Construction of polymerome-based artificial cell membrane

Hu, Zhaolong

2017

Hu, Z. (2017). Construction of polymerome-based artificial cell membrane. Doctoral thesis, Nanyang Technological University, Singapore.

<http://hdl.handle.net/10356/69959>

<https://doi.org/10.32657/10356/69959>

NANYANG
TECHNOLOGICAL
UNIVERSITY

**CONSTRUCTION OF POLYMERSOME-BASED
ARTIFICIAL CELL MEMBRANE**

HU ZHAOLONG

SCHOOL OF MATERIALS SCIENCE AND ENGINEERING

2016

CONSTRUCTION OF POLYMEROME-BASED ARTIFICIAL CELL MEMBRANE

HU ZHAOLONG

SCHOOL OF MATERIALS SCIENCE AND ENGINEERING

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2016

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research and has not been submitted for a higher degree to any other University or Institution.

19/08/2016

.....

Date

A handwritten signature in black ink, appearing to read 'Hu Zhaolong', is shown within a rectangular box. The signature is fluid and cursive.

.....

Hu Zhaolong

Abstract

Membrane Proteins (MPs) are key players in many different biological networks, such as signal transduction cascades, cell-cell communication and environment sensing. The structural complexity and range of vital biological functions make them not only of scientific interest but also important targets for therapeutic intervention. However, they are more difficult to work with as compared to soluble proteins, due to their rapid disintegration once they are removed from the embedding plasma membrane. Traditional methods of producing MPs, which include heterologous overexpression cells, extraction from cell membranes, purification, and refolding efforts to preserve native structures are labour intensive and the use of detergents during the extraction and purification may cause protein denaturation.

Since 1988, when Zozulia SA et al. first demonstrated the *in vitro* translation of visual rhodopsin (an important class of MP), cell free protein synthesis (CFPS) has been an evolving tool for MP research. Different from *in vivo* protein production in host cells, CFPS uses cell extract containing translational machinery (e.g. ribosomes, translation factors and metabolic enzymes) for protein expression. The absence of cell membrane creates an open and versatile environment that can be optimized exclusively towards quality production of a single protein product without supporting cell viability and growth. Such characteristics allow maximum modification and control of expression/isolation conditions to overcome issues in conventional *in vivo* expression systems (e.g. no expression or cell death due to target toxicity, protein denaturation in extraction), making CFPS a valuable alternative tool for challenging protein target production, e.g. MPs. Furthermore, it allows modifying the components that are amenable for expressing membrane proteins (e.g., additional chaperons to assist membrane proteins folding) with flexible strategies as well as high throughput screening when necessary.

In the research involving MPs, however, CFPS are facing similar challenges as in traditional methods such as the expressed MPs have to be stabilized and maintain structural

integrity during expression and processing. MPs are intrinsically unstable and in *in vivo* synthesis, the current approaches depend largely on trial-and-errors (e.g. expression system optimization, directed evolution, condition screening). In CFPS, such challenges are approached by supplying amphiphilic components, e.g. membrane support like liposomes (lipid-based vesicles), to mimic cellular environment and to assist MP folding and stabilization. However, a key disadvantage of conventional membrane supports is their limited repertoire of physicochemical properties. Consequently, their applications in CFPS are restricted to genetically stabilized, well-studied targets (e.g. β 2AR, AqpZ, bacteria rhodopsin) whereas the majority of challenging MPs (e.g. class B GPCRs) with wild-type sequences have not been attempted. In order to overcome these challenges, more flexible and robust biomimetic membrane systems are highly desired. In an entirely different field, self-assembled block copolymer vesicles (polymersomes) have emerged as suitable candidates in addition to lipid membranes. The chemical diversity allows tuning of membrane properties (e.g. chain length, hydrophobicity, rigidity) over a wider range than traditional membrane supports. They have been explored and show great potential as artificial cell membranes.

In this work, we systematically investigated polymersomes as native-like membrane support for MP production in CFPS. Following previous proof of concept studies, we have selected Class B GPCRs, which are important drug targets in many human diseases, including diabetes, osteoporosis and obesity. In comparison to Class A GPCRs, they are more inherently unstable and lack high-affinity ligands for stabilization thus the study of this class of GPCRs has been considerably delayed. Specifically, a type II diabetes target, the Glucagon-like peptide-1 receptor (GLP-1R) was selected because of its role in chronic metabolic disorder and the very limited information available on the characterization of this protein target. Prevalence of type II diabetes has been increasing steadily globally yet no cure is available. As a diabetes target, GLP-1R related treatments offer unique benefits that include not only control of blood glucose level but improvement of co-morbid conditions.

In this thesis, the polymersomes were used as a platform for development of artificial cell membranes with GLP-1R as a model protein. We found that the novel purification approach using tagged membranes produced correctly orientated, functional GLP-1R incorporated in vesicular polymersomes of defined size whereas traditional histidine purification produced only protein aggregates. The characterization of polymer-based membranes also suggests an association between membrane rigidity and protein functionality, thus providing guidelines for different production strategies. In addition to GLP-1R, the principle and methodology were applied to a separate class of MP, a potassium (K2P) channel TREK-2.

MPs constitute around 60% of approved drug targets and the development of therapeutic antibodies has long been hindered by the limited availability of natively folded MP antigen like GPCRs. In our work, the novel polymersome-based artificial membranes with stable and functional GLP-1R showed great potential as a solution to GPCR antibody generation. The methods presented herein are broadly applicable to other complex MPs and may serve to reduce bottlenecks such as production and presentation of drug targets for the discovery of novel ligands and antibodies to such targets.

Acknowledgements

I would like to thank my supervisor, Assistant Professor Madhavan Nallani for his guidance over my research. I am touched by his positive attitude towards life and work and inspired by his insights into material science and technology.

I would like to extend my appreciation to all my group members, Dr. Hans-Peter, Dr. Sourabh, Dr. Tom, Dr. James, Esther, Winna, and Ranjani, Hossam, and Abbas for their support and encouragement. I thank the friends and colleges at CBSS.

I would also like to thank School of Materials Science and Engineering (MSE) and Center for Biomimetic Sensor Science (CBSS).

Lastly but not least, I would like to thank my beloved family and friends.

Table of Contents

Abstract	i
Acknowledgements	v
Table of Contents	vii
Table Captions	xi
Figure Captions	xiii
Abbreviations	xxi
 Chapter 1 Introduction	1
1.1 Background and Hypothesis	2
1.2 Objectives and Scope	3
1.3 Dissertation Overview	5
1.4 Findings and Outcomes/Originality	6
References	7
 Chapter 2 Literature Review	9
2.1 Membrane Protein Structure and Function	10
2.1.1 G-protein Coupled Receptors	10
2.1.2 Ion Channels	13

2.2	Membrane Protein Production and Characterization	14
2.2.1	In Vivo Membrane Protein Expression Systems	15
2.2.2	Membrane Protein Engineering	16
2.2.3	Cell Free Protein Synthesis	18
2.2.4	Biomimetic Membrane Support.....	21
2.2.5	Membrane Supports in Cell Free Protein Synthesis	26
	References	30

Chapter 3 Experimental Methodology..... 39

3.1	Rationale for selection	40
3.2	Synthesis	41
3.2.1	Materials	41
3.2.2	Molecular Cloning - Principles and Analysis	42
3.2.3	Protein Expression and Purification - Principles and Analysis	43
3.2.4	Polymersomes Preparation - Principles and Analysis	46
3.3	Characterization	47
3.3.1	Protein Gel Electrophoresis and Western Blot - Principles and Analysis ...	47
3.3.2	Bicinchoninic Acid Assay - Principles and Analysis	48
3.3.3	Fluorescence Detection Size-exclusion Chromatography - Principles and Analysis	48
3.3.4	Biosensor analysis - Principles and Analysis.....	49
3.3.5	Transmission Electron Microscopy - Principles and Analysis	50
3.3.6	Radioligand Binding Assay - Principles and Analysis	50
3.4	Overview of Methodologies	51
	References	51

Chapter 4 Protein Synthesis and Primary Characterization 53

4.1	Introduction.....	54
4.2	Experimental Methodology	54
4.3	Principle Outcomes	56

4.3.1	GLP-1R Is Expressed and Reconstituted in Polymersomes Using High Yield Wheat Germ System.....	56
4.3.2	Protein Yield Improvement Using Bilayer Reaction	58
4.3.3	GLP-1R Purification and Primary Characterization	61
4.4	Conclusions.....	65
	References.....	66
Chapter 5 Biophysical Characterization of Proteopolymersomes.....		67
5.1	Introduction.....	68
5.2	Experimental Methodology	68
5.3	Principle Outcomes.....	70
5.3.1	Morphological Characterization of Purified Product.....	70
5.3.2	Molecular cloning and expression of GFP-GLP-1R Proteopolymersomes for FSEC Characterization	72
5.3.3	Size-exclusion chromatography Analysis of GLP-1R incorporation into polymersomes.....	74
5.4	Conclusions.....	80
	References.....	81
Chapter 6 Biochemical Characterization of Proteopolymersomes		83
6.1	Introduction.....	84
6.2	Experimental Methodology	84
6.3	Principle Outcomes.....	86
6.3.1	Determination of GLP-1R Orientation in Polymersomes and Ligand Binding by Surface Plasmon Resonance (SPR).....	86
6.3.2	Lack of GLP-1R Functionality Determined Using Radioligand Saturation Binding Assay	91
6.3.3	Assisted Folding of GLP-1R Enhances Functionality	94
6.4	Conclusions.....	102
	References.....	102

Chapter 7 Functional Production of Proteopolymersomes.....	105
7.1 Introduction.....	106
7.2 Experimental Methodology	106
7.3 Principle Outcomes	107
7.3.1 Co-translational Incorporation of GLP-1R into Polymersomes	107
7.3.2 Reconstitution of Purified GLP-1R into Polymersomes.....	111
7.4 Conclusions.....	116
References	117
 Chapter 8 Discussion and Future Work	 119
8.1 Discussion	120
8.2 Outlook	121
References	124
 APPENDIX.....	 125
Appendix I: DNA sequence of pEU-MCS.....	125
Appendix II: DNA sequence of G1LP-1R-pEU	128
Appendix III: DNA sequence of GFP-GLP-1R.....	132
Appendix IV: DNA sequence of pCMVTNT-MCS	134
Appendix V: DNA sequence of TREK-2	137
Appendix VI: DNA sequence of truncated TREK-2	139
Appendix VII: Impact of Additional Amphiphiles on Cell-Free Protein Synthesis	140
Appendix VIII: Functionality Characterization with Alternative Methods	143
Appendix IX: SPR Sensorgrams of Exendin-4 Binding.....	153
Appendix X: List of Publications	153

Table Captions

Table 3.1 Recipe of protein expression screening in different cell free expression systems.

Figure Captions

Figure 2.1 Schematic representation of the predicted topology of (A) rhodopsin-like family GPCR, the human chemokine CXCR4 (CXCR4) and (B) secretin-like family GPCR, the rat Glucagon like peptide-1 receptor (GLP-1R). The GLP-1R has long N- (126 amino acids) and C- (55 amino acids) terminal domain compared to CXCR4 (39 and 44 amino acids, respectively).

Figure 2.2 Schematic representation of the predicted topology of (A) secretin-like family GPCR protein and (B) K⁺ ion channel Kv7.1. GPCR has structural feature of seven-transmembrane domains with extracellular N-terminal and intracellular C-terminal domains. In comparison, ion channels display six-transmembrane domains with both N- and C- terminal domains in the cytosol.

Figure 2.3 (A) A typical workflow for directed evolution of GPCR proteins. DNA library constructed from randomized GPCR genes were transformed to expression host (yeast), followed by protein display, functional selection by ligand binding assay and finally selected members were recovered to embark a second round of screening. (B) Snake plot of stabilized, functional human corticotropin-releasing factor receptor 1 (CRF1). Both N- (1-103) and C- (373-404) terminus were truncated with T4L inserted at second intracellular loop. Insertions and mutations were labelled red and green.

Figure 2.4 Schematic illustration of *in vivo* expression and *in vitro* CFPS. *In vivo* expression starts with transformation of cDNA coding protein of interest to host cells, followed by cell culture. Likewise, *in vitro* expression requires addition of cDNA to cell lysate, which contains all the essential elements for RNA transcription and protein translation. Both *in vivo* and *in vitro* expressed proteins require protein purification after synthesis.

Figure 2.5 Schematic scaling of lipid vesicles and polymersomes membrane thickness with increasing copolymer molecular weight. PBD-PEO: poly (butadiene)-poly (ethylene oxide). EO: ethylene oxide. BD: polybutadiene.

Figure 4.1 Western blot analysis of GLP-1R expression in regular wheat germ (WG, lane 1), high yield wheat germ (HYWG, lane 2), S30 *E. coli* lysate (S30, lane 3), and rabbit reticulocyte lysate (RRL, lane 4) cell free expression systems with using mouse anti-C9 monoclonal (1D4) antibody. Expression was also performed in each system without adding cDNA to test the non-specific binding (background) of 1D4 antibody. Arrow indicates GLP-1R. M: molecular weight marker.

Figure 4.2 (A) Schematic illustration of size-based filtration used to separate polymersomes from free proteins and impurities. (B) Western blot analysis of fractions after filtration. Lane 1: soluble fraction of cell free expression; Lane 2: flow-through; Lane 3: retentate; Lane 4: expression without additional plasmid DNA (negative control). Arrow indicates the position of GLP-1R protein.

Figure 4.3 DNA gel electrophoresis of (A) GLP-1R-pCMVTNT digestion and (B) colony test digest. (A) Lane 1: digested GLP-1R in pCMVTNT. Lane 2: digested pEU vector. (B) Lane 1-7: different colonies picked for test digestion. (A, B) Arrows indicate digested vectors and inserted gene. (C) Western blot analysis of cell free expressed GLP-1R using bilayer format. Lane 1: soluble fraction of reaction mixture. Lane 2: filtration retentate. Arrows indicate GLP-1R.

Figure 4.4 SDS-PAGE and total protein SYPRO staining of cell free expressed GLP-1R after size based filtration. Lane; flow-through. Lane 2: filtration retentate. Lane 3: filtration retentate of expression reaction without additional mRNA template (background, negative control). M: molecular weight marker. Arrow indicates the GLP-1R.

Figure 4.5 Expression of GLP-1R in CFPS and biotin- or histidine-affinity purification. Biotin tags were introduced by mixing AB di-block polymer with DSPE-PEG biotin lipids

during vesicle formation. Biotinylated AB polymersomes were added to cell free reaction mixture followed by two purification routes.

Figure 5.1 Dynamic light scattering of AB polymersomes. Top: intensity-weighted size distribution, bottom: correlation coefficient decay. The average size of AB polymersomes was determined to be 170 nm with polydispersity index of 0.083.

Figure 5.2 Transmission electron microscopy images of (A) AB polymersomes, (B) GLP-1R proteopolymersomes before purification and (C) biotin-affinity purified GLP-1R proteopolymersomes. Arrows indicate vesicular structures. (D) Dynamic light scattering measurements of biotin- and histidine- affinity purified GLP-1R. The measured sizes are indicated in the figure with PDI of 0.310 ± 0.097 and 0.825 ± 0.303 for biotin- and histidine- affinity purified samples.

Figure 5.3 Molecular cloning of the fusion protein GFP-GLP-1R. DNA gel electrophoresis of (A) PCR amplified GFP gene from pGFP vector, (B) digested GFP gene (insert, 0.7 kb) and pEU vector containing GLP-1R before ligation, (C) test digest of GFP-GLP-1R-containing pEU vector. Lane 1-5: colonies picked. (D) Western blot analysis GFP-GLP-1R expression in HYWG cell free system. Lane 1: GFP-GLP-1R. M: molecular weight marker.

Figure 5.4 SDS-PAGE (A) and Western blot (B) analysis of GFP-GLP-1R purification using 1: biotin-affinity method and 2: histidine-affinity method. (B) Western blotting with a monoclonal antibody against the C9 tag at C-terminal domain of the GFP-GLP-1R. M: molecular weight marker. Arrow indicates GFP-GLP-1R.

Figure 5.5 Size-Exclusion Chromatography analysis of GLP-1R proteopolymersomes using Sepharose-based column. (A) UV traces of GLP-1R proteopolymersomes (unpurified expression mixture). Void volume and exclusion limit indicate column resolution. BSA: bovine serum albumin (66 kDa) was injected as reference to distinguish

incorporated and unincorporated GLP-1R (53 kDa). Tetrazole fluorescence (B) and Western blot (C) analysis of elution fractions from (A).

Figure 5.6 Fluorescence detection size-exclusion chromatography (FSEC) profile of cell free expressed GLP-1R proteopolymersomes before purification. Protein elution: column resolution for protein separation. Vesicle size indicates the elution volume of liposomes with defined sizes.

Figure 5.7 FSEC using calcein encapsulated AB polymersomes. Protein elution: column resolution for protein separation. Vesicle size indicates the elution volume of liposomes with defined sizes.

Figure 5.8 Fluorescence detection size-exclusion chromatography (FSEC) of GFP-GLP-1R proteopolymersomes. (A) Biotin- and (B) histidine-affinity purified GFP-GLP-1R proteopolymersomes. (C) Cell free protein synthesis of GFP-GLP-1R protein without supplied polymersomes. The sample was directly analyzed after synthesis without any purification. Protein elution: column resolution for protein separation. Vesicle size indicates the elution volume of liposomes with defined sizes.

Figure 6.1 Characterization of antibody conformation sensitivity and sample preparation. (A) Western blot analysis of GLP-1R using 1D4 and pAb. Lane 1: GLP-1R expressed using cell free protein synthesis (CFPS). Lane 2: CFPS without additional mRNA (negative control). (B) SDS-PAGE and Western blot of unpurified and biotin-affinity purified chemokine CXC receptor-4 (CXCR4) using anti-C9 antibody. SN: soluble fraction of cell free reaction mixture. Pure: biotin-affinity purified CXCR4 proteopolymersomes. M: molecular weight marker.

Figure 6.2 SPR sensorgrams of proteopolymersomes immobilization using antibody capture method. Immobilization of (A) unpurified and (B) purified proteopolymersomes using antibody capture. (A, B) Anti-PEG antibody was first immobilized onto biosensor chip using standard amine coupling reaction, followed by flowing proteopolymersomes

onto surface to enable antibody capture. The immobilization level was 3000 RU for unpurified and 1000 RU for biotin-affinity purified proteopolymersomes with no significant difference between GLP-1R and CXCR4.

Figure 6.3 SPR sensorgrams of NTD specific pAb binding to (A) unpurified and (B) purified GLP-1R proteopolymersomes. Sensorgram of the binding for (from bottom to top) increasing concentrations of pAb antibody. Binding curves were fitted to a 1:1 kinetic binding model and an affinity of 18.6 nM was derived from the fitting.

Figure 6.4 SPR sensorgrams of peptide ligand exendin-4 binding to proteopolymersomes. Unpurified (A) and (B) biotin-affinity purified GLP-1R and CXCR4 proteopolymersomes binding to exendin-4 and 1D4 antibody. Proteopolymersomes were immobilized onto biosensor chip using anti-PEG capture. Exendin-4 of from 0.1 to 1,000 nM was flowed over immobilized proteopolymersomes with buffer injections in between. 1D4 mAb: 1D4 monoclonal antibody targets C-terminal tag of GLP-1R and CXCR4.

Figure 6.5 GLP-1R is retained in polymersomes. Specific binding of [I-125] GLP-1 to (A) GLP-1R membrane preparation (positive control), GLP-1R AB proteopolymersomes purified using (B) filtration, (C) biotin-affinity and (D) histidine-affinity purification. Dissociation constant (K_d) of membrane preparation was calculated to be 0.24 nM, with a Bmax of binding of 2.2 pmol/mg of protein for membrane preparation (A). Top: specific binding, fitted saturation binding curves using 1:1 model. Bottom: raw binding curves. TB: total binding. NSB: non-specific binding.

Figure 6.6 Production of GLP-1R in the presence of Fos14. (A) Western blot and (B) SDS-PAGE analysis of purified proteins. (A, B) Lane 1: biotin-affinity purified GLP-1R proteopolymersomes. Lane 2: histidine-affinity GLP-1R in the absence of Fos14 without polymersomes. Lane 3: sample in lane 2 after Fos14 removal. Lane 4: non-relevant target DRD2 purified as the sample in lane 1. M: molecular weight marker. Sample 1, 3 and 4 were used for the following radioligand binding assay. (C) Tetrazole fluorescence of

fractions from biotin-affinity purification of GLP-1R proteopolymersomes produced using CFPS with supplied AB polymersomes and Fos14.

Figure 6.7 Transmission electron microscopy (TEM) images (A, B) of biotin-affinity purified GLP-1R proteopolymersomes produced using CFPS with supplied AB polymersomes and Fos14. Arrows indicate vesicular structures.

Figure 6.8 Fluorescence detection size-exclusion chromatography (FSEC) of cell free expressed GFP-GLP-1R in the presence of Fos14. Cell free reaction mixture. Analysis of (A) cell free reaction mixture before purification and (B) biotin-affinity purified proteopolymersomes. Top annotations: composition of eluent at each volume. Vesicle size indicates the elution volume of liposomes with defined sizes. Cartoons represent the composition of each fraction.

Figure 6.9 Functional characterization of GLP-1R proteopolymersomes using radioligand binding assay. Specific binding of peptide ligand [I-125] GLP-1 to biotin-affinity purified (A) GLP-1R proteopolymersomes and (B) GLP-1R produced in the absence of AB polymersomes. Top: specific and fitted binding curves using 1:1 model. Bottom: raw binding curves. TB: total binding. NSB: non-specific binding. Dissociation constant (K_d) of GLP-1R proteopolymersomes (A) was calculated to be 0.37 nM, with a Bmax of binding of 1.5 pmol/mg. In competition binding assay, the half effective concentration (EC_{50}) was determined to be 54.3 and 37.8 nM for GLP-1R proteopolymersomes and membrane preparation, respectively.

Figure 6.10 Functional characterization of non-relevant target dopamine receptor D2 (DRD2). Radioligand binding assay of [I-125] GLP-1 binding to biotin-affinity purified DRD2 proteopolymersomes produced in the presence of Fos14. Top: specific and fitted binding curves using 1:1 model. Bottom: raw binding curves. TB: total binding. NSB: non-specific binding. No specific binding was detected.

Figure 7.1 Cell free protein synthesis of GLP-1R with diblock and triblock copolymers. (A) Western blot and SDS-PAGE analysis of GLP-1R expressed in the presence of detergent, digitonin (D) and/or Fos14 (F), and copolymers (AB: diblock, ABA: triblock). Lane 1: DF-AB. Lane 2: F-AB. Lane 3: F-ABA. (B-D) Saturation binding of unpurified GLP-1R proteopolymersomes expressed in different combinations of detergent(s) and copolymers (A, lane 1-3), to radioligand [I-125] GLP-1. No specific binding was detected.

Figure 7.2 Biotin-affinity purification and functional characterization of GLP-1R proteopolymersomes. (A) Western blot and SDS-PAGE analysis of biotin-affinity purified GLP-1R expressed in the presence different detergent/polymer combinations. Lane 1: DF-AB. Lane 2: F-AB. Lane 3: F-ABA. (B, C) Saturation binding of biotin-affinity purified GLP-1R AB (B) and ABA (C) proteopolymersomes (samples in lane 2-3) to [I-125] GLP-1. Dissociation constant (K_d) of purified GLP-1R proteopolymersomes expressed in the presence of F-AB and F-ABA was determined to be 1.7 and 2.1 nM, respectively. No specific binding of (D) ABA to GLP-1 and (E) GLP-1 to filter plate was detected.

Figure 7.3 (A) SDS-PAGE of D-CF expressed and histidine-affinity purified GLP-1R (lane 1), followed by reconstitution into AB (lane 2) or ABA (lane 3). Purified GLP-1R is boxed in red. Radioligand binding assay of (B) histidine-affinity purified GLP-1R solubilized in Fos14 before reconstitution, and after reconstitution into (C) AB or (D) ABA. (E) Competition binding assays of reconstituted GLP-1R into ABA or GLP-1R derived from membrane preparation, with peptide ligand exendin-4. [I-125] GLP-1 (RL) concentration was indicated by nM.

Figure 7.4 (A) SDS-PAGE and (B) Western blot of expressed in detergent and histidine-affinity purified GFP-GLP-1R (lane 1) and DRD2 (lane 2), followed by reconstitution into ABA polymersomes. (A, B) Blue box: proteolysis of the fusion protein or alternative translation initiation at the methionine. Radioligand binding assay of GFP-GLP-1R before (C) and after (D) reconstitution into ABA. (E) Purified DRD2 after reconstitution. [I-125] GLP-1 (RL) concentration was in nM. The K_d was determined to be 1.1 nM (C) and 1.7 nM (D), respectively.

Figure 8.1 Production of ion channel TREK-2 using established protocol with polymer membrane support. (A) Western blot analysis of purified TREK-2 using size-based filtration method. The retention of TREK-2 indicates its incorporation into AB polymersomes. Lane 1: soluble fraction of reaction mixture. Lane 2: filtration retentate. (B) Western blot and SDS-PAGE analysis of biotin-affinity purification of TREK-2. SDS-PAGE gel was stained with Coomassie blue. Western blot with a specific antibody against the C9 tag at C-terminal of TREK-2. M: molecular weight marker. Arrows indicate TREK-2.

Abbreviations

AB	Poly(butadiene) ₂₂ -poly(ethylene glycol) ₁₃
ABA	Poly(2-methyloxazoline) ₂₀ -poly(dimethylsiloxane) ₅₄ -poly(2-methyloxazoline) ₂₀
β2AR	β-2 Adrenergic Receptor
CFPS	Cell Free Protein Synthesis
DDM	n-Dodecyl β-D-maltoside
CXCR4	C-X-C Chemokine Receptor-4
DRD2	Dopamine Receptor-2
DLS	Dynamic Light Scattering
ELISA	Enzyme-Linked immunosorbent assay
FP	Fluorescence Polarization
FSEC	Fluorescence Detection Size Exclusion Chromatography
GLP-1R	Glucagon like Peptide-1 Receptor
GPCR	G-protein Coupled Receptor
GFP	Green Fluorescence Protein
HYWG	High Yield Wheat Germ Extract
MP	Membrane Protein
RRL	Rabbit Reticulocyte Lysate
PEG	Polyethylene glycol
PBD	Polybutadiene
pAb	Polyclonal Antibody
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SPR	Surface Plasmon Resonance
TREK-2	The Two-Pore-Domain K ⁺ Channel
TEM	Transmission Electron Microscopy
WG	Wheat Germ Extract
1D4	Anti-Rhodopsin (C9 tag) Antibody
7TM	7 Transmembrane Domain

Chapter 1

Introduction

Membrane proteins are important components in cell membranes and play a pivotal role in cellular signaling and cell-cell communication. The research of membrane proteins comprises of several studies including structure, function (i.e., molecular interactions), assembly at the cellular membrane interface and downstream signaling cascade. However, the key question is to understand the basic structure and function for many unknown Membrane Proteins. For this, scientists employ different strategies ranging from cellular engineering to protein engineering in biology. In parallel, as an interdisciplinary approach, chemistry and its synthetic routes are adapted for membrane protein biology. In such an approach, construction of artificial cellular membranes for hosting Membrane Proteins has become a powerful tool to manipulate them in a controlled laboratory setting. Despite the enormous efforts, however, it still remains challenging mainly due to the inherent protein instability and structural complexity. In this thesis, I present how to achieve a synthetic membrane model system to understand complex Membrane Proteins. Furthermore, such amphiphilic self-assembled polymer vesicles can be exploited as a platform for understanding complex Membrane proteins. In this chapter, a general overview of such proposal and hypothesis, with specific objective and scope, are presented along with the selection of chemical constituents such as self-assembled block copolymers, for the construction of artificial cell membranes.

1.1 Background and Hypothesis

The basic structure of a cell membrane is provided and maintained by a fluid lipid bilayer assembly and mosaic organization of membrane proteins, first conceptualized by Singer and Nicolson ^[1]. The membrane proteins (MPs) span across the membrane, with their transmembrane domain(s) interacting with the hydrophobic region of the lipid bilayer and their hydrophilic domains exposed on either side of the membranes ^[2]. Such structural organization connects cytosol to extracellular environment and provide functionality to the lipid membrane, such as sensing ^[3], signaling transduction ^[4] and ion transport ^[5].

The first major bottleneck in membrane protein functional and structural characterization is the protein expression and purification ^[2]. Despite similar sequence based properties to soluble proteins, they are characterized by the alternative transmembrane regions that consist of almost exclusively hydrophobic amino acids like alanines and glycines. Compared to soluble proteins, these helical regions display closer packing (e.g. 10 % less packing angles of helix bundles) and more buried interfacial surface (about 25 %) resulting from the surrounding lipid environment ^[6]. In real practice, MPs tend to lose the original conformations once isolated from the native environment. Although isolation of endogenous MPs is possible in some cases, the general low abundance even after overexpression hinders further structural studies. In addition, MPs could also be produced using cell-free protein synthesis (CFPS) ^[7]. Unlike recombinant expression in host cells or endogenous overexpression, CFPS is independent of live cells. Instead, cell lysate, a fluid of the lysed cells containing translational machinery (e.g. ribosomes, translation factors, metabolic enzymes) is used with supplied essential elements for protein synthesis, e.g. DNA or RNA template of protein of interest, RNA polymerase cofactors, energy resources, etc. Moreover, CFPS is an open reaction system that allows more extensively modified reactions to facilitate, for instance, MP expression as compared to traditional heterologous expression systems ^[7-10]. For MP expression using CFPS, the cell lysate can be further supplemented with components necessary for stable MP production, in particular amphiphiles like detergent, lipid or block copolymer to achieve spontaneous protein anchoring (co-translational incorporation) as expression proceeds. The addition of co-

translational amphiphilic reagents can assist in protein folding and stabilization of the hydrophobic regions of the MP. However, the presence of some of these amphiphiles may inhibit protein synthesis. Furthermore, detergents and lipids exhibit an inherent physicochemical instability as reflected by the critical micelle concentration (CMC) at which amphiphiles start to self-assemble into stable vesicular structures. Typically, detergents have highest CMCs of about 10^{-3} - 10^{-6} M and least stability compared to lipid vesicles which have CMCs of about 10^{-6} - 10^{-9} M. In comparison, the CMC of bilayer forming self-assembled block polymers (polymersomes) can be easily tuned to a few orders of magnitude smaller by changing chain length or building blocks. In the *in-vivo* environment, polymer carriers display significant longer circulation time (up to over 40 h) compared to lipid vesicles (10-15 h) ^[11-12]. The superior stability and chemical tunability of polymersomes make them ideal candidates as supplementary amphiphiles in CFPS ^[13-15].

We herein hypothesize that polymersomes may act as suitable biomimetic membranes that can be used to support cell-free expressions of broad classes of MPs in their functional state. Building on existing foundation, we propose to develop a comprehensive synthetic model membrane that can assist in production of a functional multi-pass MPs with difficult class of MP (Class B GPCR-GLP-1R) as a model protein, including functional characterization to understand the complexity of these multi-pass MPs in a synthetic membrane (i.e., integration and association into polymersome membranes).

1.2 Objectives and Scope

To test our hypothesis and rigorously assess our synthetic model membrane (artificial cell membranes) firstly, membrane protein targets of interests have to be selected for analytical scale production and characterizations. In this thesis, we aim to challenge MP targets that cannot be successfully approached with conventional methods like *in-vivo* expression, detergent isolation. The glucagon like peptide-1 receptor (GLP-1R) and potassium channel subfamily K member 10 (TREK-2), representing G-protein coupled receptor (GPCR) superfamily and ion channel, respectively, were chosen. The expression of GLP-1R and

TREK-2 to date has only reached similar levels in native tissues ^[16-17]. Besides, no functional proteins were ever isolated with current methodologies, which has limited further structural and functional characterizations. Thus, the optimal expression and purification conditions have to be determined, including expression systems and the block copolymers. Lastly, purified MPs have to be shown to be functional in assays customized for each MP.

