional $GABA_{B1}$ homomers [56, 81, 82]. The structure of a pair of heterodimeric cytoplasmic coiled-coil has been solved using X-ray crystallography (Figure 1.10) [78]. The structure is 61Å long and 22 Å wide, and follows the characteristic of heptad coiled-coils [78, 83]. Packing of the heterodimer follows the "knobs-into-holes" model [84], where hydrophobic residues are found at the core of the coiled-coil. The $GABA_BR$ heterodimeric coiled-coil displays some specific hydrogen bonds at the coiled-coil interface (Figure 1.10), which cannot be satisfied in a homomeric arrangement [78]. The leucine residues from the internalisation motif are found buried at the core of the

coiled-coil interface [78]. The ER retention signal is not part of the solved coiled-coil structure but is 2 residues down at the C-terminus. It has been predicted that the retention signal is located in proximity to the coiled-coil [78] so that it becomes sterically masked upon association with $GABA_{B2}$, causing the release of heterodimers from the ER [56, 58, 78, 81, 82]. When expressed separately, the cytoplasmic helices from each subunit were found to bind to each other with high affinity, indicating that the $GABA_BR$ cytoplasmic region should dimerise *in vivo* [78].

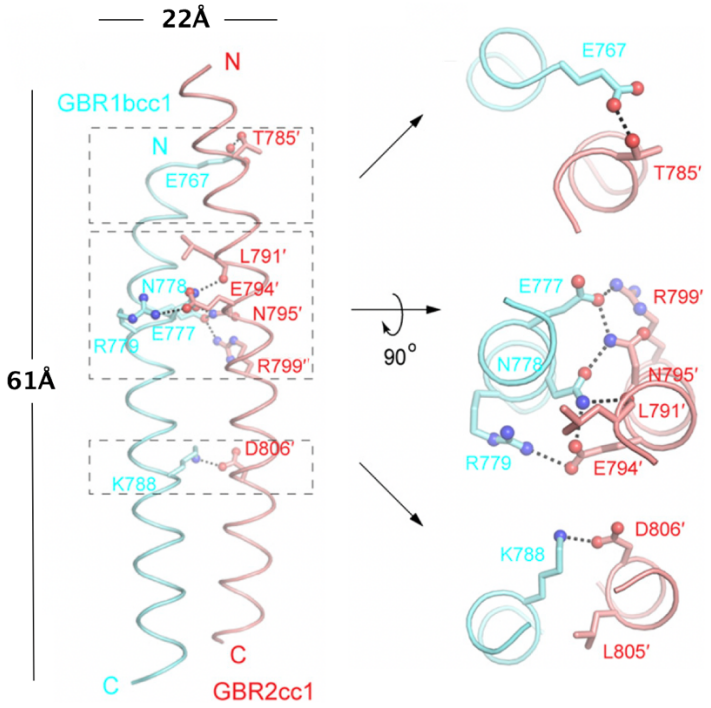


Figure 1.10 Crystal structure of a $GABA_{B1/B2}$ intracellular coiled-coil.

The reported heterodimeric coiled-coil crystal structure ($GABA_{B1}$ and $GABA_{B2}$ C-terminus labelled as GBR1bcc1 and GBR2cc1 respectively) displays characters of a typical heptad coiled-coil. Specific hydrogen bonds are found in the structure that could not be satisfied in a homomeric structure. Specifically, a central network of hydrogen bonds are initiated by $GABA_{B1}$-Asn^{778} and $GABA_{B2}$-Asn^{795}. In addition to hydrogen bond formation with each other, these two residues also form inter-subunit hydrogen bonds with adjacent residues, including $GABA_{B1}$ Asn^{778} with $GABA_{B2}$ Glu^{794} and Leu^{791}, $GABA_{B2}$ Asn^{795} with $GABA_{B1}$ Glu^{777}. There are two additional sets of hydrogen bonds at the each end of the structure. At the N-terminal side using $GABA_{B1}$ Glu^{767} and $GABA_{B2}$ Thr^{785}. At the C-terminal side using $GABA_{B1}$ Lys^{788} and $GABA_{B2}$ Asp^{806}. Figure was modified from [78] .

1.5 Pharmacology of $GABA_BR$

Malfunction of the $GABA_BR$ pathway is associated with a broad range of neurological disorders and neurodegenerative diseases [4][5–7]. Hypoactivity of the $GABA_BR$ system has implications for spasticity, epilepsy, anxiety, depression, and drug addition, whereas hyperactive $GABA_BR$ may cause schizophrenia [5–7]. Due to its physiological importance, there has been a long interest to study the possibility of targeting the $GABA_BR$ for clinical treatments. Many preclinical and clinical studies have shown therapeutic potential of $GABA_BR$ ligands [85]. For example, the use of agonists for treating drug abuse and anxiety [86, 87] and antagonists to improve cognitive performance, depression and Down syndrome [85]. However these ligands often cause undesirable side effects due to the wide distribution and highly conserved ligand recognition site of the receptor. More recently, $GABA_BR$ allosteric modulators were designed to avoid side effects associated agonists and antagonists, because allosteric modulators cannot activate the receptors on their own but can affect the potency and efficacy of agonists/antagonists [88, 89]. Sections 1.5.1 to 1.5.3 below described the development of main $GABA_BR$ ligands and their potential clinical applications.

1.5.1 $GABA_BR$ agonists

Historically, $GABA_BR$ is pharmacologically defined by its insensitivity to the $GABA_AR$ antagonist bicuculline and their sensitivity to the GABA lipophilic analogue baclofen (Figure 1.11) [85]. Baclofen is thus the first identified and best characterised $GABA_BR$ agonist. Since then, many agonists have been designed based on the GABA backbone (Figure 1.11), which have much higher potencies (IC_{50} in the low nanomolar range) than the original GABA and baclofen (Table 1.3) [90]. These agonists display a

number of pharmacological properties, including muscular relaxation, inhibition of gastric acid secretion, antitussive effects, suppression of drug-addiction and alcohol-abuse, antinociception, and so on [13, 90]. At the moment, baclofen and gamma-hydroxybutyrate (GHB) are the only FDA-approved drugs targeting the $GABA_BR$. Baclofen is used as a muscle relaxant in treating spasticity [15] and GHB is used for the treatment of narcolepsy [91]. Many other drugs are in the pipeline, such as lesogaberan (AZD-3355) for treating gastro-esophageal reflux disease [92].

1.5.2 $GABA_BR$ antagonists

The first wave of $GABA_BR$ antagonists included phaclofen and 2-hydroxysaclofen [93]. These drugs were again derived from the GABA molecule, where the carboxylic acid is replaced with sulfonic acids or phosphinic acids (Figure 1.11). While most $GABA_BR$ agonists and antagonists share a similar backbone, the size of the substituent placed on the carboxylic acid end is a key factor in determining what kind of activity it induces. In general, as the substituent grows in size, the molecule shifts from agonist, to partial agonist, and to antagonist. (This size-activity relationship is elaborated upon using structural information in section 1.6.) The initial set of designed antagonists had relatively low affinities with IC_{50} in the micromolar range (Table 1.3). In 1996, Froestl and colleagues substituted the amino end with benzyl groups (Figure 1.11) [94, 95] and achieved much higher affinity in low nanomolar range (Table 1.3). Many analog antagonists have followed since then. $GABA_BR$ antagonists have been shown to lower seizure activity, reduce depression, and facilitate cognitive functions in animal models [90].

1.5.3 $GABA_BR$ allosteric modulators

Substantial efforts have been made in the search for $GABA_BR$ allosteric modulators.. Allosteric modulators have great therapeutic potential in that they possess fewer side effects and are less likely to develop receptor desensitization and drug-tolerance than the traditional agonists [13, 95, 96]. Several classes of $GABA_BR$ allosteric modulators with different chemical scaffolds have been described (Figure 1.11). These molecules are all suggested to interact with the transmembrane region of the $GABA_{B2}$ subunit [34, 95, 97, 98]. The structure-activity relationships of allosteric modulators are less well defined, partially due to the lack of structural information in the transmembrane region.

Cations have also been reported to act as allosteric modulators for $GABA_BR$. Specifically, Ca^{2+} ions act as positive allosteric modulators [99, 100] for GABA by binding to $GABA_{B1}$ extracellular VFT and stabilising the agonist-bound conformation [99]. $GABA_{B1}$ Ser^{269}, a residue flanking the orthosteric binding site, has been found to be critical for the Ca^{2+} sensitivity of the receptors and its allosteric effects on ligand binding [99]. Under physiological conditions, the orthosteric binding site is saturated with extracellular Ca^{2+} ions [99]. It remains to be studied whether the depletion of Ca^{2+} ions under pathological conditions, such as epileptic seizure [101], affects the function of the receptor [102, 103].

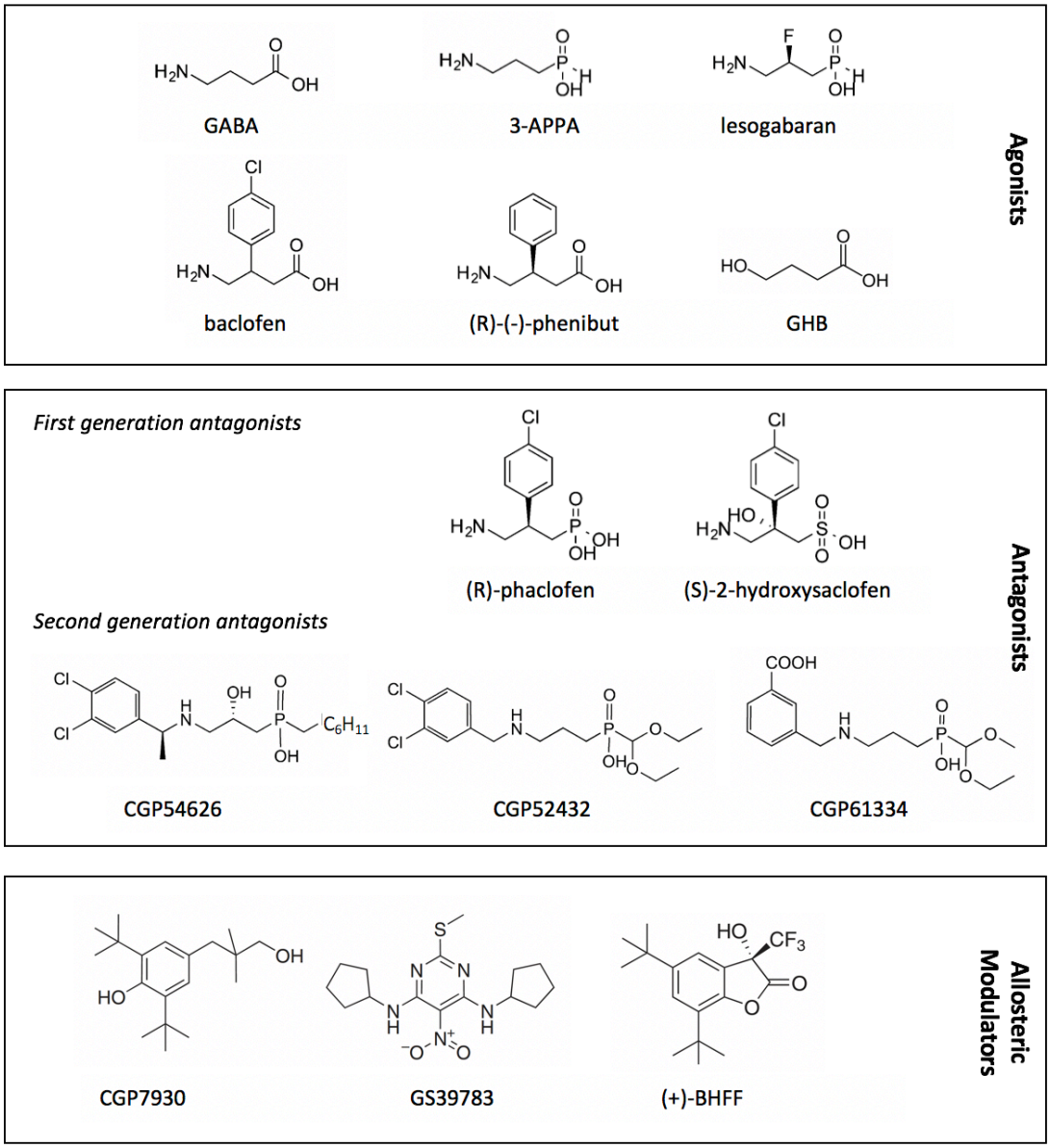


Figure 1.11 Chemical structures of representative $GABA_BR$ agonists, antagonists and allosteric modulators.

Table 1.3 Binding affinities of representative $GABA_BR$ ligands.

Ligand	Function	Affinity μM (species) *	Reference
baclofen	Agonist	0.06 (rat brain)	[104]
(R)-phenibut	Agonist	25 (rat brain)	[105]
lesogaberan	Agonist	0.008 (EC_{50}) (rat brain)	[106]
GHB	Agonist	79-126 (K_i) (recombinant receptor in *Xenopus* oocytes)	[107]
3-APPA	Agonist	0.002 (cat brain)	[104]
(R)-phaclofen	Antagonist	76 (rat brain)	[108]
2-OH-saclofen	Antagonist	11 (rat brain)	[109]
CGP54626	Antagonist	0.004 (rat brain)	[110]
CGP52432	Antagonist	0.055 (rat brain)	[110]
CGP61334	Antagonist	0.036 (rat brain)	[110]
SCH50911	Antagonist	1.1 (rat brain)	[111]
CGP7930	Allosteric Modulator	30 μM CGP7930 enhances GABA both the potency by 6-fold (Recombinant receptor in CHO cells)	[112]
(+)-BHFF	Allosteric Modulator	GABA EC_{50} is increased by 87-fold in presence of 0.3 μM BHFF (Recombinant receptor in CHO cells)	[113]
GS39783	Allosteric Modulator	30uM GS39783 enhances GABA both the potency by 8-fold (rat brain)	[114]
Ca^{2+} ions	Allosteric Modulator	Affinity for GABA and 3-APPA increased by 10-fold in presence of 1.8 mM Ca^{2+} (rat brain)	[99]

* Binding affinities as IC_{50} unless otherwise stated

1.6 $GABA_BR$ ligand recognition mechanisms

1.6.1 Structural plasticity of the $GABA_BR$ VFT

Structures of the $GABA_BR$ VFT have been solved using X-ray crystallography for $GABA_{B2}$ monomer and $GABA_{B1b/B2}$ heterodimer (Table 1.4) [55, 74].

Table 1.4 Existing crystal structures of $GABA_B$ VFT.

Domain	State	PDB ID	Reference
Heterodimeric VFT	Agonist-bound	4MS3, 4MS4	[55]
Heterodimeric VFT	*Apo*	4MQE	[55]
Heterodimeric VFT	Antagonist-bound	4MQF, 4MR7, 4MR8, 4MR9, 4MRM, 4MS1	[55]
$GABA_{B2}$ VFT	N/A	4F11, 4F12	[74]

The more recently reported heterodimeric structures had the VFT domain crystallised in multiple conformations [55]. These structures have greatly improved our understanding of $GABA_BR$'s interactions with drugs and the receptor's ECD activation mechanism. As confirmed by the crystal structures, $GABA_BR$ in the open

conformation correspond to the resting state, while the closed conformation represents activated receptors (Figure 1.12). Section 1.6.2 and 1.6.3 elaborate on the details of ligand recognition mechanisms in different states.

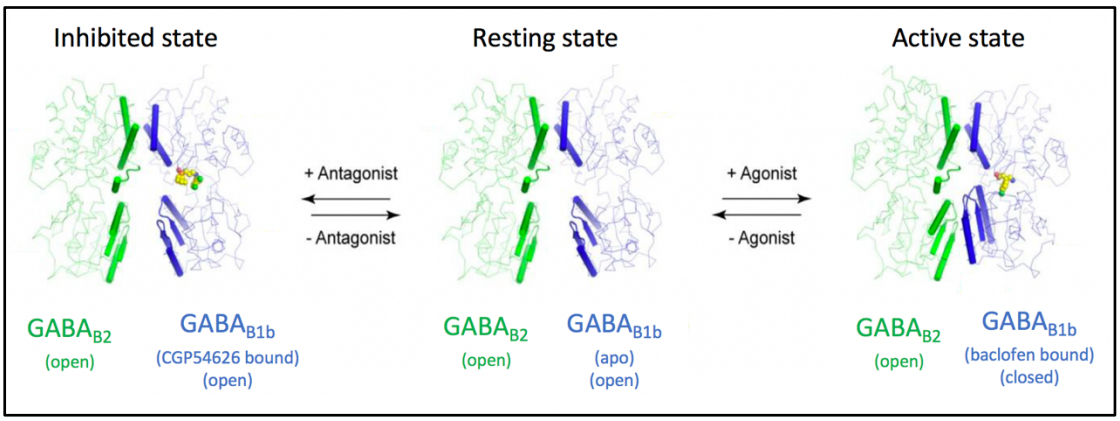


Figure 1.12 GABA$_B$R VFT oscillates between open and closed conformations.

In the resting state (middle), GABA$_{B1}$ subunit adopts an open conformation and interacts with GABA$_{B2}$ subunit at the upper lobe. This open conformation is stabilised by binding of antagonists to the orthosteric site (left). Agonist binding to the same site induces movement of the GABA$_{B1}$ lower lobe to form a close conformation (right). As a consequence, an additional interface is created between the two subunits at the lower lobe. Ligands are marked as spheres. Figure was modified from [55].

1.6.2 GABA$_B$R antagonist recognition

The *apo* crystal structure (i.e. ligand-free structure) of the heterodimeric VFT domain represents receptors in the resting state. Both the GABA$_{B1b}$ and GABA$_{B2}$ subunits have their crevices "open", when compared to the structures of mGluRs [55, 115-117]. All antagonist-bound structures reported so far look nearly identical to the *apo* resting state structure (Figure 1.13a). In these structures, antagonist molecules are bound at the GABA$_{B1}$ VFT crevice where the molecule interacts predominantly with GABA$_{B1}$ and very little with GABA$_{B2}$ (Figure 1.13b). Only certain antagonists (e.g. CGP54626 and SCH50911) can make direct contact with the GABA$_{B2}$ subunit using their large substituents (Figure 1.13c) [55]. Antagonists share a common set of hydrogen bonding interactions with GABA$_{B1}$: interacting with GABA$_{B1}$ residues Ser^{130}

and Ser^{153} at the GABA carboxyl group equivalent end, and with His^{170} and Glu^{349} at the GABA amino group equivalent end (Figure 1.13b) [55]. The Try^{65} residue also makes a universal Van der Waals interaction with all antagonists at their three-carbon linker region. As suggested in the crystal structures, the bulky substituent groups in antagonists stabilise the $GABA_{B1}$ VFT in its open conformation. As a result, the two lower lobes stay separated from each other and away from the $GABA_{B2}$ crevice.

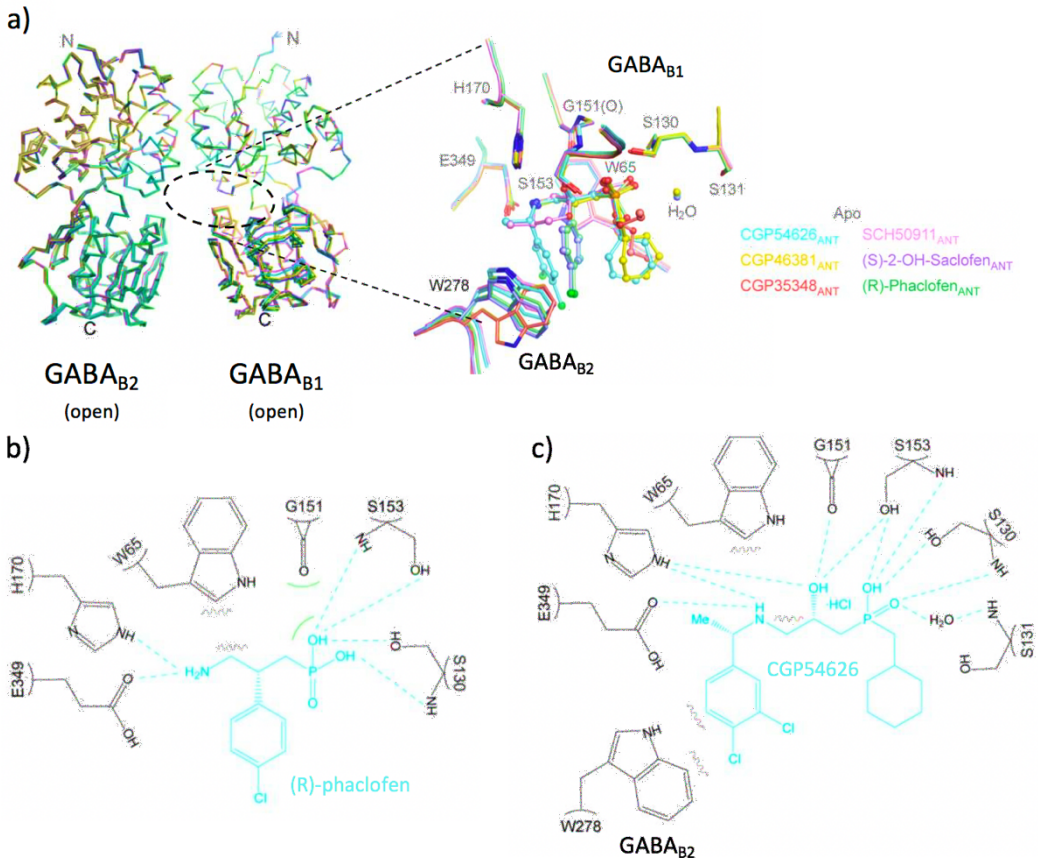


Figure 1.13 Antagonist recognition mechanism.

(a) Superimposition of *apo* and antagonist-bound structures. The structures show high similarities to each other. Antagonists predominantly interact with B1 subunit, with only CGP54626 and SCH59811 making direct contacts with B2. Alignment shown here was performed based on the $GABA_{B1}$ subunit. (b) Schematic diagrams of antagonist (R)-phaclofen binding at the $GABA_{B1}$ cleft. The amino group end of (R)-phaclofen forms hydrogen bonds with $GABA_{B1}$ Glu^{349} and $GABA_{B1}$ His ^{170}. There is a conserved van der Waals interaction site found in all antagonist and is mediated through $GABA_{B1}$ Trp^{65}. The acid group at the other end forms more hydrogen bonds with $GABA_{B1}$ Ser^{153} and $GABA_{B1}$ Ser^{130}. (c) Schematic diagrams of antagonist CGP54626 binding at the $GABA_{B1}$ cleft. In addition to the contacts described in b, CGP54626 makes more hydrogen bonds with $GABA_{B1}$ using its beta-hydroxyl substituent. CGP54626 also makes direct contact with $GABA_{B2}$Trp^{278} through its large amino-end substituent. Hydrogen bonds are shown as dashed lines in cyan. Hydrophobic interactions are marked by grey wiggled lines. Polar interactions are shown as green curved lines. Figure was modified from [55].

1.6.3 $GABA_BR$ agonist recognition

The binding modes of the endogenous GABA molecule and its derivate agonist baclofen are very similar (Figure 1.14a). Agonist recognition is again achieved via hydrogen bonds at each end of the molecule (Figure 1.14b): $GABA_{B1}$ Ser^{130} and $GABA_{B1}$ Ser^{153} at the carboxyl end, and $GABA_{B1}$ His^{170} and $GABA_{B1}$ Glu^{349} at the other end [55]. Agonists also share the Van der Waals contacts with $GABA_{B1}$ Trp^{65} [55].

Interaction with an agonist triggers movement of the $GABA_{B1}$ VFT and allows its "closure" due to the agonist's smaller size. This change in conformation brings an agonist into direct contact with the $GABA_{B2}$ subunit. Interactions with $GABA_{B2}$ enhance the binding affinity of an agonist with the receptors [34]. The agonist forms hydrogen bonds with $GABA_{B2}$ Tyr^{250} and strong Van Der Waals interactions with $GABA_{B2}$ Trp^{278} (Figure 1.14b). The $GABA_{B2}$ Trp^{278} residue has been found to display different rotameric adaptations in the crystal structure in order to accommodate different ligands (Figure 1.14a) [55]. This conformation flexibility provides a mechanism for $GABA_BR$ ligand binding specificity and affinity [55].

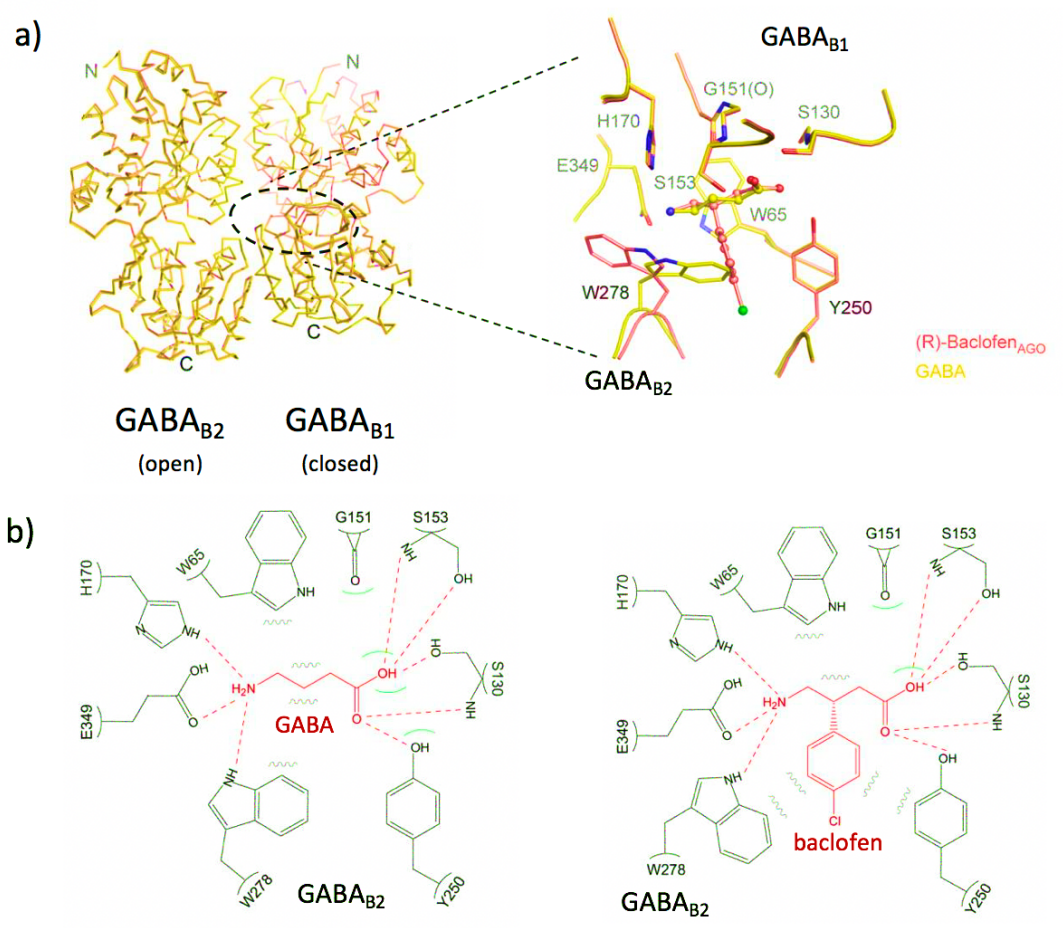


Figure 1.14 Agonist recognition mechanism.

(a) Superimposition of GABA and baclofen bound structures. The two molecules showed similar binding mode with the receptor, making contacts with both subunits. GABA$_{B2}$ Trp^{278} residue is found to adopt a different conformation in the baclofen bound structure to accommodate its chlorophenyl group. (b) Schematic diagrams showing interactions of GABA and baclofen at the binding site. The two agonists share a common set of hydrogen bonds with the receptor. Agonist interacts with GABA$_{B1}$ Glu^{349}, His^{170} and GABA$_{B2}$ Trp^{278} at the amino-group end of the molecule. At the carboxyl end, agonist forms hydrogen bonds with GABA$_{B1}$ Ser^{153}, Ser^{130} and GABA$_{B2}$ Tyr^{250}. Agonist also forms hydrophobic interactions with GABA$_{B1}$ Trp^{65} and GABA$_{B2}$ Trp^{278}. Hydrogen bonds are shown as dashed lines in red. Hydrophobic interactions are marked by grey wiggled lines. Polar interactions are shown as green curved lines. Figure was adapted from [55].

1.7 Aims

Our current knowledge about the molecular structure of GABA$_B$ receptors is limited to the extracellular region and a short intracellular region [8,9]. Crystal structures of the extracellular domain VFT have provided information regarding the mode of ligand recognition and its impact on the ectodomain conformation [9]. It has

been proposed that transitions in the VFT upon agonist binding decrease the distance between the membrane-proximal domains in the two subunits [55], which may lead to changes in the relative orientation of the transmembrane regions for full receptor activation. The integral receptor structure is thus needed to gain insight into the mechanisms of full receptor activation and subsequent signal transduction.

The aim of this thesis is to investigate the structures of human full-length $GABA_BR$. Obtaining high-resolution structures for GPCRs has been particularly challenging. Only very few full-length receptor structures were reported when this thesis commenced. Traditional methodology (i.e. X-ray crystallography) requires highly concentrated amount of receptor proteins locked in one uniform conformation. With their large extracellular region, Class C GPCRs are likely to exhibit more complicated conformational variations making them very difficult targets to obtain high quality crystals. Recent developments in Cryo-EM and single particle analysis provide an alternative approach for studying GPCR structures.

Chapter 3 of this thesis sets out to establish a pipeline for expressing the heteromeric full-length $GABA_BR$. Chapter 4 describes the works in searching for suitable purification protocols for samples preparation for Cryo-EM studies. Finally, Chapter 5 and 6 analyses and discusses Cryo-EM datasets of $GABA_BR$ reconstituted in amphipols and lipid nanodiscs.

Chapter 2. METHODS

2.1 Molecular Biology

2.1.1 Construct design

Three $GABA_B$ constructs were designed to generate the $GABA_{B1a}$, $GABA_{B1b}$ and $GABA_{B2}$ subunits. Domain boundaries were estimated based on existing structural information [55, 78]. The $GABA_{B1a}$ longer variant (Uniprot entry Q9UBS5) was split into two synthetic constructs: construct B1_sushi_EB containing natural secretion signal sequences and the sushi domain (residues 1-164), and construct B1_noS_BRIL_AK containing the rest until the intracellular coiled-coil (residues 165-920). Joining the two constructs together, using restriction site BsrGI, yielded the complete $GABA_{B1a}$ variant. The B1_noS_BRIL_AK construct alone is equivalent to the $GABA_{B1b}$ subunit. The $GABA_{B2}$ subunit construct B2_BRIL_EK (Uniprot entry O75899, residues 1-827) started from the natural secretion signal peptide until the coiled-coil at the receptor C-terminus. The Kozak consensus sequence (GCCACC) was inserted before the natural secretion signal sequence in $GABA_{B1a}$ and $GABA_{B2}$ subunits. No point mutations were introduced in the initial synthetic sequences. One thermostablised apocytochrome b_{562}RIL (PDB ID: 4EIY, residues 234-339) protein was fused at the hypothetical intracellular loop region on each subunit (Residues 794-795 in $GABA_{B1}$ subunit and residues 681-682 in $GABA_{B2}$ subunit). All sequences were codon-optimised for expression in mammalian cells. Synthetic complementary DNAs were ordered from Eurofins Scientific. Sequences of the synthetic constructs are shown in Appendix B.

2.1.2 Polymerase Chain Reactions (PCR)

The $GABA_{B1a}$ variant, and all subsequent constructs used for protein expression optimisation experiments, were generated by polymerase chain reactions (PCR) using

the synthetic DNA as template. Primers used in PCR experiments in this thesis are summarised in Appendix B. For each PCR reaction (50 µl in total), 1 µl (10 pmol/µl) forward primer, 1 µl (10 pmol/µl) reverse primer, 1 µl DNA template (100 ng), 5 µl 10x buffer (Takara), 4 µl deoxyribonucleotide triphosphates (Takara), 37.9 µl milliQ water and 0.1 µl Pyrobest polymerase (Takara) was used. PCR was performed following the generic protocol below, with settings adjusted for each particular experiment:

Step.1 Activation at 95 °C for 20 s

Step.2 DNA denaturation at 95 °C for 10 s

Step.3 Primer annealing at 55-70 °C (depending on GC contents) for 30 s

Step.4 Elongation at 72 °C for 1 min per 500 bp (product length)

Repeat step 2-4 for 30 cycles

Step.5 Extension at 72 °C for 10 min

Overlap-PCR were performed following the same protocol, where two fragments were first generated separately. The full template was then generated by fusing the two fragments together following step 1-5 for 3 cycles without primers, after which primers were added for another 30 cycles.

PCR products were cleaned using QIAquick spin column (Qiagen) to remove unwanted oligos (<100 bp), enzyme and dNTPs. PCR products were resuspended in 250 µl PB binding buffer (Qiagen) and centrifuged at 20,000 *x*g, for 30 sec in the spin column. Column filter was washed with 500 µl PE wash buffer (Qiagen) by centrifugation at 20,000 *x*g for 30 sec. Column was transferred into fresh eppendorf tube and DNA was eluted with 40 µl MilliQ water.

2.1.3 Cloning and transformation

PCR products were sub-cloned into various versions of the pHLsec vectors [118] for transient transfection. Vectors incorporating new affinity tags had been created to adapt to the needs of this project. Plasmids generated and used in this thesis are summarised in Appendix B. Inserts were cloned into the pHLsec vectors using sites EcoRI/KpnI for $GABA_{B1a}$ subunit, *AgeI/KpnI* for $GABA_{B1b}$ subunit and sites *EcoRI/ XhoI* for $GABA_{B2}$ subunit, unless otherwise stated. Inserts were cloned into the pHR-CMV-TetO2 vector using *EcoRI/ XhoI sites.* For cloning into the *pHRSIN _CSGW_vc vector,* insert in the pHLsec cassette between the *EcoRI/ XhoI sites* were replaced with restriction sites *NotI /MluI* following cloning into the vector.

Restriction digestions were prepared as follows: 0.3 µl of each restriction enzyme (NEB), 2 µl enzyme buffer, 3 µg DNA, topped up with MilliQ water to prepare a 20 µl reaction mix. Digestion was performed at 37 °C for 1 hr. Digestion products were run on 1.5% agarose gel for DNA separation. DNA band (of the target product length) was excised and purified using the QIAquick Gel Extraction Kit (Qiagen).

A 20 µl ligation mix was prepared using 10 µl 2x ligation buffer (NEB), 0.5 µl quick ligase (NEB), 9.5 µl DNA mix (at 3:1 molar ratio of vector to insert). Ligation reaction was set at room temperature for 10 min and stopped by placing the reaction tube on ice. 3 µl ligation product was transformed into 30-40 µl DH5α library efficient competent cells (Invitrogen) by heat shocking the bacteria at 42 °C for 40 sec. Heat shocked cells were put back on ice immediately after heat shock for 10 min to recover. Cells were supplemented with 100 µl SOC medium and left to shake at 37 °C, 200 rpm for 1 hr. Cells were grown on LB-agar plates containing antibiotics of choice at 37 °C

overnight. Individual bacteria colony was picked and grown in 6 ml LB culture at 37 °C, 200 rpm overnight. Bacteria were collected using centrifugation and DNA was purified using Qiagen Miniprep Kit following manufacturer's instruction. All constructs used have had the target insert sequence verified.

2.2 Mammalian Cell culture

2.2.1 Growing and passaging of adherent cells

Adherent HEK293 cells were maintained in standard T175 flasks in expression medium Dulbecco's Modified Eagle's Medium (DMEM, Sigma) supplemented with 1% L-glutamine (Gibco), 1% non-essential amino-acids (Gibco) (NEAA) and 10% fetal calf serum (Gibco) (FCS) at 37 °C with 5% CO_2. Cells were passaged as follows:

Step.1 Expression medium was removed and cells were washed with 20 ml phosphate-buffered Saline (Gibco) (PBS)

Step.2 Cells were treated with 4ml trypsin (Gibco) for 5 min at 37 ˚C

Step.3 Trypsinisation was stopped by addition of 21 ml expression medium

Step.4 Cells were gently resuspended and desired amount was transferred into new flasks.

Typically, cells were passaged every 3-4 days and kept under 90% confluency.

2.2.2 Growing and passaging of suspension cells

One T175 flask of confluent adherent cells was trypsinized, and resuspended in 20 ml Freestyle™ 293 Expression medium (Gibco) supplemented with 10% fetal calf serum (FCS, Gibco). Cells were centrifuged at 100 *x*g, for 10 min at room temperature. Supernatant was removed and cells were gently resuspended in 25 ml Freestyle

medium containing 1 % FCS, 1% NEAA, 1% L-glutamine. Cells were transferred into sterile 500 ml bottles/flasks for growth in suspension. 155 ml of freestyle medium containing 1 % fetal calf serum, 1% NEAA, 1 % L-glutamine was added and cells were left to grow for 3-4 days. Shaking incubator for suspension culture was set at 37°C, 8% CO_2, 130 rpm unless otherwise stated. Cells were split 1-in-2 with passaging medium Protein Expression Media (PEM, Invitrogen) supplemented with 1% L-glutamine, 1% NEAA and 1% fetal calf serum once 2 million/ml cell density was reached. Cells were split 1-in-2 to 1-in-6 with passaging medium as desired from here on.

2.2.3 DNA plasmids preparation for transient transfection

Plasmids for cell transfection were replicated using *E. coli* DH5α cells. Plasmids were purified using the Endotoxin-Free Plasmid Mini/Mega/Giga Kit (Qiagen) depending on the scale of the experiment. DNA for transfection was cleaned with chloroform immediately before transfection experiments.

2.2.4 Transient transfection and expression in adherent cells

Adherent cells were seeded in 6-well plates (Sigma) and grown until 90% confluency. Transfection cocktail (per well) was prepared by mixing 250 µl DMEM containing 6 µl Lipofectamine (Invitrogen) and 250 µl DMEM containing 3 µg DNA. Transfection cocktail was incubated at room temperature for 20 min and added to the cells. Cells were kept to express for 48–72 hr after transfection. Surface trafficking of C-terminal mVenus tagged constructs was checked by fluorescence microscopy 48 hr post transfection. Cells were washed twice with phosphate buffered saline and harvested. Cells were used immediately for purification experiments or snap-frozen for storage at -80 °C.

2.2.5 Transient transfection and expression in suspension cells

HEK293S-GnTI$^-$ cells [118] were grown to 2-3 million/ml density in suspension for transient transfection. Cells from 800 ml cultures were collected by centrifugation at 1400 rpm for 7min at room temperature. While cells were spinning, transfection cocktail was prepared by mixing 2 x 10 ml freestyle medium: one containing 0.4 mg DNA and one containing 1.2mg PEI Max (Polysciences) supplemented with 1% NEAA, 1% L-glutamine. After centrifugation, expression medium was removed and cell pellets were gently re-suspended in the 20 ml transfection cocktail. Cells in transfection cocktail were transferred into conical flask and topped up with 100 ml freestyle medium supplemented with 1% NEAA, 1% L-glutamine. At this point, 1mM final concentration valproic acid (Santa Cruz Biotechnology) was added. Cells were transfected for 3–6 hr at 160 rpm. Cells were replenished with 950 ml passaging medium for expression (other supplements or additives can be added at this point) at 130 rpm. Cell pellets were collected 48–72 hr post-transfection and snap-frozen for storage at −80 °C.

2.2.6 Lentiviral cell lines generation and protein expression

All procedures described below from virus particles generation, to cell infection, were carried out at Containment Level 2 following local safety requirements and guidelines.

Replication deficient HIV-derived retroviral vector particles were generated using HEK293 Lenti-X cells (Clontech). Cells were transfected with the transfection plasmid mix consisting of three vectors (10 µg each). Viral particles for constitutive expression were generated using: transducing vector pHRSIN-CSGW_vc, packaging vector pCMV-dR8.91 and envelope vector pMD2.G (all three vectors were kindly

provided by Dr. Veronica Chang). Viral particles for inducible expression were generated using transfer vector pHR-CMV- TetO2, envelope vector pMD2.G and packaging plasmid psPAX2 (these plasmids were kindly provided by Dr. Jonathan Elegheert). To infect lenti-X cells, transfection plasmid mix was resuspended in DMEM/F-12 medium (Sigma) in a total volume of 250 µl. 75 µl of PEI (1mg/ml) was resuspended in another 250 µl DMEM/F-12 separately. Transfection plasmids and polyethylenimine (PEI, Merck) were mixed together by briefly vortexing and left to incubate for 15 min.

Adherent cells were grown following protocols described in section 2.2.2, with the exception of having used DMEM/F-12 medium instead of DMEM. Cells were seeded onto a T75 flask (to achieve approximately 90% confluency on the day of transfection). Growth medium was removed and cells were washed with PBS. 11.5 ml DMEM/F-12 supplemented with 2% FBS was added, followed by the 500 µl plasmid/PEI mix prepared earlier. The flask was gently tilted to make sure all cells were covered in medium. Lenti-X cells were left to generate the viral particles for 3 days. Medium containing the lentiviral particles was collected and diluted with 12ml DMEM/F-12/2%FCS medium. The diluted medium was filtered through a 0.45 µm sterile filter (Millipore). 24 µl of polybrene (10mg/ml, Sigma) was added to the filtered medium and mixed gently, giving the infection medium. For infection, either HEK293 GnTI$^-$ cells (for constitutive expression) or HEK293 GnTI$^-$ TetR (for inducible expression) cells were used. HEK293 GnTI$^-$ TetR cells were maintained identically, except for the addition of 5 µg/mL blasticidin S hydrochloride (ThermoFisher) in the medium. The expression medium was removed from cells to be infected with viral particles, cells were washed by PBS and supplemented with 12 ml infection medium. Cells were left

in the infection medium for 3-4 days to allow the target genes to be incorporated into the host cell genome. Transduction efficiency of fluorescently-tagged constructs in HEK293 GnTI$^-$ cells was assessed by fluorescence microscopy.

The stable cell line was washed at least three times with PBS to make sure all trace of virus particles were removed. After this point cells were maintained as normal adherent cells. Aliquots of the stable cell lines were supplemented with DMSO (Merck) and cryopreserved in liquid nitrogen for long term storage. Stable cell lines were expanded and put in suspension in the same condition as described in section 2.2.4. For HEK293 GnTI$^-$ cells (constitutive expression), cells were harvested once the desired biomass was reached. For HEK293 GnTI$^-$ TetR cells, cells were grown to the desired volume and induced with doxycyclin (Merck, final concentration of 10 µg/ml). HEK293 GnTI$^-$ TetR cells were left to express for 48-72 hr before harvest.

2.2.7 Stable cell line expression in WAVE bioreactors

1 l of suspension cells at 2 million/ml density was added to a pre-inflated 10 l Cellbag (GE healthcare) connected to a WAVE bioreactor (GE healthcare). Cells were diluted with 2 l expression medium and maintained at 37 °C with 20 rpm at 6° angle. Cells were expanded using fresh medium once cell density reached 2 million/ml. Cells were left to grow until the total desired volume was reached. Cells were collected by centrifugation and snap-frozen for storage at -80°C.

2.3 Bacterial protein expression and purification

2.3.1 Expression and purification of MSP proteins

Bacteria were recovered from glycerol stocks (kind gifts from Ms. Christiane Kowatsch, protein sequences in Appendix B) in 5 ml LB containing 30 µg/ml Kanamycin.

This starting culture was grown for 8 hr and put into 250 ml of LB (supplemented with 30 µg/mL Kanamycin) for overnight growth at 37 °C, 200 rpm. 42 ml of the overnight culture was added to 600 ml of TB (supplemented with 30 µg/mL Kanamycin) in 2.5 l baffled flask. Culture was grown at 37 °C, 220 rpm until OD_{600nm} reaches 2.5 – 3.0. Culture was cooled to 28°C and induced with 1mM (final) IPTG. Bacteria are left to express for 3-4 hr at 28 °C, 220 rpm before harvesting. Cells pellets were collected by centrifugation and kept at -80 °C for long-term storage. Alternatively, 10 g of fresh cells were resuspended in 50 ml of lysis buffer (20mM sodium phosphate, pH7.4, 1% Triton-X100, EDTA-free protease inhibitor cocktail). Cells were lysed by sonication and centrifuged at 30,000 *xg* for 30 min at 4°C. The supernatant containing the MSP protein was collected and incubated with 2.5 ml Ni-NTA resin (Thermo Scientific) for 1.5 hr, 65 rpm, 10 °C. Beads were collected and washed in the subsequent steps using following solutions: 25 ml of 40 mM Tris (HCl) pH 8.0, 0.3 M NaCl, 1% Triton; 15 ml of 40 mM Tris (HCl) pH 8.0, 0.3 M NaCl, 50 mM cholate; 25 ml of 40 mM Tris (HCl) pH 8.0 and 0.3 M NaCl and finally 50 ml of 40 mM Tris (HCl) pH 8.0, 0.3 M NaCl, 50 mM imidazole for MSP1D1 or 20 mM imidazole for MSP2N2. MSP proteins were eluted with 40 mM Tris (HCl), 0.3 M NaCl, 0.3 M imidazole. Proteins were dialysed into 20mM Tris (HCl) pH7.4, 0.1M NaCl, 0.5mM EDTA overnight at 4 °C with gentle stirring. Dialysed protein was concentrate to 8-10 mg/ml and snap frozen for long-term storage at -80 °C.

2.3.2 Expression and purification of saposin proteins

Saposin A and saposin D (sequences in Appendix B) were cloned into the pET-45b+ vector using restriction enzyme sites *MluI/XhoI* (vector map in Appendix B). The

final construct contained a N-terminal His_{10x}-tag and 3C cleavage site. The plasmid was transformed into Rosetta Gami-2 (DE3) cells for protein expression. Cells were grown on LB plates supplemented with 34 μg/ml Chloramphenicol, 50 μg/ml Streptomycin, 12.5 μg/ml Tetracycline and 100 μg/ml Ampicillin overnight at 37 °C. A colony was picked to grow in 1x TB start culture supplemented with 1% glucose, 34 μg/ml Chloramphenicol, 50 μg/ml Streptomycin, 12.5 μg/ml Tetracycline and 100 μg/ml Ampicillin at 250 rpm, 37 °C until OD_{600nm} reached 0.5. 18ml of the start culture was added into 600ml 1x TB supplemented with 34 μg/ml Chloramphenicol, 50 μg/ml Streptomycin, 12.5 μg/ml Tetracycline and 100 μg/ml Ampicillin in a 2L flask. Bacteria were grown at 220 rpm, 37 °C until OD_{600nm} reached 0.5 - 1.0. Cells were cooled to 20 °C and induced with 0.7mM final concentration IPTG (Sigma). Cells were grown at 20 °C overnight for protein expression. Cells were harvested by centrifugation at 7000 *x*g for 15 min.

10 g cells were resuspended in 60 ml chilled lysis buffer (150 mM NaCl, 20 mM Tris pH 7.4, 20 mM imidazole) and lysed by sonication. Lysate was centrifuged at 26,000 *x*g for 30 min. The supernatant was transferred into a fresh tube and heated at 85 °C for 10 min (until solution turned from clear to milky). Heated supernatant was centrifuged again at 26,000 *x*g for 30 min. The supernatant was incubated with 2 ml HisPur™ Ni-NTA beads (Thermo Scientific) with gentle rotation for 1.5 -2 hr. Unbound materials were removed by gravity flow using a bio-rad column. Beads were washed with 25 ml lysis buffer and 25 ml wash buffer (150 mM NaCl, 20 mM Tris pH 7.4, 40 mM imidazole). Protein was eluted with 7 ml elution buffer (150 mM NaCl, 20 mM Tris pH 7.4, 400 mM imidazole) for 15 min at room temperature. Eluate was collected and beads were incubated with another 7 ml elution buffer for 15 min. The second elution

typically gave another 30% of protein yield. Protein buffer was exchanged into 150 mM NaCl, 20 mM Tris pH 7.4 by dialysis. His-tag was removed by incubation with 3C protease (1:80 w/w ratio) at 4 °C overnight. Cleaved sample was incubated with 500 µl HisPur™ Ni-NTA beads to remove uncleaved proteins and His-tags. The unbound material (containing the cleaved protein) was collected and concentrated to 3 mg/ml. Concentrated sample was centrifuged at 15,000 *x*g for 10 min to remove large aggregates. Supernatant was injected onto Superdex 75 16/600 Hiload column (GE Healthcare) pre-equilibrated in 150 mM NaCl, 20 mM Tris pH 7.4. Fractions containing the target protein were collected and concentrated to 1.5 mg/ml. Purified protein was snap frozen and stored at -80 °C.

2.4 $GABA_BR$ solubilisation and purification

2.4.1 Detergent purification from small-scale adherent cells

All steps described below were performed at 4 °C or on ice unless otherwise stated. Buffers and equipment were all pre-chilled to 4 °C. Thawed cell pellets from 6-well plates were resuspended in 400 µl buffer (300 mM NaCl, 20 mM Tris at pH8.0, 1% mammalian protease inhibitors). 60 µl of 10% detergent (Anatrace) of choice supplemented with 20:1 molar ratio affymetrix cholesterol hemisuccinate (αCHS, Anatrace) was added to each 400 µl cell suspension. Cells were solubilised for 1.5 - 2 hr with gentle rotation. Insoluble materials were removed by centrifugation at 15000 rpm at 4 °C for 10 min. Supernatant containing the protein in detergent micelle were transferred to fresh tubes. 200 µl buffer and 30 µl purification beads of choice were added to the supernatant. Protein was allowed to bind to beads on a rotator for 1.5 - 2 hr. Sample was centrifuged at 600 rpm for 30 sec. Supernatant containing the unbound materials was removed. Beads were washed with 1 ml wash buffer (300 mM

NaCl, 20 mM Tris at pH 8.0, detergent at 2-3 x CMC%). Wash buffer was removed and wash step was repeated twice. 100 µl elution solution was added to the washed beads to elute the protein overnight with gentle rotation. 1D4-beads elution buffer was prepared by diluting 2mM 1D4 peptide stock (Genescript) with 3-5 times volume of wash buffer. Strepavidin beads elution buffer was prepared by diluting 5mM Biotin stock (Sigma) with 3 times wash buffer. For size-exclusion chromatography analysis, protein was centrifuged at 15,000 *x*g for 15 min to remove large aggregates before injection onto the column.

2.4.2 Detergent purification from suspension cells

All experimental procedures were performed at 4 °C or on ice unless otherwise stated. Buffers and equipment were all pre-chilled to 4 °C. Cell pellets from 200 ml suspension culture were resuspended in 6 ml buffer A (300mM NaCl, 20mM Tris at pH 8.0) supplemented with 1% mammalian protease inhibitors (Sigma). 900 µl 10% detergent of choice with 20:1 αCHS was added (approx. 1.5% final) to the suspension and rotated for 1.5-2 hr for solubilisation. Insoluble materials were removed by centrifugation at 9,000 *x*g for 10 min. Supernatant was transferred to fresh tube containing 6ml buffer A and 100 µl affinity-beads of choice. Protein was allowed to bind to beads for 1.5 - 2 hr with gentle stirring. Beads were collected by gravity using Bio-rad filter column. Beads were washed with 20 ml buffer A supplemented with 0.1% detergent, followed by another wash with 10 ml buffer A supplemented with 3x CMC detergent. Beads were carefully transferred into 1.5 ml eppendorf tubes and had excess liquid removed. Protein was eluted with 200 µl elution buffer (buffer A supplemented with elution molecules corresponding to choice of beads) for 4hr - overnight. Protein was concentrated using 100kD cut-off concentrator (Sartorius) at

1000 rpm, 1-2 min intervals. Concentrated protein was centrifuged at 15,000 *x*g for 10 min to remove large aggregates. Supernatant was loaded onto Superose 6 column (GE) pre-equilibrated in 150mM NaCl, 10mM Tris-HCl pH8.0, 0.01% DDTM 20:1 αCHS. Fractions were analysed by SDS-PAGE and western blotting to assess the purity of protein. Protein was concentrated again for EM sample preparation if needed.

2.4.3 Non-Detergent Membrane solubilisation

Experiments described in this section were performed on ice or at 4 °C. 300 mg cell pellet was resuspended in 800 µl hypotonic buffer (20 mM Tris pH 8, 2mM $MgCl_2$, 2mM $CaCl_2$, 1% MPI, 3 U/ml benzonase) and lysed for 40 min. Cells were homogenised using a glass homogeniser (Jencons), followed by sonication on an ice water bath. The cell lysate was centrifuged at 7,000 *x*g for 5 min and the supernatant was taken for further ultra-centrifugation at 130,000 *x*g for 1.5 hr. All supernatant was removed. The membrane pellet was resuspended in 800 µl solubilisation buffer (150mM NaCl, 10 mM Tris pH 8, 1mM EDTA) using a homogeniser (Jencons). Small amount of detergent (below CMC) were supplemented at this point to facilitate membrane solubilisation. 180 µg Saposin D or 100 mg SMA Lipolisq® (Sigma) was added to the resuspended membrane and left to incubate overnight with gentle rotation. Solubilised samples were incubated with 30 µl streptavidin beads for 2 hr. Unbound material was removed and beads were washed with 1.5 ml wash buffer (150mM NaCl, 10 mM Tris pH 8, 1mM EDTA). Saposin D-solubilised $GABA_BR$ particles were eluted with 60 µl elution buffer (1.25 mM biotin, 150mM NaCl, 10 mM Tris pH 8) for 4 hr. Eluted sample was loaded onto Superose 6 Increase column (GE Healthcare) equilibrated in 10 mM Tris-HCl, pH 8.05, 150 mM NaCl for analysis.

2.4.4 Size Exclusion Chromatography (SEC)

SEC experiments were performed using the Shimadzu high-performance liquid chromatography system. Analysis of sample in 10-40 microliter volume was achieved using the Shimadzu high-performance liquid chromatography system. By using an automated sample injection system, accurate and reproducible liquid chromatography analysis were performed (Figure 2.1). Data was interpreted as the overall shape and height of peak of the protein of interest, as FSEC were performed to compare the quantity and homogeneity of samples expressed or purified across different conditions. Chromatography results shown in the thesis were from single sample injection runs unless otherwise stated.

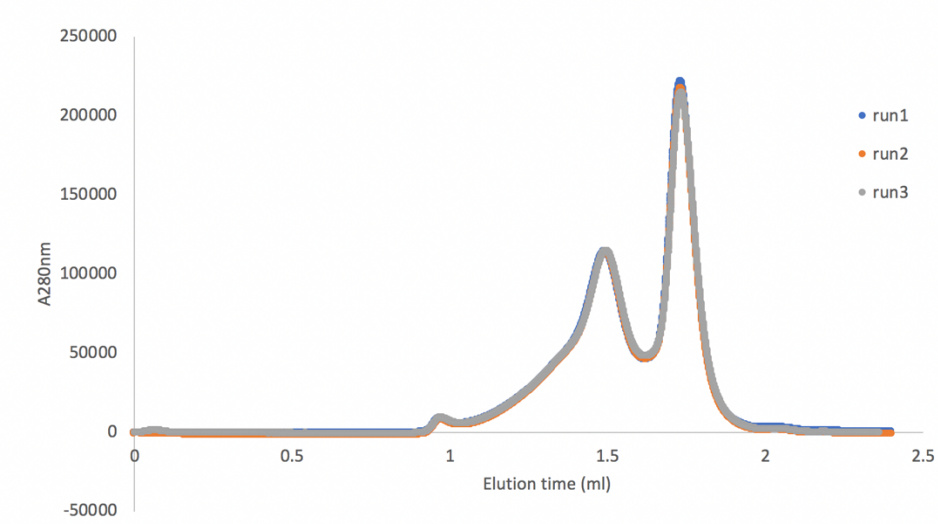


Figure 2.1 Reproducibility of Shimadzu analysis.

20 µl of the same sample was injected for each run. The output peaks were identical from three independent runs, confirming accurate sample injection and reproducibility of the data.

Typically, 20 µl sample was injected onto a Superose 6 Increase 3.2/300 column (GE healthcare) at a flow rate of 0.06 ml/min. The standard running buffer for detergent solubilised receptors is 150mM NaCl, 10mM Tris-HCl pH 8.0, 0.01% DDTM

20:1 αCHS, unless otherwise stated. Non-detergent solubilised receptors were run in 150mM NaCl, 10mM Tris-HCl pH 8.0 unless otherwise stated. Protein concentration was measured based on absorbance at 280 nm. For detection of fluorescently tagged proteins, the fluoremeter were set with an excitation wavelength at 488 nm and emission wavelength at 507 nm for the GFP-tag and an excitation wavelength at 515 nm and emission wavelength at 528 nm for the mVenus-tag.

2.4.5 Generation of 1D4-CNBr affinity beads

1 g of CnBr-activated sepharose (Sigma) was soaked with 50ml 1mM HCl for 20 min. All wash steps below were performed by gravity flow in biorad column. Beads were washed with 50 ml 1mM HCl, followed by another wash of 50 ml 500 mM NaCl, 100 mM $NaHCO_3$ pH 8.5. 14 mg antibody (UBC) was added to the beads and left to rotate 2hr-overnight at 4 °C. OD_{280nm} of the overnight solution was measured to ensure all antibody had coupled onto beads. Beads were washed with 50 ml 200 mM glycine pH 8.0 (or 500 mM NaCl, 100 mM $NaHCO_3$ pH 8.5). Beads were incubated with 50ml 200 mM glycine pH 8.0 (or 0.1 M Tris-HCl pH 8.0) at room temperature for 2 hr to block unbound activated groups on beads. Beads were washed with 50 ml 500 mM NaCl, 100 mM $NaHCO_3$ pH 8.5 (or 100 mM Tris-HCl pH 8.0, 500 mM NaCl). Followed by a wash of 50 ml 500mM NaCl, 100 mM NaOAc pH 4.0. Above two wash steps were repeated again. Beads were then washed with 50 ml 500 mM NaCl, 100 mM $NaHCO_3$ pH 8.5 (or 100 mM TrisHCl pH 8.0, 500 mM NaCl). Followed by a wash using 50 ml 100 mM Tris-HCl pH 8.0, 500mM NaCl. Beads were stored in 100 mM Tris-HCl pH8.0, 500mM NaCl at 4 °C.

2.5 Scintillation proximity assay

2mg Wheat germ agglutinin coated on yttrium silicate beads (PerkinElmer) were dissolved in 1 ml buffer (150 NaCl, 10 Tris-HCl pH 8.0, $2CaCl_2$, 0.05% DDTM 0.0025% αCHS). Protein sample was immobilized on beads by shaking for 1 hr at 4° C. For measuring nonspecific binding, another immobilization reaction was prepared in presence of 100 µM unlabeled baclofen. Protein coupled beads were then mixed with ^{3}H-baclofen and allowed to incubate for 2 hr before measurement. Emitted light at 420 nm wavelength was counted by a scintillation counter (PerkinElmer) (Figure 2.2). Data was fitted using non-linear regression in Graphpad Prism to calculate IC_{50}. In the case of a homologous competitive binding experiment, the radioactive and unlabeled ligands were assumed to have identical affinities so that K_d and K_i were the same. K_d was calculated using the formula $K_d=K_i=IC_{50}$ - [radioligand] [119].

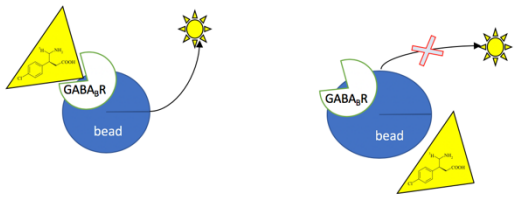


Figure 2.2 Schematic illustration of the scintillation proximity assay (SPA).

SPA measures radioligand binding to immobilised receptors. Tritium labelled particles with a short path length in aqueous solution will stimulate scintillation of the bead to which they are bound, but not adjacent beads. (left) Radioligand (yellow triangle, ^{3}H-baclofen) in close proximity to bead upon receptor binding to stimulating light emission. (right) unbound ligand is too far to activate bead and no light signal will be emitted.

2.6 $GABA_BR$ reconstitution

2.6.1 On-beads Reconstitution into amphipols/nanodisc

The procedures described below correspond to the amounts used for a 1.5 g cell pellet and were scaled linearly for large-scale experiments. All steps were performed at 4 °C or in cold room unless otherwise stated. Cells were resuspended in buffer and detergent solubilised as described in section 2.3.2. Solubilised receptors were bound to 200 µl beads of choice for 2 hr. While protein was coupling to beads, lipid mix was prepared. 5 µl of 20 mg/ml bovine brain extract (Sigma) dissolved in 1% DDTM was incubated with 25 µl of 10 mg/ml POPC (Avanti) in 3% DDM, 20 mM HEPES, 150 mM NaCl, 0.5 mM EDTA with rotation for at least 30 min at room temperature. At the end of incubation, mix was centrifuged at 15,000 *x*g for 5 min at room temperature to remove undissolved materials. Affinity beads were washed with wash buffer as described in section 2.3.2 and transferred into fresh eppendorf tubes together with 200 µl wash buffer. Supernatant from the lipid mix was added to the beads and incubated for 30 min. All supernatant was removed before the reconstitution reagent of choice was added. Typically, 1:8 w/w ratio for 0.6 mg amphipol A8-35 (Anatrace) and 0.3 mg MSP protein/saposin protein was used (expression and purification described in section 2.3.4). Sample was incubated for 1-3 hr for reconstitution. Detergent was then removed by bio-beads (Bio-rad) for 2-3 hr. Supernatant was removed and beads were washed multiple times with detergent free buffer to remove excess nanodiscs/amphipols. 2x bead volume of elution buffer was added for overnight elution. Sample was loaded onto Superose 6 Increase column (GE Healthcare) equilibrated in 10 mM Tris-HCl pH 8, 150 mM NaCl for analysis.

2.6.2 Reconstitution of purified proteins

Protein purified in detergent was concentrated to 6.5 mg/ml and supplemented with 3 mg/ml bovine brain extract (1% DDTM). Sample was incubated at room temperature for 30 min on a rotator at 20 rpm. Detergent was then removed by Bio-beads for > 30 min. Reconstituted sample was analysed by gel filtration on a Superose 6 Increase column (GE Healthcare) equilibrated in 10 mM Tris-HCl pH 8, 150 mM NaCl.

2.7 G protein complex formation

Protein was reconstituted into nanodisc as described in section 2.6.1 and gradually washed into 30mM NaCl, 20 mM Tris pH 8.0 buffer while still bound on beads. Protein was eluted overnight in 10 mM NaCl, 20 mM Tris pH 8.0 buffer supplemented with 0.5 mM biotin. Purified protein was concentrated to at least 3 mg/ml and incubated with 1.3 molar excess GFP-mini$G_{\alpha oR12}$ (kindly provided by Dr. Rony Nehmé, protein sequence in appendix B), with 10 mM $MgCl_2$, 1 µM 3-APPA (Sigma), 10 µM rac-BHFF (Sigma), 0.2 U/ml apyrase (NEB) and 0.1 mM GDP. Proteins were incubated for 8 > hr at 4 °C on a rotator. The incubation product was analysed by FSEC using a Superose 6 increase 3.2/300 column (GE Healthcare) equilibrated in 10 mM NaCl, 20 mM Tris pH 8.0, 1 µM 3-APPA and 10 µM rac-BHFF.

2.8 Chemical crosslinking

Protein complex used in cross-linking experiments were prepared following the same procedure as described in section 2.7, except the use of HEPES in buffer instead of Tris. Protein was diluted to 10 µM concentration and treated with each cross-linking reaction as specified below.

2.8.1 Chemical crosslinking using BS^3

12.5 mM BS^3 (ThermoScientific) stock solution was prepared immediately before crosslinking experiment. BS^3 solution at a range of concentrations (500 µM, 1 mM, 1.5 mM, 2 mM) was added to the protein. Sample was incubated in dark on ice for 30 min. Reaction was stopped by addition of 20 mM (final concentration) Tris buffer at pH 8.0 for 15 min at room temperature. Cross-linking reaction products were analysed by SDS-PAGE.

2.8.2 Chemical crosslinking using the CovalX kit

Crosslinking experiments using the CovalX stabilisation kit K100 and K200 (CovalX AG) were performed following manufacturer's protocol. Briefly, a 10 mg/ml stabiliser stock solution was prepared and was diluted to 2 mg/ml with stabiliser buffer just before cross-linking reaction. Protein was diluted to 10 µM concentration using the CovalX purification buffer. CovalX stabiliser in a concentration range (0.1 mg/ml, 0.2 mg/ml, 0.4 mg/ml, 0.6 mg/ml) were added the diluted protein and incubated on ice for 2 hr, or at room temperature for 60 - 180 min. Reaction was stopped by addition of 20 mM (final concentration) Tris buffer at pH 8.0. Cross-linking reaction products were analysed by SDS-PAGE.

2.9 EM sample preparation

2.9.1 Negative stain EM grid preparation

35mg Uranyl Formate powder (Polyscience) was dissolved in 5ml boiling MilliQ and vortexed for 2 min. 6.5 µl 5M NaOH was added to the side of tube without direct contact with stain solution and immediately vortexed for 2 min. The solution was filtered with 0.22 µm filter (Amicon). Prepared stain was covered in aluminium foil to

avoid exposure to light during storage. If stain was not used fresh (i.e. 20°C frozen aliquot), thawed stain was centrifuged for 30 sec before use.

Negative stain grids (Agar Scientific), with carbon side facing up, were glow-discharged for 30 sec at 30 mA using Sputter Coater S150B plasma cleaner. 4 µl of protein at 0.01-0.05 mg/ml was applied on the carbon side. Sample was left to adsorb onto the grid for 30 sec before sample liquid was removed by filter paper. The following steps were carried out with the carbon side touching each solution. The grid was rinsed twice by placing the grid onto 40 µl of sample buffer and subsequently blotted with filter paper. Then, the grid was transferred onto 20 µl of stain for a few sec and blotted. Finally, the grid was put in another 20 µl of stain for 30 sec and blotted. Grids were allowed to air-dry for 10 min before transferring into storage box to be stored at room temperature.

2.9.2 Cryo-EM grid preparation

A 3.5 µl aliquot of sample was applied on 1.2/1.3 grids (QUANTIFOIL) glow-discharged for 30 sec at 30 mA. Depending on the type of reconstitution system the protein was in, various concentration of protein sample were applied. For detergent-solubilised and saposin A salipro particles, a sample concentration of 1.5 mg/ml concentration was applied. 0.06 - 0.1 mg/ml protein concentration was used for samples reconstituted in amphipol, and 0.2 - 0.3 mg/ml for samples reconstituted in MSP nanodiscs.

The grids were vitrified by plunge freezing into liquid ethane using a Vitrobot plunger (FEI). Typical settings used for the $GABA_BR$ samples were with -15 force and 5-5.5 sec blotting time.

2.10 Cryo-EM Data processing

Cryo-EM data were processed using either the RELION software suite [120] or cryoSPARC v2 [121].

2.10.1 Single particle analysis using RELION

For RELION-based processing, movie frames were aligned by MotionCorr2 [122] to account for drift. Contrast transfer function (CTF) parameters were estimated using Gctf [123]. Movies showing minimal drift and astigmatism were used to pick particles by Gautomatch [124]. Particles were extracted by RELION and used for 2D reference free classification. 2D classes that resemble shape of a hypothetical receptor were selected and iterative rounds of 2D classification were run. For 3D classification, the initial 3D model was generated using the previously published ectodomain structure (PDB 4MQE) and the 7-helix transmembrane region of the metabotropic glutamate receptor (PDB 4OR2). Subsequent 3D classifications used a reconstruction model generated from a consensus 3D classification from experimental data as the input model. All models were filtered to at least 30 Å to reduce model bias.

Particles from 3D classification were saved to run 3D refinement with increasing local sampling rate. During auto-refinement, data set was split into two random halves to be reconstructed and refined separately. Fourier shell correlation was calculated across the two sets for resolution estimation. Masks for masked local classification were generated in UCSF Chimera [125] based on the refinement model. Masks were applied to subtract projections from areas outside the masks in all particle images. 3D classification were calculated for the subtracted data at higher T value (10-40). Multi-body refinement was run within RELION. The typical RELION-based

processing workflow is summarised in Figure 2.3. Crystal structure was flexible fitted into the ECD EM model using Phenix real space refinement [126].

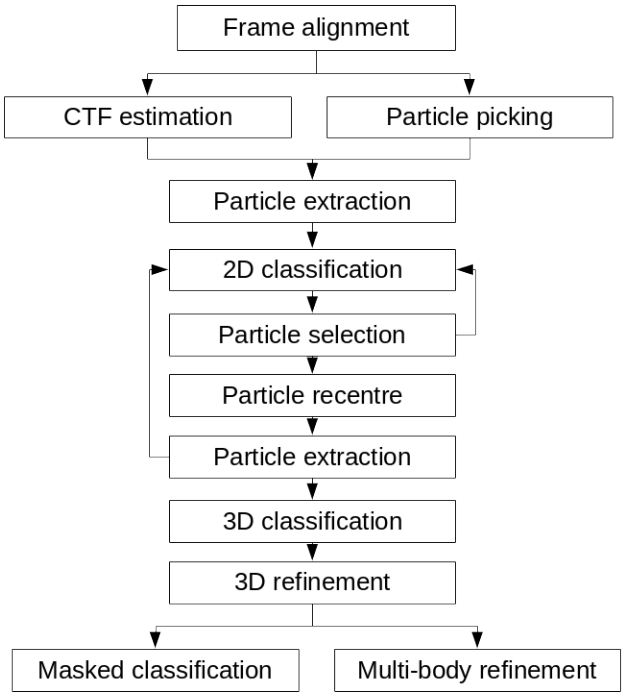


Figure 2.3 Schematic Representation of RELION-based data processing work flow

2.10.2 Single particle analysis using cryoSPARC

Processing workflow performed within the cryoSPARC System™ platform [121] is summarised in Figure 2.4. Movies were aligned using the CryoSPARC inbuilt alignment algorithm [121]. CTF estimation was performed using CTFFIND4 [127]. Initial sets of particles were manually picked for generation of templates for automated picking. Iterative rounds of 2D classifications were calculated with particle re-centering after each run. The 3D *ab initio* models were generated within the software as initial model for 3D refinements.

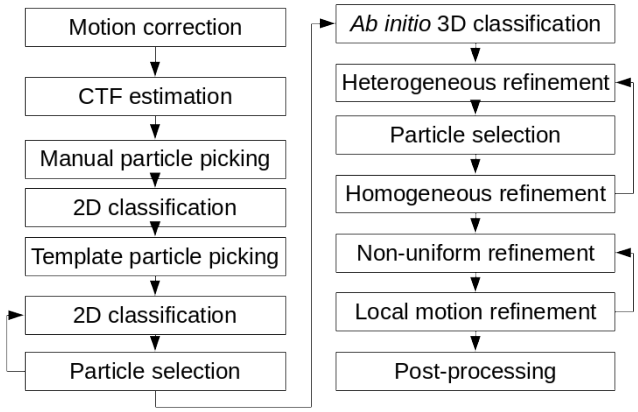


Figure 2.4 Schematic Representation of CryoSPARC data processing work flow.

Chapter 3. EXPRESSION OF HUMAN $GABA_BR$ IN MAMMALIAN CELLS

3.1 Introduction

The prerequisite to solving a protein structure is to obtain sufficient amount of protein of interest, usually in milligrams quantity. Obtaining such high quantities of protein is a particularly challenging for human membrane proteins. Their presence in the plasma membrane is normally low and expression levels are tightly regulated in the body. It had been reported that $GABA_BR$ undergoes a fast, constitutive internalisation process in the cell [128], limiting the number of fully functional receptors displayed on cell surface. Heterologous over-expression systems in bacterial or insect cells have been used for expressing individual soluble domains of the $GABA_BR$ [55, 78], but not for the transmembrane region nor the integral receptor. It is not surprising that these systems had not been able to produce the full-length human $GABA_BR$, given they lack the corresponding folding machineries and post-translational modifications found in mammalian cells. Mammalian expression systems are better alternatives as they provide the required synthesis environment for human proteins. However, expression in mammalian cells is normally at much lower level and requires more complicated tissue culture techniques. As a result, mammalian expression systems present the potential of yielding suitable protein samples but at much higher cost.