We propose specific objectives as follow:

1. Cell Free Expression of GLP-1R and TREK-2

The plasmid gene encoding GLP-1R and TREK-2 will be used in three established cell free protein synthesis systems, namely wheat germ, *E. coli*, and rabbit reticulocyte lysate-based systems as mammalian and non-mammalian protein expression. Diblock copolymers as amphiphilic reagents will be supplied to the synthesis. The one with the highest expression level as detected by Western blot will be used for further purification and characterization. Following expression, the test of incorporation will be using a size-based filtration unit that retains only large vesicular polymersomes but not free impurities or unincorporated target proteins.

2. Purification of Cell Free Expressed MPs

The expressed MPs will be purified using a size-based filtration and two affinity-purifications methods based on different locations of tagging (either on protein or on polymer membrane surface). The methods will be evaluated based on the purity, quantity, and association of protein/polymer support of purified MPs. Primary characterizations will be performed, including Western blot, SDS-PAGE and fluorescence spectroscopy.

3. Structural and Functional Characterization of MPs

Cell free expressed MPs (both unpurified and purified) will be extensively characterized using transmission electron microscopy (TEM), fluorescence detection size-exclusion chromatography (FSEC), surface plasmon resonance (SPR)

and radioligand binding assay. These characterizations will provide an overview of produced MP in terms of structural characteristics, incorporation into polymersomes and functionality.

4. Production of Functional MPs with Block Copolymers

Functional production of MPs with different block copolymers with focus on diblock (AB) and triblock (ABA) polymers. Different production strategies, including co-translational incorporation and reconstitution will be compared. Produced MP in block copolymers (proteopolymersomes) will be tested for functionality using radioligand binding assay.

1.3 Dissertation Overview

This thesis addresses the development of a comprehensive synthetic model for MP production and polymersomes as stable membrane scaffolds. Fundamental and practical issues are presented and discussed.

Chapter 1 provides a background introduction of the thesis, rationales for the research and outlines the goals and scope.

Chapter 2 reviews the literature concerning traditional methods used in MP research, recent scientific and technical progresses and challenges. Conventional *in vivo* MP expression methods, such as in bacterial, insect and mammalian cells are discussed in comparison to the novel cell free expression. This chapter also reviews the currently available artificial membrane mimetics, e.g. lipids and polymers. The chapter ends with the discussion of block copolymers as potential membrane support.

Chapter 3 discusses the principles underlying the expression/characterization/assay techniques employed, the methods of data analysis, including development, and interpretation of results and its significance.

Chapter 4 presents the results of the expression of GLP-1R in cell free system with focus on protein expression level, incorporation into polymersomes (from diblock copolymer, AB), protein purification and biophysical characterizations. The expression and incorporation of GLP-1R in different cell free systems are evaluated and compared. Further, biotin- and histidine-affinity purified GLP-1R was also characterized.

Chapter 5 presents biophysical characterization of purified GLP-1R using different methods with focus on polymersome structure and protein/polymersome association state using transmission electron microscopy (TEM) and fluorescence detection size-exclusion chromatography (FSEC) primarily.

Chapter 6 presents biochemical and functional characterization of purified GLP-1R using surface plasmon resonance (SPR) and radioligand binding assay. A modified cell free system supplemented with chemical chaperone is further tested and is shown to produce functional GLP-1R.

Chapter 7 describes the extension of such strategy for functional production of MPs and how this principle can be applied using different block copolymers in CFPS. Primary expression and characterization of CFPS expressed MP showed both diblock (AB) and triblock (ABA) copolymers are good artificial membrane supports to keep the native folding of MP.

Chapter 8 gives a concise summary of the main findings. Critical findings are discussed and outlook for future strategies and potential applications was provided.

1.4 Findings and Outcomes/Originality

This thesis led to several novel outcomes.

1. Establishment of an affinity-purification protocol for expression of MPs using cell free protein synthesis supplemented with polymersomes.

2. Expression of multi-pass MPs (GLP-1R, a GPCR) in their functional state and demonstrated production of TREK-2 (an ion channel) with the protocols developed during this study.
3. Establishment of a workflow for characterizing the biochemical and physical properties of purified GPCR proteopolymersomes, including expression level, association with polymers, orientation, structure and functionality.
4. Demonstration that MPs produced in artificial cell membranes mimic the pharmacokinetics (ability to bind and compete natural ligands) of a naturally produced MP.
5. Setup of an improved cell free protein synthesis platform, including expression, purification and characterization work-flows, for future development of functional multi-pass MPs using block copolymers as membrane support.

References

- [1] S. J. Singer, G. L. Nicolson, *Science* **1972**, 175, 720-731.
- [2] J. N. Sachs, D. M. Engelman, *Annual Review of Biochemistry* **2006**, 75, 707-712.
- [3] F. Bezanilla, *Nat Rev Mol Cell Biol* **2008**, 9, 323-332.
- [4] A. Jazayeri, J. M. Dias, F. H. Marshall, *Journal of Biological Chemistry* **2015**.
- [5] E. C. Cooper, L. Y. Jan, *Proceedings of the National Academy of Sciences of the United States of America* **1999**, 96, 4759-4766.
- [6] M. Gimpelev, L. R. Forrest, D. Murray, B. Honig, *Biophysical Journal* **2004**, 87, 4075-4086.
- [7] E. Carlson, R. Gan, C. Hodgman, M. Jewett, *Biotechnology advances* **2012**, 30, 1185-1194.
- [8] H. P. de Hoog, E. M. Lin JieRong, S. Banerjee, F. M. Decailot, M. Nallani, *PLoS One* **2014**, 9, e110847.
- [9] F. Katzen, T. C. Peterson, W. Kudlicki, *Trends in biotechnology* **2009**, 27, 455-460.
- [10] F. Bernhard, Y. Tozawa, *Current Opinion in Structural Biology* **2013**, 23, 374-380.
- [11] V. Pata, N. Dan, *Biophysical Journal* **2003**, 85, 2111-2118.

- [12] J. S. Lee, M. Ankone, E. Pieters, R. M. Schiffelers, W. E. Hennink, J. Feijen, *Journal of Controlled Release* **2011**, 155, 282-288.
- [13] H.-P. M. de Hoog, E. M. Lin JieRong, S. Banerjee, F. M. Décaillot, M. Nallani, *PLoS ONE* **2014**, 9, e110847.
- [14] S. May, M. Andreasson-Ochsner, Z. Fu, Y. X. Low, D. Tan, H.-P. M. de Hoog, S. Ritz, M. Nallani, E.-K. Sinner, *Angewandte Chemie International Edition* **2013**, 52, 749-753.
- [15] M. Nallani, M. Andreasson-Ochsner, C.-W. D. Tan, E.-K. Sinner, Y. Wisantoso, S. Geifman-Shochat, W. Hunziker, *Biointerphases* **2011**, 6, 153-157.
- [16] S. G. Brohawn, *Annals of the New York Academy of Sciences* **2015**, 1352, 20-32.
- [17] R. J. Mann, N. E. Nasr, J. K. Sinfield, E. Paci, D. Donnelly, *British Journal of Pharmacology* **2010**, 160, 1973-1984.

Chapter 2

Literature Review

Membrane proteins (MPs) are gatekeepers of cells that play an important role in a number of cellular functions. In order to maintain the functioning of cells, the MPs are elaborately regulated via intracellular mechanisms, e.g. expression, trafficking, folding and modifications, and respond to stimulus signal present in the outside environment. Structurally, the multi-pass transmembrane proteins consist of alternating hydrophobic transmembrane domains that interact with lipid bilayer membranes in cells. However, the structural complexity has made MPs extremely difficult to work with mainly because MPs present in lipid membranes are of low abundance and unstable once extracted. In order to produce large quantity MPs of high quality, enormous efforts have been put in optimizing current expression systems and developing novel strategies. Notable, cell-free protein synthesis (CFPS) emerged as a powerful tool in MP productions. As a source of native-like membrane support in CFPS, self-assembled block copolymers have demonstrated great potential. They provide several advantages over conventional vesicular or amphiphilic structures that are supplied in CFPS. In this chapter, background information is first provided about the basic cellular biology of membrane proteins, with major focus on GPCRs and ion channels. Next, currently methods and challenges in GPCR production are also discussed. Finally, polymersomes as potential biomimetic membranes are reviewed.

2.1 Membrane Protein Structure and Function

Membrane Proteins (MPs) are essential components of biological membranes in nature ^[1]. According to the Human Genome Project, it is estimated that about 20 % of total gene count code for MPs ^[2], indicating their importance for survival. MPs are classified into two broad categories, integral and peripheral, based on the nature of the membrane-protein interactions ^[3]. Integral MPs have one or more segments embedded in the phospholipid bilayers of cell membranes. Most integral proteins contain hydrophobic domains that interact with fatty acyl groups of phospholipid, thus anchoring the protein into the membranes. The peripheral proteins interact with integral MPs instead of the hydrophobic core of phospholipid. This thesis focuses on integral proteins, which have one or more membrane-spanning domains across the phospholipid bilayers. These proteins have complex conformations that bridge extracellular environment and intracellular cytosol and regulate cellular functions in a dynamic way with interactions to lipid bilayers. As two major classes of integral proteins, G-protein coupled receptors (GPCRs) and ion channels have become a focus in MP research.

2.1.1 G-protein Coupled Receptors

As the largest family of membrane proteins, GPCRs have about 1000 members in human genome ^[4]. Based on sequence homology, GPCRs are divided into six families, of which about 85 % classified as rhodopsin-like family (Class A GPCR). Since the first crystal structure of a mammalian GPCR of this family, the bovine rhodopsin, was solved in 2000, increasing knowledge has been gained in the structure-function relationship of GPCRs ^[5, 6]. To date, 35 structures have been solved and a number of these were in complex with respective ligands, e.g. β 2-adrenergic receptor (β 2AR) and chemokine CXCR4 (CXCR4) ^[7]. All GPCRs across different families feature a signature seven transmembrane topology, which shares very little genetic sequence identity (< 20 %) ^[8]. These transmembrane domains are formed by α -helical bundles made of almost exclusively hydrophobic amino acids like glycines and alanines and span across phospholipid bilayer in cell membranes, resulting in closer packing (e.g. 10 % less packing angles of helix

bundles) and more buried interfacial surface (about 25 %) as compared to soluble proteins^[9]. In addition to these features, receptors in the rhodopsin-like family are characterized a disulphide bridge that connects the first and second extracellular loops (ECLs) (Figure 2.1A). The transmembrane domains also contain “tilted” and “kinked” bundles in the second, sixth and seventh domains^[10]. The rhodopsin-like family contains receptors for odorants, small molecules such as and amines and the catecholamines^[10].

The secretin-like (Class B) GPCRs constitute the second largest family in the GPCR superfamily. In comparison to the rhodopsin-like family, the research of secretin-like GPCRs have been considerably delayed^[10]. Besides the seven-transmembrane topology and the common disulphide bridge that connects the first and second ECLs, very little information is known about the structural characteristics, especially the orientation of the transmembrane domains. Different from rhodopsin-like family, the secretin-like GPCRs feature a long N-terminal extracellular domain that contains several cysteines, which presumably form a network of disulphide bridges (Figure 2.1B)^[10]. These properties distinguish them from rhodopsin-like family and tremendously add up to the overall structural complexity and instability^[11, 12]. As a result, only limited structures were solved to date, namely the human glucagon receptor (GCGR)^[13] and human corticotropin-releasing factor receptor 1 (CRF1)^[13, 14]. In both constructs, the N-terminal domain was completely removed to stabilize the protein constructs.

The members in secretin-like family are receptors specifically for hormone sensing. The Glucagon like peptide-1 receptor (GLP-1R) is the receptor for the glucoregulatory hormone in type II diabetes^[15]. Besides the characteristics in type I diabetes, the type II is characterized by insulin deficiency/resistance, which is commonly believed to be caused a combination of genetic and environmental factors^[16]. This makes conventional treatment like insulin much less efficient as for type I diabetes. Alternatively, targeting receptors with non-insulin ligands became a promising approach. Among all the receptors being targeted, GLP-1R ligands present a unique benefit that includes not only control of blood glucose level but improvement of co-morbid conditions, such as hypertension, hyperlipidemia, and obesity^[17]. The current treatments of GLP-1R are based on active small molecule drugs,

e.g. exendin-4 and GLP-1. Both of these molecules are synthetic versions of peptides found in natural tissues but due to its short circulation time (~2.4 h), it has to be medicated twice per day via subcutaneous injection ^[18]. Meanwhile, a long lasting active drug for GLP-1R is highly desirable and would assist the diabetes treatment. In comparison, the treatment of type I diabetes has been significantly advanced with an antibody-based therapy ^[19]. The immune T-cell suppressing antibody anti-CD3 preserved of insulin production in patients with new-onset type 1 diabetes for even beyond 1 year after treatment. The sustained insulin production was accompanied by improvement of glucose level and reduced use of insulin. It is not surprising therapeutic antibodies display high efficacy given their superior specificity towards antigen epitopes compared to small molecules. In type II diabetes treatment, however, the development of therapeutic antibodies has long been hindered by the limited availability of robust antigens with stable and functional present target receptor like GLP-1R. A key challenge is, like the cases with many other membrane protein targets, the quality production of target receptors.

Figure 2.1 Schematic representation of the predicted topology of (A) rhodopsin-like family GPCR, the human chemokine CXCR4 (CXCR4) ^[20] and (B) secretin-like family GPCR, the rat Glucagon like peptide-1 receptor (GLP-1R). The GLP-1R has long N- (126 amino acids) and C- (55 amino acids) terminal domain compared to CXCR4 (39 and 44 amino acids, respectively). (Reprinted with permission from Ref. ^[20])

The complexity of GPCRs is better understood in combination with their functions in signaling transduction. From a physiological point of view, GPCRs are of scientific interest because they are important drug targets (such as GLP-1R), and they play a pivotal role in the cellular signaling network which regulates cell growth and proliferation ^[21]. In such processes, GPCRs act as gate-keepers responsible for sensing and responding to a number of stimuli within the external environment, including light, hormone, odorant, antigen, neurotransmitter or the surface of another cell ^[22]. These stimuli were mainly transduced through G-protein to a variety of downstream signaling pathways. Upon receptor activation, signal transduces across 7TM promotes the exchange of GDP/GTP associated with the $G\alpha$ subunit, causing G-protein subunits (α , β , γ) dissociation and activate downstream signaling cascades ^[23]. The intrinsic GTPase activity of $G\alpha$ converts bound GTP into GDP and returns the signaling cascade to initial state in which $G\alpha$ is bound to $G\beta\gamma$ dimer and guanosine diphosphate (GDP). Many of the downstream messengers can be activated via the G-protein signalling, including cAMP (Cyclic adenosine monophosphate), Ca^{2+} , etc. The selected target GLP-1R in the study is an important therapeutic target of type 2 diabetes.

2.1.2 Ion Channels

Ion channels, as another important family of membrane proteins, are also subject to G-protein subunit modulation initiated by GPCRs ^[24]. These pore-forming proteins have at least 250 members in human genome ^[25] and 15 of which code for potassium (K^+) channels with two pore domains (K_{2P}) ^[26]. Potassium channels participate in many biological functions, from electron potential modulation to ion homeostasis ^[26]. They are involved in a large number of neurological disorder related diseases, such as epilepsy, headache, deafness, etc. ^[27]. These channel proteins constitute dimers of pore-forming subunits responsible for producing background conductance. The structural insights of potassium channels were first reported by Doyle et al. that they contain a short signature sequence called the pore (P) domain ^[28]. Different from GPCRs, K_{2P} contain six transmembrane domains, two of which are the short pore-forming P domains, and intracellular N- and C-termini (Figure 2.2). Recently, the crystal structure of the pore-forming domain of K_{2P}

channel TREK-2 was resolved by Yin et al. at up to 3.4 angstrom resolution ^[29]. The results demonstrate how the movement of the pore-lining helices is responsible for K_{2P} channel regulation and illustrate how antidepressant drug Prozac contributes to conformation switching.

Figure 2.2 Schematic representation of the predicted topology of (A) secretin-like family GPCR protein ^[10] and (B) K⁺ ion channel Kv7.1 ^[30]. GPCR has structural feature of seven-transmembrane domains with extracellular N-terminal and intracellular C-terminal domains. In comparison, ion channels display six-transmembrane domains with both N- and C- terminal domains in the cytosol. (Reprinted with permission from Ref. ^[10, 30], Copyright: © 2012 Wrobel)

2.2 Membrane Protein Production and Characterization

The key question is to understand the basic structure and function of MPs, in particular for those that show clinical importance and whose functions are unknown or difficult to characterize. However, MP research is accompanied by challenges at all levels, especially protein production. Technically challenging recombinant expression and purification of MPs ^[7, 31], as well as their general low abundance in host cells (for direct extraction) are major hurdles for advances in MP research. Thus, production with interdisciplinary engineering approaches become important source of membrane protein materials. These challenges have been approached from various perspectives, including expression system optimization^[32], protein engineering^[33] and technical advances in protein structure biology ^[34]. A powerful tool that enables manipulation under controlled laboratory settings has also

been realized by the construction of artificial cellular membranes that host the expressed MPs.

2.2.1 *In Vivo* Membrane Protein Expression Systems

Recombinant protein expression in *E. coli*, the most well understood and established expression system, proved to be the best for prokaryotic integral membrane protein expression^[35]. It is attractive because of its easy accessibility, high growth rate, scalability and low cost of culture. However, as for all prokaryotic expression systems, they are less favored for the lack of essential cellular organelles, e.g. nucleus, endoplasmic reticulum, Golgi apparatus, etc. to express active membrane protein with correct post-translational modifications^[36]. Thus, the use of *E. coli* system to express eukaryotic protein targets bears the risk of low yield production and producing non-functional proteins. Despite these limitations, a number of eukaryotic targets were reported to have been successfully expressed in *E. coli*, including human β 2-adrenergic receptor (β 2AR, a GPCR)^[37], potassium channel Kir3.1 (an ion channel)^[38], and eukaryotic oxidoreductase (transmembrane protein)^[39].

In comparison to the prokaryotic systems, eukaryotic expression systems are more compatible mainly because MPs, with very few exceptions, require eukaryotic environment when expressed in heterologous host^[7]. The baculovirus infected insect cell system is the dominant heterologous expression system for producing eukaryotic MPs because it closely mimics mammalian expression systems^[40]. The mammalian and non-mammalian protein expression are different due the complex cellular structures and the most widely used insect cell line, Sf9 can produce sufficient amount of protein for extensive characterization including crystallization. Examples of proteins include CXCR4^[12], A2A^[41] and β 2AR^[42]. Mammalian expression systems, on the other hand, are popular recombinant production systems for its proper post-translational modification (PTM) and human protein-like folding mechanisms^[43]. Compared to the insect cell system, however, the protein yield from the mammalian expression system is typically lower compared to other systems. By transfecting the human embryonic kidney cell line (HEK293)^[44], for

example, protein yield is sufficient only for radioligand binding assays and Western blotting analysis (assuming good signal to noise ratio), but not biophysical characterizations or pre-crystallization detergent screening. The practical consequence of the difference in expression levels between insect and mammalian cells is best reflected by the numbers of crystallized membrane proteins produced in each system. As reviewed by Yuan He et al., of all the membrane structures resolved till 2014, 35 out of 63 were from insect cell system as compared to only 4 from mammalian cell system. D

2.2.2 Membrane Protein Engineering

Compared to the total number of membrane proteins encoded by human genome, the current achievement of functional recombinant expression is of limited success. For GPCRs, the 35 structures solved represent a very small portion of the almost 1000 targets, leaving the majority unexplored to date. The reasons are two folds. First, the existence of self-regulatory mechanisms in cell membranes, such as desensitization and internalization^[45], prevent exogenous membrane proteins from over accumulation by either digesting or denaturing them. It is reasonable that under such circumstances, membrane protein expression level is kept low or the expressed proteins are non-functional. Second, despite efforts in optimizing the external systems, such as host cells, plasmids and culture conditions, membrane protein itself is still limited by its intrinsic instability, defined by the inherent protein sequence^[32]. For GPCRs, for example, the third intracellular domain is known to have a flexible conformation and considered a weak point in the 7TM^[6]. In such case, it is unlikely to produce functional, stabilized proteins simply by optimizing expression system or culture condition. Thus, in pursuit of high yield and quality production with original membrane protein sequences, efforts have been made to stabilize membrane proteins on the genetic level (protein engineering)^[46].

A number of engineering solutions have led to significant advances to date^[34] which fall into two general categories, rational protein design and directed evolution. Application of rational protein design is currently limited because of insufficient structural information deposited in the Protein Data Bank (PDB) for designing stable mutants. Meanwhile, it

makes directed evolution a popular approach where random mutagenesis is applied to protein, and mutants are subject to screening for desired qualities, such as stability, high expression level and functionality (Figure 2.3 A) ^[33]. Since it does not require much information like rational design to generate mutants, this method is commonly used in GPCR protein engineering. Stabilizing mutants can be found using error-prone PCR, DNA shuffling and systematic alanine scanning mutagenesis ^[47]. In a recently reported directed evolution screening, three GPCRs, neurotensin receptor 1 (NTR1), NK-1 receptor (NK1R) and K-type opioid receptor (KOR1) were evolved ^[33]. Both NTR1 and NK1R variants were generated in yeast with up to 50 folds in expression levels as compared to the wild type. KOR1, which has never been evolved before, showed 26-fold increase in functionality in both yeast and insect cells. Besides random mutagenesis, fusion partners like T4 lysozyme (T4L) has been universally used to stabilize GPCRs by reducing flexibility and facilitate their crystallization ^[48]. Likewise, addition of antibody fragments (e.g., nanobody) is an important factor in aiding membrane protein structure studies ^[23, 42]. The combination of protein engineering approaches above has led to several structures solved since bovine rhodopsin, including CXCR4 ^[49], A2A ^[41], β 2AR ^[42].

Protein engineering is not without its drawback. Cautions have to be taken when stabilizing MPs with this approach as it may unpredictably alter their structural properties. A good example is illustrated by two β 2AR fusion proteins. Of the two stabilized structures resolved, only T4L but not nanobody (Fab) fused β 2AR resulted in functional ligand binding receptors ^[50]. As limited information on β 2AR is available, it is difficult to evaluate the structural and functional effects of protein engineering even though high resolution structures were resolved. In another case, to stabilize CRF1, a combination of T4L fusion and heavy modifications altered almost half of the full-length protein ^[51]. The extensive mutation has been shown to produce a stable and functional protein but may not reflect the native biological functions of CRF1 *in vivo*.

Figure 2.3 (A) A typical workflow for directed evolution of GPCR proteins. DNA library constructed from randomized GPCR genes were transformed to expression host (yeast), followed by protein display, functional selection by ligand binding assay and finally selected members were recovered to embark a second round of screening ^[33]. (B) Snake plot of stabilized, functional human corticotropin-releasing factor receptor 1 (CRF1) ^[51]. Both N- (1-103) and C- (373-404) terminus were truncated with T4L inserted at second intracellular loop. Insertions and mutations were labelled red and green. (Reprinted with permission from Ref.^[33, 51])

2.2.3 Cell Free Protein Synthesis

The two complementary approaches (expression system engineering and protein engineering) have led to many achievements in the past few years. As described in the previous section, however, membrane protein production is still presented with challenges in two folds. First, the existence of self-regulatory cellular mechanisms prevents high yield

of functional proteins to be expressed. Second, the intrinsic properties of membrane proteins make them unstable hence requiring external stabilizing factors, e.g. optimization of expression systems ^[32]. Technical advances to stabilize the membrane proteins involve extensive modifications, potentially result in unpredictable functional consequences, ^[33]. Therefore, alternative methodologies are highly desired to study membrane protein targets in their native sequence.

To meet the challenges, an alternative production methodology referred to as cell-free protein synthesis (CFPS), is used ^[52]. The most distinguishable difference from conventional methods is that, unlike recombinant expression in host cells, CFPS is independent of live cells (Figure 2.4). In a typical cell free expression reaction, cell lysate is used supplied with essential elements for protein translation, e.g. DNA or RNA template of protein of interest, RNA polymerase, ribosome, cofactors, energy resources, etc. The reaction is sustainable and terminated only when one or more components in the lysate fails, for example the substrates (e.g., ATP, cysteine) is depleted or enzymes inhibited by-product accumulation (e.g., inorganic phosphate) ^[52].

Cell lysate is a major component of any CFPS reaction. Currently, the sources include *E. coli* cell extract (ECE, prokaryotic) ^[53], rabbit reticulocytes lysate (RRL) ^[54], wheat germ extract (WG) ^[55], HeLa cell extract (HCL) ^[56] and insect cells extract (ICE) ^[57]. Different cell free systems retain the characteristics of respective cell type in terms of quality, yield, cost etc. ECE is undoubtedly the simplest and most cost-effective cell extract that can be prepared in large quantity. The expression system was used for development of novel broad-spectrum antibacterials in high-throughput manner to screen inhibitors of *Streptococcus pneumonia* ^[58]. The combination of cell free translation and luciferase reporter system screened 220,000 compounds and resulted approximately 50 hits with one of the compounds showed desired profile of broad-spectrum biochemical activity in bacteria. Similar strategies have been used to produce bacterial translation inhibitors as reported by Bøddeker N et al. and Kung PP et al ^[58, 59]. Later, Yutetsu Kuruma et al. developed PURE system based on cell extract of the same *E. coli* strain S30 ^[60]. The system constitutes purified components with each protein factors affinity tagged and greatly

simplify protein purification ^[61]. The product can be purified within 1 hour using affinity chromatography. The one-step “reverse” purification can be achieved with untagged target protein and was validated active dihydrofolic acid reductase (DHFR) with a yield on the order of sub milligrams in a 1 ml reaction volume. The use of ECE has been reported over 50 years and the preparation of ECE in house is possible with the well-established protocols ^[53].

Figure 2.4 Schematic illustration of *in vivo* expression and *in vitro* CFPS. *In vivo* expression starts with transformation of cDNA coding protein of interest to host cells, followed by cell culture.

Likewise, *in vitro* expression requires addition of cDNA to cell lysate, which contains all the essential elements for RNA transcription and protein translation. Both *in vivo* and *in vitro* expressed proteins require protein purification after synthesis ^[52]. (Adopted with permission from Ref. ^[52])

WG cell free expression system was developed by Yaeta Endo of Ehime University ^[62]. The current methodology is based on a bilayer diffusion system without using any membrane, first devised by Tatsuya Sawasaki et al ^[55]. Such format enables the continuous supply of substrates, together with the continuous removal of small byproducts and display more than 10 times higher yield over batch-mode. Other eukaryotic expression systems, such as ICE, and HeLa (cancer cell based), are also being developed ^[63]. They typically require more laborious extract preparation procedures and difficult to prepare in house. Commercial expression systems based on these cell extracts are currently available but yet to be validated ^[56, 57].

The nature of CFPS being an open system makes it possible for additional factors, e.g. chaperones, detergents, and finely control protein expression with flexible strategies ^[64, 65]. In comparison to *in vivo* expression systems, however, it is believed that such flexibility comes with altered protein production machinery, such as missing compartments (e.g. the endoplasmic reticulum, ER) and chaperones, during cell free extract preparation ^[66]. CFPS offers the opportunity to produce membrane proteins.

2.2.4 Biomimetic Membrane Support

A number of amphiphilic molecules have been developed to provide native-like membrane support for expressed MPs, which can be generally categorized into two classes, nanodiscs and vesicles. Both classes use the same principle that MPs are stabilized in a lipid membrane mimetic environment but with different methodologies. For nanodiscs, both lipid and polymer versions have been developed and named lipid nanodiscs and amphipols (polymeric nanodiscs). The lipid nanodiscs were first developed in Stephen Sligar lab using an encircling amphipathic helical protein belt, named membrane scaffold protein (MSP), to assemble MP targets that are transiently solubilized with a detergent in the presence of

phospholipids and MSP^[67]. MSPs are truncated forms of apolipoprotein (apo) which wrap around a patch of a lipid bilayer to form a discal particle or nanodisc. They are amphiphilic proteins with a hydrophobic surface facing the lipids, and a hydrophilic surface towards the aqueous solution. The incorporation of membrane proteins into nanodiscs is normally mediated by detergent (normally ionic detergent sodium cholate) and MSPs via reconstitution. The resulted disc structure with MP embedded in the center lipid bilayers surrounded by MSP with the size about 10 nm in diameter in which both N- and C- terminals of incorporated proteins are exposed. The nanodisc technology allows maintenance the MP soluble in solution in the absence of detergent for structural and functional studies and has been applied to a number of MPs, including bacteria rhodopsin, β 2AR, aspartate receptor TAR as well as enzymes, and conjugation factors, etc.^[68]. In CFPS, the application of lipid nanodisc has been limited by its relying on detergent solubilization whereas the successful co-translational integration of MP has been reported with only bacteriorhodopsin^[69]. However, little is known about the protein stability incorporated into lipid nanodiscs, especially for MPs and the use of additional scaffold protein MSP will limit certain structural and functional characterizations that require purified proteins. Likewise, the assembly of MPs in amphiphols also depends on detergent solubilization followed by transient reconstitution into polymeric discs^[70]. The technology was first developed by J. L Popot et al. using polyacrylate-based amphipol, A8-35 as a model system. In comparison to lipid nanodiscs, the research of amphiphols has been focused on assembling bacteriorhodopsin and investigation of improved protein stability over detergent solubilization and structural characterization of protein/polymer interactions. As expected, the amphiphols stabilized bacteriorhodopsin displays over three times increased time stability. In CFPS, a number amphiphols have been tested using bacteriorhodopsin as model protein in which the wellstudied polyacrylate-based amphipol, A8-35 inhibited protein translation. The nonionic amphipol showed good compatibility, however, no further work has been reported for its application to other targets than bacteriorhodopsin. Lipid nanodiscs and amphiphols are stable membrane supports but both rely exclusively on the selection of suitable detergent for protein solubilization, which considerably limited their applications. In addition to disc membrane structures made of single component (lipid or polymer), a hybrid lipid disc formed by styrene maleic acid (SMA) copolymer has also

been reported by Timothy J. Knowles et al. The initial study started with model protein bacteriorhodopsin ^[71]. It was demonstrated that SMA absorbs and destabilizes bacteriorhodopsin incorporated in liposomes, forming nanodiscs of 10 nm in diameter made of lipids and polymer. It was further demonstrated SMA can solubilize cell membranes expressing MP, including β 2AR ^[72] and K⁺ channel ^[73] without losing functionality. The advantage of SMA over lipid nanodiscs and amphipole is the detergent-free solubilization during all processes, which makes it a promising technology. The future research of SMA shall validate whether this technology with more challenging targets and fundamental studies of lipid/polymer/protein interactions.

In addition to nanodiscs, the vesicular membrane structures have also been developed, offering more authentic native-like environment for MP incorporation, e.g. compartmentalization. The vesicular membrane mimetics have been developed using both lipids (liposomes) and polymers (polymersomes). Compared to nanodiscs, the size of these vesicular mimetics can be controlled over a much larger range, from hundreds of nanometers (small unilamellar vesicles, SUVs) to micrometer (giant unilamellar vesicles, GUVs) vesicles via different preparation method, e.g. extrusion, electroformation ^[74]. Of the vesicular membrane mimetics, the most extensively studied is liposome. These vesicles consist of lipids that are natural components of cell membranes and different from synthetic materials, e.g. amphipols, they are usually compatible with cell free protein synthesis without causing inhibited protein translation. Different from nanodiscs, the incorporation of membrane proteins into liposomes does not require supplied MSPs or other protein chaperones. The incorporation of MPs into liposomes can be achieved mainly through two routes, reconstitution and co-translational integration. The reconstitution method starts with detergent solubilized MPs and lipids, and similar to processing of nanodiscs, MPs spontaneously assemble with lipids into the vesicular bilayer when detergent is removed, forming protein liposomes (proteoliposomes). This method has an advantage over lipid nanodisc that no impurity protein is introduced since the membrane scaffold protein required for nanodisc formation will not be required. The liposomes reconstitution technique is also more widely applied to MP research compared to nanodiscs, including GPCRs, ion channels, enzymes and virus proteins ^[75]. A key drawback of this method is

that it depends heavily on choice of detergent and target robustness and the successful reconstituted MPs to date are almost exclusively well-studied and stabilized targets, e.g. water channel protein Aqpz, bacteriorhodopsin, etc.

In an alternative approach, lipids are also supplied in CFPS in the form of vesicles (liposomes) to mimic the native like environment of cell membranes. In this approach, it is commonly believed that the proteins expressed were spontaneously assembled into the supplied liposomes, named co-translational incorporation. Although the detailed mechanism directing protein incorporation is unknown, research has been done in regards to the optimization of supplying liposomes, e.g. concentrations, lipid type ^[75]. The proteoliposomes usually have larger density than free proteins/empty liposomes, which allows the separation of proteoliposomes from reaction mixture by ultracentrifugation in a density gradient.

Figure 2.5 Schematic scaling of lipid vesicles and polymersomes membrane thickness with increasing copolymer molecular weight. PBD-PEO: poly (butadiene)-poly (ethylene oxide). EO: ethylene oxide. BD: polybutadiene. (Adopted from Ref. ^[76] with permission)

Polymersomes are self-assembled polymer vesicles composed of block copolymer amphiphiles. These synthetic amphiphiles can differ considerably in chemical and physical properties from lipids mainly in two aspects. First, the polymeric amphiphiles typically have higher molecular weight (MW), which gives the formed vesicles (polymersomes) tougher/thicker membranes and lower fluidity/permeability, e.g. due to the stronger hydrophobic interaction brought by larger number of repeating units in the hydrophobic domain (e.g. $-\text{CH}_2-\text{CH}_2-$ in poly(butadiene)-(polyethylene oxide), in diblock copolymer PBd₂₂-PEO₁₃ or AB copolymer used in this study), leading to better resistance to external strain (areal strain $\alpha=0.19\pm0.02$ compared to $\alpha\leq0.05$ for lipids) ^[77]. The MW lipid vesicles of is usually less than 1,000 and as a result, their membrane thickness is limited to 2-5 nm. In comparison, polymersomes can be easily tuned over a wide range of 8~20 nm (Figure 2.5) ^[78]. Dynamics simulation proposes a correlation between membrane thickness D , MW and a scaling exponent b , leading to

$$D \sim \text{MW}^b \quad (b \sim 0.55) \quad [79]$$

The differences in chemical properties directly affect the biophysical properties of polymeric membranes. For polymersomes of the same type (e.g. AB copolymer), the tougher/thicker membranes leads to decreased their mobility/fluidity compared to lipids membranes, as reflected by the different measured lateral diffusion coefficient D ($D_{\text{lipids}} \sim 4 \mu\text{m}^2/\text{s}$, $D_{\text{AB}} \sim 0.16 \mu\text{m}^2/\text{s}$) ^[80].

Secondly, a wider range of synthetic copolymers that can be used in forming polymersomes compared to lipids. Such diversity of copolymers becomes an advantage in tuning polymer properties to adapt different MP targets in the context of constructing artificial cell membranes. Most lipids share common structural features and hydrophobic tail of lipids is mainly made of multiple ethylene units $(-\text{CH}_2-\text{CH}_2-)_n$ (with $n=5$ to 18 typically). For copolymers, however, the choice of building blocks is almost indefinite considering the vast diversity of chemistry in polymer synthesis. Besides the PBd-PEO, the triblock copolymer PMOXA-PDMS-PMOXA (poly(2-methyloxazoline-*b*-dimethylsiloxane-*b*-2-

methyloxazoline), ABA copolymer) has been engineered as artificial membranes as polymersomes^[81]. Different from AB copolymer, the hydrophobic domain of ABA is made of PDMS (poly-dimethylsiloxane) that is more flexible than PBd because of longer bond length of Si-O bonds compared to the polyethylene backbone. This allows easier movement and conformational changes of PDMS chains and as a result, the self-assembled ABA polymersomes display higher fluidity ($D_{ABA} \sim 1.41 \mu\text{m}^2/\text{s}$ based on the formula of ABA used in this study) than AB^[82]. The same principle can be applied to decrease membrane fluidity by introducing rigid block copolymers, such as polystyrene, in which interaction is stronger due to the π -stacking between aromatic rings^[83]. In practical applications, polymersomes are currently being engineered to function as robust viral capsids that can carry, target and release actives (e.g., drugs and dyes)^[84]. Besides the biophysical properties discussed above, the increased toughness of polymersomes is also reflected by their distinct behaviors in *in vivo* studies, e.g. as drug carriers, compared to liposomes. Certain chemical groups in polymers are chemically inert, e.g. PEG. When PEG is present on the surface of polymersomes (PBD-PEO used in the thesis), it brings “stealth effect”, which refers to the minimal immune responses triggered (e.g. PBD-PEO polymersomes)^[83]. Lee et al. showed that polymersomes formed from (poly(ethylene glycol)-b-poly(D,L-lactide) (PEG-PDLLA)) harboring H³-dextran display almost 4 times longer half lifetime (from 11 h up to 50 h) and about 60 % higher accumulation in target tumors as compared to stealth liposomes with no detectable load leakage during blood circulation^[85]. Driven by higher *in-vivo* stability, the possibility of polymersomes as DNA and anti-cancer drug carriers have also been investigated^[86]. Considering such properties of polymersomes, they become a promising class of biomimetic membrane model systems.

2.2.5 Membrane Supports in Cell Free Protein Synthesis

There are three modes in cell free protein synthesis: the precipitate (P-CF) mode, the detergent (D-CF) mode or the lipid based (L-CF) mode (Figure 2.6). Amphiphilic molecules, like detergents are only able to mimic the hydrophobicity of lipid bilayers, but not its shape, lateral pressure or topology. Lipids of defined composition are commonly supplied in cell free protein synthesis to mimic native environment of lipid bilayer

interaction to acquire functional folding ^[87]. In L-CF mode, membrane proteins have the option to co-translationally insert into the lipid bilayers. In P-CF, however, no hydrophobicity or native environment is provided at all. This is not desirable for unstable targets like membrane proteins in real practice ^[88]. Thus, the following discussion will focus on D-CF and L-CF.

Detergents are small, amphiphilic molecules consisting of a hydrophilic head and hydrophobic tail. In aqueous solutions, they spontaneously form (generally) micelles ^[89]. Detergents properties of detergents are related to the structure and size of the polar head-group, charges, and length of hydrocarbon chain. Generally, they are classified into nonionic (no surface charges), zwitterionic (neutral but with both positive and negative charges), anionic, and cationic ^[35, 89, 90]. It is commonly believed that shorter detergents are better in solubilizing membrane proteins but meanwhile possess higher risk of denaturation. Nonionic detergents, such as n-Dodecyl β -D-maltoside (DDM) are considered milder than ionic detergents, which better solubilize membrane proteins. However, the optimal detergent, together with buffer systems, ionic strength, etc. for a new target have to be determined experimentally ^[90].

In D-CF mode, detergents that better solubilize target protein without inhibited translation are desired. Christian Klammt et al. expressed GPCR protein rat corticotropin releasing factor receptor (CRF) using D-CF using *E. coli* extract and screened nonionic detergents Brij 35, 58 and 78 ^[91]. Brij78 was found to be highly effective for solubilization and in less than 24 h, milligram amounts of soluble protein was generated. Cell-free expressed GPCR samples were also characterized by single particle analysis showed homogenous distribution of predominantly protein dimers. Liselotte Kaiser et al. also reported D-CF mode expression of a human olfactory receptor 17-4 (hOR17-4) using the wheat germ extract ^[92]. Optimal detergent digitonin at 0.4 % was determined to be most effective in solubilizing olfactory receptor hOR17-4 based on detergent screening. The production yield was reported to be about 1 mg protein per 3 mL of cell free reaction, sufficient for secondary structure and odorant binding studies. Purified receptor display signature α -helical structure and bind to odorant undecanal with affinity of 22 μ M.

Figure 2.6 Basic expression modes for the CFPS of MPs; the protein is either expressed without any hydrophobic agents (P-CF mode) or in the presence of detergents (D-CF mode) or lipids (L-CF mode). Below: variety of membrane proteins samples resulting from the basic expression modes ^[64]. (Reprinted with permission from Ref. ^[64])

In comparison to D-CF mode, lipids are supplied in L-CF. They are essential component of cell membranes in nature and commonly used in membrane protein research and CFPS ^[75]. They are amphiphilic molecules like detergents but do not solubilize proteins so efficiently due to their longer chain length. They also provide more stable and native support by mimicking cellular environment. In L-CF modes, lipids are supplied in the form of vesicular, bicelle and disc structures (Figure 2.6). A variety of membrane proteins were expressed in L-CF mode, including receptors, channels, transporters, pore-forming

proteins, enzymes, etc ^[64]. Here we focus on application of preformed vesicles (liposomes) in CFPS, where membrane proteins are simultaneously expressed and inserted into artificial membranes, leading to the formation of proteoliposomes ^[93]. Nobuaki Takemori et al. reported a high-throughput method for the synthesis of 263 full-length transmembrane proteins (TMPs) as proteoliposomes ^[94]. Of all the targets produced, 219 were multipath TMPs, including class A and class B GPCRs, transporters, channel proteins, etc. The synthesized proteoliposomes were used to mimic the *in-vivo* environment from which an internal standard library was constructed for targeted transmembrane proteomics by using mass spectrometry. The applicability of the system is yet to be validated. Similar strategy was employed by Hiroyuki Takeda et al. who developed monoclonal antibody (mAb) with GPCR proteoliposomes antigens using WG cell free system ^[95]. It is showed that half out of 36 mouse mAbs and 6 rabbit mAbs generated against dopamine receptor D1 (DRD1) were conformation-sensitive with two identified to recognize extracellular loop 2 of DRD1. Despite successful production of membrane proteins in L-CF, the purification remains a challenge. As a common practice, produced proteoliposomes are purified by density-based sucrose gradient ultracentrifugation ^[96]. The distribution of target protein in sucrose gradient varies among different targets and has to be experimentally determined. This method is most widely used in combination with *E. coli* CFPS systems for membrane protein production and purification, e.g. β 2AR (GPCR protein) ^[94], and gap junction channel protein Connexin-43 ^[95] but does not always yield high purity products. It is probably because the density-based purification is unable to fully separate functional incorporated protein in liposomes (proteoliposomes) from impurities with similar densities. The affinity-based purification of proteoliposomes was attempted by Norman T. Hovijitra et al. using *E. coli* cell free expressed water channel protein Aquaporin Z using biotin-functionalized synthetic lipid vesicles developed by Boris Peker et al. ^[97]. The purification yielded 507 ± 11 $\mu\text{g/mL}$ of active vesicle-associated AqpZ of 600 μL reaction. The method is not yet validated with other targets.

From a biomimetic perspective, polymersomes are similar but more stable membrane mimetics like liposomes. Previous work has shown that these self-assembled vesicles were compatible with CFPS and able to incorporate functional membrane proteins upon

translation in CFPS, e.g. Cldn2 ^[98], dopamine receptor D2 (DRD2) ^[99] and CXCR4 ^[100]. However, current studies of polymersomes in CFPS are similar to liposomes, where the extent to which most target characterization was very limited ^[64, 75]. First of all, targets were preliminarily characterized, e.g. expression condition/lipid screening, solubility studies, and barely showed purity or functionality. Such results may not reflect the actual or important properties of the preparations and can potentially lead to ambiguous conclusions. Secondly, well-established targets were often selected in CFPS, e.g. β 2AR, AqpZ, bacteria rhodopsin, and many of which were already well studied. Such targets were repeatedly characterized in CFPS with limited changes of synthesis conditions, e.g. expression systems, lipid compositions, and provide little insight into these membrane proteins. Finally, challenging and novel targets were rarely attempted using CFPS as an alternative before ample reports were released with other systems. This is best reflected by class B GPCR, GLP-1R of which no full-length nor truncated constructs were ever expressed in any cell free system with currently limited knowledge in many aspects. Thus, we decided to investigate a difficult class B GPCR target to produce in the cell free expression of GLP-1R using polymersomes as native-like membrane support.

References

- [1] A. Engel, H. E. Gaub, *Annual Review of Biochemistry* **2008**, 77, 127-148.
- [2] M. S. Almén, K. J. V. Nordström, R. Fredriksson, H. B. Schiöth, *BMC Biology* **2009**, 7, 50-50.
- [3] A. B. Harvey Lodish, S Lawrence Zipursky, Paul Matsudaira, David Baltimore, and James Darnell, *Molecular Cell Biology 4th edition*, W. H. Freeman, New York, **2000**.
- [4] A. Heifetz, G. F. X. Schertler, R. Seifert, C. G. Tate, P. M. Sexton, V. V. Gurevich, D. Fourmy, V. Cherezov, F. H. Marshall, R. I. Storer, I. Moraes, I. G. Tikhonova, C. S. Tautermann, P. Hunt, T. Ceska, S. Hodgson, M. J. Bodkin, S. Singh, R. J. Law, P. C. Biggin, *Naunyn-Schmiedeberg's Archives of Pharmacology* **2015**, 388, 883-903.
- [5] K. Palczewski, T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, I. L. Trong, D. C. Teller, T. Okada, R. E. Stenkamp, M. Yamamoto, M. Miyano, *Science* **2000**, 289, 739-745.

- [6] E. Ghosh, P. Kumari, D. Jaiman, A. K. Shukla, *Nat Rev Mol Cell Biol* **2015**, *16*, 69-81.
- [7] Y. He, K. Wang, N. Yan, *Protein & Cell* **2014**, *5*, 658-672.
- [8] V. Katritch, V. Cherezov, R. C. Stevens, *Annual review of pharmacology and toxicology* **2013**, *53*, 531-556.
- [9] M. Gimpelev, L. R. Forrest, D. Murray, B. Honig, *Biophysical Journal* **2004**, *87*, 4075-4086.
- [10] S. R. George, B. F. O'Dowd, S. P. Lee, *Nat Rev Drug Discov* **2002**, *1*, 808-820.
- [11] A. J. Venkatakrishnan, X. Deupi, G. Lebon, C. G. Tate, G. F. Schertler, M. M. Babu, *Nature* **2013**, *494*, 185-194; V. Cherezov, *Current opinion in structural biology* **2011**, *21*, 559-566.
- [12] B. Wu, E. Y. T. Chien, C. D. Mol, G. Fenalti, W. Liu, V. Katritch, R. Abagyan, A. Brooun, P. Wells, F. C. Bi, D. J. Hamel, P. Kuhn, T. M. Handel, V. Cherezov, R. C. Stevens, *Science* **2010**, *330*, 1066-1071.
- [13] F. Y. Siu, M. He, C. de Graaf, G. W. Han, D. Yang, Z. Zhang, C. Zhou, Q. Xu, D. Wacker, J. S. Joseph, W. Liu, J. Lau, V. Cherezov, V. Katritch, M.-W. Wang, R. C. Stevens, *Nature* **2013**, *499*, 444-449.
- [14] K. Hollenstein, J. Kean, A. Bortolato, R. K. Y. Cheng, A. S. Dore, A. Jazayeri, R. M. Cooke, M. Weir, F. H. Marshall, *Nature* **2013**, advance online publication.
- [15] J. J. Holst, *Physiological Reviews* **2007**, *87*, 1409-1439.
- [16] L. Prasad-Reddy, D. Isaacs, *Drugs in Context* **2015**, *4*, 212283.
- [17] D. J. Drucker, J. B. Buse, K. Taylor, D. M. Kendall, M. Trautmann, D. Zhuang, L. Porter, *The Lancet*, *372*, 1240-1250; J. B. Buse, J. Rosenstock, G. Sesti, W. E. Schmidt, E. Montanya, J. H. Brett, M. Zychma, L. Blonde, *The Lancet*, *374*, 39-47; A. B. Olokoba, O. A. Obateru, L. B. Olokoba, *Oman Medical Journal* **2012**, *27*, 269-273.
- [18] U. S. F. a. D. Administration, **2008**.
- [19] K. C. Herold, L. Taylor, *Immunologic Research* **2003**, *28*, 141-150.
- [20] S. Jähnichen, C. Blanchetot, D. Maussang, M. Gonzalez-Pajuelo, K. Y. Chow, L. Bosch, S. De Vrieze, B. Serruys, H. Ulrichs, W. Vandeveld, M. Saunders, H. J. De Haard, D. Schols, R. Leurs, P. Vanlandschoot, T. Verrips, M. J. Smit, *Proceedings of the National Academy of Sciences* **2010**, *107*, 20565-20570.

- [21] A. Lin, Z.-g. Liu, *Cell Res* **0000**, 18, 327-327.
- [22] N. Tuteja, *Plant Signaling & Behavior* **2009**, 4, 942-947.
- [23] W. I. Weis, B. K. Kobilka, *Current Opinion in Structural Biology* **2008**, 18, 734-740.
- [24] N. Dascal, *Trends in Endocrinology & Metabolism*, 12, 391-398.
- [25] T. Klassen, C. Davis, A. Goldman, D. Burgess, T. Chen, D. Wheeler, J. McPherson, T. Bourquin, L. Lewis, D. Villasana, M. Morgan, D. Muzny, R. Gibbs, J. Noebels, *Cell*, 145, 1036-1048.
- [26] S. Feliciangeli, F. C. Chatelain, D. Bichet, F. Lesage, *The Journal of Physiology* **2015**, 593, 2587-2603.
- [27] E. C. Cooper, L. Y. Jan, *Proceedings of the National Academy of Sciences of the United States of America* **1999**, 96, 4759-4766.
- [28] D. A. Doyle, J. M. Cabral, R. A. Pfuetzner, A. Kuo, J. M. Gulbis, S. L. Cohen, B. T. Chait, R. MacKinnon, *Science* **1998**, 280, 69-77.
- [29] Y. Y. Dong, A. C. W. Pike, A. Mackenzie, C. McClenaghan, P. Aryal, L. Dong, A. Quigley, M. Grieben, S. Goubin, S. Mukhopadhyay, G. F. Ruda, M. V. Clausen, L. Cao, P. E. Brennan, N. A. Burgess-Brown, M. S. P. Sansom, S. J. Tucker, E. P. Carpenter, *Science* **2015**, 347, 1256-1259.
- [30] E. Wrobel, D. Tapken, G. Seebohm, *Frontiers in Pharmacology* **2012**, 3.
- [31] F. Katzen, T. C. Peterson, W. Kudlicki, *Trends in biotechnology* **2009**, 27, 455-460.
- [32] V. Sarramegna, F. Talmont, P. Demange, A. Milon, *Cellular and Molecular Life Sciences CMLS* **2003**, 60, 1529-1546.
- [33] M. Schütz, J. Schöppe, E. Sedlák, M. Hillenbrand, G. Nagy-Davidescu, J. Ehrenmann, C. Klenk, P. Egloff, L. Kummer, A. Plückthun, *Scientific Reports* **2016**, 6, 21508.
- [34] A. Jazayeri, J. M. Dias, F. H. Marshall, *Journal of Biological Chemistry* **2015**.
- [35] R. M. Bill, P. J. F. Henderson, S. Iwata, E. R. S. Kunji, H. Michel, R. Neutze, S. Newstead, B. Poolman, C. G. Tate, H. Vogel, *Nat Biotech* **2011**, 29, 335-340.
- [36] S. Sahdev, S. K. Khattar, K. S. Saini, *Molecular and Cellular Biochemistry* **2008**, 307, 249-264.

- [37] G. Skretas, G. Georgiou, *Proceedings of the National Academy of Sciences* **2008**, *105*, 14747-14748.
- [38] M. Nishida, M. Cadene, B. T. Chait, R. MacKinnon, *The EMBO Journal* **2007**, *26*, 4005-4015.
- [39] P. Lu, D. Ma, C. Yan, X. Gong, M. Du, Y. Shi, *Proceedings of the National Academy of Sciences of the United States of America* **2014**, *111*, 1813-1818.
- [40] A. Contreras-Gómez, A. Sánchez-Mirón, F. García-Camacho, E. Molina-Grima, Y. Chisti, *Biotechnology Progress* **2014**, *30*, 1-18.
- [41] V.-P. Jaakola, M. T. Griffith, M. A. Hanson, V. Cherezov, E. Y. T. Chien, J. R. Lane, A. P. IJzerman, R. C. Stevens, *Science* **2008**, *322*, 1211-1217.
- [42] S. G. F. Rasmussen, B. T. DeVree, Y. Zou, A. C. Kruse, K. Y. Chung, T. S. Kobilka, F. S. Thian, P. S. Chae, E. Pardon, D. Calinski, J. M. Mathiesen, S. T. A. Shah, J. A. Lyons, M. Caffrey, S. H. Gellman, J. Steyaert, G. Skinotis, W. I. Weis, R. K. Sunahara, B. K. Kobilka, *Nature* **2011**, *477*, 549-555.
- [43] K. H. Khan, *Advanced Pharmaceutical Bulletin* **2013**, *3*, 257-263.
- [44] P. Thomas, T. G. Smart, *Journal of Pharmacological and Toxicological Methods* **2005**, *51*, 187-200.
- [45] S. L. Ritter, R. A. Hall, *Nat Rev Mol Cell Biol* **2009**, *10*, 819-830.
- [46] A. Jazayeri, J. M. Dias, F. H. Marshall, *Journal of Biological Chemistry* **2015**, *290*, 19489-19495.
- [47] D. J. Scott, A. Plückthun, *Journal of Molecular Biology* **2013**, *425*, 662-677; D. J. Scott, L. Kummer, P. Egloff, R. A. D. Bathgate, A. Plückthun, *Biochimica et biophysica acta* **2014**, *1838*, 2817-2824.
- [48] E. Chun, Aaron A. Thompson, W. Liu, Christopher B. Roth, Mark T. Griffith, V. Katritch, J. Kunken, F. Xu, V. Cherezov, Michael A. Hanson, Raymond C. Stevens, *Structure*, *20*, 967-976.
- [49] L. Qin, I. Kufareva, L. G. Holden, C. Wang, Y. Zheng, C. Zhao, G. Fenalti, H. Wu, G. W. Han, V. Cherezov, R. Abagyan, R. C. Stevens, T. M. Handel, *Science* **2015**, *347*, 1117-1122.
- [50] K. Hollenstein, J. Kean, A. Bortolato, R. K. Y. Cheng, A. S. Dore, A. Jazayeri, R. M. Cooke, M. Weir, F. H. Marshall, *Nature* **2013**, *499*, 438-443.

- [51] S. G. F. Rasmussen, H.-J. Choi, D. M. Rosenbaum, T. S. Kobilka, F. S. Thian, P. C. Edwards, M. Burghammer, V. R. P. Ratnala, R. Sanishvili, R. F. Fischetti, G. F. X. Schertler, W. I. Weis, B. K. Kobilka, *Nature* **2007**, *450*, 383-387.
- [52] E. Carlson, R. Gan, C. Hodgman, M. Jewett, *Biotechnology advances* **2012**, *30*, 1185-1194.
- [53] J. F. Zawada, in *Ribosome Display and Related Technologies: Methods and Protocols* (Eds.: A. J. Douthwaite, H. R. Jackson), Springer New York, New York, NY, **2012**, pp. 31-41.
- [54] R. J. Jackson, T. Hunt, in *Methods in Enzymology, Vol. Volume 96*, Academic Press, **1983**, pp. 50-74.
- [55] T. Sawasaki, Y. Hasegawa, M. Tsuchimochi, N. Kamura, T. Ogasawara, T. Kuroita, Y. Endo, *FEBS Letters* **2002**, *514*, 102-105.
- [56] R. Yadavalli, T. Sam-Yellowe, **2015**, e52772.
- [57] T. Ezure, T. Suzuki, M. Shikata, M. Ito, E. Ando, in *Cell-Free Protein Production: Methods and Protocols* (Eds.: Y. Endo, K. Takai, T. Ueda), Humana Press, Totowa, NJ, **2010**, pp. 31-42.
- [58] S. D. Pratt, C. A. David, C. Black-Schaefer, P. J. Dandliker, X. Xuei, U. Warrior, D. J. Burns, P. Zhong, Z. Cao, A. Y. C. Saiki, C. G. Lerner, L. E. Chovan, N. B. Soni, A. M. Nilius, F. L. Wagenaar, P. J. Merta, L. M. Traphagen, B. A. Beutel, *Journal of Biomolecular Screening* **2004**, *9*, 3-11.
- [59] N. Bøddeker, G. Bahador, C. Gibbs, E. Mabery, J. Wolf, L. Xu, J. Watson, *RNA* **2002**, *8*, 1120-1128.
- [60] Y. Kuruma, T. Ueda, *Nat. Protocols* **2015**, *10*, 1328-1344.
- [61] Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa, T. Ueda, *Nat Biotech* **2001**, *19*, 751-755.
- [62] K. Takai, T. Sawasaki, Y. Endo, *Nat. Protocols* **2010**, *5*, 227-238.
- [63] H.-C. Chang, C. M. Kaiser, F. U. Hartl, J. M. Barral, *Journal of Molecular Biology* **2005**, *353*, 397-409.
- [64] F. Junge, S. Haberstock, C. Roos, S. Stefer, D. Proverbio, V. Dötsch, F. Bernhard, *New Biotechnology* **2011**, *28*, 262-271.
- [65] F. Bernhard, Y. Tozawa, *Current Opinion in Structural Biology* **2013**, *23*, 374-380.

- [66] A. Zemella, L. Thoring, C. Hoffmeister, S. Kubick, *ChemBioChem* **2015**, *16*, 2420-2431.
- [67] T. H. Bayburt, S. G. Sligar, *FEBS Letters* **2010**, *584*, 1721-1727.
- [68] M. A. Schuler, I. G. Denisov, S. G. Sligar, in *Lipid-Protein Interactions: Methods and Protocols* (Ed.: H. J. Kleinschmidt), Humana Press, Totowa, NJ, **2013**, pp. 415-433.
- [69] J. A. Cappuccio, C. D. Blanchette, T. A. Sulchek, E. S. Arroyo, J. M. Kralj, A. K. Hinz, E. A. Kuhn, B. A. Chromy, B. W. Segelke, K. J. Rothschild, J. E. Fletcher, F. Katzen, T. C. Peterson, W. A. Kudlicki, G. Bench, P. D. Hoeprich, M. A. Coleman, *Molecular & Cellular Proteomics* **2008**, *7*, 2246-2253.
- [70] J. L. Popot, T. Althoff, D. Bagnard, J. L. Banères, P. Bazzacco, E. Billon-Denis, L. J. Catoire, P. Champeil, D. Charvolin, M. J. Cocco, G. Crémel, T. Dahmane, L. M. de la Maza, C. Ebel, F. Gabel, F. Giusti, Y. Gohon, E. Goormaghtigh, E. Guittet, J. H. Kleinschmidt, W. Kühlbrandt, C. Le Bon, K. L. Martinez, M. Picard, B. Pucci, J. N. Sachs, C. Tribet, C. van Heijenoort, F. Wien, F. Zito, M. Zoonens, *Annual Review of Biophysics* **2011**, *40*, 379-408.
- [71] T. J. Knowles, R. Finka, C. Smith, Y.-P. Lin, T. Dafforn, M. Overduin, *Journal of the American Chemical Society* **2009**, *131*, 7484-7485.
- [72] K. Corin, P. Baaske, S. Geissler, C. J. Wienken, S. Duhr, D. Braun, S. Zhang, *Scientific Reports* **2011**, *1*, 172.
- [73] J. M. Dörr, M. C. Koorengevel, M. Schäfer, A. V. Prokofyev, S. Scheidelaar, E. A. W. van der Crujisen, T. R. Dafforn, M. Baldus, J. A. Killian, *Proceedings of the National Academy of Sciences* **2014**, *111*, 18607-18612.
- [74] A. Akbarzadeh, R. Rezaei-Sadabady, S. Davaran, S. W. Joo, N. Zarghami, Y. Hanifehpour, M. Samiei, M. Kouhi, K. Nejati-Koshki, *Nanoscale Research Letters* **2013**, *8*, 102-102.
- [75] C. Roos, L. Kai, D. Proverbio, U. Ghoshdastider, S. Filipek, V. Dötsch, F. Bernhard, *Molecular Membrane Biology* **2013**, *30*, 75-89.
- [76] H. Bermudez, A. K. Brannan, D. A. Hammer, F. S. Bates, D. E. Discher, *Macromolecules* **2002**, *35*, 8203-8208.

- [77] B. M. Discher, Y.-Y. Won, D. S. Ege, J. C.-M. Lee, F. S. Bates, D. E. Discher, D. A. Hammer, *Science* **1999**, 284, 1143-1146; A. Jesorka, O. Orwar, *Annual review of analytical chemistry (Palo Alto, Calif.)* **2008**, 1, 801-832.
- [78] B. Harry, K. B. Aaron, A. H. Daniel, S. B. Frank, E. D. Dennis, *Macromolecules* **2002**, 35.
- [79] D. E. Discher, A. Eisenberg, *Science* **2002**, 297, 967-973.
- [80] F. Itel, M. Chami, A. Najer, S. Lörcher, D. Wu, I. A. Dinu, W. Meier, *Macromolecules* **2014**, 47, 7588-7596.
- [81] J. Habel, M. Hansen, S. Kynde, N. Larsen, S. Midtgaard, G. Jensen, J. Bomholt, A. Ogbonna, K. Almdal, A. Schulz, C. Hélix-Nielsen, *Membranes* **2015**, 5, 307.
- [82] F. Itel, A. Najer, C. G. Palivan, W. Meier, *Nano Lett* **2015**, 15.
- [83] T. Nakano, *Polym J* **0000**, 42, 103-123.
- [84] J. S. Lee, J. Feijen, *Journal of Controlled Release* **2012**, 161, 473-483.
- [85] J. Lee, M. Ankone, E. Pieters, R. Schiffelers, W. Hennink, J. Feijen, *Journal of controlled release : official journal of the Controlled Release Society* **2011**, 155, 282-288.
- [86] G. Ruan, S.-S. Feng, *Biomaterials* **2003**, 24, 5037-5044; M. Rafat, C. A. Clérout, W. G. Fong, A. N. Baker, B. C. Leonard, M. D. O'Connor, C. Tsilfidis, *Biomaterials* **2010**, 31, 3414-3421.
- [87] A. van Dalen, S. Hegger, J. A. Killian, B. de Kruijff, *FEBS Letters* **2002**, 525, 33-38.
- [88] I. Moraes, G. Evans, J. Sanchez-Weatherby, S. Newstead, P. D. S. Stewart, *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2014**, 1838, 78-87.
- [89] A. M. Seddon, P. Curnow, P. J. Booth, *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2004**, 1666, 105-117.
- [90] T. Arnold, D. Linke, in *Current Protocols in Protein Science*, John Wiley & Sons, Inc., **2001**.
- [91] C. Klammt, D. Schwarz, N. Eifler, A. Engel, J. Piehler, W. Haase, S. Hahn, V. Dötsch, F. Bernhard, *Journal of Structural Biology* **2007**, 158, 482-493.
- [92] L. Kaiser, J. Graveland-Bikker, D. Steuerwald, M. Vanberghem, K. Herlihy, S. Zhang, *Proceedings of the National Academy of Sciences* **2008**, 105, 15726-15731.