In this study, a transient transfection mammalian expression system was chosen to serve as a starting point to develop the over-expression protocol for the full-length human $GABA_BR$ [118]. This system utilises the Human Embryonic Kidney (HEK) 293 $GnTI^-$ cells, which produce proteins with short and homogenous N-linked

glycans [129]. By having a uniform glycosylation decoration of $Man_5GlcNAc_2$ [129], chemical and conformational heterogeneity is greatly reduced within a protein. For this reason, the HEK 293 GnTI$^-$ cells has become an extremely a popular cell line for structural biologists. A quick survey in the PDB database (at August 2018) revealed many membrane protein structures being solved using protein expressed in HEK 293 GnTI$^-$ cells, such as rhodopsin [130], $GABA_A$Rs [131, 132] and smoothened receptor [133]. The $GABA_BR$ contains five putative N-glycosylation sites in the $GABA_{B1}$ subunit and three in the $GABA_{B2}$ subunit. Inhibition of glycosylation has not been reported to harm $GABA_BR$ protein folding or subunit assembly [134]. Functionally, mutagenesis studies (of all five sites in the same construct) showed that N-glycosylation in the $GABA_{B1}$ subunit was not crucial for receptor pharmacology [135]. Hence HEK 293 GnTI$^-$ cells are suitable for $GABA_BR$ structural determination purposes.

Although HEK 293 GnTI$^-$ cells had been shown useful for producing high quality protein for structural studies, their expression level probably is the biggest limitation. Prior to this study, feasibility of large-scale expression of the integral human $GABA_BR$ was unclear. In this chapter, recombinant $GABA_BR$ expression levels were evaluated and optimised. The fluorescence-detection size exclusion chromatography (FSEC) was used as the primary technique to screen for key variables involved in expression levels [136]. Comparing to conventional methods like western blot, FSEC analysis provides both a quantitative measurement of the protein yield and a qualitative assessment of the protein homogeneity. FSEC screening experiments in this thesis were performed on the Shimadzu High-Performance Liquid Chromatography (HPLC) system, which allowed more rapid screening of samples at much smaller volume comparing to the conventional ÄKTA liquid chromatography instruments. The use of transient

transfection and FSEC provide a rapid and sensitive platform for evaluating protein expression levels under different conditions.

3.2 Results

3.2.1 Design of human $GABA_BR$ synthetic constructs

Synthetic constructs designed for recombinant expression resembled the $GABA_{B1a}$, $GABA_{B1b}$ and $GABA_{B2}$ subunits and covered all known key structural elements in the $GABA_BR$. The constructs protein sequences included all amino acids required for reconstitution of native-like receptor activity [58]. The sequences were attributed to three hypothetical structural domains of the receptor based on sequence alignment with existing structures, dividing them into the extracellular domain (ECD), the transmembrane hepta-helices (TMD) and the intracellular coiled-coils (CC).

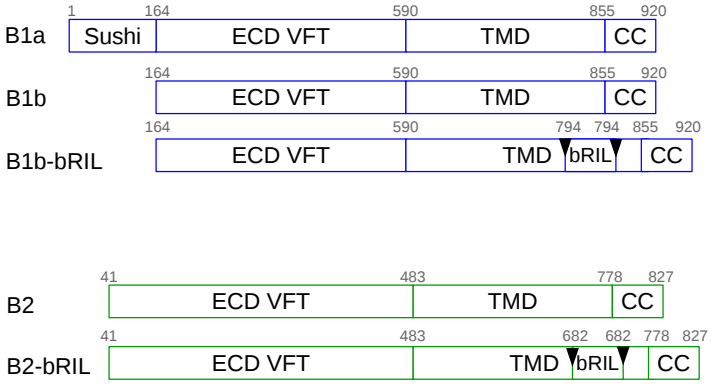


Figure 3.1 List of initial constructs generated.

$GABA_{B1}$ and $GABA_{B2}$ protein sequences were acquired using Uniprot entry Q9UBS5 and O75899 respectively. The synthetic constructs had been codon optimised for mammalian expression. Synthetic constructs DNA sequences were listed in appendix B. Domain boundaries (residue number labelled on top of each construct) were estimated by sequence alignment with exitsing structures. Locations of b_{562}RIL insertion were indicated by black triangles.

At the time when this study was proposed, both X-ray crystallography and Cryo-EM approaches were considered. For crystallisation studies, a set of constructs were designed as chimeras with the thermostabilised apocytochrome b_{562}RIL inserted in the 3rd intracellular region in both subunits (Fig 3.1). Fusion with a soluble protein such like the b_{562}RIL is a strategy commonly used to achieve successful crystallisation of GPCRs, yielding many high resolution structures such as A_{2a}AR [137], 5-HT_2B [138] and smoothened [133]. b_{562}RIL provides a well-ordered domain for crystal contacts on the intracellular side of the receptor for the highly flexible GPCR transmembrane region [139]. The sites of insertion into the $GABA_BR$ constructs for this study were determined by aligning $GABA_BR$ sequence with the solved structure of A_{2a}AR- b_{562}RIL fusion sequence [137]. A summary of the initial set of constructs generated is illustrated in Fig 3.1. These constructs were used to produce the $GABA_{B1a/B2}$ and $GABA_{B1b/B2}$ isoforms with the optional b_{562}RIL insertion.

3.2.2 $GABA_BR$ expression and surface trafficking in HEK293 $GnTI^-$ cells

Full-length receptor inserts were cloned into the pHL vector (vector map in appendix B) with a C-terminus mVenus fluorescent tag [140], followed by a 1D4 affinity tag. The fluorescent tag allowed detecting the receptor trafficking to the surface under wide-field fluorescent microscope and $GABA_BR$ specific detection in FSEC for evaluation of expression levels described in section 3.2.3 and 4.2.

Small-scale expressions were first performed to evaluate the possibility of expressing $GABA_BR$ in HEK293 $GnTI^-$ cells. The initial expression tests confirmed that cells were capable of expressing the full $GABA_BR$ subunits. Theoretical molecular

weight (MW) of the recombinant constructs were approximately 95 kDa and 94 kDa for $GABA_{B1}$ and $GABA_{B2}$ subunits respectively. Western blot analysis against 1D4-tag at receptor C-terminus detected a major band with apparent molecular weight (MW) of approximately 85 kDa (Figure 3.2) and a less defined band at 170 kDa. The two subunits were not completely separable on gel due to their similarity in size, causing their overlapped migrations. The slight discrepancy between the theoretical MW and the apparent MW observed is not uncommon for membrane proteins SDS-PAGE analysis [141]. MW of the major band was consistent with the integral recombinant $GABA_BR$ monomers and the higher band with the dimers. Strong presence of the monomer band could be due to subunits dissociation upon SDS treatment. However, it was not clear whether the receptor detected on the SDS-PAGE was homodimer or heterodimer.

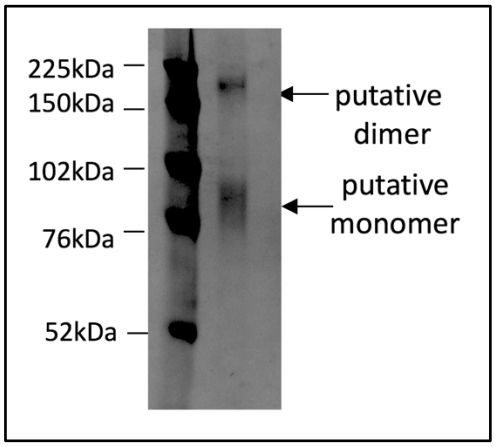


Figure 3.2 Western blot analysis of the recombinantly expressed protein from HEK293 $GnTI^-$ cells.

Receptor constructs included a mVenus fluorescent tag and a 1D4-affinity tag at the C-terminus. Bands were detected using anti-1D4 antibody against the 1D4-tag. Ladder proteins used were pre-stained and merged with western image for protein size estimation.

Formation of heterodimers was confirmed using FSEC, where signal from one subunit was detected using receptors purified using affinity tag on the other subunit. Cells in two groups were transfected with the same 1D4-tagged $GABA_{B1}$ construct but

different $GABA_{B2}$ subunit constructs. One group was expressing the mVenus-fused $GABA_{B2}$ subunit (B2_Ctmv), the other group was expressing the non-fluorescently tagged $GABA_{B2}$ (control). Proteins from both groups were purified using the 1D4 tag on B1 subunit and analysed by FSEC (Figure 3.3). Protein A_{280nm} profiles were similar in the two samples, suggesting receptors formed were of the same oligomeric state. No fluorescent signal was detected in the control group, as the receptor was not expressed with a mVenus tag. FSEC profile of the B2_Ctmv expressing cells yielded peak at same elution volume as in A_{280nm}, suggesting presence of $GABA_{B2}$ subunit in the sample. If the recombinantly expressed receptors were monomers, fluorescent signal would not be detected in either groups, since only $GABA_{B1}$ monomer could be purified. Hence the receptors expressed must be heterodimers or higher-order oligomers consisting of both subunits.

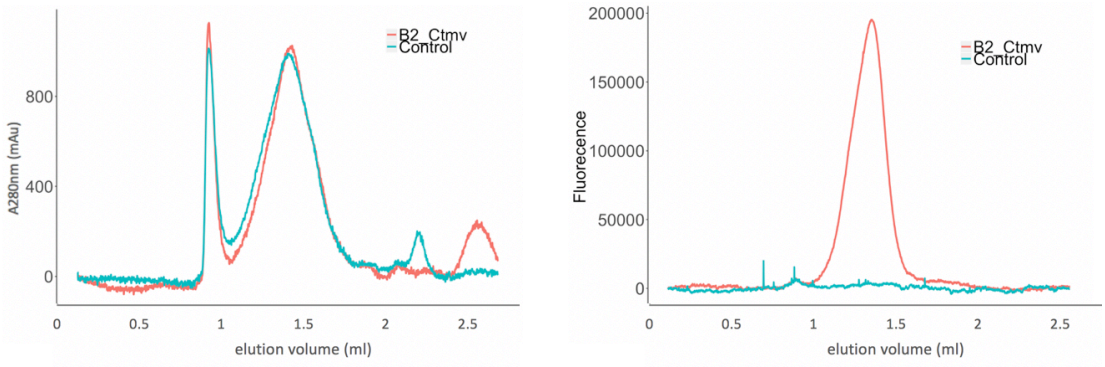


Figure 3.3 Confirmation of $GABA_{B1/B2}$ heterodimer formation.

Cells were transfected with the same $GABA_{B1}$ subunit attached to a C-terminal 1D4-tag in both groups. One group was transfected the $GABA_{B2}$ subunit linked to a C-terminal mVenus tag (B2_Ctmv) and not the other group (control). Constructs for the $GABA_{B2}$ subunits did not include a 1D4-tag. Cells from both groups were solubilised using 1.5 % DDTM detergent and purified by 1D4-affinity purification. 20 µl purified product from each was analysed by gel filtration on a Superose 6 column increase 3.2/300 column pre-equilibrated in 150 mM NaCl, 10 mM Tris pH 8.0, 0.01% DDTM.

The surface localisation of receptor constructs was examined using fluorescence microscopy 48 hr post transfection (Figure 3.4). Adherent cells transfected with fluorescently tagged $GABA_BRs$ maintained normal growth and morphology. Cells co-transfected with both subunits were transported to the plasma membrane, as indicated by the clear cell outline marked by the fluorescent tags. When expressed alone, the $GABA_{B1}$ subunit was retained in the ER while $GABA_{B2}$ subunit was able to traffic to the surface. These observations were consistent with previous studies that the $GABA_{B1}$ subunit requires association with $GABA_{B2}$ to mask its retention signal for surface transportation. $GABA_{B2}$ subunit, on the other hand, had been known to be trafficked to the cell surface independently [98]. In this study, the two subunits were always co-expressed in cells for formation of heterodimers.

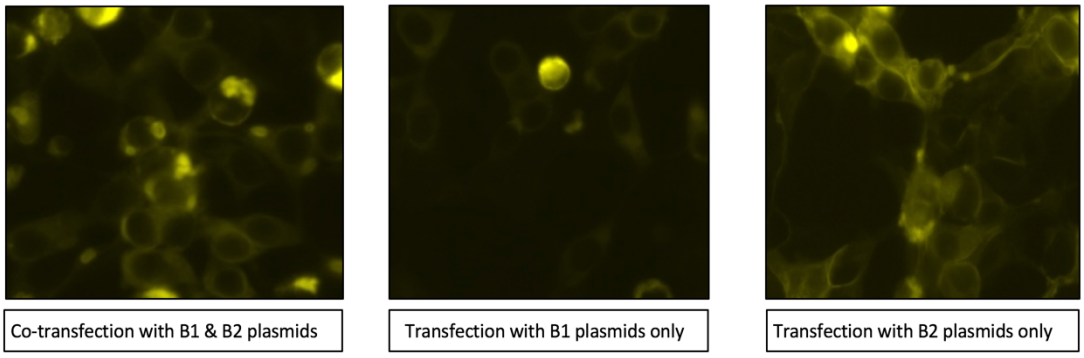


Figure 3.4 Localisation of recombinantly expressed $GABA_BRs$ in adherent cells.

Plasmids used for transfection had the C-terminal mVenus tag. Cells were transfected in 6 well plates by transient transfection. Confocal microscope images of transfected cells were recorded 48 hr post transfection.

3.2.3 Protein expression level optimisation by protein engineering

Protein engineering is routinely performed in structural biology to achieve higher yield and stability. Common engineering methods include site-directed point mutagenesis and truncations of flexible regions. This section aimed to identify

constructs giving high protein yield and monodispersity. In addition to the constructs described above, several other engineered constructs were created to investigate for their effects on protein expression levels. Table 3.1 summarised the list of constructs generated based on existing evidence in literature and their predicted impacts.

Table 3.1 Constructs designed for expression level tests.

Description	Rationale
Wild-type receptor	Native receptor
bRIL fusion at the 3rd intracellular loop in both subunits	Insertion had been shown to increase GPCR receptors stability and increase likelihood for forming crystal contacts [133, 137-139, 142]
Removal of sushi domain in $GABA_{B1}$ (i.e. generation of $GABA_{B1b}$ isoform)	$GABA_{B1a}$ and $GABA_{B1b}$ isoform are not functionally distinct [143], removal of sushi domains at N-terminal to minimise potential flexible regions
Complete removal of intracellular coiled-coils	Coiled-coils assembly are known as a rate limiting step in receptor synthesis, their removal may promote expression. Impacts on receptor functions are unknown
B2_K767R_K771R (Site-directed mutagenesis of $GABA_{B2}$ Lys^{767} and Lys^{771} to Arginine)	Lys^{767} and Lys^{771} reported to be the Lys-linked polyubiquitination sites in $GABA_{B2}$ for proteasomal degradation [144]. Removal of these sites to prevent possible protein cleavage
B1b_L889A_L890A (Site-directed mutagenesis of $GABA_{B1}$ Leu^{889} and Leu^{890} to Alanine)	$GABA_{B1}$ intracellular region is a known internalization di-leucine motif. Removal of these residues may enhance expression [81].

Cells were transiently transfected with the constructs listed in Table 3.1 and analysed by FSEC or A_{280nm} (Fig 3.5). The expression level of $GABA_{B1b/B2}$ short variant, without sushi domains, was better than the longer variant $GABA_{B1a/B2}$ based on protein peak intensity on at 1.4 ml (Figure 3.5a). (Receptors eluted at this volume has been confirmed to contain the dimeric receptor based on Cryo-EM reconstructions in Chapter 5.) A shoulder at earlier elution volume 1.2 ml was observed in FSEC profiles of both variants, suggesting that high-order oligomers might exist. The small peak at 0.9 ml was mostly likely aggregates that eluted in void volume according to manufacturer's protocol. The later peak that eluted at 1.9 ml on $GABA_{B1a}$ profile might represent the free mVenus tag that got cleaved off from the receptor and was purified

through C-terminus 1D4 affinity tag. Two constructs containing site-directed mutations were designed. “B2_K767R_K771R” construct was designed to reduce potential C-terminus cleavage [144] and “B1b_L889A_L890A” construct to promote $GABA_{B1}$ surface expression [81] (Table 3.1). The point mutated constructs did not show higher expression levels when comparing to the wildtype (Figure 3.5b). SEC profile of the coiled-coil truncation construct showed a slightly higher expression of a smaller product, which correspond to the removal of the intracellular coils (Figure 3.5c). A shoulder at 1.6 ml was observed on the truncation construct but not the full-length construct. The coiled-coil truncation constructs transfected cells had round-shaped appearance and slower growth rate, indicating that cell health was adversely affected and was causing their detaching from the flask. Cells transfected with all other constructs looked healthy. Therefore, the coiled-coil truncation construct was not proceeded further as it was not suitable for large-scale expression. The bRIL-fusion chimera construct yielded similar expression level when comparing to wildtype receptor (Figure 3.5d). The elution peak was slightly earlier for the bRIL-fusion receptors due to the increase size in fusion protein size. Wildtype receptors peak had the shoulder observed at a larger molecular weight, where bRIL-fusion receptors had a more symmetrical peak, indicating that fusion receptor may display a more uniform shape. Taken together, engineered constructs except the bRIL chimera did not show more favorable expression than the wildtype receptors. It appeared that $GABA_{B1b/B2}$ isoform with bRIL were good candidates to be proceed further on the basis of their expression level and sample homogeneity.

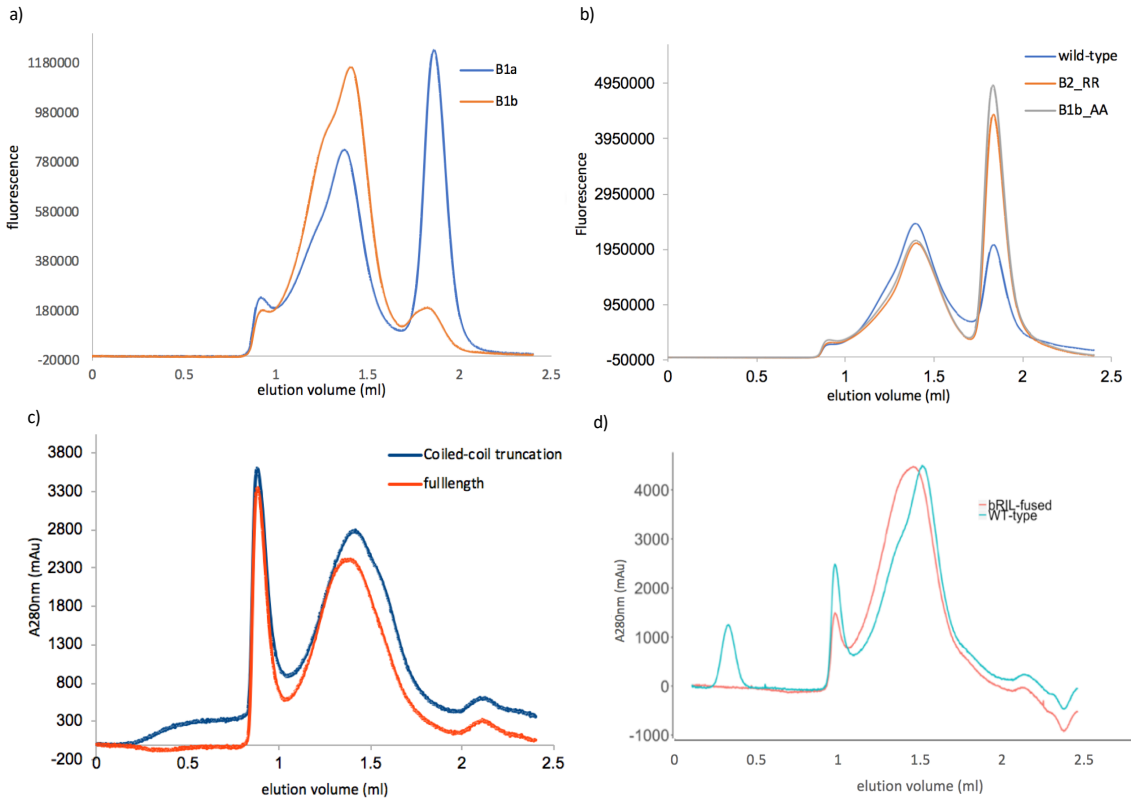


Figure 3.5 Effects of protein engineering on expression levels.

Cells were transient transfected with the following constructs and allowed to express for 48 hr before analysed by FSEC or A_{280nm} on Superose 6 Increase 3.2/300 column. Dimeric protein peak was between 1.4 - 1.5 ml elution volume depending on the construct used. Column void volume was at approximately 0.9 ml according to manufacturer protocol. Elution profiles of different constructs: (a) $GABA_{B2}$ co-expression with $GABA_{B1a}$ or $GABA_{B1b}$ isoform. (b) Site-directed mutants of residues involved in protein maturation pathways "B2_K767R_K771R" and "B1b_L889A_L890A". (c) Intracellular coiled-coils truncation. (d) bRIL-fusion chimera.

3.2.4 Improving expression in transient transfected suspension cells

In this section, the exact expression parameters were fine-tuned to best suit the recombinant $GABA_BR$. It is a critical and necessary step to identify the conditions required for each individual protein in order to achieve optimal yield. The large-scale transient transfection procedure developed in the Aricescu lab provided a general guideline for over-expression in suspension cells [118]. Trials to alter expression conditions were performed using HEK293 $GnTI^-$ cells suspension culture, so that the findings could be directly applied to large scale expressions. Protein construct used for

screening described from here on and in Chapter 4 was the $GABA_{B1b\text{-}bRIL/B2}$ heterodimer unless otherwise stated. Various aspects which might affect expression level were considered included transfection cocktail composition, tissue culture conditions, additives and modulators.

Transient transfection rate was estimated to be 35% ~ 45% by counting number of green fluorescent cells transfected with mVenus fluorescent-tagged constructs, over total cells under fluorescent light microscope. This transfection rate is typical for transient transfection of cells in suspension (personal communication with Dr. Paul Miller - STRUBI). Dilution of the transfection medium with $^1/_3$ volume of sterile MiliQ water during transfection cell resuspension was found to improve transfection efficiency by 20%. Dilution of medium might have caused cells to swell due to osmotic pressure, making them more permeable to the transfection reagents.

Effects of transient transfection using different plasmids ratios were examined (Fig 3.6). $GABA_{B1}$ and $GABA_{B2}$ plasmids were mixed at different ratios while keeping the total amount of DNA unchanged. Yield of the heterodimeric receptor was almost doubled using 2 to 1 molar ratio of $GABA_{B1}$ to $GABA_{B2}$ plasmids (Figure 3.6a). Furthermore, receptor DNA plasmids were diluted with empty pHLsec vector while keeping the same amount of total DNA in the transfection cocktail. Having twice the molar amount of empty vector to the receptor DNA was found to give strongest expression (Figure 3.6b). Further increase in the amount of empty vector started to show decreased expression level (Figure 3.6b).

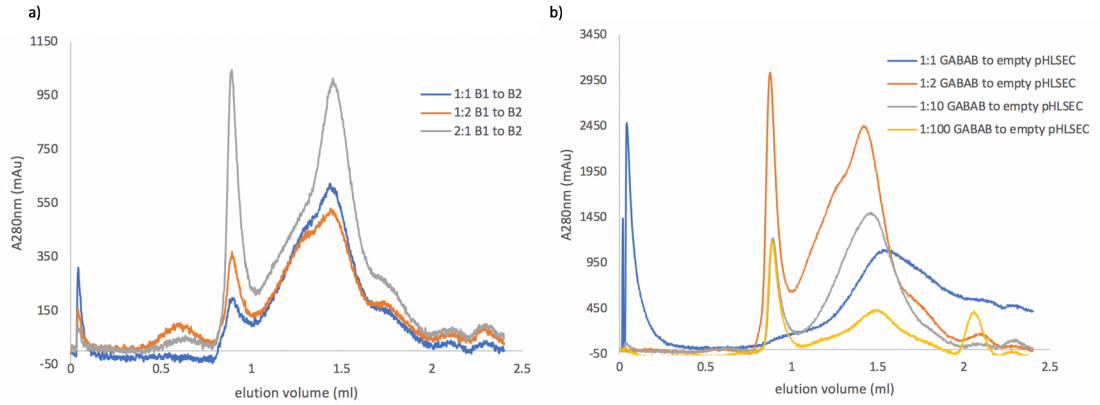


Figure 3.6 Effects on variations of DNA ratios used for transfection while keeping the same amount of total DNA for each transfection.

(a) Effects on changing DNA ratios of $GABA_{B1}$ and $GABA_{B2}$. (b) Effects on diluting receptor DNA with empty pHLsec vector during transient transfection. Ratio here was calculated by sum of $GABA_{B1}$ and $GABA_{B2}$ plasmids against empty pHLsec vector.

Some protein sample had been found to express better at lower temperature [145], therefore expression levels at 37 °C and 30 °C was investigated. In the case of $GABA_BR$, 37 °C incubation gave slight better expression level (Figure 3.7a). Other factors including expression time and different expression media were tested. It appeared that protein expression had plateaued 48 hr post-transfection and leaving them for another 24 hr (i.e. 72 hr in total) did not show significant further yield gain (Figure. 3.7b). Expression level of cells grown in different medium was compared. Freestyle medium yielded much higher expression level compared to PEM medium (Figure 3.7c). Effects of adding $GABA_BR$ endogenous ligand GABA and two histone deacetylase inhibitors [146], namely valproic acid (VPA) and sodium butyrate, were examined. Addition of 1mM VPA increased expression by 1.5-fold (Figure 3.7d). Having higher amount VPA (e.g. 4mM) did not further promote expression. Sodium butyrate was added on top of 1mM VPA to test whether expression level can be further increased. Sodium butyrate were tested in concentration ranging from 0.5 mM

to 6 mM and none showed favorable increase in expression level (Figure 3.7e). GABA supplemented at physiological extracellular concentration of 100 nM did not facilitate expression. Protein expression only appeared to be improved by approximately 10% if supplemented by up to 1mM GABA in the medium (Figure 3.7f).

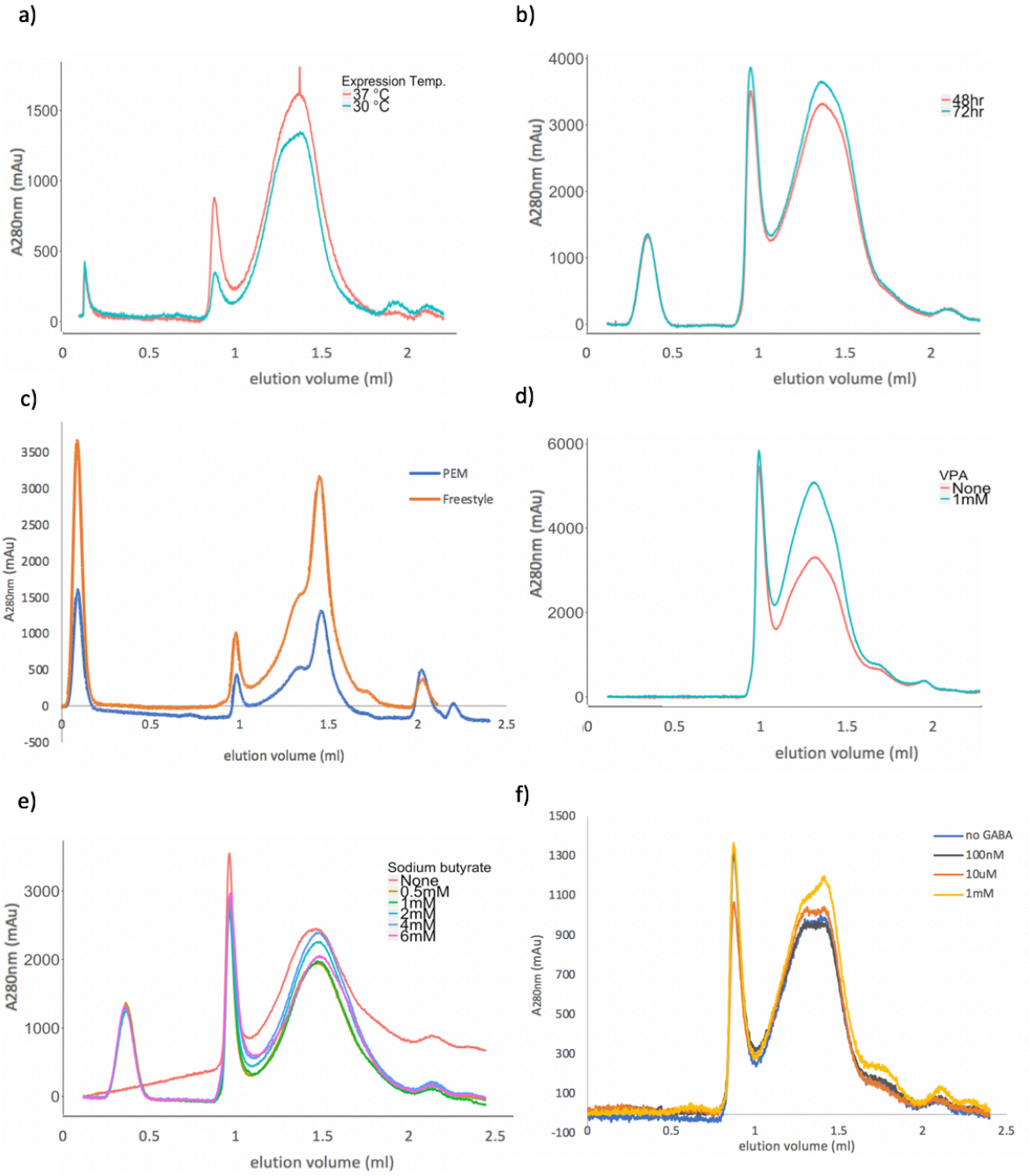


Figure 3.7 Comparison of protein expression levels in suspension culture under different conditions.

(a) Protein expression at 37 °C and 30 °C post transfection. (b) 48 hr and 72hr post-transfection at 37 °C. (c) PEM and Freestyle medium at 37 °C. (d) Effect of 1mM VPA addition during transfection at 37 °C. (e) Effects of sodium butyrate supplementations on top of 1mM VPA. (f) Effects of GABA supplementations.

3.2.5 Generation of polyclonal stable cell lines

Stable cell lines were created using lentiviral vectors in order to overcome the limited transfection efficiency associated with the transient transfection approach. Constitutively active cell line was generated in HEK 293 GnTI^{-} cells using pHRSIN-CSGW_vc (appendix B) as the transfer vector and adapted to grow in optimal suspension condition described above. Comparing to transiently transfected cells, protein yield was 3-fold higher from the same biomass of the constitutively active lentiviral cell-line (Figure 3.8a). The constitutive active cells maintained similar expression level up to 30 days of growth in suspension, after which expression started to decline (Figure 3.7b). It is a common phenomenon that expression level declines after cells were kept in culture for an extended period of time. The reason could be decline in cell health after multiple passages in suspension. Another possible reason could be the adverse effects of constitutive expression on cells health. It could be a combination of the two that caused the apparent reduce in expression level.

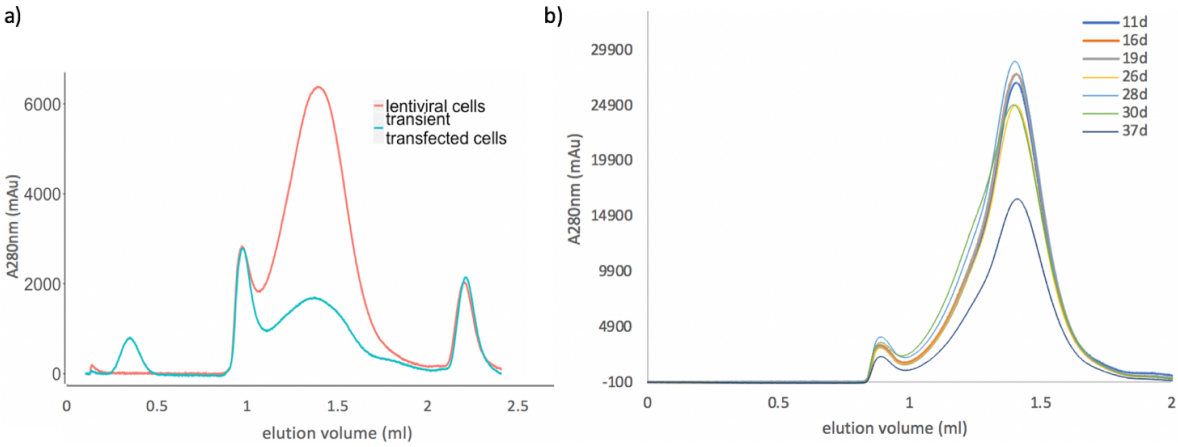


Figure 3.8 $GABA_BR$ expression levels in constitutively active cell line.

Same biomass of cells expressing the identical construct were processed in these experiments. (a) $GABA_B$ expression levels in transient transfected cells (cyan) and stable lentiviral cell line (red). (b) The constitutive active cells expression levels after 11-37 days of adaption into suspension. Expression maintained at similar levels up to 30 days of growth in suspension, after which expression started to reduce.

Next, expression level in inducible lentiviral cell line was investigated to allow cell growth and expansion in an uninduced state without the stress for overexpressing heterogenous proteins. Protein expression was only induced until a desired biomass was reached. Here, a tetracycline-inducible expression system was generated using the pHR-CMV-TetO_2 (Appendix B) plasmid suite in HEK293 GnTI$^-$ TetR cells. The HEK293 GnTI$^-$ TetR cells contain the TetR protein, which binds above the introduced gene of interest and repress its transcription [147]. TetR protein is released by induction with tetracycline to allow transgene expression [148]. HEK293 GnTI$^-$ cells and HEK293 GnTI$^-$ TetR cells were infected with receptor constructs containing N-terminal mEGFP attached on the $GABA_{B2}$ subunit to allow for yield comparison by FSEC. After cell line generation, cells were adapted to grow in suspension until the desired volume was reached. Stable cell line was grown in parallel to keep the same passage numbers and harvested at the same time as the inducible cells.

Surprisingly, expression level was found to be much lower in inducible cells compared to the same cell mass of constitutive expressing cells based on intensity of protein peak at 1.4 ml (Figure 3.9). Inducible cells were found to yield twice the amount of protein if expressed at 30 °C. Additional peaks were observed at 1.6 ml and 1.9 ml which might be the incomplete protein products. This could be an indicative for too much synthesis happening in the cells and using all available metabolic materials, as a result most protein synthesis got stuck at beginning stage. It is unknown whether the different expression behaviours observed in inducible cell line is due to a different in expression rate in different TetR cell types. Another possible explanation for the lower expression level is ineffective induction of cells. Or the inducible line generated

had lower number of plasmids incorporated in the cells comparing to in constitutive expressing cells, as infection efficiency for the inducible cell line was difficult to measure while uninduced.

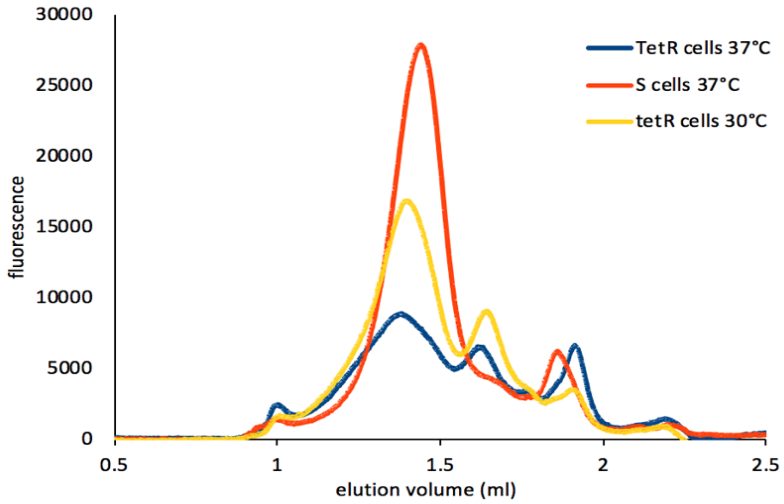


Figure 3.9 Comparison of $GABA_BR$ expression levels in constitutive active cell lines and inducible cell lines.

Same cell mass from each sample were lysed using 1.5% DDTM 20:1 αCHS for 1.5 hr, after which the unsolubilised materials were removed by ultra-centrifugation. Lysate were analysed by FSEC on a superose 6 3.2/300 increased column pre-equilibrated in 150 mM NaCl, 10 mM Tris pH 8.0, 0.01% DDTM 20:1 αCHS. Expression levels from constitutively active cell line at 37 °C was plotted in red, and inducible cell line at 37 °C in blue and at 30 °C in yellow.

Lastly, expression level of the stable cell line in the WAVEbag Bioreactor system was tested. The WAVEbag bioreactor system allows growth of large volume of cells (up to 10 litre) in a sealed bioreactor. The yield of cells in WAVEbag was ~80 % of cells grown in sterile suspension bottles (Figure 3.10). The protein from both setups displayed similar elution profiles in size exclusion chromatography (SEC). Although yield per cell mass was slightly lower in WAVEbag, the cells were able to stay at higher density without forming clumps or significant cell death in the WAVEbag, thereby keeping more cells for the same volume of medium. Thus WAVEbag system provided a cheaper alternative to suspension bottles for large scale protein production.

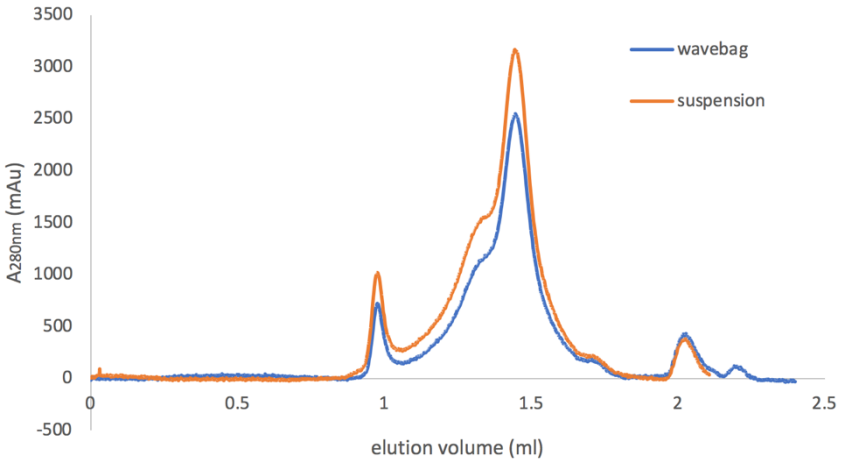


Figure 3.10 Comparison of expression levels of cells grown in suspension bottles and WAVEbag.

An aliquot of cells was used to expand both cultures and kept with the same passage numbers. Same cell mass from each was processed and analysed.

3.3 *Discussion*

One of the major bottlenecks in solving the structure of a membrane protein is the production of sufficient amount of sample. The first task of this project was to investigate the possibilities of overexpressing recombinant $GABA_BR$ using mammalian expression system. Host cells such as the HEK293 cells and CHO cells had been proven useful for production of membrane protein samples for structural studies [149, 150]. HEK293 $GnTI^-$ cells were chosen as the host cells in this study, given its ability to produce proteins with uniformed sugars due to the lack of the N-acetylglucosaminyl-transferase I in the cells. Proper expression and localisation of recombinant heterodimeric $GABA_BR$ in HEK293 $GnTI^-$ cells were confirmed at the beginning of this study. The project then proceeded to optimising over-expression of the receptor for large-scale production.

The combination of suspension cell transient transfection method and SEC using Shimadzu provided a rapid and reliable system for guiding expression optimisations. SEC profiles provided a more comprehensive assessment on protein expression comparing to conventional methods such as SDS-PAGE and western blot. In addition to the quantity of protein expressed, SEC data more accurately reflected on the quality of protein expressed in terms of sample homogeneity, degradation and aggregation. The total turn-around time for testing most parameters described in section 3.2 could be completed in less than two weeks using transient transfection. Findings from the small-scale trials could be directly applied to large-scale protein production without further modifications.

The best isoform identified so far in terms of yield and mono-dispersity was the $GABA_{B1b/B2}$. The structural elements included in both subunits were the venus-flytrap extracellular modules, the transmembrane domains and the intracellular coiled-coil helices. bRIL-fusion constructs were found to give more symmetrical SEC profile. The apparent increase in conformational homogeneity could be from stability of bRIL itself or having the insertion reduced movement in the $GABA_BR$ transmembrane region. Fusion with bRIL had been reported to increase stability of many GPCRs and facilitated their structural determination [139].

Other mutants and deletions tested so far did not appear to influence protein yields significantly comparing to the wild-type receptors. Although, the mutated constructs tested were reported to give higher surface expression, SEC profiles in this study did not show significant difference since data here reflected the quantity of all tagged proteins harvested from the cell regardless of its location. Systematic alanine

screening have been used to identify point-mutations which gave higher GPCR stability, for example in $A_{2A}R$ [151]. Such extensive screening was not performed in this study for two reasons: firstly, the yields of current constructs were amenable for structural studies; secondly, to minimise protein engineering which may affect the physiological functions of the receptors. $GABA_BR$ homologs from other species could be tested in the future. Homologs from other species might have higher intrinsic thermostability. The wildtype $GABA_{B1b/B2}$ receptor and its bRIL-fusion version were taken to the next step for large scale expression.

The optimal conditions for $GABA_BR$ expression in transient transfection were identified as the following: transfect with a DNA cocktail mix consisted of 2:1:6 ratio B1:B2:empty vector in presence of 1mM VPA supplement at 37 °C for 48-72 hr in Freestyle medium. As described in Chapter 1, $GABA_{B1}$ subunit was the rate limiting subunit in $GABA_BR$ synthesis, therefore having more $GABA_{B1}$ DNA during transient transfection allowed more heterodimer formation. Inclusion of empty vectors during transient transfection served the purpose to allow proper folding of receptors at an optimal speed in the cells. VPA functions as a histone deacetylase inhibitor in the cells, therefore addition of VPA helped to loosen DNA packing to allow better access of transcription machinery for expression [146].

Transient transfection provided a relatively quick method to identify suitable constructs and conditions for overexpression. However, this method also came with some disadvantages: the requirement of large amount of DNA, low and variable transfection efficiency, time consuming transfection procedure and etc. These drawbacks make transient transfection not ideal for the production of large quantity

the $GABA_BR$. Lentiviral system was next employed for large-scale protein expression once a suitable construct was identified. Lentiviral cell lines provided a higher yield and more reliable protein source. The use of a stable expressing cell line increased protein yield by 3-fold comparing to transient transfection. The total yield increase was more likely due to increased numbers of actively expressing cells rather than increase in protein amount produced per cell. Although creation of a cell line required 2-3 months, a cell line is more efficient on the long term once a desired construct was identified. The decline in expression level after extended period of cells growth in culture had been a commonly observed problem for constitutively active cell lines [152]. The reasons might be that over-expression of the target protein exhausted host cells metabolic resources or it adversely affects cellular function. Inducible cell lines had been developed to circumvent this problem, for example for the serotonin $5\text{-}HT_3$ [153] and b2 adrenergic receptors [154]. With cells growing in an un-induced state, only expressed for a limited time to retain host cell health and maximal protein synthesis.

The expression level of $GABA_BR$ in the inducible cells were ¼ of the constitutively active cells. Reason for the lower expression level in HEK293 $GnTI^-$ tetR inducible cells was unclear. Possible explanations included ineffective induction of the cells and lower number of genes incorporated in the inducible cell genome. Cell lines prepared in this study (stable and inducible) generated were all polyclonal cell lines, meaning that the cells are at different expression level. If looking to further improve expression, Fluorescence Activated Cell Sorting (FACS) could be used to select only the high-level expression cells [155]. FACS sort cells based on brightness of cells, signal of

which comes from a fluorescent maker that directly reflect on protein expression level. The GFP could be independently translated on the vector under uninduced state for the sorting. The selection procedure will require an estimated 3 months period and was not pursued in the study, as the protein amount obtained from current polyclonal cell-line was sufficient for Cryo-EM studies.

Stable cells could be grown either as suspension culture in bottles or in WAVEbags. The WAVEbag provided a cheaper and less labour-intensive alternative for growing large volume of culture. Growing cells adherently was not practical for the number of cells required hence not tested.

3.4 Conclusion

Having a suitable expression system and high protein yield are crucial to any structural studies. In this chapter, amenability of over-expressing of the integral human $GABA_BR$ in HEK293 $GnTI^-$ cells was investigated. $GABA_BR$ constructs starting from the ECD VFT to the intracellular coiled-coil, unmodified sequence or with bRIL-insertion, were found to be suitable for over-expression. An optimal expression condition was established for large scale protein expression in suspension culture. Constitutive active stable cell lines had been created for the suitable construct to be scaled up for structural determination.

Chapter 4. SAMPLE PREPARATION FOR CRYO-EM STUDIES

4.1 Introduction

Once sufficient amount of receptors can be expressed, the next task is to efficiently purify them from the cells. Similar to optimisation of the expression condition, every step during the purification process also has to be carefully optimised for $GABA_BR$. GPCRs are known for their inherent high degree of conformational flexibility and are normally found to be very unstable when isolated *in vitro*. Successful reconstitution of membrane proteins in solution involves two aspects: first is efficient solubilisation of receptor from cells; second is to maintain them in a functional state. Detergent is the most commonly used reagent for cell membranes solubilisation, such as the use of n-dodecyl-β-D-maltoside for structural studies of B_2AR [156] and κ-opioid receptors [157]. Choosing an appropriate detergent that meets the above criterion is crucial for high-resolution structural determination. Importantly, detergent extraction efficiency does not always correlate with its ability to maintain stability and preserve function of the solubilised receptor. Therefore, a functional assay is required to assess the stability and function of the solubilised proteins to guide in choosing the appropriate detergent(s). In this study, the scintillation proximity assay (SPA) was used to examine the ligand binding activity of detergent purified $GABA_BR$. Detergent-free membrane solubilisation was also explored in this study, as receptors isolated together with its native environment is believed to better retain protein activity and stability [158].

Once cell membranes had been solubilised, the next challenge was to efficiently purify the heterodimeric $GABA_BR$. As cells were transfected to co-express

both subunits, homodimer formation was possible and a purification strategy was necessary to remove the unwanted homomers. Constructs were designed such that the each subunits had a different affinity tag for a two-step affinity purification strategy to collect only the heterodimers. Major factors considered for choosing the suitable tags were their binding efficiency and recovery rate. In addition, delicate nature of membrane proteins imposes limitation on the choice of tags that can be used for protein purification. Common tags like His-tag were not suitable as the binding was not highly efficient and elution condition was not likely to be friendly to membrane proteins. The nine amino acid peptide Rho 1D4-tag [159] stood out in the search for appropriate tags for $GABA_BR$. The 1D4-tag had been used for purifying many membrane proteins, including various GPCRs [133, 154, 160] and ion channels [131, 161]. 1D4-tag binds to the rho-1D4 monoclonal antibody with high specificity and its binding efficiency had been shown to be unaffected by mild detergents [162], making it the ideal epitope to be put at the C-terminus of $GABA_BR$ where it is close to the membrane. Elution condition for 1D4-tag is very mild, consisting of low concentration detergent buffer and the 1D4-peptide. The other tag chosen was the Streptavidin-Binding Peptide (SBP)-tag, which is a 38 amino acids long peptide that bind to streptavidin beads. SBP-tag binds tightly streptavidin-beads with a K_d of 2.5 nM [163] and required also very mild elution condition [163]. $GABA_BR$ expressed for large-scale protein production had a N-terminal SBP-tag on the $GABA_{B1b}$ subunit and a C-terminal 1D4-tag on the $GABA_{B2}$ subunit. The purification buffer was also optimised, taking into consideration key factors such like pH and additive supplements. Again, the Shimadzu HPLC instrument was employed to evaluate the conditions tested based on intensity and shape of sample SEC profiles.

As the project developed, the single particle analysis Cryo-EM approach was prioritised over X-ray crystallography. Current popular solubilisation systems for Cryo-EM studies include detergent-micelles, amphipols and lipid nanodiscs. The ways these solubilisation reagents associate with a membrane protein are not the same, therefore protein solubilised in them were likely to show different characters. $GABA_BR$ solubilised in each of the above reagents were assessed for their suitability for cryo-EM sample preparation.

4.2 Results

4.2.1 Detergent screening

A set of detergents with different properties was evaluated in their efficiency to solubilise the $GABA_BR$. $GABA_BR$ is known to be associated with the membrane lipid raft which is enriched in lipids and cholesterol [164]. Analysis of previously solved GPCRs structures revealed that they were often purified in presence of cholesterol [133, 137, 165]. Therefore, detergent screening were performed using detergent supplemented with 20:1 molar ratio of cholesteryl hemisuccinate (αCHS), a cholesterol derivative known to stabilise GPCRs [166]. Solubilisation efficiencies of different detergents were tested by resuspending frozen cell pellets in purification buffer, dividing into aliquots of identical volume and solubilising with 1.5% final concentration of detergents (supplemented with 20:1 αCHS) at 4°C for 2 hr. Affinity purification was performed and identical volume of elutant from each condition was analysed by SEC (Figure 4.1a). n-Decyl-β-D-Maltoside (DM) and Lauryl Maltose Neopentyl Glycol (LMNG) appeared to be the most efficient in membrane solubilisation, as indicated by having the highest intensity in protein peak at 1.4 ml.

Receptors solubilised in these two detergents also had highest peak in void volume at 0.9 ml, suggesting they extracted large amount of protein but did not stabilise it. Yield of Dodecyl Maltoside (DDM), Decyl Maltose Neopentyl Glycol (DMNG) and 2,2-bis(3'-cyclohexylbutyl) propane-1,3-bis-β-D-maltopyranoside (Cymal 6 NG) were approximately 80% of DM. Receptors purified in different detergents gave different elution volumes. This variability in apparent sizes could be due to different sizes of micelle formed or different oligomeric states of the purified receptor in different detergents. In addition, most detergents had an asymmetrical protein peak, suggesting the presence of different oligomeric states in the sample. n-Dodecyl-β-D-thiomaltopyranoside (DDTM) and DM had the most symmetric peak, though, quite broad. The apparent peak symmetry could be from a homogenous sample, or due to poor separation of mixed population of oligomers. SDS-PAGE (Figure 4.1b) and Western blot (Figure 4.1c) analysis of the SEC fractions from DDTM agreed with the latter. Bands corresponding to MW of higher-order oligomers were observed at earlier eluted fractions and had overlapping elution fractions with dimers.

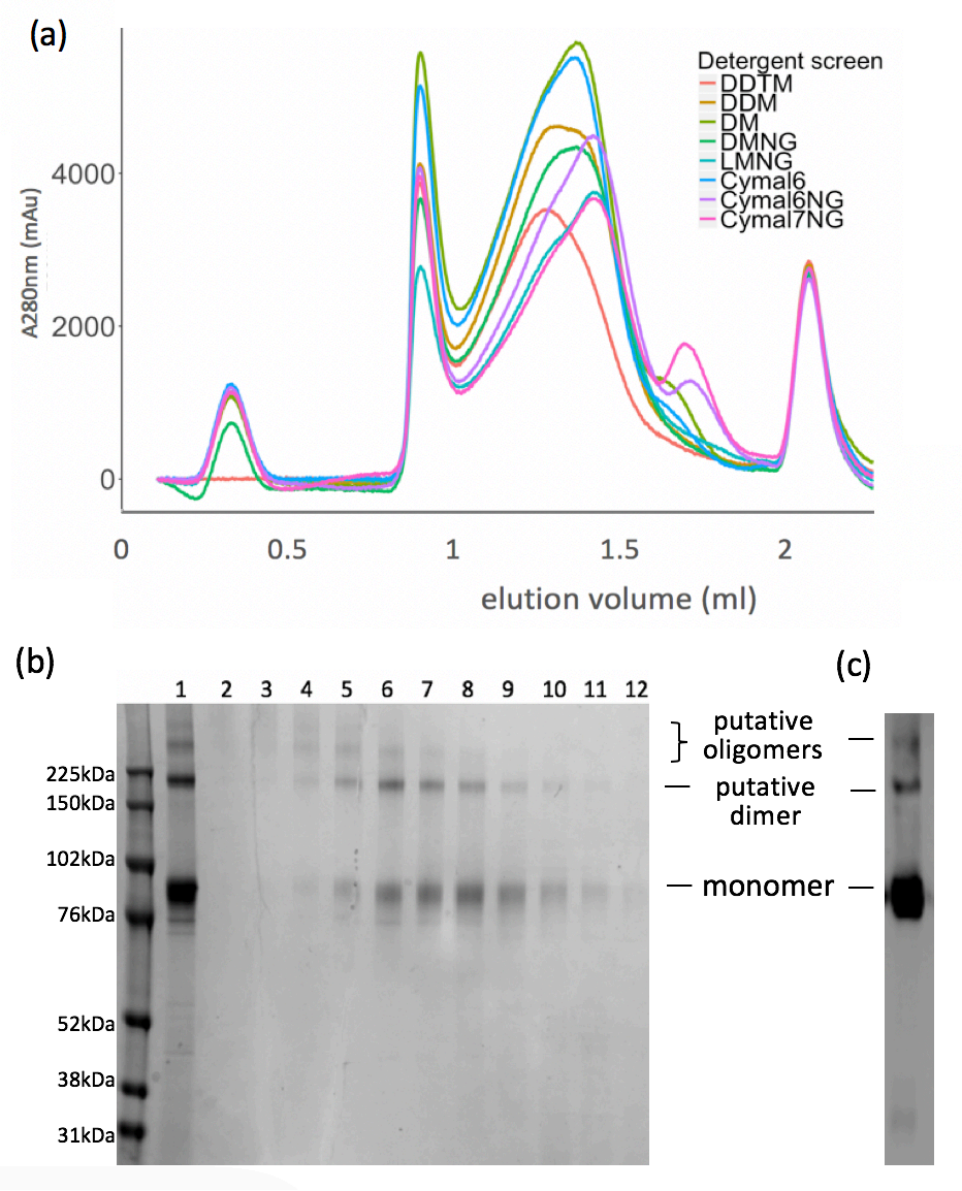


Figure 4.1 Analysis of solubilisation efficiency of various detergents.

Frozen cell pellets were resuspended in purification buffer and divided into aliquots of identical volume. Cells were solubilised with 1.5% final concentration of each detergent (supplemented with 20:1 αCHS) at 4°C for 2 hr. Cell lysate was purified and washed using buffer supplemented with the respective detergent supplemented with 20:1 αCHS. (a) SEC curves of samples in different detergents run on Superose 6 Increase 3.2/300 column equilibrated in 150 mM NaCl, 10 mM Tris-HCl pH 8.0 and 0.01% DDTM 20:1 αCHS. (b) SDS-PAGE analysis of SEC fractions from sample purified in DDTM. Lane 1 contained the sample before injection. Lane 2-12 contained fractions from the protein elution peak. Band with apparent molecular weight of approximately 180 kDa and 90 kDa represented putative dimer and monomer respectively. Bands at earlier eluted fractions above the 225kDa marker represented putative higher-order oligomers. (c) Western-blot analysis of the receptors purified in DDTM. Protein was detected using anti-1D4 antibody against the C-terminal Rho-1D4 tag on $GABA_{B2}$ subunit.

Binding activity of $GABA_BR$ solubilised in the same set of detergent was measured in scintillation proximity assay (SPA). SPA measures the light signal emitted when a radioligand binds to receptor immobilised on the scintillation beads. The radioligand chosen was ^{3}H-baclofen, which binds to the orthrosteric site at the cleft in $GABA_{B1}$ VFT. ^{3}H-baclofen was chosen as it is a well-characterised $GABA_BR$ agonist and is widely used in $GABA_BR$ radiobinding experiments [167]. Receptor solubilised in DDTM was found to maintain the highest binding activity, closely followed by sample in DM (Table 4.1). Binding of receptor in other detergent was much lower, with the 3rd detergent in ranking at 80% of DDTM and the 4th at only 50%. Taken together with SEC data above, it appeared that protein solubilised in DDTM was most functional, while DM solubilised protein exhibited best solubility and modest activity.

Table 4.1 Binding activities of $GABA_BR$ solubilised in different detergents. 20nM purified $GABA_BR$ in different detergent was couple to beads and incubated with 100nM radioligand for 1 hr before counting. Non-specific binding was measured in presence of 100 µM unlabeled ligand and 0.05% detergent. Signal recorded as counts per min.

	DDTM	**DDM**	**DM**	**DMNG**	**LMNG**	**CYMAL6**	**CYMAL6NG**	**CYMAL7NG**
Total signal	2808	1993	2747	2195	1526	2442	1622	1360
Non-specific	677	525	682	504	426	906	788	567
Specific binding	2131	1468	2065	1691	1100	1536	834	793

As mentioned earlier, proper extraction of integral membrane proteins must not sacrifice protein functionality for an apparent yield. It seemed a reasonable approach to solubilise $GABA_BRs$ using the stringent detergent DM and then exchange them into the milder detergent DDTM. In order to make the detergent exchange step as smooth as possible, the solubilisation cocktail was a detergent mixture consisted of 9:1 w/w ratio of DM/DDTM. Since DM has a relatively high CMC (0.087%, comparing to 0.0026% for DDTM), DM micelles could be easily removed with a DDTM only wash

buffer. Solubilisation profiles of using detergent mixture of DDTM with several other stringent detergents was compared (Figure 4.2). Among the detergents tested, the DDTM/DM mix showed best improved protein yield with fewer aggregates. None of the other detergents (NG, bOG, DDM) tested in this approach showed significant superior extraction to the DDTM/DM combination.

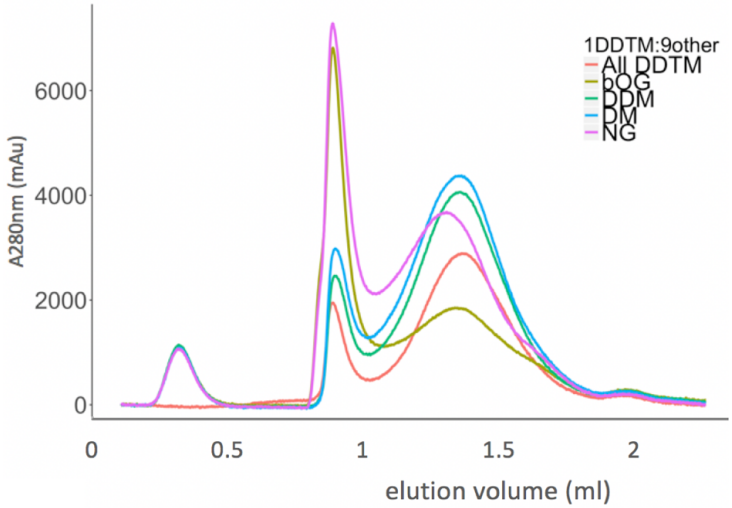


Figure 4.2 SEC analysis of protein purified using detergent mixture.

Protein was solubilised using DDTM 1:9 (w/w) to other detergent supplemented with 20:1 molar ratio αCHS. All samples were then washed with DDTM and αCHS only buffer.

Binding affinity of receptor was estimated by competition binding assay, where a constant amount of ^{3}H-baclofen was competed with increasing concentration of unlabelled baclofen. Apparent K_d of detergent solubilised $GABA_BR$ was estimated to be 1.4 μM in competition binding format (Figure 4.3a). The estimated K_d of 1.4 μM was approximately 20-fold lower than the affinity measured in brain slices [167].

The saturation binding assay format had also been tested, however binding saturation could not be achieved using up to 100 nM ^{3}H-baclofen. As the radioligand was dissolved in 70% ethanol, adding higher concentration might adversely affect protein stability and therefore was not tested. Furthermore, the effect of a positive

allosteric modulator GS39793 on baclofen binding affinity was tested. Using detergent-solubilised $GABA_BR$, no significant change in affinity was observed in presence in GS39783. Binding at different time points post-solubilisation was measured as an indication to evaluate receptor stability (Figure 4.3b). It appeared that receptor activity deteriorated quickly once solubilised out from the membrane, dropping to only 30% of the initial signal strength 26 hr post solubilisation.

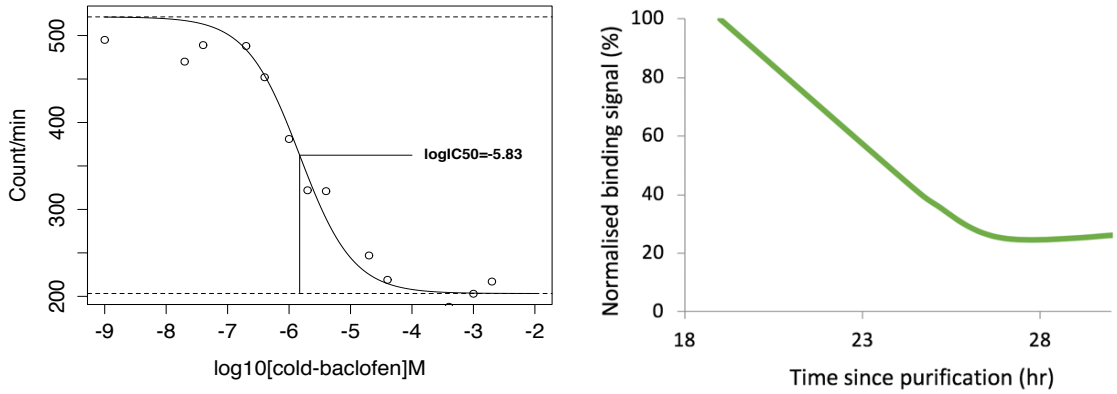


Figure 4.3 Binding activities of detergent solubilised $GABA_BR$.

(a) competition binding analysis. Purified $GABA_BR$ was incubated with constant 100nM radiolabeled-baclofen against various amount of unlabeled baclofen. (b) Receptor binding with 100nM radiolabeled-baclofen at different time points post cell solubilisation. Receptors were kept in elution buffer and only immobilised onto beads before each measurement.

Stability of solubilised receptors were also estimated by the SEC-based thermostability assay [168], utilising the advantages of Shimadzu analytical HPLC instrument discussed earlier. DDTM/DM solubilised and DDTM purified receptors were heated at various temperatures and analysed by SEC (Figure 4.4). The results followed protein denaturation status as samples were being heated, from which a melting temperature (T_m) was estimated. $GABA_BR$ incubated at 4 °C served as a reference curve, which gave a well-defined protein peak at 1.4 ml. A clear thermal shift of the peaks was observed as the incubation temperature increases. At 50 °C incubation, protein peak intensity reduced significantly and aggregation peak (at 0.9

ml) started to appear. Further increase in temperature caused more protein denaturation, as evidenced by the loss of a symmetrical protein peak and increase in void volume peak. The T_m was estimated to be approximately 65 °C, where half of the protein signal (based on height of protein elution peak) was lost.

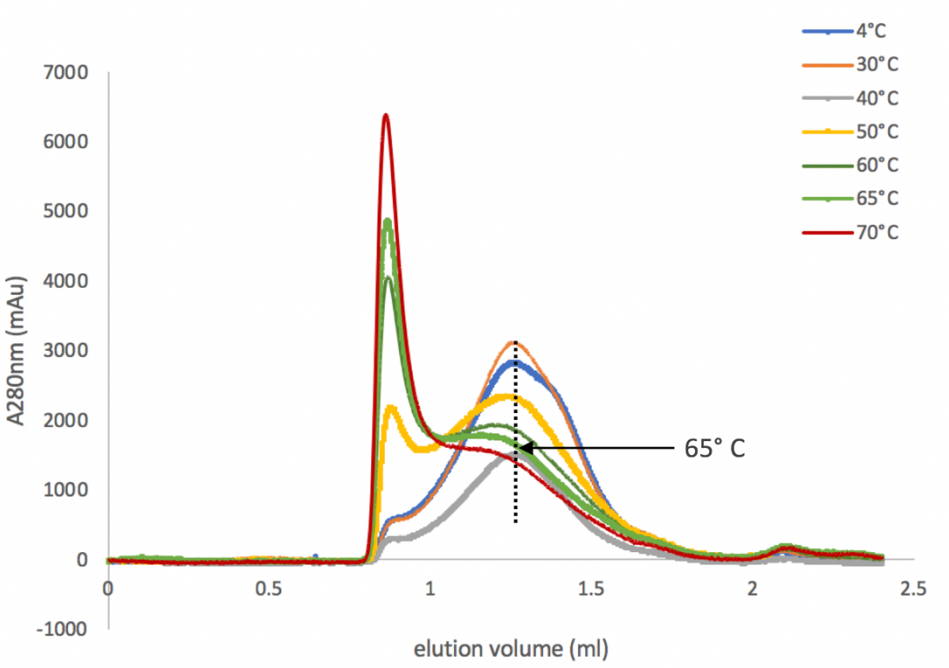


Figure 4.4 SEC-based thermostability analysis.

Detergent-solubilised samples were heated over a range of temperatures for 30 min in a thermal cycler, after which large precipitations was removed by high-speed centrifugation. Supernatant from each temperature was run onto a Superpose 6 increase 3.2/300 column.

4.2.2 Attempts in direct solubilisation

Strategies to solubilise the $GABA_BR$ together with its native surrounding environment were explored. The first candidate reagent tested was the Styrene Maleic Acid (SMA) co-polymer [169, 170]. Direct isolation of GPCR containing membrane patches using SMA co-polymer has been described before [171] and the solubilised SMA lipid particles (SMALPs) has been demonstrated to preserve protein functionality [158, 172]. Their mechanism of action has been proposed as the

following (Figure 4.5): SMA polymers make hydrophobic interactions with acyl chains exposed on membrane surface, followed by its insertion into a membrane core until polymers circularise around the membrane to form a disk structure [173].

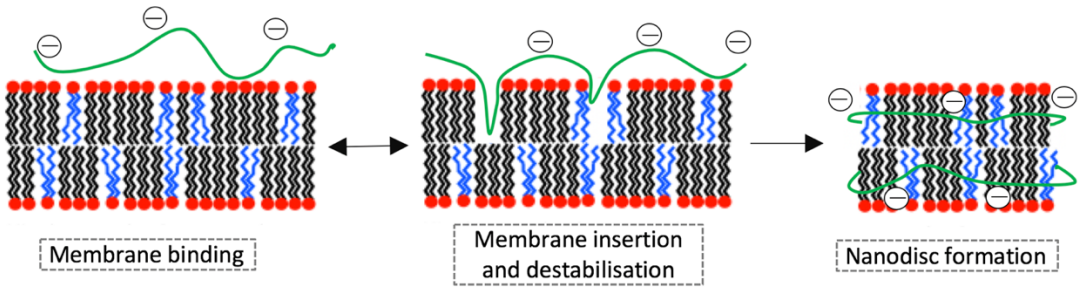


Figure 4.5 Schematic illustration of membrane solubilisation using SMA.

(1)Membrane binding through hydrophobic interactions with acyl chains exposed on surface of membrane. (2)SMA is inserted into the membrane core and destabilise the lipid bilayer. (3) circularisation of polymers to form the SMALPs.

Overnight incubation of mVenus tagged $GABA_BR$ expressing cells with SMA copolymer 3:1 was able to form native membrane particles containing folded target protein, as indicated by fluorescence observed in the lysed supernatant. Solubilisation efficiency was very low since strong fluorescence was observed in undissolved material. SMA-solubilised particles with $GABA_BR$ incorporated were purified through the N-terminal SBP-tag. $GABA_BR$ SMALP appeared as approximately 120 kDa on western blot (Figure 4.6a), which is 30kDa bigger than its usual apparent MW on gel. FSEC analysis of purified SMALPs gave multiple peaks (Figure 4.6b). Besides the void volume (peak 1) and free-mVenus peak (peak 4), there were two peaks eluted at 1.1 ml (peak 2) and 1.55 ml (peak 3) respectively. Peak 2 eluted just after the void was broad and asymmetrical, suggesting the presence of highly heterogeneous large particles. Peak 2 fraction was analysed by negative staining EM (Figure 4.6c) and showed highly heterogeneous clusters of large particles. Products in this peak might

correspond to the high MW band observed on western blotting, where migration of the large clusters was slower than detergent-solubilised particles upon gel electrophoresis. A small symmetrical peak at 1.55 ml (peak 3) could represent SMALP containing single a $GABA_BR$ heterodimer based on its elution volume. However, sample concentration in this peak was too low for characterisation by other means.

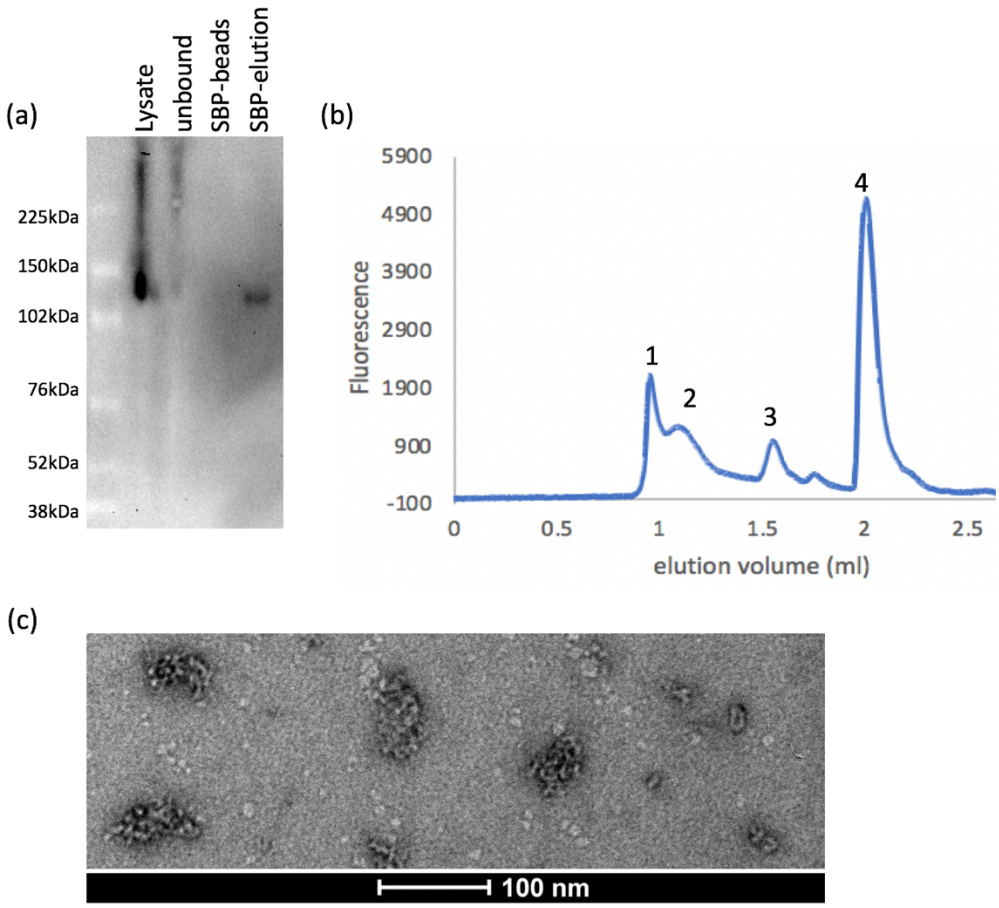


Figure 4.6 Analysis of SMA extraction products.

(a) Western blot analysis of $GABA_BR$ SMALP SBP-tag purification workflow. Protein bands were detected using anti-1D4 antibody against 1D4-tag at $GABA_{B2}$ subunit C-terminus. (b) FSEC profile of SBP-tag affinity purified product. Peak 1 at 0.9 ml represented the void volume. Peak 2 eluted at 1.1ml might contain $GABA_BR$ SMALP and was analysed by negative stain. Composition of peak 3 eluted at 1.55 ml was unclear. Peak 4 at 2 ml contained free mVenus-tag cleaved off the receptor. (c) Negative stained particles using peak 2 fraction at OD_{280nm} at 0.37 mg/ml.

Next, the possibility of direct solubilisation using the saposin D protein was tested. Saposin family proteins play roles in the degradation of sphingolipid from the

plasma membrane [5][174]. Their lipid-binding and membrane-interacting properties made them good candidates for forming native-membrane derived nanoparticles. Here, saposin D was chosen to form $GABA_BR$ containing nanoparticles since it has been previously reported to solubilise phospholipid-containing membranes [175]. Saposin D was expressed and purified from bacterial cells (purification profile of saposin protein described in appendix B).

Trial purification was performed using mVenus-$GABA_BR$ containing membrane homogenates at pH ranging from 5.5 to 7.5. Saposin D was shown capable of form $GABA_BR$ nanoparticles directly from membrane homogenates, however similar to SMA, the solubilisation efficiency was very low as suggested by the strong fluorescence in the un-solubilised material. Saposin D solubilised $GABA_BR$ was purified using SBP-tag and analysed by Western blot (Figure 4.7a). There was more receptor extracted at pH 7.5 comparing to lower pH conditions. Small amount of detergent (10-fold below CMC to prevent micelle formation) was added to reduce lateral pressure in membrane for easier penetration by saposin D. Addition of detergent during solubilisation was found to facilitate extraction, where a much stronger band of extracted $GABA_BR$ was observed on western blot (Figure 4.7a). Supplementation with ethanol was tested but no improvement in extraction efficiency was observed. FSEC profile of saposin D solubilised particles were very noisy (Figure 4.7b). Two protein peaks, at 1.4 ml (peak 3) and 1.6 ml (peak 4), were observed around the receptors usual elution volume. Products in the fractions were analysed by negative staining EM (Figure 4.7c), where particles from both peaks showed a tetrameric core with one or multiple 8 nm long Y-shaped extensions. No such tetrameric structure was observed

for negative stain particles of saposin D protein alone. Length of the extension could represent the $GABA_BR$ ECD fly-trap which was approximately 7 nm long in the crystal structure. It was suspected that saposinD-$GABA_BR$ nanoparticles formed a tetrameric centre from which the receptor ECD was displayed facing outward. However, the identity of the particles could not be confirmed since 2D classification was prevented by sample heterogeneity.