- [93] S. Habersstock, C. Roos, Y. Hoevels, V. Dötsch, G. Schnapp, A. Pautsch, F. Bernhard, *Protein expression and purification* **2012**, 82, 308-316; N. Hovijitra, J. Wu, B. Peaker, J. Swartz, *Biotechnology and bioengineering* **2009**, 104, 40-49; L. Lavinia, B. Iulia, M. Dominique, V. Michel, L. Jean-Luc, *Protein Expression and Purification* **2010**, 69.
- [94] N. Takemori, A. Takemori, K. Matsuoka, R. Morishita, N. Matsushita, M. Aoshima, H. Takeda, T. Sawasaki, Y. Endo, S. Higashiyama, *Molecular BioSystems* **2015**, 11, 361-365.
- [95] H. Takeda, T. Ogasawara, T. Ozawa, A. Muraguchi, P.-J. Jih, R. Morishita, M. Uchigashima, M. Watanabe, T. Fujimoto, T. Iwasaki, Y. Endo, T. Sawasaki, *Scientific Reports* **2015**, 5, 11333.
- [96] L. Liguori, J. L. Lenormand, in *Methods in Enzymology, Vol. Volume 465* (Ed.: D. Nejat), Academic Press, **2009**, pp. 209-223.
- [97] N. T. Hovijitra, J. J. Wu, B. Peaker, J. R. Swartz, *Biotechnology and Bioengineering* **2009**, 104, 40-49; B. Peker, J. J. Wu, J. R. Swartz, *Biotechnology Progress* **2004**, 20, 262-268.
- [98] M. Nallani, M. Andreasson-Ochsner, C.-W. Tan, E.-K. Sinner, Y. Wisantoso, S. Geifman-Shochat, W. Hunziker, *Biointerphases* **2011**, 6, 153-157.
- [99] S. May, M. Andreasson-Ochsner, Z. Fu, Y. X. Low, D. Tan, H.-P. M. de Hoog, S. Ritz, M. Nallani, E.-K. Sinner, *Angewandte Chemie International Edition* **2013**, 52, 749-753.
- [100] H. P. de Hoog, E. M. Lin JieRong, S. Banerjee, F. M. Decailot, M. Nallani, *PLoS One* **2014**, 9, e110847.

Chapter 3

Experimental Methodology

In order to produce Membrane Proteins (MPs) of interest in cell free protein protein system (CFPS), cell free expressions using different cell extracts and expression formats were performed to determine the optimal expression conditions. Following the expression, different approaches were used to produce high purity MPs, which were then subjected to biochemical characterization. In parallel, they were characterized using electron microscopy, size-exclusion liquid chromatography, surface plasmon resonance (SPR) based binding analysis and radioligand assay, which provided important information about protein/polymer association state, structure and functionality of different preparations. This chapter describes the materials and experimental methods used in the thesis, including the characterizations and assays above. Principles underlying the production/characterization techniques employed are also presented.

3.1 Rationale for selection

The selected methods include molecular cloning, protein expression and purification, preparation of polymersomes, and basic protein characterization (Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) and Western blot). The molecular cloning and protein expression define the start of protein production for any new target, which include cloning of target gene into appropriate expression vectors, followed by protein translation from the supplied target plasmid gene. Protein purification removes the impurities in expression mixture and extracts target protein of interest, thus producing high purity production. In protein expression and purification, SDS-PAGE and Western blot are used to detect the presence of target protein and to determine the purity of proteins obtained with different purification methods.

The film rehydration method is a commonly used technique to prepare vesicles, including liposomes and polymersomes. The prepared vesicles are extruded through filters with defined size and result in monodisperse size distribution. The characterization methods, including fluorescence detection size-exclusion chromatography (FSEC), transmission electron microscopy (TEM), surface plasmon resonance (SPR) and radioligand binding assay provide an overall analysis regarding protein/polymer association, functionality of target proteins and structural integrity of proteopolymersomes. The FSEC uses traditional size exclusion chromatography set-up with fluorescence detection that provides higher sensitivity and allows specific detection of fluorescently labelled molecules, e.g. proteins and polymers. TEM provides information about the structural details of vesicles. SPR detects specific interactions between molecules with high sensitivity. A number of molecules can be analyzed using SPR, including antibodies, peptides, small molecules etc. The radioligand binding assay has long been considered the “gold standard” assay for detection of functional receptors.

3.2 Synthesis

3.2.1 Materials

Plasmid pCMVTNT containing gene coding for GLP-1R was constructed by total gene synthesis and antibody recognizing peptide, the rhodopsin epitope tag (the C9 tag, N-Thr-Glu-Thr-Ser-Gln-Val-Ala-Pro-Ala-C), was added to the C-terminal of GLP-1R. Plasmid pEU containing gene coding for CXCR4 was designed and constructed by total gene synthesis. The same C9 tag was added to CXCR4 as GLP-1R in the gene design. Reagents and consumables for surface plasmon resonance measurements, including CM 5 sensor chips, 11-mercapto-undecanoic acid, 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide, N-hydroxysuccinimide, and ethanolamine were from GE Healthcare (UK). Polybutadiene-polyethylene oxide (PBD₂₂-PEO₁₃, 1200-b-600), and poly(2-methyloxazoline-b-dimethylsiloxane-b-2-methyloxazoline (PMOXA₂₀-PDMS₅₄-PMOXA₂₀, 1700-b-4000-1700) polymer was from Polymer Source Inc. (Canada) DSPE-PEG (2000) biotin was from Avanti Polar Lipids, Inc (US). Rat anti-Polyethylene glycol antibody [26A04] was from Abcam (Hong Kong, ab94764, 1.4-1.5 mg/mL) and mouse rhodopsin antibody (1D4, sc-57432, 200 µg/mL) was from Santa Cruz Biotechnology (USA). Rat GPL1R polyclonal antibody TA590112 was from Origene (USA). Softlink™ Avidin Resin was from Promega (Singapore). Ni-NTA resin was from Qiagen (Singapore). Wheat germ cell-free expression system WEPRO7240 was from CellFree Sciences Co., Ltd. (Japan). N-Dodecyl β-D-maltoside (DDM) was from Anatrace (US). I-125 labeled GLP-1 (7-36) peptide was purchased from Perkin-Elmer (Singapore) and Exendin4 from Karebay Biochem (CA, USA). SYPRO Gel Stain and Western Breeze Chemiluminescent Kit, anti-mouse were from Life Technologies (Singapore). The BODIPY® FL L-Cystine dye was from ThermoFisher Scientific (Singapore). All other reagents were from Sigma Aldrich (Singapore).

3.2.2 Molecular Cloning - Principles and Analysis

Principles

Polymer chain reaction (PCR) is a molecular biology technique to amplify from a single or small number of DNA across over several magnitude and generate sufficient DNA for downstream experiments, such as digestion and ligation ^[1]. The PCR contains template DNA used for amplification, primers that are short single chain nucleotide partially or fully overlap with template DNA, polymerase, nucleoside triphosphates, and buffer. The identification of PCR is usually performed with agarose gel electrophoresis and a DNA molecular weight marker with standard DNA of different molecular weights.

Analysis

The human glucagon-like peptide 1 receptor GLP-1R gene (GenBank: Homo sapiens glucagon-like peptide 1 receptor, mRNA; NCBI Reference Sequence: NM_002062.3) with C-terminus C9 tag was sub cloned from pCMVTNT vector into the pEU vector. The pCMVTNT vector was designed for expression of cloned genes using both in vivo and in vitro expression systems with both SP6 and T7 polymerase promoters encoded in the vector. The pEU vector was specifically developed for the wheat germ cell-free protein synthesis system with only SP6 promoter. The gene coding GLP-1R was digested with XhoI and NotI followed by a ligation reaction. Engineered green fluorescence protein (eGFP) gene sequence was PCR-amplified using primers 5'-GCAGC GATAT CATGC ATCAT CATCA TCATC ATCAT CATCA TCATA TGGTG AGCAA GGGCG AGGAG C-3' and 5'-GCTGC CTCGA GCTTG TACAG CTCGT CCATG CCG-3'. The resulting PCR product was cloned to pEU vector with GLP-1R using EcoRV and XhoI restriction enzyme sites. All constructs, including GLP-1R and eGFP-GLP-1R, were verified by sequencing before expression.

Plasmid pCMVTNT with T7 promoter was used for initial cell free expression screening among *E. coli*, insect, and wheat germ-based reaction systems. GLP-1R gene was sub-cloned to pEU vector by XhoI and NotI double digestion and ligation using T4 ligase. All cloning steps were confirmed by DNA gel electrophoresis and gene sequencing.

Subsequently, all vectors were transformed to BL21 *E. coli* for propagation and plasmid was prepped at midi scale.

3.2.3 Protein Expression and Purification - Principles and Analysis

Principles

The cell free expression refers to the method of protein synthesis independent of live cells [2]. In cells, proteins are synthesized and regulated upon the functional needs. In cell free protein expression, protein synthesis machinery was incorporated together with essential energy source, amino acids, and buffer to mimic protein translation and modification processes in cells. For coupled systems (transcription and translation in one single reaction), regular wheat germ (WG), high yield wheat germ (HYWG), *E. coli* S30 (S30) and rabbit reticulate lysate (RRL) were used. The two coupled systems, regular wheat germ (WG) uses T7 promoter and high yield wheat germ (HYWG) uses SP6 promoter. It is generally believed that with different 5' transcript site sequence, SP6 promoter has higher transcription efficiency and thus potentially increases yield in protein synthesis. For uncoupled system (cell extraction/mRNA template mixture separated from energy source and amino acids in different layers), a wheat germ system was used. Such format allows the continuous supply of substrates, together with the continuous removal of small byproducts. It was shown that the wheat germ bilayer reaction display higher yield over batch-mode [3]. The coupled and bilayer reaction systems are also different in the vector used for expression where coupled reaction uses pCVMTNT and in bilayer reactions, pEU vector was used.

Affinity purification is a method of separating biomolecules based on highly specific interactions [4]. In protein molecules, the common used strategies include poly-histidine-tag/Ni-NTA (His-tag/Ni-NTA) and biotin/avidin interactions. His-tag is a linear peptide epitope (usually 6 or 10 histidines) fused with protein target of interest and binds specifically to Nickel (Ni^{+}) ions. The Ni^{+} ions are further chelated to nitrilotriacetic acid (NTA), which is linked to solid matrix (usually agarose or polymer micron-sized beads). The target protein bound solid matrix is then physically separated with

centrifugation/column from initial mixture (cell lysate). The bound protein of interest can be eluted with competitive molecules such as imidazole from matrix. Biotin-affinity purification is of similar principles but takes advantage of specific biotin/avidin interaction. The elution uses excess free biotin instead of imidazole.

Analysis

For regular wheat germ extract (WG), high yield wheat germ (HYWG), *E. coli* S30 (S30) and rabbit reticulum lysate (RRL) -based coupled cell free reactions, the following reagents were mixed:

Table 3.1 Recipe of protein expression screening in different cell free expression systems

Component	WG*	HYWG*	RRL*	S30*
TnT Buffer (μL)	2	-	2	-
RNase inh (μL)	1	-	1	-
Lysate (μL)	25	-	25	18
A.acid –Met (μL)	0.5	-	0.5	-
A.acid – Leu (μL)	0.5	-	0.5	-
DNA template	1ug	4ug	1ug	1ug
Polymersomes (5mg/mL) (μL)	15	15	15	15
RNA polymerase (added last) (μL)	1	-	1	-
HYWG Master Mix (μL)		30		
S30 Premix (μL)				20
Nucl-free H ₂ O (μL)				
TOTAL VOL	50	50	50	50
Temp (deg C)	30, shake	25,shake	30, shake	37, shake
Time (hrs)	1.5	2	1.5	1

*WG: wheat germ extract; HYWG: high yield wheat germ, RRL: rabbit reticulum lysate; S30: *E. coli* cell extract

After mixing all the components, the reaction mixtures were incubated at respective temperature from 25 to 37 °C for 1 to 2 h, followed by harvesting the reaction for further characterization.

For expression of GLP-1R with uncoupled reaction, a pEU SP6 in wheat germ extract system in a two-step manner, i.e., RNA transcription was performed and then proceeded to protein translation. For RNA transcription, 6.5 μL of Transcription buffer, NTP mix, SP6

polymerase and RNase inhibitor with 2 ug of plasmid DNA were mixed to a final volume of 20 μ L and incubated at 37 °C for 6 h. As a next step, translation was performed by mixing 10 μ L of transcribed mRNA mixture with 10 μ L of 10 mg/mL AB polymersomes with 10 μ L wheat germ extract and carefully adding to the bottom of 220 μ L amino acid/energy mix in a 96 well plate to form bilayer (Figure 3.1). The plate was incubated at 15 °C for 20 h on a thermomixer. Translation reaction was stopped after 20 h by mixing the contents of the well with a pipette.

Bilayer reaction system

Figure 3.1 Formation of uncoupled wheat germ reaction system. Bilayer is formed by first pipetting amino acids and energy source mix into 96 well-plate followed by slow addition of wheat germ extract mixed with mRNA of interest and additives without disturbing as-formed bilayer ^[5]. (Copyright: Cell Free Sciences Co.,Ltd.)

Upon harvest of reaction, the reaction mixture was first centrifuged at 10,000 rpm for 1 h, 4 °C to spin down insoluble components and large aggregates. Soluble fraction from supernatant was collected for further purification. For size-based filtration, 200 μ L supernatant was loaded to one Durapore PVDF 1000 k MWCO centrifugal filter followed by spinning at 6,000 rpm for 30 min at 4 °C, and washed 3 times with 400 μ L 1x PBS once

the volume decreases to less than 50 μL . After the final wash, the retentate was resuspended to 200 μL .

Affinity purification of GLP-1R was performed using Ni-NTA beads. The supernatant after the centrifugation step was incubated with Ni-NTA agarose beads (100 μL slurry per 200 μL reaction mixture supernatant), followed by washing with 30 bed volumes of buffer containing 50 mM Tris-HCl (pH 6.6), 300 mM NaCl, 0.05 % DDM, and 60 mM imidazole. Elution was performed with buffer containing 300 mM imidazole. Eluted proteins were collected, and the imidazole removed by dialysis against 1x PBS in a 10 kDa Xpress Micro Dialyzer (scienova, Germany) overnight. Biotin-affinity purification was performed using SoftLink Avidin beads. The supernatant after centrifugation (and dialysis, in case of Foscholine-produced protein) was incubated with pre-equilibrated beads (100 μL slurry per 400 μL reaction mixture supernatant) overnight, followed by washing with 30 bed volumes of 1x PBS. Elution was performed using 10 mM biotin in 1x PBS. All affinity-purification steps were performed at 4 °C.

3.2.4 Polymersomes Preparation - Principles and Analysis

Principles

Film rehydration is a method of producing nano- to micron- scale vesicular structures, e.g. polymersomes and liposomes. The dried, thin lipid/polymer films form hydrated sheets during agitation and self-close up to form large vesicles through self-assembly ^[6]. The size of formed vesicles can be measured with dynamic light scattering (DLS). DLS measures the size distribution of small particles in solution. When light hits particles, the scattered light goes to all directions with intensity fluctuation. Such fluctuation can be auto-correlated and thus deduce the average size of measured particles and distribution ^[7].

Analysis

Formation of polymersomes was performed using film rehydration method. Diblock copolymer polybutadiene (PBD)-polyethylene oxide (PEO) (AB, 10 mg) or triblock copolymer poly(2-methyloxazoline)–poly(dimethylsiloxane)–poly(2-methyloxazoline)

(PMOXA–PDMS–PMOXA) (ABA, 10 mg) was dissolved in chloroform, vacuum dried under nitrogen flow, rehydrated in 1 mL of 1x PBS (phosphate buffered saline, 10 mM phosphate, 137 mM NaCl, and 2.7 mM KCl) and extruded through 0.22 μ m syringe filters and dialyzed (MWCO 50 kDa, Spectra/Por® 7, Spectrum Laboratories) against 1x PBS for at least 24 h. For biotin grafted AB polymersomes 1 wt% of DSPE-PEG (2000) biotin was added during rehydration. For fluorescent AB, 1 mol% tetrazole was added followed by UV (254 nm) cross-linking for 10 to 15 minutes after vesicle formation. Fluorescence measurement was performed using Tecan infinite 200 Pro plate reader with settings as follow: excitation/emission, 370 nm/470 nm (tetrazole), 450 nm/510 nm (eGFP).

3.3 Characterization

3.3.1 Protein Gel Electrophoresis and Western Blot - Principles and Analysis

Principles

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) is a method used in protein separation and purity analysis ^[8]. The polyacrylamide gel is used as a support medium and SDS denatures and negatively charge proteins. Under the electric field, proteins migrate towards the positive electrode (the anode) with different rate depending on the size/molecular weight of proteins. Proteins are visualized and imaged after staining, either colorimetric (coomassie blue) or fluorescent (SYPRO) dyes.

Analysis

To detect GLP-1R expression, 5 μ L of cell free reaction mixture was used for gel electrophoresis (NuPAGE 12% Bis-Tris Gel, MES running buffer, Invitrogen, 200V, 30 min) and Western Blot (iBlot2 system with PVDF membrane and Western Breeze® Chemiluminescence Kit, anti-mouse or anti-rabbit, Invitrogen) according to the supplier's manual. The blot was incubated overnight in a 1:2000 diluted mouse anti-1D4 antibody. For SDS-PAGE gel staining, SYPRO Gel Stain was used. After gel electrophoresis, the gel was fixed with fixing solution (40 % methanol, 7 % acetic acid) for 30 min followed by

overnight staining. Images were taken using ImageQuant LAS 500 (GE, Singapore) with either chemiluminescent (Western blot) or fluorescence (SYPRO) detection mode.

3.3.2 Bicinchoninic Acid Assay - Principles and Analysis

Principles

Bicinchoninic Acid (BCA) Protein Assay is based on the colorimetric detection and quantitation of total proteins using BCA [9]. This method measures the reduction of Cu^{+2} to Cu^{+1} by protein by detecting the purple-colored reaction product from Cu^{+1} and BCA. The water-soluble complex has a strong absorbance at 562 nm that is nearly linear with increasing protein concentrations at 20-2000 $\mu\text{g/mL}$. The number of peptide bonds and the presence of four particular amino acids (cystine, cysteine, tryptophan and tyrosine) are responsible for color formation with BCA.

Analysis

Protein concentrations are determined with reference to standards of bovine serum albumin (BSA). The 5 μL samples were mixed with 70 μL working solution containing BCA and incubated for 30 min at 37 °C in a thermomixer and subsequently transferred to a 96-well plate for the absorbance reading at 562 nm with a TECAN plate reader (TECAN i-Control infinite 200).

3.3.3 Fluorescence Detection Size-exclusion Chromatography - Principles and Analysis

Principles

Size exclusion chromatography (SEC) is a method to separate molecules by their different sizes, which is usually correlated to molecular weight [10]. Typically, a solution mobile phase is used to carry analytes through solid matrix in the column. The solid matrix is made of uniform, porous micro-size particles that separate analytes where large particles are first excluded first followed by small ones. The fluorescence detection size-exclusion

chromatography is modified from SEC by adding fluorescence detection upon separation [11].

Analysis

Fluorescence-detection size-exclusion chromatography characterization (FSEC) was performed using Agilent 1260 Infinity Quaternary Liquid Chromatography System using eGFP-GLP-1R construct. The soluble fraction from reaction mixture and purified receptors were loaded onto TSKgel G6000PW column (Tosoh, Japan) using 1x PBS as mobile phase at flow rate of 0.4 mL min⁻¹ at room temperature. Other than UV measurement, the eluent was analyzed with fluorescence detector set as follows: excitation/emission, 370 nm/ 470 nm (tetrazole), 450 nm/ 510 nm (eGFP), 480nm/ 520 nm (calcein).

3.3.4 Biosensor analysis - Principles and Analysis

Principles

Surface plasmon resonance (SPR) is the conduction electrons resonant oscillation at the interface. The interface is usually between a negative and positive permittivity material that can be stimulated by incident light [12]. The biosensor analysis based on SPR allows measurement of molecular interaction at high sensitivity based on mass transport on the surface of a biosensor chip.

Analysis

Surface plasmon resonance measurements (SPR) were performed on a Biacore 3000-System (GE healthcare). Rat anti-PEG antibody was immobilized on Carboxymethylated dextran (CM 5) chip with using standard amine-coupling chemistry as described before [13]. The soluble fraction was collected from centrifugation and captured by flowing at 2 μ L min⁻¹ for 30 min. Binding experiments were performed in HEPES-buffered saline (10 mM HEPES, 150 mM NaCl, pH 6.4). GLP-1R in biotinylated vesicles was measured in the same way. GLP-1R N-terminal antibody TA590112 was diluted with running buffer and injected for 1 min with flow rate of 30 μ L/min. Binding curves were fitted to a 1:1 model using T200 BIAevaluation (Biacore AB, Uppsala, Sweden).

3.3.5 Transmission Electron Microscopy - Principles and Analysis

Principles

The transmission electron microscopy (TEM) operates on the same basics as light microscopy only that electrons were used instead of light ^[14]. When an electron beam transmits through an ultra-thin specimen, e.g. on a copper grid, the contrast between transmitted and blocked beams show the size and morphology of the specimen. The use of electrons provides significantly higher resolution compared to light microscopy.

Analysis

All electron microscopy experiments were performed at the Facility for Analysis, Characterization, Testing and Simulation (FACTS) in Nanyang Technological University, Singapore. Transmission electron microscopy (TEM) was performed using Carl Zeiss Libra 120 Plus operated at 120 kV. Samples were prepared using plasma-treated copper grids followed by negative staining with 4 % osmium tetroxide and were dried in air.

3.3.6 Radioligand Binding Assay - Principles and Analysis

Principles

Radioligand binding assay uses radioisotope labeled ligand to measure binding of ligand to specific molecules, e.g. receptors ^[15-16]. Usually, radioligand binding assay is performed in form of saturation binding, where increasing amount of radioligand is mixed with receptors, followed by separation of unbound ligand. The radioactivity of bound ligand is then measured and converted to ligand concentration. At certain concentration, the bound ligand saturates all the active receptors, known as maximum binding capacity (B_{max}). Addition of excess radioligand only increases non-specific binding and specific binding remains constant. The use of radioisotopes provides highest sensitivity over any other ligand binding assays and is regarded as the gold standard in receptor functional assay.

Analysis

Saturation binding experiments were performed by incubating 20 μ L of GLP-1R proteopolymersomes after elution from the affinity beads with 5 pM-4 nM of [I-125] GLP-1 (7-36) peptide in binding buffer (50 mM HEPES pH 6.4, 100 mM NaCl, 10 mM KCl, 1 mM EDTA, 5 mM MgCl₂, 0.1% (w/v) BSA) for 1 h at room temperature with gentle shaking. Non-specific binding (NSB) counts were determined by adding excess of non-labeled Exendin-4 (2 μ M) to the corresponding wells. The reaction was terminated by separating bound from unbound ligand by applying the mixture to an Illustra Microspin G50 column and centrifuged to collect the flow through. The flow through was transferred to the wells of a Wallac Isoplate 96-well sample plate. MicroScint20 liquid scintillation cocktail was added to the wells and the plate was read using a MicroBeta Trilux Liquid Scintillation Counter (Perkin-Elmer). Specific binding was calculated by subtracting NSB counts from total binding (TB) and fitting the resultant data points with a one site-specific binding formula using GraphPad Prism.

3.4 Overview of Methodologies

The methods used in expression and purification were standard or modified from standard molecular biology and polymer methodologies. The methods were used to produce GLP-1R and TREK-2 membrane materials. Biochemical characterization of expressed and purified proteins provides information about protein expression level, molecular weight and purity. The combination of FSEC, TEM and SPR provides structural characterization of produced materials, including protein/polymer association, structure of polymersomes and topology of membrane proteins incorporated. The radioligand binding assay provides critical information regarding functionality of different preparations using cell free system and amount of functional proteins.

References

- [1] F. T. Ishmael, C. Stellato, *Annals of Allergy, Asthma & Immunology*, 101, 437-443.
- [2] F. Bernhard, Y. Tozawa, *Current Opinion in Structural Biology* **2013**, 23, 374-380.

- [3] K. Takai, T. Sawasaki, Y. Endo, *Nat. Protocols* **2010**, 5, 227-238.
- [4] W. H. Dunham, M. Mullin, A.-C. Gingras, *PROTEOMICS* **2012**, 12, 1576-1590.
- [5] <http://www.cfsciences.com/eg/>
- [6] A. Akbarzadeh, R. Rezaei-Sadabady, S. Davaran, S. W. Joo, N. Zarghami, Y. Hanifehpour, M. Samiei, M. Kouhi, K. Nejati-Koshki, *Nanoscale Research Letters* **2013**, 8, 102-102.
- [7] N. Nemoto, Y. Makita, Y. Tsunashima, M. Kurata, *Macromolecules* **1984**, 17, 2629-2633.
- [8] Maizel, J. V. Jr, *Trends in Biochemical Sciences*, 25, 590-592.
- [9] J. M. Walker, in *Basic Protein and Peptide Protocols* (Ed.: J. M. Walker), Humana Press, Totowa, NJ, **1994**, pp. 5-8.
- [10] T. Kawate, E. Gouaux, *Structure* **2006**, 14, 673-681.
- [11] P. Hong, S. Koza, E. S. P. Bouvier, *Journal of Liquid Chromatography & Related Technologies* **2012**, 35, 2923-2950.
- [12] D. G. Drescher, N. A. Ramakrishnan, M. J. Drescher, *Methods in molecular biology (Clifton, N.J.)* **2009**, 493, 323-343.
- [13] R. J. Mann, N. E. Nasr, J. K. Sinfield, E. Paci, D. Donnelly, *British Journal of Pharmacology* **2010**, 160, 1973-1984.
- [14] B. Ruozi, D. Belletti, A. Tombesi, G. Tosi, L. Bondioli, F. Forni, M. A. Vandelli, *International Journal of Nanomedicine* **2011**, 6, 557-563.
- [15] D. B. Bylund, M. L. Toews, *American Journal of Physiology - Lung Cellular and Molecular Physiology* **1993**, 265, L421-L429.
- [16] J. J. Maguire, R. E. Kuc, A. P. Davenport, in *Receptor Binding Techniques* (Ed.: P. A. Davenport), Humana Press, Totowa, NJ, **2012**, pp. 31-77.

Chapter 4

Protein Synthesis and Primary Characterizations

This chapter summarizes the first results of this thesis, including molecular cloning, test expression of target protein, purification and primary characterizations. To start with, the plasmid DNA encoding membrane protein (MP) of interest, glucagon like peptide-1 receptor (GLP-1R) was test expressed in different cell free systems, using wheat germ, rabbit reticulate and E. coli extract. From both mammalian and non-mammalian expression systems, only wheat germ cell lysates produced protein whereas the other system did not show any expression. Thus, this system was selected for further production and characterization. To ensure that the proteins are incorporated into polymersome based artificial cell membranes, a filtration-based purification method was used and that showed almost all expressed proteins were incorporated. Further analyzing these samples by protein gel electrophoresis and Western blot, it showed that simple filtration did not purify the protein incorporated polymersomes from the cell-free reaction mixture. Hence, different affinity-purification methods were further developed and resulted in high purity GLP-1R. The selected purification methods were also evaluated in terms of co-purification of supplied AB polymersomes. Primary characterization showed only one of the methods can retain both protein and polymer.

4.1 Introduction

This chapter presents the results of expression of GLP-1R in cell free system with focus on protein expression level, incorporation into polymersomes, purification and primary characterization. To start with, the plasmid DNA with gene coding target protein GLP-1R was cloned with a widely used antibody recognizing epitope (C9 tag) made of 9 amino acids at the C-terminus of bovine rhodopsin. After cloning, the plasmid DNA was purified for cell-free reaction for protein translation. To detect the expression of GLP1R, Western blot was performed using anti-C9 antibody and that confirmed the successful expression of GLP-1R and SDS-PAGE determined the purity of the expressed proteins. These two methods also provide a qualitative estimate of expression level of GLP-1R.

Furthermore, to assess the incorporation of GLP-1R into polymersome membrane, a size-based filtration assay was used to remove the unincorporated proteins and cell-free reaction mixtures by their difference of size. To obtain a purified GLP-1R incorporated polymersomes, three methods, including filtration, biotin- and histidine-affinity purification methods were used and evaluated through SDS-PAGE and Western blot. In parallel, different formats of cell free reaction were used, including coupled reaction format and bilayer two-step uncoupled reaction for increasing the yields of receptor produced. The later was reported to have increased yield although the two formats are using the same wheat germ based extract. To evaluate whether the purification process disassemble GLP-1R integration into polymersomes, fluorescently labeled vesicles were used to determine the protein association into polymersomes.

4.2 Experimental Methodology

4.2.1 Molecular Cloning

Molecular cloning was performed to generate target DNA in the expression pEU vector for bilayer cell free reaction. All cloning steps were monitored by DNA gel electrophoresis. For details, refer to chapter 3.

4.2.2 Protein Expression and Purification

Protein expression was performed using both coupled (transcription and translation in one single reaction) and bilayer reaction (separated extraction/mRNA template mixture and energy source and amino acids) formats. Purification was performed using filtration, biotin- and histidine- affinity methods. For details, refer to chapter 3.

4.2.3 Polymersomes Preparation

Diblock copolymer polybutadiene-polyethylene oxide (AB) was used for polymersomes preparation using film rehydration method. For fluorescence and biotin-tagged vesicles, tetrazole or DSPE-PEG-biotin was added during film formation. Tetrazole fluorescence was measured using Tecan 200 plate reader. For details, refer to chapter 3.

4.2.4 Protein Gel Electrophoresis and Western Blot

Protein gel electrophoresis and Western blot were performed for detection of GLP-1R protein and purity examination. For details, refer to chapter 3.

4.2.5 Bicinchoninic Acid Assay (BCA)

BCA was performed to measure protein concentration of purified proteins. For details, refer to chapter 3.

4.3 Principle Outcomes

4.3.1 GLP-1R Is Expressed and Reconstituted in Polymersomes Using High Yield Wheat Germ System

Diblock and triblock copolymers, as synthetic materials, are popular biomimetic membranes have been extensively used in *in vitro* reconstitution of membrane proteins for structural/function studies (as introduced in chapter 1). The vesicular structures of block copolymer biomimetic systems, which are made from copolymer self-assembly, display cell-like morphology and membrane bilayers. In the beginning of this study, the AB polymersomes were chosen in the first place because it is the most systematically studied polymer regarding vesicle formation and related properties (e.g. MW/membrane thickness correlation, self-assembled morphologies, etc.). The triblock copolymers will be investigated in later chapter (section 7.3). AB polymersomes were prepared using film rehydration method and added to cell free protein synthesis systems to test the expression of GLP-1R. GLP-1R gene sequence was designed, synthesized, and cloned into pCMVTNT vector, propagated and used in test expression in different cell free protein synthesis. Four systems, including regular wheat germ (WG), high yield wheat germ (HYWG), S30 *E. coli* lysate (S30), and rabbit reticulocyte lysate (RRL) systems were used for test expression of GLP-1R. The *E. coli*, WG, and HYWG represent non-mammalian and the RRL represents mammalian cell free protein synthesis system. In WG system, the T7 promoter was used and in HYWG system, the stronger SP6 promoter was used that could potentially lead to higher yield. Both T7 and SP6 promoters were encoded in the pCMVTNT vector that was being used at test expression stage. Of all the systems tested, GLP-1R was expressed only in HYWG system as detected by Western blot (Figure 4.1). Non-specific binding of mouse anti-C9 monoclonal antibody to proteins in the cell lysate was also observed in WG, S30, and RRL but only of background level. For HYWG system, expressed GLP-1R was detectable as a major band about 40 kDa. The apparent molecular weight (MW) 40 kDa was lower than the theoretical 53 kDa based on amino acid sequence. This reflects the anomalous SDS-PAGE migration of helical membrane proteins upon detergent binding, e.g. SDS, as reported previously ^[1]. A minor band was also detected,

indicating while the majority of GLP-1R was successfully reconstituted, a small fraction of it was not and formed dimerize/aggregate considering the 80 kDa band was not removed even treated with SDS before gel electrophoresis. Since the C9 tag was fused to the C-terminus of GLP-1R, the positive Western blot result indicates the complete translation of full length GLP-1R protein.