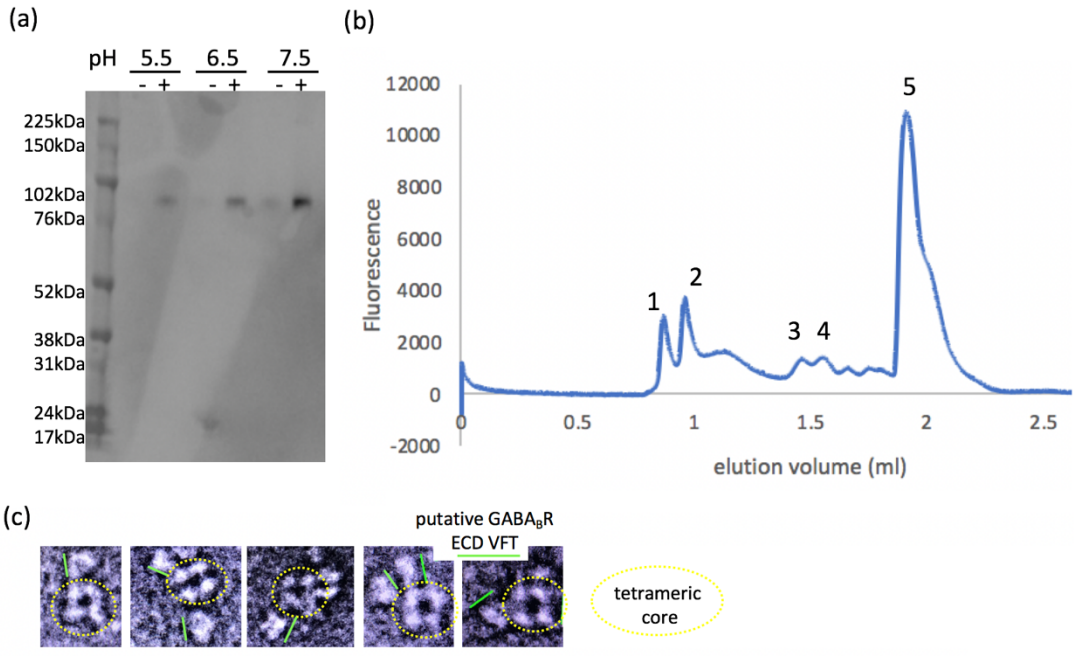


Figure 4.7 Analysis of saposin D extraction products.

(a) Western blot analysis of saposin D extraction at different pHs. Membrane homogenates were prepared and treated with saposin D overnight in different buffers. Insoluble materials were removed by ultra-centrifugation and the supernatant was purified using SBP-tag. For sample supplemented with detergent (0.001% LMNG), particles were washed into non-detergent buffer during purification steps. Lanes containing sample supplemented with detergent were marked by "+". $GABA_BR$ bands were detected using anti-1D4 antibody against $GABA_{B2}$ C-terminal 1D4-tag. (b) FSEC profile of sample extracted at pH 7.5 with saposin D and 0.001% LMNG. Sample was run on a Superpose 6 increase 3.2/300 column equilibrated with 150 mM NaCl, 10 mM Tris pH 7.5. (c) Representative negative stain EM particles of fractions from peak 3 and 4. Tetrameric core structure was circled by yellow dashed line. Y-shaped protrusions were marked by green line.

4.2.3 Optimisations on affinity purification

$GABA_BR$ affinity purification was designed as a dual-step purification with SBP-tag on the $GABA_{B1b}$ N-terminus and 1D4-tag on the $GABA_{B2}$ C-terminus to obtain pure heterodimers. Binding specificities of each affinity system were first assessed (Figure 4.8a). Detergent solubilised receptors were purified by 1D4-tag only purification, SBP-tag only purification and double-step purifications. The band with apparent MW around 90 kDa corresponded to the theoretical weight of $GABA_BR$ monomers. A minor band at around 200 kDa represented putative receptor dimers. The stronger monomer band was probably due to subunit dissociation upon SDS treatment. The identical bands were observed on western-blotting (Figure 4.8b), confirming the identity of the protein. Single step purification using the 1D4-immunoaffinity system alone yielded high purity products consisted of only the $GABA_BR$ (Figure 4.8a), suggesting high specificity of this tag. SBP alone purification product was not as clean as from 1D4-tag (Figure 4.8a), with extra bands seen across multiple molecular weights. Band at approximately 140 kDa MW had a consistent appearance in SBP purification product and was analysed by mass spectrometry. The band was identified to be HSPA5 which is a protein folding chaperone. As the SBP-tag was placed on the receptor's N-terminus, protein that was still being folded was possible to be captured by the streptavidin beads. The combination of 1D4 and SBP double-step purification gave high purity product (Figure 4.8a). The order in which the affinity-purification steps were performed was not important as it did not affect final product quality and quantity.

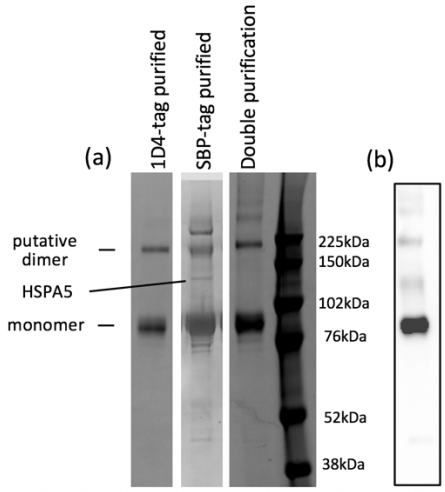


Figure 4.8 Analysis of purification specificity of each tag.

(a) SDS-PAGE analysis of $GABA_BR$ purified using 1D4-immunoaffinity tag alone, SBP-tag alone and double-purified product. (b) Western blot of the double purified product. Protein bands were detected using anti-1D4 antibody against $GABA_{B2}$ C-terminal 1D4-tag.

Binding capacities of the two types of beads were measured next: using 100 µl of each beads, 12 ml of solubilised membranes and 200 µl elution buffer as a starting point. The purification products were analysed by western-blot (Figure 4.9). SBP beads purification appeared to be very efficient, with most proteins in the elution and very little in the flow through (ft) solution or left on beads. The results suggested that current SBP-tag purification protocol was sufficient to capture and elute all protein of interest, thus no further optimisation was required. On the other hand, significant amount of proteins was lost during 1D4-beads purification. Major loss happened on the beads, where a smeared band across the sample lane was detected. The strong band signal suggested considerate amount of proteins failed to elute from the beads and was forming aggregates. There was also a small amount of protein left in the flow through solution, indicating suboptimal binding efficiency to the 1D4-antibody beads. The same batch of 1D4-antibody beads was used to purify at least 3-fold quantity of

another 1D4-tagged membrane protein, therefore the possibility of poor quality resin could be eliminated.

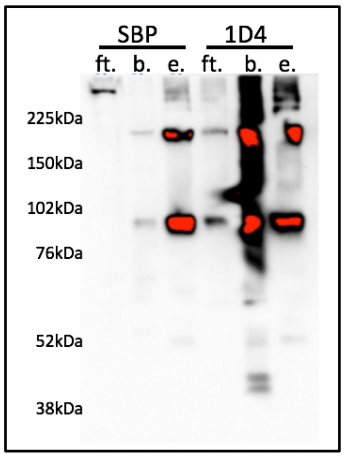
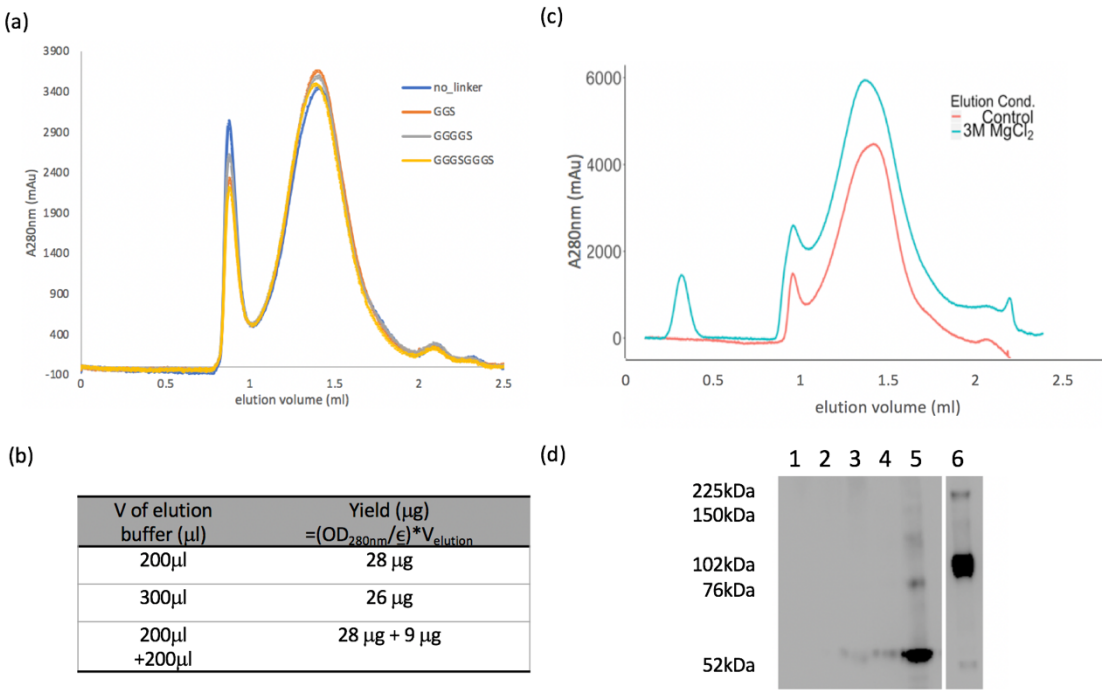


Figure 4.9 Purification efficiency analysis of SBP-tag and 1D4-tag.

Flow through (ft.), beads (b.) and elution product (e.) using each purification tag were analysed by western blot.

Next, strategies to optimise the 1D4-tag affinity purification step was tested. One possible reason for reduced efficiency of capturing tagged $GABA_BR$ could be steric hinderance by large detergent micelles as the tag was placed straight after the receptor intracellular coiled-coils. Linkers of different lengths were introduced between the C-terminal end of the protein and the 1D4-tag. Receptors carrying linkers of various lengths were expressed, purified and analysed by SEC. These constructs showed no noticeable differences in their SEC profiles (Figure 4.10a) comparing having the 1D4-tag directly linked to protein sequence. These data suggest that the 1D4-tag was not occluded by detergent micelles and is consistent with the previous report that 1D4-antibody binding is not affected by detergent. Increasing volume of elution buffer did not improved the total yield of eluted protein, where a 2^{nd} elution using equal volume of buffer eluted 30% of the amount of protein in the first elution (Figure 4.10b). Addition of 3 M $MgCl_2$ was found to improve yield by 20% without affecting protein

SEC profile (Figure 4.10c). The amount of protein left on beads was greatly reduced after adding the second elution step and $MgCl_2$ supplement (Figure 4.10d).



V of elution buffer (µl)	Yield (µg) $=(OD_{280nm}/\epsilon)*V_{elution}$
200µl	28 µg
300µl	26 µg
200µl +200µl	28 µg + 9 µg

Figure 4.10 Optimisation of the 1D4-tag purification step.

(a) SEC analysis showing effects of linkers between protein and 1D4-tag. (b) Protein yield using various amount of elution buffers. 1 g cell pellet was used for each experiment, where protein was captured using 100µl 1D4-beads. Various amount of elution buffer was used as indicated. (c) SEC analysis showing effect of adding 3M $MgCl_2$ during elution. (d) Western blot analysis of 1D4-purification products post optimisation. Lane 1-6 contained cells, lysate, 1D4 flow-through, 1D4 wash, 1D4 beads and elution product respectively.

It appeared that the standard SBP-tag purification and post-optimisation 1D4-tag purification protocol was ready for scaling up. Initial large-scale purification from large scale 2 L culture failed at 1D4-affinity purification step. Almost all protein was found to have precipitated out of solution during overnight elution. Purification of large scale 2 L and medium scale 500 ml from the same batch of culture were performed in parallel, using identical materials at the same proportions. Again, protein eluted overnight in 1 ml 1D4-beads (i.e. from 2 L culture) had almost no reading on SEC at 1.4 ml, where medium scale was purifying the expected amount (Figure 4.11).

The result suggested that having more than 250 µl 1D4-beads during elution could be the reason for protein precipitation. This was consistent with the earlier observation that protein was more likely to aggregate when bound on 1D4-beads. Similar observations had been found for large-scale purification of other membrane proteins using the 1D4-tag (personal communication with Dr. Ester Behiels - STRUBI). The 1D4 purification step was thus performed in batch splitting the sample in multiple tubes containing 500 ml culture in each. The limitation in working volume was a 1D4-tag specific problem and scaling up using SBP-tag did not show any adverse effect on protein yield.

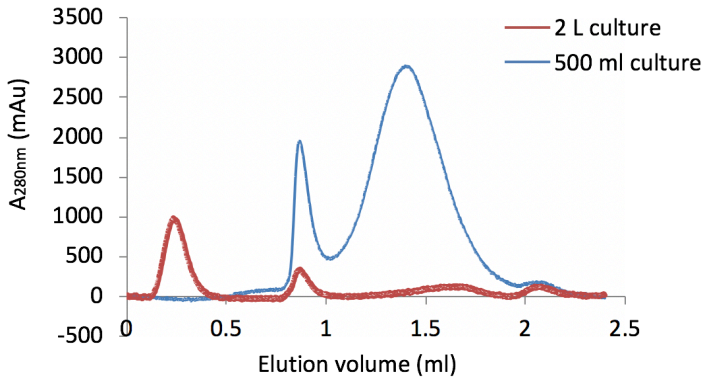


Figure 4.11 Comparison of 1D4-tag purification products from large scale 2 L purification (red) and 500 ml medium scale purification (blue).

4.2.4 Effect of different gel filtration columns

Gel filtration was performed as the last purification step to remove aggregates and higher-order oligomers. Figure 4.12 depicts the SEC experiment for separation of $GABA_BR$ in different oligomeric states. The $GABA_BR$ was eluting in a wide peak, where higher order oligomers were observed in SDS-PAGE (Figure 4.12b) of early fractions. It should be noted that the exact oligomeric state of SEC fractions could not be confirmed based on SDS-PAGE, as membrane proteins are often find to form

oligomers during gel electrophoresis experiments. Nevertheless, only later fractions which corresponded to lower molecular weight bands without oligomers were collected for subsequent analysis. It was worth noting that the choice of gel filtration column was critical for purifying $GABA_BR$. A sample from the same batch was loaded on different columns and the results suggested that Superose 6 Increase column provided the suitable matrix for the $GABA_BR$ sample separation as compared to Superdex 200 column. $GABA_BR$ was found to aggregate on Superdex 200 column, causing all protein to elute in the void volume.

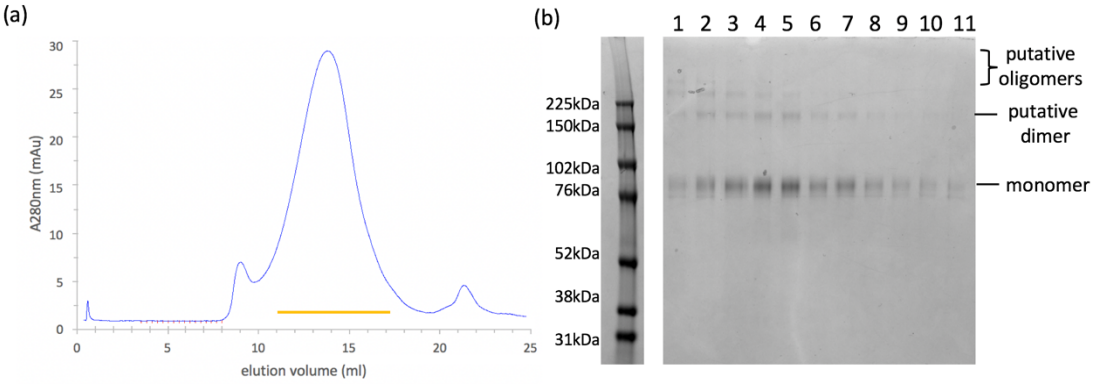


Figure 4.12 Representative SEC results from double-purified $GABA_BR$.

(a) SEC profile of protein run on a Superpose 6 increase 10/300 column equilibrated in 150 mM NaCl, 10 mM Tris pH 8.0, 0.01% DDTM 20:1 αCHS. (b) SDS-PAGE analysis of gel filtration peak fraction (underlined in yellow).

4.2.5 Optimisation of purification buffer condition

Optimisation of purification buffer started by screening buffer pH values ranging from pH 6.2 to 8.0 (Figure 4.13a). pH 8.0 (at 4 °C) was found to give the highest protein peak: 30% more than at pH 7.2. Reducing pH to 6.2 dramatically diminished protein yield to less than a quarter of the quantity at pH 8.0. The shoulder coming from potentially higher-order oligomers was seen across all pHs. Relative number of oligomers was less at pH 8.0, where the intensity of dimer protein peak was stronger.

The smaller fluorescent product that eluted at later volume was from the cleaved mVenus tag.

Next, the effect of supplementation with $GABA_BR$ endogenous ligand GABA was tested. Addition of GABA at 1mM caused 20% reduction in the protein yield comparing to non-supplemented sample (Figure 4.13c). Ca^{2+} was known to interact and stabilise the closed conformation of B1 subunit flytrap [99]. 2mM Ca^{2+} ions concentration, which was close to the Ca^{2+} level in extracellular synaptic cleft, was tested. Having Ca^{2+} ions throughout the purification did not significantly enhance protein peak intensity (Figure 4.13d). An additional small peak at 2 ml elution volume, possibly degradation product, was seen in the 2mM Ca^{2+} supplemented sample. L-arginine had been reported to protect protein C-terminal cleavage [176]. Addition of 50mM L-arginine was found to enhance protein yield by 20%, as estimated by peak intensities at identical elution volume as the control (Figure 4.13b). Gel filtration profile without L-arginine supplement had a more symmetrical yet broad peak (Figure 4.13b). As discussed earlier in section 4.2.1, the apparent peak symmetry was due to poor resolution of different oligomeric states of the receptor. Although, 50mM L-Arginine supplemented sample was asymmetrical, an increased intensity at elution volume of the dimer peak was obvious.

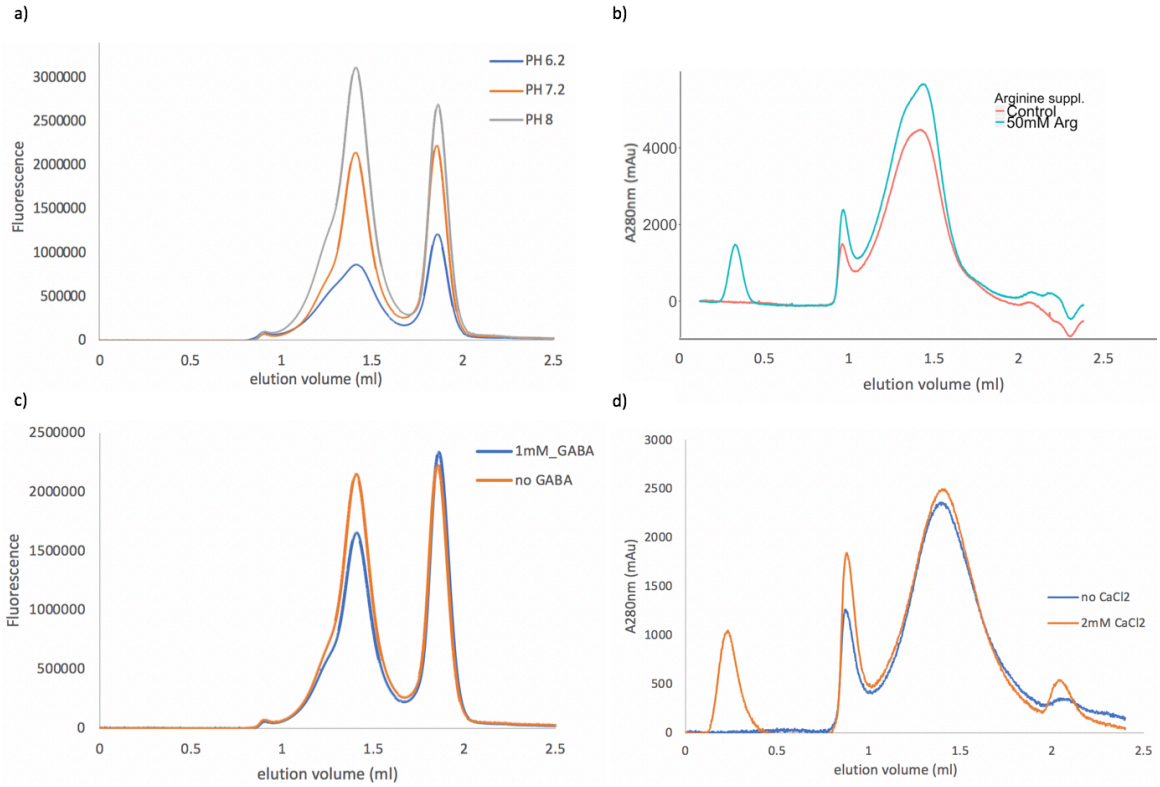


Figure 4.13 SEC profiles of $GABA_BR$ purified in different buffer conditions.

For each panel, the same cell pellet was aliquoted to be purified in parallel under the different conditions specified. (a) Cells were re-suspended and purified using buffer pH at 6.2 (blue), 7.2 (orange) and 8 (grey) at 4° C. (b) Effect of adding 50mM L-arginine in the purification buffer. (c) Effect of adding 1mM GABA in the purification buffer. (d) Effect of adding 2mM Ca^{2+} in the purification buffer. Supplements were added in buffer throughout purification and excluded in the SEC running buffer (150 mM NaCl,10 mM Tris-HCl pH 8.0 and 0.01% DDTM 20:1 αCHS).

The FSEC-based thermostability assay was used next to examine stabilising effects of a wider range of supplements, including $GABA_BR$ antagonists, agonists and other common purification additives (Figure 4.14). Receptors were purified in one batch and aliquoted for incubation with each additive before heat treatment. $GABA_BR$ at 4° C served as a positive control, which had a symmetrical protein peak and no aggregates peak. Receptor was heated to 65 °C (T_m estimated in section 4.2.1) in the absence of ligand to serve as a negative control, where the SEC peak was reduced to half of the 4° C sample. Comparing to the SEC curve without ligand, none of the additives tested here showed improved protein thermostability. Supplementation

with 10 μM GABA and 100 μM allosteric modulator GS39783 almost completely diminished protein signal. Based on the data none of these ligands was included in buffer during purification in the further experiments.

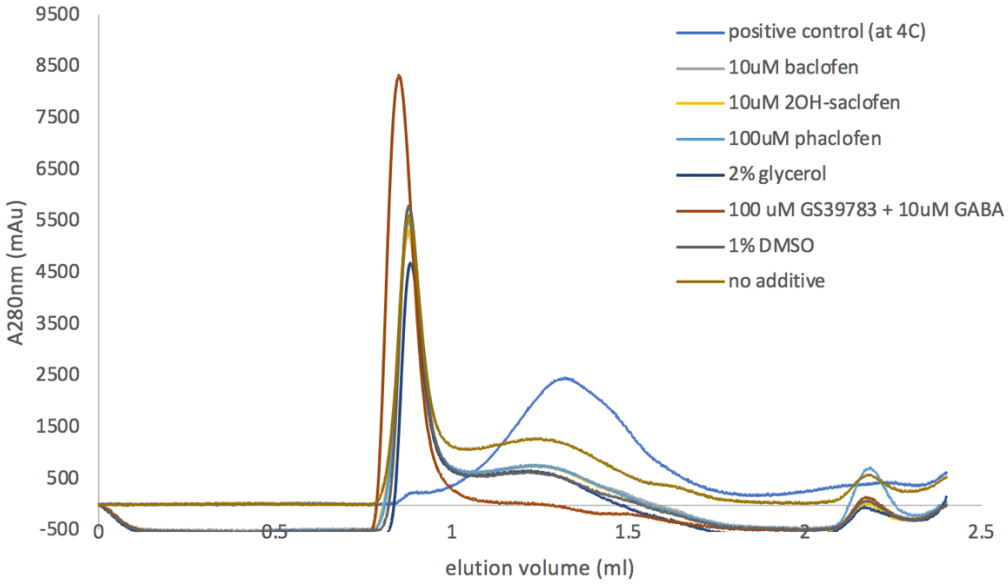


Figure 4.14 Effects of small molecule supplements on receptor thermostability.

Detergent-purified $GABA_BR$ was aliquoted and heated at T_m for 30 min in the presence of different ligands. Large precipitates were removed by high-speed centrifugation and the supernatant was analysed on SEC. Positive control (blue) was protein incubated at 4 °C. Sample was heated in the absence of ligands (brown), with GS39783 100uM and 10uM GABA (red),with baclofen 10uM (light grey),with 2-OH saclofen 10uM (yellow), with phaclofen 100uM (light blue), with 2% glycerol (dark blue), and with 1% DMSO (dark grey). All ligands were used at least 10-fold higher concentration than their reported affinity to ensure binding saturation.

4.2.6 Reconstitution of detergent solubilised receptors

Although, protein yield had been substantially improved after the optimisation steps, the quantity of protein attainable (approx. 0.05 mg per litre of culture) was still at the lower range for setting up crystallisation trials. If crystallographic approach were to be pursued, expression and purification of sufficient amount of protein would present a major rate limiting step. More importantly, receptor displayed poor stability when solubilised in detergent where data in section 4.2.1 had shown fast declining

rate in receptor's ligand binding activity. In addition, it was observed that protein solubilised in detergent had the tendency to aggregate quickly once concentrated to above 3 mg/ml. Therefore, the poorly stable detergent-solubilised receptor was unlikely to form high-quality crystals. Cryo-EM technique circumvents some of the requirements mentioned above. Current protein yield provided sufficient materials for cryo-EM grid preparation. In addition, receptors structural integrity is more likely to be better protected during cryo-EM sample preparation, as the proteins would be preserved in a hydrated state in vitreous ice immediately after purification. Given the above reasons, cryo-EM was chosen as the main technique for this project.

Cryo-EM grid was first prepared using detergent-solubilised $GABA_BR$. Grids made $GABA_BR$ in DDTM detergent micelles was of poor quality and unsuitable for data collection. Most of the protein was absorbed to the carbon support and formed large aggregations in free-standing ice (Fig.4.15).

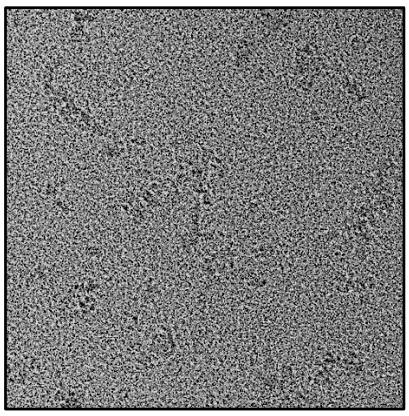


Figure 4.15 Representative micrograph of detergent-purified $GABA_BR$ particles.

Sample was prepared using receptors solubilised in 0.01% DDTM 20:1 αCHS at 1.6mg/ml on unsupported grid and imaged on a TF30 Polara microscope.

Consequently, feasibility of reconstituting detergent-solubilised receptors into alternative agents were investigated. Currently, there are two possible ways to transfer detergent solubilised protein into other reconstitution systems. One way required complete protein purification in detergent and then exchange as purified samples, as demonstrated by formation of the Salipro lipid nanoparticles [177, 178]. $GABA_BR$ reconstitution using this approach resulted in large amount of precipitation during the exchange step. As a result, reconstitution efficiency was low and highly variable. The other method is "on-bead" reconstitution, where receptors were exchanged into reconstitution agents while still bound on affinity-beads during purification [179]. The "on-beads" reconstitution approach was shown to be more suitable for $GABA_BR$, giving decent protein yield, reproducible results and stable sample. Furthermore, affinity purification efficiency was not affected after reconstitution into other systems. $GABA_BR$ reconstitution into various agents were thus further explored using the "on-bead" approach. Five reconstitution agents, namely amphipols, saposin A, saposin D, MSP1D1 and MSP2N2 were tested in parallel. Quality of the reconstitutions was evaluated by comparison of their elution profiles using SEC (Figure 4.16).

$GABA_BR$ reconstituted into saposin A, MSP 1D1 and amphipols showed similar properties upon SEC analysis. They gave symmetrical protein peak with similar elution volumes. $GABA_BR$ in amphipols gave slightly earlier peak at 1.45 ml, followed by saposin A reconstituted particles, and MSP1D1 particles appeared to be the smallest at around 1.5 ml. Peak intensity was higher for $GABA_BR$ in saposin A, followed by MSP1D1 nanodiscs then amphipols. The sample reconstituted in MSP2N2 had the

highest peak intensity, almost 2 times stronger than from MSP1D1. MSP2N2 was the only one showing a significant shoulder at earlier elution volume, indicating existence of higher-order oligomers. Saposin D reconstitution was suboptimal when compared to the others, giving a very broad peak with low intensity. Saposin D sample also had a much earlier elution volume than all other samples. The data suggested that dimensions of saposin D particles were much bigger, which is consistent with direct-solubilisation results in section 4.2.2. It was plausible to suspect that saposin D forms particles of a much bigger diameter. Cryo-EM analysis of sample in reconstitution system were proceeded in the following orders: amphipols, Saposin A, MSP1D1 and MSP2N2, based on qualities of their SEC profiles. Details of reconstitution, sample preparation, EM grid assessment and data analysis of $GABA_BR$ in amphipols and nanodiscs were elaborated in chapters 5 and 6 respectively.

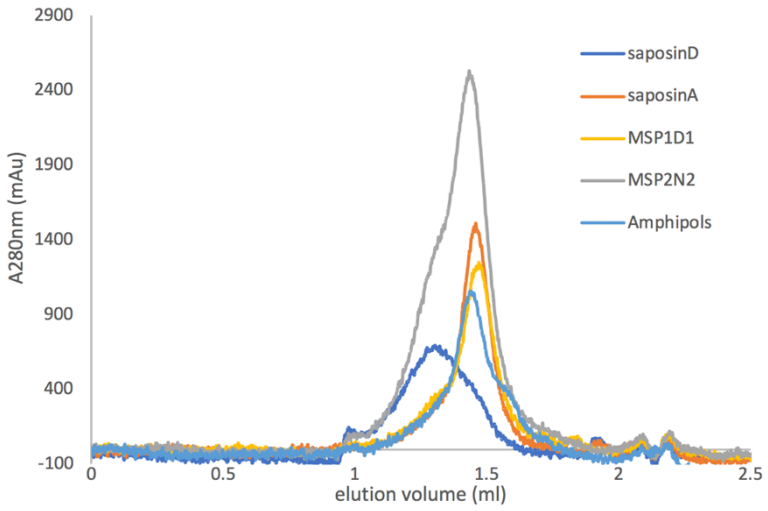


Figure 4.16 SEC profiles of $GABA_BR$ reconstituted into different systems.

The protein was solubilised with DDTM/DM detergent mix and bound onto SBP-beads for reconstitution. Subsequent purification steps were performed using the same protocol as before except the exclusion of detergents in the buffer.

4.3 Discussion

4.3.1 Establishment of a multi-step workflow for Cryo-EM sample preparation

In this chapter, the full workflow from membranes solubilisation, protein purification to receptor reconstitution was established. The protocol had been optimised and showed reproducibility over multiple sample preparations. A large-scale purification (2~4 L culture) using frozen cell pellets could be performed in 2 days. A brief summary of the workflow is as follows: cells were solubilised using detergents cocktail containing 9 to 1 mass ratio of DM to DDTM supplemented with 20:1 molar ratio αCHS. Detergent-solubilised receptor was captured using affinity beads and has the option of being reconstituted into alternative systems on beads. Receptor was purified by two-rounds of affinity purification and a final gel filtration step. Typically, 0.04 mg of pure and homogeneous heterodimeric $GABA_BR$ could be purified from 1 L culture, with cell density of 2 million cells per 1 ml culture.

Effective solubilisation using a detergent that preserves protein functionality is crucial and choosing the suitable detergent of a given membrane protein has to be determined on an individual basis. DDM, despite being the most popular detergent in GPCR studies, did not show superior characteristics in solubilising $GABA_BR$. DM, with shorter alkyl chain makes it more efficient in disrupting cell membrane integrity, was found to be the most efficient for $GABA_BR$ extraction out of the detergents tested so far. However, a protein is often completely stripped of native lipids by using such detergents, which can have two side effects for structural studies: on the positive side, this may be useful for preparing homogenous protein samples: on the other side, the loss of native lipids may lead to impaired protein function. DDTM, a milder detergent

than DM, was found better at retaining $GABA_BR$ ligand binding activity. Therefore, a mixture of both detergents was used during extraction, and later washed into DDTM only buffer for purification. The usage of a combination of detergents is not uncommon and this strategy had been successfully utilised in solving the structures of the neurotensin receptor [180] and CB_2 cannabinoid receptor [181].

It had been demonstrated that even after optimisation, receptors activity deteriorated quickly in detergent micelles. It was therefore necessary to search for non-detergent solubilisation systems. Solubilisation trials using saposin D and SMA polymers showed that these agents were capable of extracting nanoparticles containing $GABA_BR$ and its surrounding native membranes. Previously, the solubilisation effects of saposin D was reported to be more active at lower pH [175]. The disparity with our observation that pH 7.5 yield highest amount of $GABA_BR$ was probably because instability of $GABA_BR$ at lower pH, as evidenced by data from purification buffer pH screening in section 4.2.5. In addition, $GABA_BR$ was expressed in HEK293 $GnTI^-$ cells where large sugars were absence, which might have improved access of saposin D for extraction at a more physiological pH. Based on the SEC profiles, it is tempting to hypothesize that saposin D assembles into much larger particles than other types of conventional reconstitution agents. Similarly, particles of a defined size was not observed from SMALPs, although it was reported to form particles of 10 nm diameter [173, 182]. For both reagents, the biggest obstacle preventing their usage in the structural studies was the very low yield. For example, SMALPs had been used in several biochemical studies [158], but very limited applications in structural studies [169, 183]. Further optimisation would require considerate time and resources, which deviate from the main scope of this thesis, and was therefore not proceeded. Possible

optimisation steps could include increasing the ratio of solubilisation reagent to $GABA_BR$ membranes. Supplementation of anionic phospholipids could be tested to facilitate saposin D membrane solubilisation, for which saposin D has shown preferential interaction [175]. In contrast, SMA co-polymer binds un-specifically to lipids without discrimination between different lipids [173]. If sufficient amount of purified SMALPs could be obtained, it would represent the most close-to-native lipid bilayer environment.

The combination of 1D4-tag and SBP-tag utilised benefits from both systems, creating a high specificity and affinity purification protocol. SBP-tag based purification was shown to be highly efficient in capturing and eluting $GABA_BR$, with modest protein purity. Although, some contaminants were pulled using the streptavidin beads, the additional affinity purification step with 1D4-tag was able to remove the unwanted contaminants. Rho-1D4 immunoaffinity purification has been shown to yield high purity $GABA_BR$ with a single round of affinity purification. A limitation of the 1D4-tag observed was the lower yield of purified protein per volume of resins. Major defect was the large amount of residual proteins stuck on 1D4-antibody coated beads. 1D4 peptide was used as the elution agent and elution efficiency was not 100% due to its limited binding affinity to the beads. Addition of 3M $MgCl_2$ eluted 25% more protein without affecting quality of sample, where the higher salt concentration might alter peptide-antibody interactions to facilitate protein elution [184]. The high ionic strength in solution might have also helped to improve $GABA_BR$ solubility [185], increasing the amount of folded protein eluted. A second elution was helpful to collect 30% more protein, where addition of more peptide further shifted equilibrium for

elution of the target protein. Using the above strategies, most intact folded receptors could be harvested. A possible way to alleviate the limited elution efficiency was to introduce a cleavage site between the receptor and the affinity tag. TEV or 3C protease cleavage directly from affinity beads/column is a commonly used strategy in protein purification. However, these proteases required reducing environment for their activity where there is the risk of disrupting the heterodimer association. Other cleavage enzymes, such as typsin, are not specific so may result in unwanted cleavage within receptor. The protein cleavage strategy appeared not suitable for $GABA_BR$ and therefore not pursued. Alternating the 1D4-tag to other membrane protein friendly tags could be tested in the future, suitable candidates included the eGFP tag or the Flag tag.

Purification buffer conditions were also screened. A range of pH environment was tested, as proton concentration in solution had been suggested to affect membrane protein stability [186]. The most suitable environment for $GABA_BR$ was found to be at in Tris-HCl buffer pH 8 at 4 °C, which resembled physiological pH 7.17 at 37 °C. Many proteins were found to be stabilised by the binding of a small molecule. Effects of a range of supplements were tested, however none of the additive tested so far showed improved $GABA_BR$ stability. Possible points of optimisation in the future included NaCl concentration screening, as the iconic strength of solution had been reported to affect GPCR functionality [186].

4.3.2 Monitoring receptor stability and function

Stability of the purified receptors could give good indication of their likelihood to achieve high resolution structure and sometimes can guide sample preparations.

There are many options for measuring protein stability, such as CD spectrometry, sedimentation velocity or dynamic light scattering. However, these methods require milligrams quantity of sample and were therefore not suitable given the current $GABA_BR$ yield. The Shimadzu HPLC system, which had been intensively used in this study could be combined with other means to monitor changes in protein states. An example is the FSEC-based thermostability assay, where samples are heated at different temperatures to study its unfolding behaviours [168]. The apparent T_m of the detergent-solubilised $GABA_BR$ was estimated to be 65°C, a lot higher than other GPCRs analysed. The higher apparent T_m obtained here could be due to the bRIL insertion which is known to stabilise GPCRs.

Another commonly used assay in GPCR studies is the radioligand binding assay. This method has the advantage of requiring small amounts (in nanograms scale) of sample and giving very sensitive read-outs. The apparent K_d for ^{3}H-baclofen using purified $GABA_BR$ solubilised in detergent was estimated to be 1.4 μM, approximately 20-fold lower than when measured in brain slices [167]. Reduced drug binding affinity in this range has been reported in the recombinant $GABA_BR$ [40] and other detergent solubilised membrane proteins [131]. Missing native membrane components and steric restrictions from the detergent belt could be the reasons for the reduced activity. Addition of a positive allosteric modulator GS39793 did not affect the apparent K_d of baclofen. The proposed binding site of GS39793 is in the $GABA_{B2}$ transmembrane region, which might have been hindered by the detergent micelles. The “super-plus” assay format could be used in future, where the membrane is equilibrated with ligands before detergent solubilisation to allow interaction with transmembrane region [187].

4.3.3 $GABA_BR$ reconstitution for cryo-EM studies

Cryo-EM was chosen as the primary tool for the structural studies of the $GABA_BR$ for the following reasons. Firstly, Cryo-EM grid provides a more natural environment for sample preservation comparing to X-ray crystallography. In cryo-EM, hydrated sample is rapidly frozen in liquid ethane to preserve their 3D structure in vitreous ice. Where in crystallography, highly concentrated protein is left to form crystal over a period of time, requiring protein to be stable for a longer period of time. Secondly, recent advancements in EM field made structural studies of small proteins of $GABA_BR$ size possible. Finally, cryo-EM approach obviates some of the hurdle requirements for crystallographic studies that are difficult to meet for this sample, for example high protein yield and concentration.

Detergents that are powerful solubilising agent, but are found not suitable for $GABA_BR$ cryo grid preparations. Membrane protein structures solved in detergents always had to go beyond 1 mg/ml concentration as getting particles into the free-standing ice had been a common obstacle for detergent-solubilised samples [188, 189]. For $GABA_BR$ sample this was particularly challenging as the receptor tend to aggregate quickly at such high concentration. Cryo-grids made with detergent-solubilised $GABA_BR$ at >1 mg/ml were found to form aggregates in the hole of unsupported grids. In the future, grid with support such as graphene oxide coating could be tested.

Reconstitution into artificial environments other than the detergent-micelles was investigated. “On-beads” reconstitution was found to suitable for this sample as this method avoided sample concentrating and minimised $GABA_BR$ exposure in detergents. The reagents tested here associate with the membrane using different

mechanisms and showed variability in their SEC profiles. Reconstitution using amphipols, saposin A, MSP1D1 and MSP2N2 yielded more homogenous particles based on SEC results. Amphipols are expected to bind tightly around the hydrophobic region of the receptor thus not alter the apparent particles size significantly from detergent-solubilised receptor. The other four (saposin A, saposin D, MSP 1D1 and MSP 2N2) all belonged to the lipid nanodiscs category. Typically, a MSP nanodisc contains >100 lipid molecules and two MSP molecules [190]. The size of MSP nanodiscs are pre-defined and controlled by the length of the MSP used. MSP2N2 particles eluted earlier, as this MSP was bigger from a fusion construct of MSP1D1 and MSP1D2, thus generated particles with bigger diameter. It was unexpected that MSP1D1 particles had smaller apparent size than $GABA_BR$ in amphipols, assuming there were some lipid molecules incorporated into the MSP1D1 particles. Saposin A and saposin D are both derived from the prosaposin protein [174] and interact with different type of lipids. Nanoparticles formation using the saposin protein is less well characterised than the MSP. Based on their SEC profiles, saposin A particles was more suitable for single particle analysis but not saposin D.

4.4 Conclusion

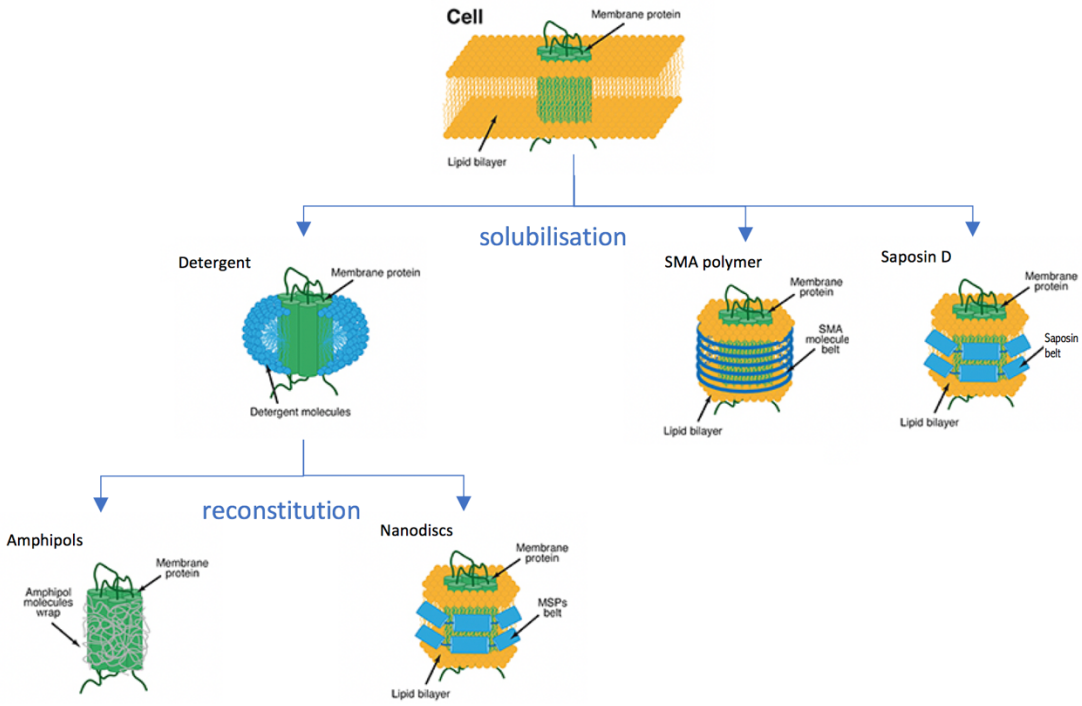


Figure 4.17 Solubilisition systems tested in this chapter.

Direct solubilisation of $GABA_BR$ containing membranes was assessed using detergent, SMA polymer and saposin D. Detergent solubilisation using a mixture of DM and DDTM were found to be suitable for $GABA_BR$ solubilisation. Extraction efficiency of SMA polymer and saposin D was too low. Detergent solubilised receptors could be later reconstituted into amphipols or nanodiscs.

In this chapter, the workflow to solubilise, purify and reconstitute the $GABA_BRs$ for cryo-EM studies was established. Membrane solubilisation step was optimised with care to retain protein function. Homogeneous heterodimers had been purified by the double-tag purification system. Figure 4.17 summarises the different solubilisation systems tested so far. Although highly desired, quantity of protein solubilised using non-detergent reagents was not sufficient for structural studies. Receptors reconstituted in amphipols and nanodiscs appeared to be suitable candidates for cryo-EM studies.

Chapter 5. CRYO-EM ANALYSIS OF $GABA_BR$ RECONSTITUTED IN AMPHIPOLS

5.1 Introduction

This chapter aims to further analyse and characterise the amphipols-reconstituted particles obtained in chapter 4. Although screening had been performed using biochemical approaches in chapter 4, it was difficult to predict whether the sample would be to able achieve high resolution reconstruction structures. To obtain a high-resolution model, the principle of single particle analysis requires the following criterion to be met: firstly, 2D projections of particles from different angles; secondly, conformational homogeneity of particles. In an ideal case, a large copy number of particles in an identical conformation would be imaged from all orientations. For small-sized particles like the $GABA_BR$, such information was not easily observed using a screening microscope and could only be properly assessed with a highly sensitive detector. This chapter describes analysis of amphipol-reconstituted cryo-EM data sets collected on a 300 kV Polara microscope with a K2 detector.

Amphipol polymers are a class of weak surfactants commonly used as stabiliser for membrane proteins [191, 192]. Amphipol A8-35 is the best characterised polymer among the amphipols family [193]. Its chemical structure consists of a hydrophilic polyacrylate chain, onto which octylamine and isopropylamine sidechains are randomly grafted. Amphipols bind tightly around the hydrophobic region of a membrane protein to protect it from aqueous solutions. Replacing detergent with amphipols was a crucial step in obtaining several high resolution cryo-EM structures, including the TRPV2 channel [194], human γ-secretase [195] and V-ATPase [196].

Amphipol reconstitution has many advantages over retaining purified protein in detergents micelles. For example, amphipols have a much slower dissociation rate and as a result are able to keep the protein stable for a longer period of time [197]. In addition, amphipols have been reported to significantly improve the biochemical properties of membrane proteins [191]. For cryo-EM data collection purpose, no free amphipols is required in solution, where excess detergents required for detergent-solubilised sample are commonly found to diminish image contrast.

5.2 Results

5.2.1 $GABA_BR$ reconstitution into amphipols

Amphipols reconstitution process was summarised as illustrated in Figure 5.1. Briefly, detergent-solubilised receptors were supplemented with 1:8 (w/w) ratio of receptor to amphipols. After incubation for >3 hr, a detergent-amphipol-receptor complex was formed. The detergent micelles were then removed by bio-beads, while receptors transmembrane region was protected by amphipols.

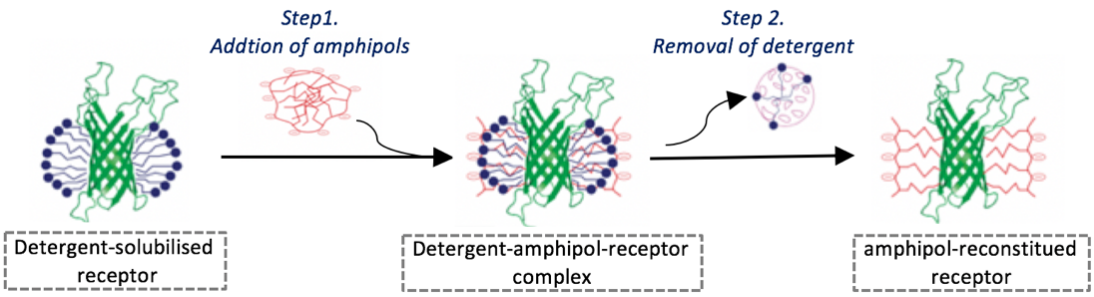


Figure 5.1 Schematic illustration of the amphipols exchange process. Detergent-solubilised receptor was supplemented with amphipols to form the detergent-amphipol-receptor complex. Detergent is then removed by absorption using porous bio-bead. Figure was modified from [198].

No significant protein loss was observed during the amphipol exchange step based on measurement of protein absorbance at 280nm before and after the reconstitution. Receptors in amphipols were also found to be less likely to aggregate when concentrating, suggesting better protein solubility. SEC profiles of the same sample before and after amphipol reconstitution were compared (Figure 5.2a and 5.2b). Detergent-solubilised receptors gave a symmetrical yet broad protein peak at 13 ml. Amphipol-exchanged receptors showed two protein peaks: one peak at same elution volume of 13ml and a new, higher peak at 14 ml. There was no peak in the void volume in the amphipol-exchanged sample, suggesting no aggregate formation. Protein compositions in both peaks of the amphipols-exchanged sample was analysed by SDS-PAGE (Figure 5.2c). Gel electrophoresis migration pattern of amphipols-reconstituted receptors was very similar to when solubilised in detergent (e.g Figure 4.2), with major bands at approx. 80 kDa and 190 kDa corresponded to the monomeric and putative dimeric $GABA_BR$. Minors bands at higher MW were also observed, which could come from higher-order oligomers or an artefact created while running the gel. Results from SDS-PAGE was more supportive for the existence of higher-order oligomers, as the peak fraction for the high MW bands was earlier than the peak of the monomer bands. The different oligomeric states were not completely separated during SEC and as a result gave overlapped bands on SDS-PAGE. Gel lanes from the new peak at 14.4 ml contained mainly bands at putative dimeric and monomeric MW and no new band. SDS-PAGE result suggested that the later peak at 14.4 ml was not from degraded receptors, although receptor's oligomeric state in this peak could not be confirmed. One possibility is the formation of an increasing population of monomers when reconstituted in amphipols. This possibility was eliminated later in

section 5.2.2, where no monomeric receptor 2D class was observed from particles imaged using this peak fraction. Another possibility is that receptor's oligomeric states were not affected by amphipols reconstitution, but rather, they now separate better during SEC. Amphipols-exchanged $GABA_BR$ showed high stability, where the reconstituted protein kept at 4°C was stable for up to 10 days based on OD measurements. Preliminary radioligand-binding using amphipol-solubilised $GABA_BR$ showed similar binding signal as in the receptor in detergent, a complete binding analysis was not performed for the amphipol-reconstituted sample.

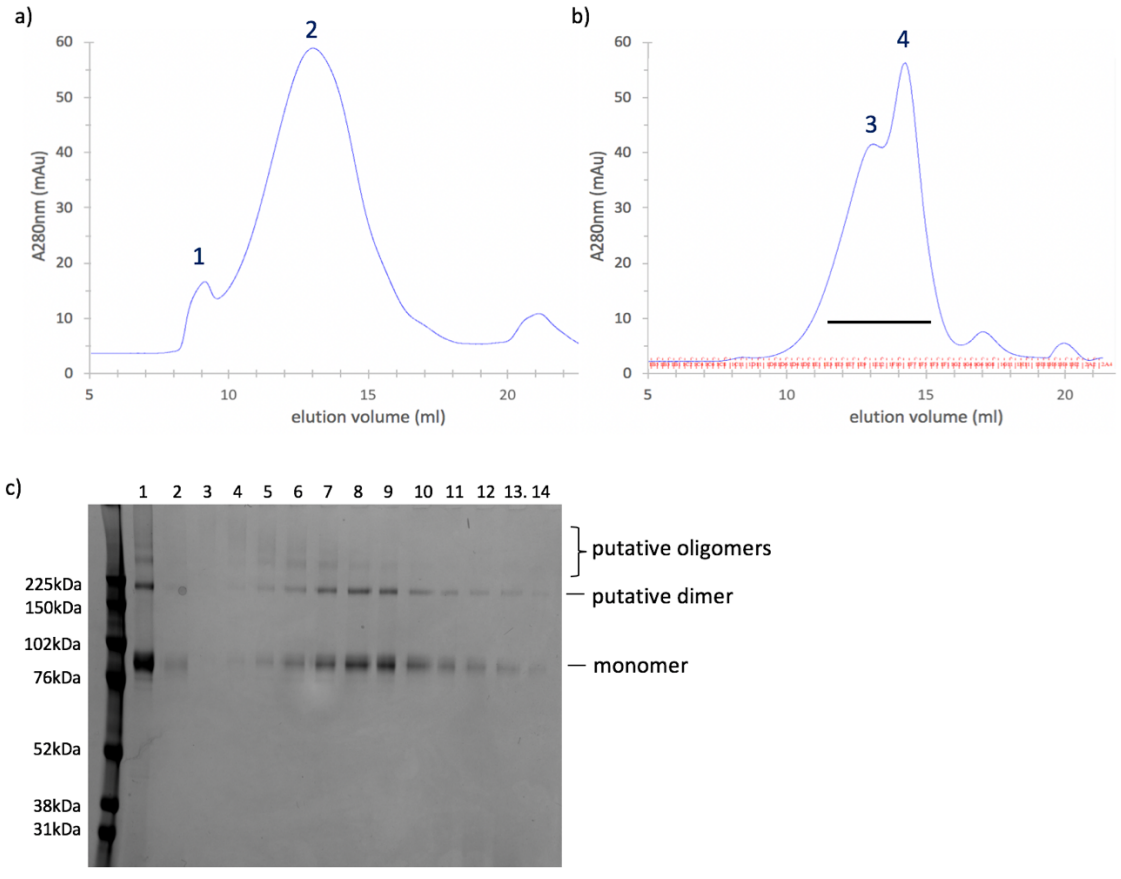


Figure 5.2 Biochemical characterisation of amphipol-reconstituted $GABA_BR$.

(a) SEC profile of detergent-solubilised $GABA_BR$. Peak 1 at 9 ml represents the column void volume according to manufacturer's protocol. Peak 2 at 13 ml contains detergent-solubilised $GABA_BR$ as confirmed by previous SDS-PAGE analysis (e.g Figure 4.12). Sample was run on a Superose 6 increase 10/300 column (GE healthcare) equilibrated in 150 mM NaCl, 10 mM Tris pH 8.0, 0.01% DDTM 20:1 αCHS. (b) SEC profile of amphipol-reconstituted $GABA_BR$. Peak 3 and peak 4 had elution volume of 13 ml and 14 ml respectively. Fractions from these peaks (underlined in black) were analysed by SDS-PAGE

in c. Sample was run on a Superose 6 increase 10/300 column (GE healthcare) equilibrated in 150 mM NaCl, 10 mM Tris pH 8.0. (c) SDS-PAGE analysis of fractions from amphipol-reconstituted SEC. Lane 1 contained concentrated sample injected onto the column. Lane 2-14 contains elution fractions from area underlined in curve b.

Amphipols-exchanged $GABA_BR$ behaved more like soluble proteins on cryo-EM grid. 0.06 - 0.12 mg/ml protein concentration range was found to yield optimal particle distribution, which was much lower than the >1 mg/ml required for detergent-solubilised receptors. $GABA_BR$-amphipols sample gave monodispersed particles in free-standing ice (Figure 5.3a). Although individual receptor domains were not recognizable on the raw micrograph, the overall shape and size of particles (Figure 5.3b) were in line with those expected for the $GABA_BR$ particles (Figure 5.4).

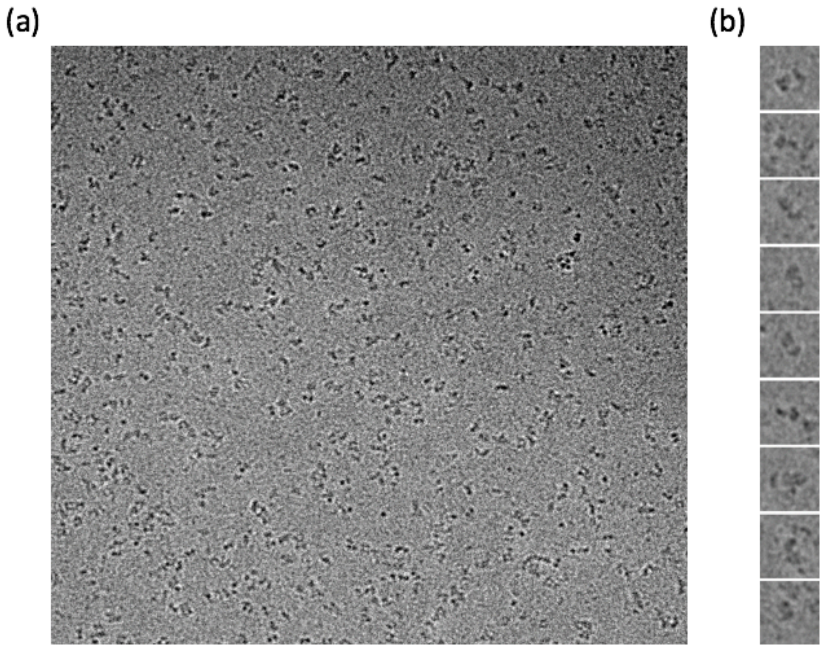


Figure 5.3 Amphipols-reconstituted $GABA_BR$ Cryo-EM sample.

(a)Representative micrograph of amphipols-reconstituted $GABA_BR$ in free-standing ice. Micrograph showed uniform distribution of single particles. Grid was prepared using protein at 0.07 mg/ml concentration. Micrograph was collected on a TF30 Polara microscope with a K2 detector at 3.5 micron defocus. (b) Selected particles corresponding to the size of $GABA_BR$. Particles were shown in a 25 nm long box.

5.2.2 Data processing and analysis of $GABA_{B1b}$-bRIL/$GABA_{B2}$-amphipols particles

The first complete cryo-EM data set was collected on a sample which had a bRIL insertion in the $GABA_{B1}$ subunit (Figure 5.4), with the intention of facilitating subunit differentiation and accurate particle alignment. A set of 922 movies were collected on a FEI TF30 Polara microscope with a K2 detector (Table 5.1).

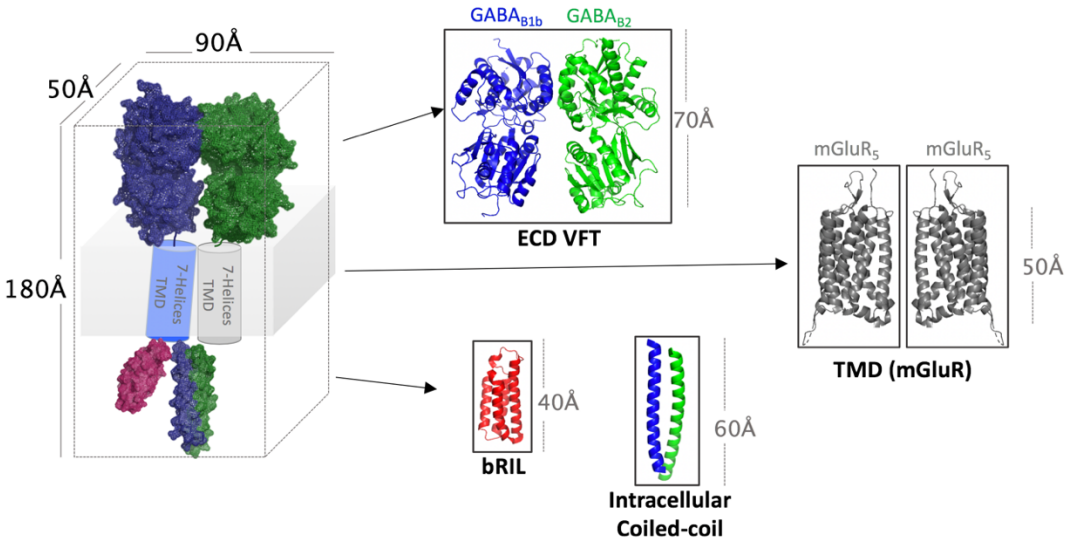


Figure 5.4 Model of the recombinant $GABA_{B1b}$-bRIL/$GABA_{B2}$ receptor.

The model was built using known structures of the $GABA_BR$ ECD VFT, $mGluR_5$ TMD, $GABA_BR$ intracellular coils and bRIL. $GABA_{B1b}$ and $GABA_{B2}$ subunits are coloured in blue and green respectively. bRIL is coloured red. Figure was generated by pyMOL using structures 4MQE, 4PAS and 4OR2.

Table 5.1 Data collection statistics for $GABA_{B1b}$-bRIL/ $GABA_{B2}$-amphipols

Microscope	FEI TF30 Polara
Voltage (kV)	300
Imaging mode	Counting
Detector	Gatan K2-Summit
Nominated Magnification	37,037x
Pixel size (Å)	1.35
Total Dose (e–/Å^2)	83
Exposure time (s)	20sec
Dose rate 2 (e$^-$/pix/s)	8
Number of frames	50
Frame length (s)	0.4
Focusing range (µm)	-1 to -4

342,439 particles were picked and extracted for 2D classifications. 2D classes showed clear densities for the ECD and TMD domains (Figure 5.5). The ECD appeared as butterfly-shaped projections, corresponding to extracellular dimeric venus fly-trap (VFT) (Figure 5.6b). The VFT domain exhibited the same arrangement as in the reported heterodimeric crystal structure where the two subunits each consisted of two lobes, sat side-by-side to each other. The bigger, single globular shaped density directly connected to the ECD represented the receptor transmembrane helices wrapped in amphipols (Figure 5.6b). The global shape of the 2D class averages confirmed the dimeric state of the amphipol-reconstituted $GABA_BR$.

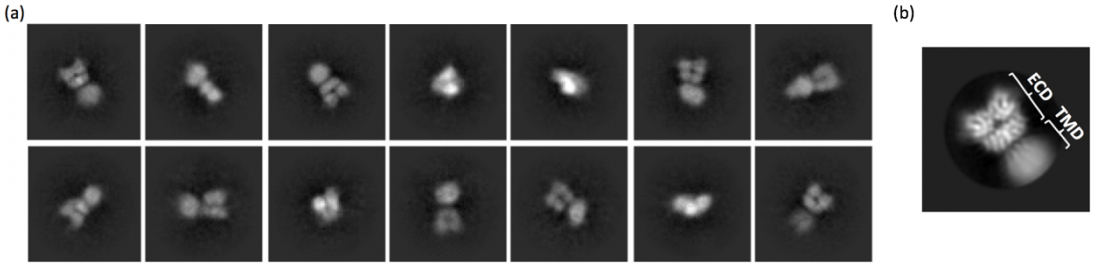


Figure 5.5 2D classification results of $GABA_{B1b}$-bRIL/ $GABA_{B2}$-amphipols particles.
(a) Representative 2D classification averages showing projections from different angles. (b) Selected 2D average showing clear ECD and TMD density.

Automatically picked particles were cleaned through iterative rounds of 2D classifications. Most class results showed homogeneous population of single particles. In some classes, density for the upper lobs of the VFT was incomplete (Figure 5.6a), indicating protein unfolding in the region. This could be due to damage occurring at the air-water-interface during grid preparation or protein degradation during purification. Particles from classes showing incomplete ECD density were removed. Projections from various views were observed, and views from three angles were the most populated. Approximately 40% of particles would fall in top/bottom views

(Figure 5.6b), 25% particles with front views (Figure 5.6c), 15% particles with side views (Figure 5.6d) and the rest coming from other views.

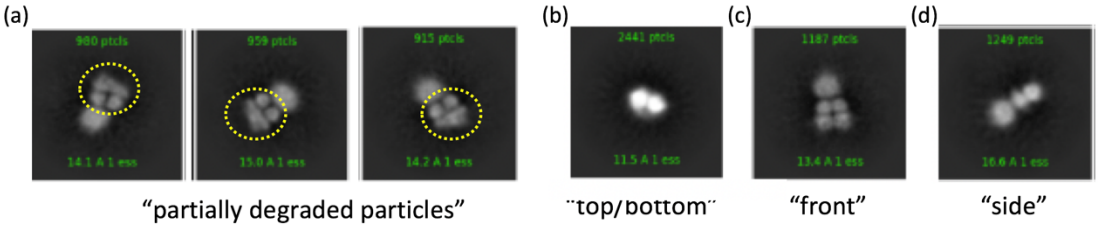


Figure 5.6 Representative 2D class averages.

(a) classes showing partial degradation in the ECD top region (circled by yellow dash lines). (b) top/bottom view. (c) front view. (d) side view.

During 2D classification, the ECD VFT density was well defined and started to show secondary structural features. Density for the TMD region however was less clear. Density for intracellular domain was not observed, suggesting very poor alignment of this region. From 2D classification a subset of 53,618 particles, with balanced number of particles across different views, were selected for 3D classification.

Initial 3D classification was run with a manually built reference model using crystal structures listed in Figure 5.4. All subsequent classifications were run using 3D models generated from the data. 3D reconstruction from consensus classification had two regions of density accounting for the receptor's ECD and TMD domain (Figure 5.7). Domain arrangement in 3D agreed with the proposed receptor model, where the ECD sits upwards above the TMD. The ECD density showed the classical fly-trap arrangement with the two subunit sitting side-by-side to each other facing opposite directions. The previously solved *apo* crystal structure (PDB ID: 4MQE) could be docked into the ECD density as a rigid-body with most residues buried in the density

(Figure 5.7). Although 2D alignment showed secondary structural features in ECD, information of the same level was not apparent during 3D reconstruction. The TMD region density appeared as a deformed-globular structure and did not show distinctive structural features. TMD density was estimated to be 6nm long and 5.5 nm wide, and 5nm deep across the membrane based on dimensions of the ECD. Docking of the 7-helices TMD structure from another Class C GPCR ($mGluR_5$, Figure 5.5) was attempted, however no reliable docking could be obtained due to the lack of features in the reconstructed density. Density for linker between ECD and TMD domains were only apparent under lower thresholds display, suggesting high flexibility in the region. Clear density for the intracellular region ($GABA_BR$ coiled-coil and bRIL) was not visible in the 3D model. Despite testing a number of 3D classification approaches, the resolution of the 3D reconstruction was restricted to 13 Å. Splitting particles into different classes (3 – 15 classes tested) gave multiple models at lower resolutions. Overall, the models were of similar quality as the one shown in Figure 5.7 and no class was distinctively better than the others.

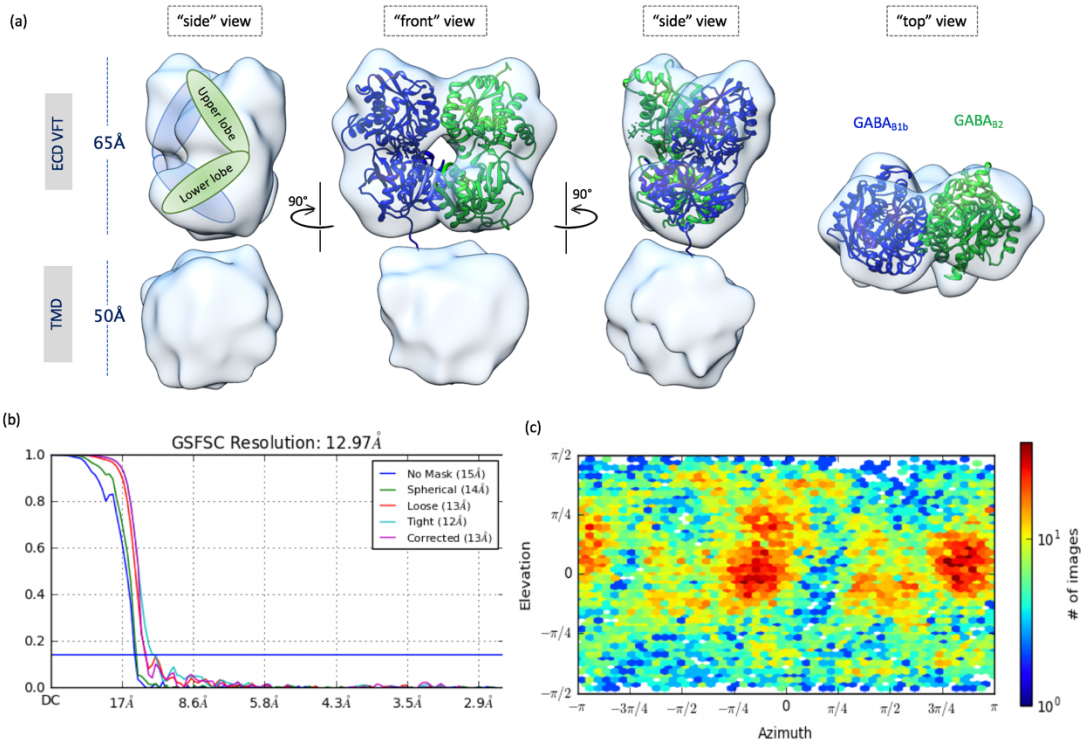


Figure 5.7 3D reconstruction model of GABA$_{B1b}$-Bril/$_{B2}$ reconstituted in amphipols

(a) Model viewed from three angles corresponding to the front, side, and top views. VFT domains from two subunits were arranged next to each other, facing opposite directions. GABA$_B$R ECD VFT crystal structure (PDB ID 4MS3) was fitted as a rigid body in UCSF Chimera. Assignment of specific subunit was random for illustration purpose. Model density was displayed at 0.28 threshold. (b) Fourier Shell Correlation (FSC) curve analysis. (c) Azimuth map showing particles distribution from different orientations.

Localised reconstruction strategies including masked local classification [199] and multi-body refinement [200] were applied in order to better resolve each domain separately. These methods did not improve reconstruction of the TMD. For the ECD region, masked classification improved its resolution to 10 Å showing improved features such as local protrusions (Figure 5.8). As the GABA$_{B1b}$ and GABA$_{B2}$ subunit structures are highly similar, quality of fitting was similar in both orientations at resolution of the model therefore density for the specific subunit was not readily distinguishable.

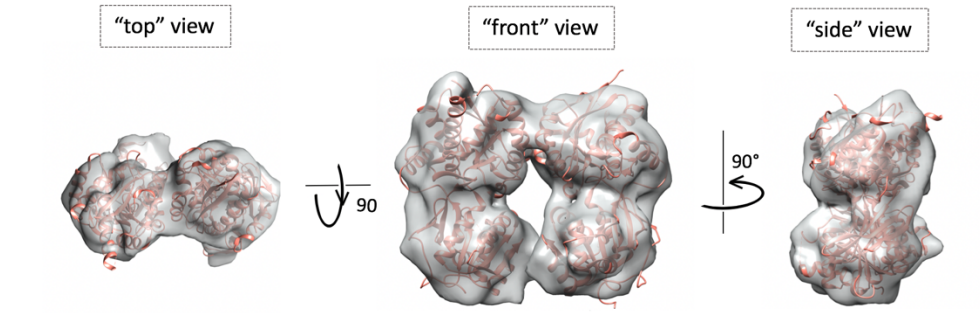


Figure 5.8 Localised reconstruction of ECD region with VFT crystal structure fitted.

(a) ECD VFT local reconstruction model shown as top, front and side view. $GABA_BR$ VFT crystal structure was fitted as a rigid body using UCSF Chimera. Model density was displayed at threshold 0.5. Model figures were generated in UCSF Chimera.

5.2.3 Data processing and analysis of $GABA_{B1b}$-bRIL/$_{B2}$-Cys-amphipols particles

To improve sample rigidity, a pair of cysteine mutations was introduced onto the B1b-bRIL/B2 receptor ($GABA_{B1b}$-bRIL/$_{B2}$-Cys). Positions of point mutation (Figure 5.9) were designed based on the active state crystal structure (PDB ID: 4MS4), with the intension to lock receptors in a closed conformation. Stabilising the protein in the closed conformation would create constitutively active receptors, impact of which has been demonstrated in cellular binding assays [55].