In summary, HYWG system has been identified as the optimal cell free system for expressing GLP-1R. Importantly, this is the first time a full-length GLP-1R is expressed in the presence of polymersomes using cell free protein synthesis.

Figure 4.1 Western blot analysis of GLP-1R expression in regular wheat germ (WG, lane 1), high yield wheat germ (HYWG, lane 2), S30 *E. coli* lysate (S30, lane 3), and rabbit reticulocyte lysate (RRL, lane 4) cell free expression systems with using mouse anti-C9 monoclonal (1D4) antibody. Expression was also performed in each system without adding cDNA to test the non-specific binding (background) of 1D4 antibody. Arrow indicates GLP-1R. M: molecular weight marker.

MPs are folded correctly only when the transmembrane domains are embedded in a hydrophobic environment. To determine if the translated GLP-1R is incorporated into the polymersomes, polymersomes were first separated from the free proteins. Polymersomes have higher MW (estimated to be a few 1,000 kDa) as compared to free proteins (less than 200 kDa) and according to previous reports, size-based filtration method managed to retain

AB polymersomes of the same size and structure without compromising structural integrity despite the shear forces of centrifugation, separating them from smaller sized molecules ^[2-4]. With such considerations, a 1,000 kDa molecular weight cutoff (MWCO) filter unit was used. The large MWCO allows the removal of unincorporated GLP-1R together with small molecule impurities (e.g. nucleotides, enzymes, amino acids, cofactors) from reaction mixture while retaining the polymersomes (Figure 4.2 A). By using western blot, the expressed GLP-1R protein was identified in the retentate, but not in the flow through (Figure 4.2 B). In the control experiments, where cell free protein synthesis was not supplied with AB, all proteins resulted in the flow-through (Appendix VII). This indicates that the GLP-1R was incorporated into AB polymersomes with high efficiency, with negligible amount of free GLP-1R.

Figure 4.2 (A) Schematic illustration of size-based filtration used to separate polymersomes from free proteins and impurities. (B) Western blot analysis of fractions after filtration. Lane 1: soluble fraction of cell free expression; Lane 2: flow-through; Lane 3: retentate; Lane 4: expression without additional plasmid DNA (negative control). Arrow indicates the position of GLP-1R protein.

4.3.2 Protein Yield Improvement Using Bilayer Reaction

For protein expression, higher yield is always desired. This is also a concern for GLP-1R expression using wheat germ extract mostly because of the limited reaction scale (about 50-200 μ L per reaction), large scale production protocol is yet to be developed. Thus, we

decided to approach with bilayer reaction format to potentially increase yield. Instead of mixing all the components in one reaction and shaking, the format is based on a bilayer diffusion system that separates cell extraction/mRNA template mixture from energy source and amino acids in different layers. The bilayer reaction was kept unshaken to maintain the layer separation while synthesizing proteins. It has been reported increased yield of protein expression can be achieved using this format [5]. The detailed reaction setup was described in Chapter 3.

In contrast to the batch method, the bilayer wheat germ is an uncoupled reaction that requires separate preparation of mRNA using pEU vector as described in Chapter 3. GLP-1R was subcloned from the existing pCMVTNT vector to the pEU vector (Figure 4.3). The successful digestion and ligation were confirmed by DNA gel electrophoresis. Sequence of ligated plasmid was confirmed by DNA sequencing.

Figure 4.3 DNA gel electrophoresis of (A) GLP-1R-pCMVTNT digestion and (B) colony test digest. (A) Lane 1: digested GLP-1R in pCMVTNT. Lane 2: digested pEU vector. (B) Lane 1-7: different colonies picked for test digestion. (A, B) Arrows indicate digested vectors and inserted gene. (C) Western blot analysis of cell free expressed GLP-1R using bilayer format. Lane 1: soluble fraction of reaction mixture. Lane 2: filtration retentate. Arrows indicate GLP-1R.

GLP-1R-pEU plasmid was tested for bilayer wheat germ cell free expression and Western blot confirmed expression (Figure 4.3 C). The GLP-1R expressed in bilayer reaction showed bands at similar positions (about 40 and 80 kDa) compared to HYWG systems.

Size-based filtration showed retention of bilayer expressed GLP-1R and confirmed its incorporation into polymersomes. At this stage, it was uncertain if bilayer reaction increased yield of GLP-1R because Western blot detection here is only qualitative due to potential saturated exposure. It can be concluded that bilayer reaction expresses at least equal amount, if not more, as HYWG system.

In addition, the efficiency of the size-based filtration was confirmed by using SYPRO staining of the SDS-PAGE gel (Figure 4.4), which showed a substantial amount of proteins was removed. In the filtration retentate, however, no dominant band can be identified as compared to the negative control without mRNA template. This indicates that the yield of GLP-1R was limited to Western blot detection level (about 0.1 ng) and not detectable by SYPRO staining (about 100 ng). In order to enrich expressed GLP-1R, further and alternative purifications were explored.

Figure 4.4 SDS-PAGE and total protein SYPRO staining of cell free expressed GLP-1R after size based filtration. Lane; flow-through. Lane 2: filtration retentate. Lane 3: filtration retentate of expression reaction without additional mRNA template (background, negative control). M: molecular weight marker. Arrow indicates the GLP-1R.

4.3.3 GLP-1R Purification and Primary Characterization

To purify GLP-1R, two affinity purification methods were investigated as detailed in Figure 4.5. A conventional approach using immobilized-metal affinity chromatography (IMAC) ^[6] was employed using poly-histidine-tagged GLP-1R (protein specific purification). This method has been widely used in free protein purification but according to our knowledge, never applied to protein/vesicle complex purification in systems like lipid mode of cell free protein synthesis. It is probably because, different from purifying soluble proteins, the misorientation of expressed proteins will limit his-tag accessibility to solid matrix, thus limiting the efficacy of histidine-affinity purification. However this requires further investigation. In parallel, we adopted a novel strategy using biotinylated polymersome membranes to take advantage of the specific high-affinity biotin/avidin interaction ^[7] (biotin-affinity purification, membrane specific). The biotinylation is more controllable and flexible as the number of biotin-tags can be tuned during vesicle formation. Specifically, biotin tags were introduced by mixing AB di-block copolymer with DSPE-PEG biotin lipids during vesicle formation. Biotinylated AB polymersomes were then added to the cell free reaction mixture. In biotin-affinity purification, the target molecule of interest has to be biotin labelled which can be difficult or undesired for protein tagging. In comparison, adding histidine tag can be readily achieved by molecular cloning and will be translated upon expression. Thus, biotin tags were introduced to membrane surface whereas histidine tags were introduced to GLP-1R protein, respectively. The elution for both methods are convenient. Biotin or imidazole solution can be used for each method. We compared the two methods in terms of purification efficiency, protein purity, and preservation (support matrix to GLP-1R).

Figure 4.5 Expression of GLP-1R in CFPS and biotin- or histidine-affinity purification. Biotin tags were introduced by mixing AB di-block polymer with DSPE-PEG biotin lipids during vesicle formation. Biotinylated AB polymersomes were added to cell free reaction mixture followed by two purification routes. (Partially adopted from: Overview of protein purification, Qiagen and Thermoal Fisher)

After cell free protein synthesis of GLP-1R, the precipitated protein/polymer were removed by centrifugation at 10,000 rpm. The soluble fraction (supernatant after centrifugation) of the cell free expression system was subjected to either the biotin- or histidine- affinity purifications and analyzed using SDS-PAGE. The result showed that with both methods, yielded higher amount of GLP-1R of higher purity (Figure 4.6), as compared to the filtration retentate (Figure 4.4). Purified receptor with both methods display a prominent band detectable on SYPRO stained gel. Histidine-affinity purified GLP-1R of slightly higher purity (90 % upon visual inspection) and less contamination of other non-specific proteins, as compared to that of biotin-affinity purified. Western blot using anti-C9 monoclonal antibody yielded a single band around 40 kDa confirmed the successful purification of GLP-1R in both methods (Figure 4.6 B). For each reaction, the yield was estimated to be 12-14 μ g.

Figure 4.6 SDS-PAGE SYPRO stained (A) and Western blot (B) analysis of GLP-1R purification using 1: biotin-affinity method and 2: histidine-affinity method. (B) Western blot with an anti-C9 monoclonal antibody against the C9 tag at C-terminal of GLP-1R. M: molecular weight marker. Arrows indicates GLP-1R.

However, these results did not address the question if polymersomes supplied were co-purified or removed. The presence of polymersomes (or polymers) is important because they are supposed to provide support and stabilize GLP-1R. It is very likely that, without co-purifying polymersomes, GLP-1R will be misfolded or aggregated. Previously, Boris Peker et al. reported quantification of liposomes recovery using tritium labelled lipids [8]. It provided high sensitivity and quantitative analysis but the use of radioisotope is not desired. Instead, we labelled AB polymersomes with fluorescent dye tetrazole using light induced cycloaddition chemistry as described by Winna Siti et al. [9-10]. The method allows easy (UV light) and fast (10-15 min) conjugation of tetrazole to double bond of poly (butadiene) in polymers in a one-step reaction. Another advantage is that only the formed

pyrazoline between tetrazole and butadiene is fluorescent but not free tetrazole. This simplifies the functionalization, allows tracing the progress of conjugation and does not require removal of free tetrazole.

Figure 4.7 Tetrazole fluorescence of fractions during biotin- and histidine-affinity purification. Tetrazole fluorescence of CFPS reaction mixture soluble fraction (SUPERNT), unbound polymersomes/receptors (UNBD) during binding with solid matrix, polymersomes/receptors came off from beads during wash process (WASH) and elution (ELUTE). Percentage of fluorescence was normalized to SUPERNT.

Using this strategy, we tracked the presence of polymersomes in the unpurified cell free expression soluble fraction and estimated the amount of polymersomes in the two purified fractions (Biotin- and Histidine-based) by monitoring the tetrazole fluorescence of during each purification step (Figure 4.7).

Using this strategy, we tracked the presence of polymersomes in the unpurified cell free expression soluble fraction and estimated the amount of polymersomes in the two purified fractions (Biotin- and Histidine-based) by monitoring the tetrazole fluorescence of during each purification step (Figure 4.7). During purification, the cell free reaction mixture was

first mixed with solid matrix (either with avidin or Ni-NTA matrix) to allow binding, followed by removal of unbound fraction by centrifugation (UNBD). The matrix was further washed with appropriate buffer to remove loosely bound or attached proteins/polymers (WASH), and finally elution (ELUTE). Each of these fractions were collected and subject to tetrazole fluorescence analysis. The fluorescence of soluble fraction (centrifugation supernatant, SUPERNT) was normalized to 100 %. About 40 % of polymers in biotin-affinity purification did not bind after incubation and for histidine-affinity purification, the unbound fraction showed up to about 60 % of total fluorescence. This indicates the binding conditions in both protocols were not optimized for maximum polymer recovery. Note that intensive washing (30 column volumes) was performed after binding to remove proteins/polymers. The last fraction of washing was collected for fluorescence analysis, which showed minimal tetrazole fluorescence and confirmed the sufficient washing. Elution of biotin-affinity purification accounts about 30 % of total intensity of reaction mixture whereas that of histidine-affinity purification is minimal (less than 3%). The fluorescence intensity/polymer amount were not strictly correlated to the GLP-1R amount in SDS-PAGE but an over 10 times difference was evidence for the presence of AB polymers only in biotin-affinity purified GLP-1R. The results suggest that histidine-affinity purification dissociated proteins from vesicles, which can be the real limiting factor of this method. It is also showed that in histidine-affinity purified GLP-1R formed aggregates (at about 150 kDa), which cannot be fully dissociated even after treated by SDS before gel electrophoresis (Figure 4.6B). There was no precipitate observed during the purification, indicating the GLP-1R only formed small aggregates/oligomers but not precipitation.

4.4 Conclusions

First, GLP-1R was successfully expressed and spontaneously incorporated into AB polymersomes using wheat germ based cell free expression system. Both coupled and bilayer reaction produced equal quality protein in terms of expression level and incorporation.

Besides, both biotin and histidine-affinity purification yielded high purity GLP-1R protein at analytical scale. Histidine-affinity purification, as a conventional approach, dissociated protein from polymersomes. We tentatively concluded such dissociation limited its application in purifying protein/vesicle complex, e.g. proteoliposomes. In comparison, biotin-affinity purification has yielded protein/polymer complex and is a more promising method to produce GLP-1R without losing its native like membrane support (polymersomes). The results also drove us to further investigate the polymersome structure and functionality of purified product.

References

- [1] A. Rath, M. Glibowicka, V. G. Nadeau, G. Chen, C. M. Deber, *Proceedings of the National Academy of Sciences* **2009**, *106*, 1760-1765.
- [2] H.-P. M. de Hoog, E. M. Lin JieRong, S. Banerjee, F. M. Décaillot, M. Nallani, *PLoS ONE* **2014**, *9*, e110847.
- [3] S. May, M. Andreasson-Ochsner, Z. Fu, Y. X. Low, D. Tan, H.-P. M. de Hoog, S. Ritz, M. Nallani, E.-K. Sinner, *Angewandte Chemie International Edition* **2013**, *52*, 749-753.
- [4] M. Nallani, M. Andreasson-Ochsner, C.-W. Tan, E.-K. Sinner, Y. Wisantoso, S. Geifman-Shochat, W. Hunziker, *Biointerphases* **2011**, *6*, 153-157.
- [5] T. Sawasaki, Y. Hasegawa, M. Tsuchimochi, N. Kamura, T. Ogasawara, T. Kuroita, Y. Endo, *FEBS Letters* **2002**, *514*, 102-105.
- [6] J. A. Bornhorst, J. J. Falke, *Methods in enzymology* **2000**, *326*, 245-254.
- [7] N. M. Green, *Biochemical Journal* **1963**, *89*, 599-609.
- [8] B. Peker, J. J. Wu, J. R. Swartz, *Biotechnology Progress* **2004**, *20*, 262-268.
- [9] W. Siti, A. K. Khan, H.-P. M. de Hoog, B. Liedberg, M. Nallani, *Organic & Biomolecular Chemistry* **2015**, *13*, 3202-3206.
- [10] H.-P. M. de Hoog, M. Nallani, B. Liedberg, *Polymer Chemistry* **2012**, *3*, 302-306.

Chapter 5

Biophysical Characterization of Proteopolymersomes

In the previous chapter, the GLP-1R was produced using wheat-germ cell free system and purified using three different methods, including size-based filtration, histidine- and biotin- affinity purification. Both affinity-purification methods yielded high purity protein but only biotin-affinity purified GLP-1R showed both protein and polymer recovery. Next, it is important to determine the structural and functional characteristics of GLP-1R produced using this method. This chapter focuses on biophysical characterization of the produced GLP-1R proteopolymersomes by TEM followed by FSEC, which allows probing the dispersity of GLP-1R proteopolymersomes, particularly the association of proteins in polymersomes.

5.1 Introduction

In the previous chapter, an optimized expression of GLP-1R based on the wheat germ cell free system with a decoupled bilayer reaction was obtained. Following that, the expressed GLP-1R was purified to >90% purity using a novel purification strategy using biotinylated polymersomes. This method was shown to be superior to conventional gel-filtration and histidine-tagged protein purification. Furthermore, based on SDS-PAGE, immunoblotting, and fluorescence spectroscopy, the biotin-affinity purification was shown to co-purify the polymersomes, an important finding not achieved by using histidine-affinity purification. Importantly, a crucial methodology involving the labeling of polymer with tetrazole, which only fluoresce when conjugated to the butadiene moiety of the polymer, makes parallel monitoring of the polymersomes and the protein possible.

In this chapter, biophysical characterizations of GLP-1R proteopolymersomes are described. Transmission electron microscopy (TEM) and dynamic light scattering (DLS) confirmed the presence of vesicular structures, indicating intact polymersomes. Further characterization with UV detection- (SEC) and fluorescence detection- size-exclusion chromatography (FSEC) confirmed GLP-1R incorporation into polymersomes. The use of FSEC further resolved two populations of polymersomes with different sizes that exist in the expression mixture and purified products, of which the second population (smaller vesicles) were identified to have the incorporated protein.

5.2 Experimental Methodology

5.2.1 Molecular Cloning

Molecular cloning was performed to generate GFP-GLP-1R in the expression pEU vector for bilayer cell free reaction. All cloning steps were monitored by DNA gel electrophoresis. For details, refer to chapter 3.

5.2.2 Protein Expression and Purification

Protein expression was performed using both coupled and bilayer reactions. Purification was performed using filtration, biotin- and histidine- affinity methods. For details, refer to chapter 3.

5.2.3 Polymersome Preparation

AB polymersomes was prepared using film rehydration. For fluorescence and biotin-tagged vesicles, tetrazole or DSPE-PEG-biotin was added during film formation. Tetrazole fluorescence was measured using Tecan 200 plate reader. For details, refer to chapter 3.

5.2.4 Protein Gel Electrophoresis and Western Blot

Protein gel electrophoresis and Western blot were performed for detection of GLP-1R protein and purity examination. For details, refer to chapter 3.

5.2.5 Bicinchoninic Acid Assay (BCA)

BCA was performed to measure protein concentration of purified proteins. For details, refer to chapter 3.

5.2.6 Transmission Electron Microscopy (TEM)

TEM was performed to identify structural characteristics of different preparations. For details, refer to chapter 3.

5.2.7 Fluorescence Detection Size-Exclusion Chromatography (FSEC)

FSEC was performed to analyze protein/polymer association before and after purification. For details, refer to chapter 3.

5.2.8 Dynamic Light Scattering (DLS)

DLS was performed to measure the size of AB polymersomes.

5.3 Principle Outcomes

5.3.1 Morphological Characterization of Purified Product

In the previous chapter, biotin-affinity purification has been shown to preserve high content of protein/polymer association. However, the characterization methods did not provide any structural information about the purified product. To exhibit optimal functional properties, GLP-1R should be maintained in a lamellar phase structure. The AB polymersomes used in cell free protein synthesis were first characterized using dynamic light scattering (Figure 5.1). The preformed vesicles showed average size of 170 nm in diameter with polydispersity of 0.083, indicating a monodisperse population of AB polymersomes.

Figure 5.1 Dynamic light scattering of AB polymersomes. Top: intensity-weighted size distribution, bottom: correlation coefficient decay. The average size of AB polymersomes was determined to be 170 nm with polydispersity index of 0.083.

Figure 5.2 Transmission electron microscopy images of (A) AB polymersomes, (B) GLP-1R proteopolymerosomes before purification and (C) biotin-affinity purified GLP-1R proteopolymerosomes. Arrows indicate vesicular structures. (D) Dynamic light scattering

measurements of biotin- and histidine- affinity purified GLP-1R. The measured sizes are indicated in the figure with PDI of 0.310 ± 0.097 and 0.825 ± 0.303 for biotin- and histidine- affinity purified samples.

We further characterized biotin-affinity purified AB proteopolymersomes using TEM and compared the structure to AB polymersomes before and after cell free reaction of GLP-1R (Figure 5.2 A-D). TEM images and DLS measurements showed that the vesicular structures remained unchanged after the reaction and biotin-affinity purification. Since all TEM was performed under dry condition, it did not reflect the actual size of vesicles in solution. For histidine-affinity purification, the DLS showed multiple populations with abnormal size distribution and the high PDI (moderate PDI=0.1~0.4) indicates the presence of aggregates (Figure 5.2 D). Thus, TEM characterization was not performed for this sample. To get more information on vesicle size distribution and to elucidate the association state(s) of protein/polymersome complex, we proceeded to FSEC. For characterization of GLP-1R with FSEC, GFP has to be cloned as a fluorescent tag.

5.3.2 Molecular cloning and expression of GFP-GLP-1R Proteopolymersomes for FSEC Characterization

Size-exclusion chromatography (SEC) has long been used for protein analysis and purification ^[1]. In normal SEC, UV detection is used to monitor protein elution and separation. A disadvantage is that it requires relatively high amount and concentration of protein to be analyzed (roughly 0.5-1 mg/mL). Besides, it is difficult to analyze unpurified protein with UV detection, e.g. in cell free reaction mixture. In our case, a chromatography method that enables the tracking of the protein as well as the polymersomes was required. We adopted the method reported by Sylviane Lesieur et al. in which SEC is coupled to fluorescence detection (FSEC) ^[2]. By expressing green fluorescence protein (GFP) fused GLP-1R (GFP-GLP-1R), it allowed us to track both the target protein and tetrazole labeled polymersomes based on fluorescence profiles during SEC elution. The tetrazole (370/470 nm) and GFP (450/520 nm) fluorescence do not interfere with each other, thus both fluorescence profiles will be analyzable.

For producing green fluorescence protein tagged GLP-1R, GFP gene from pGFP vector was amplified by PCR and inserted into the pEU vector containing GLP-1R (Figure 5.3 A-C). Correct digestion and ligation were confirmed by DNA gel electrophoresis and sequencing to make sure no mutation was generated. Expression of the new GFP-GLP-1R protein was further confirmed using HYWG bilayer cell free reaction (Figure 5.3 D). Similar to GLP-1R, the apparent molecular weight (MW) is about 70 kDa and is lower than the theoretical MW of 82.6 kDa with a minor oligomer band at about 120 kDa.

Figure 5.3 Molecular cloning of the fusion protein GFP-GLP-1R. DNA gel electrophoresis of (A) PCR amplified GFP gene from pGFP vector, (B) digested GFP gene (insert, 0.7 kb) and pEU vector containing GLP-1R before ligation, (C) test digest of GFP-GLP-1R-containing pEU vector. Lane 1-5: colonies picked. (D) Western blot analysis GFP-GLP-1R expression in HYWG cell free system. Lane 1: GFP-GLP-1R. M: molecular weight marker.

The expressed GFP-GLP-1R was further purified using the biotin- or histidine- affinity purification. Purified GFP-GLP-1R showed similar purity with both purification methods (Figure 5.4).

Figure 5.4 SDS-PAGE (A) and Western blot (B) analysis of GFP-GLP-1R purification using 1: biotin-affinity method and 2: histidine-affinity method. (B) Western blotting with a monoclonal antibody against the C9 tag at C-terminal domain of the GFP-GLP-1R. M: molecular weight marker. Arrow indicates GFP-GLP-1R.

5.3.3 Size-exclusion chromatography Analysis of GLP-1R incorporation into polymersomes

To investigate if GLP-1R was incorporated into polymersomes or membrane dissociated as free proteins, normal size-exclusion chromatography (SEC) was performed using a Sepharose-based column with resolution of 10 to 600 kDa (globular proteins). The

Sephacrose gels are commonly used in the removal of small solutes non-entrapped in liposomes^[3]. The SEC was accompanied by UV detection and to identify the presence of GLP-1R in elution, fractions were collected and examined by Western blot (Figure 5.5). The analysis of UV profile showed two major populations at 1.0-1.5 and 2.5-3.5 mL, respectively. However, since the GLP-1R proteopolymersomes were unpurified from CFPS, the UV trace represented the elution of total protein instead of GLP-1R alone (Figure 5.5 A). The bovine serum albumin (BSA, 66 kDa) was injected and the peak position of the volume fraction at which it was eluted served as a reference position for free GLP-1R (53 kDa). The tetrazole fluorescence showed that all AB polymersomes were first eluted after void volume of 1 mL (Figure 5.5 B). Western blot showed that almost all GLP-1R were eluted at the start of void volume (1 mL), indicating the majority of expressed GLP-1R were incorporated (Figure 5.5 C). This was further confirmed by that minimal GLP-1R eluted after BSA (2 mL). However, this analysis did not provide detailed information about the AB polymersomes, e.g. size distribution, and GLP-1R protein/polymer association state(s).

To further characterize the produced GLP-1R proteopolymersomes, we proceeded with fluorescence detection size-exclusion chromatography (FSEC) with a polymethacrylate-based column. It has been validated to be able to resolve lipid vesicles over a wider range (25-300 nm)^[4] and been used in further FSEC analysis with GFP-GLP-1R and tetrazole labelled AB polymersomes as tracers. The tetrazole fluorescence profile of GFP-GLP-1R expressed in the presence of AB polymersomes (unpurified) showed two population at about 6.5 and 9.2 mL (Figure 5.6). We compared it to the elution profiles of liposomes of different sizes (300, 145, 80 and 25 nm) and the two populations corresponded 300 nm (6.5 mL) and 145 nm (9.2 mL), respectively. The particle size of 145 nm was consistent with dynamic light scattering (DLS) results which showed prepared vesicles had an average size of 170 nm, considering the measured hydrodynamic diameter can be larger in DLS (Figure 5.1). The 300 nm vesicles displayed a sharp and symmetric peak in FSEC (Figure 5.6), which excluded the possibility of forming random polymer aggregation. As AB polymersomes were extruded through a 0.22 μ m filter, it is not surprising that larger structures were also formed. Besides, the cross-linking of tetrazole via photo induction may

cause membrane fluctuation, which would also change the properties of vesicles ^[5]. The GFP profile suggested that GFP-GLP-1R incorporated favorably into vesicles of 145 nm in size and were co-eluted between 8.5-10.5 mL. Besides the incorporated protein at 9-9.5 mL, the elution of target protein further extended to 10.5 mL. The column used is known to display high resolution in separating large MW vesicular structures between 25 and 300 nm but limited protein separation power (660-12 kDa from 9.5 to 10.5 mL). Thus, it is difficult to identify the elution of different GFP-GLP-1R species other than the incorporated population because both aggregated and free GFP-GLP-1R were eluted within 1 mL and the elution profile was overlapping with the small vesicles.

Figure 5.5 Size-Exclusion Chromatography analysis of GLP-1R proteopolymerosomes using Sepharose-based column. (A) UV traces of GLP-1R proteopolymerosomes (unpurified expression

mixture). Void volume and exclusion limit indicate column resolution. BSA: bovine serum albumin (66 kDa) was injected as reference to distinguish incorporated and unincorporated GLP-1R (53 kDa). Tetrazole fluorescence (B) and Western blot (C) analysis of elution fractions from (A).

Figure 5.6 Fluorescence detection size-exclusion chromatography (FSEC) profile of cell free expressed GLP-1R proteopolymersomes before purification. Protein elution: column resolution for protein separation. Vesicle size indicates the elution volume of liposomes with defined sizes, by Per Lundahl et al. ^[4].

To further characterize whether the AB polymersomes form two distinct population or is it due to the tetrazole, we analyzed the polymersomes by encapsulating calcein without any tetrazole labeling in FSEC. Calcein as a fluorescent dye is self-quenching when encapsulated at high concentration (above 30 mM) and was used in monitoring vesicle fusion and lysis. When leakage occurs by detergent lysis, released calcein leads to fluorescence increase ^[6]. We adopted this method and analyzed eluted fractions in FSEC.

Different volume fractions throughout the elution were collected and subjected to the calcein leakage test. The result showed significant increase in calcein fluorescence (2 fold increase after lysis) only in fractions corresponding to the second population (9.2 mL, the vesicle size of 145 nm) as shown in Figure 5.7. This indicates that in the prepared polymersomes, small lamellar vesicles of defined size (145 nm) were formed as shown by the calcein leakage profile. The larger structures were also formed possibly because of photo-induced cross-linking. In comparison to the smaller vesicles, they might not form lamellar phase structures as suggested by the lack of calcein encapsulation.

Figure 5.7 FSEC using calcein encapsulated AB polymersomes. Protein elution: column resolution for protein separation. Vesicle size indicates the elution volume of liposomes with defined sizes, by Per Lundahl et al. ^[4].

As shown in Chapter 4 and previous sections, both purification methods (biotin- and histidine- affinity based) yielded high purity GLP-1R. However, only biotin-affinity purification produced proteins associated with polymersome (proteopolymersomes) and

histidine-affinity purification very likely produced protein aggregates. We further characterized the purified products with these two methods using FSEC to gain information about protein/polymer association. As shown in Figure 5.8, biotin-affinity purified GFP-GLP-1R displayed a similar profile as unpurified sample in Figure 5.6, where two populations of vesicles of different sizes existed and target protein was incorporated into small vesicles. This was consistent with previous hypothesis that the absence of bilayer membrane architectures resulted in large structures, which did not facilitate protein incorporation. In comparison, small vesicles with lamellar structures were more stable and favorably assembled GFP-GLP-1R. Moreover, the small vesicles were enriched after biotin-affinity purification, indicating the defined bilayer membrane architectures facilitated protein incorporation.

The FSEC profiles of histidine-affinity purified sample showed only GFP elution without any tetrazole fluorescence (Figure 5.8 B). The elution volume in FSEC extended from 7 to 10 mL and the shape of GFP peak was identical to that of unpurified protein expressed in the absence of AB polymersomes (Figure 5.8 C). Despite the limited protein separation, this elution was beyond the upper limit of 660 kDa, indicating large protein aggregates formed without membrane support.

Figure 5.8 Fluorescence detection size-exclusion chromatography (FSEC) of GFP-GLP-1R proteopolymerosomes. (A) Biotin- and (B) histidine-affinity purified GFP-GLP-1R proteopolymerosomes. (C) Cell free protein synthesis of GFP-GLP-1R protein without supplied polymerosomes. The sample was directly analyzed after synthesis without any purification. Protein elution: column resolution for protein separation. Vesicle size indicates the elution volume of liposomes with defined sizes, by Per Lundahl et al. ^[4].

5.4 Conclusions

Extensive characterizations were performed after GLP-1R produced using WG CFPS and purified. The structural characterization using TEM confirmed the vesicular structures in biotin-affinity purified GLP-1R. Further characterizations using SEC and FSEC indicate that the proteins were incorporated into AB polymerosomes of smaller size (145 nm).

Meanwhile, large structures were formed but probably did not have bilayer membrane architectures or did not facilitate protein assembly. The histidine-affinity purified GFP-GLP-1R was confirmed to be deprived of polymers, thus formed large aggregates. Based on the characterization results using GFP-GLP-1R and GLP-1R (from previous chapter). Only biotin-affinity purification produced proteopolymersomes because the method is membrane specific. In biotin-affinity purification, 5% of biotin tags are introduced to polymersomes and the formation of vesicles and integration efficacy of tags is controllable and reliable. During the purification, the avidin functionalized solid matrix is allowed to bind to polymersomes, keep the vesicular structure intact. However, the number of accessible histidine tags is unknown and according to the purification results, it is probably less than the number of biotin tags, thus not sufficient to for purifying GLP-1R protein with intact polymersomes. In this case, the histidine tagged proteins are dissociated from the polymer membranes, thus formed aggregates in cases of GLP-1R.

References

- [1] P. Hong, S. Koza, E. S. P. Bouvier, *Journal of Liquid Chromatography & Related Technologies* **2012**, 35, 2923-2950.
- [2] T. Kawate, E. Gouaux, *Structure* **2006**, 14, 673-681.
- [3] C. Grabielle-Madelmont, S. Lesieur, M. Ollivon, *Journal of Biochemical and Biophysical Methods* **2003**, 56, 189-217.
- [4] P. Lundahl, C.-M. Zeng, C. Lagerquist Hägglund, I. Gottschalk, E. Greijer, *Journal of Chromatography B: Biomedical Sciences and Applications* **1999**, 722, 103-120.
- [5] C. K. Haluska, M. S. Baptista, A. U. Fernandes, A. P. Schroder, C. M. Marques, R. Itri, *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2012**, 1818, 666-672.
- [6] M. Noor, T. Dworeck, A. Schenk, P. Shinde, M. Fioroni, U. Schwaneberg, *Journal of Biotechnology* **2012**, 157, 31-37.

Chapter 6

Biochemical Characterization of Proteopolymersomes

This chapter focuses on biochemical characterization of glucagon like peptide-1 receptor (GLP-1R) in polymersomes, following the evidence that the expressed GLP-1R is associated with the polymersomes, as shown by size exclusion chromatograph (SEC and FSEC). To probe the receptor orientation, Surface Plasmon Resonance was used. By immobilizing the GLP-1R proteopolymersomes on the sensor surface, the receptor orientation was examined using antibodies specific to the extracellular or intracellular domains of GLP-1R. The functionality of the extracellular N-terminal domain of GLP-1R was further investigated using a radioligand saturation binding assay, a gold standard for functional protein evaluation. The result showed that the GLP-1R purified by the biotin-affinity method was not functional. To address this, a chemical chaperone was included in the cell free mixture. Biophysical and biochemical characterizations were then performed to investigate whether the additional chaperoning effect had any consequences on the insertion of GLP-1R into polymersomes and the results demonstrated that the purified GLP-1R polymersomes were structurally similar (without additional chaperones) and was also functional, as demonstrated by the ligand binding assay.

6.1 Introduction

In the previous chapter, the protein/polymer complex were co-purified by the biotin-affinity method, without compromising the vesicular structure of AB polymersomes, the membrane support of GLP-1R. This has been shown by the presence of vesicular structure by TEM, calcein encapsulation, SEC and FSEC. In this chapter, the functionality of GLP-1R was examined, in terms of the orientation of GLP-1R in the polymersomes, its ability to bind to antibody raised against it as well as its natural peptide ligand. The function was then confirmed by a radioligand saturation binding assay. To probe the receptor orientation, Surface Plasmon Resonance (SPR) was used. The GLP-1R proteopolymersomes were immobilized on the anti-PEG modified sensor surface and the receptor orientation was probed using antibodies specific to the extracellular or intracellular domains of GLP-1R. About half of the polymersomes contain correctly oriented GLP-1R. The functionality of the extracellular N-terminal domain (NTD) of GLP-1R was further investigated using the radioligand saturation binding assay. Unfortunately, the GLP-1R didn't show any functional binding. To address this, addition of chaperones into the cell free mixture were investigated. After which, the GLP-1R insertion into polymersomes were probed for integration and functionality by the ligand binding assay.

6.2 Experimental Methodology

6.2.1 Molecular Cloning

Molecular cloning was performed to generate target DNA in the expression pEU vector for bilayer cell free reaction. All cloning steps were monitored by DNA gel electrophoresis. For details, refer to chapter 3.

6.2.2 Protein Expression and Purification

Protein expression was performed using cell free protein synthesis. Purification was performed using filtration, biotin- and histidine- affinity methods. For details, refer to chapter 3.

6.2.3 Polymersome Preparation

AB polymersomes was prepared using film rehydration. For fluorescence and biotin-tagged vesicles, tetrazole or DSPE-PEG-biotin was added during film formation. Tetrazole fluorescence was measured using Tecan 200 plate reader. For details, refer to chapter 3.

6.2.4 Protein Gel Electrophoresis and Western Blot

Protein gel electrophoresis and Western blot were performed for detection of GLP-1R protein and purity examination. For details, refer to chapter 3.

6.2.5 Bicinchoninic Acid Assay (BCA)

BCA was performed to measure protein concentration of purified proteins. For details, refer to chapter 3.

6.2.6 Transmission Electron Microscopy (TEM)

TEM was performed to identify structural characteristics of different preparations. For details, refer to chapter 3.

6.2.7 Fluorescence Detection Size-Exclusion Chromatography (FSEC)

FSEC was performed to analyze protein/polymer association before and after purification. For details, refer to chapter 3.

6.2.8 Surface Plasmon Resonance (SPR)

SPR was performed to detect specific binding of GLP-1R NTD to specific antibody. For details, refer to chapter 3.

6.2.9 Radioligand Binding Assay

Radioligand binding assay was performed to detect specific binding between GLP-1R and its ligand. For details, refer to chapter 3.

6.3 Principle Outcomes

6.3.1 Determination of GLP-1R Orientation in Polymersomes and Ligand Binding by Surface Plasmon Resonance (SPR)

The accessibility of the outward facing NTD of GLP-1R is crucial for peptide ligand binding. To determine the orientation of the GLP-1R in proteopolymersomes, we first probed NTD of GLP-1R, which is solvent exposed at the exterior of the polymersomes^[1]. We used a Surface Plasmon Resonance (SPR) based Biacore system, which allows immobilization of biomolecules on a sensor surface, e.g. antibodies^[2], and monitoring binding of ligands, e.g. peptide^[3], small molecules^[4].

We first tested the conformational sensitivity of selected NTD specific polyclonal antibody (pAB) by comparing it to a monoclonal antibody (1D4), which targets the linear epitope C9-tag at the C-terminal of GLP-1R. Recognitions of GLP-1R by both antibodies were comparable, as shown by the Western blot result (Figure 6.1 A). This suggests that pAb may also recognize a contiguous sequence of GLP-1R since all proteins were denatured during gel electrophoresis and blotting.

Figure 6.1 Characterization of antibody conformation sensitivity and sample preparation. (A) Western blot analysis of GLP-1R using 1D4 and pAb. Lane 1: GLP-1R expressed using cell free protein synthesis (CFPS). Lane 2: CFPS without additional mRNA (negative control). (B) SDS-PAGE and Western blot of unpurified and biotin-affinity purified chemokine CXC receptor-4 (CXCR4) using anti-C9 antibody. SN: soluble fraction of cell free reaction mixture. Pure: biotin-affinity purified CXCR4 proteopolymersomes. M: molecular weight marker.

There are very few reports about immobilization and analysis of proteopolymersomes in Biacore system ^[5-7]. Two different preparations of GLP-1R were tested, including the unpurified GLP-1R proteopolymersomes from CFPS reaction mixture and the biotin-purified product. Both samples were captured on an anti-PEG antibody immobilized surface to minimize the GLP-1R denaturation, if any. To confirm if the capturing of proteopolymersomes on the surface was specific to the polymer moiety and independent of the expressed protein, polymersomes expressed with a non-relevant target chemokine CXC receptor-4 (CXCR4) were used. The SPR sensorgrams showed both GLP1-R and CXCR4 proteopolymersomes were immobilized using this method as shown in Figure 6.2 and the unpurified samples were immobilized to a higher level (3000 RU) than biotin-affinity purified samples (1000 RU). Besides, no significant difference of immobilization level was observed between GLP-1R and CXCR4.

Figure 6.2 SPR sensorgrams of proteopolymersomes immobilization using antibody capture method. Immobilization of (A) unpurified and (B) purified proteopolymersomes using antibody capture. (A, B) Anti-PEG antibody was first immobilized onto biosensor chip using standard amine

coupling reaction, followed by flowing proteopolymersomes onto surface to enable antibody capture. The immobilization level was 3000 RU for unpurified and 1000 RU for biotin-affinity purified proteopolymersomes with no significant difference between GLP-1R and CXCR4.

Specific binding to N-terminal extracellular domain specific antibody pAB was detected in a dose response manner from 2 to 33 ug/ml for both unpurified GLP-1R proteopolymersomes from CFPS reaction mixture and the biotin-purified product (Figure 6.3). The unpurified sample showed up to 80 RU response at the highest antibody concentration of 33 ug/mL, which was almost 8 times higher than the purified GLP-1R proteopolymersomes (10 RU). One of the difference between the two samples was the higher immobilization level of unpurified sample (3000 RU) compared to purified (1000 RU) sample that led to lower density of accessible GLP-1R on the biosensor chip surface for pAB binding. However, the 8 times difference in response indicates other factors influencing the binding. Since the measurement conditions (e.g. flow rate, antibody concentrations, and association/dissociation time) were identical for both samples, the different response is probably due to the changes of protein/polymer interaction that led to less accessible NTD of GLP-1R caused by biotin-affinity purification. The dose response binding profile allows the determination of association rate K_{on} to be $3.33^4 \text{ M}^{-1}/\text{s}$, dissociation rate K_{off} to be 6.6^{-4} s^{-1} (the rate of N-terminal recognizing antibody binding to/dissociating from GLP-1R protein), which further led to the dissociation constant K_d to be 18.6 nM according to 1:1 binding kinetic model. Thus, accessible NTD of GLP-1R was presented on AB polymersome surface, indicating that GLP-1R assumed a correct orientation.

Furthermore, the binding of exendin-4, a natural peptide ligand of GLP-1R and 1D4 to GLP-1R were examined. The response in Biacore system is directly related to the change in mass concentration on the surface, so that the absolute response are proportional to the size of molecules flowed over the surface ^[8]. In real practice, however, such correlation will make analyzing GLP-1R's peptide ligand binding difficult because of the low MW of these peptides (3 kDa) compared to the antibodies (150-180 kDa). Based on antibody binding results, the response expected for peptide binding was about 1 RU (background

level) ^[9]. Besides, the low density of receptors on membrane surface will also lower signal. Indeed, the exendin-4 binding to GLP-1R proteopolymersomes showed minimal response and no dose dependency at concentrations from 0.1 to 1,000 nM (Figure 6.4 A, B). On the other hand, the 1D4 antibody binding to C-terminal C9 tag was also tested and showed up to 140 RU response for both GLP-1R and CXCR4 (Figure 6.4 A, B), indicating the presence of reversely incorporated proteins in AB polymersomes. To further probe the functionality with higher sensitivity, we performed radioligand binding assay.

Figure 6.3 SPR sensorgrams of NTD specific pAb binding to (A) unpurified and (B) purified GLP-1R proteopolymersomes. Sensorgram of the binding for (from bottom to top) increasing

concentrations of pAb antibody. Binding curves were fitted to a 1:1 kinetic binding model and an affinity of 18.6 nM was derived from the fitting.

Figure 6.4 SPR sensorgrams of peptide ligand exendin-4 binding to proteopolymersomes. Unpurified (A) and (B) biotin-affinity purified GLP-1R and CXCR4 proteopolymersomes binding to exendin-4 and 1D4 antibody. Proteopolymersomes were immobilized onto biosensor chip using anti-PEG capture. Exendin-4 of from 0.1 to 1,000 nM was flowed over immobilized proteopolymersomes with buffer injections in between. 1D4 mAb: 1D4 monoclonal antibody targets C-terminal tag of GLP-1R and CXCR4.

6.3.2 Lack of GLP-1R Functionality Determined Using Radioligand Saturation Binding Assay

High affinity peptide binding is a hallmark of most class B GPCRs, including GLP-1R ^[10]. The functionality is an important factor of GPCRs because it indicates if the protein is

folded correctly. In applications such as therapeutic antibody generation, the functional GLP-1R proteins could generate conformational sensitive antibodies, which are more valuable than non-conformational sensitive antibodies. The functionality of different GLP-1R preparations was tested using saturation binding assay to determine the binding affinity of its cognate ligand GLP-1. Different GLP-1R preparations were tested, including membrane preparation (positive control, from transfected mammalian cells with GLP-1R), vivaspin filtration retentate, biotin-affinity and histidine-affinity purified products (Figure 6.5).

Biotin-affinity purified GLP-1R was the most promising preparation shown to be functional as compared to the other two. First, both GLP-1R protein and AB polymersomes were co-purified in this preparation and the proteopolymersome remain structurally integer after purification (TEM). Secondly, SPR showed its NTD on the membrane surface was accessible for specific antibody binding. In contrast, the histidine-affinity purified GLP-1R was the least competent preparation because of protein aggregation after dissociated from support membranes. In comparison, the vivaspin filtration preserved the majority of expressed GLP-1R proteins but showed poor purity and low concentration compared to background.

The dissociation constant (K_d) and maximum binding capacity (B_{max}) of the GLP-1R control membrane preparations were determined experimentally to be 0.24 nM and 2.2 pmol/mg respectively (Figure 6.5 A). For the rest, however, no binding was detected.

In parallel to functional characterization of GLP-1R proteopolymersomes using radioligand binding assay, a number of non-radioisotope-based methods were used, including fluorescence-based (fluorescence polarization), immunodetection-based (enzyme-linked immunosorbent assay) and Surface Plasmon Resonance-based (alternative attempts not included in main chapters) assays. However, due to the low sensitivity of these methods and insufficient amount of functional receptors, these attempts were not successful in identifying GLP-1R functionality. The detailed results are included in Appendix VIII and IX. In the next chapters, radioligand binding assay will be used in

functional characterization of GLP-1R proteopolymerosomes. To ensure the specificity of binding in this assay, negative controls (e.g. non-relevant target) and different assay format (i.e. competition binding) will be used.

Figure 6.5 GLP-1R is retained in polymerosomes. Specific binding of [I-125] GLP-1 to (A) GLP-1R membrane preparation (positive control), GLP-1R AB proteopolymerosomes purified using (B)

filtration, (C) biotin-affinity and (D) histidine-affinity purification. Dissociation constant (K_d) of membrane preparation was calculated to be 0.24 nM, with a B_{max} of binding of 2.2 pmol/mg of protein for membrane preparation (A). Top: specific binding, fitted saturation binding curves using 1:1 model. Bottom: raw binding curves. TB: total binding. NSB: non-specific binding.

6.3.3 Assisted Folding of GLP-1R Enhances Functionality

The lack of functionality can also be explained from a cellular biology perspective. Different from soluble protein, MP biosynthesis involves a series of elaborately regulated processes from the endoplasmic reticulum (ER) through the Golgi to the plasma membrane [11]. During these processes, a number of molecular chaperones were involved in MP folding such as ER residents, heat shock proteins (HSPs) [12] targeting cell membranes, ribosome guiding protein, the signal recognition particle (SRP) [13]. However, such regulation system is absent in CFPS because of the missing compartments like ER during extract preparation in WG CFPS [14]. Although its impact on protein synthesis is yet clear, for MPs, the missing chaperoning factors potentially lead to misfolding. Nonetheless, the nature of cell free reaction being an open system allows extensive adjustment and fine control of reaction conditions to mimic cellular environment and produce quality receptors. In CFPS, chemical chaperones, e.g. detergents and lipids, have been applied to assist *in vitro* refolding [15]. Compared to protein chaperones, a diverse group of chemical molecules are available and can be readily supplied to cell free reaction without adjusting production strategy for the additional molecules [16].

In the case of GLP-1R expression, we introduced chemical chaperones to mimic the function of molecular chaperones [15]. For MP like GLP-1R, the chemical chaperone required will be able to assist folding and direct folded proteins into AB polymersomes without losing functionality. With such considerations, the zwitterionic detergents, the Fos-choline (Fos) series were selected. The zwitterionic nature of Fos series is more effective in solubilizing MPs than the nonionic detergents, which are commonly used in protein structure studies [17-18] and structurally, they resemble phospholipids' native-like environment. These properties made the Fos-choline series ideal candidates.

Using the Fos-choline 14 (Fos14) detergent as chemical chaperone to assist the folding of GLP-1R, the CFPS reaction was modified by supplying it with both AB polymersomes and Fos14. We examined if the addition of Fos14 would mediate a more stable incorporation into the polymersomes. First, CFPS was supplied with both polymersomes and Fos14 to test the expression followed by removal of Fos14 via dialysis and biotin-affinity purification, thus to ensure it does not interfere with the downstream purification or characterizations. The purified sample was further analyzed using SDS-PAGE and immunoblotting. The purified GLP-1R as a major monomer band about 40 kDa with a dimer band at about 70 kD, confirming the successful biotin elution (referred as GLP-1R F-AB, Figure 6.6 A, B, lane 1). Compared to previous results where Fos14 was not supplied in CFPS (Figure 4.6), the elution of GLP-1R F-AB was of lower purity (estimated to be about 50% via visual inspection), possibly due to the effect of Fos14. . Compared to the histidine-purified GLP-1R expressed only in the presence of Fos14 (details are present in the next paragraph), the biotin eluted GLP-1R F-BD showed a slight shift up to higher MW (Figure 6.6 A, B, lane 1-3). The shift in the migration is likely due to the presence of polymer ^[19] and interference by Fos14 ^[20]. Furthermore, the similar profile of tetrazole fluorescence in the fractions from biotin-affinity purification of GLP-1R proteopolymersomes produced in the presence of Fos14 (Figure 6.6 C) compared to previous purification without addition of Fos14 in CFPS (Figure 4.7), indicating that the co-elution of AB polymers was not affected by Fos14 in the CFPS.

Figure 6.6 Production of GLP-1R in the presence of Fos14. (A) Western blot and (B) SDS-PAGE analysis of purified proteins. (A, B) Lane 1: biotin-affinity purified GLP-1R proteopolymersomes. Lane 2: histidine-affinity GLP-1R from CFPS with supplied Fos14 but not polymersomes. Lane 3: sample in lane 2 after Fos14 removal. Lane 4: non-relevant target DRD2 purified as the sample in lane 1. M: molecular weight marker. Sample 1, 3 and 4 were used for the following radioligand

binding assay. (C) Tetrazole fluorescence of fractions from biotin-affinity purification of GLP-1R proteopolymersomes produced using CFPS with supplied AB polymersomes and Fos14.

In CFPS supplied with both AB polymersomes and Fos14, the biotin-affinity purified GLP-1R proteopolymersomes was further characterized using transmission electron microscopy (TEM), which confirmed the presence of vesicular structures in the purified product (Figure 6.7).

Figure 6.7 Transmission electron microscopy (TEM) images (A, B) of biotin-affinity purified GLP-1R proteopolymersomes produced using CFPS with supplied AB polymersomes and Fos14. Arrows indicate vesicular structures.

To test the effects of Fos14 and AB polymersomes, two controls were setup. The first control was a non-relevant GPCR target dopamine receptor D2 (DRD2) that was expressed and biotin-affinity purified (DRD2 proteopolymersomes) in the same way as GLP-1R. This was a positive control showing successful expression in the presence of Fos14 and will be used as a negative control for the radioligand binding assay in the following characterization. The second control was GLP-1R expressed supplied with Fos14 but not AB polymersomes, followed by histidine-affinity purification and removal of Fos14. This was a control to testify the Fos14 mediated GLP-1R folding without polymersomes in radioligand binding assay. Both controls were examined by SDS-PAGE and

immunoblotting. Similar to GLP-1R, the biotin-affinity purified DRD2 proteopolymersomes were difficult to be identified on SDS-PAGE but the Western blot confirmed the successful biotin elution (Figure 6.6 A, B lane 4). The histidine-affinity purified GLP-1R migrated normally and displayed two major bands on both SDS-PAGE and Western blot (Figure 6.6 A, B lane 2, 3).

To further characterize if GFP-GLP-1R was stably inserted in the polymersomes, the existence of co-elution of polymersomes and the fluorescence detection size-exclusion chromatography (FSEC) profile of GFP-GLP-1R expressed in the presence of Fos14 was monitored.

The FSEC elution profiles of new preparations were changed. In the presence of Fos14, the larger structures at 6.5 mL no longer existed. Previously, we attributed its formation of such structure due to non-lamellar phase which didn't facilitate GLP-1R incorporation. The disappearance of large structures indicates that the additional chemical chaperone Fos14 not only assisted GLP-1R folding but also helped in self-assembly of AB vesicles into more monodisperse vesicles that favorably integrated target protein.

The FSEC of unpurified GFP-GLP-1R also showed the small vesicles were eluted from 8.0 to 9.5 mL and displayed three peaks at 8.0, 8.5, 9.5 mL, respectively. The second population (8.5 mL) was overlapping GFP profile and contained vesicles estimated to be slightly larger than 145 nm (Figure 6.9). The other two peaks possibly represented polymersomes of different sizes around 145 nm. The presence of GFP fluorescence in the FSEC after biotin-affinity purification indicates both AB polymersomes and target protein were co-purified (Figure 6.8 B). After purification, however, only the second (8.5 mL) and third (9.5 mL) populations were eluted and the protein association with second population was not altered.

Figure 6.8 Fluorescence detection size-exclusion chromatography (FSEC) of cell free expressed GFP-GLP-1R in the presence of Fos14. Cell free reaction mixture. Analysis of (A) cell free reaction mixture before purification and (B) biotin-affinity purified proteopolymersomes. Top annotations: composition of eluent at each volume. Vesicle size indicates the elution volume of liposomes with defined sizes, by Per Lundahl et al. ^[13]. Cartoons represent the composition of each fraction.

For functionality test, the radioligand saturation binding assay was performed. Binding was detected under conditions where the CFPS was supplemented with Fos14 and after biotin-affinity purification. This result indicates that the complex folding of GLP-1R was assisted by the lipid-like detergent and upon the removal of Fos14, such folding was preserved after co-translationally inserted into polymersomes. As positive control, the crude GLP-1R extracted from native tissues was used in radioligand binding assay as present in previous paragraph (Figure 6.5 A). As shown in Figure 6.10 A, binding of [I-125] GLP-1 to cell-free expressed with Fos14 and biotin eluted GLP-1R AB proteopolymersomes is both dose-dependent and saturable, with a detected binding affinity of 0.37 nM and the maximal binding ($B_{\max}$) of 1.5 pmol/mg. This result compares very favorably with GLP-1R membrane preparations where the K_d and $B_{\max}$ were determined experimentally to be 0.24 nM and 2.2 pmol/mg previously (Figure 6.9 A). The specific binding was further validated by performing a competition binding assay in which [I-125] GLP-1 was used as tracer and

competed by peptide ligand exendin-4 (Figure 6.9 C). The half effective concentration (EC_{50}) of GLP-1R AB proteopolymersomes was determined to be 54.3 nM, which is comparable to that of GLP-1R membrane preparation (37.8 nM).

Figure 6.9 Functional characterization of GLP-1R proteopolymersomes using radioligand binding assay. Specific binding of peptide ligand [I-125] GLP-1 to biotin-affinity purified (A) GLP-1R proteopolymersomes and (B) GLP-1R produced in the absence of AB polymersomes. Top: specific and fitted binding curves using 1:1 model. Bottom: raw binding curves. TB: total binding. NSB: non-specific binding. Dissociation constant (K_d) of GLP-1R proteopolymersomes (A) was calculated to be 0.37 nM, with a Bmax of binding of 1.5 pmol/mg. In competition binding assay, the half effective concentration (EC_{50}) was determined to be 54.3 and 37.8 nM for GLP-1R proteopolymersomes and membrane preparation, respectively.

We used two negative controls to ensure that this binding is specific to GLP-1R proteopolymersomes. In control experiment in which polymersomes were not provided to the Fos14-containing cell free system, dose-dependent binding was not observed, strongly indicating that stably insertion of functional GLP-1R into polymersomes was realized by the addition of Fos14 into the cell free system (Figure 6.9 B). Moreover, when a non-relevant GPCR target, DRD2 proteopolymersomes were purified in a similar manner and no binding was observed (Figure 6.10).

Figure 6.10 Functional characterization of non-relevant target dopamine receptor D2 (DRD2). Radioligand binding assay of [I-125] GLP-1 binding to biotin-affinity purified DRD2 proteopolymersomes produced in the presence of Fos14. Top: specific and fitted binding curves using 1:1 model. Bottom: raw binding curves. TB: total binding. NSB: non-specific binding. No specific binding was detected.

6.4 Conclusions

The produced GLP-1R was characterized with focus on orientation and functionality. The SPR analysis showed the presence of correctly orientated GLP-1R in the AB polymersomes by probing the N-terminal extracellular domain of GLP-1R. For future studies of AB polymer/GLP-1R association behavior, TEM can be used to characterize the orientation of protein inserted when combined with immunostaining with gold nanoparticle conjugated antibodies. The TEM analysis can provide information about both structural information of polymersomes and its association with GLP-1R proteins associated with membranes. The choice of antibody can be GLP-1R N-terminal specific antibody (used in SPR experiments) and 1D4 antibody (used in Western blot in previous chapters).

Production of functional GLP-1R proteopolymersomes required both native-like membrane support (AB polymersomes) and additional folding mechanisms, in this case, zwitterionic detergent Fos14 was added. The biotin-affinity purified GLP-1R proteopolymersomes produced in the presence of Fos14 displayed ligand binding capability comparable to native protein expressed in cells. Further characterizations using TEM and FSEC confirmed the vesicular structure of the product and protein/polymer association.

References

- [1] S. Runge, S. Schimmer, J. Oschmann, C. B. Schiødt, S. M. Knudsen, C. B. Jeppesen, K. Madsen, J. Lau, H. Thøgersen, R. Rudolph, *Biochemistry* **2007**, *46*, 5830-5840.
- [2] I. Navratilova, J. Sodroski, D. G. Myszka, *Analytical Biochemistry* **2005**, *339*, 271-281.
- [3] K. Schröder-Tittmann, E. Bosse-Doenecke, S. Reedtz-Runge, C. Ihling, A. Sinz, K. Tittmann, R. Rudolph, *Biochemistry* **2010**, *49*, 7956-7965.
- [4] D. G. Myszka, *Analytical Biochemistry* **2004**, *329*, 316-323.

- [5] S. May, M. Andreasson-Ochsner, Z. Fu, Y. X. Low, D. Tan, H.-P. M. de Hoog, S. Ritz, M. Nallani, E.-K. Sinner, *Angewandte Chemie International Edition* **2013**, 52, 749-753.
- [6] H. P. de Hoog, E. M. Lin JieRong, S. Banerjee, F. M. Decailot, M. Nallani, *PLoS One* **2014**, 9, e110847.
- [7] M. Nallani, M. Andreasson-Ochsner, C. W. Tan, E. K. Sinner, Y. Wisantoso, S. Geifman-Shochat, *Biointerphases* **2011**, 6.
- [8] D. G. Drescher, N. A. Ramakrishnan, M. J. Drescher, *Methods in molecular biology (Clifton, N.J.)* **2009**, 493, 323-343.
- [9] C. Spiess, Q. Zhai, P. J. Carter, *Molecular Immunology* **2015**, 67, 95-106.
- [10] K. Hollenstein, C. de Graaf, A. Bortolato, M.-W. Wang, F. H. Marshall, R. C. Stevens, *Trends in Pharmacological Sciences*, 35, 12-22.
- [11] C. Dong, C. M. Filipeanu, M. T. Duvernay, G. Wu, *Biochimica et biophysica acta* **2007**, 1768, 853-870.
- [12] J. Hou, H. Tang, Z. Liu, T. Österlund, J. Nielsen, D. Petranovic, *FEMS Yeast Research* **2014**, 14, 481-494.
- [13] P. Lundahl, C.-M. Zeng, C. Lagerquist Hägglund, I. Gottschalk, E. Greijer, *Journal of Chromatography B: Biomedical Sciences and Applications* **1999**, 722, 103-120.
- [14] K. Takai, T. Sawasaki, Y. Endo, *Nat. Protocols* **2010**, 5, 227-238.
- [15] E. Papp, P. Csermely, in *Molecular Chaperones in Health and Disease* (Eds.: K. Starke, M. Gaestel), Springer Berlin Heidelberg, Berlin, Heidelberg, **2006**, pp. 405-416.
- [16] in *Production of Recombinant Proteins*, Wiley-VCH Verlag GmbH & Co. KGaA, **2005**, pp. i-xxv.
- [17] H. Ren, D. Yu, B. Ge, B. Cook, Z. Xu, S. Zhang, *PLoS ONE* **2009**, 4, e4509.
- [18] B. L. Cook, D. Steuerwald, L. Kaiser, J. Graveland-Bikker, M. Vanberghem, A. P. Berke, K. Herlihy, H. Pick, H. Vogel, S. Zhang, *Proceedings of the National Academy of Sciences* **2009**, 106, 11925-11930.
- [19] S. Seyedmohammad, D. Born, H. Venter, *Protein Expression and Purification* **2014**, 101, 138-145.
- [20] X. Wang, K. Corin, C. Rich, S. Zhang, *Scientific Reports* **2011**, 1, 102.

Chapter 7

Functional Production of Proteopolymersomes

In the previous chapter, the protocol of functional production of GLP-1R proteopolymersomes was developed using AB polymersomes as membrane support and Fos14 as folding mechanism. The product display ligand binding capability, which compared favorably to native protein expressed in cells. Further characterizations confirmed the association state was not affected by the introduction of Fos14. These results suggest that polymersomes as membrane supports provide native-like environment for maintaining the correct folding of membrane protein GLP-1R. We foresee that this principle can be applied other classes of block copolymers.

This chapter demonstrated the versatility of the cell free expression system with polymersomes. A few important modifications were examined and compared, including the use of triblock copolymer, detergent, and reconstitution strategy for GLP-1R membrane insertion. Biotin- and histidine- affinity purifications were also compared in these experimental variants. Polymer-dependent functional production of GLP-1R was achieved with both biotin-affinity purification of co-translational incorporation and reconstitution of purified GLP-1R.

7.1 Introduction

As compared to conventional liposomes, the polymersomes are more stable and can be flexibly tuned by changing polymer chemistry. As reviewed recently by Joachim Habel et al. ^[1], a number of MPs have been studied, e.g. OmpF, AqpZ, in a variety of copolymers with different building blocks. Both AB and ABA of a number of formulas were tested for the incorporation of MPs. The previous results of GLP-1R expression, purification and functional characterization led us to take advantage of the diversity of block copolymers. In this chapter, we demonstrated that both diblock and triblock copolymers can be used as membrane support in the production of functional GLP-1R. Besides, different reconstitution methods were also attempted as an alternative approach to incorporate MPs into polymersomes. The produced proteopolymersomes with different polymers were characterized using radioligand binding assay. The results also indicated that the molecular differences between different classes of block copolymers are important in functional incorporation of GLP-1R.

7.2 Experimental Methodology

7.2.1 Protein Expression and Purification

Protein expression was performed using both coupled (transcription and translation in one single reaction) and bilayer reaction (separated extraction/mRNA template mixture and energy source and amino acids) formats. Purification was performed using filtration, biotin- and histidine- affinity methods. For details, refer to chapter 3.

7.2.2 Polymersomes Preparation

Diblock copolymer polybutadiene-polyethylene oxide (AB) and triblock copolymer Polymethyloxazoline-Polydimethylsiloxane-Polymethyloxazoline (ABA) was used for polymersomes preparation using film rehydration method. For fluorescence and biotin-

tagged vesicles, tetrazole or DSPE-PEG-biotin was added during film formation. Tetrazole fluorescence was measured using Tecan 200 plate reader. For details, refer to chapter 3.

7.2.3 Protein Gel Electrophoresis and Western Blot

Protein gel electrophoresis and Western blot were performed for detection of GLP-1R protein and purity examination. For details, refer to chapter 3.

7.2.4 Bicinchoninic Acid Assay (BCA)

BCA was performed to measure protein concentration of purified proteins. For details, refer to chapter 3.

7.2.5 Radioligand Binding Assay

Radioligand binding assay was performed to detect specific binding between GLP-1R and its ligand. For details, refer to chapter 3.

7.3 Principle Outcomes

7.3.1 Co-translational Incorporation of GLP-1R into Polymersomes

In the functionality assay, the production protocol was developed using wheat germ cell free system supplied AB polymersomes and zwitterionic detergent, Fos14. Produced GLP-1R displayed binding capability to peptide ligand GLP-1 comparably to native protein expressed in cells. In order to expand our toolbox, alternative production strategies by varying block copolymers and using detergent combinations were investigated. We introduced a triblock copolymer that has been extensively used as MP support ^[1], PMOXA₂₀-PDMS₅₄-PMOXA₂₀ (ABA) and another commonly used non-ionic detergent in MP research, digitonin in cell free expression. A few polymersome/detergent combinations were tested in GLP-1R expression, including digitonin+Fos14+AB (DF-AB), Fos14+AB

(F-AB) and Fos14+ABA (F-ABA) (Figure 7.1 A). All three combinations showed no obvious inhibition of protein translation as detected by Western blot. SDS-PAGE showed multiple bands with no distinguished band of target protein GLP-1R. In GLP-1R saturation binding assay, no binding was detected, probably because of the large amount of impurities in the samples.