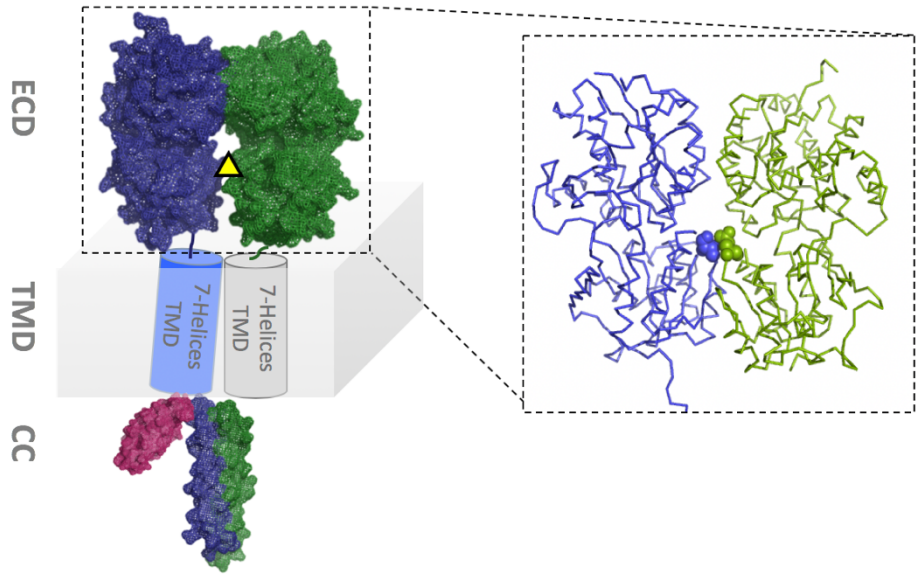


Figure 5.9 Location of point directed mutations (yellow triangle) in the receptor.

Zoomed in ECD model showing mutations (spheres, B1-Thr^{198}Cys and B2 -Gln^{206}Cys) located in the lower lob in the flytrap. $GABA_{B1b}$ and $GABA_{B2}$ subunit colored in blue and green respectively.

Suspension cells co-transfected with the mutant constructs grew slower comparing to the wildtype. Expression level of the cysteine mutant receptors was approximately 70% of wildtype based upon protein A_{280nm} peak in SEC (Figure 5.10a). Elution volume of the mutant receptor remained the same as the mutation was not imposing any change on the overall dimensions of the receptor. Protein was again reconstituted into amphipols and analysed by SEC (Figure 5.10b). Similar to section 5.2.1, peak elution time for this amphipols-reconstituted sample had a slight delay in volume (at 1.45 ml) compared to when solubilised in detergent-micelles (at 1.4 ml). Protein composition of the peak was analysed by SDS-PAGE (Figure 5.10c) and showed high degree of sample purity, revealing bands corresponding only to $GABA_BR$. The fresh fraction showing no oligomer bands was taken to make Cryo-EM grid directly

without further concentration. Grids were prepared at 0.13 mg/ml protein concentration and gave suitable particle distribution on grid (Figure 5.10d).

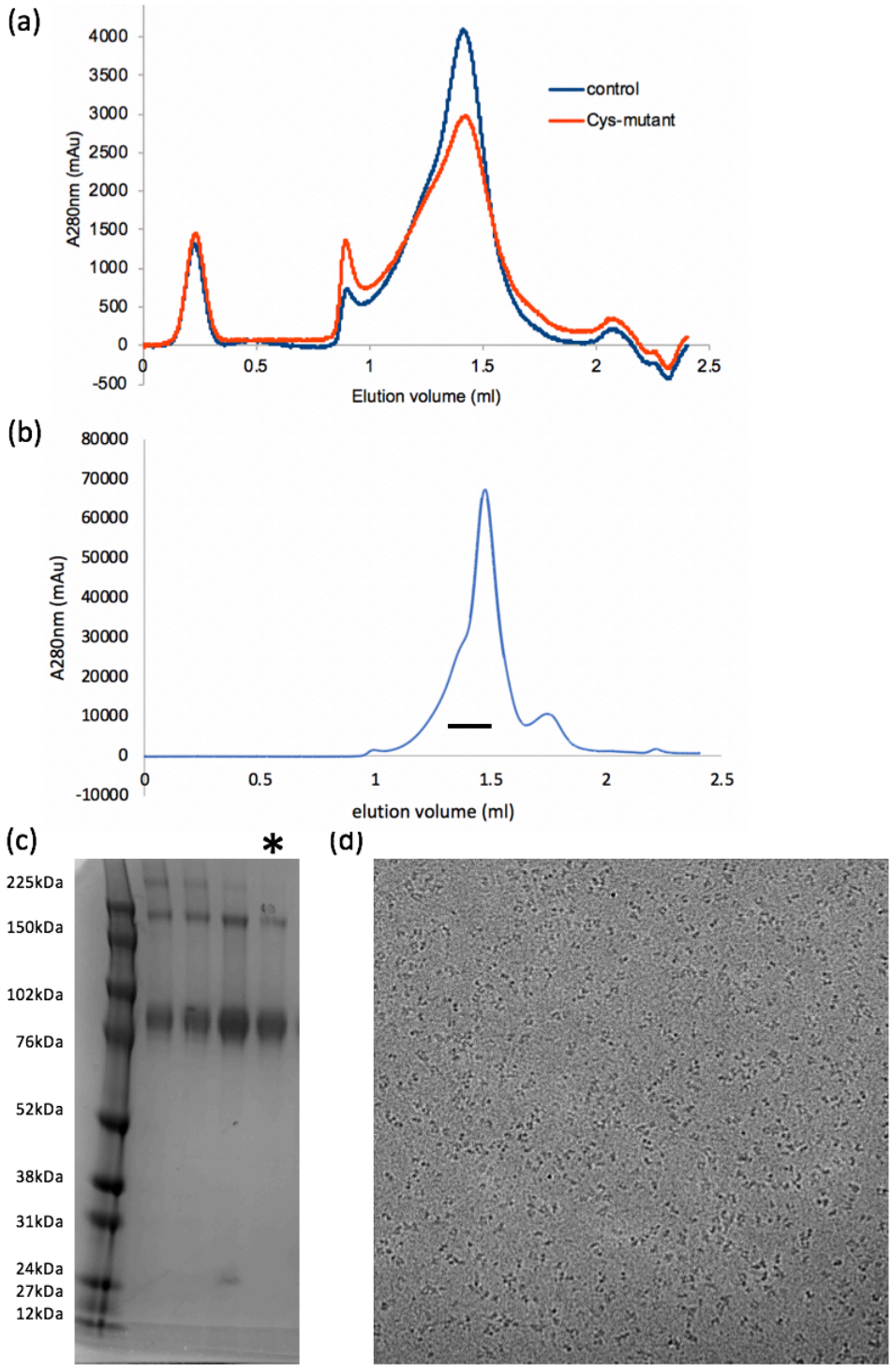


Figure 5.10 Characterisation of $GABA_{B1b}$-bRIL/$_{B2}$-cys mutant receptor.
(a) Yield comparison of $GABA_{B1b}$-bRIL/$_{B2}$ and the cys-mutant receptor. (b) SEC profile of amphipols-reconstituted $GABA_{B1b}$-bRIL/$_{B2}$-cys mutant receptor. (c) SDS-PAGE analysis of peak fractions underline on curve in b. Fraction with no sign of higher order oligomers (marked by *) was taken to prepare EM grids. (d) Micrograph of Cryo-EM grid prepared using protein sample at 0.13 mg/ml concentration. Sample was imaged on a TF30 Polara microscope with K2 detector.

Mutant receptors showed 4-fold higher affinity to ^{3}H-Baclofen in preliminary radio-binding experiments, confirming stabilisation of receptors in the closed (i.e. active) conformation. Baclofen supplementation at 100 μM was used to further stabilise the closed receptor conformation. The concentration used was well above the reported K_d (0.06 μM) and the same amount of baclofen had been used in obtaining the activated state ECD crystal structure [55]. A set of 1054 movies was collected on this sample on a TF30 Polara Microscope with K2 detector (Table 5.2).

Table 5.2 Data collection statistics for $GABA_{B1b}$-bRIL/$GABA_{B2}$-Cys data set

Microscope	FEI TF30 Polara
Detector	Gatan K2-Summit
Voltage (kV)	300
Imaging mode	Counting
Slit (eV)	20eV
Nominated Magnification	37037 x
Pixel size (Å)	1.35
Total Dose (e−/Å^2)	61
Exposure time (s)	14 sec
Dose rate 1 (e−/Å^2/s)	10.7
Number of frames	35
Frame length (s)	0.4
Focusing range (μm)	-0.8 to -2.6

624,355 particles were picked using representative 2D averages from section 5.2.2 as templates. When compared with 2D classification results for non-cysteine linked particles in section 5.2.2, Cys-mutant receptor particles displayed similar overall shape and orientations (Figure 5.11). However, 2D averages from the mutant receptor particles were less resolved, as indicated by the fuzzy outlines observed in the 2D classes. The ECD of mutant receptor was also less defined and did not show clear secondary structural features.

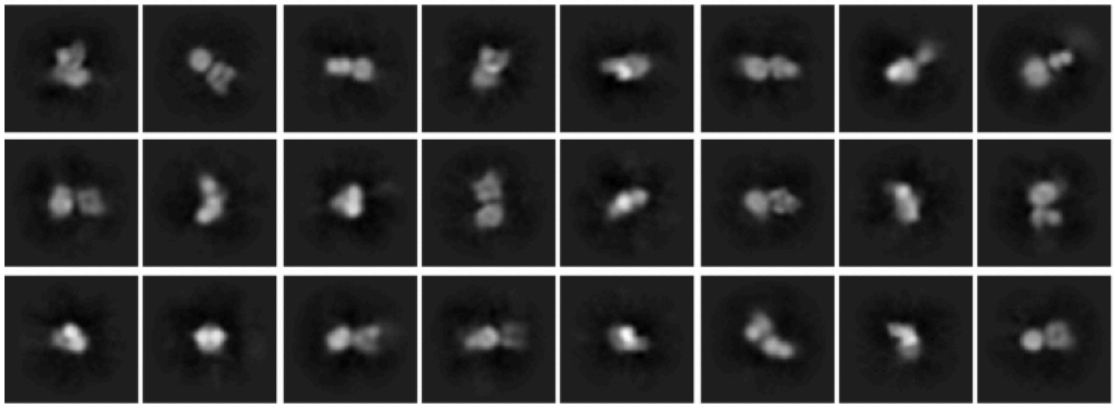


Figure 5.11 Representative 2D classification averages of $GABA_{B1b}$-bRIL/$_{B2}$-cys particles.

A subset of 77,577 particles from various views was selected for 3D classification. The 3D reconstruction of the mutant receptor had a lower nominal resolution of 13.9 Å. The overall arrangement was the same as in the previous model, having the ECD fly-trap sitting above the TMD density (Figure 5.12a). Although the density of the fly-trap was less defined in this model, the active state crystal structure (PDB ID: 4MS4) could still be docked into the density as a rigid body. Shape of the TMD density remained irregular and attempts to fit the $mGluR_5$ TMD structure were not successful. 3D models of the apo and cys-mutant samples were superimposed by rigid-body movement (Figure 5.12b). The cys-mutant sample had extra density at the core of fly-trap, which was consistent with the expectation of the mutant receptor forming a the closed confirmation. More density was reconstructed in the TMD in cys-mutant model, where the increased density expanded across the membrane direction. The emergence of density could come from enhanced structural rigidity in the region.

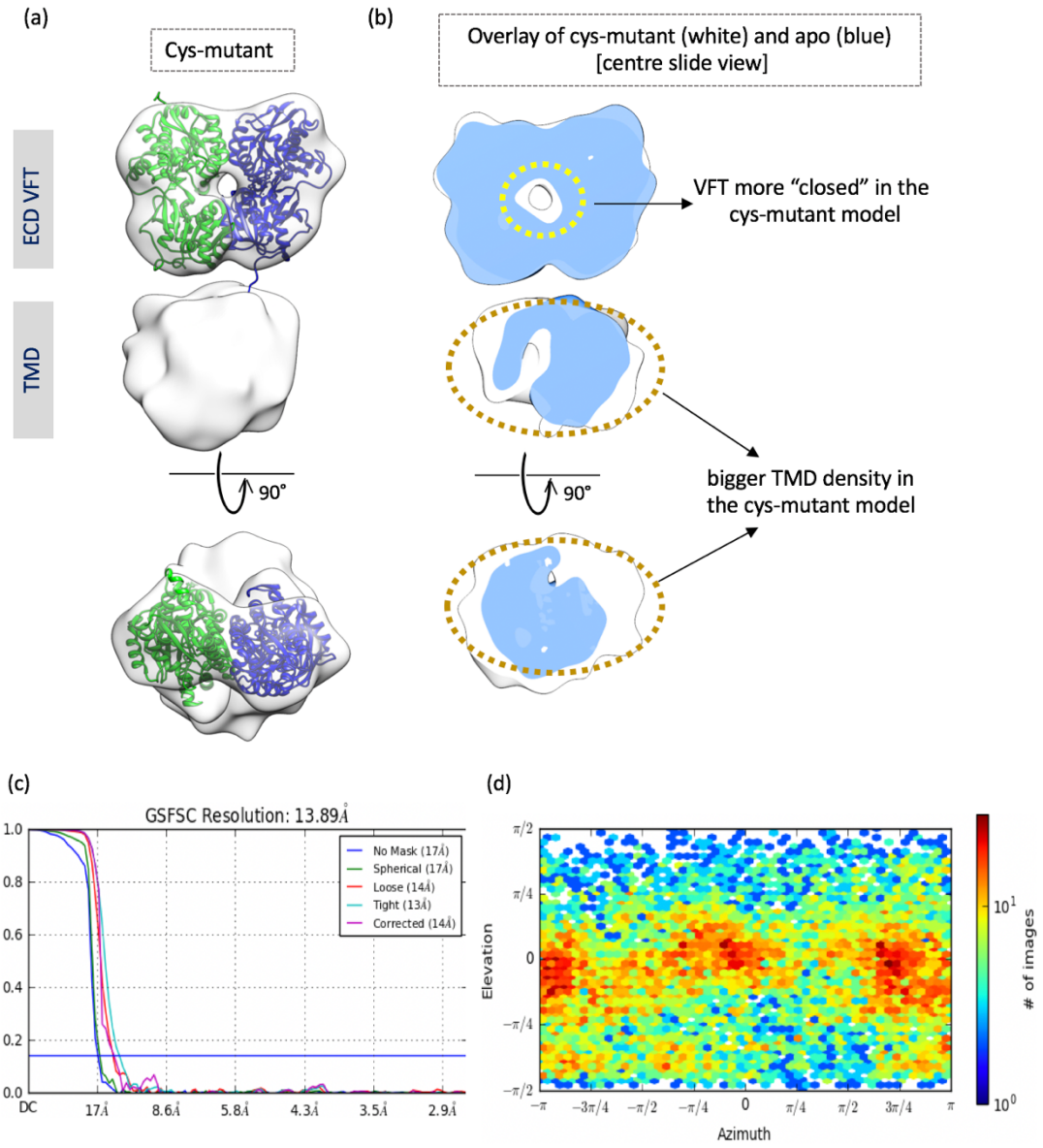


Figure 5.12 3D reconstruction of $GABA_{B1b}$-$bRIL_{/B2}$-cys particles.

(a) Reconstruction model shown as front and top views. The ECD VFT crystal structure in closed conformation with baclofen bound (PDB ID 4MS4) was fitted as a rigid-body in UCSF Chimera. Density was displayed at 0.28 threshold. (b) Superimposition of 3D models from $GABA_{B1b}$-$Bril_{/B2}$ (blue) and $GABA_{B1b}$-$bRIL_{/B2}$-cys (white). Cross-section view of the central slide was shown. The two models were superimposed by rigid-body fitting in UCSF Chimera. Both models are displayed at a threshold of 0.28. (c) Fourier Shell Correlation curve analysis. (d) Azimuth map showing particles distribution from different orientations.

5.3 Discussion

5.3.1 Amphipols reconstitution for cryo-EM studies

Based on results in Section 5.2.1, amphipols was capable to efficiently replace detergent without causing protein aggregation. A new protein peak, at a later elution volume relative to the original peak, appeared in the amphipol-reconstituted sample SEC profile. EM results confirmed that particles from the new peak were dimeric $GABA_BR$. The apparent smaller size might be due to tighter interactions of amphipols with the receptor. $GABA_BR$ was found to be more stable when stabilised in amphipols, potentially as a result of amphipol's more stable association with the receptor transmembrane region [201]. In contrast, detergent micelle are dynamic where individual detergent molecules are on constant exchange in and out of the micelle. Amphipols thus better protect the hydrophobic transmembrane domains from the aqueous environment to prevent protein precipitation. Replacing detergent with amphipols also altered the behaviour of $GABA_BR$ in solution for EM sample preparation. Amphipol-solubilised $GABA_BR$ particles were much more likely to stay in free-standing ice and had a uniform distribution in the grid hole. Amphipols, being a much weaker surfactant than detergent, might have altered solution properties such like surface tension. In addition, there was no free amphipol molecule present in solution, unlike in detergent-solubilised sample where excess detergent micelles are always around.

Issues observed that might be associated with amphipols reconstitution included particle preferential orientations. Protein projection from a few angles were more predominant than others, indicating that amphipols-exchanged $GABA_BR$

preferentially adapt those orientations in ice. Preferred orientation of amphipols-reconstituted particles had been reported for other membrane proteins, such as the RyR1/A8-35 complex [202]. The exact mechanism of why $GABA_BR$ exhibited such character was unknown. Trace amount of surfactant, such as fluorinated detergents, could be tested in future experiments for their effects in reducing protein preferential orientations [203].

Some 2D averages were found to have missing densities at the ECD periphery region, which could be due to protein degradation at the air-water-interface. One possible way to prevent this problem could be to use a physical support, such as graphene oxide coating, to completely remove contact with air at one side [204]. Removal of protein contacts with the air-water-interface might also allow the receptor to adapt different views [205].

5.3.2 Model reconstruction of $GABA_{B1b}$-bRIL/$_{B2}$ particles

3D model of amphipols reconstituted receptors had the two subunits arranged side-by-side facing opposite directions. Although it was difficult to definitively assign the density to each specific subunit at resolution of the current reconstructions, given the sample had been affinity-purified for both subunits the dimeric receptor model reconstructed here should represent the heterodimer.

Clear density for the ECD and TMD domain were observed in the model, where the ECD was linked to the TMD by a flexible peptide linker. The ECD fly-trap was expected to be more rigid compared to the TMD, based the fact that multiple crystals could be obtained for the ECD VFT alone [55] but not for the transmembrane domain

alone. As expected, the ECD showed better alignment during classification and had better resolved density. Identity of the ECD domain density was confirmed by successful rigid-body docking of the VFT crystal structures.

TMD region alignment was limited and did not show structural features in the density. This was not surprising as the TMD of GPCRs was known to be highly dynamic. The apparent size of reconstructed TMD density was slightly smaller than the theoretical $GABA_BR$ TMD. Estimated dimension of the reconstructed TMD density was approximately 5nm deep, 6nm long and 5.5 nm wide. While the theoretical size of $GABA_BRs$ 14-helices was estimated to be 4.5nm deep, 7nm long and 3nm wide based on the $mGluR_5$ TMD structure [206]. The small TMD density could be a result of weak alignment from the a highly flexible sample.

Clear density for the intracellular region was not visible in current reconstructions. The coiled-coils were expected to be difficult to visualise given their small size. However, the relatively larger bRIL was not visible either. Failing to align the bRIL was an indication of high flexibility around the area where it was inserted. One possibility was that $GABA_{B1}$ intracellular loop 3 was highly mobile, causing rigid-body movements of bRIL relative to the receptor. The intrinsic dynamic nature of transmembrane helices discussed above would further exacerbate poor alignment of bRIL. Based on the data, fusion with bRIL was not likely to help future sample alignment and the insertion was later removed for sample preparations in chapter 6.

Linker densities between domains were difficult to visualise, indicating high flexibility in the linker region. As a consequence, accurate alignment of the whole

receptor was prevented by inter-domain movements. Localised reconstruction of the individual domains was challenging, because the signal-to-noise ratio in data of one single domain was not sufficient for accurate alignment. The mass of one single $GABA_BR$ domain is ~100 kDa each, which was too small for gaining enough contrast using the current data collection setup.

5.3.3 Model reconstruction of $GABA_{B1b}$-bRIL/$_{B2}$-cys particles

Receptors of the first dataset were likely to exhibit an ensemble of conformations across different states, as sample was the wildtype construct and was not supplemented with any ligands. The cys-mutant receptor was thus created with the aim to lock protein in a uniform conformation. As expected, mutant receptors exhibited increased drug binding affinity. Mutant receptors expression level was reduced to 70%, which was in agreement with the extend of reduced maximal activity measured in cellular assays [55]. It is not uncommon to obtain reduced yields when expressing constitutively active receptors, whose activity might adversely affect on cells health.

The reconstruction model of the mutant receptor showed similar overall architecture as the apo, with the butterfly-shaped ECD VFT sitting above a TMD. Introduction of cysteine bridge mutation sites appeared to have to stabilise the ECD fly-trap in the closed confirmation. Receptor TMD rigidity seemed to be improved given that more density was reconstructed in this model. Assuming that the introduction of cystine-mutation and addition of agonist had stabilised the receptor ECD in an activated state, local conformational changes between states were not visible when comparing the current models due to the limited resolution.

Although the shape of the model was in agreement with a closed conformation receptor, the overall alignment was poorer for the mutant sample than the *apo* sample. 2D alignment did not show ECD secondary structural features and the 3D model had a lower nominal resolution. The mutant data were collected with lower electron dose and therefore has a lower signal-to-noise ratio, which was likely the reason for the reduced alignment accuracy. It was also possible that constitutively active receptors have reduced stability, since receptor activated state had been reported to be less stable in some GPCRs such as in the $A_{2A}R$ [207].

5.4 Conclusion

By our knowledge, models presented here were the first reconstructions for the integral human $GABA_BR$. Analysis of data collected so far suggested that $GABA_BR$ in amphipols had an ordered ECD VFT and a flexible TMD. Despite the attempt to stabilise receptor conformation by site-directed mutagenesis and addition of ligands, resolution of model reconstructed did not show significant improvement. At this stage the best approach of finding a better sample was still an empirical trial-and-error approach. Given that there was no dramatic improvements of amphipols-reconstituted particles, sample preparation was shifted to reconstituting the receptor into lipid nanodiscs, another popular reconstitution system for membrane proteins. The next chapter described works involving $GABA_BR$ reconstitution into nanodiscs.

Chapter 6. CRYO-EM DATA ANALYSIS OF $GABA_BR$ RECONSTITUTED IN LIPID NANODISCS

6.1 Introduction

This chapter continues the characterisation of $GABA_BR$ reconstitution system candidates for cryo-EM studies. Possibilities of $GABA_BR$ reconstitution into three types of nanodiscs, namely the saposin-lipoprotein salipro particles and two membrane scaffold proteins *(*MSP) nanodiscs MSP1D1 and MSP2N2, were assessed using SEC in chapter 4. This chapter presents the EM sample preparations and data analysis of receptors reconstituted in the three reagents.

A number of recent structural studies of membrane proteins have investigated the use of lipid nanodiscs as an alternative to detergents. Lipid nanodiscs are soluble disc-shaped structures that mimic the phospholipid bilayer. Given that membrane composition is critical for the correct functioning of membrane proteins [208], the use of lipid nanodiscs is expected to better preserve the protein in a more stable state [209]. Compared to amphipols, receptors reconstituted in nanodiscs showed improved local resolution at the protein-lipid interface [210]. As a result, an increasing number of high-resolution structures have been solved in lipid nanodiscs, such as the ryanodine receptor [177], TRVP1 channel [210] and PKD2 channel [211].

Based on analysis of various receptor constructs in chapter 5, the bRIL insertion did not appear to assist particle alignment and was therefore removed to allow $GABA_BR$ to restore to its natural conformation. The protein construct used in this chapter was therefore the wild-type sequences corresponding to the $GABA_{B1b/B2}$

heterodimer, including all known structural elements from the extracellular VFT domain to the end of the intracellular coiled-coils.

Furthermore, cryo-EM data acquired in this chapter was on a Titan Krios microscope equipped with a Volta Phase Plate (VPP). As discussed in chapter 5, accurate alignment of $GABA_BR$ particles had been challenging due to their small sizes and consequently the low signal-to-noise ratio in the recorded micrographs. Defocus phase contrast is typically imposed in order to improve the image contrast. However, collection at defocus will lead to loss of data at certain frequency and often compromise reconstruction resolution. The VPP is a device designed to efficiently enhance phase contrast in TEM [212, 213]. VPP creates a phase shift between the scattered and un-scattered waves at the diffraction plane, thereby allowing close-to-focus data acquisition while maintaining image contrast [213]. High-resolution structures of several GPCRs had been reported since the introduction of VPP [11, 214]. Data presented in this chapter utilised the VPP for enhancement in image contrast.

6.2 Results

6.2.1 Reconstitution into salipro nanodiscs

Saposin A lipoprotein particles, also known as salipro, are a new class of lipid nano-particles for membrane protein reconstitution [177]. Their versatility lie in the ability to adapting different disc sizes depending on the size of the membrane protein incorporated [215]. Detergent-solubilised $GABA_BR$ was supplemented with bovine lipid extract and reconstituted into saposin A, followed by detergent removal using Bio-beads absorption. Saposin A reconstituted particles were analysed by SEC (Figure 6.1a) where three peaks were obtained. The earliest peak at 0.9 ml represented large aggregates eluted in the void volume. The broad peak at 1.4 ml elution volume with a shoulder at 1.3 ml was identified to contain $GABA_BR$-salipro particles by SDS-PAGE (Figure 6.1b). $GABA_BR$-salipro particles peak had an earlier elution volume than detergent-solubilised or amphipols reconstituted particles which elute at approx. 1.5 ml. The significant shoulder in the peak was suggestive of a heterogenous sample. SDS-PAGE analysis showed a major $GABA_BR$ monomer band across all fractions with some higher-order oligomer bands present in earlier fractions. The highest peak at 1.8 ml corresponded to the empty salipro particles as confirmed by SDS-PAGE (Figure 6.1b). The fresh fraction from the $GABA_BR$-salipro particle peak with no sign of oligomer band was diluted to 0.007 mg/ml and analysed by negative stain EM. Visual inspection of the micrograph (Figure 6.1c) showed a homogenous population of particles corresponding to $GABA_BR$ size.

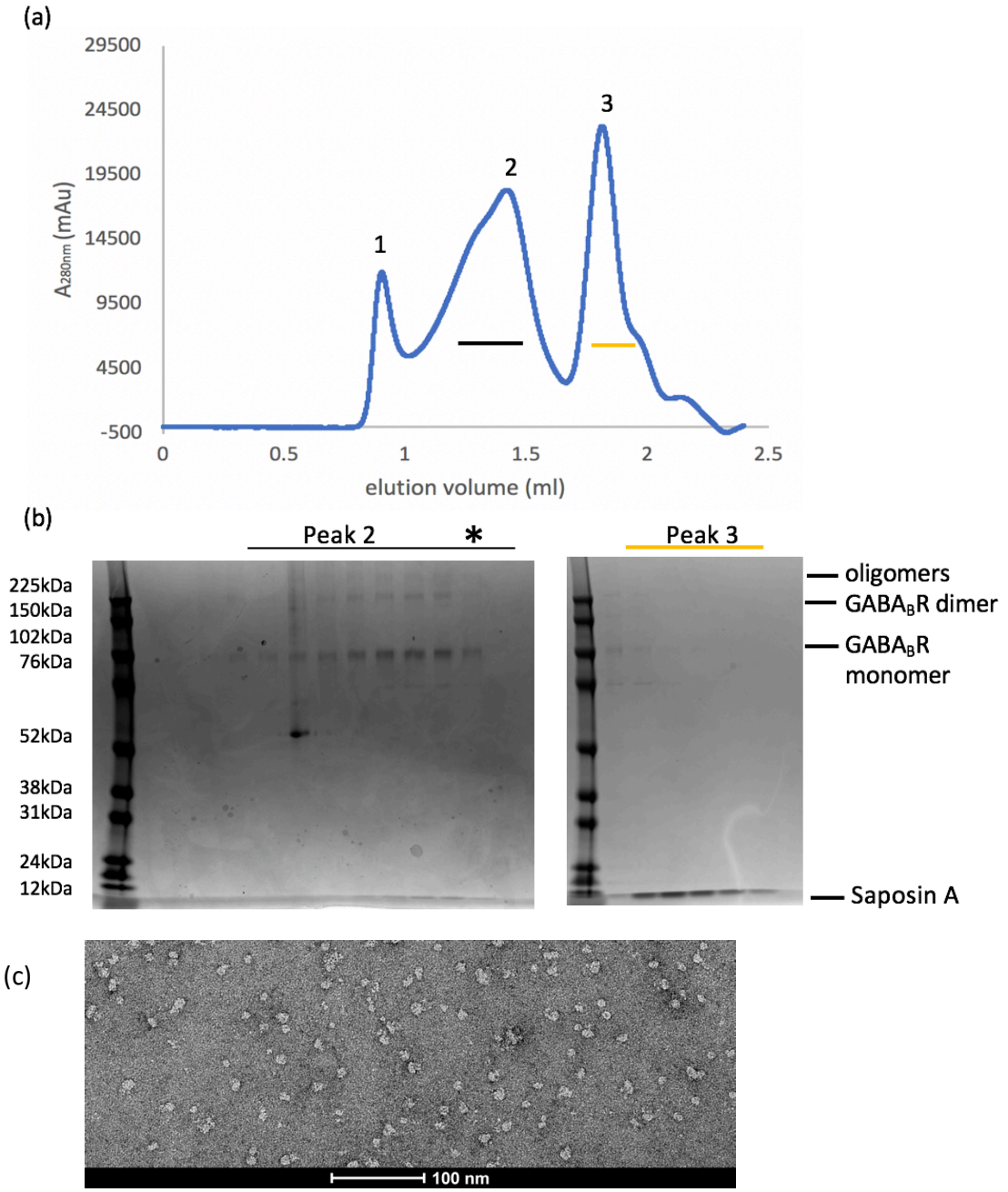


Figure 6.1 Analysis of $GABA_BR$-salipro particle reconstitution.

(a) SEC profile of the wild-type $GABA_BR$ reconstituted into salipro particles. Peak 1 represented aggregates eluted at void volume (0.9 ml). Peak 2 and peak 3 contained $GABA_BR$ reconstituted in Salipro and empty salipro particles. (b) SDS-PAGE analysis of fractions from peak 2 and 3. Fraction marked by * was taken to negative-stain EM sample preparation shown in d. (d) Micrograph of $GABA_BR$-salipro particles negative stain grid. Grid was prepared using sample at 0.007mg/ml concentration and imaged on a F20 microscope.

This fraction was further concentrated to 1.5 mg/ml for cryo-EM grid preparation. Micrograph of the $GABA_BR$-salipro sample in free-standing ice (Figure 6.2a) showed a mixed population of particles. There were some single particles and many clusters of aggregates. It was unclear whether particle aggregation occurred during the concentrating step or grid freezing step. 2D classification averages (Figure 6.2b) of particles picked from these micrographs showed poor alignment and high variability in disc dimensions. Based on the 2D average results, it was suspected that even single receptor reconstituted salipro particles had highly variable sizes. Salipro reconstitutions were not used further due to the limited quality of classification results, suggesting that such particles were not likely to yield high-resolution reconstructions.

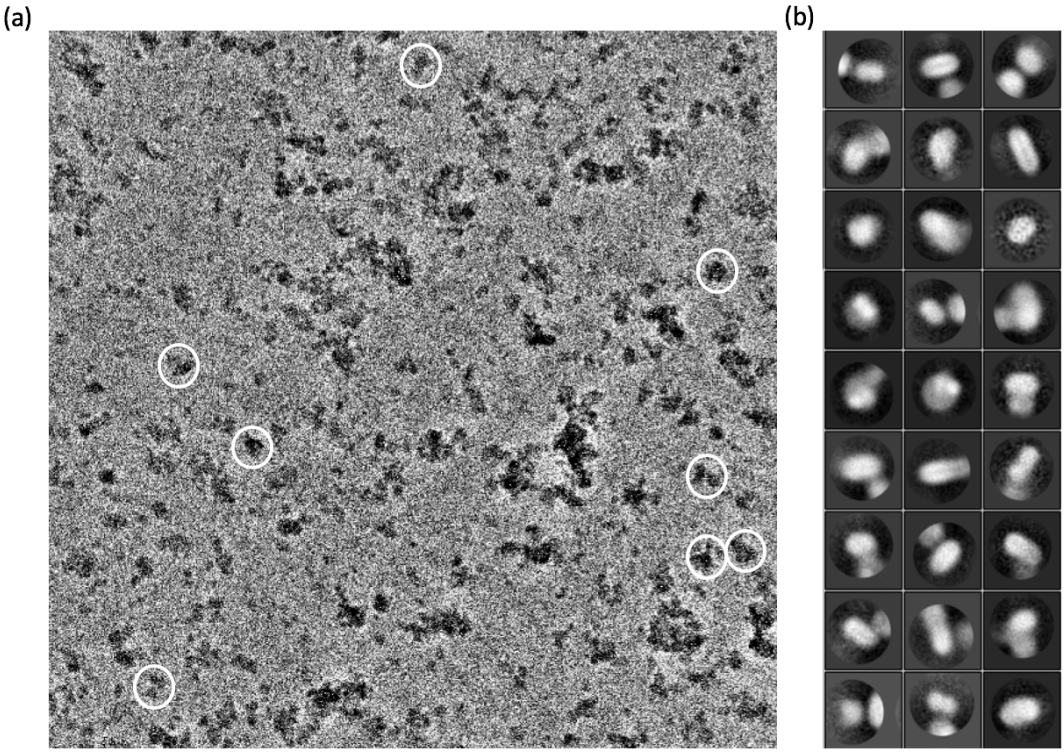


Figure 6.2 Analysis of $GABA_BR$ Salipro particle Cryo-EM sample.

(a) Representative micrograph of $GABA_BR$-salipro particles in free-standing ice. Grid was prepared using sample at 1.5 mg/ml concentration and imaged on Titan Krios microscope with VPP. Potential single particles corresponding to the size of $GABA_BR$ were circled in white. (b) 2D classification results using potential single particles extracted as exemplified in a .

6.2.2 Reconstitution into MSP1D1 nanodiscs

Next, nanodiscs formation using membrane scaffolding proteins (MSP) were assessed. MSPs are engineered proteins of short amphipathic helices, derived from the apolipoprotein A-1 [216]. The size of MSP nanodiscs are pre-defined and controlled by the length of the MSP used. A typical MSP nanodisc is assembled by two MSPs arranged in a parallel or antiparallel manner, forming a belt structure to protect the hydrophobic core.

MSP1D1 was tested in this section, which has a theoretical diameter of ~10 nm [217] and was expected to closely incorporate the $GABA_BR$ dimer. Detergent solubilised $GABA_BR$ was supplemented with a mixture of synthetic phospholipid 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC) and bovine brain lipid extract before reconstitution into MSP1D1 nanodiscs. POPC is a common lipid used for reconstituting membrane proteins [203, 218] and has a phase transition temperature of -2°C [2] which was very close to the temperature $GABA_BR$ reconstitution experiment was performed. Bovine natural lipid extract was included since $GABA_BR$ is known to be associated with the lipid rafts [164]. MSP1D1 empty nanodiscs were found to stick non-specifically onto reconstituted $GABA_BR$ nanodiscs. Multiple wash steps and a gel filtration step were required to remove the empty discs. SEC profile (Figure 6.3a) of MSP1D1 reconstituted particles yielded two protein peaks, at 1.53 ml and 1.8 ml respectively. SDS-PAGE analysis (Figure 6.3b) of peak fractions suggested that the peak at 1.53 ml contained $GABA_BR$-MSP1D1 nanodiscs. Elution volume of MSP1D1 reconstituted $GABA_BR$ particles were slightly later than its usual elution volume (which was later shown in section 6.2.3 to be the result of receptor dissociation at the domain

interface). The peak was symmetrical with a slight shoulder at earlier elution volume, suggesting a relative homogenous population of particles. Peak eluted later at 1.8 ml contained empty nanodiscs as confirmed by SDS-PAGE (Figure 6.3c).

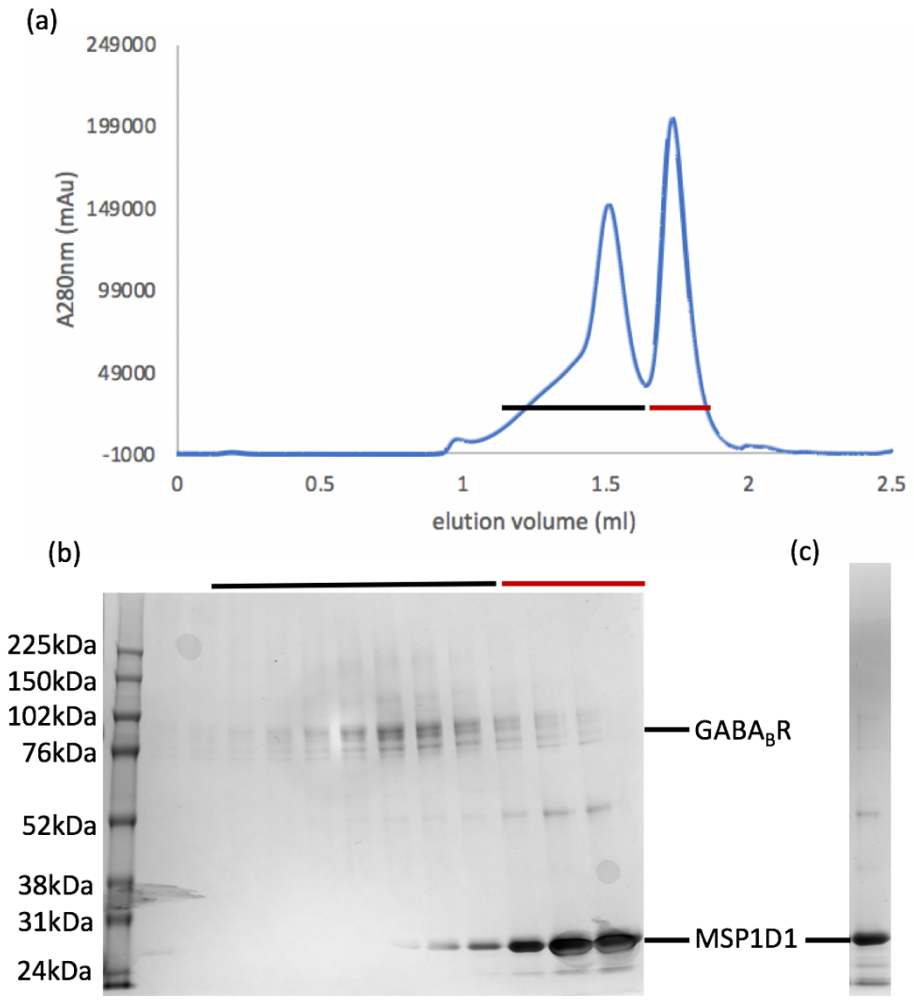


Figure 6.3 GABA$_B$R reconstitution into MSP1D1 nanodiscs.

(a) SEC profile of MSP1D1 reconstitution products. Peak at 1.53 ml and 1.8 ml contained GABA$_B$R-MSP1D1 nanodiscs and empty MSP1D1 nanodiscs respectively. (b) SDS-PAGE analysis of fraction from peak at 1.53ml. Band at approximately 90kDa apparent MW represented GABA$_B$R and the band at approximately 26 kDa represented MSP1D1. (c) SDS-PAGE analysis of fraction from peak at 1.8 ml. Band corresponding to MSP1D1 were observed.

The peak fraction did not show any bands of oligomers on SDS-PAGE and was thus used to prepare Cryo-EM grid at a concentration of 0.25 mg/ml. $GABA_BR$-MSP1D1 complex gave monodispersed particles on Cryo-EM grid (Figure 6.4a). Particles resembling the $GABA_BR$ size were automatically picked for 2D classifications. 2D averages (Figure 6.4b) showed homogenous densities consisted of two equal-sized domains adjacent to each other, which was consistent with the overall topology of the $GABA_BR$. The two domains could potentially represent the ECD and TMD respectively.

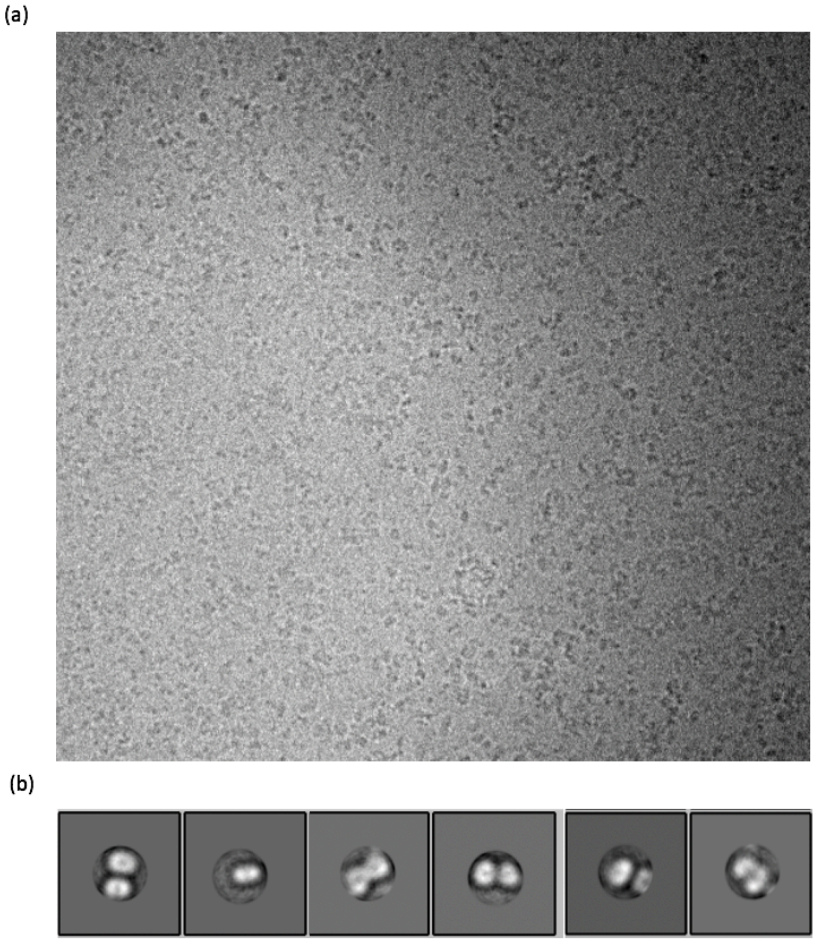


Figure 6.4 Analysis of MSP1D1 reconstituted $GABA_BR$ particles.

(a) Representative micrograph of $GABA_BR$-MSP1D1 nanodiscs at 0.25 mg/ml. Image was collected on a F20 microscope. (b) 2D class averages of automatically picked particles.

6.2.3 Data processing and analysis of the MSP1D1 nanodiscs particles

Based on 2D classification results in section 6.2.2, $GABA_BRs$-MSP1D1 particles appeared to be suitable for large scale data collection and 313 movies were collected on the Titan Krios Microscope with VPP (Table 6.1). A set of templates (Figure 6.5a) from different views was generated using manual picked particles. Domain arrangement in the low-resolution templates agreed with the receptor general topology. 75,012 particles were picked in an automated manner using the above templates. Figure 6.5b shows a micrograph with representative picking results.

Table 6.1 Data collection statistics for $GABA_BR$_WT_MSP1D1

Microscope	Titan Krios
Voltage (kV)	300
Imaging mode	Counting
Detector	Falcon III
Pixel size (Å)	1.07
Total Dose ($e^-/Å^2$)	30
Exposure time (s)	60
Dose rate 2 (e^-/pix/s)	0.56
Number of frames	75
Frame length (s)	0.8
Focusing range (µm)	-0.5 to -0.7 with Volta phase plate

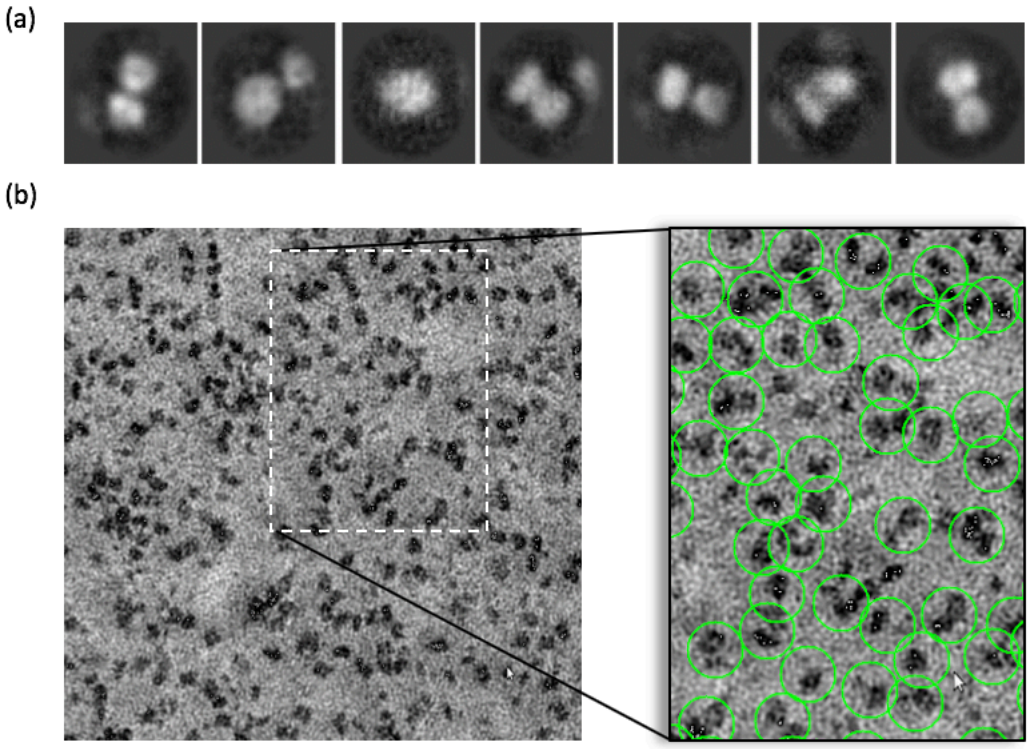


Figure 6.5 Particle picking for MSP1D1 reconstituted $GABA_BR$.

(a) Representative 2D averages from manually picked particles. Around. 500 particles were picked manually using boxsize of 256 pixel (24 nm). A circular mask of 18 nm diameter was applied to the particles during 2D classification. (b) Example of automated picked particles using the templates. Green circles (18nm in diameter) marked the coordinates of the automated picked particles. Particles were extracted as a box of 24nm.

2D classification of the whole dataset revealed a high degree of heterogeneity in the auto-picked particles. Only a small subset of 2D classes (accounting for 9% of total particles) showed intact particles with ECD density adjacent to the TMD density (Figure 6.6a). Most 2D classes had two or more density masses, where the relative distances and orientations between each other was highly variable (Figure 6.6b). A number of different picking parameters and 2D classification strategies were tested and gave similar looking 2D averages, confirming the observations were not due to inaccurate particle picking or alignment. $GABA_BR$ particles were suspected to had broken apart at domain interfaces, as the distance between densities were not consistent across the class averages. Although, there was no evidence of such cleavage

during sample preparation and characterization (e.g. SDS-PAGE and SEC profile), protein dissociation at domain interface in MSP1D1 particles was confirmed later after comparison with MSP2N2 reconstituted particles in section 6.2.5. Visual inspection of the raw particles on the micrograph from the two data set showed significant differences. Intact particles in MSP2N2 had the different domains closely connected to each other. In comparison, distance between domains were much wider and non-uniform in the MSP1D1 reconstituted data. One possibility is that the dissociation happened during cryo-EM sample freezing.

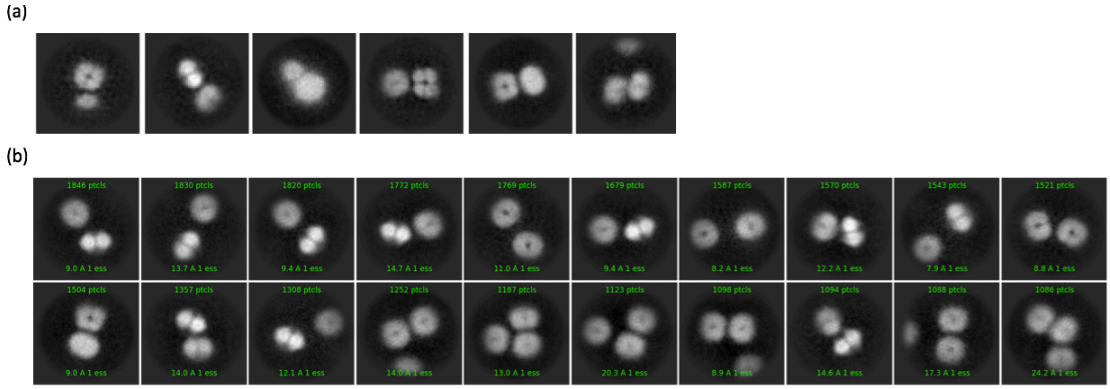


Figure 6.6 2D averages of MSP1D1 reconstituted particles.

(a) Selected 2D class averages showing intact particles (b) Representative 2D classification results from automated picked particles showing high sample heterogeneity.

Given that image contrast was greatly enhanced by the VPP, reconstruction of single domains was performed again to investigate the structures of isolated domains. Particle picking was performed using a 13 nm box (approx. half of the intact particle). 2D averages from the re-extracted “half” particles showed many well-aligned ECD classes, and some less well-defined classes possibly representing the TMD density

(Figure 6.7). Potential classes corresponding to each domain were saved separately and independently processed.

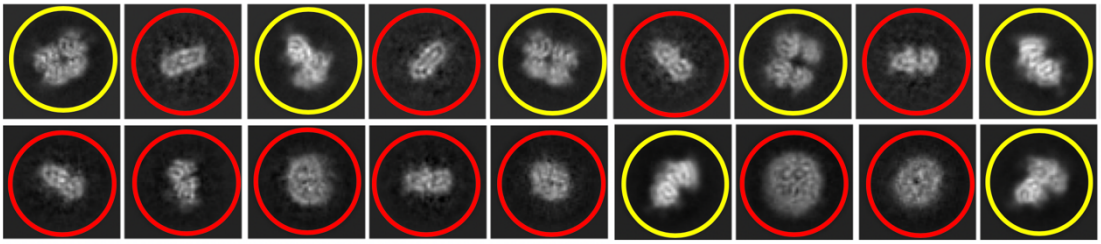


Figure 6.7 Representative 2D class averages of half-sized particles. Classes corresponding to the ECD and TMD are marked in yellow and red respectively.

Further 2D classifications of the potential TMD particles did not improve particle alignment. Class averages remained as poorly defined elliptical or circular shaped densities. No structural features corresponding to the $GABA_BR$ TMD could be observed within the density. In contrast, 2D averages of ECD particles were well aligned and had clear secondary structural features (Figure 6.8). Particles were further cleaned by 3D classifications (Figure 6.8) and 34,768 particles were selected for the final reconstruction. The model at a nominal resolution of 3.6 Å showed clear densities for α-helices although the β-sheet areas are less well resolved (Figure 6.8).

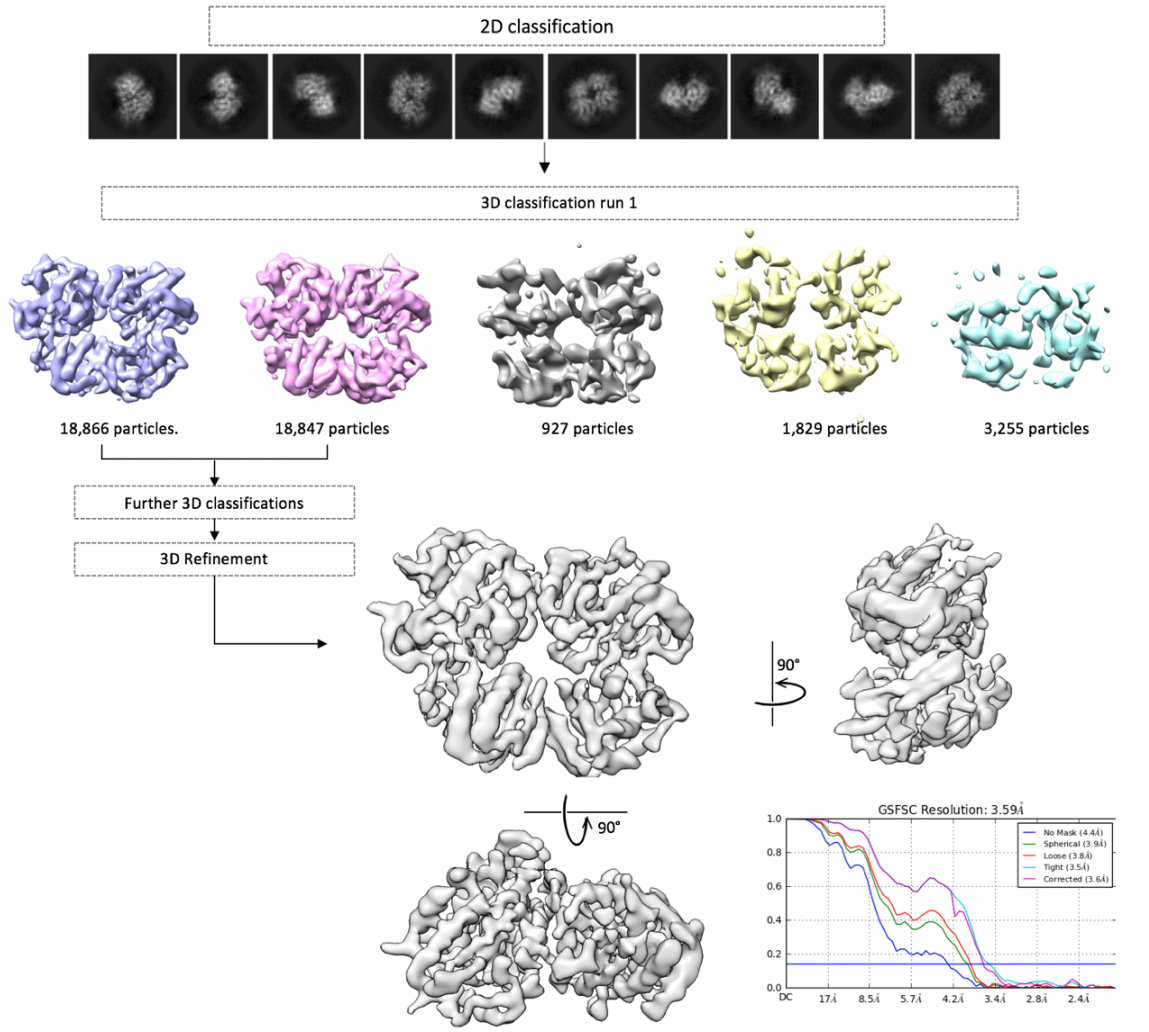


Figure 6.8 Data processing of $GABA_BR$ ECD particles.

60,551 particles were selected from classes showing features corresponding to the ECD VFT for 2D classification. 43,714 particles were selected from 2D classes showing intact VFT. Particles were further cleaned through several rounds of 3D classifications. 34,768 particles were used in the final reconstruction. The model was reconstructed to a nominal resolution of 3.6 Å based on FSC calculation.

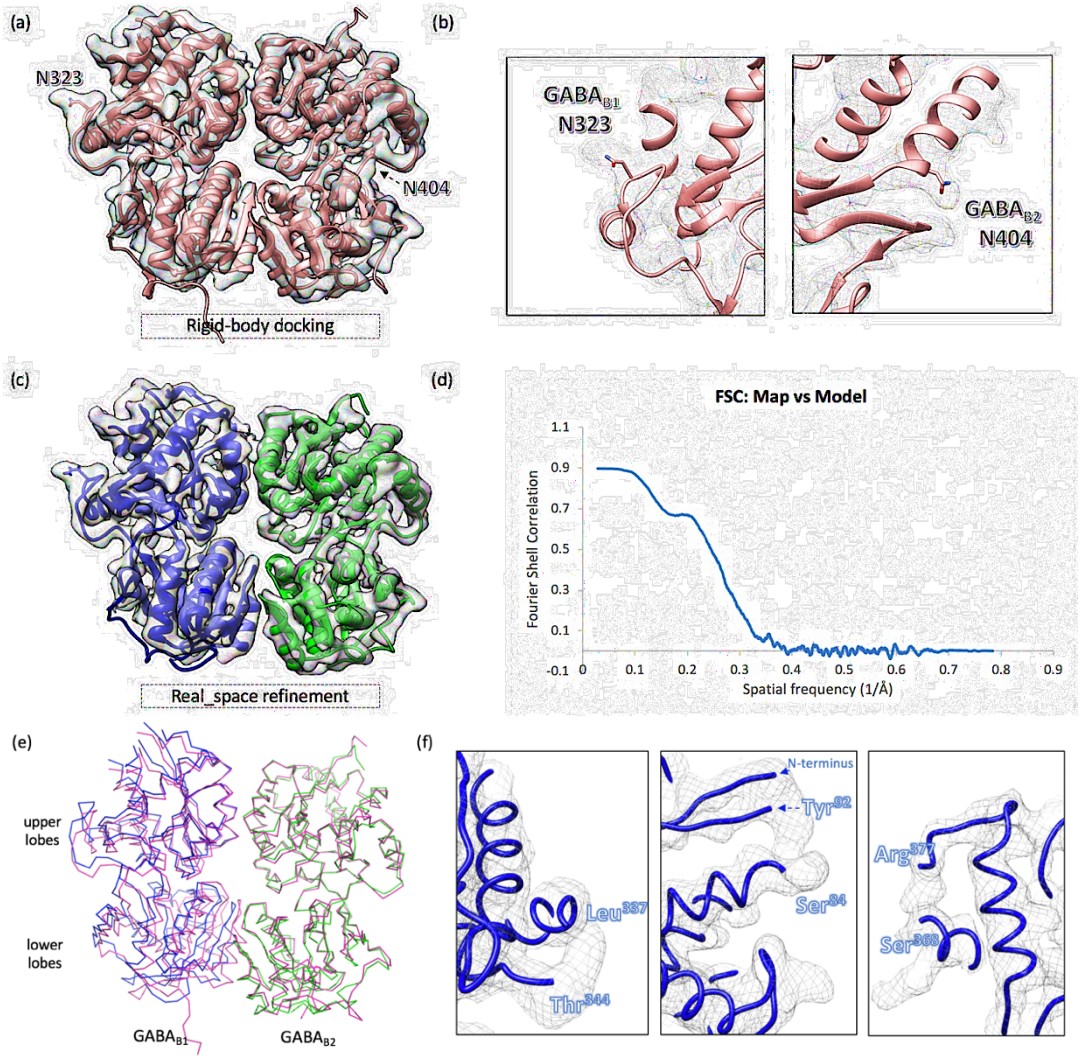


Figure 6.9 Heteromeric ECD VFT EM model.

(a) Rigid-body docking of GABA$_B$R ECD VFT crystal structure (PDB ID 4MS3) into the reconstructed density map. (b) Fitting at GABA$_{B1}$ Asn^{323} and GABA$_{B2}$ Asn^{404}. (c) EM model with refined model. GABA$_{B1}$ and GABA$_{B1}$ subunits are coloured blue and green respectively. (d) Map-to-model fit plot of the refined model. (e) Superimposition of crystal structure (pink) and real space refined model (blue and green). Overlay of models was calculated in pyMOL with a root mean square deviation of 0.945. (f) Locations of missing residues in the crystal structure.

The ECD VFT crystal structure (PDB ID: 4MS3) was first docked as a rigid-body (Figure 6.9a). The crystal structure fitted well into contours of the EM reconstruction, where most secondary structural elements fit within the density. GABA$_{B1}$ and GABA$_{B2}$ subunit could be readily differentiated based on density for subunit specific residues: GABA$_{B1}$ Asn^{323} and GABA$_{B2}$ Asn^{404} (Figure 6.9b). Real space refinement [126] was

performed next using the docked crystal structure as a starting model. The model was well-refined where almost all residues are fitted into the density except some peripheral loops (Figure 6.9c). Compared to the crystal structure, $GABA_{B1}$ of the fitted model adopted a more "open" confirmation where its lower lobe is further apart from $GABA_{B2}$ (Figure 6.9e). In addition, continuous EM densities were observed for the 3 missing fragments in the crystal structure (Figure 6.9f). Accurate model building in these areas was precluded due to lack of clear side chain features.

6.2.4 Reconstitution into MSP2N2 nanodiscs

$GABA_BR$ was next reconstituted into MSP2N2 nanodiscs following the same procedure as described in section 6.2.2. MSP2N2, which is made by fusing two smaller MSPs [219], was expected to form a bigger disc with a theoretical diameter of 15-16 nm [217]. SEC profile (Figure 6.10a) of MSP2N2 reconstituted particles had a $GABA_BR$ particles peak at 1.5 ml and an empty nanodisc peak at 1.74 ml. The $GABA_BR$ particles peak was symmetrical, with a slight shoulder at an earlier elution volume. Fractions in this peak were analysed by SDS-PAGE (Figure 6.10b) and showed very faint band at MW corresponding to putative $GABA_BR$ oligomers. Besides $GABA_BR$ and MSP2N2 bands, small amount of contaminant bands was observed at lower MW which showed co-elution pattern with MSP2N2. Although elution volume of MSP2N2 particles was the same as amphipols-reconstituted sample and earlier than MSP1D1 reconstituted particles, whether the $GABA_BR$ domains had dissociated or not could not be confirmed based on SDS-PAGE analysis results.

Peak fraction at a concentration of 0.2 mg/ml were directly applied to cryo-EM grids and plunge frozen. Similar to the MSP1D1 reconstituted sample, MSP2N2 reconstituted particles gave monodispersed distribution in free-standing ice (Figure 6.11). 2D classification was run on the particles (Figure 6.10c) and the result was compared to MSP1D1 reconstituted particles in section 6.2.3. 2D averages from MSP2N2 nanodiscs had similar dimensions as the MSP1D1 reconstituted particles and also showed two domains of similar mass adjacent to each other. Again, It was unclear based on the results here whether the ECD and TMD remained connected together or not.

In order to obtain high resolution information, data were collected on a Titan Krios microscope with VPP using the same parameters as in Table 6.1. Data processing and analysis are described in section 6.2.5 below.

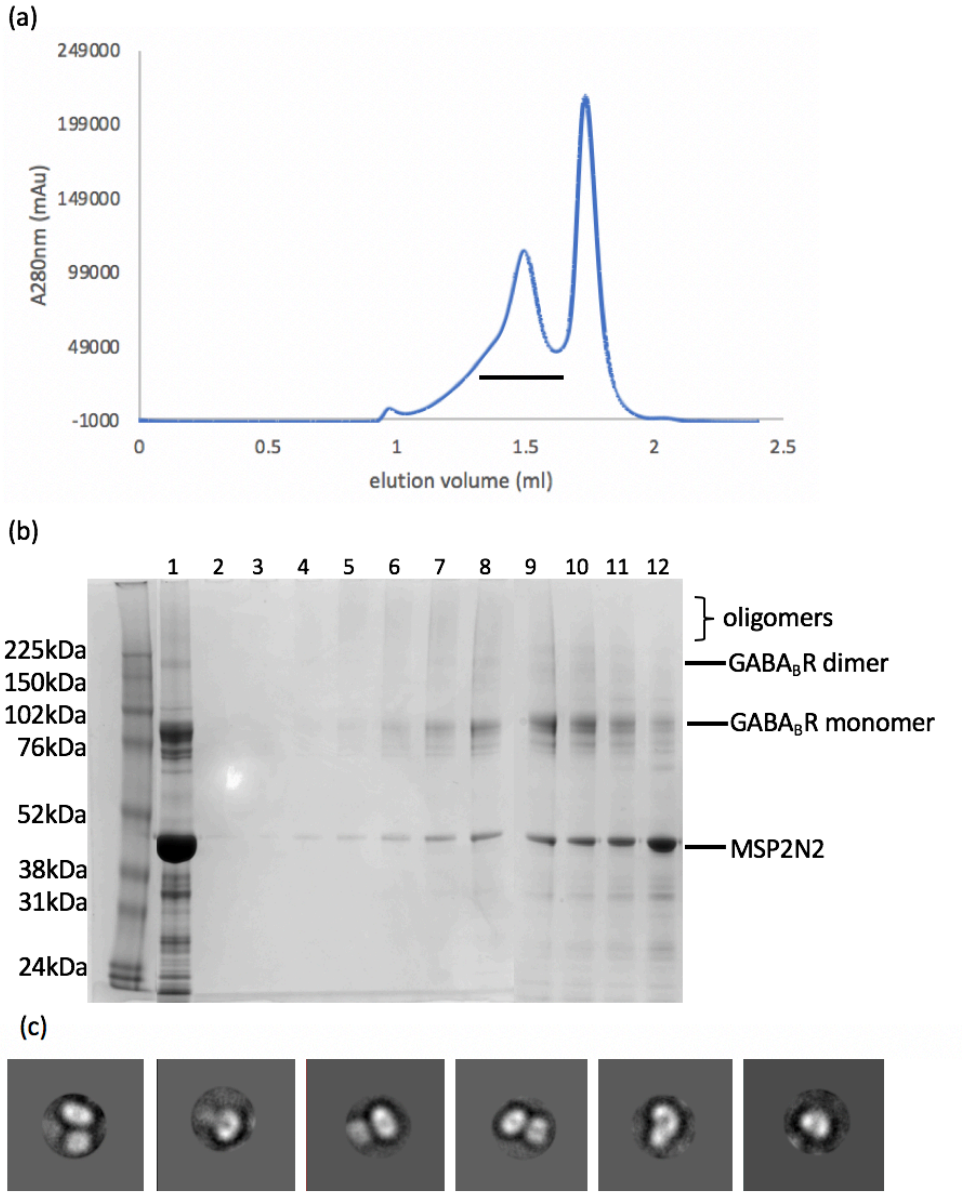


Figure 6.10 Analysis of MSP2N2 reconstituted particles.

(a) SEC profile of MSP2N2 reconstitution products. Peak at 1.5 ml and 1.74 ml corresponded to GABA$_B$R nanodisc particles and empty nanodisc particles respectively. (b) SDS-PAGE analysis of fractions from the GABA$_B$R particles peak (underlined in a). Lane 1 contained concentrated sample before injection onto the column, lane 2-12 contained fractions of peak. (c) Representative 2D class averages from MSP2N2 reconstituted particles. Data was collected on F20 microscope and processed in RELION.

6.2.5 Data processing and analysis of the GABA$_B$R-WT-2N2 particles

GABA$_B$R-MSP2N2 nanodisc particles could be readily identified in the raw micrograph with contrast-enhanced data collected using VPP (Figure 6.11a). Particles

appeared intact as evident by “front” view projections showing all three domains of the receptor (Figure 6.11b). Distances between domains were uniform and closer than in the MSP1D1 reconstituted sample. Furthermore, particle projections from wide range of orientations were observed (Figure 6.11c). On top of the three common views seen previously (discussed in section 5.2.2), more particles were found to adapt a tilted view.

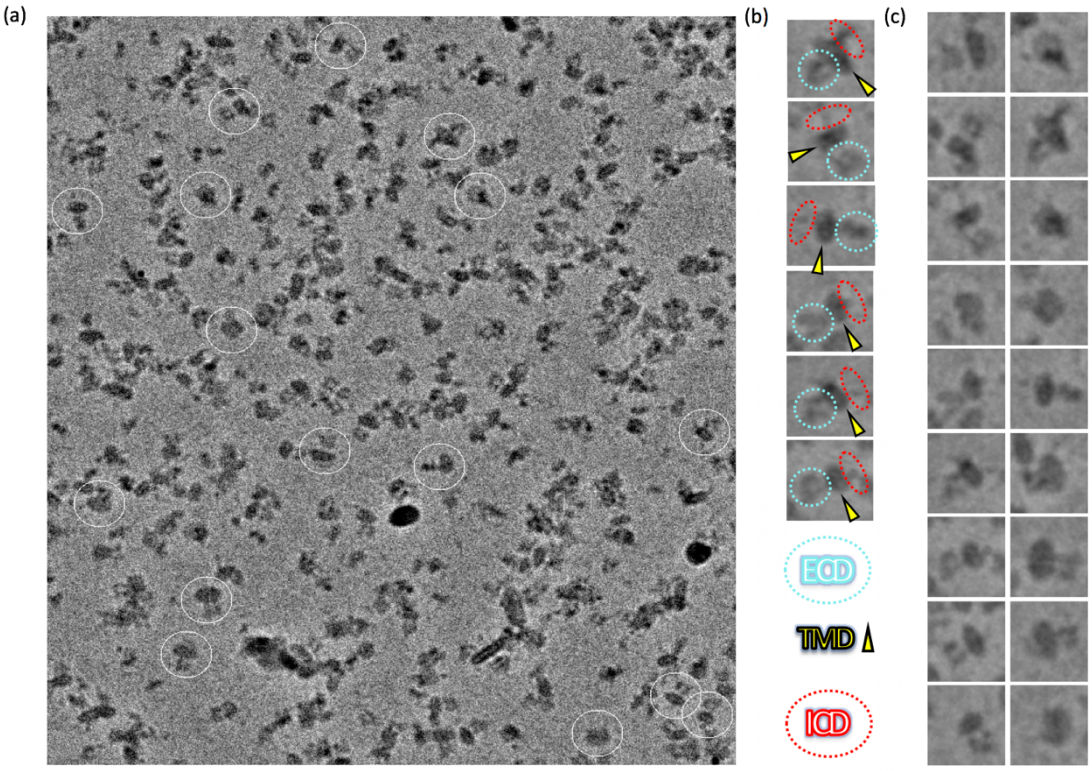


Figure 6.11 Micrograph of MSP2N2-reconstituted particles.

(a) Particle distribution in free-standing ice. Grid was prepared using sample at 0.2 mg/ml concentration. (b) Representative “front” view of intact MSP2N2-reconstituted particles showing all 3 domains. Particles showing the ECD VFT domain (blue dashed line) connected to a circular lipid bilayer (yellow triangle) and a small intracellular tail (red dashed line). (c) Representative particles from different views. Particles were displayed in a 24 nm box.

In order to confirmed the integrity of MSP2N2 reconstituted particles, 2D classification was run on a small set of particles. 2D classification results showed clear density for the ECD and TMD in close proximity to each other (Figure 6.12). In addition, projections of intact particles from various orientations were observed. The 2D averages generated here were used as templates for automated picking on the entire dataset, giving 448,415 picked particles in total.

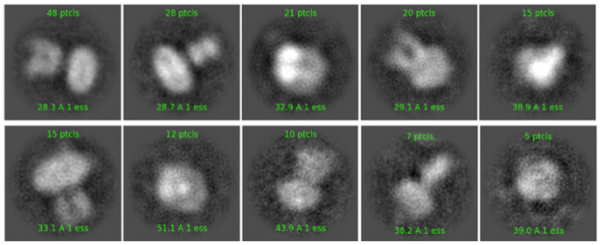


Figure 6.12 Confirmation of intact $GABA_BR$ reconstituted in MSP2N2 nanodiscs.

Around 400 particles (boxsize 24nm) was used in the 2D classification. Majority of classes had densities representing the intact $GABA_BR$ where the ECD and TMD were clearly identified. A set of classes showing projections from different views are displayed here.

Auto-picked particles were "cleaned" through iterative rounds of 2D classification. Some 2D classes showed two or multiple receptors in one MSP2N2 reconstituted particle (Figure 6.13a) and particles from such classes were removed. 56,506 particles were kept after particle cleaning during 2D classification stage. 2D class averages showed projections from a wide range of views (Figure 6.13b), including many particles in tilted angled classes. Density for the ECD and TMD could be readily differentiated, where the ECD displayed the fly-trap topology and the TMD as a planar circular disc. Density for the linker peptide between two domains could be found in some classes, mostly from side view classes. Density for the intracellular domain was not visible in 2D averages, although density for the coiled-coil could be identified upon visual inspection of the raw micrograph. Larger sampling on the micrographs revealed

that the coils were highly mobile and rotate around relative to the main receptor body (as exemplified by a selection of particles in Figure 6.11b). As a consequence, density of the coiled-coil was averaged out during classification.

Most classes had the TMD density in the centre of the average (Figure 6.13b), suggesting that alignment of MSP2N2 particles was more dominant by the nanodisc region. The nanodisc region density had a clear disc-shaped outline, possibly contributed by the scaffolding protein MSP2N2. There was also weak density inside the disc, possibly coming from the $GABA_BR$'s transmembrane helices. This was the first-time that receptor's transmembrane density appeared and could be an indication of improved rigidity of the region. In comparison, alignment of ECD was compromised and did not show clear secondary structural features as before.

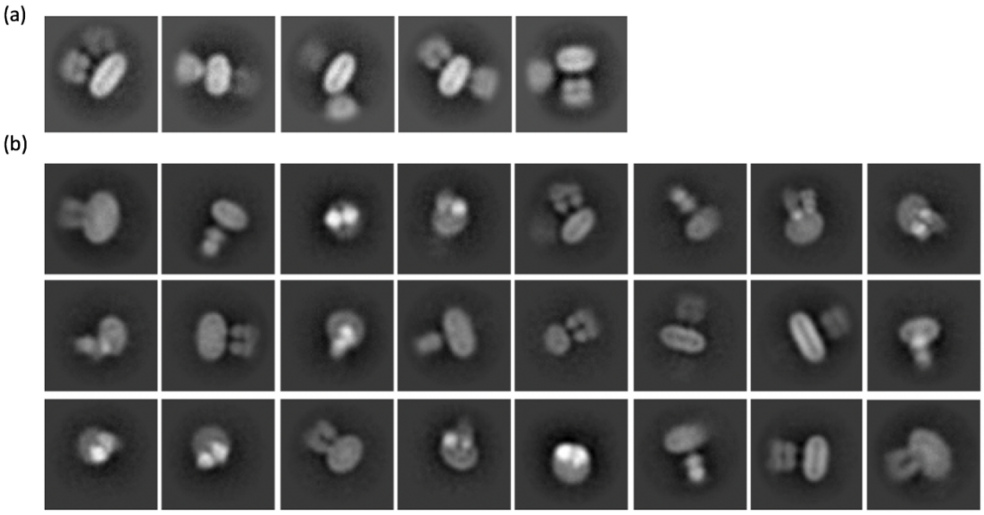


Figure 6.13 2D classification results of MSP2N2-reconstituted particles.
(a) 2D averages showing multiple receptors reconstituted into one disc. (b) Representative 2D class averages of the MSP2N2-reconstituted particles from various views.

3D classification was performed using a low-pass filtered model reconstructed from amphipols-reconstituted data in section 5.2.2 as a starting reference model.

Subsequent 3D classifications were run using a model of MSP2N2-reconstituted particles generated in the first round of 3D classification. Reconstruction with the best reported nominal resolution was at 10.1 Å (Figure 6.14a), using 56,506 particles. The model had the ECD VFT sitting on an elliptical shaped disc and no density for the intracellular domain. Density for the lipid nanodisc region was estimated to be 5 nm thick, and 8 nm x 10 nm in diameter viewing from the top. The volume of the TMD density was sufficient to accommodate the $GABA_BR$ 14 helices as estimated in section 5.3.2.

Consistent with the 2D classification result, the nanodisc region was found to be placed in the centre of the reconstruction map. Based on these observations, the TMD density of MSP2N2 reconstituted particles appeared to dominant in alignment. This was different from amphipols-solubilised samples where the ECD VFT was the stronger feature. As $GABA_BR$ ECD VFT is a stable domain, the changes in particle alignment patterns could be an indication of increased TMD rigidity in the MSP2N2 nanodisc.

The ECD region displayed an overall fly-trap shape, however it was less well-resolved than in the amphipols-reconstituted models (Figure 6.14a). Density was weaker in the upper lobes where volume of the upper lobes is shorter than the ones below. The ECD was found to sit at a tilted angle above the TMD, in contrast to $GABA_BR$-amphipols particles models where the ECD was perpendicular up to the TMD. Linker density connecting the ECD to the TMD was observed for the subunit located in the center of the disc but not the subunit reconstructed at the peripheral (Figure 6.14a).

$GABA_BR$ ECD and TMD domains are linked through one peptide linker on each subunit. The linker is estimated to be 14 residues long in $GABA_{B1}$ subunit (Asp^{459}-Lys^{473}) and 13 residues long in $GABA_{B2}$ subunit (Asp^{467}-Pro^{480}) based on domain boundaries of existing ECD VFT crystal structure and structure-based sequence alignment of TMD with mGluR (Appendix B). Rigid body movements between domains could have occurred at the linker region. Tilting of ECD in the reconstructed model was speculated to be a result of flexible movements at the domain interface. A more careful analysis on the raw micrographs and 2D averages (Figure 6.14d) agreed with this hypothesis where the ECD was found to rotate relative to the TMD. Taken together, it appeared that the apparent poor reconstruction of ECD in this data set is the result of a shift in alignment priority to the TMD, compounded by particle flexibility at domain interface .

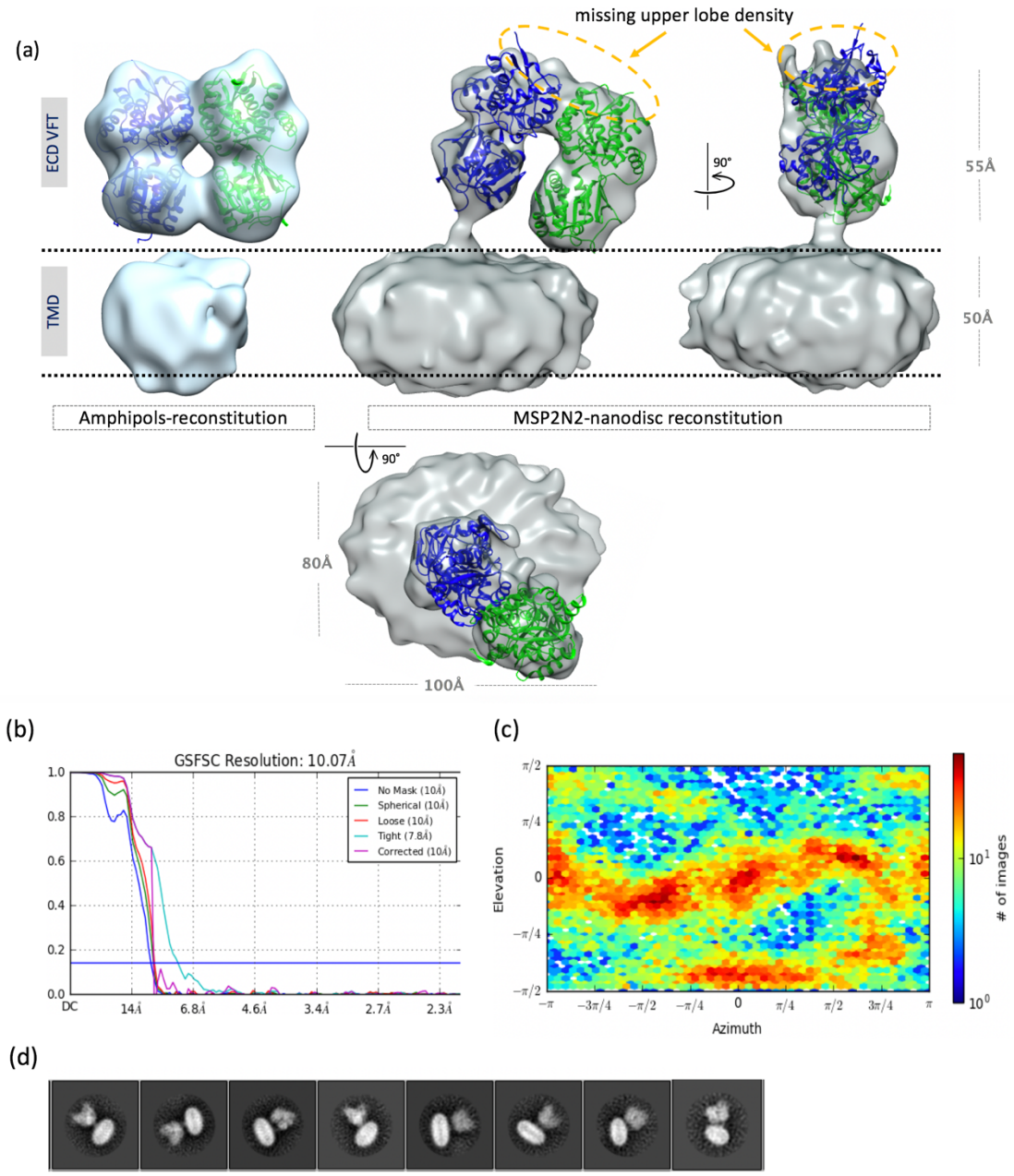


Figure 6.14. GABA$_B$R-MSP2N2 nanodisc particles reconstruction.

(a) Comparison between reconstructions from amphipols-reconstituted particles (light blue) and MSP2N2-reconstituted particles (grey). The ECD VFT crystal structure (PDB ID 4MQE) was fitted into the extracellular region density of both models as a rigid-body. (b) Fourier Shell Correlation analysis of MSP2N2 reconstituted reconstruction map. (c) Azimuth map showing particles distribution from different orientations. (d) Selected 2D class averages demonstrating the relative movements between ECD and TMD domains.

6.2.6 Forming the GPCR-G protein complex

Receptors reconstituted in MSP2N2 nanodiscs started to show transmembrane density during 2D classification. However, the alignment was still limited and clear secondary structures was not resolved. Investigation of the TMD region was prioritised in this project since no structure had been reported for this domain. As demonstrated by the high-resolution reconstruction of ECD alone in section 6.2.3, reconstruction of particle of equivalent size to a single domain was possible with current data collection setup given a rigid sample. In order to stabilise the TMD in one specific conformation, the $GABA_BR$ complexed with its endogenous binding partner, i.e., G protein, was investigated. The newly developed mini$G_{\alpha o}$ protein was chosen to stabilise $GABA_BR$ in a fully active state. mini-$G_{\alpha o}$ is an engineered Gα subunits containing only the GTPase domain [220]. It displays improved stability over the native G-protein and induces the same signalling effect when coupled to GPCRs [220]. Furthermore, the mini-$G_{\alpha o}$ could be fused with a GFP at its N-terminus for detection of complex formation on FSEC [220]. The improved stability and ease of detection makes GFP-mini$G_{\alpha o}$ an ideal tool for structural studies.

Separately purified $GABA_BR$-MSP2N2 and 1.3-fold molar ratio GFP-mini$G_{\alpha o}$ (kind gift from Dr. Rony Nehmé – MRC LMB) were incubated overnight in the presence of excess amount of agonist 3-APPA and a positive allosteric modulator rac-BHFF. Complex formation was screened using $GABA_BR$-MSP2N2 purified under various NaCl concentrations. At 10mM or lower NaCl conditions, $GABA_BRs$-MSP2N2/GFP-mini$G_{\alpha o}$ complex was formed and could be purified by gel filtration. A clear shift in GFP-mini$G_{\alpha o}$ fluorescent signal was observed under the low Na^+ conditions, yielding protein peaks

at 1.4ml and 1.7ml elution volume (Figure 6.15a). Fractions of the peak at 1.4 ml were analysed by western-blot and SDS-PAGE (Figure 6.15b and c) and showed bands for both $GABA_BR$ and GFP-mini$G_{\alpha o}$. Co-elution of the two binding partners suggested that stable protein complex had been formed. The later peak at 1.7ml contained GFP-mini$G_{\alpha o}$ only based on SDS-PAGE analysis (Figure 6.15d), although the reason for the increase in its apparent MW on SEC from GFP-mini$G_{\alpha o}$ alone sample injection was unclear.

Intensity of the GFP-mini$G_{\alpha o}$ only peak was much higher (approx. 5-fold) than the $GABA_BR$/GFP-mini$G_{\alpha o}$ complex peak. The differences in peak intensities suggested that majority of mini$G_{\alpha o}$ added was not bound to the receptor, given a theoretical stoichiometry of 1 to 1 in this complex and mini$G_{\alpha o}$ was supplemented at 1.3 molar ratio of $GABA_BR$. The relative small signal in co-elution peak indirectly suggested that not all $GABA_BR$ particles had a mini$G_{\alpha o}$ bound.

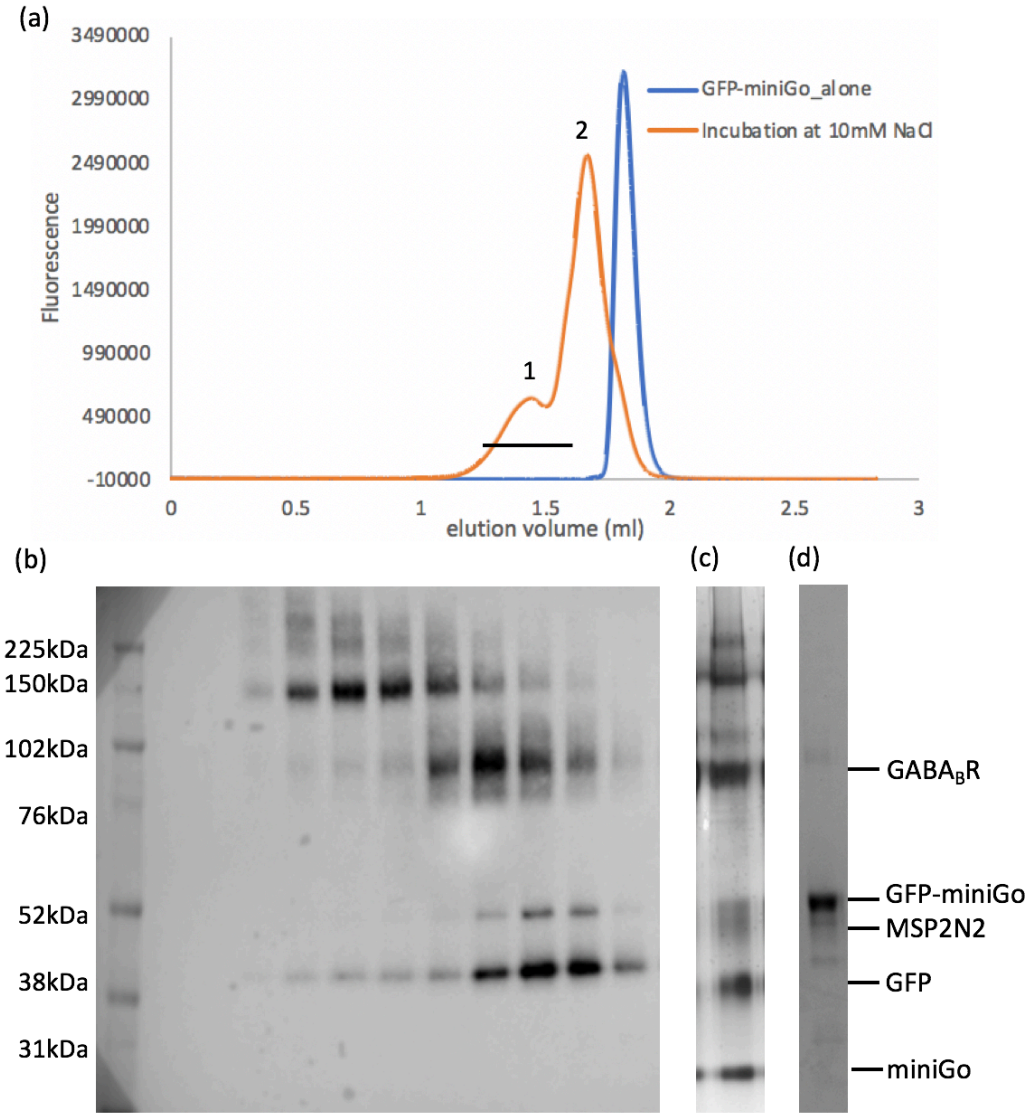


Figure 6.15 GABA$_B$R-miniG$_{\alpha o}$ complex formation under low Na^+ environment.

Affinity purified GABA$_B$R-MSP2N2 particle was eluted in a buffer containing 10 mM NaCl and 20 mM Tris pH 8.0. GABA$_B$R was incubated with 1.3 molar ratio of GFP-miniG$_{o1\text{-}R12}$ overnight in presence of 1mM 3-APPA, 100 uM rac-BHFF, 0.1 mM GDP and 0.2 U/ml apyrase. Complex was analysed and purified by SEC using a Superose 6 increase column 3.2/300 equilibrated in 10mM NaCl, 20mM Tris pH 8.0, 10 mM MgCl2, 1mM 3-APPA and 100 uM rac-BHFF. (a) FSEC analysis of incubation products (orange) in comparison to GFP-miniG alone (blue). Peak 1 contained the GABA$_B$R-miniG$_{\alpha o}$ complex and peak 2 contained only GFP-miniG$_{\alpha o}$. (b) Western blotting analysis of fractions from peak 1. Protein bands were detected by anti-GFP antibody. (c) SDS-PAGE of the co-elute peak fraction at 1.4ml (i.e. peak 1). MSP2N2 and GFP-miniG$_{o1\text{-}R12}$ were not well separated on the gel due to their similarity in size. (d) SDS-PAGE of peak fraction from peak at 1.7 ml. Presence of GFP-miniG$_{o1\text{-}R12}$ was confirmed from the fluorescence signal in this peak.

In order to stabilise the complex, possibility to crosslink the complex using chemical reagent was tested. Successful crosslinking will be indicated by emergence of a protein band at the corresponding MW of the complex (approx. 280 kDa) on SDS-PAGE. Three candidate reagents (CovalX K100, CovalX K200 and BS^3) were tested.

Protein complex sample incubated overnight was treated with each reagent separately and had the cross-linked products analysed by SDS-PAGE (Figure 6.16). The cross-linking reagents tested so far appeared to cause protein aggregation, as evidenced by migration of protein bands to the higher MW smeary bands on the gel. A band corresponding to stable cross-linked complex (theoretical MW of approx. 280 kDa) was not clearly visible on the gel.

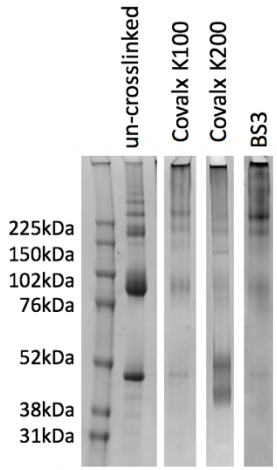


Figure 6.16 Effects of chemical cross-linking treatments on the $GABA_BR$-mini$G_{\alpha o}$ incubated sample.

For each reagent, crosslinking experiments were performed using a constant concentration of protein complex with the cross-linker under a range of screening concentrations following manufacturer's protocol. Figure contains one representative SDS-PAGE lane result of each reagent for comparison. The untreated sample lane served as a negative control showing clear protein bands.

6.2.7 Data processing and analysis of the WT-GABA$_B$R-MSP2N2-GFPminiGo particles

Although receptor binding occupancy was estimated to be low, GABA$_B$R-miniG$_{\alpha o}$ complex particles and GABA$_B$R only particles should be separable by classification if sufficient number of particles were imaged. Cryo-EM grids were therefore prepared using the SEC fraction containing GABA$_B$R-MSP2N2/GFP-miniG$_{\alpha o}$ complex and dataset was collected on a Titan Krios microscope with VPP (Table 6.2).

Table 6.2 Data collection statistics for WT GABA$_B$R-MSP2N2-GFPminiGo

Microscope	Titan Krios
Voltage (kV)	300
Imaging mode	Counting
Detector	Falcon III
Slit (eV)	N/A
Pixel size (Å)	1.06
Total Dose (e−/Å^2)	30
Exposure time (s)	69.3sec
Dose rate 2 (e$^-$/pix/s)	0.56
Number of frames	75
Frame length (s)	0.8
Focusing range (μm)	-0.5 to -0.7 with Volta Phase Plate

115,134 particles were picked automatically using representative views from section 6.2.5 as templates. In addition to 2D classes representing the GABA$_B$R particles, initial classification results showed class averages of much smaller size than the intact GABA$_B$R (Figure 6.18a circled). Dimensions and contour of these 2D averages presumably corresponded to miniG$_{\alpha o}$-GFP which has dissociated from the complex. There were 26,740 particles in such class, while there were 35,894 particles in GABA$_B$R containing classes.

Classes corresponding to the size of $GABA_BR$ particles were selected and classified further. Given the results in section 6.2.5, classes showing particles with severe tilting between domains were removed to minimise sample conformational heterogeneity. 2D averages (Figure 6.18b) of the selected particles showed good alignment for both the ECD and TMD. Both peptide linkers were visible in projections viewed from the front. Density accounting for the intracellular domain or the G protein was not visible. Either the miniG bound intracellular region was still too flexible to be seen or that the $GABA_BR$-mini$G_{\alpha o}$-GFP complex had fallen apart. Taken together with results of the smaller 2D class averages observed earlier, the latter was more likely to be the reason.

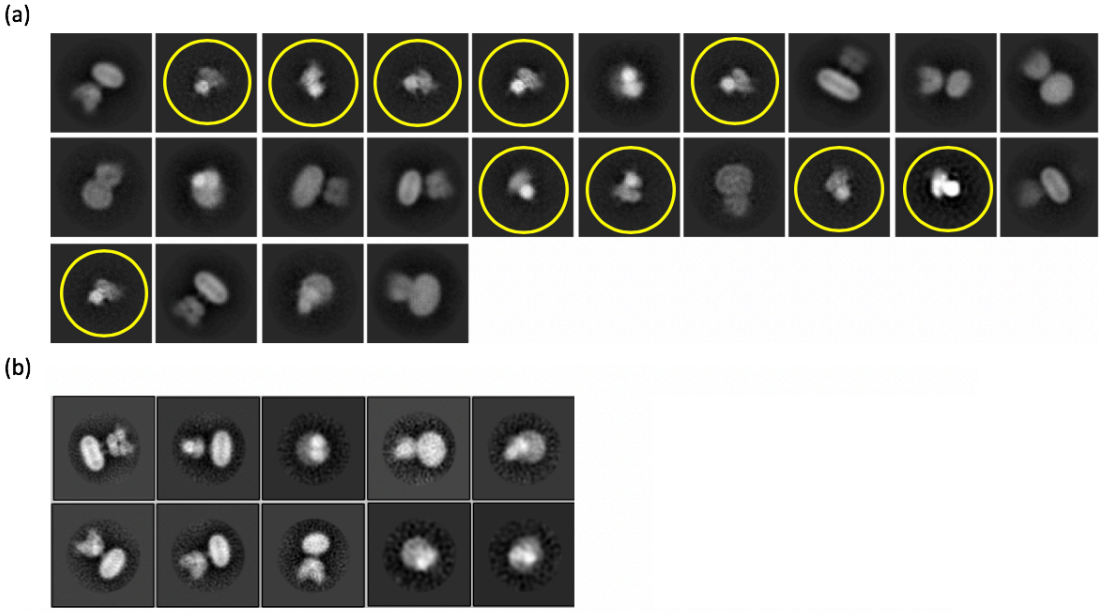


Figure 6.17 Representative 2D classification results from non-crosslinked $GABA_BR$/mini$G_{\alpha o}$ sample.

(a) 2D averages of initial rounds of classifications. Particles were masked with a 220 pixel (20.6nm) diameter circular mask. 2D classes circled in yellow represent the potential miniG only particles. (b) 2D class averages after of the selected particles after iterative rounds of cleaning.

Nevertheless, it was possible to generate a 3D reconstruction at a nominal resolution of 12.3 Å (Figure 6.18) using 21,918 particles. The model showed only partial density for the ECD VFT where density for the upper lobes were missing (Figure 6.18a). As the VFT density was shown to be intact in the selected 2D averages, the weakened density in 3D model was probably caused by poor alignment of the domain during 3D reconstruction. Density for both linkers were clearly reconstructed, with an estimated length of 1.7 nm.

TMD appeared as a 6nm thick oval disc, with an estimated diameter of 11.5 nm x 7nm. Disc reconstructed in this model is bigger and more regular than the model in section 6.2.5 (Figure 6.18a). Assuming G-protein was not bound to stabilise a particular conformation, the better reconstructed TMD was more likely as a result of more homogenous particles selected during 2D classification. Some cylinder densities were observed inside the disc (Figure 6.18b), which might come from $GABA_BR$ transmembrane helices. Docking using crystal structure of the $mGluR_5$ 7-helices was attempted (PDB ID: 4OR2) , however features available was not strong enough to give reliable fitting.

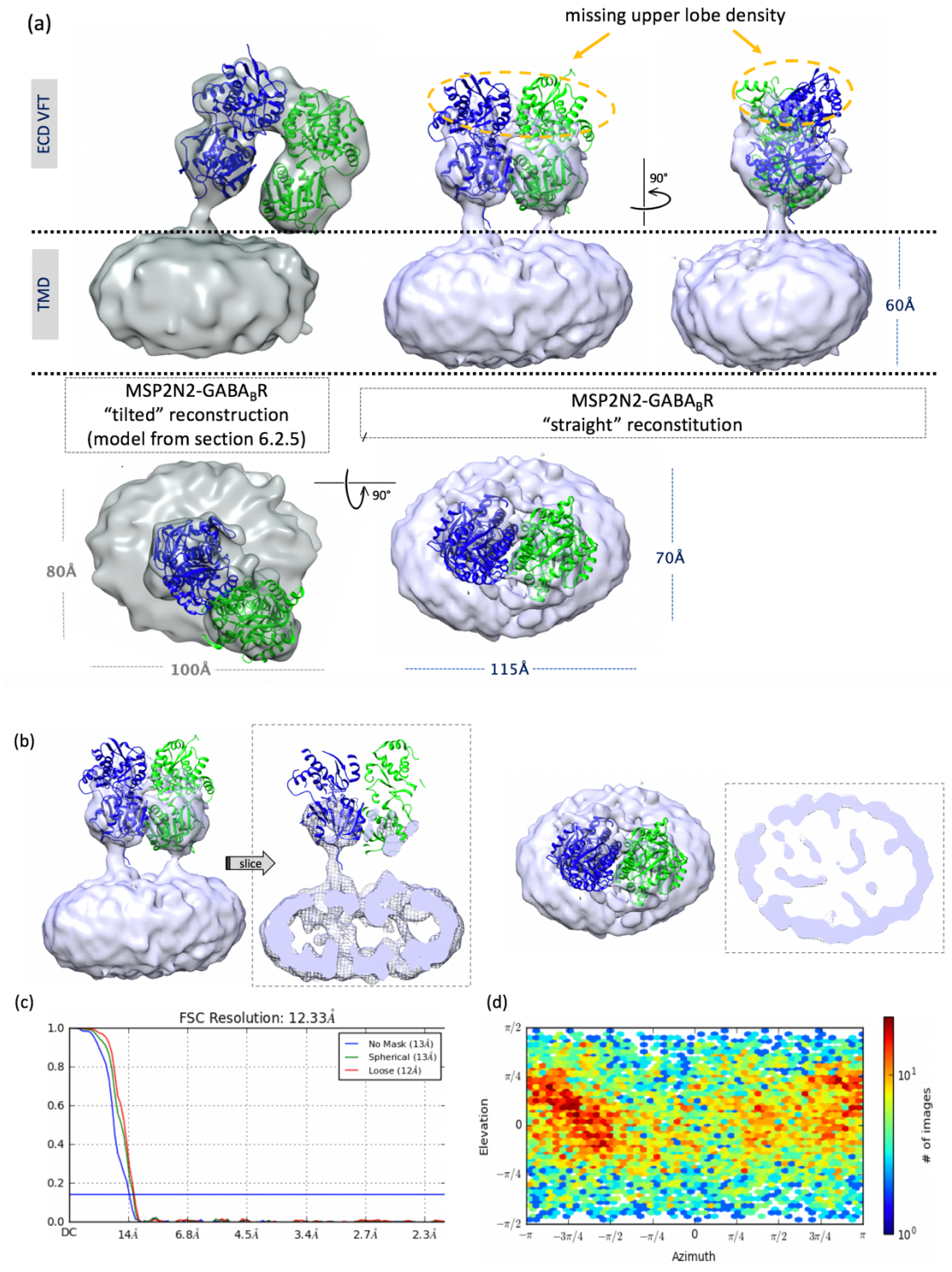


Figure 6.18 3D reconstruction model of MSP2N2 reconstituted particles.

(a) Comparison between the model reconstructed here (light blue) and in section 6.2.5 (grey). The 3D model had missing density for the VFT upper lobes (area highlighted in yellow dashed line). (b) Cross-sectional slices view from front and top. (c) Fourier Shell Correlation analysis of the reconstruction map. (d) Azimuth map showing particles distribution from different orientations.

6.3 Discussion

6.3.1 Choosing the suitable nanodisc reconstitution system

In this chapter, reconstitution of detergent solubilised $GABA_BR$ in to different nanodisc particles was investigated. MSP2N2 reconstituted $GABA_BR$ particles showed better properties when compared against the other two nanodisc systems tested, yielding reconstituted $GABA_BR$ in nanodisc particles of a suitable and homogenous size. As shown in section 6.2.1, size of saposin A reconstituted particles were not precisely controlled and gave a broad SEC curve. The highly variable disc sizes caused very poor particle alignment during image processing. This was somewhat consistent the salipro particle's feature of its ability to adapt different disc sizes from conventional nanodiscs. Comparing to other lipid nanodisc systems, salipro particles are not yet as well established and no high-resolution structure has been reported using the system to date.

The MSP nanodisc system, which provides a pre-defined disc size was assessed next. MSP1D1 reconstituted $GABA_BR$ particles on EM-grid were found to be isolated single domains of either ECD or TMD, with only small percentage of intact receptors. MSP1D1-$GABA_BR$ particles were reconstituted while still bound on the first affinity beads. The reconstituted particles must have remained as heterodimers during affinity purification, since they were purified using a N-terminal tag on $GABA_{B1}$ subunit and a C-terminal tag on $GABA_{B2}$ subunit. Particles broken at domain interface would only contain one tag and get removed during the double purification. One possibility is that the dissociation happened during the concentrating step before gel filtration, e.g. due to proteolytic cleavage at the linker region. When compared to the later

tested MSP2N2 nanodisc, gel filtration elution volume of MSP1D1 particles were slightly earlier.

A bigger nanodisc scaffolding protein, the MSP2N2, was tested next. As confirmed in section 6.2.5, MSP2N2 reconstitution yielded intact $GABA_BR$ particles. Dimensions of the MSP2N2 disc was suitable to accommodate $GABA_BR$ transmembrane region which could be aligned during processing. The apparent oval shape of the disc might be a result of MSP2N2 adapting to the shape of the $GABA_BR$'s TMD. Particles had mostly single receptor reconstitution, with some incorporating more than one receptor per disc. Formation of $GABA_BR$ tetra/oligo-mers is possible and has been reported under the condition where $GABA_BR$ local density were high [221]. However oligomers were believed to be formed by association of heterodimers through their $GABA_{B1}$ subunit [221], which was not consistent with the observations in the 2D averages here. Therefore, reconstitution of multiple receptors per particle here was more likely to be an artefact during the reconstitution process. Larger size nanodiscs were not tested, as they may cause "floating" of the receptor within the disc and higher percentage of multiple receptors reconstitution per disc.

Reconstruction using lipid nanodiscs particles started to show signs of transmembrane helices while no such feature was observed from bigger amphipol reconstitution datasets. The improved TMD reconstruction could be an indication of enhanced structural rigidity in the region in nanodisc-reconstituted particles. The nanodisc provides a more regular structure, as compared to amphipol polymers wrapping around a proteins hydrophobic region non-structurally. Nanodiscs have shown to be useful in functional reconstitution of GPCR, as exemplified by the study

of β2-adrenergic receptors [209]. Another possible contributing factor was the change of receptor construct. Removal of the bRIL to restore $GABA_BR$ wild-type sequence might help to return the receptors to a more stable state. Effect of different constructs could not be directly compared since the data was collected on different microscopes using different reconstitution systems. In addition, particles from a wider range of views were observed in 2D classification in nanodisc-reconstituted sample. Again, it was difficult to confirm the contributing factor as the data was collected on different microscopes. Possible reasons included that nanodiscs particles adapt more views and the contrast enhance phase plate data allowed more accurate alignment.

6.3.2 $GABA_BR$ exhibited large degree of inter-domain flexibility

$GABA_BR$ particles preserved in Cryo-EM grids were found to display high flexibility at the domain interface. It was clear upon visual inspection of the raw micrograph and 2D class averages that $GABA_BR$ individual domains rotate relative to each other. Flexibility between domains caused difficulty in aligning the whole particle, which was indicated in the 2D classification results where in one given class only one domain showed accurate alignment and signal from the other domains were averaged out. In addition, poor global particle alignment had prevented accurate 3D reconstruction even when secondary structural features were observed during 2D classification stage.

$GABA_BR$ is the only Class C GPCR lacking a cysteine rich domain where the ligand-binding domain (i.e. $GABA_BR$ ECD VFT) is directly connected to the TMD. The two domains are connected by only one peptide linker of estimated 13-14 residues long in each subunit. The two linkers would allow space for rotational movements

between domains relative to each other. Another source of the inter-domain movements might arise during Cryo-EM sample preparation stage. $GABA_BR$ ECD, but not yet the TMD, was found to prefer certain orientations based on 2D classification results. The differences in preferential orientations between domains might have caused receptor particles to twist at the linker region in free-standing ice and induced artificial inter-domain rotations.

There are several ways to reduce the movements between domains. One possibility was to shorten the length of the linker peptide by site-directed mutagenesis. However, removal of the linker had been shown to prevent surface expression of the $GABA_{B2}$ subunit [222]. This approach was thus not pursued to avoid affecting receptor functionality. Another possible way was to have other proteins to bind at the domain interface. Binding at the interface has two effects: one is to stabilise the receptor conformation, another is to act as steric hindrance to block the inter-domain movements. Ilama nanobodies are potentially good binding partners, since they bind only to specific conformation often found at cavities with very high affinity [223]. Therefore a Ilama nanobody library has been raised against the wild-type integral human $GABA_BR$ (in collaboration with Dr Jan Steyaert, VIB, Belgium). $GABA_BR$ specific nanobodies will be characterised using biophysical techniques such as SPR or Octet and used in future EM sample preparations.

Assuming rigid-body movement between domains, another approach could be to solve each domain separately. The isolated ECD fly-trap domain appeared well-ordered and could be individually reconstructed in section 6.2.3 using images collected with VPP. ECD from MSP1D1 was able to be individually reconstructed for

two reasons. Firstly, the implementation of VPP greatly enhanced image contrast to allow accurate alignment of the small particles. Secondly, the separation from the TMD removed the otherwise closely adjacent interference signals. In contrast, the transmembrane domain was challenging as its alignment was driven by the nanodisc scaffolding protein instead of $GABA_BR$ transmembrane signal, preventing accurate reconstruction of the TMD helices.

6.3.3 Stabilisation of the receptor TMD by G-protein

In order to improve rigidity of the transmembrane helices, the engineered mini-$G_{\alpha o}$ was employed to stabilise the receptor in a fully activated state conformation. Mini-$G_{\alpha o}$ binding would also add distinctive features to the TMD region to facilitate particle alignment. $GABA_BR$/Mini-$G_{\alpha o}$ complex was found to only form under low sodium conditions (<10 mM). This result is consistent with the previous report that Na^+ ions act as negative allosteric modulators for GPCR [137], where a Na^+ binding site has been identified in the receptor's transmembrane region and is proposed to stabilise the ligand-free receptor conformation [224]. Therefore, presence of higher Na^+ concentrations in the $GABA_BR$ sample might have prevented Mini-$G_{\alpha o}$ binding through a similar mechanism.

Although some $GABA_BR$-mini$G_{\alpha o}$ complex could be purified by gel filtration, mini$G_{\alpha o}$ binding occupancy was estimated that to be very low percentage based on FSEC profile. GPCR-G protein complexes are often found to be transient with weak binding affinities [2]. The relatively low affinity for the binding might have caused some $GABA_BR$-mini$G_{\alpha o}$ complex to dissociate during SEC. The similarity in sizes of $GABA_BR$ alone and $GABA_BR$-mini$G_{\alpha o}$ complex made it difficult to separate the two

species on SEC. Therefore Cryo-EM specimen prepared was likely to contain a mixtures of $GABA_BR$ alone particles and $GABA_BR$-mini$G_{\alpha o}$ complex particles. According to 2D classification results, the complex applied on grid had disassembled in free-standing ice. One possible reason for the dissociation could be due to contacts at the air-water interface.

The combination of sub-optimum receptor binding occupancy and interactions with the air-water interface resulted in loss of $GABA_BR$/mini$G_{\alpha o}$ particles in free standing ice. Attempts to stabilise the complex by chemical cross-linking with the reagents tested so far appeared to cause protein aggregation. Another approach could be to create a fusion construct connecting the G protein to the C-terminus of $GABA_{B2}$ subunit. Such design would bring the $G_{\alpha o}$ protein to close proximity to the receptor to promote complex formation. Alternatively, EM sample preparation strategy surpassing the air-water interface problem could be tested [225]. For example to provide a physical support using graphene oxide coated grid or continues carbon grid [225].

6.4 Conclusion

The usage of Krios microscope with VPP implementation greatly enhanced image contrast to allow more accurate $GABA_BR$ particle picking and alignment. Artificial lipid nanodiscs provided by MSP2N2 were identified to be the best reconstitution system for $GABA_BR$ in which the receptor showed binding activity to G protein. In this chapter, new insights about the receptor architecture and dynamic were gained. The high degree of interdomain flexibility presented a major obstacle for accurate alignment of the whole particle.

Chapter 7. Concluding remarks and future work

So far, expression and purification of the integral human $GABA_BR$ had been carefully optimised. Functional $GABA_BR$ had been successfully reconstituted into lipid nanodiscs for cryo-EM analysis. The extracellular domain alone was reconstructed to modest resolution to reveal its structural arrangements in a close-to-native environment. The transmembrane and intracellular region reconstruction remained challenging due to their intrinsic dynamics, compounded by the flexibility at domain interfaces. In order to rigidify the transmembrane region, complex with G protein were formed using $GABA_BR$ embedded in MSP2N2 nanodisc. However, the complex was found to be unstable and dissociate during EM sample preparations. Future works should focus on stablislisation of the $GABA_BR$-G protein complex. Strategies such as chemical cross-linking or construct fusion could be explored.

Recently, a new class of cytoplasmic proteins called K^+channel tetramerization domain (KCTDs) were reported to interact with the $GABA_BRs$ [39]. KCTDs are believed to constitutively bind to the C-terminal domain of $GABA_{B2}$ subunit [67, 226]. In addition, KCTDs bind the $G_{\beta\gamma}$ and tether the G-protein at the receptor to accelerates G-protein signalling [227]. This class of proteins have the potential to form KCTD/$GABA_BR$ or KCTD/$GABA_BR$/G-protein complexes. Solving the structure of such complex will provide insights regarding G-protein activation and signal transduction.

Complex with other proteins could also be sought, with ideal candidates exhibiting high affinity and slow off rate. Therefore a large library of Ilama nanobodies was raised against the recombinantly expressed wild-type human $GABA_BR$ (Appendix

B). Nanobodies only bind to conformational epitopes, thereby stabilising the receptor in a particular conformational state. Furthermore, they had been shown to bind with very high affinity, forming tight complexes to stabililise receptor's transient conformation during sample preparation. Screening and characterisation of nanobodies using biophysical approaches, such as the octet and SPR, will be the next step of this project. Identification of nanobodies binding to domain interface might prevent interdomain flexibility.

Getting the higher resolution reconstructions were not possible without the recent technical developments for cryo-EM sample imaging and data analysis. A more stable microscope, improved direct electron detectors and development of the Volta Phase Plate all greatly enhanced the quality of images collected. The higher signal-to-noise ratio images collected provided better input data for downstream processing. As a result of the better contrasted data, accuracy of image alignment, particle picking and classifications were improved. On the data processing aspect, more software packages and analysis algorithms became available. These algorithms provided alternative processing strategies to resolve and validate structures reconstructed. For example, the cryoSPARC suite used in this thesis provided an alternative which employed different classification and model building algorithms than the RELION suite.

Together with the rapidly advanced cryo-EM microscope imaging and processing techniques, high-resolution structural determination of the $GABA_BRs$ should be achieved once factors to rigidify the receptor are identified.

References

1. Rosenbaum, D.M., S.G. Rasmussen, and B.K. Kobilka, *The structure and function of G-protein-coupled receptors.* Nature, 2009. **459**(7245): p. 356-63.
2. Khoor, A., et al., *Giant cell interstitial pneumonia in patients without hard metal exposure: analysis of 3 cases and review of the literature.* Hum Pathol, 2016. **50**: p. 176-82.
3. Foord, S.M., et al., *International Union of Pharmacology. XLVI. G protein-coupled receptor list.* Pharmacol Rev, 2005. **57**(2): p. 279-88.
4. Kolakowski, L.F., Jr., *GCRDb: a G-protein-coupled receptor database.* Receptors Channels, 1994. **2**(1): p. 1-7.
5. Fredriksson, R., et al., *The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints.* Mol Pharmacol, 2003. **63**(6): p. 1256-72.
6. Lagerstrom, M.C. and H.B. Schioth, *Structural diversity of G protein-coupled receptors and significance for drug discovery.* Nat Rev Drug Discov, 2008. **7**(4): p. 339-57.
7. Schulte, G., *International Union of Basic and Clinical Pharmacology. LXXX. The class Frizzled receptors.* Pharmacol Rev, 2010. **62**(4): p. 632-67.
8. Palczewski, K., et al., *Crystal structure of rhodopsin: A G protein-coupled receptor.* Science, 2000. **289**(5480): p. 739-45.
9. Pandy-Szekeres, G., et al., *GPCRdb in 2018: adding GPCR structure models and ligands.* Nucleic Acids Res, 2018. **46**(D1): p. D440-D446.
10. Zhang, D., Q. Zhao, and B. Wu, *Structural Studies of G Protein-Coupled Receptors.* Mol Cells, 2015. **38**(10): p. 836-42.
11. Garcia-Nafria, J., et al., *Cryo-EM structure of the serotonin 5-HT1B receptor coupled to heterotrimeric Go.* Nature, 2018. **558**(7711): p. 620-623.
12. Liang, Y.L., et al., *Phase-plate cryo-EM structure of a class B GPCR-G-protein complex.* Nature, 2017. **546**(7656): p. 118-123.
13. Bettler, B., et al., *Molecular structure and physiological functions of GABA(B) receptors.* Physiol Rev, 2004. **84**(3): p. 835-67.
14. Bowery, N.G., et al., *Baclofen: a selective agonist for a novel type of GABA receptor proceedings.* Br J Pharmacol, 1979. **67**(3): p. 444P-445P.
15. Bowery, N.G., et al., *(-)Baclofen decreases neurotransmitter release in the mammalian CNS by an action at a novel GABA receptor.* Nature, 1980. **283**(5742): p. 92-4.
16. Owens, D.F. and A.R. Kriegstein, *Is there more to GABA than synaptic inhibition?* Nat Rev Neurosci, 2002. **3**(9): p. 715-27.
17. Chebib, M., *GABAC receptor ion channels.* Clin Exp Pharmacol Physiol, 2004. **31**(11): p. 800-4.
18. Martinez-Delgado, G., et al., *An Update on GABArho Receptors.* Curr Neuropharmacol, 2010. **8**(4): p. 422-33.
19. Bormann, J., *Electrophysiology of GABAA and GABAB receptor subtypes.* Trends Neurosci, 1988. **11**(3): p. 112-6.
20. Preininger, A.M. and H.E. Hamm, *G protein signaling: insights from new structures.* Sci STKE, 2004. **2004**(218): p. re3.

21. Coleman, D.E., et al., *Structures of active conformations of Gi alpha 1 and the mechanism of GTP hydrolysis.* Science, 1994. **265**(5177): p. 1405-12.
22. Rasmussen, S.G., et al., *Crystal structure of the beta2 adrenergic receptor-Gs protein complex.* Nature, 2011. **477**(7366): p. 549-55.
23. Gilman, A.G., *G proteins: transducers of receptor-generated signals.* Annu Rev Biochem, 1987. **56**: p. 615-49.
24. Ross, E.M. and T.M. Wilkie, *GTPase-activating proteins for heterotrimeric G proteins: regulators of G protein signaling (RGS) and RGS-like proteins.* Annu Rev Biochem, 2000. **69**: p. 795-827.
25. Syrovatkina, V., et al., *Regulation, Signaling, and Physiological Functions of G-Proteins.* J Mol Biol, 2016. **428**(19): p. 3850-68.
26. Asano, T. and N. Ogasawara, *Uncoupling of gamma-aminobutyric acid B receptors from GTP-binding proteins by N-ethylmaleimide: effect of N-ethylmaleimide on purified GTP-binding proteins.* Mol Pharmacol, 1986. **29**(3): p. 244-9.
27. Campbell, V., N. Berrow, and A.C. Dolphin, *GABAB receptor modulation of Ca2+ currents in rat sensory neurones by the G protein G(0): antisense oligonucleotide studies.* J Physiol, 1993. **470**: p. 1-11.
28. Greif, G.J., et al., *Altered regulation of potassium and calcium channels by GABA(B) and adenosine receptors in hippocampal neurons from mice lacking Galpha(o).* J Neurophysiol, 2000. **83**(2): p. 1010-8.
29. Menon-Johansson, A.S., N. Berrow, and A.C. Dolphin, *G(o) transduces GABAB-receptor modulation of N-type calcium channels in cultured dorsal root ganglion neurons.* Pflugers Arch, 1993. **425**(3-4): p. 335-43.
30. Morishita, R., K. Kato, and T. Asano, *GABAB receptors couple to G proteins Go, Go* and Gi1 but not to Gi2.* FEBS Lett, 1990. **271**(1-2): p. 231-5.
31. Franek, M., et al., *The heteromeric GABA-B receptor recognizes G-protein alpha subunit C-termini.* Neuropharmacology, 1999. **38**(11): p. 1657-66.
32. Wettschureck, N. and S. Offermanns, *Mammalian G proteins and their cell type specific functions.* Physiol Rev, 2005. **85**(4): p. 1159-204.
33. Kulik, A., et al., *Subcellular localization of metabotropic GABA(B) receptor subunits GABA(B1a/b) and GABA(B2) in the rat hippocampus.* J Neurosci, 2003. **23**(35): p. 11026-35.
34. Brown, K.M., et al., *Activation of the gamma-Aminobutyric Acid Type B (GABA(B)) Receptor by Agonists and Positive Allosteric Modulators.* J Med Chem, 2015. **58**(16): p. 6336-47.
35. Wu, L.G. and P. Saggau, *Presynaptic inhibition of elicited neurotransmitter release.* Trends Neurosci, 1997. **20**(5): p. 204-12.
36. Dolphin, A.C., et al., *G protein modulation of voltage-dependent calcium channels and transmitter release.* Biochem Soc Trans, 1993. **21**(2): p. 391-5.
37. Zylbergold, P., N. Ramakrishnan, and T. Hebert, *The role of G proteins in assembly and function of Kir3 inwardly rectifying potassium channels.* Channels (Austin), 2010. **4**(5): p. 411-21.
38. Noebels JL, A.M., Rogawski MA, et al., editors., *Jasper's Basic Mechanisms of the Epilepsies*. 2012.

39. Schwenk, J., et al., *Modular composition and dynamics of native GABAB receptors identified by high-resolution proteomics.* Nat Neurosci, 2016. **19**(2): p. 233-42.
40. Kaupmann, K., et al., *Expression cloning of GABA(B) receptors uncovers similarity to metabotropic glutamate receptors.* Nature, 1997. **386**(6622): p. 239-46.
41. Billinton, A., et al., *Advances in the molecular understanding of GABA(B) receptors.* Trends Neurosci, 2001. **24**(5): p. 277-82.
42. Biermann, B., et al., *The Sushi domains of GABAB receptors function as axonal targeting signals.* J Neurosci, 2010. **30**(4): p. 1385-94.
43. Craig, M.T., et al., *Distinct roles of GABAB1a- and GABAB1b-containing GABAB receptors in spontaneous and evoked termination of persistent cortical activity.* J Physiol, 2013. **591**(4): p. 835-43.
44. Billinton, A., N. Upton, and N.G. Bowery, *GABA(B) receptor isoforms GBR1a and GBR1b, appear to be associated with pre- and post-synaptic elements respectively in rat and human cerebellum.* Br J Pharmacol, 1999. **126**(6): p. 1387-92.
45. Hawrot, E., et al., *Demonstration of a tandem pair of complement protein modules in GABA(B) receptor 1a.* FEBS Lett, 1998. **432**(3): p. 103-8.
46. O'Leary, O.F., et al., *GABAB(1) receptor subunit isoforms differentially regulate stress resilience.* Proc Natl Acad Sci U S A, 2014. **111**(42): p. 15232-7.
47. Ng, G.Y., et al., *Cloning of a novel G-protein-coupled receptor GPR 51 resembling GABAB receptors expressed predominantly in nervous tissues and mapped proximal to the hereditary sensory neuropathy type 1 locus on chromosome 9.* Genomics, 1999. **56**(3): p. 288-95.
48. Ng, G.Y., et al., *Identification of a GABAB receptor subunit, gb2, required for functional GABAB receptor activity.* J Biol Chem, 1999. **274**(12): p. 7607-10.
49. Kaupmann, K., et al., *GABA(B)-receptor subtypes assemble into functional heteromeric complexes.* Nature, 1998. **396**(6712): p. 683-7.
50. Jones, K.A., et al., *GABA(B) receptors function as a heteromeric assembly of the subunits GABA(B)R1 and GABA(B)R2.* Nature, 1998. **396**(6712): p. 674-9.
51. White, J.H., et al., *Heterodimerization is required for the formation of a functional GABA(B) receptor.* Nature, 1998. **396**(6712): p. 679-82.
52. Havlickova, M., et al., *The intracellular loops of the GB2 subunit are crucial for G-protein coupling of the heteromeric gamma-aminobutyrate B receptor.* Mol Pharmacol, 2002. **62**(2): p. 343-50.
53. Marshall, F.H., et al., *GABAB receptors - the first 7TM heterodimers.* Trends Pharmacol Sci, 1999. **20**(10): p. 396-9.
54. Galvez, T., et al., *Mutagenesis and modeling of the GABAB receptor extracellular domain support a venus flytrap mechanism for ligand binding.* J Biol Chem, 1999. **274**(19): p. 13362-9.
55. Geng, Y., et al., *Structural mechanism of ligand activation in human GABA(B) receptor.* Nature, 2013. **504**(7479): p. 254-9.
56. Couve, A., et al., *Intracellular retention of recombinant GABAB receptors.* J Biol Chem, 1998. **273**(41): p. 26361-7.
57. Robbins, M.J., et al., *GABA(B2) is essential for g-protein coupling of the GABA(B) receptor heterodimer.* J Neurosci, 2001. **21**(20): p. 8043-52.

58. Margeta-Mitrovic, M., Y.N. Jan, and L.Y. Jan, *Function of GB1 and GB2 subunits in G protein coupling of GABA(B) receptors.* Proc Natl Acad Sci U S A, 2001. **98**(25): p. 14649-54.
59. Benke, D. and H.U. Zeilhofer, *Divorce of obligatory partners in pain: disruption of GABA(B) receptor heterodimers in neuralgia.* EMBO J, 2012. **31**(15): p. 3234-6.
60. Pin, J.P., T. Galvez, and L. Prezeau, *Evolution, structure, and activation mechanism of family 3/C G-protein-coupled receptors.* Pharmacol Ther, 2003. **98**(3): p. 325-54.
61. Li, X., et al., *Human receptors for sweet and umami taste.* Proc Natl Acad Sci U S A, 2002. **99**(7): p. 4692-6.
62. Xu, H., et al., *Different functional roles of T1R subunits in the heteromeric taste receptors.* Proc Natl Acad Sci U S A, 2004. **101**(39): p. 14258-63.
63. Pandya, N.J., et al., *Group 1 metabotropic glutamate receptors 1 and 5 form a protein complex in mouse hippocampus and cortex.* Proteomics, 2016. **16**(20): p. 2698-2705.
64. Yin, S., et al., *Selective actions of novel allosteric modulators reveal functional heteromers of metabotropic glutamate receptors in the CNS.* J Neurosci, 2014. **34**(1): p. 79-94.
65. Pin, J.P. and B. Bettler, *Organization and functions of mGlu and GABAB receptor complexes.* Nature, 2016. **540**(7631): p. 60-68.
66. Comps-Agrar, L., et al., *The oligomeric state sets GABA(B) receptor signalling efficacy.* EMBO J, 2011. **30**(12): p. 2336-49.
67. Schwenk, J., et al., *Native GABA(B) receptors are heteromultimers with a family of auxiliary subunits.* Nature, 2010. **465**(7295): p. 231-5.
68. Calebiro, D., et al., *Single-molecule analysis of fluorescently labeled G-protein-coupled receptors reveals complexes with distinct dynamics and organization.* Proc Natl Acad Sci U S A, 2013. **110**(2): p. 743-8.
69. Jin, R., et al., *Crystal structure and association behaviour of the GluR2 amino-terminal domain.* EMBO J, 2009. **28**(12): p. 1812-23.
70. Sack, J.S., M.A. Saper, and F.A. Quiocho, *Periplasmic binding protein structure and function. Refined X-ray structures of the leucine/isoleucine/valine-binding protein and its complex with leucine.* J Mol Biol, 1989. **206**(1): p. 171-91.
71. Chun, L., W.H. Zhang, and J.F. Liu, *Structure and ligand recognition of class C GPCRs.* Acta Pharmacol Sin, 2012. **33**(3): p. 312-23.
72. Frangaj, A. and Q.R. Fan, *Structural biology of GABAB receptor.* Neuropharmacology, 2017.
73. Blein, S., et al., *Structural analysis of the complement control protein (CCP) modules of GABA(B) receptor 1a: only one of the two CCP modules is compactly folded.* J Biol Chem, 2004. **279**(46): p. 48292-306.
74. Geng, Y., et al., *Structure and functional interaction of the extracellular domain of human GABA(B) receptor GBR2.* Nat Neurosci, 2012. **15**(7): p. 970-8.
75. *The PyMOL Molecular Graphics System*. [1.2r3pre].
76. Liu, J., et al., *Molecular determinants involved in the allosteric control of agonist affinity in the GABAB receptor by the GABAB2 subunit.* J Biol Chem, 2004. **279**(16): p. 15824-30.

77. Rondard, P., et al., *Functioning of the dimeric GABA(B) receptor extracellular domain revealed by glycan wedge scanning.* EMBO J, 2008. **27**(9): p. 1321-32.
78. Burmakina, S., et al., *Heterodimeric coiled-coil interactions of human GABAB receptor.* Proc Natl Acad Sci U S A, 2014. **111**(19): p. 6958-63.
79. Kuner, R., et al., *Role of heteromer formation in GABAB receptor function.* Science, 1999. **283**(5398): p. 74-7.
80. Monnier, C., et al., *Trans-activation between 7TM domains: implication in heterodimeric GABAB receptor activation.* EMBO J, 2011. **30**(1): p. 32-42.
81. Margeta-Mitrovic, M., Y.N. Jan, and L.Y. Jan, *A trafficking checkpoint controls GABA(B) receptor heterodimerization.* Neuron, 2000. **27**(1): p. 97-106.
82. Pagano, A., et al., *C-terminal interaction is essential for surface trafficking but not for heteromeric assembly of GABA(b) receptors.* J Neurosci, 2001. **21**(4): p. 1189-202.
83. Lupas, A.N., J. Bassler, and S. Dunin-Horkawicz, *The Structure and Topology of alpha-Helical Coiled Coils.* Subcell Biochem, 2017. **82**: p. 95-129.
84. Crick, F.H.C., *The packing of alpha-helices: simple coiled-coils.* Acta Crystallogr D Biol Crystallogr, 1953(6): p. 689-697.
85. Froestl, W., *Chemistry and Pharmacology of GABAB Receptor Ligands.* 2010. **58**: p. 19-62.
86. Agabio, R. and G. Colombo, *[GABAB receptor as therapeutic target for drug addiction: from baclofen to positive allosteric modulators].* Psychiatr Pol, 2015. **49**(2): p. 215-23.
87. Cryan, J.F. and K. Kaupmann, *Don't worry 'B' happy!: a role for GABA(B) receptors in anxiety and depression.* Trends Pharmacol Sci, 2005. **26**(1): p. 36-43.
88. Urwyler, S., *Allosteric modulation of family C G-protein-coupled receptors: from molecular insights to therapeutic perspectives.* Pharmacol Rev, 2011. **63**(1): p. 59-126.
89. Pin, J.P. and L. Prezeau, *Allosteric modulators of GABA(B) receptors: mechanism of action and therapeutic perspective.* Curr Neuropharmacol, 2007. **5**(3): p. 195-201.
90. Bowery, N.G., et al., *International Union of Pharmacology. XXXIII. Mammalian gamma-aminobutyric acid(B) receptors: structure and function.* Pharmacol Rev, 2002. **54**(2): p. 247-64.
91. Szabadi, E., *GHB for cataplexy: Possible mode of action.* J Psychopharmacol, 2015. **29**(6): p. 744-9.
92. Boeckxstaens, G.E., et al., *Effects of lesogaberan on reflux and lower esophageal sphincter function in patients with gastroesophageal reflux disease.* Gastroenterology, 2010. **139**(2): p. 409-17.
93. Ong, J., D.I. Kerr, and G.A. Johnston, *Alfaxalone potentiates and mimics GABA-induced contractile responses in the guinea-pig isolated ileum.* Br J Pharmacol, 1988. **95**(1): p. 33-8.
94. Ong, J., D.I. Kerr, and A.R. Johnston, *Differing actions of beta-phenyl-GABA and baclofen in the guinea pig isolated ileum.* Neurosci Lett, 1987. **77**(1): p. 109-12.
95. Froestl, W., *Chemistry and pharmacology of GABAB receptor ligands.* Adv Pharmacol, 2010. **58**: p. 19-62.

96. Gjoni, T. and S. Urwyler, *Receptor activation involving positive allosteric modulation, unlike full agonism, does not result in GABAB receptor desensitization.* Neuropharmacology, 2008. **55**(8): p. 1293-9.
97. Urwyler, S., et al., *Mechanisms of allosteric modulation at GABAB receptors by CGP7930 and GS39783: effects on affinities and efficacies of orthosteric ligands with distinct intrinsic properties.* Neuropharmacology, 2005. **48**(3): p. 343-53.
98. Binet, V., et al., *The heptahelical domain of GABA(B2) is activated directly by CGP7930, a positive allosteric modulator of the GABA(B) receptor.* J Biol Chem, 2004. **279**(28): p. 29085-91.
99. Galvez, T., et al., *Ca(2+) requirement for high-affinity gamma-aminobutyric acid (GABA) binding at GABA(B) receptors: involvement of serine 269 of the GABA(B)R1 subunit.* Mol Pharmacol, 2000. **57**(3): p. 419-26.
100. Wise, A., et al., *Calcium sensing properties of the GABA(B) receptor.* Neuropharmacology, 1999. **38**(11): p. 1647-56.
101. Pumain, R., I. Kurcewicz, and J. Louvel, *Fast extracellular calcium transients: involvement in epileptic processes.* Science, 1983. **222**(4620): p. 177-9.
102. Lazarewicz, J.W., *Calcium transients in brain ischemia: role in neuronal injury.* Acta Neurobiol Exp (Wars), 1996. **56**(1): p. 299-311.
103. Heinemann, U., et al., *Extracellular calcium and potassium concentration changes in chronic epileptic brain tissue.* Adv Neurol, 1986. **44**: p. 641-61.
104. Froestl, W., et al., *Phosphinic acid analogues of GABA. 1. New potent and selective GABAB agonists.* J Med Chem, 1995. **38**(17): p. 3297-312.
105. Ong, J., et al., *R-(-)-beta-phenyl-GABA is a full agonist at GABAB receptors in brain slices but a partial agonist in the ileum.* Eur J Pharmacol, 1993. **233**(1): p. 169-72.
106. Lehmann, A., et al., *(R)-(3-amino-2-fluoropropyl) phosphinic acid (AZD3355), a novel GABAB receptor agonist, inhibits transient lower esophageal sphincter relaxation through a peripheral mode of action.* J Pharmacol Exp Ther, 2009. **331**(2): p. 504-12.
107. Mathivet, P., et al., *Binding characteristics of gamma-hydroxybutyric acid as a weak but selective GABAB receptor agonist.* Eur J Pharmacol, 1997. **321**(1): p. 67-75.
108. Frydenvang, K., et al., *GABAB antagonists: resolution, absolute stereochemistry, and pharmacology of (R)- and (S)-phaclofen.* Chirality, 1994. **6**(7): p. 583-9.
109. Kerr, D.I., et al., *2-Hydroxy-saclofen: an improved antagonist at central and peripheral GABAB receptors.* Neurosci Lett, 1988. **92**(1): p. 92-6.
110. Froestl, W., et al., *Phosphinic acid analogues of GABA. 2. Selective, orally active GABAB antagonists.* J Med Chem, 1995. **38**(17): p. 3313-31.
111. Bolser, D.C., et al., *The pharmacology of SCH 50911: a novel, orally-active GABA-beta receptor antagonist.* J Pharmacol Exp Ther, 1995. **274**(3): p. 1393-8.
112. Urwyler, S., et al., *Positive allosteric modulation of native and recombinant gamma-aminobutyric acid(B) receptors by 2,6-Di-tert-butyl-4-(3-hydroxy-2,2-dimethyl-propyl)-phenol (CGP7930) and its aldehyde analog CGP13501.* Mol Pharmacol, 2001. **60**(5): p. 963-71.

113. Malherbe, P., et al., *Characterization of (R,S)-5,7-di-tert-butyl-3-hydroxy-3-trifluoromethyl-3H-benzofuran-2-one as a positive allosteric modulator of GABAB receptors.* Br J Pharmacol, 2008. **154**(4): p. 797-811.
114. Urwyler, S., et al., *N,N'-Dicyclopentyl-2-methylsulfanyl-5-nitro-pyrimidine-4,6-diamine (GS39783) and structurally related compounds: novel allosteric enhancers of gamma-aminobutyric acidB receptor function.* J Pharmacol Exp Ther, 2003. **307**(1): p. 322-30.
115. Kunishima, N., et al., *Structural basis of glutamate recognition by a dimeric metabotropic glutamate receptor.* Nature, 2000. **407**(6807): p. 971-7.
116. Tsuchiya, D., et al., *Structural views of the ligand-binding cores of a metabotropic glutamate receptor complexed with an antagonist and both glutamate and Gd3+.* Proc Natl Acad Sci U S A, 2002. **99**(5): p. 2660-5.
117. Muto, T., et al., *Structures of the extracellular regions of the group II/III metabotropic glutamate receptors.* Proc Natl Acad Sci U S A, 2007. **104**(10): p. 3759-64.
118. Aricescu, A.R., W. Lu, and E.Y. Jones, *A time- and cost-efficient system for high-level protein production in mammalian cells.* Acta Crystallogr D Biol Crystallogr, 2006. **62**(Pt 10): p. 1243-50.
119. Weiss, J.N., *The Hill equation revisited: uses and misuses.* FASEB J, 1997. **11**(11): p. 835-41.
120. Scheres, S.H., *RELION: implementation of a Bayesian approach to cryo-EM structure determination.* J Struct Biol, 2012. **180**(3): p. 519-30.
121. Punjani, A., et al., *cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination.* Nat Methods, 2017. **14**(3): p. 290-296.
122. Zheng, S.Q., et al., *MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy.* Nat Methods, 2017. **14**(4): p. 331-332.
123. Zhang, K., *Gctf: Real-time CTF determination and correction.* J Struct Biol, 2016. **193**(1): p. 1-12.
124. Zhang, D.K. *Gautomatch*. Available from: http://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch/.
125. Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research and analysis.* J Comput Chem, 2004. **25**(13): p. 1605-12.
126. P.V. Afonine, J.J.H., T.C. Terwilliger & P.D. Adams, *Computational Crystallography Newsletter.* 4, 2013: p. 43-44.
127. Rohou, A. and N. Grigorieff, *CTFFIND4: Fast and accurate defocus estimation from electron micrographs.* J Struct Biol, 2015. **192**(2): p. 216-21.
128. Benke, D., K. Zemoura, and P.J. Maier, *Modulation of cell surface GABA(B) receptors by desensitization, trafficking and regulated degradation.* World J Biol Chem, 2012. **3**(4): p. 61-72.
129. Reeves, P.J., et al., *Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line.* Proc Natl Acad Sci U S A, 2002. **99**(21): p. 13419-24.
130. Singhal, A., et al., *Insights into congenital stationary night blindness based on the structure of G90D rhodopsin.* EMBO Rep, 2013. **14**(6): p. 520-6.
131. Miller, P.S. and A.R. Aricescu, *Crystal structure of a human GABAA receptor.* Nature, 2014. **512**(7514): p. 270-5.

132. Miller, P.S., et al., *Structural basis for GABAA receptor potentiation by neurosteroids.* Nat Struct Mol Biol, 2017. **24**(11): p. 986-992.
133. Byrne, E.F.X., et al., *Structural basis of Smoothened regulation by its extracellular domains.* Nature, 2016. **535**(7613): p. 517-522.
134. Engidawork, E., J. Aradska, and G. Lubec, *Neurotransmitter receptor complexes: methods for bioanalysis, their potentials and limitations.* Rev Neurosci, 2016. **27**(2): p. 111-33.
135. Deriu, D., et al., *Determination of the minimal functional ligand-binding domain of the GABAB1b receptor.* Biochem J, 2005. **386**(Pt 3): p. 423-31.
136. Chaudhary, S., et al., *Efficient expression screening of human membrane proteins in transiently transfected Human Embryonic Kidney 293S cells.* Methods, 2011. **55**(4): p. 273-80.
137. Liu, W., et al., *Structural basis for allosteric regulation of GPCRs by sodium ions.* Science, 2012. **337**(6091): p. 232-6.
138. Wacker, D., et al., *Structural features for functional selectivity at serotonin receptors.* Science, 2013. **340**(6132): p. 615-9.
139. Chun, E., et al., *Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors.* Structure, 2012. **20**(6): p. 967-76.
140. Rekas, A., et al., *Crystal structure of venus, a yellow fluorescent protein with improved maturation and reduced environmental sensitivity.* J Biol Chem, 2002. **277**(52): p. 50573-8.
141. Rath, A., et al., *Detergent binding explains anomalous SDS-PAGE migration of membrane proteins.* Proc Natl Acad Sci U S A, 2009. **106**(6): p. 1760-5.
142. Zemoura, K. and D. Benke, *Proteasomal degradation of gamma-aminobutyric acidB receptors is mediated by the interaction of the GABAB2 C terminus with the proteasomal ATPase Rtp6 and regulated by neuronal activity.* J Biol Chem, 2014. **289**(11): p. 7738-46.
143. *<Neurology-2012-Benarroch-GABAb review.pdf>.*
144. Zemoura, K., et al., *Endoplasmic reticulum-associated degradation controls cell surface expression of gamma-aminobutyric acid, type B receptors.* J Biol Chem, 2013. **288**(48): p. 34897-905.
145. Drew, D., et al., *Optimization of membrane protein overexpression and purification using GFP fusions.* Nat Methods, 2006. **3**(4): p. 303-13.
146. Gottlicher, M., et al., *Valproic acid defines a novel class of HDAC inhibitors inducing differentiation of transformed cells.* EMBO J, 2001. **20**(24): p. 6969-78.
147. Gossen, M., et al., *Transcriptional activation by tetracyclines in mammalian cells.* Science, 1995. **268**(5218): p. 1766-9.
148. Orth, P., et al., *Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system.* Nat Struct Biol, 2000. **7**(3): p. 215-9.
149. Aricescu, A.R. and R.J. Owens, *Expression of recombinant glycoproteins in mammalian cells: towards an integrative approach to structural biology.* Curr Opin Struct Biol, 2013. **23**(3): p. 345-56.
150. Gherardi, E., et al., *Structural basis of hepatocyte growth factor/scatter factor and MET signalling.* Proc Natl Acad Sci U S A, 2006. **103**(11): p. 4046-51.
151. Strege, A., et al., *Strategy for the Thermostabilization of an Agonist-Bound GPCR Coupled to a G Protein.* Methods Enzymol, 2017. **594**: p. 243-264.

152. Grisshammer, R. and C.G. Tate, *Overexpression of integral membrane proteins for structural studies.* Q Rev Biophys, 1995. **28**(3): p. 315-422.
153. Hassaine, G., et al., *X-ray structure of the mouse serotonin 5-HT3 receptor.* Nature, 2014. **512**(7514): p. 276-81.
154. Chelikani, P., et al., *The synthesis and high-level expression of a beta2-adrenergic receptor gene in a tetracycline-inducible stable mammalian cell line.* Protein Sci, 2006. **15**(6): p. 1433-40.
155. Fu, A.Y., et al., *A microfabricated fluorescence-activated cell sorter.* Nat Biotechnol, 1999. **17**(11): p. 1109-11.
156. Cherezov, V., et al., *High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor.* Science, 2007. **318**(5854): p. 1258-65.
157. Wu, H., et al., *Structure of the human kappa-opioid receptor in complex with JDTic.* Nature, 2012. **485**(7398): p. 327-32.
158. Rehan, S., V.O. Paavilainen, and V.P. Jaakola, *Functional reconstitution of human equilibrative nucleoside transporter-1 into styrene maleic acid co-polymer lipid particles.* Biochim Biophys Acta, 2017. **1859**(5): p. 1059-1065.
159. Takayama, H., et al., *High-level expression, single-step immunoaffinity purification and characterization of human tetraspanin membrane protein CD81.* PLoS One, 2008. **3**(6): p. e2314.
160. Standfuss, J., et al., *The structural basis of agonist-induced activation in constitutively active rhodopsin.* Nature, 2011. **471**(7340): p. 656-60.
161. Hansen, S.B., X. Tao, and R. MacKinnon, *Structural basis of PIP2 activation of the classical inward rectifier K+ channel Kir2.2.* Nature, 2011. **477**(7365): p. 495-8.
162. Molday, L.L. and R.S. Molday, *1D4: a versatile epitope tag for the purification and characterization of expressed membrane and soluble proteins.* Methods Mol Biol, 2014. **1177**: p. 1-15.
163. Keefe, A.D., et al., *One-step purification of recombinant proteins using a nanomolar-affinity streptavidin-binding peptide, the SBP-Tag.* Protein Expr Purif, 2001. **23**(3): p. 440-6.
164. Becher, A., J.H. White, and R.A. McIlhinney, *The gamma-aminobutyric acid receptor B, but not the metabotropic glutamate receptor type-1, associates with lipid rafts in the rat cerebellum.* J Neurochem, 2001. **79**(4): p. 787-95.
165. Hanson, M.A., et al., *A specific cholesterol binding site is established by the 2.8 A structure of the human beta2-adrenergic receptor.* Structure, 2008. **16**(6): p. 897-905.
166. Zocher, M., et al., *Cholesterol increases kinetic, energetic, and mechanical stability of the human beta2-adrenergic receptor.* Proc Natl Acad Sci U S A, 2012. **109**(50): p. E3463-72.
167. Bowery, N.G., D.R. Hill, and A.L. Hudson, *[3H](-)Baclofen: an improved ligand for GABAB sites.* Neuropharmacology, 1985. **24**(3): p. 207-10.
168. Hattori, M., R.E. Hibbs, and E. Gouaux, *A fluorescence-detection size-exclusion chromatography-based thermostability assay for membrane protein precrystallization screening.* Structure, 2012. **20**(8): p. 1293-9.

169. Postis, V., et al., *The use of SMALPs as a novel membrane protein scaffold for structure study by negative stain electron microscopy.* Biochim Biophys Acta, 2015. **1848**(2): p. 496-501.
170. Orwick, M.C., et al., *Detergent-free formation and physicochemical characterization of nanosized lipid-polymer complexes: Lipodisq.* Angew Chem Int Ed Engl, 2012. **51**(19): p. 4653-7.
171. Logez, C., et al., *Detergent-free Isolation of Functional G Protein-Coupled Receptors into Nanometric Lipid Particles.* Biochemistry, 2016. **55**(1): p. 38-48.
172. Charlton, J., *SOLUBILISATION AND CHARACTERISATION OF G- PROTEIN-COUPLED RECEPTORS USING STYRENE MALEIC ACID POLYMER*, in *School of Biosciences*. 2015, University of Birmingham.
173. Scheidelaar, S., et al., *Molecular model for the solubilization of membranes into nanodisks by styrene maleic Acid copolymers.* Biophys J, 2015. **108**(2): p. 279-90.
174. Kolter, T. and K. Sandhoff, *Principles of lysosomal membrane digestion: stimulation of sphingolipid degradation by sphingolipid activator proteins and anionic lysosomal lipids.* Annu Rev Cell Dev Biol, 2005. **21**: p. 81-103.
175. Ciaffoni, F., et al., *Saposin D solubilizes anionic phospholipid-containing membranes.* J Biol Chem, 2001. **276**(34): p. 31583-9.
176. Olsen, J.V., S.E. Ong, and M. Mann, *Trypsin cleaves exclusively C-terminal to arginine and lysine residues.* Mol Cell Proteomics, 2004. **3**(6): p. 608-14.
177. Frauenfeld, J., et al., *A saposin-lipoprotein nanoparticle system for membrane proteins.* Nat Methods, 2016. **13**(4): p. 345-51.
178. Flayhan, A., et al., *Saposin Lipid Nanoparticles: A Highly Versatile and Modular Tool for Membrane Protein Research.* Structure, 2018. **26**(2): p. 345-355 e5.
179. Puthenveetil, R. and O. Vinogradova, *Optimization of the design and preparation of nanoscale phospholipid bilayers for its application to solution NMR.* Proteins, 2013. **81**(7): p. 1222-31.
180. White, J.F., et al., *Structure of the agonist-bound neurotensin receptor.* Nature, 2012. **490**(7421): p. 508-13.
181. Locatelli-Hoops, S.C., et al., *Expression, surface immobilization, and characterization of functional recombinant cannabinoid receptor CB2.* Biochim Biophys Acta, 2013. **1834**(10): p. 2045-56.
182. Lee, S.C., et al., *A method for detergent-free isolation of membrane proteins in their local lipid environment.* Nat Protoc, 2016. **11**(7): p. 1149-62.
183. Parmar, M., et al., *Using a SMALP platform to determine a sub-nm single particle cryo-EM membrane protein structure.* Biochim Biophys Acta, 2018. **1860**(2): p. 378-383.
184. Hodges, R.S., et al., *Antigen-antibody interaction. Synthetic peptides define linear antigenic determinants recognized by monoclonal antibodies directed to the cytoplasmic carboxyl terminus of rhodopsin.* J Biol Chem, 1988. **263**(24): p. 11768-75.
185. Structural Genomics, C., et al., *Protein production and purification.* Nat Methods, 2008. **5**(2): p. 135-46.
186. Otzen, D.E., *Protein unfolding in detergents: effect of micelle structure, ionic strength, pH, and temperature.* Biophys J, 2002. **83**(4): p. 2219-30.