Figure 7.1 Cell free protein synthesis of GLP-1R with diblock and triblock copolymers. (A) Western blot and SDS-PAGE analysis of GLP-1R expressed in the presence of detergent, digitonin (D) and/or Fos14 (F), and copolymers (AB: diblock, ABA: triblock). Lane 1: DF-AB. Lane 2: F-

AB. Lane 3: F-ABA. (B-D) Saturation binding of unpurified GLP-1R proteopolymersomes expressed in different combinations of detergent(s) and copolymers (A, lane 1-3), to radioligand [I-125] GLP-1. No specific binding was detected.

It has been determined in Chapter 4 and 5 that of all the three purification methods, i.e. vivaspin filtration, histidine- and biotin-affinity purification, only the biotin-affinity purified GLP-1R showed positive functionality without dissociating from polymer membrane support. Thus, the new preparations were purified using the same method. The SDS-PAGE showed all of the elutions under these conditions showed multiple bands without any distinct band of GLP-1R (Figure 7.2 A). This could probably because of Fos-14 interfering the biotin streptavidin purification.. However, in Western blot, both F-AB and F-ABA showed protein bands corresponding to GLP-1R, indicating the successful elution but addition of digitonin (DF-AB) has interfered with purification, possibly due to the suboptimal detergent removal in dialysis.

In the radioligand binding assay, purified GLP-1R proteopolymersomes expressed in the presence of F-AB and F-ABA showed positive binding with the dissociation constant (K_d) determined to be 1.7 and 2.1 nM, respectively (Figure 6.2 B, C). In terms of polymer building blocks, the choice of diblock or triblock copolymers did not affect the functionality of GLP-1R. Meanwhile, no non-specific binding of ABA polymer to GLP-1 or GLP-1 to filter plate used in the assay was detected, indicating the binding of GLP-1 to GLP-1R proteopolymersomes was specific (Figure 7.2 D, E).

Figure 7.2 Biotin-affinity purification and functional characterization of GLP-1R proteopolymersomes. (A) Western blot and SDS-PAGE analysis of biotin-affinity purified GLP-1R expressed in the presence different detergent/polymer combinations. Lane 1: DF-AB. Lane 2: F-AB. Lane 3: F-ABA. (B, C) Saturation binding of biotin-affinity purified GLP-1R AB (B) and ABA (C) proteopolymersomes (samples in lane 2-3) to [I-125] GLP-1. Dissociation constant (K_d) of purified GLP-1R proteopolymersomes expressed in the presence of F-AB and F-ABA was determined to be 1.7 and 2.1 nM, respectively. No specific binding of (D) ABA to GLP-1 and (E) GLP-1 to filter plate was detected.

7.3.2 Reconstitution of Purified GLP-1R into Polymersomes

As an alternative to co-translational incorporation, reconstitution of purified protein to vesicles, e.g. liposomes, nanodiscs, is widely used for constructing biomimetic membrane systems with MPs [2]. In the case of GLP-1R, we also attempted using CFPS with Fos14. Using this approach, polymersomes were not added to the cell free expression system. Consequently, the GLP-1R was histidine-affinity purified and reconstituted into polymersomes in the presence of Fos14. Meanwhile, the biotin-affinity purification was not possible as polymersomes were not present. SDS-PAGE and Western blot confirmed the successful purification of GLP-1R, which showed high purity comparable to previous histidine-affinity purifications (Figure 7.3 A, lane 1). In the radioligand binding assay, the dissociation constant (K_d) of purified GLP-1R in Fos14 was determined to be 0.8 nM, indicating the correct protein folding (Figure 7.3 B).

To incorporate the purified GLP-1R into polymersomes, they were mixed with either diblock copolymer AB or triblock copolymer ABA, followed by removal of Fos14 by dialysis for reconstitution. Different from solubilized proteins, the migration of incorporated GLP-1R in SDS-PAGE was changed. For reconstitution into diblock AB polymersomes, the SDS-PAGE showed smearing from 37 to about 100 kDa and only small amount of GLP-1R was detected in Western blot at about 98 kDa, indicating the reconstituted protein was not successfully transferred during blotting (Figure 7.3 A, lane 2). The reconstituted GLP-1R into triblock ABA polymersomes showed similar migration profile, where only a small amount of protein entered the gel during gel electrophoresis (Figure 7.3 A, lane 3). The GLP-1R reconstituted in ABA can only be detected in Western blot but not SDS-PAGE. Although the SDS-PAGE analysis showed different migration profile than previously, it was a good indication of protein incorporation and Western blot confirmed the presence of GLP-1R in reconstituted samples.

In radioligand binding assay, both reconstituted GLP-1R samples were tested. Different from co-translational incorporation into AB, the reconstitution of GLP-1R into AB lost the binding capability (Figure 7.3 C). In comparison, reconstituted GLP-1R in ABA showed

specific binding with K_d determined to be 0.67 nM and compared favorably to the binding affinity before reconstitution (0.8 nM) and native GLP-1R expressed in cells (0.24 nM), indicating the reconstitution did not compromise the functionality (Figure 6.3 D). The specificity of binding was further validated using competition binding mode with exendin-4, which also showed similar half effective concentration (EC_{50}) of 93.2 nM compared to native GLP-1R (37.8 nM) as shown in Figure 6.3 E.

Figure 7.3 (A) SDS-PAGE of D-CF expressed and histidine-affinity purified GLP-1R (lane 1), followed by reconstitution into AB (lane 2) or ABA (lane 3). Purified GLP-1R is boxed in red. Radioligand binding assay of (B) histidine-affinity purified GLP-1R solubilized in Fos14 before

reconstitution, and after reconstitution into (C) AB or (D) ABA. (E) Competition binding assays of reconstituted GLP-1R into ABA or GLP-1R derived from membrane preparation, with peptide ligand exendin-4. [I-125] GLP-1 (RL) concentration was indicated by nM.

To further validate the reconstitution approach and its applicability, the fluorescent GLP-1R construct, GFP-GLP-1R was tested (Figure 7.4). The GFP-GLP-1R together with a negative control, DRD2 were histidine-affinity purified and reconstituted into triblock copolymer ABA in the same way as GLP-1R. Both GFP-GLP-1R and DRD2 were detected as a monomer band with smearing in Western blot and SDS-PAGE (Figure 7.4 A, lane 1 and 2). The purified GFP-GLP-1R also showed a minor band at 30 kDa which may arise either from proteolysis of the fusion protein or from translation initiation at the methionine residue at the beginning of the GFP coding sequence ^[3]. Moreover, a substantial amount of incorporated proteins were stuck in the wells and did not enter the gel during gel electrophoresis, affecting protein detection and subsequent transfer in Western blot. After reconstitution, more proteins were lost during migration, indicating the successful incorporation of GFP-GLP-1R and DRD2 into ABA (Figure 7.4 A, lane 3 and 4). In Western blot, a very low amount of target proteins were detected as smearing. Binding was detected in Fos14 solubilized and ABA reconstituted GFP-GLP-1R, with K_d of 1.1 and 1.7 nM, respectively. The similar affinities indicate no functionality was lost during reconstitution. The K_d of GFP-GLP-1R determined was slightly lower than the GLP-1R construct (0.37 nM) produced in the same way, which possibly reflected the effect of fusion partner GFP on the GLP-1R ligand binding capability.

In co-translational incorporation of GLP-1R, the use of both diblock AB and triblock ABA copolymers managed to yield functional receptors whereas in the second route of reconstitution, only ABA resulted in functional GLP-1R. This suggested the different building blocks of copolymer membranes can affect MP incorporation and functionality.

Figure 7.4 (A) SDS-PAGE and (B) Western blot of D-CF expressed and histidine-affinity purified GFP-GLP-1R (lane 1) and DRD2 (lane 2), followed by reconstitution into ABA polymersomes. (A, B) Blue box: proteolysis of the fusion protein or alternative translation initiation at the methionine. Radioligand binding assay of GFP-GLP-1R before (C) and after (D) reconstitution into ABA. (E) Purified DRD2 after reconstitution. [I-125] GLP-1 (RL) concentration was in nM. The K_d was determined to be 1.1 nM (C) and 1.7 nM (D), respectively.

We proposed a theory that such difference was resulted from the superior flexibility of ABA polymer in compared to AB. For membrane protein incorporation, the functionality of proteins in the membranes is related to the physical environment. The membrane rigidity

is an important property and commonly used to describe the nature of the environment membranes provide to incorporated proteins because it controls how much flexibility is allowed for membrane proteins to move around. The rigidity has to be within a certain range so that membrane proteins can anchor stably in the bilayer via hydrophobic interaction and meanwhile, their folding is not altered by the interaction with the membrane. The hydrophobic domain of AB, poly-butadiene (PBD) is more rigid than ABA, which contains polydimethylsiloxane (PDMS). The flexibility is also reflected the low glass transition temperature of PDMS (-120 °C) compared to PBD (about -21 °C). In MP reconstitution, the more flexible PDMS better accommodated GLP-1R and maintained its correct folding when Fos14 was removed. For PBD, the high rigidity did not favorably interact with the transmembrane domains of GLP-1R and caused mis-folding or orientation during reconstitution. In other words, the functional incorporation of Fos14 solubilized GLP-1R depends on the fluidity of polymer membranes, which is correlated to the stiffness of hydrophobic layer formed. This was supported by the diffusion studies of polymeric vesicles made from diblock and triblock copolymers. In the fluid mosaic model proposed by Singer and Nicholson ^[4], the membrane constituents can diffuse freely in a two dimensional membrane surface. The diffusion coefficient D ($\mu\text{m}^2/\text{s}$) is used to describe membrane fluidity or integral membrane protein mobility, and can be measured using techniques based on the light microscope ^[5] such as fluorescence correlation spectroscopy ^[6]. The measured diffusion coefficient D of ABA with similar formula ($\text{A}_{12}\text{B}_{63}\text{A}_{12}$) was reported to be $1.41 \pm 0.04 \mu\text{m}^2/\text{s}$ ^[7] whereas the D of AB was estimated to be $0.16 \mu\text{m}^2/\text{s}$ ^[8]. The difference is significant considering the diffusion coefficient of liposomes is $12.5 \pm 0.5 \mu\text{m}^2/\text{s}$ measured using the same method. Although the polymer reconstitution, protein folding and their interactions are very likely to be much more complicated, the polymer membrane rigidity is likely to be a factor influencing functional incorporation of GLP-1R.

In the co-translational incorporation method, both AB and ABA showed successful incorporation of functional GLP-1R. Such results were different from reconstitution of purified GLP-1R where only ABA adopted correctly folded proteins but not AB. These indicate there were other factors influencing GLP-1R functionality besides the effect of polymer chain flexibility. This is most likely due to the folding mechanism in WG extract

that was used in the CFPS. Although it is known that the ER, which is known to play a critical role in MP folding, trafficking after translation, was removed during WG extract preparation, we anticipate the remaining folding machines were able to assist the modification in CFPS.

7.4 Conclusions

In the previous chapter, it was demonstrated that both diblock AB and triblock ABA copolymers could be used as native-like membrane support in functional production of GPCR protein GLP-1R in CFPS. In this chapter, the applicability of this principle was validated using different production strategies, co-translational incorporation and reconstitution.

In the pursuit of functional production of GLP-1R proteopolymersomes, the successful cases are using AB copolymer with co-translational incorporation, and ABA copolymer with reconstitution method. The other attempts with AB/reconstitution and ABA/co-translational insertion proved to be unsuccessful. These results show that the functional incorporation of GLP-1R depends on two crucial factors, namely the type of copolymers and method of incorporation. We correlate the results with the incorporation mechanisms and concluded with the following model: a) in co-translational insertion of GLP-1R, the concentration of free FC14 (detergent that mediated protein folding) is low due to the high concentration of proteins in the extract. In case AB copolymer was used, the thin membrane (est. 5.4 nm) is a dominant factor rather than fluidity and allows GLP-1R insertion with the aid of translational machinery of wheat germ extract (e.g. ribosome) instead of detergent which is the driving force. In case of ABA copolymer with reconstitution, the system contains purified GLP-1R protein, polymers and detergent, thus the concentration of free FC14 is high and it does not contain any translational machinery. The high concentration detergent led to the destabilization of ABA membranes and formed mixture of polymer/protein/detergent and the GLP-1R incorporation was further triggered by removal of detergent and formed proteopolymersomes. In this case, the ABA membrane is fluid enough for detergent destabilization. For the unsuccessful cases where AB copolymer was

combined with reconstitution method, the low fluidity resisted detergent penetration and becomes a limiting factor. When ABA copolymer was used with co-translational incorporation, the thick membrane excluded the insertion from the machinery of wheat germ extract. We also proposed that the thickness of copolymer membranes play a role in the incorporation and further investigations are required to elucidate its effects on functional incorporation, especially when copolymers with different molecular configurations are taken into consideration.

According to the results so far, ABA is a better candidate compared to AB because the incorporated GLP-1R is of higher purity using reconstitution method compared to GLP-1R in AB using co-translational insertion method. It requires further study of their in vivo performance such as capability to generate therapeutic antibodies.

Polymer membranes are important tools that provide native-like environment for membrane proteins, especially considering the diversity of unexplored polymer building blocks. It is anticipated that the same principle can be applied to essentially all types of amphiphilic polymers and in return, acts as an enabling methodology in MP research.

References

- [1] M. Perry, S. Madsen, T. Jørgensen, S. Braekevelt, K. Lauritzen, C. Hélix-Nielsen, *Membranes* **2015**, *5*, 685.
- [2] J.-L. Rigaud, D. Lévy, in *Methods in Enzymology, Vol. Volume 372*, Academic Press, **2003**, pp. 65-86.
- [3] T. Kawate, E. Gouaux, *Structure* **2006**, *14*, 673-681.
- [4] S. J. Singer, G. L. Nicolson, *Science* **1972**, *175*, 720-731.
- [5] Y. Chen, B. C. Lagerholm, B. Yang, K. Jacobson, *Methods* **2006**, *39*, 147-153.
- [6] S. Ramadurai, A. Holt, V. Krasnikov, G. van den Bogaart, J. A. Killian, B. Poolman, *Journal of the American Chemical Society* **2009**, *131*, 12650-12656.
- [7] F. Itel, M. Chami, A. Najer, S. Lörcher, D. Wu, I. A. Dinu, W. Meier, *Macromolecules* **2014**, *47*, 7588-7596.

- [8] J. C. M. Lee, M. Santore, F. S. Bates, D. E. Discher, *Macromolecules* **2002**, *35*, 323-326.

Chapter 8

Discussion and Future Work

In the previous chapters, the principle of polymer-based membrane system was presented with validation by establishing step-by-step workflow to produce and characterize the proteopolymersomes. This chapter includes the final conclusions and future outlook of this thesis. It starts with a brief summary of this thesis, including the major outcomes and their impact in polymer biomimetic systems and membrane protein research. Finally, the future directions of these research projects were proposed.

8.1 Discussion

As described earlier, membrane proteins (MPs) are essential components of biological membranes that plays crucial role for cellular transport ^[1]. The development of membrane mimetic materials can help realize applications in relation to drug discovery and therapies. We approached a major challenge in MP research by taking advantage of block copolymers self-assembling property to stabilize MPs in a native-like membrane environment. Specifically, we developed a polymersome-based artificial cell membrane system and demonstrated its broad application to a Class B G-protein coupled receptor (GPCR), the glucagon like peptide-1 receptor (GLP-1R). Biophysical and biochemical characterizations regarding the structure and function of protein/polymer complex (proteopolymersomes) were carried out. In the construction of the polymersome-based membrane system, we established a basic production protocol (chapter 4), biophysical and biochemical characterization workflow (chapter 5 and 6), and setup an improved protein synthesis platform (chapter 6 and 7).

Chapter 4 described efforts to express and purify GLP-1R expressed in cell free protein synthesis system (CFPS) supplemented with diblock copolymer (AB). A novel purification method using affinity tags on the polymer membrane surface was used and the purified material contained GLP-1R as well as polymer membrane. Chapter 5 and 6 describe a series of biophysical characterization on the purified material. Electron microscopy and size-exclusion chromatography (SEC) confirmed the presence of nano-scaled vesicular polymer membrane structure (polymersomes), which shown to also encapsulate small molecule dye, in support of a fully functional membrane support. In parallel, incorporation of GLP-1R was evident from two different types of SEC performed. The characterizations in chapter 4 and 5 constitute the first part of the workflow of membrane systems with focus on membrane biophysics. Subsequently, biochemical characterizations, including different binding assays to probe functional GLP-1R epitopes, were described, an important aspect to assess established expression and purification protocol. GLP-1R epitopes was successfully probed by an antibody specific to the extracellular NTD, but not the widely recognized radioligand binding, which prompted further optimization of the protocol. An

addition of a chemical chaperone, Fos14, which assist in the correct folding of GLP-1R, in the protocol was shown to produce fully functional epitopes, as assessed by the radioligand assay. A methodology for producing and characterizing functional GLP-1R proteopolymersomes has been now established. Using the established system, in chapter 7, we further tested its versatility by comparing the effects of polymer vesicles' molecular diversity, detergent types and MP insertion methods, using GLP-1R as a model. GLP-1R proteopolymersomes were produced with two different types of polymer building blocks and both supported functional protein expression, thus demonstrating the robustness of polymer membrane systems.

Taken together, the research presented in this thesis have (1) advanced the application of polymer membrane systems in amphiphile assisted membrane protein production in cell free protein synthesis, (2) standardized the workflow of protein/polymer membrane complex production and characterization, broadly applicable to other complex membrane proteins and may serve to reduce bottlenecks for the discovery of novel ligands and antibodies to such targets. (3) bridged the gap between membrane biology and polymer chemistry. In a broader sense, the work presented here has improved the understanding of polymer membrane systems in the context complex membrane protein synthesis and assembly.

8.2 Outlook

The great potential of polymersomes-based artificial system led us to further studies in two major areas. First, the established production protocol was extensively applied to another family of MPs, an ion channel TREK-2. The retention of TREK-2 after size-based filtration indicates successful incorporation and the biotin-affinity purified product showed high purity comparable to that of GLP-1R (Figure 8.1).

Figure 8.1 Production of ion channel TREK-2 using established protocol with polymer membrane support. (A) Western blot analysis of purified TREK-2 using size-based filtration method. The retention of TREK-2 indicates its incorporation into AB polymersomes. Lane 1: soluble fraction of reaction mixture. Lane 2: filtration retentate. (B) Western blot and SDS-PAGE analysis of biotin-affinity purification of TREK-2. SDS-PAGE gel was stained with Coomassie blue. Western blot with a specific antibody against the C9 tag at C-terminal of TREK-2. M: molecular weight marker. Arrows indicate TREK-2.

A key advantage of the polymer vesicles is they are generally more stable as compared to conventional vesicular nanocarriers, e.g. liposomes. Therefore, they were selected as native-like support membrane in this thesis. As functional GLP-1R was produced in polymersomes, it is still an open question as to the stability of the proteopolymersomes. The question can be addressed from different perspectives, e.g. short term/long term stability under *in vitro/in vivo* conditions. The *in vivo* stability is especially relevant in this case because in drug development, current methods of therapeutic antibody generation is limited by the availability of natively folded MP antigen like GPCRs. In type I diabetes, for example, the treatment with T-cell suppressing antibody anti-CD3 has shown encouraging outcome in phase II trials ^[2]. However, for type II diabetes the process has

been considerably delayed. Thus, the functional GLP-1R proteopolymersomes, a type II diabetes target, are highly valuable in antibody discovery and therapeutics.

In the meantime, we expect the stability measurement to be technically challenging. Unlike established targets like rhodopsin, for which the stability can be easily measured by changes in absorption spectrum^[3], reliable GPCR protein stability could only be monitored using radioligand binding assay. Alternative assays are highly desirable to avoid the use of radioisotopes such as fluorescence-based assays like Förster resonance energy transfer (FRET). In the future, we aim to optimize current production protocol by increasing yield for non-radioisotope-based assays.

As part of the current characterization workflow, the fluorescence-detection size-exclusion chromatography (FSEC) also needs further investigation. An interesting observation in the FSEC was the presence of large non-lamellar structures. These large structures were believed to be consisted of non-lamellar structures, which are undesirable for protein incorporation. In the future, we propose to systematically investigate the formation of these large structures using various labeling methods, including fluorescence and radioisotopes, in combination with biophysical techniques, such as DLS and TEM. Small lamellar vesicles with incorporated protein may be compared to elucidate the effect of different molecular interactions between polymer vesicle and protein on protein assembly and functionality.

From the radioligand binding experiments, it clearly shows that the polymersomes are indeed stabilizing the receptors. However, further extensive characterization would reveal the extent of stability provided by polymersomes for receptors. For studying this property, thermal shift assay has to be established, in which, one could compare the thermal stability of MPs between polymersomes with other membrane supports, e.g., liposomes and nanodiscs. Based on that, polymers can be tuned to increase the stability without perturbing the biological function of MPs.

References

- [1] A. Engel, H. E. Gaub, *Annual Review of Biochemistry* **2008**, 77, 127-148.
- [2] C. Wilson, *Nat Rev Endocrinol* **2010**, 6, 242-242.
- [3] J.-L. Popot, T. Althoff, D. Bagnard, J.-L. Banères, P. Bazzacco, E. Billon-Denis, L. J. Catoire, P. Champeil, D. Charvolin, M. J. Cocco, G. Crémel, T. Dahmane, L. M. d. l. Maza, C. Ebel, F. Gabel, F. Giusti, Y. Gohon, E. Goormaghtigh, E. Guittet, J. H. Kleinschmidt, W. Kühlbrandt, C. L. Bon, K. L. Martinez, M. Picard, B. Pucci, J. N. Sachs, C. Tribet, C. v. Heijenoort, F. Wien, F. Zito, M. Zoonens, *Annual Review of Biophysics* **2011**, 40, 379-408.

APPENDIX

Appendix I: DNA sequence of pEU-MCS

The pEU vector with multiple cloning sites (pEU-MCS) was used for wheat germ cell free expression of GLP-1R, GFP-GLP-1R, and TREK-2 with bilayer format. The vector is 3747 bp containing an SP6 promoter, a multiple cloning sites and an anti-ampicillin sequence.

```

ATTTAGGTGACACTATAGAACTCACCTATCTCCCCAACACCTAATAACATTCA
ATCACTCTTTCCACTAACCACCTATCTACATCACCAAGATATCACTAGTTCTC
GAGCTCGGTACCTGTCCGCGGTCTGCGACGTACGCGGGCGGCCGCCATAAATT
GGATCCATATATAGGGCCCCGGGTATAATTACCTCAGGTCTGACGTCCCATGGT
TTTGTATAGAATTTACGGCTAGCGCCGGATGCGACGCCGGTCTGCGTCTTATCC
GGCCTTCCTATATCAGGCGGTGTTTAAGACGCCGCCGCTTCGCCCAAATCCTT
ATGCCGGTTCGACGACTGGACAAAATACTGTTTATCTTCCCAGCGCAGGCAG
GTTAATGTACCACCCCAGCAGCAGCCGGTATCCAGCGCGTATATACCTTCCG
GCGTACCTTTGCCCTCCAGCGATGCCCAGTGACCAAAGGCGATGCTGTATTCT
TCAGCGACAGGGCCAGGAATCGCAAACACGGTTTCAGTGGGGCAGGGGCC
TCTTCCGGCGATTCTTACTAGCTAGTATGCATAGGTGCTGAAATATAAAGTTT
GTGTTTCTAAAACACACGTGGTACGTACGATAACGTACAGTGTTTTTCCCTCC
ACTTAAATCGAAGGGTAGTGTCTTGGAGCGCGCGGAGTAAACATATATGGTT
CATATATGTCCGTAGGCACGTAAAAAAGCGAGGGATTCGAATTCCTCCCGGA
ACCTCCGGTTGGGGCCACGCCTCGATCGAGCAAAAAAAAAAAAAAAAAAGAAA
AAAAAAAAAAAAAAAAAGCTTTCCTCGCGGCCAGCTTGGCGTAATCATGGTCAT
AGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGA
GCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCA
CATTAAATTGCGTTGCGCTCACTGCCCCGCTTTCAGTCGGGAAACCTGTCGTGC
CAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTG
GGCGCTCTTCCGCTTCCTCACTCACTGACTCGCTGCGCTCGGTCGCTCGGCTG
CGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAAT

```

CAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCA
GGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCT
GACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACA
GGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC
TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAA
GCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTC
GTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCT
GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTA
TCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAG
GCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAG
AACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGA
GTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT
TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCT
TTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGG
GATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATT
AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGA
CAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTT
GTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAG
GGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCAC
CGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCA
GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGG
GAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCAT
TGCTACAGGCATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTCAGCT
CCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAA
AGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAG
TGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCA
TCCGTAAGATGCTTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGA
ATAGCGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAAT
ACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTT
CGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTA
ACCCACTCGTGCACCCAAGTATCTTCAGCATCTTTTACTTTCACCAGCGTTT

CTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAATAAGGG
CGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGC
ATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAA
AAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGAC
GTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCA
CGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACAC
ATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCA
GACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCT
TAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATTTCGACGCTC
TCCCTTATGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCG
TTGAGCACCGCCGCCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCAAC
AGTCCCCCGGCCACGGGGCCTGCCACCATACCACGCCGAAACAAGCGCTCA
TGAGCCCGAAGTGGCGAGCCCGATCTTCCCCATCGGTGATGTCGGCGATATA
GGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTCCG
GCGTAGAGGATCTGGCTAGCGATGACCCTGCTGATTGGTTCGCTGACCATTTC
CGGGTGCGGGACGGCGTTACCAGAACTCAGAAGGTTTCGTCCAACCAAACCG
ACTCTGGCGGCAGTTTACGAGAGAGATGATAGGGTCTGCTTCAGTAAGCCAG
ATGCTACACAATTAGGCTTGTACATACTGTCGTTAGAACGCGGCTACAATTAA
TACATAACCTTATGTATCATAACATACG

Appendix II: DNA sequence of G1LP-1R-pEU

The construct was used for wheat germ cell free expression of GLP-1R with bilayer format. The plasmid contains pEU vector and GLP-1R gene, with 6x Histidine tag and C9 (1D4) tag. The full length GLP-1R is 1389 bp encoding for 463 amino acids.

```

ATTAGGTGACACTATAGAACTCACCTATCTCCCCAACACCTAATAACATTCA
ATCACTCTTTCCACTAACCACCTATCTACATCACCAAGATATCACTAGTTCTC
GAGATGGCCGGCGCCCCCGGCCGCTGCGCCTTGCGCTGCTGCTGCTCGGGA
TGGTGGGCAGGGCCGGCCCCCGCCCCAGGGTGCCACTGTGTCCCTCTGGGA
GACGGTGCAGAAATGGCGAGAATACCGACGCCAGTGCCAGCGCTCCCTGACT
GAGGATCCACCTCCTGCCACAGACTTGTTCTGCAACCGGACCTTCGATGAAT
ACGCCTGCTGGCCAGATGGGGAGCCAGGCTCGTTCGTGAATGTCAGCTGCCC
CTGGTACCTGCCCTGGGCCAGCAGTGTGCCGCAGGGCCACGTGTACCGGTTC
TGCACAGCTGAAGGCCTCTGGCTGCAGAAGGACAACTCCAGCCTGCCCTGGA
GGGACTTGTCGGAGTGCGAGGAGTCCAAGCGAGGGGAGAGAAGCTCCCCGG
AGGAGCAGCTCCTGTTCTCTACATCATCTACACGGTGGGCTACGCACTCTCC
TTCTCTGCTCTGGTTATCGCCTCTGCGATCCTCCTCGGCTTCAGACACCTGCAC
TGCACCAGGAACTACATCCACCTGAACCTGTTTGCATCCTTCATCCTGCGAGC
ATTGTCCGTCTTCATCAAGGACGCAGCCCTGAAGTGGATGTATAGCACAGCC
GCCCAGCAGCACCAGTGGGATGGGCTCCTCTCCTACCAGGACTCTCTGAGCT
GCCGCCTGGTGTTTCTGCTCATGCAGTACTGTGTGGCGGCCAATTACTACTGG
CTCTTGGTGGAGGGCGTGTACCTGTACACACTGCTGGCCTTCTCGGTCTTCTC
TGAGCAATGGATCTTCAGGCTCTACGTGAGCATAGGCTGGGGTGTTCCCCTGC
TGTTTGTTGTCCCCTGGGGCATTGTCAAGTACCTCTATGAGGACGAGGGCTGC
TGGACCAGGAACTCCAACATGAACTACTGGGCTCATTATCCGGCTGCCCATTCT
CTTTGGCATTGGGGTGAACCTTCCTCATCTTTGTTTCGGGTCATCTGCATCGTGGT
ATCCAAACTGAAGGCCAATCTCATGTGCAAGACAGACATCAAATGCAGACTT
GCCAAGTCCACGCTGACACTCATCCCCCTGCTGGGGACTCATGAGGTCATCTT
TGCTTTTGATGATGGACGAGCACGCCCGGGGGACCCTGCGCTTCATCAAGCTG
TTTACAGAGCTCTCCTTCACCTCCTTCCAGGGGCTGATGGTGGCCATATTATA

```

CTGCTTTGTCAACAATGAGGTCCAGCTGGAATTTTCGGAAGAGCTGGGAGCGC
TGGCGGCTTGAGCACTTGCACATCCAGAGGGACAGCAGCATGAAGCCCCTCA
AGTGTCCCACCAGCAGCCTGAGCAGTGGAGCCACGGCGGGCAGCAGCATGT
ACACAGCCACTTGCCAGGCCTCCTGCAGCACAGAGACCAGCCAAGTGGCGCC
TGCCTGAGCGGCGCCATAAATTGGATCCATATATAGGGCCCCGGGTTATAAT
TACCTCAGGTCGACGTCCCATGGTTTTGTATAGAATTTACGGCTAGCGCCGGA
TGCGACGCCGGTCGCGTCTTATCCGGCCTTCCTATATCAGGCGGTGTTTAAGA
CGCCGCGCTTCGCCCAAATCCTTATGCCGGTTCGACGACTGGACAAAATAC
TGTTTATCTTCCCAGCGCAGGCAGGTAAATGTACCACCCCAGCAGCAGCCGG
TATCCAGCGCGTATATACCTTCCGGCGTACCTTTGCCCTCCAGCGATGCCAG
TGACCAAAGGCGATGCTGTATTCTTCAGCGACAGGGCCAGGAATCGCAAACC
ACGGTTTCAGTGGGGCAGGGGCCTCTTCCGGCGATTCTTACTAGCTAGTATGC
ATAGGTGCTGAAATATAAAGTTTGTGTTTCTAAAACACACGTGGTACGTACG
ATAACGTACAGTGTTTTTCCCTCCACTTAAATCGAAGGGTAGTGTCTTGGAGC
GCGCGGAGTAAACATATATGGTTCATATATGTCCGTAGGCACGTAAAAAAG
CGAGGGATTCTGAATTCCCCCGGAACCCCCGGTTGGGGCCCACGCCTCGATCG
AGCAAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAGCTTTCCCGCGG
CCAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCG
CTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGG
GTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCT
TTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCG
CGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCACTCACTGAC
TCGCTGCGCTCGGTGCGCTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGG
CGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGT
GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGG
CGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCA
AGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAACAGGCGTTTCCCC
CTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATAC
CTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTG
TAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACG
AACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAG

TCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA
GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTG
GCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGA
AGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAAC
CACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAA
AAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAG
TGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGA
TCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGT
ATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCAC
CTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCG
TGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAAT
GATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAG
CCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCA
TCCAGTCTATTAATTGTTGCCGGAAGCTAGAGTAAGTAGTTCGCCAGTTAAT
AGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTC
GTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACAT
GATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTT
GTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGC
ATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAG
TACTCAACCAAGTCATTCTGAGAATAGCGTATGCGGCGACCGAGTTGCTCTTG
CCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTG
CTCATCATTGGAAAACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCT
GTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCAT
CTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGC
CGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTC
CTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATA
CATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTT
CCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAA
CCTATAAAAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGAT
GACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTC
TGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTG

TTGGCGGGTGTCTGGGGCTGGCTTAAGTATGCGGCATCAGAGCAGATTGTACT
GAGAGTGCACCATTCGACGCTCTCCCTTATGCGACTCCTGCATTAGGAAGCA
GCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCGCGCAAGGAATGGTGCA
TGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCCTGCCACCATAC
CCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCC
ATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTG
ATGCCGGCCACGATGCGTCCGGCGTAGAGGATCTGGCTAGCGATGACCCTGC
TGATTGGTTCGCTGACCATTTCCGGGTGCGGGACGGCGTTACCAGAACTCA
GAAGGTTCGTCCAACCAAACCGACTCTGGCGGCAGTTTACGAGAGAGATGAT
AGGGTCTGCTTCAGTAAGCCAGATGCTACACAATTAGGCTTGACATACTGTC
GTTAGAACGCGGCTACAATTAATACATAACCTTATGTATCATACACATACG

Appendix III: DNA sequence of GFP-GLP-1R

The construct was used for wheat germ cell free expression of GFP-GLP-1R with bilayer format. The plasmid contains pEU vector and GFP-GLP-1R gene, with 6x Histidine tag and C9 (1D4) tag.