187. Shibata, Y., et al., *Thermostabilization of the neurotensin receptor NTS1.* J Mol Biol, 2009. **390**(2): p. 262-77.
188. Zhang, Y., et al., *Cryo-EM structure of the activated GLP-1 receptor in complex with a G protein.* Nature, 2017. **546**(7657): p. 248-253.
189. Garcia-Nafria, J., et al., *Cryo-EM structure of the adenosine A2A receptor coupled to an engineered heterotrimeric G protein.* Elife, 2018. **7**.
190. Efremov, R.G., C. Gatsogiannis, and S. Raunser, *Lipid Nanodiscs as a Tool for High-Resolution Structure Determination of Membrane Proteins by Single-Particle Cryo-EM.* Methods Enzymol, 2017. **594**: p. 1-30.
191. Popot, J.L., et al., *Amphipols from A to Z.* Annu Rev Biophys, 2011. **40**: p. 379-408.
192. Tribet, C., R. Audebert, and J.L. Popot, *Amphipols: polymers that keep membrane proteins soluble in aqueous solutions.* Proc Natl Acad Sci U S A, 1996. **93**(26): p. 15047-50.
193. Liao, M., et al., *Single particle electron cryo-microscopy of a mammalian ion channel.* Curr Opin Struct Biol, 2014. **27**: p. 1-7.
194. Huynh, K.W., et al., *Structure of the full-length TRPV2 channel by cryo-EM.* Nat Commun, 2016. **7**: p. 11130.
195. Lu, P., et al., *Three-dimensional structure of human gamma-secretase.* Nature, 2014. **512**(7513): p. 166-170.
196. Mazhab-Jafari, M.T., et al., *Atomic model for the membrane-embedded VO motor of a eukaryotic V-ATPase.* Nature, 2016. **539**(7627): p. 118-122.
197. Sanders, C.R., et al., *French swimwear for membrane proteins.* Chembiochem, 2004. **5**(4): p. 423-6.
198. Anatrace Products, L., *Amphipol A8-35: Amphipathic Polymer for Membrane Protein Studies*.
199. Ilca, S.L., et al., *Localized reconstruction of subunits from electron cryomicroscopy images of macromolecular complexes.* Nat Commun, 2015. **6**: p. 8843.
200. Nakane, T., et al., *Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION.* Elife, 2018. **7**.
201. Zoonens, M., et al., *Dynamics of membrane protein/amphipol association studied by Forster resonance energy transfer: implications for in vitro studies of amphipol-stabilized membrane proteins.* Biochemistry, 2007. **46**(36): p. 10392-404.
202. Baker, M.R., G. Fan, and Serysheva, II, *Single-particle cryo-EM of the ryanodine receptor channel in an aqueous environment.* Eur J Transl Myol, 2015. **25**(1): p. 35-48.
203. Efremov, R.G., et al., *Architecture and conformational switch mechanism of the ryanodine receptor.* Nature, 2015. **517**(7532): p. 39-43.
204. Palovcak, E., et al., *A simple and robust procedure for preparing graphene-oxide cryo-EM grids.* J Struct Biol, 2018. **204**(1): p. 80-84.
205. Kedrov, A., et al., *Structural Dynamics of the YidC:Ribosome Complex during Membrane Protein Biogenesis.* Cell Rep, 2016. **17**(11): p. 2943-2954.
206. Dore, A.S., et al., *Structure of class C GPCR metabotropic glutamate receptor 5 transmembrane domain.* Nature, 2014. **511**(7511): p. 557-62.

207. Magnani, F., et al., *Co-evolving stability and conformational homogeneity of the human adenosine A2a receptor.* Proc Natl Acad Sci U S A, 2008. **105**(31): p. 10744-9.
208. Phillips, R., et al., *Emerging roles for lipids in shaping membrane-protein function.* Nature, 2009. **459**(7245): p. 379-85.
209. Leitz, A.J., et al., *Functional reconstitution of Beta2-adrenergic receptors utilizing self-assembling Nanodisc technology.* Biotechniques, 2006. **40**(5): p. 601-2, 604, 606, passim.
210. Gao, Y., et al., *TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action.* Nature, 2016. **534**(7607): p. 347-51.
211. Shen, P.S., et al., *The Structure of the Polycystic Kidney Disease Channel PKD2 in Lipid Nanodiscs.* Cell, 2016. **167**(3): p. 763-773 e11.
212. Danev, R., D. Tegunov, and W. Baumeister, *Using the Volta phase plate with defocus for cryo-EM single particle analysis.* Elife, 2017. **6**.
213. Danev, R. and W. Baumeister, *Cryo-EM single particle analysis with the Volta phase plate.* Elife, 2016. **5**.
214. Liang, Y.L., et al., *Phase-plate cryo-EM structure of a biased agonist-bound human GLP-1 receptor-Gs complex.* Nature, 2018. **555**(7694): p. 121-125.
215. Lyons, J.A., et al., *Saposin-Lipoprotein Scaffolds for Structure Determination of Membrane Transporters.* Methods Enzymol, 2017. **594**: p. 85-99.
216. Ritchie, T.K., et al., *Chapter 11 - Reconstitution of membrane proteins in phospholipid bilayer nanodiscs.* Methods Enzymol, 2009. **464**: p. 211-31.
217. Bayburt, T.H. and S.G. Sligar, *Membrane protein assembly into Nanodiscs.* FEBS Lett, 2010. **584**(9): p. 1721-7.
218. Gatsogiannis, C., et al., *Membrane insertion of a Tc toxin in near-atomic detail.* Nat Struct Mol Biol, 2016. **23**(10): p. 884-890.
219. Grinkova, Y.V., I.G. Denisov, and S.G. Sligar, *Engineering extended membrane scaffold proteins for self-assembly of soluble nanoscale lipid bilayers.* Protein Eng Des Sel, 2010. **23**(11): p. 843-8.
220. Nehme, R., et al., *Mini-G proteins: Novel tools for studying GPCRs in their active conformation.* PLoS One, 2017. **12**(4): p. e0175642.
221. Maurel, D., et al., *Cell-surface protein-protein interaction analysis with time-resolved FRET and snap-tag technologies: application to GPCR oligomerization.* Nat Methods, 2008. **5**(6): p. 561-7.
222. Margeta-Mitrovic, M., Y.N. Jan, and L.Y. Jan, *Ligand-induced signal transduction within heterodimeric GABA(B) receptor.* Proc Natl Acad Sci U S A, 2001. **98**(25): p. 14643-8.
223. Rasmussen, S.G., et al., *Structure of a nanobody-stabilized active state of the beta(2) adrenoceptor.* Nature, 2011. **469**(7329): p. 175-80.
224. Miller-Gallacher, J.L., et al., *The 2.1 A resolution structure of cyanopindolol-bound beta1-adrenoceptor identifies an intramembrane Na+ ion that stabilises the ligand-free receptor.* PLoS One, 2014. **9**(3): p. e92727.
225. Edoardo D'Imprima, D.F., Mirko Joppe, Ricardo S´anchez, Martin and a.W.K.u. Grininger, *The deadly touch: protein denaturation at the water-air interface and how to prevent it.* 2018.

226. Ivankova, K., et al., *Up-regulation of GABA(B) receptor signaling by constitutive assembly with the K+ channel tetramerization domain-containing protein 12 (KCTD12).* J Biol Chem, 2013. **288**(34): p. 24848-56.
227. Turecek, R., et al., *Auxiliary GABAB receptor subunits uncouple G protein betagamma subunits from effector channels to induce desensitization.* Neuron, 2014. **82**(5): p. 1032-44.

Appendix A

Summary

Appendix A contains two journal papers published based on the results obtained during the first rotation (Oct, 2014- April, 2015) completed in Prof. Juha Huiskonen's Lab.

The first paper investigated the structure of the bacteriophage phi6. Phi6 is one of the most established models for studying segmented, double-stranded RNA(dsRNA) virus replication. The group of viruses possessing dsRNA genomes includes several important pathogens such as the rotavirus and bluetongue virus. In this study, structure of the virus protein capsids was solved to 3.9 Å using single particle analysis, revealing folding of the previously unknown protein P8. This allowed us to propose a model, i.e. *bona fide* domain swapping for the outer shell assembly. This led us to extend the theory of domain swapping from the level of monomeric subunits and multimers to closed spherical shells, and to hypothesize a mechanism by which closed protein shells may arise in evolution.

The second paper was on investigation of the conformational changes of the viral surface protein of the Lassa Virus under different environments. Lassa virus is a fatal, fever-causing human pathogen. This project aims to study the interaction between viral glycoproteins and the host cell receptor. My contribution in the project was to reconstruct a model of the glycoproteins at pH3 using Electron Cryo-Tomography. This structure, together with others solved under different conditions, showed specific structural rearrangements in the glycoprotein in various environments, providing a molecular-level rationale for this important stage of host cell entry.

nature COMMUNICATIONS

ARTICLE

Received 2 Nov 2016 | Accepted 26 Jan 2017 | Published 13 Mar 2017

DOI: 10.1038/ncomms14814 OPEN

Double-stranded RNA virus outer shell assembly by *bona fide* domain-swapping

Zhaoyang Sun[1,*], Kamel El Omari[1,*], Xiaoyu Sun[2], Serban L. Ilca[1], Abhay Kotecha[1], David I. Stuart[1], Minna M. Poranen[2] & Juha T. Huiskonen[1,2]

Correct outer protein shell assembly is a prerequisite for virion infectivity in many multi-shelled dsRNA viruses. In the prototypic dsRNA bacteriophage ϕ6, the assembly reaction is promoted by calcium ions but its biomechanics remain poorly understood. Here, we describe the near-atomic resolution structure of the ϕ6 double-shelled particle. The outer $T=13$ shell protein P8 consists of two alpha-helical domains joined by a linker, which allows the trimer to adopt either a closed or an open conformation. The trimers in an open conformation swap domains with each other. Our observations allow us to propose a mechanistic model for calcium concentration regulated outer shell assembly. Furthermore, the structure provides a prime exemplar of *bona fide* domain-swapping. This leads us to extend the theory of domain-swapping from the level of monomeric subunits and multimers to closed spherical shells, and to hypothesize a mechanism by which closed protein shells may arise in evolution.

[1] Division of Structural Biology, Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK. [2] Department of Biosciences, University of Helsinki, Viikinkaari 9, Helsinki 00014, Finland. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to M.M.P. (email: minna.poranen@helsinki.fi) or to J.T.H. (email: juha@strubi.ox.ac.uk).

The protein shells of icosahedral viruses assemble with remarkable accuracy from smaller subunits[1–3]. Accumulating evidence suggests that such viruses can be grouped into a fairly small number, perhaps less than a dozen, of viral lineages, each sharing a set of common assembly principles[4,5]. However, explaining how such closed, highly symmetric shells may have arisen during evolution poses a challenge, as formation of a closed shell first requires a network of complementary subunit-subunit interactions with roughly 50% of the surfaces buried from solvent[6]. One of the most complex examples of such interactions is the capsid of bacteriophage HK97 where the capsid proteins are intertwined to form a topologically linked 'chain mail' organization[7]. A simpler question to ask is how do multimeric proteins evolve from monomers? It has been suggested that this can be achieved by a mechanism called 'domain swapping'[8,9]. In the resulting multimer, the domain-swapped subunits in an 'open' conformation recapitulate the domain-domain interactions of the original 'closed' conformation of the monomer. This concept explains how mutations increasing the flexibility of a linker region of the monomer can facilitate domain swapping to give rise to the multimer.

Domain-swapping may play a role not just in the assembly and evolution of multimeric proteins, but also of viral shells[10–12]. The structure of rice yellow mottle virus (RYMV; genus *Sobemovirus*), and its comparison to the related southern cowpea mosaic virus (SCPMV), provides an interesting exemplar[12]. In SCPMV, the $T=3$ icosahedral capsid is formed in an usual way by trimers in a compact, closed conformation[13]. In RYMV, however, the trimers are atypical and adopt an open conformation, swapping domains with the neighbouring trimers and creating long range interactions, to which the increased stability of the capsid has been attributed[12]. However, it remains unclear whether RYMV trimers can exist in both open and closed conformations.

We use bacteriophage ϕ6, a member of the *Cystoviridae* family, as a model system to study the assembly of a complicated multi-shelled virus. The inner $T=1$ shell, composed of protein P1, forms a dodecahedral polymerase complex (PC) harbouring the three viral dsRNA segments (S, M, and L)[14,15], the viral polymerase P2 (refs 16,17), the hexameric packaging protein P4 (refs 18,19), and the assembly cofactor P7 (refs 20,21). Unlike eukaryotic dsRNA viruses, ϕ6 and other cystoviruses undergo a dramatic conformational change during RNA packaging and replication, from an empty collapsed PC to a fully expanded, packaged PC[22,23]. The expanded PC is covered by a layer of P8, forming the outer $T=13$ shell[23]. This double-shelled particle, or nucleocapsid (NC), is itself covered by a protein-lipid envelope during virion assembly[24,25]. A remarkable aspect of the ϕ6 system is that PC and NC particles can be assembled *in vitro* from purified protein and RNA components[26]. In addition to ϕ6 bacteriophage (*Cystovirus*), many other dsRNA viruses, such as aquareovirus (*Aquareovirus*)[27], mammalian reovirus (*Orthoreovirus*)[28], blue tongue virus (*Orbivirus*)[29,30], rice dwarf virus (*Phytoreovirus*)[31,32], rotavirus (*Rotavirus*)[33], and Banna virus (*Seadornavirus*)[34], form a multi-shelled particle with an inner $T=1$ and one or more outer $T=13$ shells. While structures for some of these have been solved to near-atomic resolution shedding light into molecular interactions in the $T=1$ and $T=13$ shells[27,29,30,33], assembly of the $T=13$ shell in ϕ6 NC has remained elusive.

Here we present the structure of ϕ6 NC to near-atomic resolution using electron cryomicroscopy and single particle analysis. The structure of the outer NC shell reveals an intricate domain-swapped architecture created by P8 trimers. The majority of the P8 subunits (540 of 600) are in an open conformation, swapping domains with their neighbours. The inter-trimer interactions recapitulate the intra-trimer interactions of the closed conformation that were also observed within the NC. These observations not only provide a possible explanation for the calcium-regulated assembly of the ϕ6 NC particle, but they also constitute an exemplar of *bona fide* domain swapping in a viral shell, allowing us to propose a mechanism for how such complex shells may have arisen during evolution.

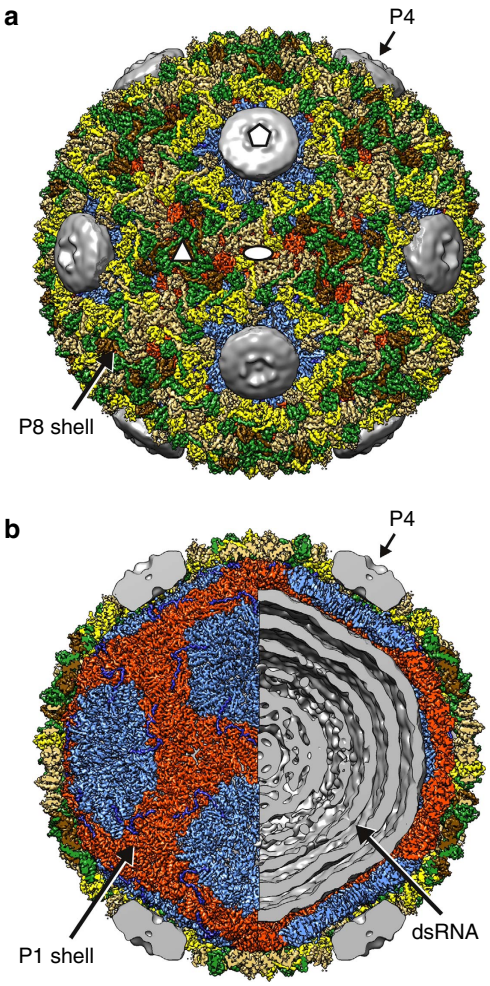


Figure 1 | Cryo-EM structure of bacteriophage ϕ6 nucleocapsid. (**a**,**b**) Three-dimensional reconstruction of the NC at 4.0-Å resolution. The P8 trimers forming the outer protein shell are coloured in yellow ($P8_Q$), green ($P8_R$), gold ($P8_S$) and brown ($P8_T$). Half of the P8 density was removed to reveal in **b** the inner protein shell composed of $P1_A$–$P1_B$ asymmetric dimers, coloured in light blue ($P1_A$) and red ($P1_B$). The elongated density corresponding to the P4 C-terminus is shown in dark blue. Residual density, not assigned to the aforementioned icosahedrally ordered components was filtered to 15-Å resolution to visualize in **a**,**b** the symmetry-mismatched P4 hexamers at the icosahedral five-fold vertices and the dsRNA in **b**. See also Supplementary Fig. 1.

Results

Near-atomic resolution structure of the ϕ6 nucleocapsid. We determined the structure of ϕ6 double-shelled particle by electron cryomicroscopy (cryo-EM; Fig. 1; Table 1). A high-resolution cryo-EM data set was acquired by imaging purified ϕ6 NC samples, some of which had spontaneously lost the RNA genome (Supplementary Fig. 1a). Two-dimensional image classification revealed the morphology of NC particles

Table 1 | Nucleocapsid cryo-EM data collection parameters.

Parameter	Value
Voltage (kV)	300
Magnification (×)	37,037
Defocus (μm)	0.3–3.0
Dose rate (e^- per pixel per s)	5
Frames	22
Frame length (s)	0.2
Total dose (e^- per Å^2)	16
Micrographs	903

cryo-EM, electron cryomicroscopy.

(Supplementary Fig. 1b). Some particles had lost either fully or partially the outer P8 shell (Supplementary Fig. 1a,b). Complete double-shelled particles were used to calculate a three-dimensional reconstruction using a 'gold-standard' refinement approach (see Methods; Table 1). The overall resolution of the icosahedrally symmetric areas of the reconstruction was 4.0 Å, as determined by Fourier shell correlation (FSC, threshold = 0.143; Supplementary Fig. 1c). In many areas the local resolution was approaching 3.0 Å, while some areas, especially solvent exposed loops and areas around the icosahedral five-fold vertices in the outer protein shell were resolved to a lower resolution (Supplementary Fig.1d–i). The resolution was sufficient to not only trace the course of the polypeptide chain but also to align the sequence to the structure and refine the structure, for nearly all of P1 and P8 residues (see Methods). The P4 hexamers located at the icosahedral five-fold axes of symmetry were not resolved due to the symmetry mismatch (Fig. 1). Density corresponding to the three linear dsRNA genome segments appeared as concentric layers spaced ~3 nm apart but the detailed organization of the segments could not be addressed using icosahedral reconstruction (Fig. 1b).

Architecture of the nucleocapsid outer shell. The NC structure revealed, consistent with our earlier observations[23], a total of 200 P8 trimers forming the outer protein shell, following $T = 13laevo$ quasi-symmetry (Fig. 1a). Four types of trimers ($P8_Q$, $P8_R$, $P8_S$ and $P8_T$) occupy the different positions in the shell: Q-type trimers around the icosahedral five-fold axis of symmetry, S-type trimers adjacent to the two-fold axis, T-type trimers on the three-fold axis and R-type trimers between the other trimers. Each of the 60 asymmetric units of the shell thus consists of 10 P8 chains, as the $P8_T$ trimers contribute only one chain to each asymmetric unit.

The $T = 13$ shell revealed a remarkable domain swapped organization of P8 trimers (Fig. 2). The relatively high resolution of the reconstruction allowed modelling the fold of the P8 structure (residues 3–147; Fig. 2a,b; Supplementary Fig. 2). The structure is α-helical and can be divided into two domains that are connected by an elongated linker (Fig. 2c,d): an N-terminal peripheral domain (PD; residues 1–83) is followed by a 10-residue long elongated loop (residue 84–93) and a C-terminal core domain (CD; residues 94–149). The loop ends in a proline residue (Pro93), which breaks the preceding helix α6 of the CD (Supplementary Fig. 2). The CDs intertwine together to form the core of the P8 trimer. The PDs, due to the elongated nature of the linker, are located a considerable distance (~50 Å) from the CD. Instead of interacting with the CD of the trimer they belong to, for the majority of the P8 molecules the PDs swap between the trimers, each PD interacting with the CDs and PDs of adjacent trimers via heterotypic interactions (Fig. 2e,f). Nine such structures are observed in the icosahedral asymmetric unit of the NC.

In addition to the 'open' conformation described above and adopted by 9 out of 10 of the P8 chains, a fundamentally different conformation is also found on the shell. This 'closed', more compact, conformation is observed in one of the three chains of the peripentonal $P8_Q$ trimer (Fig. 2g; Supplementary Fig. 2). The two conformations differ in degree of bending of the elongated linker that acts as a hinge. The change in relative position between the two conformations is ~75 Å and involves a rotation of 71 degrees (Fig. 2h). In the closed conformation, the PD is bound to its own CD via a homotypic interaction (Fig. 2i). This interaction is characterised by an interface that is analogous to the one created between the PD and a CD of an adjacent trimer in heterotypic interactions (Fig. 2j).

P4 hexamer is anchored to the vertices of the P1 shell. In the members of the subfamily *Sedoreovirinae* of the family *Reoviridae*, such as blue tongue virus (*Orbivirus*)[29] and rice dwarf virus (*Phytoreovirus*)[32], P-type trimers occupy the five-fold vertices. In contrast, in members of *Cystoviridae*, such as ϕ6 (ref. 23) and *Pinareovirinae*, such as aquareoviruses[27], the P-type trimers are absent leaving room for turret-like structures at the icosahedral five-fold vertices. In ϕ6, the vertices are occupied by the P4 packaging hexamer (Fig. 1). The presence of P4 hexamers at the five-folds provides also a rationale for the observed closed conformation of certain peripentonal P8 proteins (Supplementary Fig. 3); if the type $P8_Q$ monomers were all in the open conformation, significant clashes would occur between $P8_Q$ and P4 (Supplementary Fig. 3b).

The six-fold symmetry of the P4 hexamer and five-fold symmetry of the icosahedral vertex beneath creates a symmetry mismatch, so that the structure of P4 cannot be resolved using conventional icosahedral reconstruction[23]. To address this we applied the localized reconstruction method[17] to P4 hexamers. Local areas corresponding to P4 hexamers were extracted from the NC images and reconstructed as single particles with six-fold symmetry ('P4 hexamer reconstruction') or without any symmetry ('NC vertex reconstruction'). This allowed a low-resolution structure of the P4 hexamer to be determined *in situ* (Fig. 3). At a resolution of 9.1 Å (FSC = 0.143), fitting of a P4 crystal structure (PDB:4BLO)[19] into the density was unambiguous (Fig. 3a–c) and allowed us to orient the hexamer relative to the NC vertex (Fig. 3d,e).

A disorder prediction algorithm suggests that the C-terminal residues of P4 are likely to be disordered (Supplementary Fig. 4)[35]. Indeed, they are missing in the crystal structure and this disordered region has been proposed to attach the P4 to the underlying P1 shell[19]. However, in earlier low-resolution cryo-EM reconstructions, the P4 hexamers appeared to float on top of the five-fold vertexes of the P1 shell[23,36,37] and the connections between these symmetry-mismatched components have remained unresolved. Our localized reconstruction of P4 hexamers that correctly dealt with the symmetry mismatch revealed elongated densities extending from 3 of the 6 P4 C-terminal parts (Fig. 3f). We hypothesised that these may correspond to the C-terminal parts that are disordered in the crystal structure. Consistent with this hypothesis, in the high-resolution icosahedral NC reconstruction we observed sixty copies of elongated density that lacked any apparent secondary structure stretching ~115 Å on top of each of the asymmetric P1 dimers (Figs 1b and 4c–e). At the vertex distal part of this density, densities consistent with the side-chains 324-ProArgArg-326

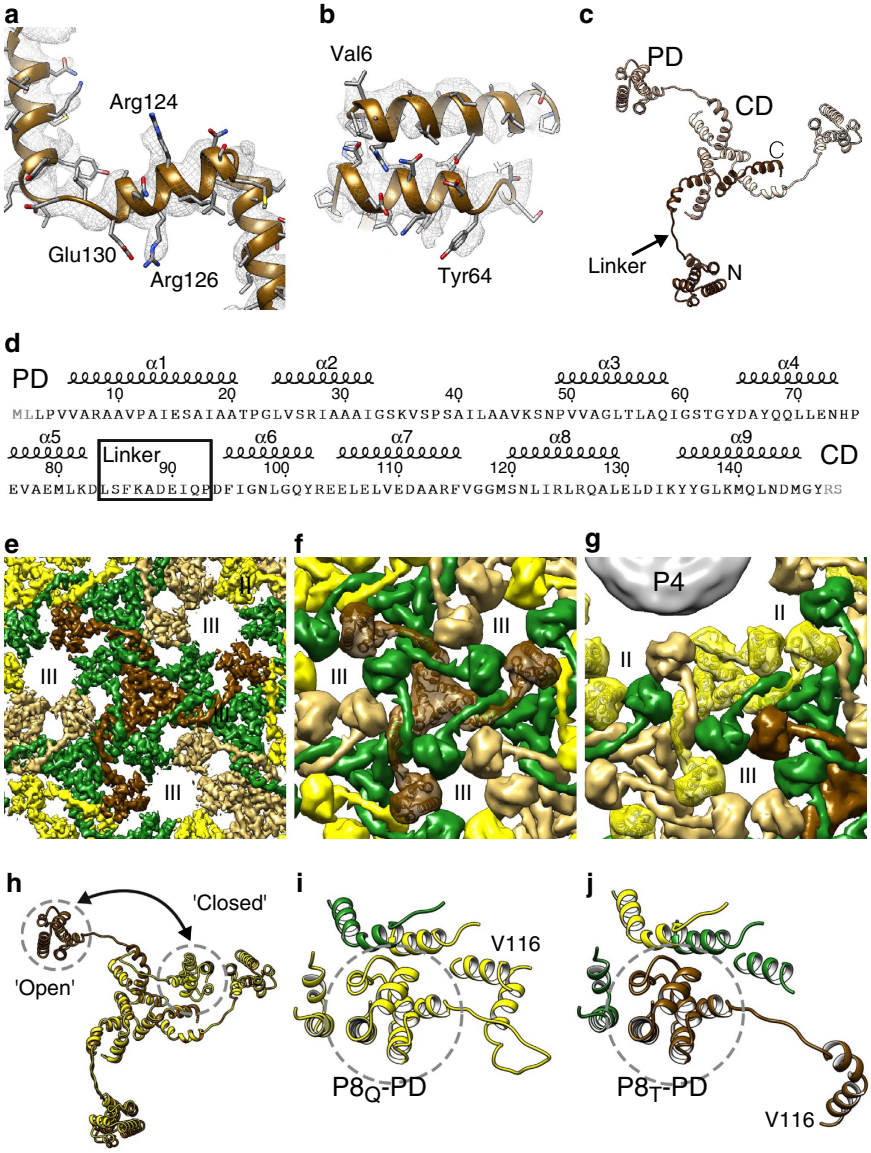


Figure 2 | Domain swapped architecture of the outer ϕ6 P8 shell. (**a**,**b**) Close-ups of the atomic model of the T-type trimer ($P8_T$) and corresponding density (mesh). Some side chains are labelled. (**c**) Atomic model of the P8 trimer. The three different chains are shown in different shades. The peripheral domain (PD) and the core domain (CD) are labelled and the corresponding areas of the structure circled. The linker (arrow), N terminus (N) and C terminus (C) are labelled for one chain (dark brown). (**d**) Sequence of the P8 with secondary structure elements, PD, CD and the linker region labelled. Amino-acid residues not resolved in the structure are in grey. (**e**) A close-up of the P8 shell density viewed along the three-fold axis of symmetry. The linkers from different trimers intertwine, leading to domain swapping of the peripheral domains between the trimers. (**f**) Same view as in **e** but a rough molecular surface is shown for clarity. $P8_T$ is shown as a ribbon. (**g**) Same as **f** but a view showing the Q-type trimers ($P8_Q$, ribbon) around the P4 hexamers. P4 is shown as a grey surface. (**h**) Structural alignment of $P8_T$ and $P8_Q$ trimers. The location of the peripheral domain in 'open' and 'closed' conformation of the trimer is indicated. (**i**) The mode of interaction of $P8_Q$-PD in closed conformation is shown. The interactions are mainly of homotypic nature with other parts of the same trimer. (**j**) The mode of interaction of the $P8_T$-PD in open conformation is shown. The interactions are of heterotypic nature with other types of trimers. The P8 trimers are coloured in yellow ($P8_Q$), green ($P8_R$), gold ($P8_S$) and brown ($P8_T$) and the type II and type III holes[23] are labelled. See also Supplementary Figs 2 and 3.

of the P4 C terminus were identified (Fig. 4f), allowing us to specifically assign this density to the C-terminal residues 292–332 of P4. The sixth, unbound C-terminal region of P4 was not resolved in any of the reconstructions, and is likely to be either disordered or unresolved in our maps due to limited resolution of the localized reconstructions.

Binding of up to five of the six flexible C-terminal tails of P4 provides an elegant solution to the symmetry mismatch. These tails are likely to be disordered in solution (Supplementary Fig. 4) and to order only upon binding to P1 during assembly. The P4 tails are analogous to the 'stay-cables' observed earlier in adenovirus fibres[38]. As such cables have now been observed in at least two unrelated viruses, adenovirus (a dsDNA virus)[38] and ϕ6 (a dsRNA virus, this study), they may be a general solution to the problem of attaching fibres and other symmetry-mismatched appendices to vertices of icosahedral virus capsids.

P4 C-terminal tail solidifies the dimeric P1 building block. The inner P1 shell is composed of 60 copies of asymmetric P1 dimers,

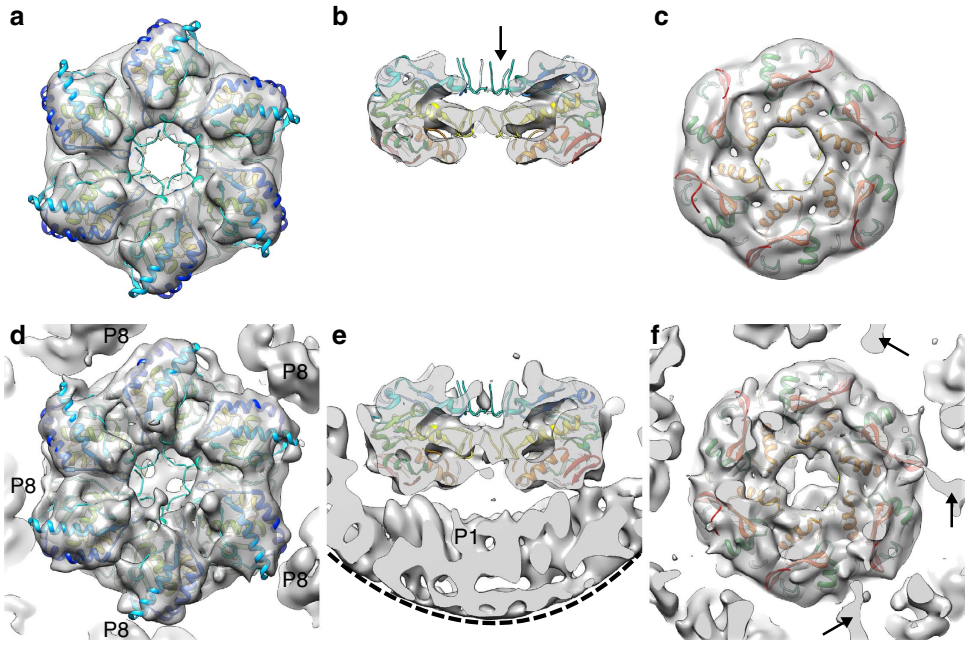


Figure 3 | Hexameric packaging ATPase P4. (**a**–**c**) Localized reconstruction of ϕ6 P4 with six-fold symmetry is shown as a grey transparent surface from the top (**a**), side (**b**) and below (**c**). P4 crystallographic structure[19] is coloured from red (C-terminus) to blue (N-terminus) and was fitted into the reconstruction as a rigid body. Flexible loops that were only partially resolved in the crystallographic structure were not resolved in the localized reconstruction (arrow). (**d**–**f**) Same views as in **a**–**c** showing the localized reconstruction without symmetry. The edge of a mask cutting through the density under the P4 hexamer is indicated with a dashed line in **e**. Three densities possibly connecting the P1 shell to the P4 C-termini are indicated with arrows.

formed by $P1_A$ and $P1_B$ (see Fig. 1b)[23]. A large-scale expansion of the P1 shell occurs during RNA packaging to empty particles and subsequent RNA replication[22,23]. To model the structural transitions in the P1 shell during this expansion, we refined the existing atomic coordinates of the P1 monomer (PDB:4K7H)[39] into our EM density map of the NC (Fig. 4a,b) and compared them to P1 coordinates flexibly fitted to an unexpanded PC reconstruction (PDB:4BTG)[39]. The angle between the two subunits in the $P1_A$–$P1_B$ asymmetric dimer changes relatively little (11°) showing that the $P1_A$–$P1_B$ dimer behaves as quite a rigid block during the expansion (Supplementary Fig. 5a,b; Supplementary Movie 1). In contrast to these intradimer angles, the interdimer angles change significantly, mainly at the $P1_B$–$P1_B$ interfaces[39]. The interdimer angle changes 57° at the two-fold interface and 49° at the three-fold interface (Supplementary Fig. 5c,d). The root-mean-square deviation (RMSD) calculated between the two states was 2.6 and 6.0 Å for the $P1_A$ and $P1_B$ C-alpha coordinates, respectively. This shows that relatively small changes occur within the $P1_A$ subunits, while $P1_B$ subunits adjust to their local environment in the two very different P1 shell expansion states.

The relative rigidity of the $P1_A$–$P1_B$ dimer interface is now explained by our observation that the C-terminal part of P4 bridges the $P1_A$ and $P1_B$ monomers (Fig. 4c,d) whilst still allowing interdimer angles to change. Thus, the P4 C-terminus seems not only to link P4 hexamers to the P1 shell, but also to play a role in P1 dimerization. A dimer of the major inner capsid protein could indeed be a conserved intermediate in the assembly pathways of dsRNA viruses. This notion is supported by the structure of the *Penicillium chrysogenum* virus in which the equivalent building block is formed by a covalently linked major capsid protein dimer, produced by gene duplication[40]. In ϕ6, we propose that the C-terminal tail of P4 provides an analogous linkage.

Discussion

When combined with earlier structural and biochemical results, our NC structure allows us to propose an assembly model for ϕ6 (Fig. 5; Supplementary Movie 2). The assembly of the P1 shell is initiated *in vitro* by P4 hexamers[26] and the C-terminus of P4 is essential for the formation of recombinant PC *in vivo*[41]. We now have a structural rationale for these observations, since the P4 C-terminus cements the P1 asymmetric dimer together (Fig. 4c,d). Accordingly, we suggest that the ϕ6 PC assembly starts with the formation of a $P1_A$–$P1_B$ dimer stabilized by the C-terminus of P4 (Fig. 5a). As observed in the P1 crystal structure, individual P1 chains tend to create a pentamer in a conformation that resembles a group of five $P1_A$ chains in the unexpanded conformation of PC[39]. Our modelling experiments also showed that six $P1_A$–$P1_B$ dimers per P4 would create steric clashes (Supplementary Movie 2). It is thus likely that each P4 grabs five and not more $P1_A$–$P1_B$ dimers. Dimers attached to distinct P4 hexamers then interact to eventually form a closed P1-shell (Fig. 5b). The assembly of the compact empty PC is facilitated by the assembly cofactor P7 (refs 20,26). P7, together with P2, is located at the three-fold symmetry axis of the empty PC[21], but the order of P2 and P7 binding to the PC remains unclear. The PC subsequently packages and replicates the RNA genome segments leading to expansion of the PC (Fig. 5c)[23,39] and finally formation of the double-shelled particles by incorporation of the P8 layer (Fig. 5d).

P8 shells form spontaneously in the presence of 0.1–1.0 mM Ca^{2+} and in the absence of any other structural proteins[42]. Conversely, depleting calcium by a chelating agent or exposure to low pH leads to P8 shell disassembly[42,43]. Raman spectroscopy has revealed that the assembled shell stabilizes the mostly α-helical conformation of P8 (ref. 44). It is thus clear that Ca^{2+} regulates the conformational state of P8. In the light of our observation of a closed and open state of the P8 trimer in

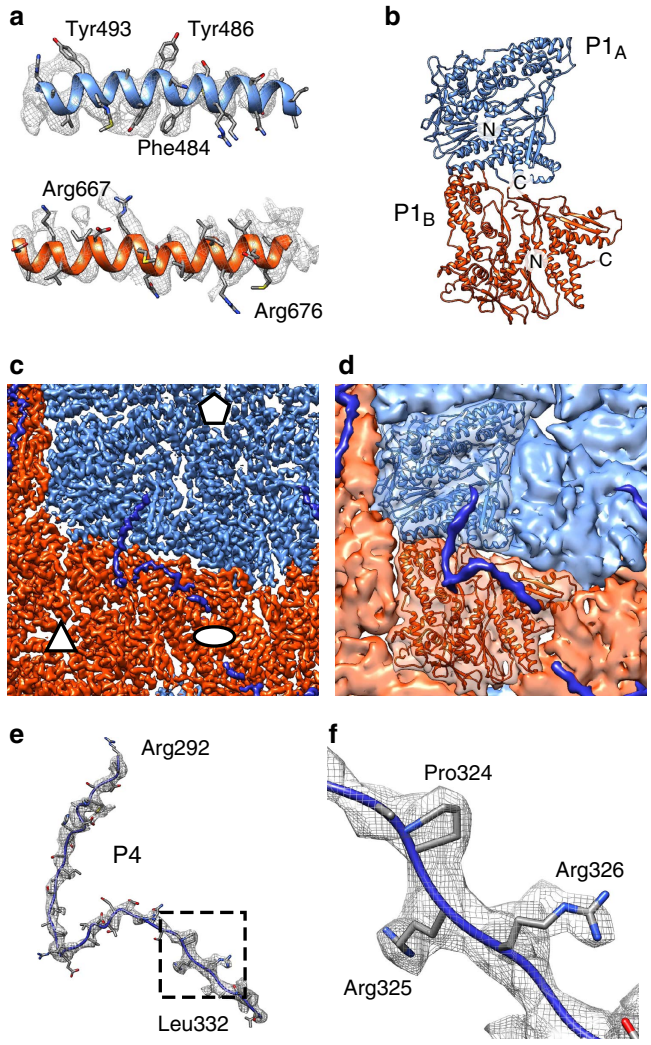


Figure 4 | Structure of the inner P1 shell. (**a**) Close-ups of the atomic models for the ϕ6 $P1_A$ (top) and $P1_B$ (bottom) subunits and corresponding density (mesh). (**b**) Atomic model of the asymmetric P1 dimer. The two subunits and the N- and C-termini are labelled. (**c**) Close-up of the density map. Icosahedral symmetry axes are labelled with an ellipse (two-fold), a triangle (three-fold) and a pentagon (five-fold). (**d**) Same view as in **c** but a rough molecular surface is shown for clarity and one P1 dimer is shown as a ribbon. $P1_A$ is blue, $P1_B$ is red, and density assigned to P4 C-terminal part is dark blue. (**e**) Atomic model for the P4 C-terminal part (R292–L332) and the corresponding density (mesh) are shown. (**f**) Close-up of the area indicated in **e** showing the 324-ProArgArg-326 motif that allowed assignment of this density to P4. See also Supplementary Fig. 4 and Supplementary Movie 1.

the NC, we hypothesize that Ca^{2+} induces transition of the P8 trimer from the closed to the open conformation and this further facilitates outer shell assembly by domain swapping (Fig. 5d,e). Due to steric clashes, the P8 monomers closest to the P4 hexamers are forced to remain in closed conformation (Supplementary Fig. 3). Ca^{2+} regulation of P8 conformation, possibly in addition to acidic pH, might also be important to facilitate the delivery of the PC to the host cell upon infection.

The conformational space sampled by P8 trimers in solution and at different Ca^{2+} concentrations remains unknown. Insolubility of purified P8 and its tendency to readily form large assemblies even in the presence of Ca^{2+}-chelating agents has impeded its solution structure studies by small angle X-ray scattering (Z.S. and J.T.H., unpublished observations). The P8 CD domain is preceded by Pro93 that breaks the helix α6 of the CD. Proline residues are abundant in the linker regions of domain swapped proteins and may play a role in modulating the equilibrium between an open and a closed state[45]. It is thus possible that Pro93 has a functional role in adding strain to the backbone of the linker region but this hypothesis remains to be tested.

The observed P8 shell assembly allows us to extend the concept of domain swapping to viral shells[8–12]. Domain swapping is considered *bona fide*, when the same subunit can still adopt both closed and open conformations, as opposed to the subunits adopting just the open conformation in a multimer and the closed conformation in a different, albeit orthologous, monomeric protein[8]. Following this nomenclature, our ϕ6 outer shell structure provides an exemplar of *bona fide* domain swapping between P8 trimers that exhibit both closed and open conformations.

Analogously to monomer to multimer transition during the evolution of multimeric proteins, our results allow us to propose a hypothetical model how closed viral protein shells may further evolve from such proteins. A primordial, relatively simple capsomer may have some tendency to form larger assemblies, but initially the protein-protein interfaces required to make a spherical shell are not optimized. Mutations in a linker region can increase its flexibility to allow for the formation of an open conformation of the multimeric capsomer. These capsomers in their open conformations would then be able to further link together by recapitulating the domain-domain interactions of the closed conformation. A pre-existing inner protein shell, genome, or internal lipid bilayer present in some viruses, may play a role in increasing the fidelity of shell assembly[46]. This is also evident in the case of ϕ6, where in the absence of the inner P1 shell, P8 self-assembles *in vitro* into aberrant, shell-like structures that vary in size and are not topologically closed[43,47]. In some viruses, the capsomers may have subsequently lost the ability to form both open and closed conformations. To what extent the outer protein shells of ϕ6 and other dsRNA viruses are homologous remains an open question, as domain-swapping is not observed in other dsRNA viruses. It is conceivable, however, that linker region deletions and domain-shuffling may have played a further role in stabilizing protein shells. In the case of ϕ6, the ability of P8 to still adopt both conformations may have been preserved in evolution as this may play a role in calcium-regulated outer shell assembly/disassembly and also the presence of P4 hexamer at the vertices necessitates a closed conformation of P8.

In conclusion, our structural characterization of the ϕ6 double-shelled particle revealed interactions of the inner P1 shell and symmetry-mismatched P4 packaging hexamers. Most importantly, the study led to the discovery of two states of the outer shell forming protein P8, 'open' and 'closed'. Trimers in the 'open' conformation swap domains between each other. This process is possibly regulated by calcium ion concentration but further studies are required to reveal the exact mechanism of Ca^{2+} regulation. The derived outer shell assembly model not only provides insights into assembly in this particular model system, but also suggests how closed protein shells may have arisen during evolution. Further studies are required to assess to what extent domain swapping may have played a role in other viral systems.

Methods

Cryo-EM sample preparation. Bacteriophage ϕ6 was grown on its host *Pseudomonas syringae pv.phaseolicola* HB10Y[48] and purified as described previously[49]. The viral envelope was removed by Triton X-114 extraction[49] and the resulting aqueous mixture containing the NC loaded onto a CIM DEAE-1

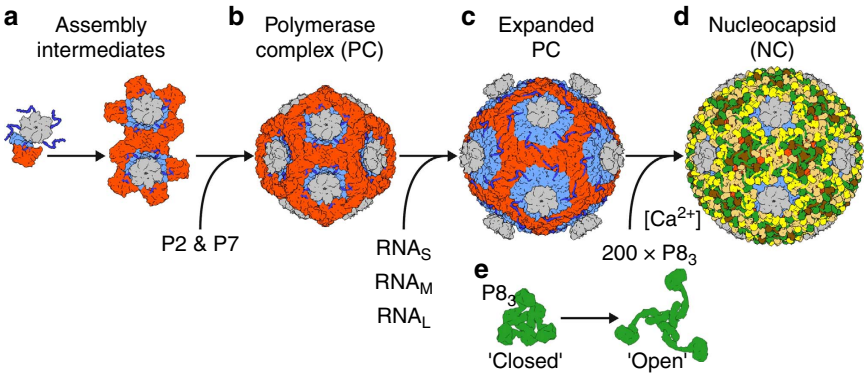


Figure 5 | Assembly model. (**a**) The ϕ6 inner capsid is composed of protein P1, P2, P4 and P7. Preassembled P4 hexamers (grey) nucleate the assembly of the P1 shell by bringing together two P1 subunits that will form the asymmetric $P1_A$–$P1_B$ dimer (light blue and red) in the final assembly. The dimerization of P1 is facilitated by the C-terminal tail of P4 (dark blue). Up to five P1 dimers can assemble around a P4 hexamer. (**b**) P1 dimers connected to distinct P4 hexamers interact bringing together the precursors of the icosahedral vertices. This is facilitated by protein P7, which acts as an assembly cofactor. (**c**) The self-assembled empty capsids package and replicate the single-stranded RNA precursors of the ϕ6 dsRNA genome (S, M and L) and the capsid adopts the expanded conformation. (**d**) Trimers of P8 (green, yellow, gold and brown) then assemble around the dsRNA containing inner capsid to form the nucleocapsid. (**e**) The assembly of the P8 shell is induced by Ca^{2+} ions and involves major conformational changes in the linker regions of P8 subunits to mediate the transition from the closed to the open conformation. See also Supplementary Movie 2.

Table 2 | Data processing parameters.

Parameter	NC particle	P4 hexamer	NC vertex
Particles*	13,291 (16,466)	56,448 (159,492)	56,448 (159,492)
Box size (pixels)	512	128	128
Pixel size (Å)	1.35	1.35	1.35
Symmetry	I1	C6	C1
Resolution (Å)	4.0	9.1	9.1
B-factor (Å^2)†	−177	−600‡	−600‡

NC, nucleocapsid.
*Number of particles (or subparticles in the case of P4 hexamer and NC vertex) chosen for averaging the final map is given. The number of all particles extracted from the micrographs (or all subparticles extracted from particles) is given in parenthesis.
†Inverse B-factor used for sharpening of the map.
‡Value of B-factor determined *ad hoc*.

tube monolithic column (BIA Separations, Slovenia) in 20 mM potassium phosphate pH 7.2, 1 mM $MgCl_2$, 0.1 mM $CaCl_2$, 150 mM NaCl from which the bound NC particles were eluted using a 0.15–2 M NaCl complex gradient (in 20 mM potassium phosphate pH 7.2, 1 mM $MgCl_2$, 0.1 mM $CaCl_2$). The NC containing fractions were concentrated using an ultra-centrifugal filter with a 100-kDa cutoff (Amicon; EMD Millipore, Billerica, MA, USA). A 3-µl aliquot of sample, diluted in buffer (20 mM KPO_4, pH 7.2, 1 mM $MgCl_2$, 0.1 mM $CaCl_2$, 150 mM NaCl) to concentration of 3 mg ml^{-1}, was applied to a glow-discharged grid (C-flat; Protochips, Raleigh, NC) and vitrified by plunge-freezing into liquid ethane using a Vitrobot (FEI, Hilsboro, OR).

Data acquisition and processing. Data were acquired using a 300-kV transmission electron microscope (Tecnai F30 'Polara'; FEI) equipped with an energy filter (slit width 20 eV; GIF Quantum LS, Gatan, Pleasanton, CA) and a direct electron detector (K2 Summit, Gatan). Movies (22 frames, each frame 0.2 s) were collected at in electron counting mode at dose rate of 5 e^- per pixel per s at calibrated magnification of 37,037 × resulting in a total dose of ∼16 e^- per Å^2 and pixel size of 1.35 Å. Data acquisition statistics are summarized in Table 1.

Movie frames were aligned to account for drift[50] and contrast transfer function (CTF) parameters were estimated[51]. A total of ∼900 movies, showing minimal drift and astigmatism were used to pick 16,466 particles. Particles were extracted for 2D reference free classification in Relion 1.3 (ref. 52). Particles from classes showing complete P1 and P8 layers were selected for 3D classification using a previously published NC structure (EMD-1,206) as an initial model. A subset of 13,291 particles extracted from movies was refined using standard refinement and particle polishing protocols implemented in Relion.

The structure of the P4 hexamer was calculated using localized reconstruction (www.opic.ox.ac.uk/localrec)[17]. Briefly, twelve 'subparticles', each corresponding to a projection of the P4 hexamer in each of the particle images, were extracted (total of 159,463 subparticles) to a smaller box and treated as independent single particles. To assign the orientation for each subparticle, one of the five possible triplets of Euler angles that were equivalent for the corresponding five-fold symmetric vertex was chosen randomly so as not to bias orientations in further processing. Two types of reconstructions were calculated. For the first reconstruction ('P4 hexamer'), partial signal subtraction was used to remove the contribution of unwanted NC density in the particle images[53]. We modified our earlier localized reconstruction method[17] to allow subtracting 'all-but-one' P4 hexamer density (available from http://github.com/OPIC-Oxford/localrec), in addition to the other unwanted densities (P1 and P8 shells and RNA) from the icosahedral reconstruction. Only subparticles from the edge of the particle (total of 103,482; maximum 40° deviation from a side view) were chosen for further analysis to reduce the overlap of the projected hexamer density with the RNA component of the virion that could not be accurately subtracted from the particle images. The structure of the P4 hexamer was then determined using Relion and 3D classification, using a localized reconstruction of the hexamer, calculated from a subset of subparticles and filtered to 40 Å resolution, as a starting model. 3D classification revealed the hexamer in different orientations relative to the five-fold vertex. Particles from classes showing clear hexamer density were combined and six-fold symmetry was applied on the reconstruction. For the second reconstruction ('NC vertex'), subparticles extracted from the original particle images and orientation parameters from the P4 hexamer reconstruction were used and no symmetry was applied.

Resolutions of the icosahedral NC reconstruction and P4 hexamer localized reconstruction were estimated by using FSC using 0.143 cutoff. The NC vertex reconstruction was low-pass filtered to the resolution of P4 hexamer reconstruction (9.1 Å) and both localized reconstructions were sharpened by applying an inverse B-factor of −600 Å^2 (determined by trial and error). Local resolution of the NC reconstruction was estimated using ResMap[54]. Reconstruction statistics are summarized in Table 2.

Atomic model building and refinement. As no P8 atomic model was available, a polyalanine model was built *de novo* using Coot. Visually identified densities

Table 3 | Atomic model refinement statistics.

Parameter	P1	P8
Refinement		
Resolution (Å)	4.0	
Map CC	0.715	
RMS deviations		
Bonds (Å)	0.09	
Angles (°)	1.0	
Molprobity validation		
Clashscore, all atoms (percentile)	13.8 (55th)	10.4 (70th)
Molprobity score	1.92 (80th)	1.53 (94th)
Favoured rotamers (poor) (%)	98.4 (0.1)	97.3 (0.0)
Ramachandran (% favoured)	96	99.0
Ramachandran (% allowed)	99.7	100
Ramachandran Plot (% outliers)	0.3	0

RMS, root mean square.

for bulky residues facilitated building the full atomic model of P8 (L3-Y147) and the C terminus of P4 (R292-L332) with side-chains. The crystal structure of ϕ6 P1 (PDB:4K7H) was fitted in the NC map using COOT[55] as a rigid body in two different positions corresponding to subunits $P1_A$ and $P1_B$. $P1_A$ and $P1_B$ main-chains and side-chains were adjusted using manual and real space fitting in COOT. The crystal structure of ϕ6 P4 (PDB:4BLO)[19] was fitted in the P4 hexamer and NC vertex localized reconstructions as a rigid body in UCSF Chimera[56].

Atomic models for all of the chains in the asymmetric unit ($P1_A$, $P1_B$, P4 C terminus and ten copies of P8) were refined in Phenix.real_space_refine applying secondary structure, rotamer, and Ramachandran plot restraints. In addition, non-crystallographic symmetry restraints were applied on P8 chains[57]. The models were validated with MOLPROBITY[58]. Atomic model refinement statistics are summarized in Table 3. Disordered regions in P4 sequence (UniProt:P11125) were predicted with RONN[35] and changes in subunit-subunit angles calculated in UCSF Chimera[59].

Data availability. Density maps and atomic models that support the findings of this study have been deposited in the Electron Microscopy Database and in the Protein Databank with the accession codes EMD-3571 (NC reconstruction), PDB 5MUU (NC atomic model), EMD-3572 (P4 hexamer reconstruction), PDB 5MUV (P4 atomic model fitted in the P4 hexamer reconstruction), EMD-3573 (NC vertex reconstruction), and PDB 5MUW (P4 atomic model fitted in the NC vertex reconstruction).

References

1. Aksyuk, A. A. & Rossmann, M. G. Bacteriophage assembly. *Viruses* **3,** 172–203 (2011).
2. Zlotnick, A., Porterfield, J. Z. & Wang, J. C.-Y. To build a virus on a nucleic acid substrate. *Biophys. J.* **104,** 1595–1604 (2013).
3. Perlmutter, J. D. & Hagan, M. F. Mechanisms of virus assembly. *Annu. Rev. Phys. Chem.* **66,** 217–239 (2015).
4. Bamford, D. H. Do viruses form lineages across different domains of life? *Res. Microbiol.* **154,** 231–236 (2003).
5. Abrescia, N. G. A., Bamford, D. H., Grimes, J. M. & Stuart, D. I. Structure unifies the viral universe. *Annu. Rev. Biochem.* **81,** 795–822 (2012).
6. Shepherd, C. M. & Reddy, V. S. Extent of protein-protein interactions and quasi-equivalence in viral capsids. *Proteins* **58,** 472–477 (2005).
7. Wikoff, W. R. *et al.* Topologically linked protein rings in the bacteriophage HK97 capsid. *Science* **289,** 2129–2133 (2000).
8. Bennett, M. J., Schlunegger, M. P. & Eisenberg, D. 3D domain swapping: a mechanism for oligomer assembly. *Protein Sci.* **4,** 2455–2468 (1995).
9. Liu, Y. & Eisenberg, D. 3D domain swapping: as domains continue to swap. *Protein Sci.* **11,** 1285–1299 (2002).
10. Simpson, A. A., Chipman, P. R., Baker, T. S., Tijssen, P. & Rossmann, M. G. The structure of an insect parvovirus (*Galleria mellonella* densovirus) at 3.7 Å resolution. *Structure* **6,** 1355–1367 (1998).
11. Tate, J. *et al.* The crystal structure of cricket paralysis virus: the first view of a new virus family. *Nat. Struct. Biol* **6,** 765–774 (1999).
12. Qu, C. *et al.* 3D domain swapping modulates the stability of members of an icosahedral virus group. *Structure* **8,** 1095–1103 (2000).
13. Silva, A. M. & Rossmann, M. G. Refined structure of southern bean mosaic virus at 2.9 Å resolution. *J. Mol. Biol.* **197,** 69–87 (1987).
14. Semancik, J. S., Vidaver, A. K. & Van Etten, J. L. Characterization of segmented double-helical RNA from bacteriophage ϕ6. *J. Mol. Biol.* **78,** 617–625 (1973).
15. Butcher, S. J., Dokland, T., Ojala, P. M., Bamford, D. H. & Fuller, S. D. Intermediates in the assembly pathway of the double-stranded RNA virus ϕ6. *EMBO J.* **16,** 4477–4487 (1997).
16. Butcher, S. J., Grimes, J. M., Makeyev, E. V., Bamford, D. H. & Stuart, D. I. A mechanism for initiating RNA-dependent RNA polymerization. *Nature* **410,** 235–240 (2001).
17. Ilca, S. L. *et al.* Localized reconstruction of subunits from electron cryomicroscopy images of macromolecular complexes. *Nat. Commun.* **6,** 8843 (2015).
18. Pirttimaa, M. J., Paatero, A. O., Frilander, M. J. & Bamford, D. H. Nonspecific nucleoside triphosphatase P4 of double-stranded RNA bacteriophage ϕ6 is required for single-stranded RNA packaging and transcription. *J. Virol.* **76,** 10122–10127 (2002).
19. El Omari, K. *et al.* Tracking in atomic detail the functional specializations in viral RecA helicases that occur during evolution. *Nucleic Acids Res.* **41,** 9396–9410 (2013).
20. Poranen, M. M., Butcher, S. J., Simonov, V. M., Laurinmäki, P. & Bamford, D. H. Roles of the minor capsid protein P7 in the assembly and replication of double-stranded RNA bacteriophage ϕ6. *J. Mol. Biol.* **383,** 529–538 (2008).
21. Nemecek, D., Qiao, J., Mindich, L., Steven, A. C. & Heymann, J. B. Packaging accessory protein P7 and polymerase P2 have mutually occluding binding sites inside the bacteriophage ϕ6 procapsid. *J. Virol.* **86,** 11616–11624 (2012).
22. Nemecek, D. *et al.* Stepwise expansion of the bacteriophage ϕ6 procapsid: possible packaging intermediates. *J. Mol. Biol.* **414,** 260–271 (2011).
23. Huiskonen, J. T. *et al.* Structure of the bacteriophage ϕ6 nucleocapsid suggests a mechanism for sequential RNA packaging. *Structure* **14,** 1039–1048 (2006).
24. Stitt, B. L. & Mindich, L. Morphogenesis of bacteriophage ϕ6: a presumptive viral membrane precursor. *Virology* **127,** 446–458 (1983).
25. Jäälinoja, H. T., Huiskonen, J. T. & Butcher, S. J. Electron cryomicroscopy comparison of the architectures of the enveloped bacteriophages ϕ6 and ϕ8. *Structure* **15,** 157–167 (2007).
26. Poranen, M. M., Paatero, A. O., Tuma, R. & Bamford, D. H. Self-assembly of a viral molecular machine from purified protein and RNA constituents. *Mol. Cell* **7,** 845–854 (2001).
27. Zhang, X., Jin, L., Fang, Q., Hui, W. H. & Zhou, Z. H. 3.3 A cryo-EM structure of a nonenveloped virus reveals a priming mechanism for cell entry. *Cell* **141,** 472–482 (2010).
28. Dryden, K. A. *et al.* Early steps in reovirus infection are associated with dramatic changes in supramolecular structure and protein conformation: analysis of virions and subviral particles by cryoelectron microscopy and image reconstruction. *J. Cell Biol.* **122,** 1023–1041 (1993).
29. Grimes, J. M. *et al.* The atomic structure of the bluetongue virus core. *Nature* **395,** 470–478 (1998).
30. Zhang, X. *et al.* Atomic model of a nonenveloped virus reveals pH sensors for a coordinated process of cell entry. *Nat. Struct. Mol. Biol.* **23,** 74–80 (2016).
31. Zhou, Z. H. *et al.* Electron cryomicroscopy and bioinformatics suggest protein fold models for rice dwarf virus. *Nat. Struct. Biol.* **8,** 868–873 (2001).
32. Nakagawa, A. *et al.* The atomic structure of rice dwarf virus reveals the self-assembly mechanism of component proteins. *Structure* **11,** 1227–1238 (2003).
33. Settembre, E. C., Chen, J. Z., Dormitzer, P. R., Grigorieff, N. & Harrison, S. C. Atomic model of an infectious rotavirus particle. *EMBO J.* **30,** 408–416 (2011).
34. Mohd Jaafar, F. *et al.* The structure and function of the outer coat protein VP9 of Banna virus. *Structure* **13,** 17–28 (2005).
35. Yang, Z. R., Thomson, R., McNeil, P. & Esnouf, R. M. RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. *Bioinformatics* **21,** 3369–3376 (2005).
36. de Haas, F., Paatero, A. O., Mindich, L., Bamford, D. H. & Fuller, S. D. A symmetry mismatch at the site of RNA packaging in the polymerase complex of dsRNA bacteriophage ϕ6. *J. Mol. Biol.* **294,** 357–372 (1999).
37. Huiskonen, J. T., Jäälinoja, H. T., Briggs, J. A. G., Fuller, S. D. & Butcher, S. J. Structure of a hexameric RNA packaging motor in a viral polymerase complex. *J. Struct. Biol.* **158,** 156–164 (2007).
38. Liu, H., Wu, L. & Zhou, Z. H. Model of the trimeric fiber and its interactions with the pentameric penton base of human adenovirus by cryo-electron microscopy. *J. Mol. Biol.* **406,** 764–774 (2011).
39. Nemecek, D. *et al.* Subunit folds and maturation pathway of a dsRNA virus capsid. *Structure* **21,** 1374–1383 (2013).

40. Luque, D. *et al.* The T = 1 capsid protein of *Penicillium chrysogenum* virus is formed by a repeated helix-rich core indicative of gene duplication. *J. Virol.* **84,** 7256–7266 (2010).
41. Paatero, A. O., Mindich, L. & Bamford, D. H. Mutational analysis of the role of nucleoside triphosphatase P4 in the assembly of the RNA polymerase complex of bacteriophage ϕ6. *J. Virol.* **72,** 10058–10065 (1998).
42. Ojala, P. M., Romantschuk, M. & Bamford, D. H. Purified ϕ6 nucleocapsids are capable of productive infection of host cells with partially disrupted outer membranes. *Virology* **178,** 364–372 (1990).
43. Cvirkaite-Krupovic, V., Poranen, M. M. & Bamford, D. H. Phospholipids act as secondary receptor during the entry of the enveloped, double-stranded RNA bacteriophage ϕ6. *J. Gen. Virol.* **91,** 2116–2120 (2010).
44. Tuma, R., Bamford, J. K., Bamford, D. H. & Thomas, G. J. Assembly dynamics of the nucleocapsid shell subunit (P8) of bacteriophage ϕ6. *Biochemistry* **38,** 15025–15033 (1999).
45. Rousseau, F., Schymkowitz, J. W., Wilkinson, H. R. & Itzhaki, L. S. Three-dimensional domain swapping in p13suc1 occurs in the unfolded state and is controlled by conserved proline residues. *Proc. Natl Acad. Sci. USA* **98,** 5596–5601 (2001).
46. Gil-Carton, D. *et al.* Insight into the assembly of viruses with vertical single β-barrel major capsid proteins. *Structure* **23,** 1866–1877 (2015).
47. Olkkonen, V. M., Ojala, P. M. & Bamford, D. H. Generation of infectious nucleocapsids by *in vitro* assembly of the shell protein on to the polymerase complex of the dsRNA bacteriophage ϕ6. *J. Mol. Biol.* **218,** 569–581 (1991).
48. Vidaver, A. K., Koski, R. K. & Van Etten, J. L. Bacteriophage ϕ6: a lipid-containing virus of *Pseudomonas phaseolicola*. *J. Virol.* **11,** 799–805 (1973).
49. Bamford, D. H., Ojala, P. M., Frilander, M., Walin, L. & Bamford, J. K. H. in *Methods in Molecular Genetics* Vol. 6 (ed. Adolph, K. W.) 455–474 (Academic Press, 1995).
50. Li, X. *et al.* Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* **10,** 584–590 (2013).
51. Mindell, J. A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* **142,** 334–347 (2003).
52. Scheres, S. H. W. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180,** 519–530 (2012).
53. Bai, X.-C., Rajendra, E., Yang, G., Shi, Y. & Scheres, S. H. Sampling the conformational space of the catalytic subunit of human γ-secretase. *Elife* **4,** e11182 (2015).
54. Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11,** 63–65 (2014).
55. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132 (2004).
56. Goddard, T. D., Huang, C. C. & Ferrin, T. E. Visualizing density maps with UCSF Chimera. *J. Struct. Biol.* **157,** 281–287 (2007).
57. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 213–221 (2010).
58. Davis, I. W. *et al.* MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35,** W375–W383 (2007).
59. Goddard, T. D., Huang, C. C. & Ferrin, T. E. Software extensions to UCSF chimera for interactive visualization of large molecular assemblies. *Structure* **13,** 473–482 (2005).

Acknowledgements

We thank Alistair Siebert for electron microscopy support and Riitta Tarkiainen for technical help. The OPIC electron microscopy facility was founded by a Wellcome Trust JIF award (060208/Z/00/Z) and is supported by a WT equipment grant (093305/Z/10/Z). The EU ESFRI Instruct Centre for Virus Production (ICVIR) facility used in this study is supported by the Academy of Finland (grant 272853) and University of Helsinki and this is a collaboration with the Oxford Instruct Centre. This work was funded by Academy of Finland (218080 and 263677 to J.T.H; 283192, 250113, and 272507 to M.M.P.), Sigrid Jusélius Foundation, Medical Research Council (MR/N00065X/1 to K.E.O.), European Research Council under the European Union's Horizon 2020 research and innovation programme (649053 to J.T.H.) and Wellcome Trust Core Award Grant Number 090532/Z/09/Z. The authors acknowledge the support of employees and the use of experimental resources of Instruct, a Landmark ESFRI project.

Author contributions

Z.S., S.L.I. and J.T.H. collected and analysed cryo-EM data. K.E.O., M.M.P., D.I.S. and J.T.H. designed the study. X.S. prepared the samples. Z.S., K.E.O. and A.K. built and refined the atomic models. Z.S. and J.T.H. wrote the manuscript and all authors read and commented on the manuscript.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Sun, Z. *et al.* Double-stranded RNA virus outer shell assembly by *bona fide* domain-swapping. *Nat. Commun.* **8,** 14814 doi: 10.1038/ncomms14814 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

© The Author(s) 2017

PLOS | PATHOGENS

RESEARCH ARTICLE

Acidic pH-Induced Conformations and LAMP1 Binding of the Lassa Virus Glycoprotein Spike

Sai Li[1], Zhaoyang Sun[1], Rhys Pryce[1], Marie-Laure Parsy[1], Sarah K. Fehling[2], Katrin Schlie[2], C. Alistair Siebert[1], Wolfgang Garten[2], Thomas A. Bowden[1], Thomas Strecker[2]*, Juha T. Huiskonen[1]*

1 Division of Structural Biology, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom, **2** Institute of Virology, Philipps Universität Marburg, Marburg, Germany

* strecker@staff.uni-marburg.de (TS); juha@strubi.ox.ac.uk (JTH)

CrossMark
click for updates

OPEN ACCESS

Citation: Li S, Sun Z, Pryce R, Parsy M-L, Fehling SK, Schlie K, et al. (2016) Acidic pH-Induced Conformations and LAMP1 Binding of the Lassa Virus Glycoprotein Spike. PLoS Pathog 12(2): e1005418. doi:10.1371/journal.ppat.1005418

Editor: Ted C. Pierson, National Institute of Health, UNITED STATES

Received: September 24, 2015

Accepted: January 5, 2016

Published: February 5, 2016

Copyright: © 2016 Li et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Sub-tomogram averages have been deposited in the Electron Microscopy Databank with the following accession codes: EMD-3290 (GP spike from fixed virions), EMD-3291 (GP spike from VLPs at pH 7), EMD-3292 (GP spike from VLPs at pH 5), EMD-3293 (GP spike from VLPs at pH 5 in complex with LAMP1 fragment) and EMD-3294 (GP spike from VLPs at pH 3). Fitted atomic coordinates of GP1 have been deposited in the Protein Data Bank with PDB ID code 5FT2.

Funding: This work was funded by Medical Research Council (MR/J007897/1 to TAB and JTH, and MR/L009528/1 to TAB; http://www.mrc.ac.uk/),

Abstract

Lassa virus is an enveloped, bi-segmented RNA virus and the most prevalent and fatal of all Old World arenaviruses. Virus entry into the host cell is mediated by a tripartite surface spike complex, which is composed of two viral glycoprotein subunits, GP1 and GP2, and the stable signal peptide. Of these, GP1 binds to cellular receptors and GP2 catalyzes fusion between the viral envelope and the host cell membrane during endocytosis. The molecular structure of the spike and conformational rearrangements induced by low pH, prior to fusion, remain poorly understood. Here, we analyzed the three-dimensional ultrastructure of Lassa virus using electron cryotomography. Sub-tomogram averaging yielded a structure of the glycoprotein spike at 14-Å resolution. The spikes are trimeric, cover the virion envelope, and connect to the underlying matrix. Structural changes to the spike, following acidification, support a viral entry mechanism dependent on binding to the lysosome-resident receptor LAMP1 and further dissociation of the membrane-distal GP1 subunits.

Author Summary

Lassa virus is a zoonotic, hemorrhagic fever-causing pathogen. Because the virus can spread as an aerosol and there are no approved vaccines or specific antiviral drugs currently available, it poses a major impact on human health, affecting annually up to half a million people in West-African countries. Entry of the virus into the cells of the infected individual is the first step where the virus could be stopped. When we try to determine the molecular mechanism of virus entry, information on the structure of the virus and its components can be highly valuable. However, in the case of Lassa virus, this remains poorly understood. Here, we used high-resolution electron cryomicroscopy and tomography techniques to image chemically-inactivated Lassa virus. Computational image reconstruction allowed determination of the three-dimensional structure of the virus, and allowed us to localize the protein, lipid bilayer, and RNA genome components. Our analysis of non-

Academy of Finland (218080 and 263677 to JTH; http://www.aka.fi/en/), Von Behring-Röntgen-Stiftung (to SKF and TS; http://www.br-stiftung.de/) and the German Research Council (SFB 1021 TP B05 to TS; http://www.dfg.de/en/). The OPIC electron microscopy facility was founded by a Wellcome Trust JIF award (060208/Z/00/Z) and is supported by a WT equipment grant (093305/Z/10/Z; http://www.wellcome.ac.uk/). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

infectious virus-like particles at different pH's and in the presence of a functional cellular receptor, human LAMP-1, revealed specific structural rearrangements in the surface protein spike, providing a molecular-level rationale for this important stage of host cell entry.

Introduction

Lassa virus (LASV) is an enveloped, ambi-sense, bi-segmented RNA virus endemic throughout Western Africa and is the most lethal of all known Old World arenaviruses. Due to the high mortality rates amongst hospitalized patients (~15%), ability of the virus to be spread by aerosol, and absence of licensed protective vaccines or therapeutics to treat acute infection, LASV has been classified as a biosafety level (BSL) 4 pathogen [1].

The LASV RNA genome encodes an RNA-dependent RNA polymerase (L), nucleoprotein (NP), matrix protein (Z), and a highly-glycosylated membrane glycoprotein (GP). GP is synthesized as an inactive precursor preGPC, which is co-translationally cleaved by signal peptidase into GPC and the stable signal peptide (SSP) [2]. Post-translational maturation cleavage of GPC by host protease SKI-1/S1P yields the receptor-binding subunit GP1 and the membrane-spanning fusion subunit GP2 [3–5]. SSP is not only critical by functioning as a *trans*-acting maturation factor [6], but also associates with the GP2 subunit, resulting in a tripartite mature GP complex on the viral surface [7,8].

Previous chemical cross-linking studies and sucrose density gradient analysis have revealed that GP complexes assemble on the virion surface as trimeric spikes [9]. Crystallographic analysis of GP1 has recently identified a putative binding surface for the lysosome-associated membrane protein 1 (LAMP1) [10], an essential intracellular receptor [11]. Structural studies on glycoproteins from related arenaviruses have further revealed that the GP1 forms a novel α/β fold [12,13] and that the GP2 is a class I fusion protein [3,5,14]. However, no three-dimensional (3D) structures exists to date for the whole LASV spike and there is thus paucity in our understanding the molecular architecture of LASV which in turn is important for understanding the mechanism of LASV entry.

Here, we have studied the ultrastructure of fixed Lassa virions by using electron cryomicroscopy and tomography. By averaging 6,500 spike densities, we derived the structure of the full-length GP spike in its pre-fusion conformation at 14 Å resolution. At this resolution, the trimeric arrangement and contacts to the underlying matrix were resolved. Low pH structures of GP trimers, derived from virus-like particles (VLPs), allowed assessing the conformational changes upon acidification. Addition of purified LAMP1 fragment allowed mapping its binding site on the spike and fitting of the GP1 X-ray structure [10] to the EM density allowed determination of its putative membrane-distal location. At pH below 5.0, the GP1 subunit was shed, possibly priming the GP2 for fusion. Taken together, these results help to map the architecture of Lassa virions and shed light on the conformational changes and internal receptor binding occurring during endosomal entry.

Results

Ultrastructure of Lassa virus

We produced live LASV in African green monkey kidney epithelial cells under BSL-4 containment conditions. Cell supernatants containing LASV virions were subsequently inactivated by chemical cross-linking using 4% paraformaldehyde (PFA). Prior to structural analysis, we

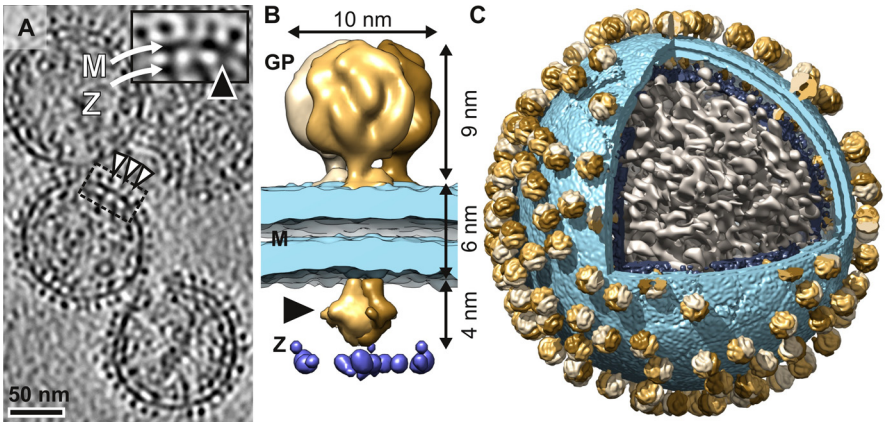


Fig 1. Structure of Lassa virus solved using electron cryotomography and sub-tomogram averaging. (A) A slice through a 3D tomogram of LASV. Three glycoprotein are spikes indicated (white arrowheads). The inset shows a magnified view of the area indicated (dotted line). The membrane (M), matrix (Z), and density bridging a spike to the Z layer (black arrowhead) are indicated. (B) LASV GP spike structure segmented to three spike protomers (different shades of brown), membrane (light blue), and the Z layer (dark blue). Intra-viral tails connecting the spike to the underlying matrix are indicated (arrowhead). (C) One LASV virion is shown cut open highlighting the irregular organization of the glycoprotein spikes, membrane, Z layer, and RNP density (gray).

doi:10.1371/journal.ppat.1005418.g001

optimized a gradient ultracentrifugation based protocol for virion purification (see Materials and Methods).

Tomographic 3D reconstructions ('tomograms') calculated from electron cryomicroscopy tilt series of purified virions revealed both the exterior and interior density of LASV virions (Fig 1A). The virions were roughly spherical (S1 Fig, LASV) with diameter of 132±22 nm (S2 Fig, LASV). Consistent with previous observations [15], a density plot calculated normal to the membrane (S3 Fig, LASV) revealed i. an outer layer, comprising the GP spikes, ii. two middle layers, comprising the two leaflets of the lipid bilayer, and iii. two internal layers ('inner track 1' and 'inner track 2') [15], the first comprising GP intra-viral tails (see below) and the second comprising Z and possibly ribonucleoprotein (RNP), which also partially fills the virion interior (Fig 1A).

Glycoprotein spike trimers connect to the underlying matrix and are randomly distributed on the virion surface

By averaging ~6,500 volumes ('sub-tomograms') of GP spikes extracted from tomograms of 48 inactivated LASV virions, we solved the 3D structure of the GP spike (Fig 1B and 1C and Table 1) to 14 Å resolution, as estimated by the Fourier shell correlation (FSC; S4 Fig). The structure revealed a trimeric assembly, measuring ~9 nm in height and ~10 nm in width (Fig 1B). Three separate densities ('legs'), spaced by ~4 nm, anchor the spike to the membrane surface. By analogy to the trimeric structures of the HIV-1 [16] and Ebola virus glycoprotein [17] spikes, where the class I fusion glycoproteins subunits are membrane-proximal, we putatively assign these membrane proximal legs to the GP2 glycoprotein. This is supported by the presence of a trans-membrane (TM) region at the C-terminus of the LASV GP2, as predicted by sequence analysis. As SSP interacts with GP2 [8,18], it may also contribute to these leg densities. Earlier biochemical studies have suggested several topologies for SSP [8,19–21]. However, at 14-Å resolution we were unable to resolve the structure of SSP, making structural arrangements and orientation of SSP within the membrane proximal region difficult to predict.

Table 1. Data acquisition and processing statistics.

	LASV	VLP pH 7.3	VLP pH 5.2	VLP+LAMP1 pH 5.5	VLP pH 3.0
Data acquisition					
Magnification	37,037	37,037	37,037	37,037	37,037
Tilt series	27	31	16	30	10
Tilt range (°)	-45–45	-45–45	-45–45	-45–45	-45–45
Interval (°)	5	5	5	5	5
Frames per tilt	8	8	8	8	8
Dose (e^{-}/Å^2)	~60	~60	~60	~60	~60
Defocus (um)[a]	2.3–3.9	1.7–3.7	2.8–6.7	1.4–4.6	3.0–3.9
Data processing					
Virions/VLPs	48	112	22	113	23
Seeds [b]	n/a	n/a	9,406	n/a	7,821
Sub-tomograms	6,496	2,764	2,578	8,527	1,454
Box size (pixels)	128	128	128	128	128
Pixel size (Å)	2.7	2.7	2.7	2.7	2.7
Resolution (Å) [c]	13.6	13.9	16.4	14.8	16.7

[a]Positive defocus denotes underfocus.

[b]Number of 'seeds' created on the virion/VLP surfaces. If n/a is indicated, manual picking was used instead of seeds.

[c]Resolution (Fourier shell correlation = 0.5).

doi:10.1371/journal.ppat.1005418.t001

Furthermore, at this resolution the TM regions cannot be resolved and were thus not visible in our structure.

Visual inspection of the tomographic slices revealed connections bridging the spike to the Z layer underneath the membrane (Fig 1A, inset). Density corresponding to these connections was ordered in the averaged GP structure (Fig 1B). The measured volume of this intra-viral density corresponded to molecular mass of ~18 kDa. This mass is consistent with the three 41-residue long intra-viral tails of GP2 (14 kDa in total), indicated by sequence analysis. We thus putatively assign the 18-kDa intra-viral density collectively to the three tails of GP2. In addition, SSP may contribute additional mass to these tails.

Template matching allowed mapping the distribution of spikes on the virions (N = 48). The spikes (273±22) almost fully covered the virion surface (Fig 1C). No higher-order clustering was observed (S5 Fig), confirming the well-accepted notion that LASV is not icosahedrally symmetric and suggesting that no specific interactions between the GP trimers exist.

Structure of the native LASV GP from virus-like particles

To account for possible effects of chemical fixation, we exploited a previously reported stable cell line expressing non-infectious virus-like particles (VLPs) displaying the full-length LASV GP but lacking the Z [22]. Similar to the fixed virions, VLPs purified from cell culture supernatants were covered with GP spikes (S1 Fig; VLP pH 7). However, size measurements (S2 Fig) showed that VLPs were significantly smaller (diameter 82±15 nm) than the virions (132±22 nm) and also their shape was more variable (S1 Fig). This is most likely due to the absence of Z, which modulates virion budding and RNP packaging [23].

Following the same sub-tomogram averaging approach as for the virions, we solved the structure of the native GP spike from the VLPs (Fig 2, pH 7; Table 1) to 14-Å resolution

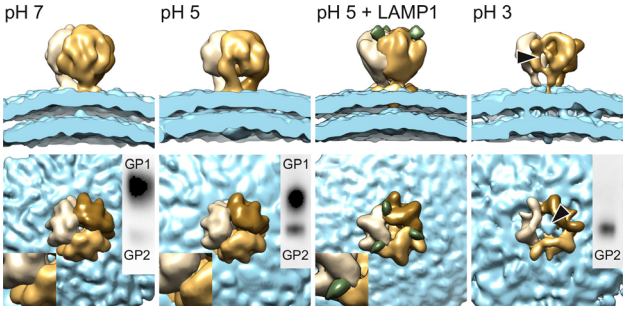


Fig 2. Structural transitions and LAMP1 binding of the LASV GP spikes upon acidification. (A) GP structures at different pHs are shown from side (top row) and top (bottom row). All volumes were filtered to 17-Å resolution, rendered at molecular threshold corresponding to the expected molecular mass, and colored as in Fig 1. Residual density corresponding to LAMP1 is colored in green. Inserts in the lower left corners show a close-up of the interface between two spike monomers. Insets in the top right corners show Western blot analysis of GP1 and GP2 subunits. The arrowheads indicate the missing density in the central top part and side of the pH 3 structure.

doi:10.1371/journal.ppat.1005418.g002

(S4 Fig). Visual inspection revealed that the VLP-derived spike ectodomain structure shared the same tripodal architecture and structural features with the virion-derived spike structure (S6A Fig). To provide a quantitative measure for similarity, we performed cross-FSC analysis, which indicated that the two ectodomains agreed to 16-Å resolution (S6B Fig).

The density profile calculated across the VLP membrane was practically identical to that of the virion for the density corresponding both to the GP ectodomain and GP2 intra-viral tails. As expected, density assigned to Z in the virion was absent in the VLPs that lacked this component (S3 Fig, VLP). Surprisingly, the GP2 intra-viral tails that were ordered in the virions (Fig 1B) were disordered in the VLP-derived GP structure (S6A Fig). While it is possible to attribute the ordering of the intra-viral tails to the GP–Z interaction [22,24], we cannot exclude the possible effect of chemical fixation in ordering the intra-viral tails.

The direct comparison of the virion-derived and VLP-derived GP structures supports our assignment of the density in the first density layer under the membrane (denoted earlier as 'inner track 1') [15] to the GP2 intra-viral tails. We note that this assignment differs from a previous assignment, where the 'inner track 1' was assigned to the Z layer in the two-dimensional cryo-EM images of three other arenaviruses, namely Pichinde, Tacaribe and lymphocytic choriomeningitis virus [15]. We note, however, that we cannot exclude structural differences between LASV and the other studied arenaviruses. It is also possible that Z contributes some density to the inner track 1 in LASV. In conclusion, by comparison to the VLP system, we could verify the LASV GP spike structure and our assignment of the intra-viral GP2 density.

Glycoprotein spike binding to LAMP1 at endosomal pH

Following virus internalization into the host cell through GP attachment to α-dystroglycan [25], DC-SIGN [26,27], or other cellular receptors [26], fusion of viral and host cell membranes occurs at an unusually low pH of 3–4.5 [28], a process mediated by recognition of the host lysosomal receptor, LAMP1 [11]. Receptor binding is modulated by pH; GP binds to α-dystroglycan at pH 8.0 but not at pH 6.0 and below, and, conversely, LAMP1 binding occurs only at pH 6.0 and below [11].

To study GP conformation at endosomal pH, we determined the VLP-derived GP structure at pH 5.2 (Table 1 and Fig 2, pH 5, and S1 Fig VLP pH 5). The structure was resolved at 16-Å resolution (S4 Fig) and revealed conformational differences when compared to the same structure at pH 7.3 (Fig 2, pH 7). These can be attributed to a conformational difference at the

membrane-proximal legs and a slight opening up of the spike, creating crevices at the top of the spike between the GP subunits.

To map the LAMP1 binding site on the GP spike, we repeated the electron cryomicroscopy-based structure analysis after addition of purified LAMP1 at pH 5.5 (Fig 2, pH 5 + LAMP1). First we expressed and purified the predicted membrane distal, highly glycosylated β-prism domain of human LAMP1 (residues 27–194) [29]. This domain was biologically active binding to VLPs at pH 5.0 (S7 Fig). The GP–LAMP1 complex structure was resolved to 15-Å resolution (S4 Fig). Additional density attributed to LAMP1 was evident at the top of the spike and located close to the crevices observed in the VLP GP structure in absence of LAMP1 (Fig 2; pH 5 + LAMP1). The volume of LAMP1 density was less than expected based on the molecular mass of the LAMP1 glycoprotein fragment used for analysis (S7 Fig), suggesting that LAMP1 either bound to VLP GP in sub-stoichiometric amounts and/or that the binding was flexible. In conclusion, the location of this additional residual density allowed mapping of the LAMP1 binding site to the membrane distal part of the spike.

GP1 subunit is shed from the spike at lysosomal pH

In line with previous reports [11], our Western blot analysis revealed that the GP1 subunit was shed from the VLPs at pH 3.0 and 4.0 (S8 Fig). To determine the LASV GP structure in the absence of GP1, we performed electron cryotomography and sub-tomogram averaging of VLP-derived GPs at pH 3.0 (Table 1 and S1 Fig, VLP pH 3). The structure was missing density in the membrane-distal part of the spike (Fig 2, pH 3), suggesting that the GP1 occupies this region.

In the absence of a target membrane, the pH 3 structure may constitute SSP in complex with either a non-productive post-fusion conformation of GP2 or an intermediate conformation prior to fusion. Indeed, it is strikingly different from the canonical post-fusion ectodomain structures of other class I fusion proteins, which display a central trimeric alpha-helical coiled coil (S9 Fig) [3,30]. However, we also cannot exclude the possibility that a fraction of GP2 has mediated fusion of some VLPs with one another. The resulting post-fusion conformation species in this fraction may have been excluded from our analysis. Despite these limitations, the pH 3 structure allowed putative localization the GP1 subunit to the membrane distal region of the spike. This is consistent with the mapped LAMP1 binding site in the same region, as LAMP1 binds to the GP1 subunit [10].

Fitting of GP1 structure into the GP spike density

The structure of LASV GP1 has been recently solved by X-ray crystallography [10]. To localize the GP1 and determine its orientation relative to the LAMP1 binding site, we used volumetric correlation to fit the GP1 crystal structure into our GP reconstruction (Fig 3 and S10 Fig). As the GP1 has been crystallized at pH 5 and acidic pH has been suggested to have an effect on GP1 conformation [10], we used our GP structure derived from VLPs at pH 5 for fitting. The fitting was unambiguous (S10C Fig; cross-correlation coefficient 0.86) and placed the GP1 structure to the membrane distal part of the spike (Fig 3). Interestingly, the triad of histidines that has been indicated in LAMP1 binding [10] points towards the crevice between GP1 subunits (Fig 3), the putative LAMP1 binding side (Fig 2). The histidines may act as a pH sensor [10] opening up the spike at pH 5 to create a binding site for LAMP1, but further studies are needed to test this hypothesis.

Discussion

Viral surface glycoproteins mediate entry of enveloped viruses into host cells through receptor binding and subsequent fusion of viral and cellular membranes. Therefore, the structural

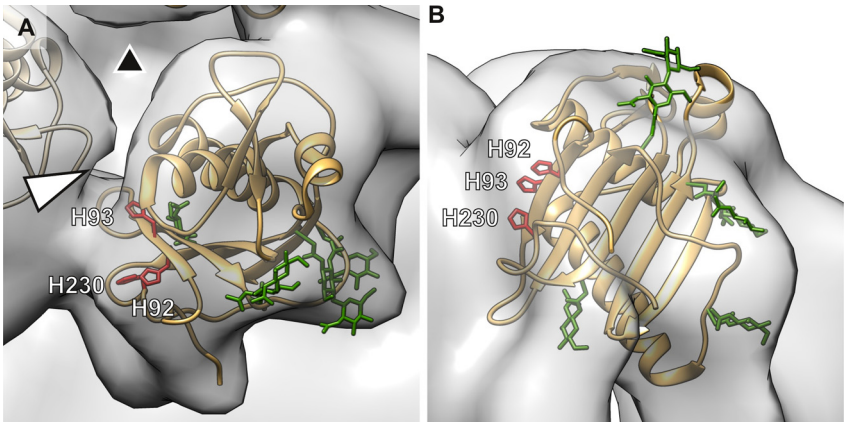


Fig 3. Location of the GP1 subunit in the glycoprotein spike. (A) Crystallographic structure [10](brown ribbon; PDB:4ZJF) was fitted into density (transparent surface) derived by electron cryomicroscopy and subtomogram averaging of GP spikes from virus-like particles at pH 5. The glycans are represented with sticks and colored in green. The histidines indicated in binding of an intracellular receptor (lysosome-associated membrane protein 1, LAMP1) are labeled and colored in red. This triad of histidines faces a crevice (white arrowhead) between the spike subunits. The three-fold symmetry axis is indicated with a black triangle. (B) Same fitting as in *A* is shown from the side.

doi:10.1371/journal.ppat.1005418.g003

characterization of viral glycoproteins is of fundamental interest to understand the molecular mechanisms underlying host cell infection. LASV uses a two-step process to enter target cells during infection. Upon receptor binding at the cell surface, LASV enters the cell through late endosomes/multivesicular bodies, where GP undergoes a pH-induced switch to engage the intracellular receptor LAMP1 [11,31].

We analyzed the three-dimensional ultrastructure of authentic LASV using electron cryotomography. Sub-tomogram averaging yielded a structure of the glycoprotein spike at 14-Å resolution. To our knowledge, the GP structures presented here constitute the highest resolution viral glycoprotein spike structures solved *in situ*. The structure revealed a trimeric glycoprotein architecture, which confirms previous chemical cross-linking studies, demonstrating the formation of trimeric mature glycoprotein spikes on the surface of LASV particles [9].

Our analysis of VLPs harboring GP revealed that the VLPs were significantly smaller in size (diameter 82±15 nm) than the virions (132±22 nm) and that their shape was more variable; possibly due to the absence of other viral components. Despite these morphological differences, comparison of the GP spike structures between virions and non-infectious VLPs showed that they share high degree of structural similarity. It is thus likely that the GP-harboring VLPs mimic at least to some extent the immunological properties of the authentic virions, which is critical for the use of VLP-based vaccine platforms for prophylactic protection against Lassa fever [32].

Although the outer spike domain on VLPs appeared structurally similar to LASV virions, differences were observed for the intra-viral domain architecture. The intra-viral GP2/SSP tails were ordered in LASV virions, however, in the VLP-derived GP structure they were disordered. The arenaviral matrix protein Z plays an important role in viral assembly and budding, acting as a bridging factor between the virus envelope spike complex and the viral RNP [23]. Thus, the assembly of infectious LASV involves an interaction between Z and the cytoplasmic portion of the GP complex, which is supported by our finding that Z is positioned beneath the viral bilayer and connected with the spike in LASV virions. Interestingly, arenaviruses are unique in that Z associates via its N-terminal myristylation with SSP, even in the absence of other subunits of the GP complex [24]. However, the detailed organization of SSP in the tripartite GP

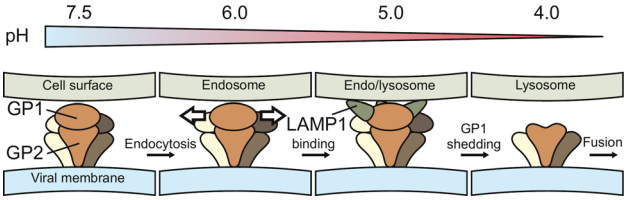


Fig 4. Model for Lassa virus entry. A schematic representation of a hypothetical entry model, derived from the structures determined in this study, is shown. An approximate range of pH in different cellular compartments is depicted as a color gradient from neutral (blue) to very acidic (red). Lassa virus glycoprotein (GP) spike trimer is depicted in three shades of brown and the two different subunits (GP1 and GP2) are labeled. Lysosome-associated membrane protein 1 (LAMP1), an intracellular Lassa virus receptor, is labeled and colored in green. Viral membrane is colored in blue. Different cellular membranes are labeled and colored in gray. Formation of the crevices between GP1 subunits is indicated with white arrows. See text for full description of the entry model.

doi:10.1371/journal.ppat.1005418.g004

complex and the interactions between the GP complex and Z remain elusive. Furthermore, how Z potentially contributes to the stabilization of the GP intra-viral domain remains to be determined.

In addition to providing an independent control for our structural analysis, the VLP system allowed us to probe the structural states of the GP in acidic conditions, similar to those encountered by LASV during endocytotic entry. In another group of membraneous viruses, alphaviruses, the E1 fusion glycoprotein is shielded by E2 glycoprotein preventing premature fusion [33]. Here, the membrane-distal GP1 remained bound to the membrane-anchored GP2 at pH 5.0, possibly providing an analogous shielding function and at the same time creating a binding site for the intra-cellular receptor LAMP1.

To rationalize the findings from our structural analysis, we propose a hypothetical entry model (Fig 4). In this model, GP1 structure opens up at endosomal pH, creating a crevice between the membrane distal GP1 subunits and exposing a binding site for LAMP1. Interestingly, based on the GP1 crystal structure a pH dependent conformational change in GP1 has been suggested recently, preventing α-dystroglycan and enabling LAMP1 binding [10]. It is tempting to hypothesize that LAMP1 promotes detachment of GP1 from GP2 but the role of LAMP1 in promoting LASV fusion still remains unclear. We further suggest that detachment of GP1 at lysosomal pH primes the GP2/SSP complex for membrane-fusion [7,28,34]. Prior to membrane fusion, GP2 is expected to embed the hydrophobic fusion peptide into the target membrane [35] and further to adopt a post-fusion trimeric alpha-helical coiled coil characteristic of class I fusion proteins [3,30]. The organization of SSP during this process remains unknown.

Interestingly, the observed three-legged architecture of LASV GP differs from that reported for a novel arenavirus-like virus infecting snakes [36], belonging to the genus Reptarenavirus within the *Arenaviridae* family, which is surprisingly more similar to the glycoprotein spike of Ebola virus [17] (Fig 5). This observation may reflect both the structural diversity amongst mammalian and reptilian arenaviruses and an evolutionary link to filoviruses [37]. In conclusion, our structures provide a blueprint for the surface of mammalian arenaviruses and a model for the structural rearrangements that occur during host cell entry.

Materials and Methods

Preparation, inactivation and purification of LASV

All work with infectious LASV was performed under the highest safety precautions in the biosafety level-4 (BSL-4) facility at the Institute of Virology in Marburg. Vero cells were infected

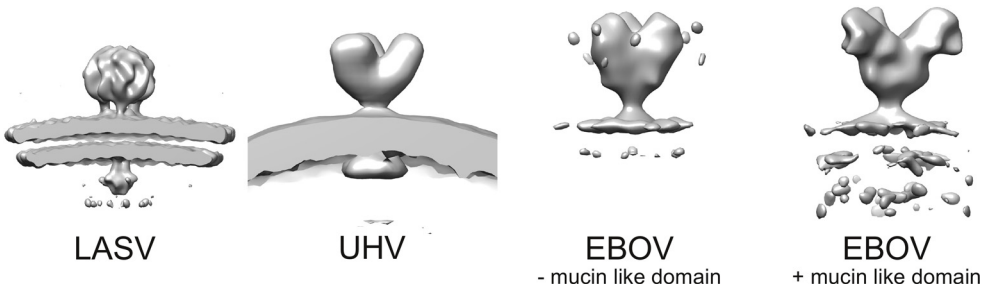


Fig 5. Comparison of arenavirus and filovirus glycoprotein spike structures. Structures of Lassa virus spike (LASV, this study), spike of the arenenavirus-like virus infecting snakes (University of Helsinki Virus, UHL; EMD-2424), and Ebola virus spike (EBOV) with (EMD-6003) and without (EMD-6004) the mucin-like domain are shown at the same scale for comparison.

doi:10.1371/journal.ppat.1005418.g005

with LASV (strain Josiah) at a multiplicity of infection of 0.1. Infected cells were maintained in Dulbecco's modified Eagle medium (DMEM; Gibco, Thermo Fischer Scientific, Waltham, MA) supplemented with 2% fetal calf serum (FCS; Gibco, Thermo Fischer Scientific, Waltham, MA), penicillin (100 U/ml), streptomycin (100 mg/ml), and L-glutamine (2 mmol/l) (Invitrogen, Thermo Fischer Scientific, Waltham, MA) at 37°C under 5% CO_2. Supernatant was collected 5 days after infection and cleared twice by centrifugation at 5,000 g for 30 min. Supernatant containing LASV was then subjected to virus inactivation using PFA (4% final concentration) in DMEM for 24 h. The sample was removed from the BSL-4 facility and after additional 24 h processed for further experiments.

Supernatants containing PFA-fixed LASV (30 ml) were again cleared twice by low-speed centrifugation (5,000 g for 30 min) to remove potential PFA-induced aggregates and subsequently pelleted through 20% sucrose cushion by ultracentrifugation (100,000 g for 2 h). Pellets were resuspended in 150 μl phosphate-buffered saline (PBS) for 12 h at 4°C. The samples were further purified by pelleting twice through 20% sucrose cushion by ultracentrifugation (100,000 g for 2 h) and resuspended in PBS. The sample was laid onto a 10%–60% sucrose gradient and centrifuged (100,000 g for 12 h). The gradient was fractionated using a gradient fractionator (Biocomp, Fredericton, NB, Canada) and fractions were collected using a fraction collector (Gilson, Middleton, WI) while monitoring the UV absorbance at 254 nm using a BioProbe (Biocomp, Fredericton, NB, Canada) together with EM-1 UV monitor (Bio-Rad Laboratories, Hercules, CA). Ten 100-μl fractions collected around the absorbance peak were pooled, pelleted and resuspended as described above.

Expression and purification of virus-like particles

Madin–Darby canine kidney (MDCK-II) cells stably expressing LASV GP protein [22] were cultured in DMEM supplemented with 10% FCS for 96 h. The medium was replaced by DMEM with 2% FCS for VLP expression. After 96 h, the cell supernatant was cleared twice by centrifugation (5,000 g for 30 min). Virus-like particles (VLPs) were pelleted through 20% sucrose cushion by ultracentrifugation (125,000 g for 3 h). Pellets were resuspended in 150 μl PBS for 12 h at 4°C.

Solutions of succinic acid, sodium dihydrogen phosphate and glycine were mixed in 2:7:7 molar ratio to make 50 mM SPG buffer. The buffer was adjusted to the desired pH by adding 0.5 M sodium hydroxide or 1 M hydrogen chloride. For Western blot analysis of particles at different pH, 300 μl of sample was incubated with 5 ml SPG buffer at the desired pH for 5 min at room temperature, then pelleted through sucrose cushion and resuspended as described

above. Protein concentration of purified and different pH treated VLPs was measured using Nanodrop (Thermo Fischer Scientific).

Western blot analysis

An aliquot corresponding to 8 μg protein from each sample was separated by SDS-PAGE on a NuPAGE 4–12% Bis-Tris gel (Life Technologies, Thermo Fischer Scientific) and subsequently transferred to nitrocellulose membranes by using a dry-blotting system (iBlot; Life Technologies). Membranes were blocked in 5% milk in PBS buffer for 2 hours. Membranes were incubated with 1/3,000 diluted primary antibodies (monoclonal mouse antibody AC1 against GP-C and GP1 [kindly provided by Marie-Claude Georges-Courbot]; polyclonal rabbit antibody GP4 recognizing GP-C and GP2 [38]) for 1 h followed by an incubation with 1/1,000 diluted horseradish peroxidase- conjugated secondary antibodies against mouse and rabbit (Qiagen, Crawley, UK) for 1 h. The membrane was soaked in 5 ml of enhanced chemiluminescence reagent mixture (GE Healthcare, Buckinghamshire, UK) and visualized in a ChemiDoc MP gel imager (Bio-Rad Laboratories, Hercules, CA). A 1-μl aliquot of pre-stained protein maker (Benchmark; Life Technologies) was used to determine the protein mass.

Expression and purification of human LAMP1 membrane distal domain

A fragment of human LAMP1 (residues 27–194; UniProt accession no. P11279), predicted to correspond to the membrane-distal β-prism domain [29], was cloned into the pHLsec mammalian expression vector [39]. The LAMP1 domain was transiently expressed in HEK293T cells in the presence of the α-mannosidase inhibitor, kifunensine [40], with 2 mg of DNA transfected per litre of cell culture. Cell culture supernatant was collected 3 days post-transfection, filtered and diafiltrated against 10 mM Tris (pH 8.0) and 150mM NaCl. Purification comprised a two-step immobilized metal-affinity and size-exclusion procedure. Size-exclusion purification was carried out using a Superdex 200 10/30 column equilibrated against 10 mM Tris (pH 8.0) and 150 mM NaCl.

Binding of human LAMP1 membrane distal domain to VLPs

Enzyme-linked immunosorbent assays were performed in 96-well Nunc-Immuno PolySorp plates (Sigma) to which VLPs were bound overnight at a concentration of 10 μg/ml. Plates were blocked with PBS containing 5% milk powder and incubated with varying concentrations of purified LAMP1 membrane distal domain at different pH. LAMP1 detection was mediated using a rabbit derived anti-6His primary (Abcam, UK), and goat-anti rabbit-HRP (Abcam) antibody combination. Binding was assessed on the basis of optical absorbance at 405 nm after 30 min following HRP substrate addition.

Cryo-EM sample preparation

An aliquot (3 μl) of purified LASV or VLP sample was mixed with 3 μl of colloidal 6-nm gold particles coupled to bovine serum albumin (Aurion, Wageningen, The Netherlands) and applied on a plasma cleaned EM grids coated with holey carbon (C-flat; Protochips, Raleigh, NC). For low pH treatment of VLPs, the grid was floated for three minutes on top of a 2-ml droplet of SPG buffer at the desired pH. The grids were vitrified using a plunger device (CP3; Gatan, Pleasanton, CA) by blotting the sample for 3 s followed by plunge-freezing into a mixture of liquid ethane (37%) and propane (63%) [41].

Electron cryomicroscopy and tomography

All electron cryomicroscopy data were collected using a 300-kV transmission electron microscope (TF30 'Polara'; FEI, Eindhoven, Netherlands) operated at liquid nitrogen temperature. SerialEM [42] was used for acquiring low dose, single axis tilt series (from –45° to 45° with 5° angular sampling). A direct electron detector camera (K2 Summit; Gatan, Pleasanton, CA) mounted behind a post-column energy filter (QIF Quantum LS; Gatan, Pleasanton, CA) was used to acquire data at zero energy-loss mode (slit width 20 eV). At each tilt, a movie consisting of eight frames with total exposure time of 1.6 s was collected at a calibrated magnification of ×37,000 in electron counting super resolution mode, corresponding to a pixel size of 0.675 Å. Tilt series were recorded at nominal defocus values ranging from 2.0 to 3.5 μm under focus and the total electron dose was ~60 $e^{-}/Å^{2}$.

Image pre-processing and tomogram generation

Drift correction [43] was used to correct for the electron beam induced motion by averaging eight frames for each tilt and applying 2× binning in Fourier space. IMOD package [44] was used to reconstruct three-dimensional tomograms from the stacks of tilted images. Images were aligned using the 6-nm gold beads as fiducial markers (S11 Fig). Prior to reconstruction, contribution from the gold beads was computationally removed, contrast transfer function (CTF) parameters were manually estimated (S12 Fig) and the effects of the CTF were corrected in the images by phase flipping [45]. For CTF background correction, noise images were recorded and processed the same was as actual tilt images. The amplitudes were not modified during CTF correction. Further 2× binning was applied, resulting in the final pixel size of 2.7 Å. LASV virions and VLPs were extracted from the full tomograms for further analysis.

Sub-tomogram averaging

Sub-tomogram alignment and averaging was carried out in Dynamo, which takes into account the missing wedge in tomographic data and allows parallel and graphics processing unit accelerated computations [46]. For initial template generation, extracted virion and VLP volumes were low-pass filtered to 80 Å using Bsoft [47]. Spikes were manually picked from the volumes of LASV and VLPs at pH 7 and pH 5 using Dynamo [46]. Dynamo *tomoview* program was used to create a surface model approximating the viral membrane, assuming either spherical or ellipsoidal geometry. Surface normals of the model surface provided initial orientations of the picked spikes. Spikes were extracted from the unfiltered tomograms into boxes of 128×128×128 voxels and averaged together. Cylindrical symmetry and low-pass filter to 35-Å resolution were applied on the average to reduce template bias.

The orientations were refined using the standard refinement strategy in Dynamo that implements an adaptive low-pass ('push-back') filter to further mitigate effects of template bias. Iterative refinement was performed in several stages (Supplementary Methods in S1 Text) [48–50]. Initially, no symmetry was applied to not bias the structure towards any particular symmetry. Only after three-fold appearance became evident in the asymmetric average (S13A Fig; C1), three-fold symmetry was applied in subsequent iterations (S13A Fig; C3). We further tested the effects of the starting model on the resulting structure. Even when the refinement was started with an initial model with wrong symmetry, it converged to a trimeric structure (S13B Fig). Typically a total 20 iterations were run until the resolution stopped improving (S13C Fig). In addition to the standard 'push-back' refinement, we carried out further tests with an alternative, so called 'gold-standard' refinement strategy using the VLP pH 7 data set. Both strategies yielded a trimeric spike at comparable resolutions (S14 Fig; Supplementary Methods in S1 Text).

To increase the size of the VLP pH 5 dataset and to locate all spikes in the VLP pH 3 data, we used template matching restricted on the VLP membrane surface. First, evenly distributed and oriented pseudo-particles (denoted here as 'seeds') were created on the surface of the particles [49,50] using Dynamo *tomoview* function using an ellipsoidal or a spherical surface model [46]. Seeds were defined with a seed-to-seed spacing of 24 pixels, resulting on average 430 and 340 seeds for pH 5 and pH 3 VLPs, respectively. As the seeds can correspond both to membrane areas that have spikes (true positives) and to areas that are devoid of spikes (false positives), these need to be differentiated. We thus used two templates in a multi-reference refinement run. First, a low-pass filtered cylindrically averaged spike, generated from the earlier VLP pH 5 average, served as a template for the true positives. Second, a spike-free membrane model served as a template for the false positives. The template that correlated the best determined whether the sub-tomogram in the given 'seed' position was a true or false positive. Shifts of 20 pixels were allowed lateral to the membrane. Unique spike sub-tomograms were identified by removing seeds that were too close to another seed. The included sub-tomograms were then subjected to structure refinement similar to the manually picked spikes. No spikes selected for the refinement were excluded after this stage.

The positions of the seeds were then iteratively refined, similarly as described above for manually picked spikes. Low-pass filtered cylindrical average of the densities extracted at seed positions was used as an initial template. The seeds were allowed mainly to shift only in the direction normal to the membrane. A smoothness criterion was implemented as a custom Dynamo function to weight these shifts by the shifts of the neighboring seeds. Similarly to motion correction of particles in single particle processing [51], the weight of each neighbor was adjusted by Gaussian function taking into account the distance to the seed in question, the neighbors farther away from the seed contributing less than the neighbors closer to the seed.

The resolution of the final maps was estimated by FSC, by splitting the sub-tomograms into two half datasets and using 0.5 as threshold [52]. As sub-tomograms extracted from the same virion (or VLP) may not be fully independent, each virion (or VLP) contributed sub-tomograms to one of the half datasets, but not both.

Structure analysis

To putatively assign densities to different structural components, the averaged density maps were computationally segmented using Segger [53] in UCSF Chimera [54]. The isosurface threshold in the spike averages was set to match the expected volume of the spike ectodomain (290,000 $Å^3$), calculated from the estimated spike ectodomain trimer mass (230 kDa, including glycosylations) assuming density of 0.81 Da / $Å^3$. For the spike at pH 3 lagging GP1, the volume was matched to correspond to three copies of GP2 (86 kDa; 106,000 $Å^3$). Spike ectodomain densities were then segmented. The volume corresponding to the LASV GP intra-viral segments was measured in UCSF Chimera.

To create a composite model for the LASV virion, sub-tomogram refinement parameters were first converted from Dynamo tables to a STAR file compatible with Jsubtomo[50]. The segmented densities for the spike protomers, membrane and matrix layer were then placed on their original places relative to the virion by using *jsubtomo_create_model.py*. An isosurface representation, also showing the RNP density inside of the virion was rendered in UCSF Chimera.

Fitting of atomic coordinates

We fitted the atomic coordinates of LASV GP1 (PDB:4ZJF) [10] to our VLP pH 5 spike using Segger [53]. Each segment assigned to a GP subunit was split into four smaller segments

(S10A and S10B Fig). Each of the four chains present in the GP1 crystal structure were fitted to the four GP segments. From the total of 1,000 evenly rotated fits, unique fits were created by further optimization and by clustering fits that were less than 5.0 Å and 3.0 degrees apart (S10C Fig).

Supporting Information

S1 Fig. Electron cryotomography of Lassa virus (LASV) and virus-like particles (VLP) harboring GP. 8-nm thick slices through three-dimensional tomographic reconstructions of LASV and VLPs at different pH low-pass filtered to 60-Å resolution are shown. Scale bars, 50 nm.
(TIF)

S2 Fig. Size distribution of LASV and virus-like particles (VLPs) at different pH. Particle size histograms are plotted for the LASV virion and VLPs at different pH. Mean diameter ±standard deviation and the number of measurements (N) are indicated for each sample. The area of each fitted Gaussian distribution was normalized to 1.
(TIF)

S3 Fig. Density distribution of LASV and virus like particle (VLP). Plots showing the average density distribution in LASV and VLP at pH 7. The insets show a slice through the averaged density, from which the plots were calculated. The glycoprotein layer (GP), membrane (M) and matrix layer (Z) are indicated. The GP intra-viral tails are marked with an asterisk. Notice that the peak assigned to the intra-viral tails is present in both LASV and VLP, but whereas the peak assigned to the Z-layer in present in LASV it is absent in VLP as expected. Both plots have been centered to the middle of the lipid bilayer.
(TIF)

S4 Fig. Resolution estimation of the structures. Fourier shell correlation (FSC) is plotted for the LASV GP structure, in addition to the VLP structures at different pH and GP–LAMP1 complex. The resolutions at which the FSC drops below the threshold (0.5, dashed line) are indicated.
(TIF)

S5 Fig. Glycoprotein spikes are distributed randomly on the virions. Histogram of pairwise distances of 13,060 spikes found by template-matching on 48 LASV virions is shown. Each bin of the histogram is 3 nm wide. The smallest pairwise distance (~10 nm) reflects the closest spike-to-spike packing corresponding to the size of the spike. The largest pairwise distance (~170 nm) reflects the size of the largest virions in the data set.
(TIF)

S6 Fig. Comparison of glycoprotein spike structures from fixed LASV and native virus-like particles at pH 7. (A) An isosurface representation is shown for the GP structure derived from fixed virions (LASV) and from virus-like particles (VLP pH 7) from the side and from the top. (B) Fourier shell correlation (FSC) calculated between the spike ectodomain parts of the two structures is plotted. The resolution at which the FSC drops below the threshold (0.25, dashed line) is indicated and represents the resolution up to which the two structures share significant signal above the noise level. The two structures were in good agreement up to 16-Å resolution.
(TIF)

S7 Fig. Purification of human LAMP1 membrane-distal domain and binding to VLPs. (A) SDS-PAGE analysis shows two duplicate fractions (1,2) from LAMP1 purification in addition

to the molecular size marker (M). The theoretical size of the LAMP1 construct used is indicated. The smearing of LAMP1 bands is attributed to heavy N-linked glycosylation. (B) Binding of purified LAMP1 to VLPs was tested at pH 8 and pH 5. Binding at pH 8 was only slightly above the background level whereas significant binding was observed at pH 5. Error bars denote standard deviation. Statistical significance was calculated using unpaired T-test with Welch's correction. Significance $P<0.01$ is indicated with **.
(TIF)

S8 Fig. Western blot of VLP samples at different pH. A molecular size maker (M) is shown and the sizes of the bands are indicated on the left. The sizes corresponding to GP1 and GP2 are indicated on the right. Notice the absence of GP1 band at pH 3.0 and pH 4.0.
(TIF)

S9 Fig. Comparison of LASV spike structure at pH 3 to the crystal structure of GP2 ectodomain in post-fusion conformation. Post-fusion structure of lymphocytic choriomeningitis virus (LCMV) GP2 (PDB:3MKO) is shown as a blue ribbon and was filtered to 17-Å resolution for comparison to the LASV spike structure derived from virus-like particles (VLP) at pH 3. Membranes are shown as schematic representations and labeled.
(TIF)

S10 Fig. Fitting of GP1 atomic coordinates into the GP spike EM density. (A,B) GP spike, determined from VLPs at pH 5.0, is shown from the top (A) and from the side (B). Density segments considered in the fitting are numbered (1–4). Insets show the structure of the GP1 (green; PDB:4ZJF). In addition to the ribbon representation, a surface corresponding to GP1 atomic structure filtered to 16-Å resolution is shown. (C) Box-and-whisker-plots indicating the cross-correlation coefficients for 1,000 evenly rotated fits of GP1 atomic coordinates to the segments 1–3 (segment 4 was omitted as it was embedded in the lipid bilayer). The four duplicates (labeled A to D) correspond to the four different chains in the GP1 in the crystal (PDB:4ZJF) and gave similar cross-correlation coefficients as expected. Circles indicate the unique optimized fits. The final best fit with the highest cross-correlation coefficient (0.860) is indicated with an arrow.
(TIF)

S11 Fig. Accuracy of tomographic tilt series alignment. Fiducial alignment error mean is plotted for every tilt series in pixels. Most of the tilt series had an alignment error close to 1 pixel.
(TIF)

S12 Fig. Contrast transfer function correction of tomographic tilt series. (A) Amplitude is plotted as a function of spatial frequency for one image extracted from a tilt series. The contrast transfer function (CTF, purple) was fitted to the radial power spectrum (cyan) calculated from several tiles extracted from each tilted image. The frequency range used for fitting is indicated. (B) Defocus, estimated from the fitted CTF illustrated in *A*, is plotted for all the tilts in one tilt series, collected at nominal defocus of 3.5 μm (dashed line). Positive values denote underfocus. Error bars represent minimized error between the actual and fitted curve for each tilt, calculated for the range shown in *A*.
(TIF)

S13 Fig. Effect of the starting model and resolution evolution in refinement of the glycoprotein spike. (A) The glycoprotein spike structure derived from virus like particles at pH 7 is shown from the side and from the top both before (C1) and after (C3) applying three-fold symmetry. Note that the trimeric nature of the spike is evident even before C3 symmetry is applied

(C1). (B) In addition to using an unbiased starting model with cylindrical averaging (C360), no symmetry (C1) or with the correct symmetry (C3), we tested the effect of wrong starting model symmetry (red dashed boxes) on the resulting average. Results are shown after running 12 iterations of refinement without imposing any symmetry and limiting the resolution to 29-Å resolution. The three-fold, trimeric appearance of the spike was retained in all cases. (C) Four different stages of the refinement are shown. Asymmetric, low-resolution starting model (iteration 0) was used to refine the structure first without symmetry (iteration 4). After the three-fold symmetry had become evident, it was applied (iteration 5), and refinement was run to reach the final resolution of 14 Å (iteration 20).
(TIF)

S14 Fig. Comparison of two different refinement strategies. (A) In addition to the standard Dynamo 'push-back' refinement strategy, GP averages from the virions were calculated using a custom 'gold-standard' refinement strategy. The final results from both refinements are shown from the top and from the side. (B) Fourier shell correlation (FSC) is plotted for averages shown in *A*. Both strategies indicated similar nominal resolution; push-back (FSC = 0.5 threshold) 13.6 Å and gold-standard (FSC = 0.143 threshold) 12.3 Å. However, the lower frequencies were stronger in the map calculated using the push-back strategy (shaded area) providing a better quality map.
(TIF)

S1 Text. Supplementary methods. Further detail on the refinement of the sub-tomogram averages.
(DOCX)

Acknowledgments

Daniel Castaño-Díez is thanked for Dynamo support and Marie-Claude Georges-Courbot for providing mouse monoclonal LASV antibody. We also thank Petra Neubauer-Rädel for expert technical assistance and Gotthard Ludwig and Michael Schmidt for technical support with biosafety level 4 procedures. We would like to acknowledge the use of University of Oxford Advanced Research Computing facilities (http://dx.doi.org/10.5281/zenodo.22558).

Author Contributions

Conceived and designed the experiments: SL WG TAB TS JTH. Contributed reagents/materials/analysis tools: CAS. Wrote the paper: SL TAB TS JTH. Prepared the virus and VLP samples: SKF TS KS SL. Purified the virus and VLP samples: SL ZS MLP. Prepared and purified the LAMP1 sample: RP. Collected and analyzed the EM data: SL ZS MLP.

References

1. McCormick JB, King IJ, Webb PA, Johnson KM, O'Sullivan R, Smith ES, et al. A case-control study of the clinical diagnosis and course of Lassa fever. J Infect Dis. 1987; 155: 445–455. PMID: 3805772
2. Eichler R, Lenz O, Strecker T, Garten W. Signal peptide of Lassa virus glycoprotein GP-C exhibits an unusual length. FEBS Lett. 2003; 538: 203–206. PMID: 12633879
3. Igonet S, Vaney M-C, Vonhrein C, Bricogne G, Stura EA, Hengartner H, et al. X-ray structure of the arenavirus glycoprotein GP2 in its postfusion hairpin conformation. P Natl Acad Sci Usa. 2011; 108: 19967–19972.
4. Lenz O, Meulen ter J, Klenk HD, Seidah NG, Garten W. The Lassa virus glycoprotein precursor GP-C is proteolytically processed by subtilase SKI-1/S1P. P Natl Acad Sci Usa. 2001; 98: 12701–12705.

5. Parsy M-L, Harlos K, Huiskonen JT, Bowden TA. Crystal structure of Venezuelan hemorrhagic fever virus fusion glycoprotein reveals a class 1 postfusion architecture with extensive glycosylation. J Virol. 2013; 87: 13070–13075. doi: 10.1128/JVI.02298-13 PMID: 24049182
6. Eichler R, Lenz O, Strecker T, Eickmann M, Klenk HD, Garten W. Identification of Lassa virus glycoprotein signal peptide as a trans-acting maturation factor. EMBO Rep. 2003; 4: 1084–1088. PMID: 14555961
7. York J, Romanowski V, Lu M, Nunberg JH. The signal peptide of the Junín arenavirus envelope glycoprotein is myristoylated and forms an essential subunit of the mature G1-G2 complex. J Virol. 2004; 78: 10783–10792. PMID: 15367645
8. Bederka LH, Bonhomme CJ, Ling EL, Buchmeier MJ. Arenavirus stable signal peptide is the keystone subunit for glycoprotein complex organization. MBio. 2014; 5: e02063. doi: 10.1128/mBio.02063-14 PMID: 25352624
9. Schlie K, Maisa A, Lennartz F, Ströher U, Garten W, Strecker T. Characterization of Lassa virus glycoprotein oligomerization and influence of cholesterol on virus replication. J Virol. 2010; 84: 983–992. doi: 10.1128/JVI.02039-09 PMID: 19889753
10. Cohen-Dvashi H, Cohen N, Israeli H, Diskin R. Molecular mechanism for LAMP1 recognition by Lassa Virus. J Virol. 2015.
11. Jae LT, Raaben M, Herbert AS, Kuehne AI, Wirchnianski AS, Soh TK, et al. Lassa virus entry requires a trigger-induced receptor switch. Science. 2014; 344: 1506–1510. doi: 10.1126/science.1252480 PMID: 24970085
12. Bowden TA, Crispin M, Graham SC, Harvey DJ, Grimes JM, Jones EY, et al. Unusual molecular architecture of the machupo virus attachment glycoprotein. J Virol. 2009; 83: 8259–8265. doi: 10.1128/JVI.00761-09 PMID: 19494008
13. Abraham J, Corbett KD, Farzan M, Choe H, Harrison SC. Structural basis for receptor recognition by New World hemorrhagic fever arenaviruses. Nat Struct Mol Biol. 2010; 17: 438–444. doi: 10.1038/nsmb.1772 PMID: 20208545
14. Eschli B, Quirin K, Wepf A, Weber J, Zinkernagel R, Hengartner H. Identification of an N-terminal trimeric coiled-coil core within arenavirus glycoprotein 2 permits assignment to class I viral fusion proteins. J Virol. 2006; 80: 5897–5907. PMID: 16731928
15. Neuman BW, Adair BD, Burns JW, Milligan RA, Buchmeier MJ, Yeager M. Complementarity in the supramolecular design of arenaviruses and retroviruses revealed by electron cryomicroscopy and image analysis. J Virol. 2005; 79: 3822–3830. PMID: 15731275
16. Lyumkis D, Julien J-P, de Val N, Cupo A, Potter CS, Klasse P-J, et al. Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer. Science. 2013; 342: 1484–1490. doi: 10.1126/science.1245627 PMID: 24179160
17. Tran EEH, Simmons JA, Bartesaghi A, Shoemaker CJ, Nelson E, White JM, et al. Spatial localization of the ebola virus glycoprotein mucin-like domain determined by cryo-electron tomography. J Virol. 2014; 88: 10958–10962. doi: 10.1128/JVI.00870-14 PMID: 25008940
18. York J, Nunberg JH. Intersubunit interactions modulate pH-induced activation of membrane fusion by the Junin virus envelope glycoprotein GPC. J Virol. 2009; 83: 4121–4126. doi: 10.1128/JVI.02410-08 PMID: 19224989
19. Schrempf S, Froeschke M, Giroglou T, Laer von D, Dobberstein B. Signal peptide requirements for lymphocytic choriomeningitis virus glycoprotein C maturation and virus infectivity. J Virol. 2007; 81: 12515–12524. PMID: 17804515
20. Eichler R, Lenz O, Strecker T, Eickmann M, Klenk HD, Garten W. Lassa virus glycoprotein signal peptide displays a novel topology with an extended endoplasmic reticulum luminal region. J Biol Chem. 2004; 279: 12293–12299. PMID: 14709548
21. Agnihothram SS, York J, Trahey M, Nunberg JH. Bitopic membrane topology of the stable signal peptide in the tripartite Junín virus GP-C envelope glycoprotein complex. J Virol. 2007; 81: 4331–4337. PMID: 17267481
22. Schlie K, Maisa A, Freiberg F, Groseth A, Strecker T, Garten W. Viral protein determinants of Lassa virus entry and release from polarized epithelial cells. J Virol. 2010; 84: 3178–3188. doi: 10.1128/JVI.02240-09 PMID: 20071570
23. Fehling SK, Lennartz F, Strecker T. Multifunctional nature of the arenavirus RING finger protein Z. Viruses. 2012; 4: 2973–3011. doi: 10.3390/v4112973 PMID: 23202512
24. Capul AA, Perez M, Burke E, Kunz S, Buchmeier MJ, la Torre de JC. Arenavirus Z-glycoprotein association requires Z myristoylation but not functional RING or late domains. J Virol. 2007; 81: 9451–9460. PMID: 17581989

25. Cao W, Henry MD, Borrow P, Yamada H, Elder JH, Ravkov EV, et al. Identification of alpha-dystroglycan as a receptor for lymphocytic choriomeningitis virus and Lassa fever virus. Science. 1998; 282: 2079–2081. PMID: 9851928

26. Shimojima M, Ströher U, Ebihara H, Feldmann H, Kawaoka Y. Identification of cell surface molecules involved in dystroglycan-independent Lassa virus cell entry. J Virol. 2012; 86: 2067–2078. doi: 10.1128/JVI.06451-11 PMID: 22156524

27. Goncalves A-R, Moraz M-L, Pasquato A, Helenius A, Lozach P-Y, Kunz S. Role of DC-SIGN in Lassa virus entry into human dendritic cells. J Virol. 2013; 87: 11504–11515. doi: 10.1128/JVI.01893-13 PMID: 23966408

28. Klewitz C, Klenk HD, Meulen ter J. Amino acids from both N-terminal hydrophobic regions of the Lassa virus envelope glycoprotein GP-2 are critical for pH-dependent membrane fusion and infectivity. J Gen Virol. 2007; 88: 2320–2328. PMID: 17622638

29. Wilke S, Krausze J, Büssow K. Crystal structure of the conserved domain of the DC lysosomal associated membrane protein: implications for the lysosomal glycocalyx. BMC Biol. 2012; 10: 62. doi: 10.1186/1741-7007-10-62 PMID: 22809326

30. Bowden TA, Jones EY, Stuart DI. Cells under siege: viral glycoprotein interactions at the cell surface. J Struct Biol. 2011; 175: 120–126. doi: 10.1016/j.jsb.2011.03.016 PMID: 21440638

31. Pasqual G, Rojek JM, Masin M, Chatton J-Y, Kunz S. Old world arenaviruses enter the host cell via the multivesicular body and depend on the endosomal sorting complex required for transport. PLoS Pathog. 2011; 7: e1002232. doi: 10.1371/journal.ppat.1002232 PMID: 21931550

32. Branco LM, Grove JN, Geske FJ, Boisen ML, Muncy IJ, Magliato SA, et al. Lassa virus-like particles displaying all major immunological determinants as a vaccine candidate for Lassa hemorrhagic fever. Virol J. 2010; 7: 279. doi: 10.1186/1743-422X-7-279 PMID: 20961433

33. Voss JE, Vaney M-C, Duquerroy S, Vonrhein C, Girard-Blanc C, Crublet E, et al. Glycoprotein organization of Chikungunya virus particles revealed by X-ray crystallography. Nature. 2010; 468: 709–712. doi: 10.1038/nature09555 PMID: 21124458

34. Messina EL, York J, Nunberg JH. Dissection of the role of the stable signal peptide of the arenavirus envelope glycoprotein in membrane fusion. J Virol. 2012; 86: 6138–6145. doi: 10.1128/JVI.07241-11 PMID: 22438561

35. Harrison SC. Viral membrane fusion. Nat Struct Mol Biol. 2008; 15: 690–698. doi: 10.1038/nsmb.1456 PMID: 18596815

36. Hetzel U, Sironen T, Laurinmäki P, Liljeroos L, Patjas A, Henttonen H, et al. Isolation, identification, and characterization of novel arenaviruses, the etiological agents of boid inclusion body disease. J Virol. 2013; 87: 10918–10935. doi: 10.1128/JVI.01123-13 PMID: 23926354

37. Gallaher WR, DiSimone C, Buchmeier MJ. The viral transmembrane superfamily: possible divergence of Arenavirus and Filovirus glycoproteins from a common RNA virus ancestor. BMC Microbiol. 2001; 1: 1. PMID: 11208257

38. Lenz O, Meulen ter J, Feldmann H, Klenk HD, Garten W. Identification of a novel consensus sequence at the cleavage site of the Lassa virus glycoprotein. J Virol. 2000; 74: 11418–11421. PMID: 11070044

39. Aricescu AR, Lu W, Jones EY. A time- and cost-efficient system for high-level protein production in mammalian cells. Acta Crystallogr D Biol Crystallogr. 2006; 62: 1243–1250. PMID: 17001101

40. Elbein AD, Tropea JE, Mitchell M, Kaushal GP. Kifunensine, a potent inhibitor of the glycoprotein processing mannosidase I. J Biol Chem. 1990; 265: 15599–15605. PMID: 2144287

41. Tivol WF, Briegel A, Jensen GJ. An improved cryogen for plunge freezing. Microsc Microanal. 2008; 14: 375–379. doi: 10.1017/S1431927608080781 PMID: 18793481

42. Mastronarde DN. Automated electron microscope tomography using robust prediction of specimen movements. J Struct Biol. 2005; 152: 36–51. 7 PMID: 16182563

43. Li X, Mooney P, Zheng S, Booth CR, Braunfeld MB, Gubbens S, et al. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. Nat Methods. 2013; 10: 584–590. doi: 10.1038/nmeth.2472 PMID: 23644547

44. Kremer J, Mastronarde D, McIntosh J. Computer visualization of three-dimensional image data using IMOD. J Struct Biol. 1996; 116: 71–76. PMID: 8742726

45. Xiong Q, Morphew MK, Schwartz CL, Hoenger AH, Mastronarde DN. CTF determination and correction for low dose tomographic tilt series. J Struct Biol. 2009; 168: 378–387. doi: 10.1016/j.jsb.2009.08.016 PMID: 19732834

46. Castaño-Díez D, Kudryashev M, Arheit M, Stahlberg H. Dynamo: a flexible, user-friendly development tool for subtomogram averaging of cryo-EM data in high-performance computing environments. J Struct Biol. 2012; 178: 139–151. doi: 10.1016/j.jsb.2011.12.017 PMID: 22245546

47. Heymann JB, Belnap DM. Bsoft: image processing and molecular modeling for electron microscopy. J Struct Biol. 2007; 157: 3–18. PMID: 17011211

48. Huiskonen JT, Hepojoki J, Laurinmäki P, Vaheri A, Lankinen H, Butcher SJ, et al. Electron cryotomography of Tula hantavirus suggests a unique assembly paradigm for enveloped viruses. J Virol. 2010; 84: 4889–4897. doi: 10.1128/JVI.00057-10 PMID: 20219926

49. Bowden TA, Bitto D, McLees A, Yeromonahos C, Elliott RM, Huiskonen JT. Orthobunyavirus ultrastructure and the curious tripodal glycoprotein spike. PLoS Pathog. 2013; 9: e1003374. doi: 10.1371/journal.ppat.1003374 PMID: 23696739

50. Huiskonen JT, Parsy M-L, Li S, Bitto D, Renner M, Bowden TA. Averaging of Viral Envelope Glycoprotein Spikes from Electron Cryotomography Reconstructions using Jsubtomo. J Vis Exp. 2014.

51. Scheres SH. Beam-induced motion correction for sub-megadalton cryo-EM particles. Elife. 2014; 3: e03665. doi: 10.7554/eLife.03665 PMID: 25122622

52. Rosenthal PB, Henderson R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. J Mol Biol. 2003; 333: 721–745. PMID: 14568533

53. Pintilie GD, Zhang J, Goddard TD, Chiu W, Gossard DC. Quantitative analysis of cryo-EM density map segmentation by watershed and scale-space filtering, and fitting of structures by alignment to regions. J Struct Biol. 2010; 170: 427–438. doi: 10.1016/j.jsb.2010.03.007 PMID: 20338243

54. Goddard TD, Huang CC, Ferrin TE. Visualizing density maps with UCSF Chimera. J Struct Biol. 2007; 157: 281–287. PMID: 16963278

Appendix B

```
GABAB1a LLLLLLAPLFLRPPGAGGAQTPNATSEGCQIIHPPWEGGIRYRGLTRDQVKAINFLPVDY 60
GABAB1b ------------------------------------------------------------
GABAB2  ------------------------------------------------------------


GABAB1a EIEYVCRGEREVVGPKVRKCLANGSWTDMDTPSRCVRICSKSYLTLENGKVFLTGGDLPA 120
GABAB1b ----------------------------------------------MGPGAPF-------- 7
GABAB2  ---------------------------MASPRSS-------------G----QPGPPPP 15
                                                        *


GABAB1a L--DGARVDFRCDPDFHLVGSSRSICSQGQWSTPKPHCQVNRTPHSERRAVYIGALFPMS 178
GABAB1b -----ARVGWPLPLLVVMAAGVAPVWASHSPHLPRPHSRVPPHPSSERRAVYIGALFPMS 62
GABAB2  PPPPPARLL------LLLLLPLLLPLAPGAWGWARGA---PRPPP-SSPPLSIMGLMPLT 65
             **:       . :        :       :        *  .   : * .*:*::


GABAB1a GGW---PGGQACQPAVEMALEDVNSRRDILPDYELKLIHHDSKCDPGQATKYLYELLYND 235
GABAB1b GGW---PGGQACQPAVEMALEDVNSRRDILPDYELKLIHHDSKCDPGQATKYLYELLYND 119
GABAB2  KEVAKGSIGRGVLPAVELAIEQIRNE-SLLRPYFLDLRLYDTECDNAKGLKAFYDAIKYG 124
                *:.  ****:*:*::... .:*  * *.*  :*::** .:. * :*: :  .


GABAB1a PIKIILM-PGCSSVSTLVAEAARMWNLIVLSYGSSSPALSNRQRFPTFFRTHPSATLHNP 294
GABAB1b PIKIILM-PGCSSVSTLVAEAARMWNLIVLSYGSSSPALSNRQRFPTFFRTHPSATLHNP 178
GABAB2  PNHLMVFGGVCPSVTSIIAESLQGWNLVQLSFAATTPVLADKKKYPYFFRTVPSDNAVNP 184
        * :::::   * **::::**: : ***: **:.:::*.*::::::* **** ** .  **


GABAB1a TRVKLFEKWGWKKIATIQQTTEVFTSTLDDLEERVKEAGIEITFRQSFFSDPAVPVKNLK 354
GABAB1b TRVKLFEKWGWKKIATIQQTTEVFTSTLDDLEERVKEAGIEITFRQSFFSDPAVPVKNLK 238
GABAB2  AILKLLKHYQWKRVGTLTQDVQRFSEVRNDLTGVLYGEDIEISDTESFSNDPCTSVKKLK 244
        : :**:::: **::.*: * .: *:.. :**   :   .***:  :** .**.. **:**


GABAB1a RQDARIIVGLFYETEARKVFCEVYKERLFGKKYVWFLIGWYADNWFKIYDPS---INCTV 411
GABAB1b RQDARIIVGLFYETEARKVFCEVYKERLFGKKYVWFLIGWYADNWFKIYDPS---INCTV 295
GABAB2  GNDVRIILGQFDQNMAAKVFCCAYEENMYGSKYQWIIPGWYEPSWWEQVHTEANSSRCLR 304
         :*.***:* * :. * **** .*:*.::*.** *:: ***  .*::  . .    .*


GABAB1a DEMTEAVEGHITTEIVMLNPANTRSISNMTSQEFVEKLTKRLKRHPEETGGFQEAPLAYD 471
GABAB1b DEMTEAVEGHITTEIVMLNPANTRSISNMTSQEFVEKLTKRLKRHPEETGGFQEAPLAYD 355
GABAB2  KNLLAAMEGYIGVDFEPLSSKQIKTISGKTPQQYEREYN----NKRSGVGPSKFHGYAYD 360
        .::  *:**:* .::  *.  : ::**. * *:: .: .    .: . .*  :    ***


GABAB1a AIWALALALNKTSG--GGGRSGVRLEDFNYNNQTITDQIYRAMNSSSFEGVSGHVVFDAS 529
GABAB1b AIWALALALNKTSG--GGGRSGVRLEDFNYNNQTITDQIYRAMNSSSFEGVSGHVVFDAS 413
GABAB2  GIWVIAKTLQRAMETLHASSRHQRIQDFNYTDHTLGRIILNAMNETNFFGVTGQVVFRN- 419
        .**.:* :*:::     ..    *::****.::*:   * .***.:.* **:*:***


GABAB1a GSRMAWTLIEQLQGGSYKKIGYYDSTKDDLSWS-KTDKWIGGSPPADQTLVIKTFRFLSQ 588
GABAB1b GSRMAWTLIEQLQGGSYKKIGYYDSTKDDLSWS-KTDKWIGGSPPADQTLVIKTFRFLSQ 472
GABAB2  GERMGTIKFTQFQDSREVKVGEYNAVADTLEIINDTIRFQGSEPPKDKTIILEQLRKISL 479
        *.**.   : *:*..   *:* *::. * *.   .* :: *..** *:*:::: :* :*


GABAB1a KLFISVSVLSSLGIVLAVVCLSFNIYNSHVRYIQNSQPNLNNLTAVGCSLALAAVFPLGL 648
GABAB1b KLFISVSVLSSLGIVLAVVCLSFNIYNSHVRYIQNSQPNLNNLTAVGCSLALAAVFPLGL 532
GABAB2  PLYSILSALTILGMIMASAFLFFNIKNRNQKLIKMSSPYMNNLIILGGMLSYASIFLFGL 539
         *:  :*.*: **:::* . * *** * : : *: *.* :***  :*  *: *::* :**


GABAB1a DGYHIGRNQFPFVCQARLWLLGLGFSLGYGSMFTKIWWVHTVFTKKEEKKEWRKTLEPWK 708
GABAB1b DGYHIGRNQFPFVCQARLWLLGLGFSLGYGSMFTKIWWVHTVFTKKEEKKEWRKTLEPWK 592
GABAB2  DGSFVSEKTFETLCTVRTWILTVGYTTAFGAMFAKTWRVHAIFKNVKMKKK---IIKDQK 596
        ** .:..: *  :* .* *:* :*:: .:*:**:* * **::*.: : **:    ::  *
```

Figure B.1 Sequence alignment of $GABA_{B1a}$, $GABA_{B1b}$ and $GABA_{B2.}$ Alignment of the human $GABA_{B1a}$ (Uniport Q9UBS5), $GABA_{B1b}$ (Uniport Q9UBS5 Isoform 1b) and $GABA_{B2}$ (Uniprot ID O75899) sequences was performed using Clustal Omega (Zimmermann, Stephens et al. 2018) (Sievers, Wilm et al. 2011). * marks conserved residues. : and .marks strongly and weakly similar residues.

```
GABAB1a LYATVGLLVGMDVLTLAIWQIVDPLHRTIETFAKEEPKEDIDVSILPQLEHCSSRKMNTW 768
GABAB1b LYATVGLLVGMDVLTLAIWQIVDPLHRTIETFAKEEPKEDIDVSILPQLEHCSSRKMNTW 652
GABAB2  LLVIVGGMLLIDLCILICWQAVDPLRRTVEKYSMEPDPAGRDISIRPLLEHCENTHMTIW 656
        * . ** :: :*:  *  ** ****:**:*.:: *    . *:** * ****.. :*. *
GABAB1a LGIFYGYKGLLLLLGIFLAYETKSVSTEKINDHRAVGMAIYNVAVLCLITAPVTMILSSQ 828
GABAB1b LGIFYGYKGLLLLLGIFLAYETKSVSTEKINDHRAVGMAIYNVAVLCLITAPVTMILSSQ 712
GABAB2  LGIVYAYKGLLMLFGCFLAWETRNVSIPALNDSKYIGMSVYNVGIMCIIGAAVSFLTRDQ 716
        ***.*.*****:*:* ***:**:.**   :** : :**::***.::*:* * *:::  .*

GABAB1a QDAAFAFASLAIVFSSYITLVVLFVPKMRRLITRGEWQSEAQDT----------MKTG-S 877
GABAB1b QDAAFAFASLAIVFSSYITLVVLFVPKMRRLITRGEWQSEAQDT----------MKTG-S 761
GABAB2  PNVQFCIVALVIIFCSTITLCLVFVPKLITLRTNPDAATQNRRFQFTQNQKKEDSKTSTS 776
         :. *.:.:*.*:*.* *** ::****:  * *. :  :: :             **. *

GABAB1a STNNNEEEKSR--LLEKENRELEKIIAEKEERVSELRHQLQSRQQLRSR---RHPP---- 928
GABAB1b STNNNEEEKSR--LLEKENRELEKIIAEKEERVSELRHQLQSRQQLRSR---RHPP---- 812
GABAB2  VTSVNQASTSRLEGLQSENHRLRMKITELDKDLEEVTMQLQDTPEKTTYIKQNHYQELND 836
         *. *: ..**   *:.**:.*.  *:* :: :.*:  ***.  :  :    .*

GABAB1a -----T-PPEPSGGL--------------------------PRG----PPEPPDRLSCDG 952
GABAB1b -----T-PPEPSGGL--------------------------PRG----PPEPPDRLSCDG 836
GABAB2  ILNLGNFTESTDGGKAILKNHLDQNPQLQWNTTEPSRTCKDPIEDINSPEHIQRRLSLQL 896
            .   . .**                            *      * .   *** :

GABAB1a SRVHLLYK------------------------------------ 960
GABAB1b SRVHLLYK------------------------------------ 844
GABAB2  PILHHAYLPSIGGVDASCVSPCVSPTASPRHRHVPPSFRVMVSGL 941
          :*  *
```

Figure B.1 (continued) Sequence alignment of $GABA_{B1a}$, $GABA_{B1b}$ and $GABA_{B2.}$ Alignment of the human $GABA_{B1a}$ (Uniport Q9UBS5), $GABA_{B1b}$ (Uniport Q9UBS5 Isoform 1b) and $GABA_{B2}$ (Uniprot ID O75899) was performed using Clustal Omega (Sievers, Wilm et al. 2011, Zimmermann, Stephens et al. 2018). * marks conserved residues. : and .marks strongly and weakly similar residues.

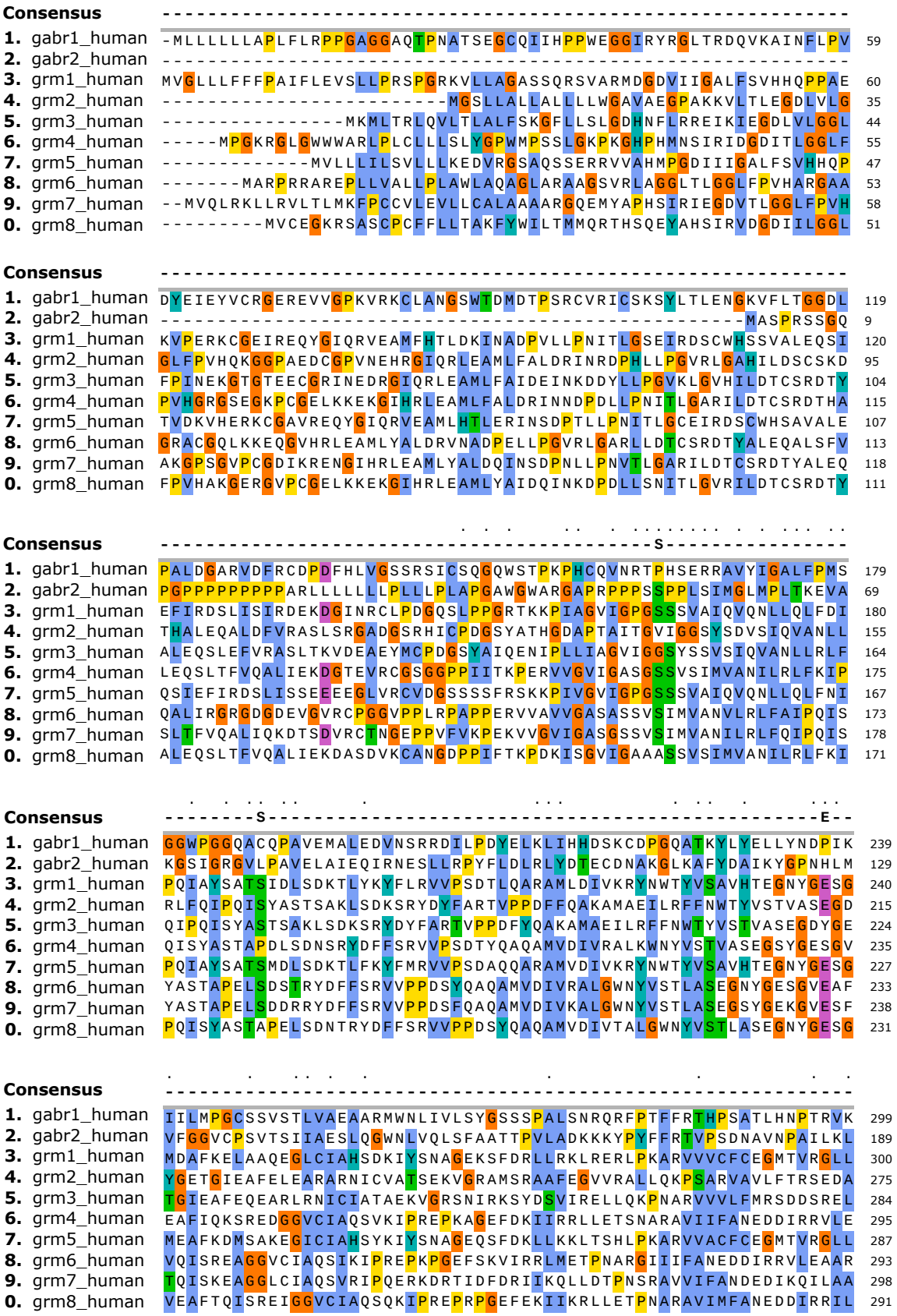


Figure B.2 Sequence alignment of $GABA_{B1}$, $GABA_{B2}$ and $mGluR_{1\text{-}8}$. Alignment was calculated by GPCRdb (Pandy-Szekeres, Munk et al. 2018). Figure was generated using Snapgene.