GATATCATGCATCATCATCATCATCATCATCATATGGTGAGCAAGGG
CGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGAC
GTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCT
ACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCGTGCC
CTGGCCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCT
ACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCCGAAGG
CTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACC
CGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGA
AGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGT
ACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACG
GCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCA
GCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTG
CTGCCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCA
ACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGAT
CACTCTCGGCATGGACGAGCTGTACAAGCTCGAGATGGCCGGCGCCCCCGGC
CCGCTGCGCCTTGCGCTGCTGCTGCTCGGGATGGTGGGCAGGGCCGGCCCCC
GCCCCCAGGGTGCCACTGTGTCCCTCTGGGAGACGGTGCAGAAATGGCGAGA
ATACCGACGCCAGTGCCAGCGCTCCCTGACTGAGGATCCACCTCCTGCCACA
GACTTGTTCTGCAACCGGACCTTCGATGAATACGCCTGCTGGCCAGATGGGG
AGCCAGGCTCGTTCGTGAATGTCAGCTGCCCCCTGGTACCTGCCCTGGGCCAG
CAGTGTGCCGCAGGGCCACGTGTACCGGTTCTGCACAGCTGAAGGCCTCTGG
CTGCAGAAGGACAACTCCAGCCTGCCCTGGAGGGACTTGTCGGAGTGCGAGG
AGTCCAAGCGAGGGGAGAGAAGCTCCCCGGAGGAGCAGCTCCTGTTCTCTA
CATCATCTACACGGTGGGCTACGCACTCTCCTTCTCTGCTCTGGTTATCGCCT
CTGCGATCCTCCTCGGCTTCAGACACCTGCACTGCACCAGGAAGTACATCCAC

CTGAACCTGTTTGCATCCTTCATCCTGCGAGCATTGTCCGTCTTCATCAAGGA
CGCAGCCCCTGAAGTGGATGTATAGCACAGCCGCCAGCAGCACCAGTGGGAT
GGGCTCCTCTCCTACCAGGACTCTCTGAGCTGCCGCCTGGTGTTTCTGCTCAT
GCAGTACTGTGTGGCGGCCAATTACTACTGGCTCTTGGTGGAGGGCGTGTAC
CTGTACACACTGCTGGCCTTCTCGGTCTTCTCTGAGCAATGGATCTTCAGGCT
CTACGTGAGCATAGGCTGGGGTGTTCCCCTGCTGTTTGTTGTCCCCTGGGGCA
TTGTCAAGTACCTCTATGAGGACGAGGGCTGCTGGACCAGGAACTCCAACAT
GAACTACTGGCTCATTATCCGGCTGCCCATCTCTTTGGCATTGGGGTGA
TCCTCATCTTTGTTTCGGGTCATCTGCATCGTGGTATCCAACTGAAGGCCAAT
CTCATGTGCAAGACAGACATCAAATGCAGACTTGCCAAGTCCACGCTGACAC
TCATCCCCCTGCTGGGGACTCATGAGGTCATCTTTGCCTTTGTGATGGACGAG
CACGCCCCGGGGGACCCTGCGCTTCATCAAGCTGTTTACAGAGCTCTCCTTCAC
CTCCTTCCAGGGGCTGATGGTGGCCATATTATACTGCTTTGTCAACAATGAGG
TCCAGCTGGAATTTTCGGAAGAGCTGGGAGCGCTGGCGGCTTGAGCACTTGCA
CATCCAGAGGGACAGCAGCATGAAGCCCCTCAAGTGTCCCACCAGCAGCCTG
AGCAGTGGAGCCACGGCGGGCAGCAGCATGTACACAGCCACTTGCCAGGCCT
CCTGCAGCACAGAGACCAGCCAAGTGGCGCCTGCCTGAGCGGCCGC

Appendix IV: DNA sequence of pCMVTNT-MCS

The pCMVTNT vector with multiple cloning sites (pCMVTNT-MCS) was used for cell free expression of GLP-1R with coupled reaction format as expression screening. The vector is 4050 bp containing both SP6 and T7 promoters, a multiple cloning sites and an anti-ampicillin sequence.

TCAATATTGGCCATTAGCCATATTATTCATTGGTTATATAGCATAAATCAATA
 TTGGCTATTGGCCATTGCATACGTTGTATCTATATCATAATATGTACATTTATA
 TTGGCTCATGTCCAATATGACCGCCATGTTGGCATTGATTATTGACTAGTTAT
 TAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCG
 CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCC
 CGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGA
 CTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCA
 GTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGG
 TAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTA
 CTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTT
 TGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAA
 GTCTCCACCCCATTGACGTCAATGGGAGTTTGTGTTTGGCACCAAAATCAACGG
 GACTTTCCAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTA
 GGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCA
 GATCACTAGAAGCTTTATTGCGGTAGTTTATCACAGTTAAATTGCTAACGCAG
 TCAGTGCTTCTGACACAACAGTCTCGAACTTAAGCTGCAGAAGTTGGTCGTG
 AGGCACTGGGCAGGTAAGTATCAAGGTTACAAGACAGGTTTAAGGAGACCA
 ATAGAACTGGGCTTGTCGAGACAGAGAAGACTCTTGCGTTTCTGATAGGCA
 CCTATTGGTCTTACTGACATCCACTTTGCCTTTCTCTCCACAGGTGTCCACTCC
 CAGTTCAATTACAGCTCTTAAGGCTAGAGTACTTAATACGACTCACTATAGGC
 TAGCATTTAGGTGACACTATAGAATACAAGCTACTTGTTCTTTTTGCACTCGA
 GAATTCACGCGTGGTACCTCTAGAGTCGACCCGGGCGGCCGCTTCGAGCAGA
 CATGATAAGATACATTGATGAGTTTGGACAAACCACAAGTGAATGCAGTGA
 AAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATT

ATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCA
GGTTCAGGGGGAGATGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAA
TGTGGTAAAATCGATAAGGATCCGGGCTGGCGTAATAGCGAAGAGGCCCCGA
CCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCC
TGTAGCGGCGCATTAAGCGCGGGCGGGTGTGGTGGTTACGCGCAGCGTGACCG
CTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTC
TCGCCACGTTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTA
GGGTTCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAACTTGATTAGGG
TGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGA
CGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACA
CTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTTCG
GCCTATTGGTTAAAAAATGAGCTGATTTAACA AAAAATTTAACGCGAATTTTA
ACAAAATATTAACGCTTACAATTTCTGATGCGGTATTTTCTCCTTACGCATC
TGTGCGGTATTTACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGAT
GCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCT
GACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCC
GGGAGCTGCATGTGTCAGAGGTTTTTCACCGTCATCACCGAAACGCGCGAGAC
GAAAGGGCCTCGTGATACGCCTATTTTTATAGGTAAATGTCATGATAATAATG
GTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTAT
TTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAAC
CCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACA
TTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCTTCCTGTTTTTGCT
CACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCAC
GAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTT
TCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTG
GCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTGCGCGCAT
ACACTATTCTCAGAATGACTTGTTGAGTACTACCAGTCACAGAAAAGCAT
CTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGA
GTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGA
GCTAACCGCTTTTTTGACACAACATGGGGGATCATGTAACCTCGCCTTGATCGTT
GGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGA

TGCCTGTAGCAATGGCAACAACGTTGCGCAAACCTATTAACCTGGCGAACTACT
TACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTT
GCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAA
ATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCA
GATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAA
CTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAA
GCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAA
AACTTCATTTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTTGATAATCTC
ATGACCAAAATCCCTTAACGTGAGTTTTTCGTTCCACTGAGCGTCAGACCCCGT
AGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCT
GCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCA
AGAGCTACCAACTCTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATA
CCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTC
TGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTG
CCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACC
GGATAAGGCGCAGCGGTCTGGGCTGAACGGGGGGTTCGTGCACACAGCCCAG
CTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGA
GAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGC
GGCAGGGTTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCC
TGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTT
TTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGG
CCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGGCTCGACAGA
TCT

Appendix V: DNA sequence of TREK-2

The ion channel TREK-2 full length DNA sequence is 1614 bp. The TREK-2 gene in pEU vector was used for expression using wheat germ cell free system with bilayer format. Both 10x histidine-tag and 1D4 tag were added for purification and detection.

CGAATTCCTCGAGATGGAGGATGGATTTAAGGGGGACAGGACTGAAGGCTGT
 CGCAGTGATTTCAGTGGCCGTTCCCGCAGCAGCACCGGTGTGCCAGCCCAAGA
 GCGCCACTAACGGGCAACCCCCGGCTCCGGCTCCGACTCCAACTCCGCGCCT
 GTCCATTTCTCTCCCGAGCCACAGTGGTAGCCAGGATGGAAGGCACCTCCCAA
 GGGGGCTTGCAGACCGTCATGAAGTGGAAGACGGTGGTTGCCATCTTTGTGG
 TTGTGGTGGTCTACCTTGTCACTGGCGGTCTTGTCTTCCGGGCATTGGAGCAG
 CCCTTTGAGAGCAGCCAGAAGAATACCATCGCCTTGGAGAAGGCGGAATTCC
 TGCGGGATCATGTCTGTGTGAGCCCCCAGGAGCTGGAGACGTTGATCCAGCA
 TGCTCTTGATGCTGACAATGCGGGAGTCAGTCCAATAGGAAACTCTTCCAAC
 AACAGCAGCCACTGGGACCTCGGCAGTGCCTTTTTCTTTGCTGGAAGTGTGAT
 TACGACCATAGGGTATGGGAATATTGCTCCGAGCACTGAAGGAGGCAAAATC
 TTTTGTATTTTATATGCCATCTTTGGAATTCCACTCTTTGGTTTCTTATTGGCTG
 GAATTGGAGACCAACTTGGAACCATCTTTGGGAAAAGCATTGCAAGAGTGGA
 GAAGGTCTTTCGAAAAAAGCAAGTGAGTCAGACCAAGATCCGGGTCATCTCA
 ACCATCCTGTTTCATCTTGGCCGGCTGCATTGTGTTTGTGACGATCCCTGCTGTC
 ATCTTTAAGTACATCGAGGGCTGGACGGCCTTGGAGTCCATTTACTTTGTGGT
 GGTCACCTCTGACCACGGTGGGCTTTGGTGATTTTGTGGCAGGGGGAAACGCT
 GGCATCAATTATCGGGAGTGGTATAAGCCCCTAGTGTGGTTTTTGGATCCTTGT
 TGGCCTTGCCTACTTTGCAGCTGTCCTCAGTATGATCGGAGATTGGCTACGGG
 TTCTGTCCAAAAAGACAAAAGAAGAGGTGGGTGAAATCAAGGCCCATGCGG
 CAGAGTGGAAGGCCAATGTACGGCTGAGTTCCGGGAGACACGGCGAAGGC
 TCAGCGTGGAGATCCACGATAAGCTGCAGCGGGCGGCCACCATCCGCAGCAT
 GGAGCGCCGGCGGCTGGGCCTGGACCAGCGGGCCCACTCACTGGACATGCTG
 TCCCCCGAGAAGCGCTCTGTCTTTGCTGCCCTGGACACCGGCCGCTTCAAGGC
 CTCATCCCAGGAGAGCATCAACAACCGGCCCAACAACCTGCGCCTGAAGGGG

CCGGAGCAGCTGAACAAGCATGGGCAGGGTGCGTCCGAGGACAACATCATC
AACAAAGTTCGGGTCCACCTCCAGACTCACCAAGAGGAAAAACAAGGACCTC
AAAAAGACCTTGCCCGAGGACGTTTCAGAAAATCTACAAGACCTTCCGGAATT
ACTCCCTGGACGAGGAGAAGAAAGAGGAGGAGACGGAAAAGATGTGTAAC
CAGACAACTCCAGCACAGCCATGCTGACGGACTGTATCCAGCAGCACGCTGA
GTTGGAGAACGGAATGATACCCACGGACACCAAAGACCGGGAGCCGGAGAA
CAACTCATTACTTGAAGACAGAAACGCAGAGAACCTCTACTTCCAATCGCAC
CATCATCACCATCACCATCACCACCATACAGAGACCAGCCAAGTGGCGCCTG
CCTGAGCGGCCGCGG

Appendix VI: DNA sequence of truncated TREK-2

The truncated TREK-2 contains the pore formation domain of ion channel, from glycine⁶⁷ to glutamate³⁴⁰, with a tobacco etch virus (TEV) protease cleavage site, a 10x His purification sequence. The construct was in pEU and expressed in D-CF mode for development of MP stability assay.

TTAAGAAGGAGATATACTATGGGCTTGCAGACCGTCATGAAGTGGAAGACGG
TGGTTGCCATCTTTGTGGTTGTGGTGGTCTACCTTGTCAGTGGCGGTCTTGTCT
TCCGGGCATTGGAGCAGCCCTTTGAGAGCAGCCAGAAGAATACCATCGCCTT
GGAGAAGGCGGAATTCCTGCGGGATCATGTCTGTGTGAGCCCCCAGGAGCTG
GAGACGTTGATCCAGCATGCTCTTGATGCTGACAATGCGGGAGTCAGTCCAA
TAGGAAACTCTTCCAACAACAGCAGCCACTGGGACCTCGGCAGTGCCTTTTT
CTTTGCTGGAAGTGTCAATTACGACCATAGGGTATGGGAATATTGCTCCGAGCA
CTGAAGGAGGCAAAATCTTTTGTATTTTATATGCCATCTTTGGAATTCCACTC
TTTGGTTTCTTATTGGCTGGAATTGGAGACCAACTTGAACCATCTTTGGGAA
AAGCATTGCAAGAGTGGAGAAGGTCTTTCGAAAAAAGCAAGTGAGTCAGAC
CAAGATCCGGGTCATCTCAACCATCCTGTTTCATCTTGGCCGGCTGCATTGTGT
TTGTGACGATCCCTGCTGTCATCTTTAAGTACATCGAGGGCTGGACGGCCTTG
GAGTCCATTTACTTTGTGGTGGTCACTCTGACCACGGTGGGCTTTGGTGATTT
TGTGGCAGGGGGAAACGCTGGCATCAATTATCGGGAGTGGTATAAGCCCCTA
GTGTGGTTTTGGATCCTTGTGGCCTTGCCTACTTTGCAGCTGTCCTCAGTATG
ATCGGAGATTGGCTACGGGTCTGTCCAAAAAGACAAAAGAAGAGGTGGGT
GAAGCAGAGAACCTCTACTTCCAATCGCACCATCATCACCATCACCATCACC
ACCATGATTACAAGGATGACGACGATAAGTGAGGATCC

Appendix VII: Impact of Additional Amphiphiles on Cell-Free Protein Synthesis

In CFPS, the additional molecules risk altering protein production and one of the most concerns is inhibited translation. We investigated such effects using an established GPCR protein, C-X-C Chemokine Receptor-4 (CXCR4). The expression was performed using pEU vector and bilayer wheat germ cell free reaction as described for GLP-1R. Diblock copolymer AB (Polymer A) and triblock copolymer ABA (Polymer B), lipids, and two common maltoside/zwitterionic detergents used in membrane proteins purification were tested for translational inhibition. Polymersomes, despite the type used, and maltoside detergent are compatible with cell free expression and cause no inhibition like lipids (Figure S1). In the presence of zwitterionic detergent, however, translation was completely inhibited. It is not surprising that inhibition happens only to zwitterionic detergents, as nonionic detergents are considered mild whereas charged zwitterionic detergents are denaturing. The filtration with 1,000 MWCO filter unit showed that self-assembled diblock copolymersomes AB were better in retaining incorporated CXCR4 whereas liposomes (lipid vesicles) and ABA polymersomes showed less retention. This is probably because of the size differences between AB, liposomes and ABA polymersomes. Besides, maltoside detergent showed little but detectable amount in Western blot probably due to the formation of micelles with CXCR4. This study is a good indication of polymersomes' general compatibility in term of the minimal effects on translation in CFPS.

Figure S1 Western blot analysis of cell free expressed CXCR4 in the presence of various amphiphilic molecules. (A) Cell free reaction mixture. (B) Filtration retentate using 1,000 kDa MWCO filter unit. Di-block (lane 1) and tri-block (lane 2) copolymers, lipid (lane 3), and maltoside

(lane 4) and zwitterionic detergent (lane 5) were tested for translation inhibition. Lane 6: expression without any additional molecules. Lane 7-9 are cell free reaction without target cDNA template (negative controls).

Appendix VIII: Functionality Characterization with Alternative Methods

In Chapter 5, GLP-1R orientation of GLP-1R proteopolymersomes was probed using SPR based Biacore system. Anti-PEG antibody capture mode was used for immobilization of prepared proteopolymersomes. Both 1D4 (recognizes C-terminal C9 epitope) and N-terminal domain specific pAb showed positive binding to incorporated GLP-1R. Qualitative analysis of binding response in SPR showed the presence of both correctly and reversely incorporated proteins and it was likely the mis-orientated proteins were dominant. This is further supported by the observations in peptide ligand exendin-4 binding experiments which showed minimal response (background level). The overall low signal probably reflected the low receptor density on membrane surface and the number of accessible NTD of GLP-1R. Besides the anti-PEG capture immobilization, direct EDC/NHS immobilization of proteopolymersomes were also attempted for binding analysis, although unsuccessful as below.

In the antibody capture mode, anti-PEG antibody was first immobilized onto biosensor chip using standard EDC/NHS amine coupling method followed by antibody binding capture of proteopolymersomes with PEG chain on membrane surface. The method proved to be successful in capture AB polymersomes with immobilization level up to 3000 RU. In direct capture mode, the size-based filtration purified proteopolymersomes were coupled onto biosensor chip using EDC/NHS method and NH₂ functionalized AB polymers. As shown in Figure S1, the vivaspin purified samples were successfully immobilized onto CM5 biosensor chip with similar immobilization level (1500 RU) for both GLP-1R and CXCR4.

Figure S2 SPR sensorgrams of EDC/NHS coupling of GLP-1R (orange) and CXCR4 (grey) proteopolymersomes onto Biacore sensor chip. Typical immobilization level was about 1500 RU.

The peptide exendin-4 was flowed over biosensor chip and no obvious specific binding was detected with exendin-4 concentrations ranging from 1 to 100 nM in two continuous rounds of injections (appendix IX). The sensorgrams showed background level response which was mostly like to be because either the absence of functional GLP-1R incorporated into AB polymersomes or the density of functional receptors accessible to exendin-4 was too low to be detected with Biacore 3000. With an alternative approach, we attempted to analyze GLP-1R functionality with specific epitope recognizing antibodies. Since most antibodies have much larger MW (typically 150 to 180 kDa) compared to peptide (MW of exendin-4 is 4 kDa), we expected to obtain higher response with antibody binding in SPR. First of all, the four commercially available GLP-1R antibodies, including mAB 2814 (antigen: full-length GLP-1R, R&D system), 1H4 Ab166987 mAB (antigen: DNA vaccine, Abcam), and anti-C9 tag 1D4 mAB were tested in Western blot. Although SPR is a

powerful tool in analyzing specific molecular interactions, however, with the currently available GLP-1R ligands/antibodies, very little information can be acquired. It is known that the 1D4 mAB targets a specific linear epitope (C9 tag) fused at the C-terminal of GLP-1R which is not conformational sensitive. For pAb TA590112 (Origene), it is known that the antibody targets epitope(s) at the N-terminal of GLP-1R but the specific sequence is unknown. The mAB 2814 has been validated with flow cytometry (RnD SYSTEMS, Catalog # MAB2814), which possibly indicates the mAB recognizing certain epitopes in the extracellular domain of GLP-1R. The 1H4 antibody was generated by DNA vaccine and has been validated by immunostaining of GLP-1R expressing pancreas tissue (Abcam, Catalog #ab166987). Although the specific DNA sequence is unknown, the validation indicates that like mAB 2814, the 1H4 also targets extracellular epitopes. In Western blot, 1H4 and mAB 2814 were positive whereas TA590112 was negative. Although, such analysis did not confirm the conformational sensitivity, it did indicate certain conformations of extracellular epitopes in GLP-1R are important for antibody binding which can be destroyed by protein denaturation. Thus, we tested the binding of these antibodies to produced GLP-1R purified with size-based filtration method immobilized on biosensor chip with EDC/NHS method (Figure S2). The SPR sensorgrams of all three antibodies showed positive specific binding despite the different responses (Figure S3). Given that the immobilized GLP-1R AB proteopolymersomes were of the same level, the responses of each antibody in SPR will be comparable. All the three antibodies mAB 2814, pAb TA590112 and 1H4 AB166987 bind to GLP-1R with responses within 40 RUs. The binding profile of mAB 2814 was similar to 1H4 AB166987 in which the response dropped from 35 to about 20 RU quickly after injection at about 100 s, indicating a fast dissociation and relatively low binding affinity (Figure S3, B, and D). In comparison, the pAb TA590112 showed slow dissociation and the response stayed almost unchanged after injection. The SPR binding profile was similar to 1D4 where slow dissociation was observed. Although the 1D4 mAB binding to C9 epitope was not thoroughly characterized, its high binding affinity was reflected by its various applications in purification and characterization membrane/soluble proteins, e.g. immunoprecipitation. In the second injections, mAB 2814 and 1H4 AB166987 showed similar level of responses whereas the response of pAb

TA590112 increased from 10 to 20 RU, indicating the slow dissociation has caused pAB accumulation on the sensor chip surface from the first injection.

Figure S3 SPR binding analysis of GLP-1R AB proteopolymersomes binding to epitope specific antibodies. (A) Western blot analysis and (B, C) SPR sesorgrams of GLP-1R AB

proteopolymersomes binding to epitope specific antibodies TA590112 NTD specific pAB, mAB 2814 (antigen: full-length GLP-1R, R&D system), 1H4 Ab166987 mAB (antigen: DNA vaccine, Abcam), and anti-C9 tag 1D4 mAB. (A) G1: GLP-1R. NC: cell free reaction without adding mRNA template (negative control). All antibodies were diluted to 0.5 ug/mL in blotting. (B) All sensorgrams were subtracted by the binding of non-relevant target CXCR4 purified and immobilized in the same way. In binding analysis, antibodies of 2 ug/mL were injected in two rounds separately. (C) Anti-C9 1D4 mAB is a known linear epitope recognizing antibody that does not sense any specific protein conformation.

Functional characterization of GLP-1R was also attempted using fluorescence-based and immunodetection-based methods, namely fluorescence polarization (FP) and enzyme-linked immunosorbent assay (ELISA). In FP, association of a fluorescent ligand (FAM-labelled exendin-4) to a large molecule (GLP-1R proteopolymersome) was measured with plane-polarized excitation light. The free fluorescent ligands' rapid tumbling depolarizes emission light and in immobile state, where ligand and large molecule are associated, the emission light becomes polarized. The correlation between ligand concentration and percentage of polarized emission light provides the foundation of measuring molecular interactions with FP (Figure S4). The proteopolymersomes were purified with filtration method and cell free reaction without supplied mRNA was used as negative control (NC). In the direct binding, both GLP-1R and NC showed decreased polarization level of about 150 mP because of light scattering from AB polymers (Figure S4). Competition with cold non-labelled ligand was performed to valid the binding between GLP-1R and exendin-4 (Figure S4). However, no significant competition was observed. Such observations were consistent with radioligand binding assay results, where in vivaspin filtration purified samples, no functional receptors were present. Despite the non-functional proteins produced, the sensitivity of FP is much lower than radioisotope labelling and very unlikely to be able to detect active binding at current level. Furthermore, polymer scattering seem to be an issue for fluorescence-based assays like FP .

Figure S4 Fluorescence polarization analysis of GLP-1R proteopolymersomes binding to peptide ligand exendin-4. (A) Expression and purification of GLP-1R using size-based filtration method. Lane 1: cell free expression of GLP-1R. Lane 2: cell free expression without supplied mRNA. (B) Binding of proteopolymersomes to FAM-labelled exendin-4 (10 nM). (C) Competition binding using non-labelled exendin-4 at concentrations from 0.1 to 1,000 nM. NS: no sample added. NC: cell free expression without supplied mRNA.

Immunodetection-based ELISA was also attempted to probe the ligand/antibody binding capability of GLP-1R preparations. Two different preparations, including histidine-affinity and vivaspin filtration (1,000 MWCO) purified GLP-1R in AB polymersomes were tested

together with non-relevant target CXCR4 as negative controls. For ELISA, GLP-1R and CXCR4 preparations were immobilized onto Maxisorb 96 well pre-coated with rat/rabbit anti-PEG antibody, followed by washing, antibody binding and colorimetric detection. Primary antibodies including TA590112 NTD specific pAB, mAB 2814 (antigen: full-length GLP-1R, R&D system), 1H4 Ab166987 mAB (antigen: DNA vaccine, Abcam), and 12G5 ab45001 mAB (CXCR4 extracellular conformation antibody) were tested. Anti-C9 tag 1D4 mAB was also tested in order to roughly determine the level of protein capture for different preparations. In no primary antibody control, secondary antibody was added directly after capture followed by colorimetric detection to determine background signal.

Figure S5 ELISA analysis of GLP-1R preparations antibody binding using rat anti-PEG antibody capture. (A-E) 1H4 Ab166987 mAB (antigen: DNA vaccine, Abcam), mAB 2814 (antigen: full-length GLP-1R, R&D system), 12G5 ab45001 mAB (CXCR4 extracellular conformation antibody), TA590112 NTD specific pAB and 1D4 mAB binding to histidine-affinity and vivaspin filtration purified GLP-1R and CXCR 4. (F) No primary antibody used (negative control).

It is known that from Chapter 4, the histidine-affinity purified proteins contained minimal tetrazole fluorescence (3 % of total fluorescence), indicating the absence of AB polymers in comparison to biotin-affinity purified proteins (30 %). However, the ELISA results with

1D4 as primary antibody showed both vivaspin filtration (VR) and histidine-affinity (H) purified receptors contain accessible PEG chain that can be captured by anti-PEG antibody (Figure S5, E). The histidine-affinity purified GLP-1R and CXCR4 even showed higher signal than the VR samples at all antibody concentrations from 0.1 to 10 ug/mL. Such results were not very consistent with tetrazole fluorescence of elution fraction in histidine-affinity purification. It was possible that small amount of non-fluorescent AB polymers were associated with purified receptors, leading to certain level of anti- capture in ELISA with PEG antibody. In the no primary antibody negative control, background level signal (about 0.25 OD) was observed across all the samples tested (-H and -VR), indicating neither the immobilized anti-PEG nor the captured GLP-1R/CXCR4 were causing increased colorimetric signal. Thus, it is probably because of the non-specific binding of 1D4 antibody to the well plate surface that was causing higher signal of GLP-1R/CXCR4-H than VR samples. This can be best validated by setting up immobilized rabbit anti-PEG with no sample capture controls using 1D4 antibody in ELISA. This may also explain the little difference observed for TA590112 pAB, mAB 2814, 1H4 mAB and 12G5 mAB between GLP-1R and CXCR4 (-H or -VR) in respective controls. The only positive result was showed in GLP-1R NTD specific antibody TA590112 pAB binding, where the histidine-affinity purified GLP-1R displayed higher signal in compared to CXCR4 at various concentrations from 0.1 to 10 ug/mL (Figure S5 D). Considering that the 1D4 binding showed CXCR4 had higher immobilization level, it probably reflected the specific TA590112 pAB binding to the N-terminal epitope of GLP-1R rather than CXCR4. However, for GLP-1R/CXCR4-VR, there was no such difference despite the 1D4 signals were similar (Figure S5, D, E). In comparison to the positive and specific TA590112 pAB binding to GLP-1R in SPR analysis, it was probably because of the processing (blocking/washing) in ELISA that has caused protein denaturation or dissociation from well-plate, given that both methods are of similar sensitivity.

In a parallel experiment, capture antibody rabbit anti-PEG was tested for the detection of 12G5 antibody binding to CXCR4-VR and -H (Figure S6). The no sample capture control showed minimal signal, indicating that the rabbit antibody did not interfere with colorimetric detection. The 1D4 binding showed despite the target or method of

purification, the overall protein immobilization level was lower than rat anti-PEG capture. Besides, the GLP-1R-H and -VR showed lower level of immobilization than CXCR4-H and -VR. Despite such difference, the 12G5 binding was almost equal to GLP-1R and CXCR4, -H and -VR, indicating no specific binding detected (Figure S6).

Figure S6 ELISA analysis of GLP-1R preparations antibody binding using rabbit anti-PEG antibody capture. (A, B) 12G5 ab45001 mAb (CXCR4 extracellular conformation antibody) and 1D4 mAb binding to histidine-affinity and vivaspin filtration purified GLP-1R and CXCR 4. (C) No primary antibody used (negative control).

Appendix IX: SPR Sensorgrams of Exendin-4 Binding

Figure S7 SPR sensorgrams of GLP-1R and CXCR4 binding to exendin-4. (A-D) GLP-1R and CXCR4 proteopolymersomes binding to peptide ligand exendin-4 from 1 to 100 nM. Both GLP-1R and CXCR4 were expressed using CFPS, purified using size-based filtration method and immobilized onto biosensor chip but displayed no detectable binding.

Appendix X: List of Publications

Journal Articles

1. **Hu Z**; Banerjee S; Cornell T; de Hoog H-PM; Nallani M. *Functional Production Strategy of Membrane Protein-Polymersomes: a Class B GPCR (Glucagon-Like Peptide-1 Receptors)* (submitted).
2. **Hu Z**; Ho CS; Nallani M. *Synthetic (Polymer) Biology (Membrane): Functionalization of Polymer Scaffolds for Membrane Proteins* (accepted). 2017, **Current Opinion in Biotechnology**. 2017 Feb 7;46:51-56.

Conference Presentation

Hu Z; Banerjee S; de Hoog H-PM; Cornell T; Nallani M. *Construction of Polymersome-based Artificial Cell Membrane* (poster presentation). 2016, **Gordon Research Conference (Biointerface Science)**.