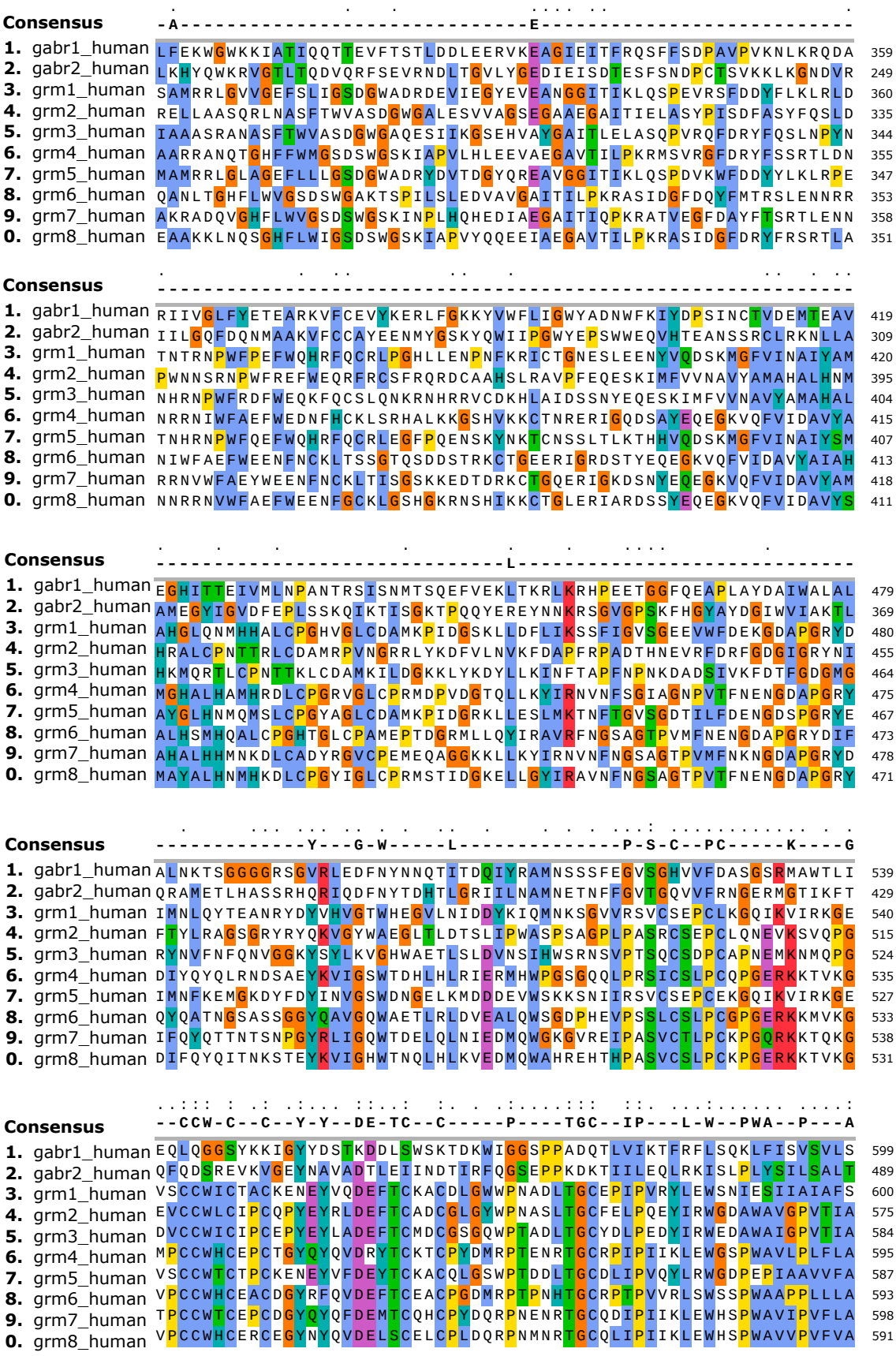


Figure B.2 (continued) Sequence alignment of $GABA_{B1}$, $GABA_{B2}$ and $mGluR_{1\text{-}8}$. Alignment was calculated by GPCRdb (Pandy-Szekeres, Munk et al. 2018). Figure was generated using Snapgene.

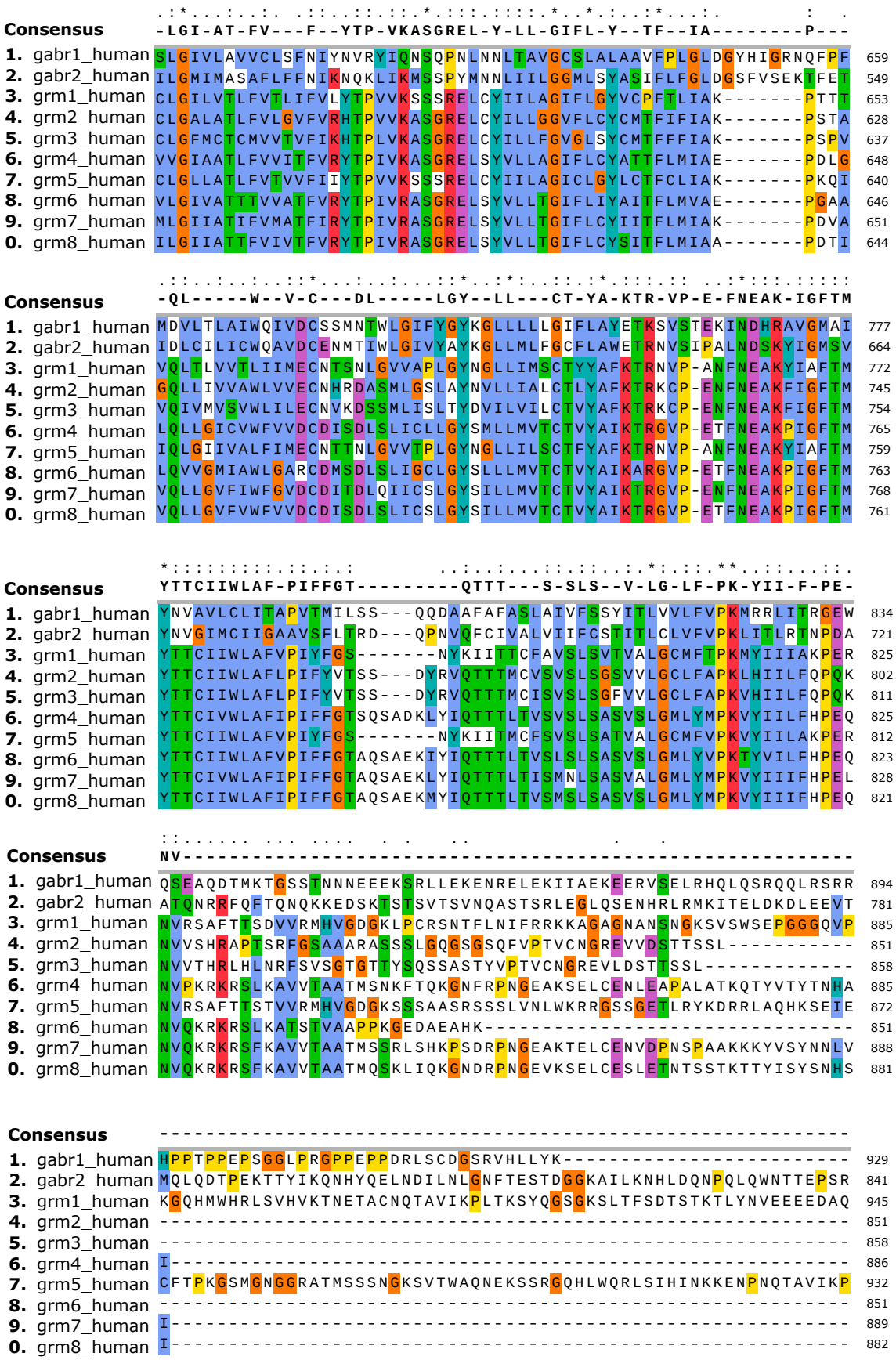


Figure B.2 (continued) Sequence alignment of $GABA_{B1}$, $GABA_{B2}$ and $mGluR_{1\text{-}8}$. Alignment was calculated by GPCRdb (Pandy-Szekeres, Munk et al. 2018). Figure was generated using Snapgene.

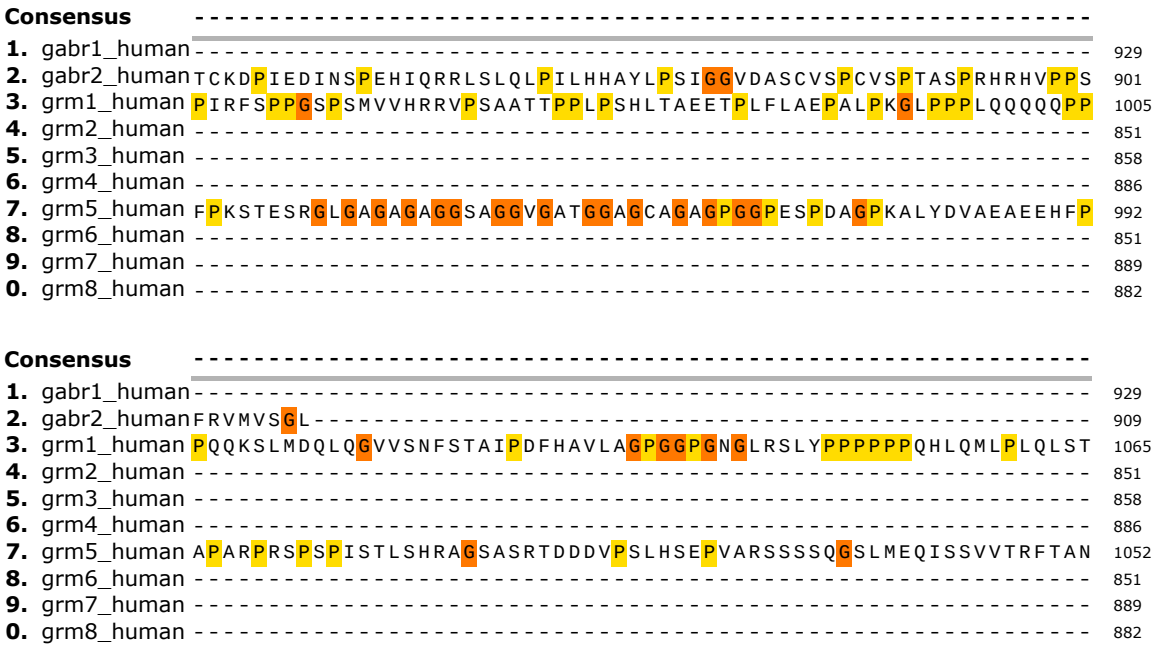


Figure B.2 (continued) Sequence alignment of $GABA_{B1}$, $GABA_{B2}$ and $mGluR_{1-8}$. Alignment was calculated by GPCRdb (Pandy-Szekeres, Munk et al. 2018). Figure was generated using Snapgene.

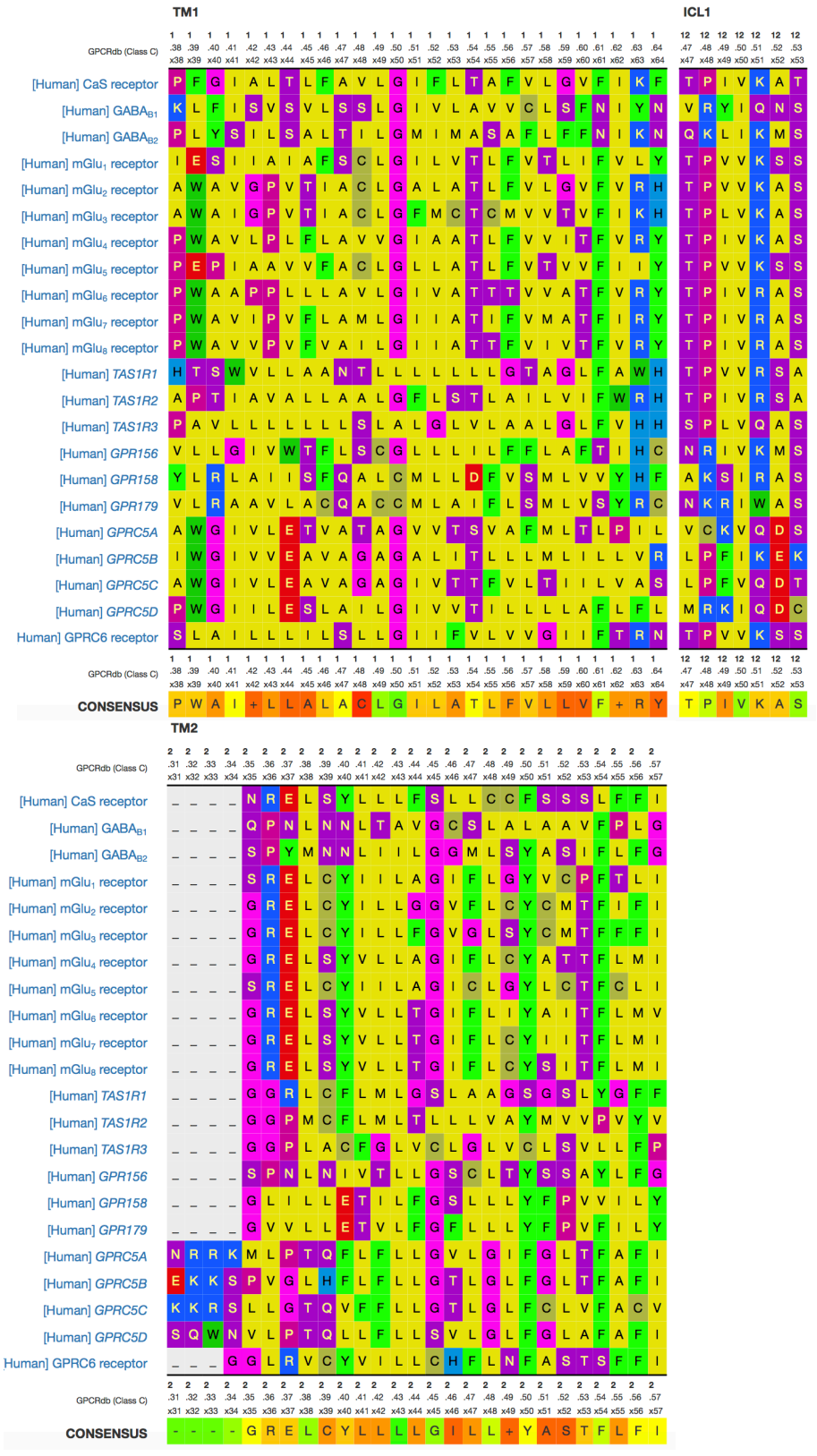


Figure B.3 Structural-based sequence alignment of the transmembrane grain with other Class C GPCRs. Alignment data was retrieved from the GPCRdb (Pandy-Szekeres, Munk et al. 2018) .

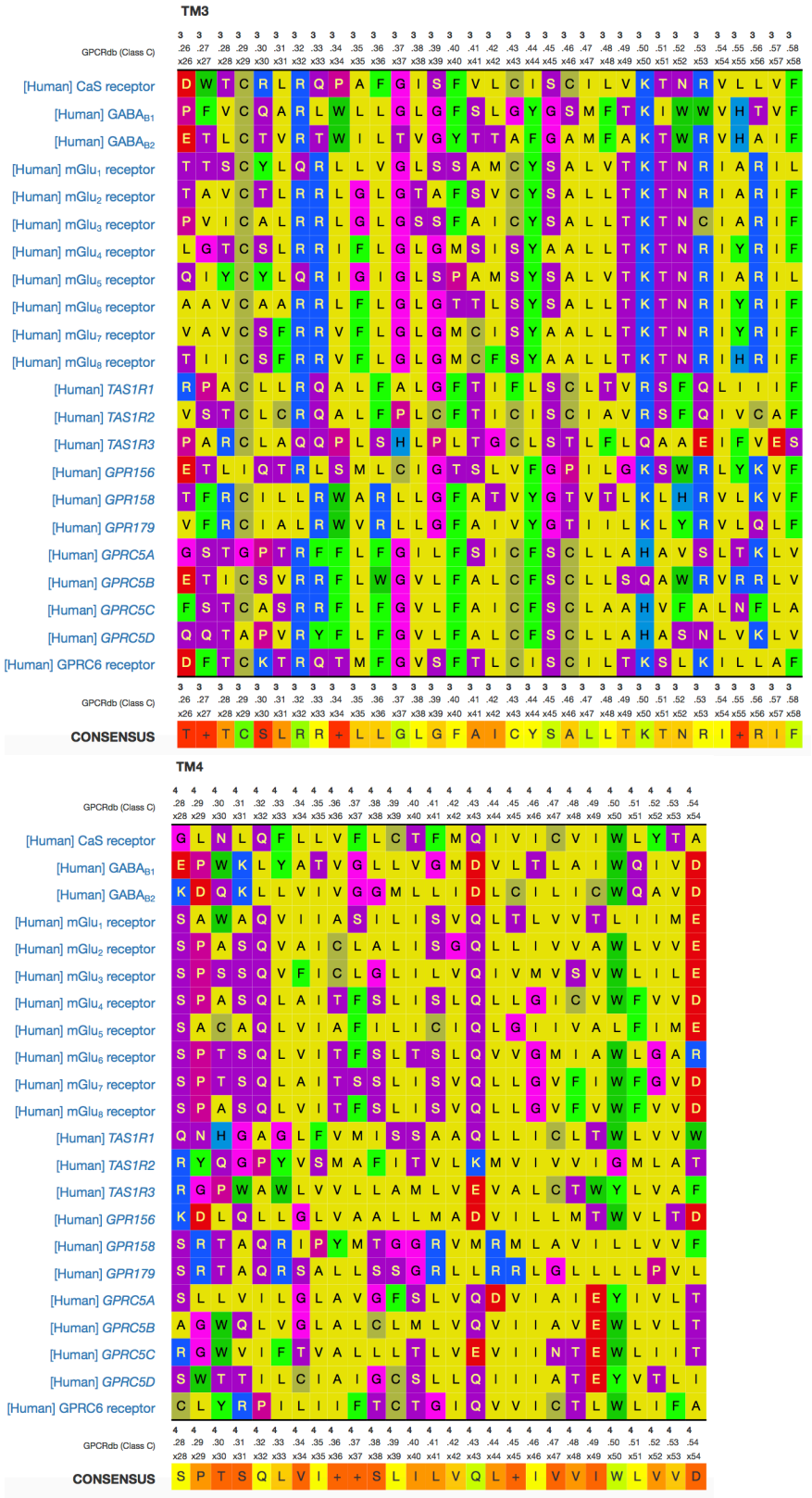


Figure B.3 (continued) Structural-based sequence alignment of the transmembrane grain with other Class C GPCRs. Alignment data was retrieved from the GPCRdb (Pandy-Szekeres, Munk et al. 2018) .



Figure B.3 (continued) Structural-based sequence alignment of the transmembrane grain with other Class C GPCRs. Alignment data was retrieved from the GPCRdb (Pandy-Szekeres, Munk et al. 2018) .



Figure B.3 (continued) Structural-based sequence alignment of the transmembrane grain with other Class C GPCRs. Alignment data was retrieved from the GPCRdb (Pandy-Szekeres, Munk et al. 2018) .

5’GAATTCGCCACCATGCTGCTCTTGCTGCTGTTGGCTCCGCTCTTCCTTAGACCACCCGGCGCCGGAG
GGGCTCAAACCCCAAATGCGACATCCGAAGGGTGCCAGATTATTCACCCTCCCTGGGAGGGCGGCA
TCCGCTATAGAGGCTTGACACGCGACCAGGTAAAGGCAATTAACTTCCTTCCCGTGGACTATGAAAT
CGAATATGTGTGCAGAGGCGAGCGAGAAGTTGTTGGACCCAAAGTCCGCAAATGTCTGGCCAATGG
GTCTTGGACTGATATGGATACACCTTCCCGGTGTGTGCGGATATGCTCTAAGTCTTATCTGACCCTGG
AGAACGGCAAAGTCTTTCTTACCGGCGGTGATTTGCCAGCCCTTGATGGCGCCAGAGTAGACTTTCG
GTGCGACCCTGATTTTCACCTGGTCGGATCAAGCAGATCTATCTGTTCCCAAGGACAGTGGTCTACTC
CAAAGCCGCATTGTCAGGTGAATCGCACTCCACATAGCGAGCGGAGGGCCGTGTATATTGGGGCTT
TGTTCCCCATGAGTGGCGGCTGGCCCGGCGGACAGGCCTGCCAGCCAGCCGTGGAGATGGCCCTGG
AGGATGTTAACTCTCGACGAGATATCTTGCCCGATTATGAGCTGAAACTTATTCATCACGACTCAAAG
TGCGACCCTGGCCAAGCTACGAAGTACCTTTACGAGCTTCTCTACAACGATCCGATCAAGATCATTCT
GATGCCTGGATGCAGCAGCGTGAGTACACTGGTTGCAGAAGCTGCTCGTATGTGGAACTTGATTGTT
CTGAGCTACGGTTCCTCCTCTCCTGCACTCTCCAATCGACAACGTTTCCCCACTTTCTTTCGGACACAC
CCATCCGCGACGCTGCACAATCCAACCCGCGTCAAACTCTTCGAAAAGTGGGGATGGAAGAAAATC
GCTACTATCCAACAGACAACTGAAGTGTTCACCAGCACCCTGGACGACCTGGAGGAAAGGGTGAAG
GAGGCGGGAATCGAGATTACATTCAGACAAAGCTTCTTCAGTGACCCAGCAGTCCCCGTGAAGAACC
TCAAACGCCAAGACGCCCGGATAATAGTCGGGCTGTTTTATGAGACTGAGGCCAGGAAGGTCTTTTG
CGAAGTGTACA3’

Figure B.4 Synthetic $GABA_{B1a}$ subunit sushi domain construct (B1_Sushi_EB). Codon optimised for expression in mammalian cells.

5’ACCGGTAGCGAGCGGAGGGCCGTGTATATTGGGGCTTTGTTCCCCATGAGTGGCGGCTGGCCCG
GCGGACAGGCCTGCCAGCCAGCCGTGGAGATGGCCCTGGAGGATGTTAACTCTCGACGAGATATCT
TGCCCGATTATGAGCTGAAACTTATTCATCACGACTCAAAGTGCGACCCTGGCCAAGCTACGAAGTA
CCTTTACGAGCTTCTCTACAACGATCCGATCAAGATCATTCTGATGCCTGGATGCAGCAGCGTGAGTA
CACTGGTTGCAGAAGCTGCTCGTATGTGGAACTTGATTGTTCTGAGCTACGGTTCCTCCTCTCCTGCA
CTCTCCAATCGACAACGTTTCCCCACTTTCTTTCGGACACACCCATCCGCGACGCTGCACAATCCAACC
CGCGTCAAACTCTTCGAAAAGTGGGGATGGAAGAAAATCGCTACTATCCAACAGACAACTGAAGTG
TTCACCAGCACCCTGGACGACCTGGAGGAAAGGGTGAAGGAGGCGGGAATCGAGATTACATTCAG
ACAAAGCTTCTTCAGTGACCCAGCAGTCCCCGTGAAGAACCTCAAACGCCAAGACGCCCGGATAATA
GTCGGGCTGTTTTATGAGACTGAGGCCAGGAAGGTCTTTTGCGAAGTGTACAAGGAGCGGTTGTTT
GGTAAGAAGTACGTTTGGTTTCTGATCGGGTGGTATGCGGACAATTGGTTCAAAATTTACGACCCGT
CTATCAACTGCACCGTCGATGAAATGACAGAAGCCGTCGAAGGACACATAACAACCGAGATCGTCAT
GCTCAACCCCGCTAATACCAGATCCATATCAAACATGACGTCACAGGAATTTGTCGAAAAGCTGACT
AAGCGCCTGAAGAGGCATCCCGAAGAAACTGGTGGGTTTCAGGAAGCCCCTTTGGCTTATGACGCA
ATTTGGGCCTTGGCACTCGCTCTGAACAAAACATCAGGAGGTGGAGGTAGGTCCGGGGTGAGGCTG
GAGGATTTCAATTACAACAATCAGACCATCACCGATCAGATCTATAGGGCTATGAACAGCAGCTCCTT
CGAAGGTGTGAGCGGCCATGTCGTCTTTGATGCCTCCGGGTCTAGGATGGCCTGGACACTGATAGA
GCAGCTGCAGGGTGGCAGTTACAAGAAAATTGGTTATTACGACTCTACTAAAGACGATCTCTCATGG
TCCAAGACCGACAAGTGGATTGGCGGCTCTCCTCCAGCCGATCAGACCCTGGTGATCAAAACATTTA
GATTCCTGAGCCAGAAGCTGTTTATAAGTGTGAGCGTGCTGAGTTCCCTCGGAATAGTTCTCGCAGT
CGTTTGCCTGAGCTTTAACATCTACAACTCACACGTGCGATACATTCAGAACTCTCAGCCTAACCTTAA
TAATTTGACCGCTGTAGGCTGTAGTCTCGCACTGGCAGCCGTGTTTCCCCTGGGCCTGGACGGCTAC
CACATTGGCAGGAATCAGTTTCCTTTTGTGTGTCAGGCTAGACTCTGGCTCCTTGGGCTGGGTTTCTC
ACTGGGTTATGGGAGTATGTTCACCAAGATTTGGTGGGTGCACACAGTGTTTACAAAGAAAGAGGA
GAAGAAAGAGTGGCGCAAGACTCTGGAACCTTGGAAACTGTATGCGACCGTGGGGCTGTTGGTGG
GAATGGACGTACTTACCCTCGCCATCTGGCAGATCGTAGACCCATTGCATCGGACGATCGAGACATT
CGCTAAAGAGGAGCCTAAAGAGGACATTGACGTGAGCATCCTGCCGCAGCTGGAGCACTGTAGCTC
CCGTAAAATGAATACATGGTTGGGAATATTCTACGGATACAAAGGCCTGCTCCTTCTCCTGGGGATC
TTCCTGGCCTACGAAACGAAATCTGTAGCAGACCTTGAAGATAATTGGGAGACACTGAATGACAATC
TGAAAGTGATCGAGAAAGCCGACAATGCTGCCCAAGTGAAAGACGCGCTTACCAAAATGAGAGCCG
CAGCACTGGATGCACAGAAAGCTACTCCACCAAAGCTGGAGGATAAGAGCCCCGACTCACCCGAAA
TGAAGGATTTCCGCCACGGATTTGATATTCTGGTTGGGCAGATCGATGACGCCCTGAAACTGGCCAA
CGAGGGGAAGGTGAAGGAAGCTCAGGCCGCTGCGGAGCAGCTTAAGACGACCCGCAATGCCTATA
TCCAGAAATACCTGAGCACAGAGAAGATCAATGATCATAGGGCGGTAGGAATGGCCATATACAACG
TGGCCGTTCTGTGTCTCATTACTGCTCCCGTCACTATGATTCTCTCCAGTCAACAGGATGCCGCATTTG
CTTTCGCATCACTGGCAATCGTTTTCTCCAGCTATATCACCCTCGTGGTTCTGTTCGTACCTAAGATGC
GACGGCTGATTACCAGAGGGGAATGGCAAAGCGAGGCCCAGGATACCATGAAAACGGGCAGTTCA
ACTAACAACAACGAAGAGGAAAAGAGCAGACTCCTGGAGAAGGAGAACAGGGAGCTGGAAAAGAT
CATTGCAGAGAAGGAGGAGCGGGTGAGCGAACTGCGGCATCAACTGCAGAGTCGTCAGGGTACC
3 ’

Figure.B.5 Synthetic $GABA_{B1b}$ subunit constructs (B1_noS_BRIL_AK). Codon optimised for expression in mammalian cells.

5’GAATTCGCCACCATGGCTTCTCCAAGAAGCTCCGGGCAACCTGGACCTCCCCCCCCACCACCTCCG
CCACCCGCAAGGCTGCTCTTGTTGCTCCTTCTTCCACTGCTTCTGCCCCTGGCCCCGGGTGCATGGGG
ATGGGCTCGTGGGGCCCCGCGCCCTCCACCTAGTAGTCCTCCCCTCAGTATCATGGGACTGATGCCTC
TGACGAAAGAAGTGGCAAAAGGCTCAATAGGCAGGGGCGTTCTGCCCGCTGTTGAACTCGCCATTG
AACAAATCCGCAACGAGAGTCTGCTGCGCCCTTACTTTCTGGATCTGCGGCTGTACGACACTGAATG
CGACAACGCCAAGGGGCTGAAAGCCTTTTATGACGCAATCAAATATGGGCCTAATCACTTGATGGTC
TTTGGAGGAGTGTGTCCCTCAGTCACTAGCATCATTGCCGAGTCACTGCAGGGATGGAATCTGGTGC
AACTTTCCTTCGCCGCCACAACTCCGGTCCTGGCCGACAAGAAGAAATACCCATACTTCTTTCGCACA
GTGCCCAGCGACAACGCCGTCAATCCCGCAATTCTCAAGCTGCTCAAACACTACCAGTGGAAACGGG
TCGGAACGCTCACCCAAGATGTGCAGCGGTTCAGCGAGGTGCGGAATGACCTTACCGGAGTCCTGT
ATGGTGAAGATATCGAGATCTCCGATACGGAATCATTCTCTAATGATCCATGTACATCAGTTAAGAA
GCTGAAGGGCAATGATGTGCGGATTATTCTGGGCCAATTTGACCAGAACATGGCTGCTAAGGTGTTC
TGCTGTGCGTATGAGGAGAATATGTATGGTAGTAAGTACCAGTGGATCATTCCCGGCTGGTACGAG
CCGTCTTGGTGGGAACAGGTGCATACCGAAGCTAACAGCTCCCGATGCCTTAGGAAGAACTTGCTTG
CAGCCATGGAGGGCTACATCGGGGTTGACTTTGAACCCCTCTCTTCTAAGCAGATAAAGACTATTTCA
GGGAAAACACCCCAGCAATACGAAAGAGAGTACAACAACAAACGTAGCGGGGTAGGTCCCAGCAA
GTTCCACGGGTATGCCTATGATGGAATTTGGGTGATCGCCAAAACTCTGCAGAGAGCCATGGAGACT
CTTCATGCGTCCTCACGTCATCAGCGAATCCAGGACTTCAACTACACAGATCACACACTGGGCAGAAT
CATTCTCAATGCCATGAACGAAACCAACTTCTTCGGCGTGACAGGTCAGGTGGTCTTTCGAAACGGC
GAGAGAATGGGAACGATCAAATTCACACAGTTTCAGGACAGTAGAGAGGTGAAGGTTGGAGAGTA
CAACGCGGTAGCTGACACACTGGAAATAATCAATGACACCATTCGCTTCCAGGGGTCTGAACCACCA
AAGGATAAAACCATTATACTGGAGCAGTTGAGGAAAATAAGCTTGCCTCTGTACTCTATTCTCTCAGC
CCTTACTATCCTGGGCATGATTATGGCTTCCGCTTTTCTGTTCTTTAACATTAAGAACCGGAATCAGAA
ACTGATCAAAATGAGCAGCCCATATATGAATAACCTGATTATCCTGGGAGGCATGTTGAGCTACGCG
AGCATCTTCCTTTTCGGTCTGGATGGGTCCTTCGTAAGCGAGAAAACCTTCGAGACTCTCTGTACCGT
TAGGACGTGGATTCTGACCGTGGGCTATACTACTGCCTTTGGTGCTATGTTCGCCAAAACATGGAGG
GTCCACGCCATCTTTAAGAATGTGAAAATGAAGAAGAAGATCATAAAGGATCAGAAACTGCTCGTG
ATCGTTGGAGGGATGTTGCTGATTGACCTGTGCATTCTGATTTGTTGGCAGGCAGTTGACCCGCTGA
GGCGTACAGTGGAGAAGTATTCTATGGAGCCTGATCCCGCTGGTCGGGACATCTCCATCAGACCATT
GCTGGAGCACTGCGAAAATACACATATGACCATATGGCTGGGCATTGTGTACGCATATAAGGGATT
GCTGATGTTGTTTGGGTGCTTCCTGGCGTGGGAGACAAGGAATGTGGCCGACCTTGAGGATAATTG
GGAGACATTGAACGATAATCTGAAAGTCATCGAGAAGGCCGATAATGCAGCTCAGGTCAAAGATGC
CCTGACTAAGATGAGGGCGGCAGCCCTCGACGCCCAGAAGGCTACCCCTCCTAAGCTGGAGGATAA
ATCCCCAGATAGTCCCGAAATGAAAGACTTCCGACACGGCTTTGACATACTCGTTGGGCAGATCGAC
GACGCACTGAAACTTGCTAACGAGGGCAAAGTCAAGGAAGCGCAAGCTGCAGCTGAGCAGCTCAA
GACCACCCGAAACGCCTACATTCAGAAGTATCTGTCTATTCCCGCCCTCAATGATAGCAAGTATATCG
GCATGTCCGTGTATAATGTTGGTATCATGTGCATTATAGGGGCTGCAGTGAGCTTTCTCACACGCGA
CCAGCCTAATGTACAATTCTGTATCGTCGCCCTGGTGATAATATTTTGTAGCACGATCACGCTGTGCT
TGGTATTTGTACCCAAGCTCATCACTCTGAGAACTAATCCAGACGCTGCAACCCAAAACCGGCGCTTT
CAGTTCACCCAAAACCAGAAGAAGGAGGACTCCAAAACCTCAACCTCTGTGACAAGCGTCAACCAG
GCAAGTACATCCAGACTGGAAGGCCTCCAGTCTGAAAACCATCGCCTGCGGATGAAAATCACTGAAC
TTGATAAGGACCTGGAAGAGGTGACCATGCAACTGCAGGATACCCCTGAAAAGACCACCTACATCA
AAGGTACC 3’

Figure B.6 Synthetic $GABA_{B2}$ subunit construct (B2_BRIL_EK). Codon optimised for expression in mammalian cells.

Table B.1 PCR primers used in this thesis

Primer name	sequence
mGABB_B1x001f	5’-cgcaccggtagcgagcggagggcc-3’
mGABB_B1x001r	5’-cgcggtaccctgacgactctgcag-3’
mGABB_B2x001f	5’-cgcgaattcgccaccatggcttct-3’
mGABB_B2x001r	5’-cgcggtacctttgatgtaggtggt-3’
mGABA_B1x002r	5’-GCGGGTACCTTATCACTGACGACTCTGCAG-3’
mGABab_b1.X01f	5’-GAAAAGAGCAGAGCCGCGGAGAAGGAGAAC3’
mGABab_b1.X01r	5’-GTTCTCCTTCTCCGCGGCTCTGCTCTTTTC3’
mGABab_b2.X01f	5’-ACCCAAAACCAGAGGAAGGAGGACTCCAGAACCTCAACCTCT3’
mGABab_b2.X01r	5’-AGAGGTTGAGGTTCTGGAGTCCTCCTTCCTCTGGTTTTGGGT3’
mGABA_B1x003r	5’-cgccgcGGTACCACTGCCCGTTTTCATGGTAT3’
mGABA_B2x003r	5’-cgccgcGGTACCTGTCACAGAGGTTGAGGTTTTG3’
mGABA_B2x004f	5’-Gcgtgggagacaaggaatgtgtctattcccgccctcaatgat3’
mGABA_B2x004r	5’-atcattgagggcgggaatagacacattccttgtctcccacgC3’
mGABA_B1x004f	5’-gcctacgaaacgaaatctgtaagcacagagaagatcaatgat3’
mGABA_B1x004r	5’-atcattgatcttctctgtgcttacagatttcgtttcgtaggc3’
mGABA_B2x005r	5’- GCGGGTACCTTATCATTTGATGTAGGTGGT -3’
mGABA_B1x006r	5’- gcggcgGGTACCTTATCAACTGCCCGTTTTCATGG3’
mGABA_B2x006r	5’- gcggcgGGTACCTTATCATGTCACAGAGGTTGAG -3’
mGABA_B1x007f	5'- cgcGAATTCGCCACCATGCTGCTCTTG-3'
mGABA_B2x007r	5’gcgCTCGAGTTATCATGCAGGTGCTACCTGAGACGTCTCAGTTGAGCCTCCGGTACC TGTCACAGAGGTTGAGGTTTTGGA
mGABA_B2x008r	5’gcgCTCGAGTTATCATGCAGGTGCTACCTGAGACGTCTCAGTGGAACCTCCACCTCC GGTACCTGTCACAGAGGTTGAGGTTTTGGA
mGABA_B2x009r	5’gcgCTCGAGTTATCATGCAGGTGCTACCTGAGACGTCTCAGTGGACCCACCACCACT TCCTCCTCCGGTACCTGTCACAGAGGTTGAGGTTTTGGA
mGABA_B2x010f	5’-cgccgcGAATTCACGCGTGCCACCATGGCTTCTCCAAGAAGC-3’
mGABA_B2x010r	5’-gcggcgCTCGAGGCGGCCGCTTATCATGCAGGTGCTACCTGAGA-3’
mGABA_B1x008f	5’-cgccgcGAATTCACGCGTGCCACCatGGGCATCCTGCCCAGC-3’
mGABA_B1x008r	5’- gcggcgGGTACCGCGGCCGCTTATCACTGACGACTCTGCAGTTG-3’
mGABA_B1x009r	5’-gcggcgGGTACCGCGGCCGCTTATCAACTGCCCGTTTTCATGG-3’
mGABA_B1bT198Cf	5’actatccaacagtgcactgaagtgttc-3’
mGABA_B1bT198Cr	5’gaacacttcagtgcactgttggatagt-3’
mGABA_B2Q206Cf	5’ Acccaagatgtgtgtcggttcagcgag 3’
mGABA_B2Q206Cr	5’ ctcgctgaaccgacacacatcttgggt 3’
mGABA_B2x011f	5’GATACCCCTGAAAAGACCACCtACATCAAAGGTACCGGCGGCAGCGGAGGATCTC TGGAAGTTCTGTTTCAGGGCCCCGACTACAAGGACGACGACGACAAGGTGGACAAC AAGTTCAACAAAGAGCAGCAGAACGCCTTCTACGAGATCCTGCATCTGCCCAACCTG AACGAGGAACAGAGAAACGCCTTCATCCAGAGCCTGAAGGACGACCCTAGCCAGAG CGCTAATCTGCTGGCCGAGGCCAAGAAACTGAACGGCGCTCAGGCCCCTAAAGTCG ATAACAAGTTTAACAAAGAACAACAAAATGCCTTTTATGAGATTCTGCACCTTCCGAA TCTCAATGAGGAACAGCGGAATGCCTTTATTCAGTCCCTGAAGGATGATCCCTCTCAG TCCGCCAACCTGCTGGCTGAAGCTAAAAAGCTGAATGGTGCCCAGGCTCCTAAGTGA TGAGCGGCCGCCTCGAGcgc -3’
mGABA_B2x012f	5’GATACCCCTGAAAAGACCACCt 3’
mGABA_B2x012r	5’ gcgCTCGAGGCGGCCGCTCAT3’

Table B.1 (continued) PCR primers used in this thesis

Primer name	sequence
mGABA_B2x013f	5'cgcGGTACCGGAGGATCTCTTGTGCCAAGGGGTAGTGACTACAAGGACGACGACG ACAAG3'
mGABA_B2x014f	5' CATCAAACAGAACCACTACCAGGAGCTGA3'
mGABA_B2x015r	5'cgccgcCTCGAGGCGGCCGCTTATCACAGTCCTGACACCAT3'
mGABA_B2x016f	5'cgccgcACGCGTGAATTC GCCACCATGGGAATT3'
mGABA_B2x016r	5'cgccgcACCGGTCCCGGAACCGCCTTTATAC 3'
mGABA_B2x017f	5'cgccgcACCGGTCCACCTAGTAGTCCTCCC3'
mGABA_B2x018f	5'- cgccgcACCGGTAGTCCTCCCCTCAGTATCATGGG 3'
mGABA_B2x018r	5'- cgccgcCTCGAGTTATCcGCCAATACTGGGAAGATAG 3'

Table.B.2 Sequencing primers used in this thesis

Primer name	Sequence	aim
mGABB_B1_SeqF00	5' GTACCTTTACGAGCTTCTCTACA 3'	B1_noS_Bril_AK sushi domain
mGABB_B1_SeqF01	5'-ggagcggttgtttggtaaga-3'	B1_noS_Bril_AK 654 forward
mGABB_B1_SeqF02	5'-caactcacacgtgcgataca-3'	B1_noS_Bril_AK 1359 forward
mGABB_B1_SeqF03	5'-gactcacccgaaatgaagga-3'	B1_noS_Bril_AK 2056 forward
mGABB_B1_SeqR01	5'-ggtgcagttgatagacgggt-3'	B1_noS_Bril_AK 747 reverse
mGABB_B1_SeqR02	5'-ctgccagtgcgagactacag-3'	B1_noS_Bril_AK 1441 reverse
mGABB_B1_SeqR03	5'-agcttccttcaccttcccct-3'	B1_noS_Bril_AK 2157 reverse
mGABB_B2 _SeqF01	5'-accggagtcctgtatggtga-3'	B2_Bril_EK 658 forward
mGABB_B2 _SeqF02	5'-cggtagctgacacactggaa-3'	B2_Bril_EK 1343 forward
mGABB_B2 _SeqF03	5'-gccgaccttgaggataattg-3'	B2_Bril_EK 2056 forward
mGABB_B2 _SeqR01	5'-tcattgcccttcagcttctt-3'	B2_Bril_EK 752 reverse
mGABB_B2 _SeqR02	5'-gaggcaagcttattttcctcaa-3'	B2_Bril_EK 1453 reverse
mGABB_B2 _SeqR03	5'-cagggcatctttgacctgag-3'	B2_Bril_EK 2145 reverse
PHLSEC_REV	5'- CACCAGCCACCACCTTCTGAT -3'	phLSEC vector end reverse
PHLSEC_Fwd	5'-GCTCTAGAGCCTCTGCTAACC-3'	phLSEC vector forward
pHR_Fwd	5'- tgcttctcgcttctgttcg – 3'	phR vector forward
pHR_Rev	5'- ccacatagcgtaaaaggagc – 3'	phR vector reverse
SBP_rev	5'-GCCGGTGGTTTTCTCGTCCAT-3'	Reverse in SBP tag
SBP_fwd	5' GCA TCC TGC CCA GCC CTG 3'	SBP tag forward
mGABB_B2 _SeqF04	GTACCCAAGCTCATCAC	B2 C-terminus forward
Seq_mEGFP_fwd	CAAGATTCGCCATAACA	mEFGP tag forward
Seq_SNAP_fwd	CCGCCGTGAAAACAGCCC	SNAP tag forward
Seq_pet45b_rev	CCAGACTCGAGTGCGGCCG	pet45b+ vector forward
Seq_pet45b_fwd	CCTCTAGAAATAATTTTG	pet45b+ vector reverse

Table B.3 Summary of constructs generated in this thesis

Template	Forward primer	Reverse primer	Insert	Restriction sites	vector
B1_noS_BRIL_AK	mGABB_B1x001f	mGABB_B1x001r	B1b_BRIL	Age-Kpn	pHLsec
B1_noS_BRIL_AK	mGABB_B1x001f	mGABB_B1x001r	B1b_BRIL	Age-Kpn	pHLsec-1D4
B1_noS_BRIL_AK	mGABB_B1x001f	mGABB_B1x001r	B1b_BRIL	Age-Kpn	pHLsec-Ctmv
B1_noS_BRIL_AK	mGABB_B1x001f	mGABB_B1x001r	B1b_BRIL	Age-Kpn	pHLsec-CtmRuby
B2_BRIL_EK	mGABB_B2x001f	mGABB_B2x001r	B2_BRIL	EcoRI-Kpn	pHLsec
B2_BRIL_EK	mGABB_B2x001f	mGABB_B2x001r	B2_BRIL	EcoRI-Kpn	pHLsec-1D4
B2_BRIL_EK	mGABB_B2x001f	mGABB_B2x001r	B2_BRIL	EcoRI-Kpn	pHLsec-Ctmv
B2_BRIL_EK	mGABB_B2x001f	mGABB_B2x001r	B2_BRIL	EcoRI-Kpn	pHLsec-CtmRuby
B1_noS_BRIL_AK	mGABB_B1x001f	mGABB_B1x001r	B1b_BRIL	Age-Kpn	pHLsec-Hisx6
B2_BRIL_EK	mGABB_B2x001f	mGABB_B2x001r	B2_BRIL	EcoRI-Kpn	pHLsec-Hisx6
B1_noS_BRIL_AK	mGABB_B1x001f	mGABA_B1x002r	B1b_BRIL**	Age-Kpn	SBP-pHLsec
Sushi domain			B1a_BRIL	EcoRI-BsrGI	pHLsec-1D4
B1_noS_BRIL_AK	mGABB_B1x001f mGABA_B1x004f	mGABA_B1x004r mGABB_B1x002r	B1b**	Age-Kpn	SBP-pHLsec
B2_BRIL_EK	mGABB_B2x001f mGABA_B2x004f	mGABA_B2x004r mGABB_B2x001r	B2	EcoRI-Kpn	pHLsec-1D4
B2_BRIL_EK	mGABB_B2x001f	mGABB_B2x003r	B2_bRIL_noCC	EcoRI-Kpn	pHLsec-1D4
B2_BRIL_EK	mGABB_B2x001f	mGABA_B2x007r	B2_BRIL_GGS	EcoRI- XhoI	pHLsec-1D4
B2_BRIL_EK	mGABB_B2x001f	mGABA_B2x008r	B2_BRIL_GGGGS	EcoRI- XhoI	pHLsec-1D4
B2_BRIL_EK	mGABB_B2x001f	mGABA_B2x009r	B2_BRIL_GGGSGGGS	EcoRI- XhoI	pHLsec-1D4
B1b_BRIL**	mGABB_B1x001f mGABA_B1bT198Cf	mGABA_B1bT198Cr mGABB_B1x002r	B1b_T198C_BRIL**	Age-Kpn	SBP-pHLsec
B2	mGABB_B2x001f mGABA_B2Q206Cf	mGABA_B2Q206Cr mGABB_B2x001r	B2_Q206C	EcoRI-Kpn	pHLsec-1D4
B1b**	mGABB_B1x001f mGABA_B1bT198Cf	mGABA_B1bT198Cr mGABB_B1x002r	B1b_T198C**	Age-Kpn	SBP-pHLsec
B1_noS_BRIL_AK	mGABB_B1x001f mGABab_b1.X01f	mGABB_B1x001r mGABab_b1.X01r	B1b_BRIL_AA	Age-Kpn	SBP-pHLsec
B2_BRIL_EK	mGABB_B2x001f	mGABB_B2x001r	B2_BRIL_RR	EcoRI-Kpn	pHLsec-1D4
B1_noS_BRIL_AK	mGABB_B1x001f	mGABB_B2x003r	B1b_bRIL_noCC	Age-Kpn	SBP-pHLsec
B1b_BRIL**	mGABB_B1x008f	mGABB_B1x008r	B1b_BRIL**	MluI-NotI	pHR
B1b**	mGABB_B1x008f	mGABB_B1x008r	B1b**	MluI-NotI	pHR
B2	mGABB_B2x010f	mGABB_B2x010r	B2	MluI-NotI	pHR

** a double stop codon

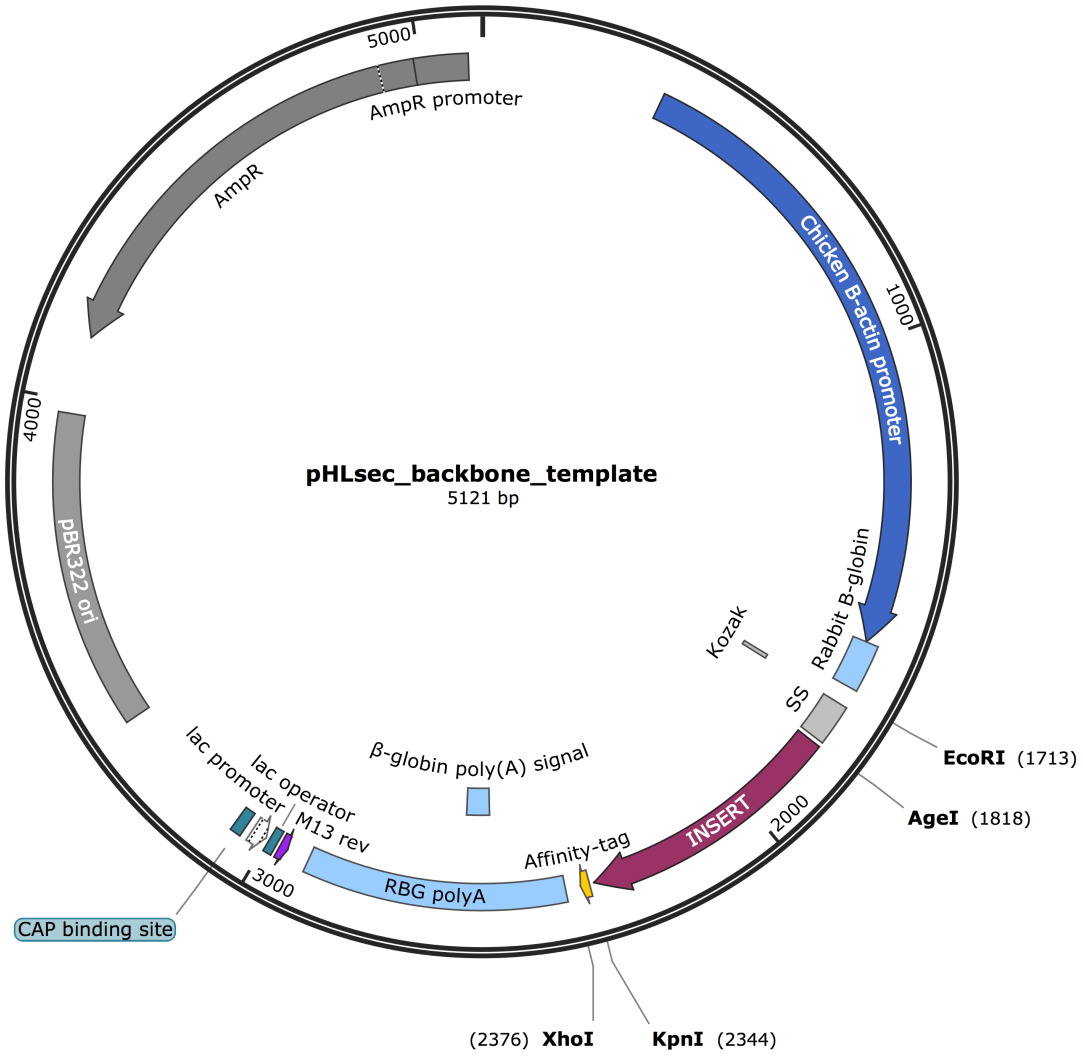


Figure B.7 pHLsec plasmic vector map.

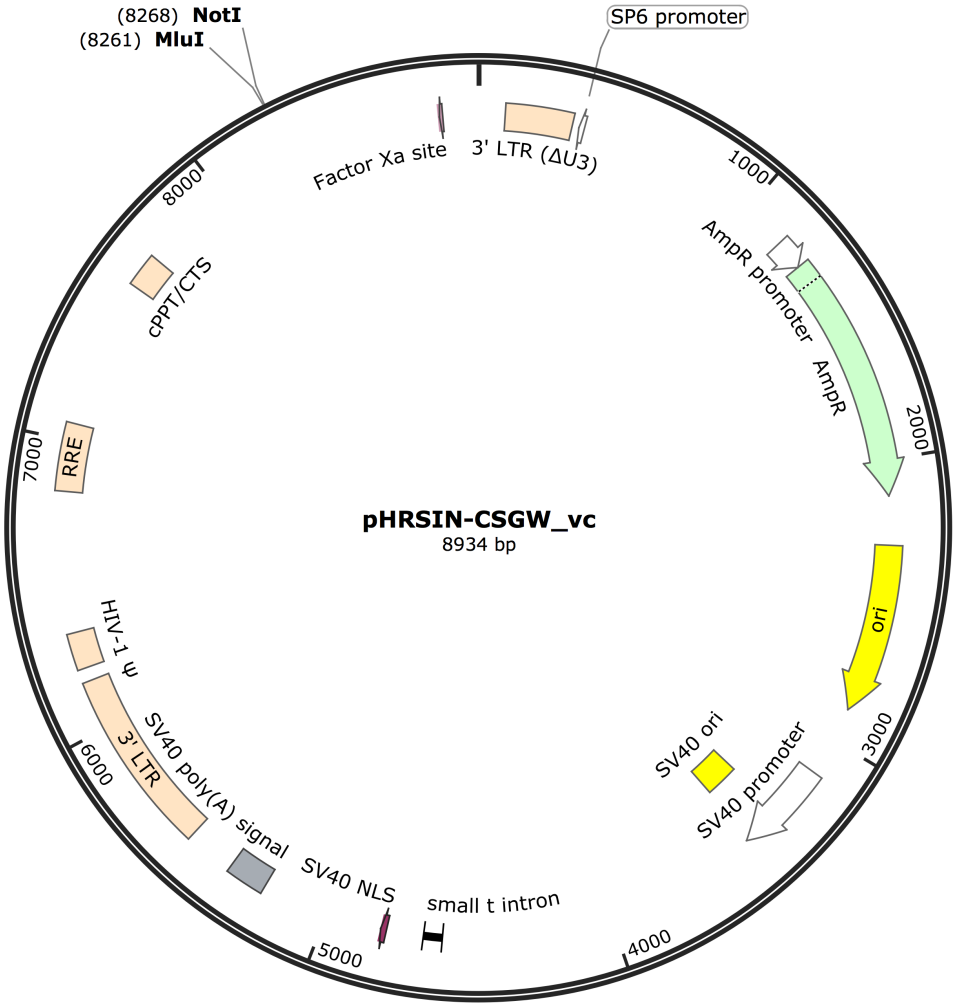


Figure B.8 pHRSIN-CSGW_vc vector map.

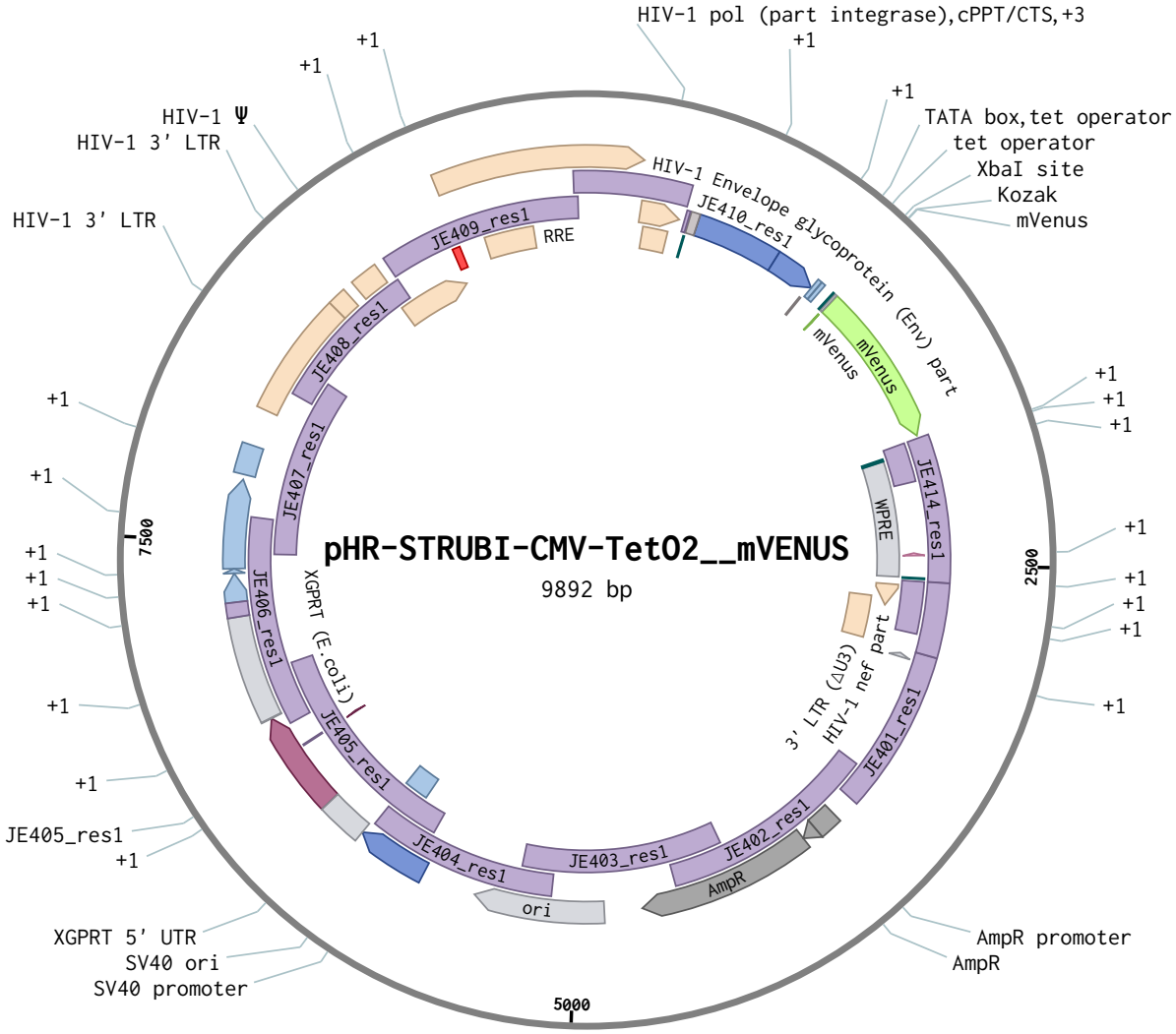


Figure B.9 pHR-CMV-TetO$_2$ vector map.

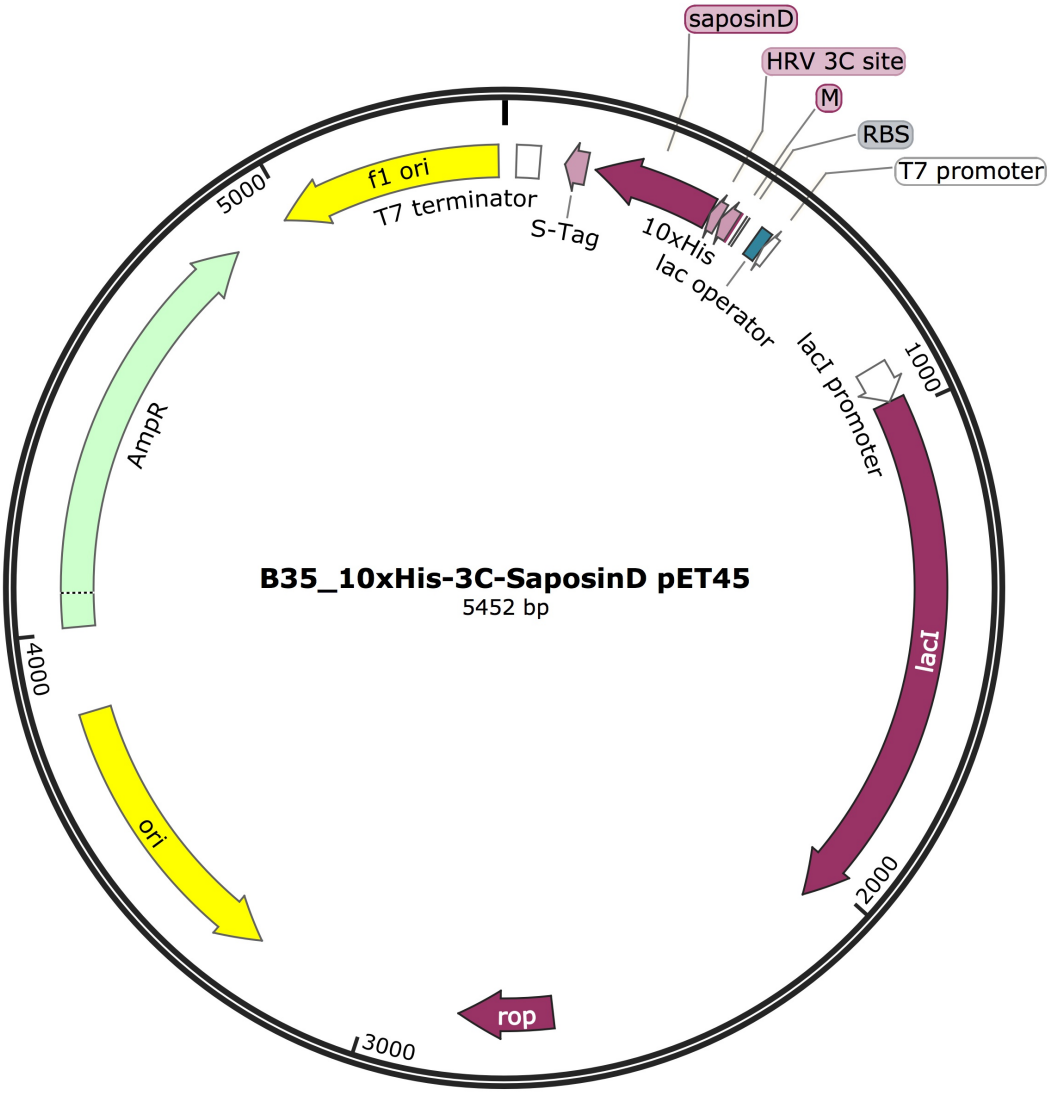


Figure B.10 pET45+ vector map.

MDEKTTGWRGGHVVEGLAGELEQLRARLEHHPQGQREPDYDIPTTENLYFQGTGSERRAVYIGALFPM
SGGWPGGQACQPAVEMALEDVNSRRDILPDYELKLIHHDSKCDPGQATKYLYELLYNDPIKIILMPGCSS
VSTLVAEAARMWNLIVLSYGSSSPALSNRQRFPTFFRTHPSATLHNPTRVKLFEKWGWKKIATIQQTTEV
FTSTLDDLEERVKEAGIEITFRQSFFSDPAVPVKNLKRQDARIIVGLFYETEARKVFCEVYKERLFGKKYVWF
LIGWYADNWFKIYDPSINCTVDEMTEAVEGHITTEIVMLNPANTRSISNMTSQEFVEKLTKRLKRHPEET
GGFQEAPLAYDAIWALALALNKTSGGGGRSGVRLEDFNYNNQTITDQIYRAMNSSSFEGVSGHVVFDAS
GSRMAWTLIEQLQGGSYKKIGYYDSTKDDLSWSKTDKWIGGSPPADQTLVIKTFRFLSQKLFISVSVLSSL
GIVLAVVCLSFNIYNSHVRYIQNSQPNLNNLTAVGCSLALAAVFPLGLDGYHIGRNQFPFVCQARLWLLGL
GFSLGYGSMFTKIWWVHTVFTKKEEKKEWRKTLEPWKLYATVGLLVGMDVLTLAIWQIVDPLHRTIETF
AKEEPKEDIDVSILPQLEHCSSRKMNTWLGIFYGYKGLLLLLGIFLAYETKSVSTEKINDHRAVGMAIYNVA
VLCLITAPVTMILSSQQDAAFAFASLAIVFSSYITLVVLFVPKMRRLITRGEWQSEAQDTMKTGSSTNNNE
EEKSRLLEKENRELEKIIAEKEERVSELRHQLQSRQ **

Fig B.11 Amino acid sequence of SBP_GABA$_{B1b}$_WT construct. N-terminal SBP-tag is highlighted in yellow.

MASPRSSGQPGPPPPPPPPPARLLLLLLLPLLLPLAPGAWGWARGAPRPPPSSPPLSIMGLMPLTKEVAK
GSIGRGVLPAVELAIEQIRNESLLRPYFLDLRLYDTECDNAKGLKAFYDAIKYGPNHLMVFGGVCPSVTSIIA
ESLQGWNLVQLSFAATTPVLADKKKYPYFFRTVPSDNAVNPAILKLLKHYQWKRVGTLTQDVQRFSEVR
NDLTGVLYGEDIEISDTESFSNDPCTSVKKLKGNDVRIILGQFDQNMAAKVFCCAYEENMYGSKYQWIIP
GWYEPSWWEQVHTEANSSRCLRKNLLAAMEGYIGVDFEPLSSKQIKTISGKTPQQYEREYNNKRSGVGP
SKFHGYAYDGIWVIAKTLQRAMETLHASSRHQRIQDFNYTDHTLGRIILNAMNETNFFGVTGQVVFRNG
ERMGTIKFTQFQDSREVKVGEYNAVADTLEIINDTIRFQGSEPPKDKTIILEQLRKISLPLYSILSALTILGMIM
ASAFLFFNIKNRNQKLIKMSSPYMNNLIILGGMLSYASIFLFGLDGSFVSEKTFETLCTVRTWILTVGYTTAF
GAMFAKTWRVHAIFKNVKMKKKIIKDQKLLVIVGGMLLIDLCILICWQAVDPLRRTVEKYSMEPDPAGRD
ISIRPLLEHCENTHMTIWLGIVYAYKGLLMLFGCFLAWETRNVSIPALNDSKYIGMSVYNVGIMCIIGAAVS
FLTRDQPNVQFCIVALVIIFCSTITLCLVFVPKLITLRTNPDAATQNRRFQFTQNQKKEDSKTSTSVTSVNQA
STSRLEGLQSENHRLRMKITELDKDLEEVTMQLQDTPEKTTYIKGTTETSQVAPA

Fig B.12 Amino acid sequence of $GABA_{B2}$_WT_1D4 construct. C-terminal 1D4-tag is highlighted in yellow.

Saposin A Sequence:
SLPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDIIKGEMSRPGEV
CSALNLCES

Saposin D Sequence:
GDGGFCEVCKKLVGYLDRNLEKNSTKQEILAALEKGCSFLPDPYQKQCDQFVAEYEPVLIEILVEVMDPSF
VCLKIGACPSAH

MSP1D1 Sequence:
GHHHHHHHDYDIPTTENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYL
DDFQKKWQEEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDE
LRQRLAARLEALKENGGARLAEYHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLNT
Q

MSP2N2 Sequence:
GHHHHHHHDYDIPTTENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYL
DDFQKKWQEEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDE
LRQRLAARLEALKENGGARLAEYHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLNT
QGTPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELYRQKVEPLRAELQEG
ARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELRQRLAARLEALKENGGARLAEYHAKATE
HLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLNTQ

GFP-mini-G_{o1R12} Sequence
MGHHHHHHHENLYFQGVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPV
PWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIE
LKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVL
LPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYKGGGGSIEKNLKEDGISAAKDVKLLLLG
ADNSGKSTIVKQMKIIHGGSGGSGGTTGIVETHFTFKNLHFRLFDVGGQRSERKKWIHCFEDVTAIIFCVD
LSDYNRMHESLMDFDSICNNKFFIDTSIILFLNKKDLFGEKIKKSPLTICFPEYTGPNTYEDAAAYIQAQFESK
NRSPNKEIYCHMTCATDTNNAQVIFDAVTDIIIANNLRGCGLY

Fig B.13 Amino acid sequence of saposin A, saposin D, MSP1D1, MSP2N2 and GFP-mini-G_{o1R12} sequence.

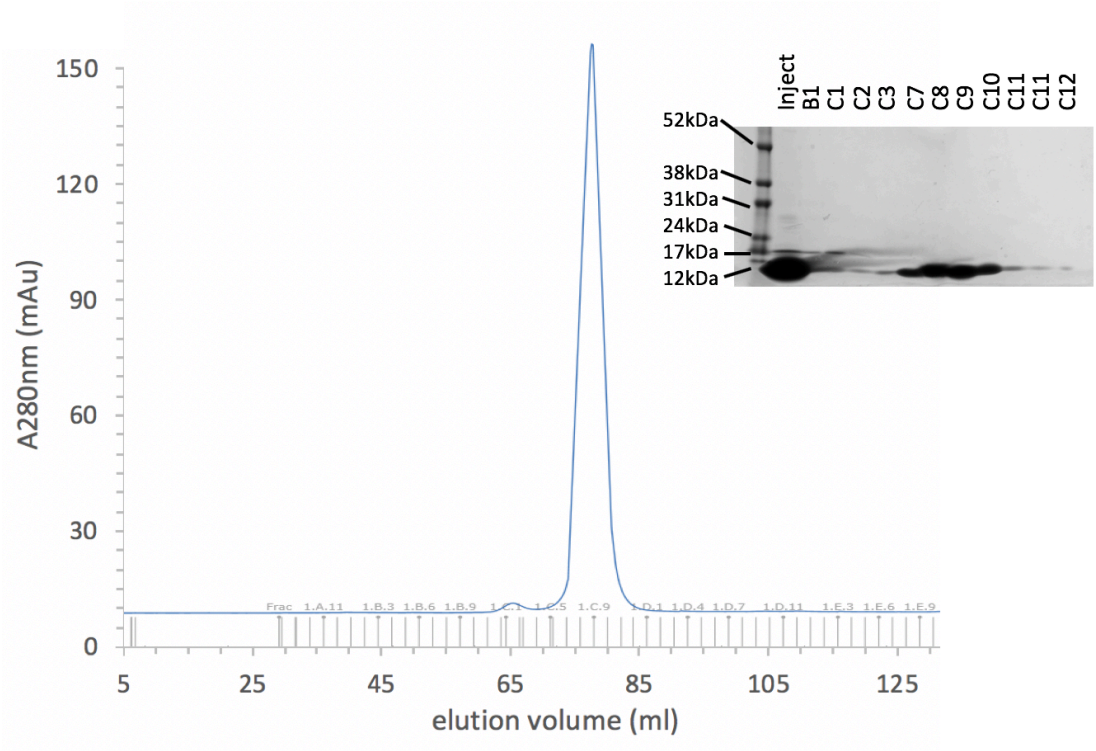


Fig B.14 Representative purification profile of saposin protein. (a) Gel filtration profile of saposin D from 1 L culture. (b) SDS-PAGE analysis of peak fractions.Typical yield from 1 L culture was 5-7 mg.

Table B.4 Ilama nanobodies generated against the wild-type $GABA_{B1b/B2}$ receptor.

Immunization with Human Gamma-aminobutyric acid Type B receptor									
Reference No. (1)	Family (2)	Occur-rence (3)	Selection conditions in which this family of Nanobodies is found (4)					Vector (5)	Clone ID (6)
			solid phase coated GABAb_WT in absence of 3APPA or CGP54626	GABAb_WT in amphipols trapped via neutravidin in absence of 3APPA or CGP54626	GABAb_WT in amphipols trapped via 1D4:				
					in absence of 3APPA or CGP54626	in presence of 100nM 3APPA	in presence of 100nM CGP54626		
CA13890	1	7			x	x	x	pMESy4	MP1350_H1
CA13891									MP1352_C3
CA13938									MP1349_G7
CA13886	2	8			x	x			MP1350_B5
CA13887									MP1350_B1
CA13888									MP1352_E4
CA13889									MP1352_G1
CA13892	3	4			x	x			MP1349_E3
CA13893									MP1350_E1
CA13894									MP1350_F5
CA13895									MP1352_G3
CA13896	4	3			x	x			MP1350_B3
CA13897									MP1352_D3
CA13898									MP1350_C6
CA13899	5	3			x	x			MP1350_C3
CA13900	6	2				x	x		MP1349_A3
CA13901									MP1349_B8
CA13902	7	2				x	x		MP1352_E1
CA13903	8	1	x						MP1348_F4
CA13904	9	1		x					MP1348_F11
CA13905	10	1		x					MP1348_A10
CA13906	11	1		x					MP1348_C12
CA13907	12	1		x					MP1348_G10
CA13908	13	1		x					MP1348_F8
CA13909	14	1		x					MP1348_D8
CA13910	15	1		x					MP1348_F12
CA13911	16	1		x					MP1348_H7
CA13912	17	1		x					MP1348_F10
CA13913	18	1			x				MP1350_C1
CA13914	19	1			x				MP1350_A2
CA13915	20	1			x				MP1350_H5
CA13916	21	1			x				MP1350_G3
CA13917	22	1			x				MP1350_A6
CA13918	23	1			x				MP1350_C2
CA13919	24	3				x			MP1349_D1
CA13920									MP1349_B5
CA13921	25	1				x			MP1352_A6
CA13922	26	1				x			MP1352_B2
CA13923	27	1				x			MP1349_H1
CA13924	28	1				x			MP1352_F2
CA13925	29	1				x			MP1349_D2
CA13926	30	1				x			MP1352_E3
CA13927	31	1				x			MP1349_E6
CA13928	32	1				x			MP1349_G3
CA13929	33	1				x			MP1349_D5
CA13930	34	1				x			MP1352_E6
CA13931	35	1				x			MP1349_E1
CA13932	36	1				x			MP1352_E5
CA13933	37	1				x			MP1349_B2
CA13934	38	1				x			MP1352_G5
CA13935	39	1				x			MP1349_C1
CA13936	40	1				x			MP1349_A2
CA13937	41	1				x			MP1352_B4
CA13939	42	2					x		MP1349_G10
CA13940									MP1352_C11
CA13941	43	1					x		MP1349_E10
CA13942	44	1					x		MP1349_F8
CA13943	45	1					x		MP1349_B10
CA13944	46	1					x		MP1349_F11
CA13945	47	1					x		MP1349_G9
CA13946	48	1					x		MP1349_H7
CA13947	49	1					x		MP1349_H9